hexagon:2611.00209v1

IMPROVED SCALING IN THE METRIC COTYPE INEQUALITY

MANOR MENDEL

ABSTRACT. Let $2 \leqslant q < \infty$. It is proved that every Banach space with Rademacher cotype q satisfies the metric cotype q inequality with scaling parameter $m = O(n^{1/2+1/q})$. The improved bound on the scaling is achieved by reducing the metric cotype problem to a smoothing and approximation criterion on the continuous torus. For a single unconditional kernel, the criterion is controlled by its mean ℓ_∞ radius and its variation in the main diagonal direction. The construction fits the universal smoothing and approximation scheme of Giladi, Mendel, and Naor. The proofs were generated by GPT 6 Astra with the help of the author.

CONTENTS

1. Introduction 1
2. Preliminaries 3
3. General convolution kernels 4
4. A simple product kernel 7
5. Limitations of the argument 9
A. Sharp metric cotype for K-convex spaces 12
References 17

1. INTRODUCTION

1.1. **Metric cotype and its scaling parameter.** A Banach space X has *Rademacher cotype q*, where $2 \leqslant q < \infty$, if there is $C < \infty$ such that, for every $n \in \mathbb{N}$ and $v_1, \ldots, v_n \in X$,

$$\Big(\sum_{j=1}^n \|v_j\|^q\Big)^{1/q} \leqslant C\Big(\mathbb{E}_\varepsilon \Big\|\sum_{j=1}^n \varepsilon_j v_j\Big\|^q\Big)^{1/q}. \tag{1.1}$$

Here and below, ε is uniform on $\{-1,1\}^n$. See the references in [8, § 1] for background on this important linear property.

The metric cotype inequality replaces the vectors in (1.1) by long coordinate increments of a map on a discrete torus. Write $\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z}$, and let $e_1, \ldots, e_n$ be the coordinate vectors.

Definition 1.1 (Metric cotype [8])**.** Let (M,d) be a metric space and $q > 0$. We say that M has *metric cotype q* if there is a constant $\Gamma > 0$ such that, for every $n \in \mathbb{N}$, there is an even positive integer m for which every map $f : \mathbb{Z}_m^n \to M$ satisfies the *metric cotype inequality*

$$\sum_{j=1}^n \mathbb{E}_x d\Big(f\Big(x + \tfrac{m}{2} e_j\Big), f(x)\Big)^q \leqslant \Gamma^q m^q \mathbb{E}_{x,\varepsilon} d\big(f(x+\varepsilon), f(x)\big)^q. \tag{1.2}$$

Here x and ε are independent and uniform on $\mathbb{Z}_m^n$ and $\{-1,1\}^n$, respectively.[1]

Date: November 8, 2026.

2020 *Mathematics Subject Classification.* Primary 46B07; Secondary 46B09, 46B85, 51F30.

Key words and phrases. Metric cotype, Rademacher cotype, convolution kernels, Fisher information.

[1] The original definition of metric cotype in [8] averages ε over $\{-1,0,1\}^n$. The two definitions are equivalent up-to the metric cotype constant [4, § 4].

For maps $f:\mathbb{Z}_m^n \to X$ into a Banach space and $q \geqslant 2$, we use the notation

$$\mathscr{L}(f) = \Big(\sum_{j=1}^{n} \mathbb{E}_x \Big\| f\Big(x+\tfrac{m}{2}e_j\Big) - f(x)\Big\|^q\Big)^{1/q}, \qquad \mathscr{E}(f) = \big(\mathbb{E}_{x,\varepsilon}\|f(x+\varepsilon)-f(x)\|^q\big)^{1/q}. \tag{1.3}$$

Thus (1.2) reads $\mathscr{L}(f) \leqslant \Gamma m \mathscr{E}(f)$.

It was proved in [8] that metric cotype q and Rademacher cotype q coincide for Banach spaces, with comparable constants. This qualitative equivalence guarantees an admissible m for every n. The parameter m is called the *scaling parameter* of the metric cotype inequality. Obtaining a small m is a separate quantitative problem which is the subject of this paper.

The scaling parameter is important for applications as it determines the sizes of finite ℓ_∞ grids to which the inequality can be applied, and hence the resulting distortion bounds [8, § 7]. A two-point test [8, Lemma 2.3] shows that a non-singleton metric space with metric cotype constant Γ requires $m \gtrsim_\Gamma n^{1/q}$. We therefore say that it has *sharp metric cotype q* if there are constants $A, \Gamma < \infty$, independent of n, such that for every n one can choose an even $m \leqslant An^{1/q}$ for which (1.2) holds for every f. Sharp scaling yields obstructions to coarse and uniform embeddings: if a metric space has sharp metric cotype q, with $q \geqslant 2$, then ℓ_p admits no coarse or uniform embedding into it for any $p > q$ [8]. The *sharp metric cotype conjecture* asks whether every Banach space of Rademacher cotype q, for $2 \leqslant q < \infty$, has sharp metric cotype q [8, § 8].

For general Banach spaces of Rademacher cotype q, the original proof in [8] gave metric cotype q with $m = O(n^{2+1/q})$. It was later improved to $m = O(n^{1+1/q})$ [5]. Here we further improve the bound on the scaling parameter.

Theorem 1.2. *Let $2 \leqslant q < \infty$, and suppose that X satisfies* (1.1) *with constant C. For every $n \in \mathbb{N}$, there is a positive integer m, divisible by four, with $m \leqslant 16n^{1/2+1/q}$, such that $\mathscr{L}(f) \leqslant (1+C)m\mathscr{E}(f)$ for every $f:\mathbb{Z}_m^n \to X$.*

1.2. **Further results and the relation to previous work.** A universal smoothing and approximation scheme for metric cotype in Banach spaces was isolated in [5] (already implicit in [8]): convolution operators approximate the original map while controlling randomized sums of their coordinate differences. Their construction has approximation constant $A = O(n)$ and bounded smoothing constant S. They also proved the obstruction $AS \gtrsim_q \sqrt{n}$, under a lower bound on m/A. The present paper gives a matching upper bound: our construction has $A = O(\sqrt{n})$ and $S = 4$, so $AS = O(\sqrt{n})$. Thus it attains the optimal order for the universal scheme, as discussed in § 5.3.

The sharp metric cotype conjecture remains unresolved for general Banach spaces of Rademacher cotype q. It is known for K-convex spaces [8], and for L_1 [3] and the Schatten trace class [7]. In Appendix A we give two continuous-torus proofs of the K-convex case, using the Rademacher projection directly and type–cotype duality.

Beyond Banach spaces, metric cotype has been studied for other classes of metric spaces; see [11, 4, 1, 2]. For a wider perspective on the relevance of the metric cotype inequality to metric geometry, see [8, 4].

1.3. **Outline of the proof of Theorem 1.2.** The proof of Theorem 1.2 follows the smoothing and approximation strategy of [8, 5]: average the map, apply Rademacher cotype to coordinate differences of the averages, and telescope to obtain long increments. The quantitative task is to keep both the approximation error and the Rademacher differences small in comparison to $\mathscr{E}(f)$.

We implement this strategy on the continuous torus $\mathbb{T} = (\mathbb{R}/(m/2)\mathbb{Z})^n$. Extend each parity class $\eta \in \{0,1\}^n$ of f to the step function $F_\eta(u) = f(2\lfloor u \rfloor + \eta)$ for $u \in \mathbb{T}$. Set $G_\eta = F_\eta * K$, where K is an unconditional probability density. Write

$$M(K) = \int_{\mathbb{R}^n} \|z\|_\infty K(z)\,\mathrm{d}z, \qquad D(K) = \Big\| \sum_{j=1}^{n} \partial_j K \Big\|_{L_1(\mathbb{R}^n)}.$$

In Proposition 3.4 we will prove the following estimates:

$$\big(\mathbb{E}_{\eta,u}\|F_\eta(u)-G_\eta(u)\|^q\big)^{1/q}\leqslant 2(1+M(K))\mathscr{E}(f),\qquad \Big(\mathbb{E}_{\eta,u,\varepsilon}\Big\|\sum_{j=1}^n \varepsilon_j\partial_j G_\eta(u)\Big\|^q\Big)^{1/q}\leqslant (2+D(K))\mathscr{E}(f).$$

Here η,u,ε are independent, with η uniform on $\{0,1\}^n$, ε uniform on $\{-1,1\}^n$ and u Haar-uniform on $\mathbb{T}$. Apply Rademacher cotype to the sequence of vectors $\partial_1 G_\eta(u),\dots,\partial_n G_\eta(u)$, for every $u\in\mathbb{T}$. Then, integrate $\partial_j G_\eta$ along coordinate segments of length $m/4$, and add the two endpoint errors. This yields

$$\mathscr{L}(f)\lesssim\big((1+M(K))n^{1/q}+Cm(1+D(K))\big)\mathscr{E}(f).$$

The key new observation is the directional smoothing estimate above. We slice the kernel along diagonal lines and use the one-dimensional cancellation estimate (Lemma 3.3). Unconditionality gives the same variation $D(K)$ in every sign direction.

For the product kernel

$$K_r(z)=r^{-n}\prod_{j=1}^n\rho\big(z_j/r\big),\qquad \rho(t)=\tfrac{3}{2}\big(1-|t|\big)_+^2,$$

the support gives $M(K_r)\leqslant r$, while independence and centering of the logarithmic derivatives evaluated at independent samples give $D(K_r)\leqslant\sqrt{12n}/r$. Taking $r=2\sqrt{n}$ therefore gives approximation error $O(\sqrt{n})\mathscr{E}(f)$ and a bounded smoothing constant. This convolution construction supplies the improvement within the strategy of [8, 5].

