METRIC $X_p$ INEQUALITIES

ASSAF NAOR AND GIDEON SCHECHTMAN

ABSTRACT. We show that if $m, n \in \mathbb{N}$ and $k \in \{1, \ldots, n\}$ satisfy $m \geq n^{3/2}/\sqrt{k}$ then for every $p \in [2, \infty)$ and $f : \mathbb{Z}_{4m}^n \to \mathbb{R}$ we have

$$\frac{1}{\binom{n}{k}} \sum_{\substack{S \subseteq \{1, \ldots, n\} \atop |S| = k}} \mathbb{E} \left[ \left| f(x + 2m \sum_{j \in S} \varepsilon_j e_j) - f(x) \right|^p \right] \leq \frac{k}{n} \sum_{j=1}^n \mathbb{E} \left[ \left| f(x + e_j) - f(x) \right|^p \right] + \left( \frac{k}{n} \right)^{p/2} \mathbb{E} \left[ \left| f(x + \varepsilon) - f(x) \right|^p \right],$$

(1)

where the expectation is with respect to $(x, \varepsilon) \in \mathbb{Z}_{4m}^n \times \{-1, 1\}^n$ chosen uniformly at random and $e_1, \ldots, e_n$ is the standard basis of $\mathbb{Z}_{4m}^n$. Inequality (1) is a nonlinear extension of a linear inequality for Rademacher sums that was proved by Johnson, Maurey, Schechtman and Tzafriri in 1979. We show that for (1) to hold true it is necessary that $m$ tends to infinity with $n$. The formulation (and proof) of (1) completes the long-standing search for bi-Lipschitz invariants that serve as an obstruction to the nonembeddability of $L_p$ spaces into each other, the previously understood cases of which were metric notions of type and cotype, which fail to certify the nonembeddability of $L_p$ into $L_q$ when $2 < q < p$. Among the consequences of (1) are new quantitative restrictions on the bi-Lipschitz embeddability into $L_q$ of snowflakes of $L_q$ and integer grids in $\ell^q_p$, for $2 < q < p$. As a byproduct of our investigations, we also obtain results on the geometry of the Schatten $p$ trace class $S_p$ that are new even in the linear setting.

1. INTRODUCTION

The ensuing discussion uses standard notation and terminology from Banach space theory, as in [31]. In particular, for $p \in [1, \infty]$ and $n \in \mathbb{N}$, the space $\ell^p_n$ (respectively $\ell^p_n(\mathbb{C})$) denotes the vector space $\mathbb{R}^n$ (respectively $\mathbb{C}^n$), equipped with the standard $\ell^p$ norm, and the space $L_p$ refers to the Lebesgue function space $L_p(0, 1)$. Banach spaces are assumed to be over real scalars unless stated otherwise, though our results hold true mutatis mutandis for complex Banach spaces as well.

We shall also use standard notation and terminology from the theory of metric embeddings, as in [34, 40]. In particular, a metric space $(X, d_X)$ is said to admit a bi-Lipschitz embedding into a metric space $(Y, d_Y)$ if there exists $s \in (0, \infty)$, $D \in [1, \infty)$ and a mapping $f : X \to Y$ such that

$$\forall x, y \in X, \quad s d_X(x, y) \leq d_Y(f(x), f(y)) \leq D s d_X(x, y)$$

(2)

When this happens we say that $(X, d_X)$ embeds into $(Y, d_Y)$ with distortion at most $D$. Given $f : X \to Y$, the infimum over those $D \in [1, \infty)$ for which there exists $s \in (0, \infty)$ such that (2) holds true is called the distortion of $f$ and is denoted $\text{dist}(f)$. If no such $D$ exists set $\text{dist}(f) = \infty$. We denote by $c_{(Y,d_Y)}(X, d_X)$ (or simply $c_Y(X)$ if the metrics are clear from the context) the infimum over those $D \in [1, \infty]$ for which $(X, d_X)$ embeds into $(Y, d_Y)$ with distortion at most $D$. If $(X, d_X)$ does not admit a bi-Lipschitz embedding into $(Y, d_Y)$ then we set $c_{(Y,d_Y)}(X, d_X) = \infty$. When $Y = L_p$ we use the shorter notation $c_{L_p}(X, d_X) = c_p(X, d_X)$.