1.4. **AI statement.** The proofs in this paper were obtained by GPT 6 Astra. Most of the writing was done by GPT 6 Astra and GPT 6.1 Sol. The author's contributions consist of posing the questions; guiding the AI toward improved bounds, simpler arguments, and clearer exposition; manually editing the text at some points, and verifying the correctness of the proofs. The author does not claim credit for the result. This paper is intended to be posted online and will not be submitted to a mathematical journal.

2. Preliminaries

From now on, all Banach spaces are real. This is without loss of generality: a complex Banach space and its underlying real space have the same norm, Rademacher sums, and metric, so their Rademacher type and cotype, K-convexity, and metric cotype constants coincide.

All finite averages and all Haar measures on tori are normalized. Coordinate sums, such as the sum in (1.3), are not normalized. Convolution of a periodic function with a kernel on $\mathbb{R}^n$ means

$$(h*K)(u)=\int_{\mathbb{R}^n}h(u-z)K(z)\,\mathrm{d}z.$$

For $v,z\in\mathbb{R}^n$, write $v\odot z=(v_1z_1,\dots,v_nz_n)$. A function $h:\mathbb{R}^n\to\mathbb{R}$ is *unconditional* if it is invariant under coordinate sign changes, i.e., $h(\varepsilon\odot v)=h(v)$ for every $v\in\mathbb{R}^n$ and $\varepsilon\in\{-1,1\}^n$.

Fix n, an integer m divisible by four, and $f:\mathbb{Z}_m^n\to X$. For $S\subseteq\{1,\dots,n\}$, let $\varepsilon_S=\sum_{j\in S}\varepsilon_j e_j$.

Lemma 2.1. *For every fixed $S\subseteq\{1,\dots,n\}$,*

$$\big(\mathbb{E}_{x,\varepsilon}\|f\big(x+2\varepsilon_S\big)-f(x)\|^q\big)^{1/q}\leqslant 2\mathscr{E}(f).\tag{2.1}$$

Let $Y=(Y_1,\dots,Y_n)$ be a random vector taking values in $\mathbb{Z}^n$. Assume that its law is uncondtional. If $\|Y\|_\infty\leqslant L$ almost surely, where L is a nonnegative integer, then, for a uniform $x\in\mathbb{Z}_m^n$ independent of Y,

$$\big(\mathbb{E}_{x,Y}\|f(x+2Y)-f(x)\|^q\big)^{1/q}\leqslant 2L\mathscr{E}(f).\tag{2.2}$$

Proof. Let $\varepsilon'_j=\varepsilon_j$ for $j\in S$ and $\varepsilon'_j=-\varepsilon_j$ otherwise. Both ε and ε' are uniform sign vectors, and $\varepsilon+\varepsilon'=2\varepsilon_S$. Split the increment into these two steps and apply Minkowski's inequality. Translation invariance of x gives $\mathscr{E}(f)$ for the norm of each step, proving (2.1). The two sign vectors need not be independent.

For (2.2), condition on all the magnitudes $|Y_j|$. The signs of the nonzero coordinates are independent uniform signs; we can supply additional independent signs in the zero coordinates. Thus, conditionally, Y has the form

$$Y = \sum_{\ell=1}^{L} \varepsilon_{S_\ell}, \qquad S_\ell = \{j : |Y_j| \geqslant \ell\}.$$

Keep the magnitudes fixed and telescope along the partial sums $2\sum_{k=1}^{\ell} \varepsilon_{S_k}$. For each fixed ε, translating the uniform x removes the preceding partial sum from each term. Thus Minkowski's inequality and (2.1) give

$$\big(\mathbb{E}_{x,\varepsilon} \|f(x+2Y) - f(x)\|^q\big)^{1/q} \leqslant \sum_{\ell=1}^{L} \big(\mathbb{E}_{x,\varepsilon} \|f\big(x+2\varepsilon_{S_\ell}\big) - f(x)\|^q\big)^{1/q} \leqslant 2L\mathscr{E}(f).$$

Average the q-th power of this bound over the magnitudes to obtain (2.2). □

Let $\mathbb{T} = (\mathbb{R}/(m/2)\mathbb{Z})^n$ be the n-dimensional continuous torus of girth $m/2$. For each parity vector $\eta \in \{0,1\}^n$, define $F_\eta : \mathbb{T} \to X$ by

$$F_\eta(u) = f(2\lfloor u \rfloor + \eta), \qquad u \in \mathbb{T}, \tag{2.3}$$

where the floor is taken coordinatewise and the argument of f is reduced modulo m. Throughout the proof, η, u, and ε are independent and uniform on $\{0,1\}^n$, $\mathbb{T}$, and $\{-1,1\}^n$, respectively. Thus, for a family of functions $h_\eta : \mathbb{T} \to X$,

$$\mathbb{E}_{\eta,u} \|h_\eta(u)\|^q = \frac{1}{2^n (m/2)^n} \sum_{\eta \in \{0,1\}^n} \int_{[0,m/2)^n} \|h_\eta(u)\|^q \, \mathrm{d}u.$$

An additional subscript ε means that we also average over the 2^n sign vectors. We keep these averages explicit: $u \in \mathbb{T}$ records the position, $\eta \in \{0,1\}^n$ the parity class, and $\varepsilon \in \{-1,1\}^n$ the direction in the Rademacher average.

Lemma 2.2. *For every $v \in \mathbb{R}^n$,*

$$\big(\mathbb{E}_{\eta,u,\varepsilon} \|F_\eta(u + \varepsilon \odot v) - F_\eta(u)\|^q\big)^{1/q} \leqslant 2\lceil \|v\|_\infty \rceil \mathscr{E}(f). \tag{2.4}$$

Proof. Write $u = y + t$, where y, t are independent and uniform on $\{0, \dots, m/2-1\}^n$ and $[0,1)^n$, respectively, and independent of η, ε. Then

$$F_\eta(u + \varepsilon \odot v) - F_\eta(u) = f(x + 2Y) - f(x), \qquad x = 2y + \eta, \quad Y = \lfloor t + \varepsilon \odot v \rfloor.$$

Here x is uniform on $\mathbb{Z}_m^n$ and independent of Y. The coordinates of Y are independent, with $|Y_j| \leqslant \lceil \|v\|_\infty \rceil$. Replacing (t_j, ε_j) by $(1 - t_j, -\varepsilon_j)$ changes Y_j to $-Y_j$ almost surely, so each coordinate is symmetric. Apply (2.2). □

3. General convolution kernels

Let $K = K_n$ be an unconditional probability density such that

$$K \in W^{1,1}\big(\mathbb{R}^n\big), \qquad M(K) = \int_{\mathbb{R}^n} \|z\|_\infty K(z) \, \mathrm{d}z < \infty. \tag{3.1}$$

Here $W^{1,1}$ means that the function and all its first weak partial derivatives are integrable. We write

$$D(K) = \|\partial_{\mathbf{1}} K\|_{L_1(\mathbb{R}^n)}, \qquad \partial_{\mathbf{1}} K = \sum_{j=1}^{n} \partial_j K. \tag{3.2}$$

The following bound on the scaling parameter in terms of $M(K)$ and $D(K)$ is main result of the section.

Theorem 3.1. *Suppose that X satisfies* (1.1) *with constant C, and that K is an unconditional probability density satisfying* (3.1)*. For every positive integer m divisible by four and every $f : \mathbb{Z}_m^n \to X$,*

$$\mathscr{L}(f) \leqslant \Big(4(1+M(K))n^{1/q} + \frac{Cm}{4}(2+D(K))\Big)\mathscr{E}(f). \tag{3.3}$$

In particular, there is an admissible scale

$$m = 4\Big\lceil \Big(1 + \frac{M(K)D(K)}{2}\Big) n^{1/q} \Big\rceil \tag{3.4}$$

for which $\mathscr{L}(f) \leqslant (1+C)m\mathscr{E}(f)$ for every f.

Before proving Theorem 1.2, we first state a more general continuous smoothing and approximation criterion. Here, the kernels may vary with the coordinate.

Proposition 3.2. *Let $2 \leqslant q < \infty$, let $n \in \mathbb{N}$, and let m be a positive integer divisible by four. Suppose that X satisfies* (1.1) *with constant C. Let $K_1, \dots, K_n$ be probability densities on $\mathbb{R}^n$ such that the weak derivative $\partial_j K_j$ belongs to $L_1(\mathbb{R}^n)$ for each j. For $f : \mathbb{Z}_m^n \to X$, put $G_{\eta,j} = F_\eta * K_j$. Suppose that $A, S \geqslant 0$ satisfy*

$$\Big(\frac{1}{n}\sum_{j=1}^n \mathbb{E}_{\eta,u}\|F_\eta(u) - G_{\eta,j}(u)\|^q\Big)^{1/q} \leqslant A\mathscr{E}(f), \tag{3.5}$$

$$\Big(\mathbb{E}_{\eta,u,\varepsilon}\Big\|\sum_{j=1}^n \varepsilon_j \partial_j G_{\eta,j}(u)\Big\|^q\Big)^{1/q} \leqslant S\mathscr{E}(f) \tag{3.6}$$

for every $f : \mathbb{Z}_m^n \to X$. Then every such f satisfies

$$\mathscr{L}(f) \leqslant \Big(2An^{1/q} + \frac{CmS}{4}\Big)\mathscr{E}(f). \tag{3.7}$$

Proof. Since F_η is bounded, convergence of the difference quotients of K_j in L_1 gives $\partial_j G_{\eta,j} = F_\eta * \partial_j K_j$. This derivative is continuous by continuity of translations in L_1. Apply (1.1) and (3.6),

$$\Big(\sum_{j=1}^n \mathbb{E}_{\eta,u}\|\partial_j G_{\eta,j}(u)\|^q\Big)^{1/q} \leqslant C\Big(\mathbb{E}_{\eta,u,\varepsilon}\Big\|\sum_{j=1}^n \varepsilon_j \partial_j G_{\eta,j}(u)\Big\|^q\Big)^{1/q} \leqslant CS\mathscr{E}(f). \tag{3.8}$$

Since $m/4$ is an integer, the definition of F_η gives

$$\mathscr{L}(f) = \Big(\sum_{j=1}^n \mathbb{E}_{\eta,u}\Big\|F_\eta\Big(u + \tfrac{m}{4}e_j\Big) - F_\eta(u)\Big\|^q\Big)^{1/q}.$$

Replace F_η by $G_{\eta,j}$ at both endpoints in the j-th summand. Minkowski's inequality and translation invariance of Haar measure give

$$\mathscr{L}(f) \leqslant \Big(\sum_{j=1}^n \mathbb{E}_{\eta,u}\Big\|G_{\eta,j}\Big(u + \tfrac{m}{4}e_j\Big) - G_{\eta,j}(u)\Big\|^q\Big)^{1/q} + 2\Big(\sum_{j=1}^n \mathbb{E}_{\eta,u}\|F_\eta(u) - G_{\eta,j}(u)\|^q\Big)^{1/q}. \tag{3.9}$$

The last term is at most $2An^{1/q}\mathscr{E}(f)$ by (3.5). For the first term, the fundamental theorem of calculus, Minkowski's inequality, and translation invariance in each summand yield

$$\begin{aligned}\Big(\sum_{j=1}^n \mathbb{E}_{\eta,u}\Big\|G_{\eta,j}\Big(u + \tfrac{m}{4}e_j\Big) - G_{\eta,j}(u)\Big\|^q\Big)^{1/q} &\leqslant \int_0^{m/4}\Big(\sum_{j=1}^n \mathbb{E}_{\eta,u}\|\partial_j G_{\eta,j}(u + te_j)\|^q\Big)^{1/q}\mathrm{d}t \\ &= \frac{m}{4}\Big(\sum_{j=1}^n \mathbb{E}_{\eta,u}\|\partial_j G_{\eta,j}(u)\|^q\Big)^{1/q} \leqslant \frac{CmS}{4}\mathscr{E}(f),\end{aligned} \tag{3.10}$$

where the last inequality follows from (3.8). Combining these estimates proves (3.7). □

We next specialize to a single unconditional kernel and prove Theorem 1.2.

Lemma 3.3. *Let X be a Banach space, let $\phi \geqslant 0$ belong to $W^{1,1}(\mathbb{R})$, and let $h : \mathbb{R} \to X$ be bounded and strongly measurable. If*

$$\|h(t) - h(s)\|_X \leqslant A(|t-s|+1) \qquad (s,t \in \mathbb{R}),$$

then

$$\Big\| \int_{\mathbb{R}} \phi'(t) h(t)\,\mathrm{d}t \Big\|_X \leqslant A\Big(\int_{\mathbb{R}} \phi(t)\,\mathrm{d}t + \frac{1}{2}\int_{\mathbb{R}} |\phi'(t)|\,\mathrm{d}t \Big). \tag{3.11}$$

Proof. First suppose that ϕ is nonnegative, compactly supported, and piecewise linear with finitely many pieces. For $s > 0$, write its superlevel set, consisting of the points where ϕ exceeds s, as a finite union of disjoint open intervals:

$$\{t \in \mathbb{R} : \phi(t) > s\} = \bigcup_k \big(a_k(s), b_k(s)\big).$$

To change variables from t to $s = \phi(t)$, first consider one nonconstant linear piece (a,b). For s strictly between $\phi(a)$ and $\phi(b)$, let $t(s)$ be the unique point of this piece with $\phi(t(s)) = s$. Since $\mathrm{d}s = \phi'(t)\,\mathrm{d}t$,

$$\int_a^b \phi'(t) h(t)\,\mathrm{d}t = \begin{cases} \displaystyle\int_{\phi(a)}^{\phi(b)} h(t(s))\,\mathrm{d}s, & \phi(a) < \phi(b), \\ \displaystyle -\int_{\phi(b)}^{\phi(a)} h(t(s))\,\mathrm{d}s, & \phi(a) > \phi(b). \end{cases}$$

Thus an increasing piece contributes with a plus sign and a decreasing piece with a minus sign; constant pieces contribute zero. Except at the finitely many levels attained at breakpoints, each $a_k(s)$ lies on an increasing piece and each $b_k(s)$ lies on a decreasing piece. These are all the points where $\phi(t) = s$. Summing the contributions of the linear pieces therefore gives

$$\int_{\mathbb{R}} \phi'(t) h(t)\,\mathrm{d}t = \int_0^\infty \sum_k \big(h\big(a_k(s)\big) - h\big(b_k(s)\big) \big)\,\mathrm{d}s. \tag{3.12}$$

The layer-cake identity and the same change of variables give

$$\int_0^\infty \sum_k \big(b_k(s) - a_k(s) \big)\,\mathrm{d}s = \int_{\mathbb{R}} \phi, \qquad \int_0^\infty \sum_k 1\,\mathrm{d}s = \frac{1}{2}\int_{\mathbb{R}} |\phi'|.$$

The triangle inequality now proves (3.11).

Every nonnegative $\phi \in W^{1,1}(\mathbb{R})$ is a limit in $W^{1,1}$ of nonnegative compactly supported piecewise linear functions: multiply by nonnegative compactly supported cutoffs, smooth by a nonnegative mollifier, and interpolate on a fine partition. Since h is bounded, L_1 convergence of the derivatives allows passage to the limit on the left-hand side of (3.11). Convergence in $W^{1,1}$ allows passage to the limit in both terms on the right-hand side as well. □

For a differentiable function $h : \mathbb{T} \to X$, write $\partial_\varepsilon h(u) = \sum_{j=1}^n \varepsilon_j \partial_j h(u)$. Derivatives are taken in u, with ε held fixed. We use the same notation for functions on $\mathbb{R}^n$.

Proposition 3.4. *Let K be an unconditional probability density satisfying* (3.1)*, and put $G_\eta = F_\eta * K$. Then G_η is continuously differentiable, and*

$$\big(\mathbb{E}_{\eta,u} \|F_\eta(u) - G_\eta(u)\|^q \big)^{1/q} \leqslant 2(1 + M(K))\mathscr{E}(f), \tag{3.13}$$

$$\big(\mathbb{E}_{\eta,u,\varepsilon} \|\partial_\varepsilon G_\eta(u)\|^q \big)^{1/q} \leqslant (2 + D(K))\mathscr{E}(f). \tag{3.14}$$

Proof. The regularity of G_η and the identity $\partial_j G_\eta = F_\eta * \partial_j K$ follow from the difference-quotient argument in the proof of Proposition 3.2.

Let V have density K. Unconditionality gives $G_\eta(u) = \mathbb{E}_{V,\varepsilon} F_\eta(u - \varepsilon \odot V)$. Minkowski's inequality, followed by Jensen's inequality in ε and (2.4), yields

$$\begin{aligned}\big(\mathbb{E}_{\eta,u}\|F_\eta(u) - G_\eta(u)\|^q\big)^{1/q} &\leqslant \mathbb{E}_V\big(\mathbb{E}_{\eta,u,\varepsilon}\|F_\eta(u-\varepsilon\odot V) - F_\eta(u)\|^q\big)^{1/q} \\ &\leqslant 2\mathbb{E}\lceil\|V\|_\infty\rceil\mathscr{E}(f) \leqslant 2(1+M(K))\mathscr{E}(f).\end{aligned}$$

For smoothing, identify $\mathbb{R}^{n-1}$ with $\{w \in \mathbb{R}^n : w_n = 0\}$ and put $\phi_w(t) = K(w + t\mathbf{1})$. The linear change of variables $z = w + t\mathbf{1}$ has Jacobian one. For almost every w, the section ϕ_w is nonnegative and belongs to $W^{1,1}(\mathbb{R})$, with $\phi_w'(t) = \partial_{\mathbf{1}} K(w + t\mathbf{1})$. Differentiating $K(\varepsilon \odot z) = K(z)$ in the weak sense gives $(\partial_j K)(\varepsilon \odot z) = \varepsilon_j \partial_j K(z)$ for almost every z. Hence

$$\big(\partial_\varepsilon K\big)(\varepsilon \odot z) = \sum_{j=1}^n \varepsilon_j^2 \partial_j K(z) = \partial_{\mathbf{1}} K(z).$$

Changing variables first by the coordinate sign changes and then by $z = w + t\mathbf{1}$ therefore gives

$$\partial_\varepsilon G_\eta(u) = \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} \phi_w'(t) F_\eta(u - \varepsilon \odot (w + t\mathbf{1}))\,\mathrm{d}t\,\mathrm{d}w.$$

For fixed w, regard

$$h_w(t)(\eta, u, \varepsilon) = F_\eta(u - \varepsilon \odot (w + t\mathbf{1}))$$

as an $L_q(\eta, u, \varepsilon; X)$-valued function. It is bounded and strongly measurable. For fixed $s, t \in \mathbb{R}$, substitute $u' = u - \varepsilon \odot (w + s\mathbf{1})$. Conditional on each sign vector, u' remains Haar-uniform, and the difference becomes $F_\eta(u' - (t-s)\varepsilon) - F_\eta(u')$. Thus (2.4) gives

$$\|h_w(t) - h_w(s)\|_{L_q(\eta,u,\varepsilon;X)} \leqslant 2\lceil|t-s|\rceil\mathscr{E}(f) \leqslant 2(|t-s|+1)\mathscr{E}(f).$$

Apply Lemma 3.3 and then Minkowski's inequality in w. Since

$$\int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} \phi_w(t)\,\mathrm{d}t\,\mathrm{d}w = 1, \qquad \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} |\phi_w'(t)|\,\mathrm{d}t\,\mathrm{d}w = D(K),$$

we obtain (3.14). □

Proof of Theorem 3.1. Apply Proposition 3.2 with $K_j = K$ for every j. By Proposition 3.4, we may take $A = 2(1 + M(K))$ and $S = 2 + D(K)$, which gives (3.3).