Among the simplest and most basic questions that one could ask in the context of metric embeddings is to determine those $p, q \in [1, \infty)$ for which $L_q$ admits a bi-Lipschitz embedding into

A. N. was supported by the NSF, the BSF, the Packard Foundation and the Simons Foundation.
G. S. was supported by the ISF and the BSF.
that if Banach spaces. Chapter 12 of Banach’s book [6] is devoted to this topic. Banach proved there
that if \( L_q \) is isomorphic to a subspace of \( L_p \) then necessarily either \( p < q \leq 2 \) or \( 2 \leq q < p \), and that \( L_2 \) is isomorphic to a subspace of \( L_p \) for all \( p \in [1, \infty) \). Banach also conjectured [6, page 205] that \( L_q \) is isomorphic to a subspace of \( L_p \) if \( p < q < 2 \) or \( 2 < q < p \). In the range \( p < q < 2 \), Banach’s question was answered affirmatively by Kadec [27], who showed that in this case \( L_q \) is linearly isometric to a subspace of \( L_p \). When \( 2 < q < p \), Banach’s question was answered negatively by Paley [47], i.e., \( L_q \) is not isomorphic to a subspace of \( L_p \) when \( 2 < q < p \). Our goal here is to obtain a nonlinear version of Paley’s theorem, i.e., the formulation of a bi-Lipschitz invariant that serves as an obstruction to the embeddability \( L_q \) into \( L_p \) when \( 2 < q < p \). This invariant allows us to obtain nonembeddability results that were beyond the reach of previously available methods, and in addition it leads to interesting open questions. Our new invariant thus completes a long line of work on the bi-Lipschitz classification of \( L_p \) spaces, because the remaining cases, namely the bi-Lipschitz nonembeddability of \( L_q \) into \( L_p \) when either \( q \in [1, 2) \) and \( p > q \), or \( q \in (2, \infty) \) and \( p < q \), were previously understood through notions of metric type and cotype that were introduced over the past four decades (see below for more on this topic).

Our main result is the following theorem, in which (and in what follows) for every \( n \in \mathbb{N} \) and \( j \in \{1, \ldots, n\} \) we let (as usual) \( e_j \in \mathbb{R}^n \) denote the vector whose \( j \)th coordinate equals 1 and the rest of whose coordinates vanish, and for \( S \subseteq \{1, \ldots, n\} \) and \( \varepsilon \in \{-1, 1\}^n \) we denote \( \varepsilon_S = \sum_{j \in S} \varepsilon_j e_j \).

**Theorem 1.1** (Metric \( X_p \) inequality). Fix \( p \in [2, \infty) \). Suppose that \( m, n \in \mathbb{N} \) and \( k \in \{1, \ldots, n\} \) satisfy

\[
m \geq \frac{n^{3/2} \log p}{\sqrt{k}} + pn.
\]

Then for every \( f : \mathbb{Z}_{4m}^n \to L_p \) we have

\[
\left( \frac{p}{\log p} \right)^{-p} \frac{1}{\binom{n}{k}} \sum_{S \subseteq \{1, \ldots, n\}, |S| = k} \mathbb{E} \left[ \| f(x + 2m\varepsilon_S) - f(x) \|_p^p \right] \leq_p \frac{k}{n} \sum_{j=1}^n \mathbb{E} \left[ \| f(x + e_j) - f(x) \|_p^p \right] + \left( \frac{k}{n} \right)^2 \mathbb{E} \left[ \| f(x + \varepsilon) - f(x) \|_p^p \right], \tag{3}
\]

where the expectation is with respect to \( (x, \varepsilon) \in \mathbb{Z}_{4m}^n \times \{-1, 1\}^n \) chosen uniformly at random.

**Asymptotic notation.** In Theorem 1.1 and in what follows, we use the (somewhat nonstandard) convention that for \( a, b \in [0, \infty) \) and \( p \in [1, \infty) \) the notation \( a \lesssim_p b \) (respectively \( a \gtrsim_p b \)) stands for \( a \leq c^p b \) (respectively \( a \geq c^p b \)) for some universal constant \( c \in (0, \infty) \). The notation \( a \lesssim b \) (respectively \( a \gtrsim b \)) stands for \( a \leq cb \) (respectively \( a \geq cb \)) for some universal constant \( c \in (0, \infty) \). The notation \( a \asymp b \) stands for \( (a \lesssim b) \land (b \lesssim a) \). At times our discussion will be in the presence of an auxiliary Banach (or metric) space \( X \), in which case the notation \( a \lesssim_X b \) will stand for \( a \leq c(X)b \), where \( c(X) \in (0, \infty) \) is allowed to depend only on \( X \) (in fact, \( c(X) \) will always depend on certain numerical geometric invariants of \( X \) that will be clear from the context).

The term \( p/\log p \) in the left-hand side of (3) is sharp up to a universal constant factor. We defer the explanation of why (3) is called a metric \( X_p \) inequality to the ensuing discussion. Note that
since \([3]\) involves the \(p\)’th power of \(L_p\) norms, it suffices to prove its validity when \(f\) is real-valued, but we stated Theorem \([3]\) for functions with values in \(L_p\) since this is the way by which we will apply it to prove new nonembeddability results. The fact that in Theorem \([3]\) the function \(f\) is assumed to be defined on the discrete torus \(\mathbb{Z}_{4m}^n\) rather than on \(\mathbb{Z}_m^n\) is not important: for notational reasons it is beneficial to work with \(\mathbb{Z}_m^n\) when the modulus \(m\) is divisible by 4, and this suffices for all of the applications of \([3]\) that we can imagine. However, it is straightforward to modify our proof of Theorem \([1.1]\) so as to obtain variants of \([3]\) for functions defined on discrete tori whose modulus is not necessarily divisible by 4.