For $r > 0$, the dilation $K_r(z) = r^{-n} K(z/r)$ satisfies

$$M\big(K_r\big) = r\,M(K), \qquad D\big(K_r\big) = D(K)/r. \tag{3.15}$$

Moreover, $D(K) > 0$ (otherwise almost every integrable line section of K would be constant, hence zero). Take $r = D(K)/2$. Then $D(K_r) = 2$, and (3.3) applied to K_r gives

$$\mathscr{L}(f) \leqslant \Big(4\Big(1 + \frac{M(K)D(K)}{2}\Big)n^{1/q} + Cm\Big)\mathscr{E}(f).$$

By (3.4), $4(1 + M(K)D(K)/2)n^{1/q} \leqslant m$, so the last bound is at most $(1 + C)m\mathscr{E}(f)$. □

4. A SIMPLE PRODUCT KERNEL

The following lemma gives a one-dimensional sufficient condition for $M(K_n)D(K_n) = O(\sqrt{n})$. For an absolutely continuous density ρ, denote

$$I(\rho) = \int_{\{\rho>0\}} \frac{\rho'(t)^2}{\rho(t)}\,\mathrm{d}t.$$

$I(\rho)$ is called in statistics the Fisher information of the parametrized densities $p_\theta(x) = \rho(x - \theta)$.

Lemma 4.1. *Let $\rho \in W^{1,1}(\mathbb{R})$ be an even probability density supported on $[-a,a]$, and assume $I(\rho) < \infty$. Then $K_n(z) = \prod_{j=1}^n \rho(z_j)$ satisfies*

$$M\big(K_n\big) \leqslant a, \qquad D\big(K_n\big) \leqslant \sqrt{nI(\rho)}. \tag{4.1}$$

In particular, every fixed such ρ satisfies

$$M\big(K_n\big)D\big(K_n\big) = O(\sqrt{n}). \tag{4.2}$$

Proof. Let $T_1, \dots, T_n$ be independent random variables with density ρ, and put $S_j = \rho'(T_j)/\rho(T_j)$, defining the ratio to be zero where $\rho = 0$. These values do not affect the law of S_j, since $\Pr(\rho(T_j) = 0) = 0$. The weak product rule gives

$$\partial_j K_n(z) = \rho'\big(z_j\big) \prod_{k \neq j} \rho\big(z_k\big).$$

Moreover, $\rho' = 0$ almost everywhere on $\{\rho = 0\}$, so summing over j yields, for almost every $z \in \mathbb{R}^n$,

$$\partial_{\mathbf{1}} K_n(z) = \sum_{j=1}^n \rho'\big(z_j\big) \prod_{k \neq j} \rho\big(z_k\big) = K_n(z) \sum_{j=1}^n \frac{\rho'\big(z_j\big)}{\rho\big(z_j\big)}.$$

Since $(T_1, \dots, T_n)$ has joint density K_n, it follows that

$$D\big(K_n\big) = \int_{\mathbb{R}^n} |\partial_{\mathbf{1}} K_n(z)| \, \mathrm{d}z = \int_{\mathbb{R}^n} \Big| \sum_{j=1}^n \frac{\rho'\big(z_j\big)}{\rho\big(z_j\big)} \Big| K_n(z) \, \mathrm{d}z = \mathbb{E}\Big| \sum_{j=1}^n S_j \Big|. \tag{4.3}$$

The integrable density $\rho \in W^{1,1}(\mathbb{R})$ tends to zero at both ends of the line. Consequently,

$$\mathbb{E}\big[S_j\big] = \int_{\mathbb{R}} \frac{\rho'(t)}{\rho(t)} \rho(t) \, dt = 0, \quad \text{while} \quad \mathbb{E}\big[S_j^2\big] = \int_{\mathbb{R}} \frac{\rho'\big(t\big)^2}{\rho\big(t\big)^2} \rho(t) \, dt = I(\rho).$$

Hence independence and Cauchy–Schwarz imply

$$D\big(K_n\big)^2 = \Big(\mathbb{E}\Big| \sum_{j=1}^n S_j \Big|\Big)^2 \leqslant \mathbb{E}\Big[\Big(\sum_{j=1}^n S_j\Big)^2\Big] = \mathbb{E}\Big[\sum_{j=1}^n S_j^2\Big] = nI(\rho).$$

The support gives $M(K_n) \leqslant a$. □

We choose the squared-tent probability density

$$\rho(t) = \frac{3}{2}\big(1 - |t|\big)_+^2, \qquad K_r(z) = r^{-n} \prod_{j=1}^n \rho\big(z_j/r\big). \tag{4.4}$$

It is even, belongs to $W^{1,1}(\mathbb{R})$, and satisfies

$$\int_{\mathbb{R}} \rho(t) \, \mathrm{d}t = 3 \int_0^1 \big(1 - t\big)^2 \, \mathrm{d}t = 1, \qquad \rho'(t) = -3 \operatorname{sgn}(t)(1 - |t|) \quad (0 < |t| < 1).$$

Consequently,

$$\frac{|\rho'(t)|^2}{\rho(t)} = 6 \quad (0 < |t| < 1), \qquad I(\rho) = 12. \tag{4.5}$$

For $r = 2\sqrt{n}$, Lemma 4.1 and (3.15) give

$$M\big(K_r\big) \leqslant 2\sqrt{n}, \qquad D\big(K_r\big) \leqslant \sqrt{3} < 2. \tag{4.6}$$

Although ρ is not differentiable at zero, Proposition 3.4 ensures that $F_\eta * K_r$ is continuously differentiable.

Proof of Theorem 1.2. By (3.3) and (4.6),

$$\mathscr{L}(f) \leqslant \big(4(1+2\sqrt{n})n^{1/q} + Cm\big)\mathscr{E}(f) \leqslant \big(12n^{1/2+1/q} + Cm\big)\mathscr{E}(f).$$

Take $m = 4\lceil 3n^{1/2+1/q}\rceil$. Then $m \geqslant 12n^{1/2+1/q}$, which gives $\mathscr{L}(f) \leqslant (1+C)m\mathscr{E}(f)$, and

$$m \leqslant 12n^{1/2+1/q} + 4 \leqslant 16n^{1/2+1/q}. \qquad \square$$

Remark 4.2 (Choice of the density). The squared tent has the least polynomial degree in the following class: even probability densities supported on $[-1,1]$, given by one polynomial on $(0,1)$, whose zero extension belongs to $W^{1,1}(\mathbb{R})$ and has finite $I(\cdot)$. Indeed, the polynomial must vanish at 1. A simple zero would make ρ'^2/ρ comparable to $(1-t)^{-1}$ near 1, whose integral diverges. Hence the zero has multiplicity at least two. Degree two is attained uniquely by $c(1-t)^2$, and normalization gives $c = 3/2$.

Remark 4.3. The kernel $K_n = \rho^{\otimes n}$ where $\rho(t) = \cos^2(\pi t/2)\mathbf{1}_{[-1,1]}(t)$, gives $M(K_n)D(K_n) \leqslant \sqrt{2\pi n}$. We omit the calculation.

5. Limitations of the argument

5.1. **Bounds for the moment and variation criterion.**

Proposition 5.1. *Every n-dimensional unconditional probability density K_n satisfying* (3.1) *obeys*

$$M\big(K_n\big)D\big(K_n\big) \geqslant (1-o(1))\sqrt{\pi n/8}. \tag{5.1}$$

Proof. Write $K = K_n$. Let $c_n = \mathbb{E}_\varepsilon[|\varepsilon_1 + \cdots + \varepsilon_n|]$. Let $Z = (Z_1, \ldots, Z_n)$ have density K. Write $\ell(z) = n^{-1}\sum_{j=1}^n z_j$, and put $Y = \ell(Z)$. For every $z \in \mathbb{R}^n$, convexity in each coordinate gives

$$\mathbb{E}_\varepsilon\Big|\sum_{j=1}^n \varepsilon_j z_j\Big| \leqslant c_n\|z\|_\infty.$$

The maximum of this convex function on $[-\|z\|_\infty, \|z\|_\infty]^n$ occurs at a vertex. Unconditionality therefore gives

$$\mathbb{E}[|Y|] = \frac{1}{n}\mathbb{E}_Z\mathbb{E}_\varepsilon\Big|\sum_{j=1}^n \varepsilon_j Z_j\Big| \leqslant \frac{c_n}{n}M(K). \tag{5.2}$$

Let p be the density of $Y = \ell(Z)$. Let $\chi \in C_c^\infty(\mathbb{R}^n)$ be an arbitrary "smooth cutoff" function that satisfies $0 \leqslant \chi \leqslant 1$, $\chi = 1$ on $\{\|z\|_2 \leqslant 1\}$, and $\chi = 0$ on $\{\|z\|_2 \geqslant 2\}$. For $R > 0$ and $\psi \in C_c^\infty(\mathbb{R})$, the function $\chi(z/R)\psi(\ell(z))$ is a smooth compactly supported test function on $\mathbb{R}^n$. Since $\partial_{\mathbf{1}}\ell = 1$, weak integration by parts gives

$$\int_{\mathbb{R}^n} K(z)\chi(z/R)\psi'(\ell(z))\,\mathrm{d}z = -\int_{\mathbb{R}^n} \partial_{\mathbf{1}}K(z)\chi(z/R)\psi(\ell(z))\,\mathrm{d}z - \frac{1}{R}\int_{\mathbb{R}^n} K(z)\big(\partial_{\mathbf{1}}\chi\big)(z/R)\psi(\ell(z))\,\mathrm{d}z.$$

The last term tends to zero, because $\int K = 1$ and

$$\frac{1}{R}\Big|\int_{\mathbb{R}^n} K(z)\big(\partial_{\mathbf{1}}\chi\big)(z/R)\psi(\ell(z))\,\mathrm{d}z\Big| \leqslant \frac{\|\psi\|_\infty\|\partial_{\mathbf{1}}\chi\|_\infty}{R} \longrightarrow 0.$$

Since $K, \partial_{\mathbf{1}}K \in L_1(\mathbb{R}^n)$, dominated convergence in the remaining terms yields

$$\int_{\mathbb{R}} p(t)\psi'(t)\,\mathrm{d}t = \int_{\mathbb{R}^n} K(z)\psi'(\ell(z))\,\mathrm{d}z = -\int_{\mathbb{R}^n} \partial_{\mathbf{1}}K(z)\psi(\ell(z))\,\mathrm{d}z.$$

The pushforward of $\partial_{\mathbf{1}}K(z)\,\mathrm{d}z$ under ℓ has an L_1 density, which this identity identifies with p'. Consequently, $p \in W^{1,1}(\mathbb{R})$ and $\|p'\|_1 \leqslant D(K)$.

Put $H = \|p\|_\infty$. Since $p \geqslant 0$ tends to zero at both ends of the line, $2H \leqslant \|p'\|_1 \leqslant D(K)$. Also $\Pr(|Y| \leqslant t) \leqslant 2Ht$, so

$$\mathbb{E}[|Y|] \geqslant \int_0^{1/(2H)} (1-2Ht)\,\mathrm{d}t = \frac{1}{4H} \geqslant \frac{1}{2D(K)}. \tag{5.3}$$

Combining (5.2) and (5.3) proves

$$D(K)M(K) \geqslant \frac{n}{2c_n}.$$

Put $S_n = n^{-1/2}\sum_{j=1}^n \varepsilon_j$. Since $\mathbb{E}S_n^2 = 1$, the random variables $|S_n|$ are uniformly integrable. The central limit theorem therefore gives

$$\frac{c_n}{\sqrt{n}} = \mathbb{E}|S_n| \longrightarrow \mathbb{E}|G| = \sqrt{\frac{2}{\pi}}, \qquad G \sim N(0,1).$$

Thus $c_n = (1+o(1))\sqrt{2n/\pi}$, which proves the claim. □

5.2. **A characterization for product kernels.**

Proposition 5.2. *Let $\rho \in W^{1,1}(\mathbb{R})$ be a fixed even probability density, independent of n, and put $K_n(z) = \prod_{j=1}^n \rho(z_j)$. Then*

$$M\big(K_n\big)D\big(K_n\big) = O(\sqrt{n})$$

if and only if ρ has bounded support and $I(\rho) < \infty$.

Proof. Sufficiency is Lemma 4.1. For necessity, let $(T_j)_{j\geqslant 1}$ be independent with density ρ, and set $S_j = \rho'(T_j)/\rho(T_j)$. Then

$$M\big(K_n\big) = \mathbb{E}\max_{j\leqslant n}|T_j|, \qquad D\big(K_n\big) = \mathbb{E}\Big|\sum_{j=1}^n S_j\Big|.$$

The variables S_j are symmetric, and $0 < \mathbb{E}|S_j| = \|\rho'\|_1 < \infty$. Conditional on their magnitudes, the scalar Khintchine inequality gives

$$D\big(K_n\big) \geqslant \frac{1}{\sqrt{3}}\mathbb{E}\Big(\sum_{j=1}^n S_j^2\Big)^{1/2} \geqslant \sqrt{\frac{n}{3}}\mathbb{E}|S_1|. \tag{5.4}$$

Therefore the assumed bound forces $M(K_n)$ to remain bounded. But $\max_{j\leqslant n}|T_j|$ increases almost surely to the essential supremum of $|T_1|$, so ρ has bounded support.

Since $M(K_n) \geqslant \mathbb{E}|T_1| > 0$, the assumed bound also gives $D(K_n) \leqslant B\sqrt{n}$ for some finite B. For every $R > 0$, (5.4) implies

$$B \geqslant \frac{1}{\sqrt{3}}\mathbb{E}\Big(\frac{1}{n}\sum_{j=1}^n \min\{S_j^2, R\}\Big)^{1/2}.$$

The law of large numbers and bounded convergence, followed by monotone convergence as $R \to \infty$, give $I(\rho) = \mathbb{E}S_1^2 \leqslant 3B^2 < \infty$. □

5.3. **Relation to the smoothing and approximation scheme.** We verify that the general construction fits the smoothing and approximation scheme from [5, Section 1.1]. Fix $n \in \mathbb{N}$ and $m \in 4\mathbb{N}$. Choose probability measures $\nu_1, \dots, \nu_n$ on $\mathbb{Z}_m^n$, and probability laws β_1, β_2 on pairs (x, y) with $x - y \in \{-1, 0, 1\}^n$, interpreted modulo m. For $T_j f = f * \nu_j$, the scheme requires, for every Banach space X and every $f : \mathbb{Z}_m^n \to X$,

$$\frac{1}{n}\sum_{j=1}^n \mathbb{E}_x\|T_j f(x) - f(x)\|^q \leqslant A^q \int \|f(x) - f(y)\|^q \,\mathrm{d}\beta_1(x,y), \tag{5.5}$$

$$\mathbb{E}_{x,\varepsilon}\Big\|\sum_{j=1}^n \varepsilon_j\big(T_j f\big(x+e_j\big) - T_j f\big(x-e_j\big)\big)\Big\|^q \leqslant S^q \int \|f(x) - f(y)\|^q \,\mathrm{d}\beta_2(x,y). \tag{5.6}$$

Proposition 5.3. *Every unconditional probability density K satisfying* (3.1) *gives a scheme with*

$$A = 3 + 2M(K), \qquad S = 2 + D(K).$$

One can take β_2 to be the law of $(x, x+\varepsilon)$.

Proof. Let V have density K, and let U be independent and uniform on $[0,1)^n$. Let $U^{(j)}$ agree with U except that its j-th coordinate is zero. Define

$$T_j f(x) = \mathbb{E}_{U,V} f\big(x + e_j + 2\lfloor U^{(j)} - V \rfloor\big). \tag{5.7}$$

This is convolution with the probability law of $-e_j - 2\lfloor U^{(j)} - V\rfloor$, reduced modulo m.

Write $x = 2y + \eta$. From the definition of F_η and $G_\eta = F_\eta * K$,

$$T_j f\big(x + e_j\big) = \mathbb{E}_U G_\eta\big(y + U^{(j)} + e_j\big), \qquad T_j f\big(x - e_j\big) = \mathbb{E}_U G_\eta\big(y + U^{(j)}\big).$$

The fundamental theorem of calculus gives

$$T_j f\big(x + e_j\big) - T_j f\big(x - e_j\big) = \int_{[0,1)^n} \partial_j G_\eta(y + t)\,\mathrm{d}t. \tag{5.8}$$

For uniform $x = 2y + \eta$ and independent uniform $t \in [0,1)^n$, the point $u = y + t$ is Haar-uniform on $\mathbb{T}$ and independent of η. Since the integral in (5.8) is over the same cube for every j, Jensen's inequality and (3.14) give

$$\begin{aligned}\mathbb{E}_{x,\varepsilon}\Big\| \sum_{j=1}^n \varepsilon_j \big(T_j f\big(x + e_j\big) - T_j f\big(x - e_j\big)\big)\Big\|^q &= \mathbb{E}_{\eta,y,\varepsilon}\Big\| \int_{[0,1)^n} \sum_{j=1}^n \varepsilon_j \partial_j G_\eta(y + t)\,\mathrm{d}t\Big\|^q \\ &\leqslant \mathbb{E}_{\eta,y,t,\varepsilon}\|\partial_\varepsilon G_\eta(y + t)\|^q = \mathbb{E}_{\eta,u,\varepsilon}\|\partial_\varepsilon G_\eta(u)\|^q \leqslant \big(2 + D(K)\big)^q \mathscr{E}\big(f\big)^q.\end{aligned}$$

This proves (5.6) with $S = 2 + D(K)$ and β_2 the law of $(x, x+\varepsilon)$.