**Remark 1.2.** If one makes the weaker assumption \(m \geq n^{3/2}/\sqrt{K}\) in Theorem \([1.1]\) then \([3]\) holds true with the (sharp) term \(p/\log p\) in the left-hand side replaced by \(p^2/\log p\). This, and additional tradeoffs of this type, can be deduced from an inspection of our proof of Theorem \([1.1]\).

### 1.1. Quantitative nonembeddability.

The above classification of those \(p, q \in [1, \infty)\) for which \(L_q \) admits a bi-Lipschitz embedding into \(L_p\) is based on an abstract reduction to linear embeddings, and as such it fails to yield a metric invariant that serves as an obstruction to bi-Lipschitz embeddings. This argument also does not imply various quantitative estimates that are inherently nonlinear and cannot be deduced from the linear theory. For example, given a metric space \((X, d_X)\) and \(\theta \in (0, 1]\), the \(\theta\)-snowflake of \((X, d_X)\) is defined (see e.g. [14]) to be the metric space \((X, d_X^\theta)\). A natural quantitative refinement of the assertion that \(L_q\) does not admit a bi-Lipschitz embedding into \(L_p\) is that if the \(\theta\)-snowflake of \(L_q\) admits a bi-Lipschitz embedding into \(L_p\) then necessarily \(\theta\) must be bounded away from 1 by a definite constant (depending on \(p, q\)). While such statements are known (through the theory of metric type and cotype; see below) when either \(q \in [1, 2)\) and \(p > q\), or \(p \in (2, \infty)\) and \(q > p\), in the range \(2 < q < p\) no such quantitative refinement of bi-Lipschitz nonembeddability was previously known. For \(2 < p < q\), in Theorem \([1.6]\) below we obtain, as a consequence of Theorem \([1.1]\) an explicit \(\delta(p, q) \in (0, 1)\) such that if the \(\theta\)-snowflake of \(L_q\) admits a bi-Lipschitz embedding into \(L_p\) then necessarily \(\theta \leq 1 - \delta(p, q)\). In Section \([6]\) we formulate a conjectural convolution inequality that is shown to yield the sharp value \(\delta(p, q)\) in this context. Since H"older mappings need not be differentiable anywhere, and moreover continuous linear mappings are necessarily Lipschitz, it seems impossible to obtain a restriction on those snowflakes of \(L_q\) that embed into \(L_p\) via a reduction to linear embeddings as above.

Another natural quantitative refinement of the bi-Lipschitz nonembeddability of \(L_q\) into \(L_p\) is, given \(m, n \in \mathbb{N}\), to ask for a lower bound on \(c_p(|m|_q^n)\), where here, and in what follows, \(|m|_q^n\) denotes the grid \(\{0, \ldots, m\}^n \subseteq \mathbb{R}^n\), equipped with the metric inherited from \(c_q^n\). While such an estimate can be obtained from general principles, namely Bourgain’s discretization theorem \([9]\) (see Remark \([3.2]\) below), in Theorem \([1.9]\) we obtain, as a consequence of Theorem \([1.1]\) the best known lower bound on \(c_p(|m|_q^n)\) when \(2 < q < p\). The convolution inequality that is conjectured in Section \([6]\) is shown to imply an asymptotically sharp evaluation of \(c_p(|m|_q^n)\), exhibiting a striking phase transition when \(m \asymp n^{(p-q)/(q(p-2))}\); see Theorem \([1.12]\) below.

### 1.2. Local invariants.

Suppose that \(p, q \in [1, \infty)\) are such that \(L_q\) does not admit a bi-Lipschitz embedding into \(L_p\). This assertion is local in the sense that the smallest possible distortion of a linear embedding of \(c_q^n\) into \(L_p\) tends to \(\infty\) with \(n\). Thus, there is a finite dimensional linear obstruction (which will be stated explicitly in Section \([1.3]\) below) showing that no \(n\)-dimensional subspace of \(L_p\) can be close to \(c_q^n\). An important rigidity theorem of Ribe \([53]\) suggests that such finite dimensional linear obstructions can be reformulated while only referring to distances between pairs of points. This is the basis for the Ribe program \([8, 40, 4]\), and our work constitutes a completion of this program for \(L_p\) spaces, the previously missing case being when \(2 < q < p\). The next step in the Ribe program, a step that has proven in the past to be useful for various questions in metric geometry, would be to study metric spaces that satisfy \([3]\). However, unlike previous
advances in the Ribe program, in the present setting it seems that it is more natural for the linear theory to be developed further before its metric counterpart is investigated; we discuss this matter and formulate some related open problems in Section 1.5 below.