For approximation, compare each T_j with the common averaging operator

$$T f(x) = \mathbb{E}_{U,V} f\big(x + 2\lfloor U - V\rfloor\big) = \int_{[0,1)^n} G_\eta(y + t)\,\mathrm{d}t, \qquad x = 2y + \eta.$$

Since for $t \in [0,1)^n$,

$$F_\eta(y + t) = f\big(2\lfloor y + t\rfloor + \eta\big) = f(2y + \eta) = f(x),$$

Jensen's inequality and (3.13) give

$$\begin{aligned}\big(\mathbb{E}_x \| T f(x) - f(x)\|^q\big)^{1/q} = \big(\mathbb{E}_{\eta,y}\| \int_{[0,1)^n} \big(G_\eta - F_\eta\big)(y + t)\,\mathrm{d}t\|^q\big)^{1/q} \\ \leqslant \big(\mathbb{E}_{\eta,u}\|G_\eta(u) - F_\eta(u)\|^q\big)^{1/q} \leqslant 2(1 + M(K))\mathscr{E}(f).\end{aligned}$$

To compare T_j and T, use the same U, V in both averages. Since $0 \leqslant U_j < 1$, we have $\lfloor U_j - V_j\rfloor - \lfloor -V_j\rfloor \in \{0,1\}$, and hence

$$e_j + 2\lfloor U^{(j)} - V\rfloor - 2\lfloor U - V\rfloor \in \big\{-e_j, e_j\big\}.$$

Thus their arguments differ by a single coordinate step. Jensen's inequality and translation invariance of the uniform x imply

$$\begin{aligned}\big(\mathbb{E}_x \| T_j f(x) - T f(x)\|^q\big)^{1/q} &\leqslant \big(\mathbb{E}_{x,U,V}\| f\big(x + e_j + 2\lfloor U^{(j)} - V\rfloor\big) - f\big(x + 2\lfloor U - V\rfloor\big)\|^q\big)^{1/q} \\ &= \big(\mathbb{E}_x \| f\big(x + e_j\big) - f(x)\|^q\big)^{1/q},\end{aligned}$$

where either sign of the coordinate step has the same averaged cost. The triangle inequality now gives

$$\begin{aligned}\big(\mathbb{E}_x \| T_j f(x) - f(x)\|^q\big)^{1/q} &\leqslant \big(\mathbb{E}_x \| T_j f(x) - T f(x)\|^q\big)^{1/q} + \big(\mathbb{E}_x \| T f(x) - f(x)\|^q\big)^{1/q} \\ &\leqslant (A - 1)\mathscr{E}(f) + \big(\mathbb{E}_x \| f\big(x + e_j\big) - f(x)\|^q\big)^{1/q},\end{aligned} \tag{5.9}$$

since $A - 1 = 2(1 + M(K))$.

Take β_1 to be the following mixture: with probability $(A-1)/A$, sample $(x, x+\varepsilon)$; with probability $1/A$, sample $(x, x + e_J)$, where J is uniform on $\{1, \dots, n\}$. All these choices are independent. Apply $((A-1)a + b)^q \leqslant A^{q-1}((A-1)a^q + b^q)$ to (5.9) and average over j. This is (5.5). □

Whenever $m > cA$ for a sufficiently large universal constant c, [5, Proposition 4.1] gives $AS \gtrsim_q \sqrt{n}$. For our choice $m = 4\lceil 3n^{1/2+1/q}\rceil$ and $A = 3 + 4\sqrt{n}$, this hypothesis holds for all sufficiently large n, for each fixed q. Thus the construction attains the optimal order in that universal scheme. The elementary continuous bound (5.1) expresses the corresponding limitation of the moment and variation criterion directly.

Appendix A. Sharp metric cotype for K-convex spaces

We give two different continuous-torus proofs of the sharp metric cotype theorem for K-convex Banach space from [8, Theorem 4.1]. We retain the notation $q, C, \mathscr{L}(f), \mathscr{E}(f), \mathbb{T}, F_\eta$ from (1.1), (1.3), and (2.3). The first proof uses the definition of K-convexity directly, with a different convolution kernel for each coordinate. The second proof uses type–cotype duality and one squared-tent product kernel. Pisier's theorem identifies K-convexity with nontrivial Rademacher type, so these are two proofs for the same class of spaces.

A.1. **Direct use of K-convexity.** For $h : \{-1,1\}^n \to X$, its Walsh–Fourier coefficient corresponding to the singleton $\{j\}$ is

$$\hat{h}_{\{j\}} = \mathbb{E}_\delta\big[\delta_j h(\delta)\big],$$

where δ is a uniform sign vector. Define the Rademacher projection by

$$\big(\mathrm{Rad}_n h\big)(\varepsilon) = \sum_{j=1}^{n} \varepsilon_j \hat{h}_{\{j\}}.$$

For $s \geqslant 2$, write

$$\kappa_s(X) = \sup_{n\geqslant 1} \|\mathrm{Rad}_n\|_{L_s(\{-1,1\}^n;X)\to L_s(\{-1,1\}^n;X)}.$$

The space X is *K-convex* if $\kappa_2(X) < \infty$. For $q \geqslant 2$, the Kahane–Khintchine inequality gives

$$\|\mathrm{Rad}_n h\|_{L_q} \leqslant b_q \|\mathrm{Rad}_n h\|_{L_2} \leqslant b_q \kappa_2(X) \|h\|_{L_2} \leqslant b_q \kappa_2(X) \|h\|_{L_q},$$

where b_q depends only on q. Thus $\kappa_q(X) < \infty$ for K-convex X, with

$$\kappa_q(X) \leqslant b_q \kappa_2(X). \tag{A.1}$$

See also [8, § 2].

Theorem A.1 ([8])**.** *Let $2 \leqslant q < \infty$, and suppose that X is K-convex and satisfies* (1.1) *with constant C. For every $n \in \mathbb{N}$, every positive integer m divisible by four, and every $f : \mathbb{Z}_m^n \to X$,*

$$\mathscr{L}(f) \leqslant \big(4n^{1/q} + C\kappa_q(X)m\big)\mathscr{E}(f). \tag{A.2}$$

In particular, for

$$m = 4\lceil n^{1/q}\rceil \leqslant 8n^{1/q},$$

one has $\mathscr{L}(f) \leqslant (1 + C\kappa_q(X))m\mathscr{E}(f)$. Consequently, X has sharp metric cotype q.

Proof. Let F_η be the extension in (2.3), and define

$$K_j(z) = \big(1 - |z_j|\big)_+ \prod_{k\neq j} \frac{1}{2}\mathbf{1}_{[-1,1]}\big(z_k\big), \qquad G_{\eta,j} = F_\eta * K_j, \qquad 1 \leqslant j \leqslant n. \tag{A.3}$$

Each K_j is an unconditional probability density supported in $[-1,1]^n$. Fix j, and let V_j have density K_j, independently of η, u, ε. By convolution, $G_{\eta,j}(u) = \mathbb{E}_{V_j} F_\eta(u - V_j)$. Jensen's inequality and unconditionality, which implies that $\varepsilon \odot V_j$ has the same law as V_j, give

$$\mathbb{E}_{\eta,u}\|F_\eta(u) - G_{\eta,j}(u)\|^q \leqslant \mathbb{E}_{V_j}\mathbb{E}_{\eta,u}\|F_\eta(u) - F_\eta\big(u - V_j\big)\|^q = \mathbb{E}_{V_j}\mathbb{E}_{\eta,u,\varepsilon}\|F_\eta(u) - F_\eta\big(u - \varepsilon\odot V_j\big)\|^q.$$

For each fixed $v \in [-1,1]^n$, Lemma 2.2 gives

$$\big(\mathbb{E}_{\eta,u,\varepsilon}\|F_\eta(u) - F_\eta(u - \varepsilon\odot v)\|^q\big)^{1/q} \leqslant 2\lceil \|v\|_\infty\rceil \mathscr{E}(f) \leqslant 2\mathscr{E}(f),$$

since $-\varepsilon$ is also a uniform sign vector. Averaging the q-th power of this bound over V_j and taking the q-th root proves

$$\big(\mathbb{E}_{\eta,u}\|F_\eta(u)-G_{\eta,j}(u)\|^q\big)^{1/q}\leqslant 2\mathscr{E}(f). \tag{A.4}$$

The weak derivative in coordinate j is

$$\partial_j K_j(z)=-2\operatorname{sgn}\big(z_j\big)2^{-n}\mathbf{1}_{[-1,1]^n}(z). \tag{A.5}$$

Thus the kernels satisfy the regularity assumption of Proposition 3.2.

Let Z be uniform on $[-1,1]^n$, independently of η,u,ε, and put $\delta=(\operatorname{sgn}(Z_1),\dots,\operatorname{sgn}(Z_n))$. Then δ is a uniform sign vector, and (A.5) gives

$$\partial_j G_{\eta,j}(u)=-2\mathbb{E}_Z\big[\delta_j\big(F_\eta(u-Z)-F_\eta(u)\big)\big].$$

For fixed η,u,

$$-\frac{1}{2}\sum_{j=1}^n\varepsilon_j\partial_j G_{\eta,j}(u)=\sum_{j=1}^n\varepsilon_j\mathbb{E}_Z\big[\delta_j\big(F_\eta(u-Z)-F_\eta(u)\big)\big]=\operatorname{Rad}_n\big(\mathbb{E}_Z\big[F_\eta(u-Z)-F_\eta(u)\mid\delta\big]\big).$$

Thus, the Rademacher projection bound and Jensen's inequality give

$$\Big(\mathbb{E}_{\eta,u,\varepsilon}\Big\|\sum_{j=1}^n\varepsilon_j\partial_j G_{\eta,j}(u)\Big\|^q\Big)^{1/q}\leqslant 2\kappa_q(X)\big(\mathbb{E}_{\eta,u,Z}\|F_\eta(u-Z)-F_\eta(u)\|^q\big)^{1/q}\leqslant 4\kappa_q(X)\mathscr{E}(f), \tag{A.6}$$

where the last step is Lemma 2.2, averaged over $(|Z_1|,\dots,|Z_n|)$.

By (A.4) and (A.6), Proposition 3.2 applies with $A=2$ and $S=4\kappa_q(X)$, proving (A.2). Taking $m=4\lceil n^{1/q}\rceil$ gives the asserted sharp scaling. □

A.2. **Using Rademacher type and cotype duality.** For the second proof, put $q'=q/(q-1)$, so that $1<q'\leqslant 2$ and $1/q+1/q'=1$. We write X^* for the space of bounded real-linear functionals $x^*:X\to\mathbb{R}$, with its operator norm, and $\langle x^*,x\rangle=x^*(x)\in\mathbb{R}$.

A Banach space Y has *Rademacher type p*, where $1\leqslant p\leqslant 2$, if there is a finite constant T such that

$$\Big(\mathbb{E}_\varepsilon\Big\|\sum_{j=1}^N\varepsilon_j y_j\Big\|^p\Big)^{1/p}\leqslant T\Big(\sum_{j=1}^N\|y_j\|^p\Big)^{1/p} \tag{A.7}$$

for every finite family (y_j) in Y. The least such constant is denoted by $T_p(Y)$. Type is *nontrivial* if $T_p(Y)<\infty$ for some $p>1$. The Rademacher projection and the constants $\kappa_s(X)$ were introduced in § A.1. Pisier's theorem states that X has nontrivial Rademacher type if and only if it is K-convex [9]; see also [8, § 2] and [6]. Together with (A.1), this gives $\kappa_q(X)<\infty$ whenever X has nontrivial type.

The next lemma records the classical duality implication from cotype q of a K-convex space to type q' of its dual; see Pisier [10] and [6] for the type–cotype duality theory. We include the finite-cube proof to keep track of the constants and the exponent.

Lemma A.2. *If X satisfies* (1.1) *with constant C and $\kappa_q(X)<\infty$, then X^* has type $q'=q/(q-1)$, with*

$$T_{q'}\big(X^*\big)\leqslant C\kappa_q(X). \tag{A.8}$$

Proof. For $h:\{-1,1\}^N\to X$, cotype and the definition of $\kappa_q(X)$ give

$$\Big(\sum_{j=1}^N\|\hat h_{\{j\}}\|^q\Big)^{1/q}\leqslant C\|\operatorname{Rad}_N h\|_{L_q(X)}\leqslant C\kappa_q(X)\|h\|_{L_q(X)}. \tag{A.9}$$

Now fix $x_1^*,\dots,x_N^*\in X^*$. By the definition of $\hat h_{\{j\}}$, followed by Hölder's inequality for the finite sum and (A.9),

$$\Big|\mathbb{E}_\delta\Big\langle\sum_{j=1}^N\delta_j x_j^*,h(\delta)\Big\rangle\Big|=\Big|\sum_{j=1}^N\langle x_j^*,\hat h_{\{j\}}\rangle\Big|$$
$$\leqslant\Big(\sum_{j=1}^N\|x_j^*\|^{q'}\Big)^{1/q'}\Big(\sum_{j=1}^N\|\hat h_{\{j\}}\|^q\Big)^{1/q}\leqslant C\kappa_q(X)\Big(\sum_{j=1}^N\|x_j^*\|^{q'}\Big)^{1/q'}\|h\|_{L_q(X)}.$$

To compute the supremum over $\|h\|_{L_q(X)}\leqslant 1$, put

$$g(\delta)=\sum_{j=1}^N\delta_j x_j^*,\qquad L=\big(\mathbb{E}_\delta\|g(\delta)\|^{q'}\big)^{1/q'}.$$

Hölder's inequality gives

$$|\mathbb{E}_\delta\langle g(\delta),h(\delta)\rangle|\leqslant\big(\mathbb{E}_\delta\|g(\delta)\|^{q'}\big)^{1/q'}\big(\mathbb{E}_\delta\|h(\delta)\|^q\big)^{1/q}\leqslant L\qquad\big(\|h\|_{L_q(X)}\leqslant 1\big).$$

Assume without loss of generality that $L>0$. For $0<\theta<1$, the definition of the dual norm allows us to choose, for each of the finitely many δ, a vector $u_\delta\in X$, changing its sign if necessary, such that

$$\|u_\delta\|\leqslant 1,\qquad\langle g(\delta),u_\delta\rangle\geqslant(1-\theta)\|g(\delta)\|.$$

Take $u_\delta=0$ when $g(\delta)=0$, and define

$$h_\theta(\delta)=\frac{\|g(\delta)\|^{q'-1}}{L^{q'-1}}u_\delta.$$

Since $(q'-1)q=q'$, this choice satisfies

$$\|h_\theta\|_{L_q(X)}^q\leqslant\frac{\mathbb{E}_\delta\|g(\delta)\|^{(q'-1)q}}{L^{(q'-1)q}}=\frac{\mathbb{E}_\delta\|g(\delta)\|^{q'}}{L^{q'}}=1,$$

and

$$\mathbb{E}_\delta\langle g(\delta),h_\theta(\delta)\rangle=\frac{\mathbb{E}_\delta\big[\|g(\delta)\|^{q'-1}\langle g(\delta),u_\delta\rangle\big]}{L^{q'-1}}\geqslant(1-\theta)\frac{\mathbb{E}_\delta\|g(\delta)\|^{q'}}{L^{q'-1}}=(1-\theta)L.$$

Letting $\theta\downarrow 0$ proves the identity

$$\sup_{\|h\|_{L_q(X)}\leqslant 1}\Big|\mathbb{E}_\delta\Big\langle\sum_{j=1}^N\delta_j x_j^*,h(\delta)\Big\rangle\Big|=L=\Big(\mathbb{E}_\delta\Big\|\sum_{j=1}^N\delta_j x_j^*\Big\|^{q'}\Big)^{1/q'}.$$

Taking this supremum in the preceding coefficient bound gives the type inequality (A.7) for X^*, with the constant in (A.8). □

Lemma A.3. *Suppose that $T_{q'}(X^*)\leqslant T$. Let $\xi_1,\dots,\xi_N$ be independent symmetric real random variables in $L_{q'}$ on a probability space, and let Y be a finite-valued X-valued random variable on that same probability space. Then, for every $a\in X$,*

$$\Big(\sum_{j=1}^N\|\mathbb{E}\big[\xi_j Y\big]\|^q\Big)^{1/q}\leqslant T\max_j\|\xi_j\|_{L_{q'}}\big(\mathbb{E}\|Y-a\|^q\big)^{1/q}.\tag{A.10}$$

No independence between Y and the family (ξ_j) is required.

Proof. Symmetry means that ξ_j and $-\xi_j$ have the same law. Together with independence, this implies that inserting an independent sign in every coordinate preserves the joint law of $(\xi_1,\dots,\xi_N)$. Hence, for $x_j^*\in X^*$,

$$\mathbb{E}\Big\|\sum_{j=1}^N\xi_j x_j^*\Big\|^{q'}=\mathbb{E}_\xi\mathbb{E}_\varepsilon\Big\|\sum_{j=1}^N\varepsilon_j\xi_j x_j^*\Big\|^{q'}\leqslant T^{q'}\sum_{j=1}^N\mathbb{E}\big[|\xi_j|^{q'}\big]\|x_j^*\|^{q'}\leqslant T^{q'}\max_j\|\xi_j\|_{L_{q'}}^{q'}\sum_{j=1}^N\|x_j^*\|^{q'}.$$

Since $q' > 1$, symmetry and integrability imply $\mathbb{E}\xi_j = 0$. We may therefore replace Y by $Y - a$ in every coefficient. Scalar Hölder's inequality and the preceding bound give

$$\Big|\sum_{j=1}^{N} \langle x_j^*, \mathbb{E}[\xi_j Y]\rangle\Big| = \Big|\mathbb{E}\Big\langle \sum_{j=1}^{N} \xi_j x_j^*, Y - a\Big\rangle\Big|$$
$$\leqslant \big(\mathbb{E}\|Y-a\|^q\big)^{1/q}\Big(\mathbb{E}\Big\|\sum_{j=1}^{N}\xi_j x_j^*\Big\|^{q'}\Big)^{1/q'} \leqslant T\max_j \|\xi_j\|_{L_{q'}}\big(\mathbb{E}\|Y-a\|^q\big)^{1/q}\Big(\sum_{j=1}^{N}\|x_j^*\|^{q'}\Big)^{1/q'}.$$

To justify the duality step explicitly, put

$$v_j = \mathbb{E}\big[\xi_j Y\big], \qquad B = \Big(\sum_{j=1}^{N}\|v_j\|^q\Big)^{1/q}.$$

These expectations exist because Y has finite range and the ξ_j are integrable. For any $x_1^*, \dots, x_N^* \in X^*$ with $\sum_j \|x_j^*\|^{q'} \leqslant 1$, Hölder's inequality gives

$$\Big|\sum_{j=1}^{N}\langle x_j^*, v_j\rangle\Big| \leqslant \Big(\sum_{j=1}^{N}\|x_j^*\|^{q'}\Big)^{1/q'}\Big(\sum_{j=1}^{N}\|v_j\|^q\Big)^{1/q} \leqslant B.$$

We now exhibit functionals for which equality holds. If $B = 0$, all the v_j are zero and the assertion is immediate. If $B > 0$, the real Hahn–Banach theorem gives $\phi_j \in X^*$ such that

$$\|\phi_j\| \leqslant 1, \qquad \langle \phi_j, v_j\rangle = \|v_j\|;$$

take $\phi_j = 0$ when $v_j = 0$. Define

$$x_j^* = \frac{\|v_j\|^{q-1}}{B^{q-1}}\phi_j.$$

Since $(q-1)q' = q$, these functionals satisfy

$$\sum_{j=1}^{N}\|x_j^*\|^{q'} \leqslant \frac{\sum_{j=1}^{N}\|v_j\|^{(q-1)q'}}{B^{(q-1)q'}} = \frac{\sum_{j=1}^{N}\|v_j\|^q}{B^q} = 1,$$

whereas

$$\sum_{j=1}^{N}\langle x_j^*, v_j\rangle = \frac{\sum_{j=1}^{N}\|v_j\|^{q-1}\langle\phi_j, v_j\rangle}{B^{q-1}} = \frac{\sum_{j=1}^{N}\|v_j\|^q}{B^{q-1}} = B.$$

Consequently,

$$\sup_{\sum_{j=1}^{N}\|x_j^*\|^{q'}\leqslant 1}\Big|\sum_{j=1}^{N}\langle x_j^*, v_j\rangle\Big| = \Big(\sum_{j=1}^{N}\|v_j\|^q\Big)^{1/q}.$$

Taking this supremum in the preceding estimate proves (A.10). □

Theorem A.4. *Let X be a Banach space such that X^* has Rademacher type $q' = q/(q-1)$, and put $T = T_{q'}(X^*)$. For every $n \in \mathbb{N}$, every $m \in 4\mathbb{N}$, and every $f : \mathbb{Z}_m^n \to X$,*

$$\mathscr{L}(f) \leqslant \big(4n^{1/q} + \sqrt{3}\,Tm\big)\mathscr{E}(f). \tag{A.11}$$

Consequently $m = 4\lceil n^{1/q}\rceil \leqslant 8n^{1/q}$ is admissible for sharp metric cotype q, with $\Gamma = 1 + \sqrt{3}\,T$.

Proof. Use the extension F_η from (2.3) and the normalized torus averages. Set

$$\rho(t) = \tfrac{3}{2}\big(1-|t|\big)_+^2, \qquad K(z) = \prod_{j=1}^n \rho\big(z_j\big), \qquad G_\eta(u) = \big(F_\eta * K\big)(u) = \mathbb{E}_Z F_\eta(u-Z), \tag{A.12}$$

where Z has density K and is independent of η, u. Since ρ is even and supported on $[-1,1]$, the law of Z is unconditional and supported on $[-1,1]^n$. Inserting an independent uniform sign vector into Z and applying (2.4) for each fixed $Z = z$ gives

$$\mathbb{E}_{\eta,u,Z}\|F_\eta(u-Z) - F_\eta(u)\|^q = \mathbb{E}_Z\mathbb{E}_{\eta,u,\varepsilon}\|F_\eta(u-\varepsilon\odot Z) - F_\eta(u)\|^q \leqslant 2^q\mathscr{E}\big(f\big)^q. \tag{A.13}$$

Here $\lceil\|z\|_\infty\rceil \leqslant 1$ on the support of K, and $-\varepsilon$ has the same law as ε. Jensen's inequality for the expectation over Z now gives

$$\|F - G\|_{L_q(\eta,u)} \leqslant 2\mathscr{E}(f). \tag{A.14}$$

We next justify differentiation of G_η. The density ρ is absolutely continuous on $\mathbb{R}$, with

$$\rho'(t) = -3\operatorname{sgn}(t)(1-|t|) \quad (0<|t|<1),$$

and derivative zero outside $[-1,1]$, almost everywhere. Consequently $\partial_j K(z) = \rho'(z_j)\prod_{k\neq j}\rho(z_k)$ belongs to $L_1(\mathbb{R}^n)$. The fundamental theorem of calculus and continuity of translations in L_1 show that the coordinate difference quotients of K converge to $\partial_j K$ in L_1. Since F_η is bounded, convolution turns this convergence into uniform convergence. Moreover, $F_\eta * \partial_j K$ is continuous, because its change under a translation is bounded by $\|F_\eta\|_\infty$ times the L_1 change of the translated $\partial_j K$. Thus G_η is continuously differentiable and

$$\partial_j G_\eta(u) = \int_{\mathbb{R}^n} F_\eta(u-z)\partial_j K(z)\,\mathrm{d}z.$$

These vector integrals lie in the finite-dimensional span of the range of f.

Define the score variables by

$$s_j(Z) = \frac{\rho'\big(Z_j\big)}{\rho\big(Z_j\big)} = -\frac{2\operatorname{sgn}\big(Z_j\big)}{1-|Z_j|} \quad \big(0<|Z_j|<1\big),$$

setting their values to zero on the remaining null sets. They are independent, since the coordinates of Z are independent, and symmetric, since ρ is even and the displayed score is odd. Direct integration gives

$$\mathbb{E}s_j = 0, \qquad \mathbb{E}|s_j|^2 = \int_{-1}^1 \frac{\rho'\big(t\big)^2}{\rho(t)}\,\mathrm{d}t = \int_{-1}^1 6\,\mathrm{d}t = 12.$$

In particular, $q' \leqslant 2$ implies $\|s_j\|_{L_{q'}} \leqslant \sqrt{12}$. The derivative formula and $\mathbb{E}s_j = 0$ imply

$$\partial_j G_\eta(u) = \mathbb{E}_Z\big[s_j(Z)F_\eta(u-Z)\big] = \mathbb{E}_Z\big[s_j(Z)\big(F_\eta(u-Z) - F_\eta(u)\big)\big]. \tag{A.15}$$

For fixed η, u, apply Lemma A.3 on the probability space of Z, with $\xi_j = s_j(Z)$, $Y = F_\eta(u-Z)$, and $a = F_\eta(u)$. It follows that

$$\Big(\sum_{j=1}^n \|\partial_j G_\eta(u)\|^q\Big)^{1/q} \leqslant \sqrt{12}\,T\big(\mathbb{E}_Z\|F_\eta(u-Z) - F_\eta(u)\|^q\big)^{1/q}.$$

Raise this inequality to the qth power, average over η, u, and use (A.13) to get

$$\Big(\sum_{j=1}^n \mathbb{E}_{\eta,u}\|\partial_j G_\eta(u)\|^q\Big)^{1/q} \leqslant 2\sqrt{12}\,T\mathscr{E}(f). \tag{A.16}$$

To finish, use only the endpoint integration steps in the proof of Proposition 3.2, without applying cotype again. Specifically, (3.9) with $G_{\eta,j} = G_\eta$ and the first two steps of (3.10) give

$$\mathscr{L}(f) \leqslant 2n^{1/q}\big(\mathbb{E}_{\eta,u}\|F_\eta(u) - G_\eta(u)\|^q\big)^{1/q} + \frac{m}{4}\Big(\sum_{j=1}^n \mathbb{E}_{\eta,u}\|\partial_j G_\eta(u)\|^q\Big)^{1/q} \leqslant \big(4n^{1/q} + \sqrt{3}\,Tm\big)\mathscr{E}(f),$$

where the last step uses (A.14) and (A.16). This proves (A.11). Finally, for $m = 4\lceil n^{1/q}\rceil$, the term $4n^{1/q}$ is at most m, and $\lceil n^{1/q}\rceil \leqslant 2n^{1/q}$. □

Corollary A.5 (Mendel–Naor)**.** *Let X be a Banach space with nontrivial Rademacher type satisfying* (1.1) *with constant C. Then X has sharp metric cotype q. More precisely, for every $n \in \mathbb{N}$, $m \in 4\mathbb{N}$, and $f : \mathbb{Z}_m^n \to X$,*

$$\mathscr{L}(f) \leqslant \left(4n^{1/q} + \sqrt{3}\,C\kappa_q(X)m\right)\mathscr{E}(f). \tag{A.17}$$

One may take $m = 4\lceil n^{1/q}\rceil \leqslant 8n^{1/q}$ and $\Gamma = 1 + \sqrt{3}\,C\kappa_q(X)$.

Proof. Pisier's theorem and (A.1) imply $\kappa_q(X) < \infty$. By Lemma A.2, $T_{q'}(X^*) \leqslant C\kappa_q(X)$. Substitute this bound in Theorem A.4. □

References

[1] Alexandr Andoni, Assaf Naor, and Ofer Neiman. "Snowflake universality of Wasserstein spaces". In: *Ann. Sci. Éc. Norm. Supér. (4)* 51.3 (2018), pp. 657–700. DOI: 10.24033/asens.2363.

[2] Q. Cheng, L. Sun, and Y. Zhang. "The optimal scaling parameter in metric cotype for Alexandrov spaces of nonnegative curvature". Preprint. 2026. arXiv: 2609.08754 [math.FA].

[3] Qingjin Cheng, Yue Wang, and Bo Xiang. "Sharp metric cotype inequalities for L_1 via nonlinear cut smoothing". Preprint. 2026. arXiv: 2609.08749 [math.FA].

[4] Alexandros Eskenazis, Manor Mendel, and Assaf Naor. "Nonpositive curvature is not coarsely universal". In: *Invent. Math.* 217.3 (2019), pp. 833–886. DOI: 10.1007/s00222-019-00878-1.

[5] Ohad Giladi, Manor Mendel, and Assaf Naor. "Improved bounds in the metric cotype inequality for Banach spaces". In: *J. Funct. Anal.* 260.1 (2011), pp. 164–194. DOI: 10.1016/j.jfa.2010.08.015. arXiv: 1003.0279 [math.FA].

[6] Tuomas Hytönen, Jan van Neerven, Mark Veraar, and Lutz Weis. *Analysis in Banach Spaces. Volume II: Probabilistic Methods and Operator Theory*. Vol. 67. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. Cham: Springer, 2017.

[7] Yong Jiao, Sijie Luo, and Lian Wu. "The sharp scale inequality for the Schatten trace class". Preprint. 2026. arXiv: 2610.21445 [math.FA].

[8] Manor Mendel and Assaf Naor. "Metric cotype". In: *Ann. of Math. (2)* 168.1 (2008), pp. 247–298. DOI: 10.4007/annals.2008.168.247.

[9] Gilles Pisier. "Holomorphic semi-groups and the geometry of Banach spaces". In: *Ann. of Math. (2)* 115.2 (1982), pp. 375–392. DOI: 10.2307/1971396.

[10] Gilles Pisier. "On the duality between type and cotype". In: *Martingale Theory in Harmonic Analysis and Banach Spaces*. Vol. 939. Lecture Notes in Mathematics. Berlin: Springer, 1982, pp. 131–144. DOI: 10.1007/BFb0096265.

[11] E. Veomett and K. Wildrick. "Spaces of small metric cotype". In: *J. Topol. Anal.* 2.4 (2010), pp. 581–597. DOI: 10.1142/S1793525310000422. arXiv: 1001.3326 [math.MG].

(M. M.) Mathematics and CS Department, The Open University of Israel, 1 University Road, Raanana 43107, Israel

Email address: manorme@openu.ac.il