1.3. Type, cotype and symmetric structures. For \( r, s \in [1, \infty) \), a Banach space \((X, \| \cdot \|_X)\) is said to have Rademacher type \( r \) and cotype \( s \) if for every \( n \in \mathbb{N} \) and \( x_1, \ldots, x_n \in X \) we have

\[
\mathbb{E} \left[ \left\| \sum_{j=1}^n \varepsilon_j x_j \right\|_X^r \right] \lesssim_X \sum_{j=1}^n \|x_j\|_X^r \quad \text{and} \quad \sum_{j=1}^n \|x_j\|_X^s \lesssim_X \mathbb{E} \left[ \left\| \sum_{j=1}^n \varepsilon_j x_j \right\|_X^s \right],
\]

where the expectation is with respect to \( \varepsilon \in \{-1, 1\}^n \) chosen uniformly at random. See [35] and the references therein for more on these important notions. It suffices to say here that if \( T : \ell_q^n \to L_p \) is an invertible linear operator then necessarily

\[
\text{dist}(T) = \|T\| \cdot \|T^{-1}\| \gtrsim \kappa(p) \cdot \begin{cases} \frac{n^{1/2}}{2} & \text{if } 1 \leq q \leq p \leq 2, \\ \frac{n^{1/2}}{2} & \text{if } 1 \leq q \leq 2 \leq p < \infty, \\ \frac{n^{1/2}}{2} & \text{if } 2 \leq p \leq q, \\ \frac{n^{1/2}}{2} & \text{if } 1 \leq p \leq 2 \leq q. \end{cases}
\]

(5) follows from an application of (4) with \( X = L_p, \ r = \min \{p, 2\}, \ s = \max \{p, 2\} \) and \( x_j = T e_j \).

The bounds in (5) cannot be improved up to the value of \( \kappa(p) \). Thus, type and cotype constitute the finite dimensional linear invariants that were alluded to in Section 1.2, i.e., they certify (in a sharp way) that if either \( q \in [1, 2) \) and \( p > q \) or \( p \in (2, \infty) \) and \( q > p \), then any linear embedding of \( \ell_q^n \) into \( L_p \) incurs large distortion.

The usefulness of the notions of Rademacher type and cotype goes far beyond their relevance to embeddings of \( L_p \) spaces. For this reason (in addition to the intrinsic geometric interest arising from the Ribe program) there has been considerable effort to reformulate these notions while using only distances between pairs of points rather than linear combinations of vectors as in (4), thereby understanding when a metric space has type \( r \) and cotype \( s \). We will quickly recall now a very small part of what is known in this direction, stating only those results that are needed for the present discussion on metric \( X_p \) inequalities.

Following Enflo [17], a metric space \((X, d_X)\) is said to have Enflo type \( r \in [1, \infty) \) if for every \( n \in \mathbb{N} \) and \( f : \{-1, 1\}^n \to X \),

\[
\mathbb{E} [d_X(f(\varepsilon), f(-\varepsilon))] \lesssim_X \sum_{j=1}^n \mathbb{E} [d_X(f(\varepsilon), f(\varepsilon_1, \ldots, \varepsilon_{j-1}, -\varepsilon_j, \varepsilon_{j+1}, \ldots, \varepsilon_n))] \quad \text{for } \varepsilon \in \{-1, 1\}^n \text{ chosen uniformly at random.}
\]

where the expectation is with respect to \( \varepsilon \in \{-1, 1\}^n \) chosen uniformly at random. Note that if \( X \) is a Banach space then (6) coincides with the leftmost inequality in (4) when \( f \) is the linear function given by \( f(\varepsilon) = \sum_{j=1}^n \varepsilon_j x_j \). For \( p \in [1, \infty) \), \( L_p \) actually has Enflo type \( r = \min \{p, 2\} \), i.e., \( X = L_p \) satisfies (6) with \( f : \{-1, 1\}^n \to L_p \) allowed to be an arbitrary mapping rather than only a linear mapping. This statement was first proved for \( p \in [1, 2] \) in [16] and for \( p \in (2, \infty) \) in [43].

One is tempted to define when a metric space \((X, d_X)\) has cotype \( s \in (0, \infty) \) by reversing the inequality in (6) (with \( r \) replaced by \( s \)). But, note that if \( d_X(f(\varepsilon), f(\delta)) = 1 \) for every \( \varepsilon, \delta \in \{-1, 1\}^n \) (this can occur even if \( X \) is a Hilbert space) then the right-hand side of (6) grows linearly with \( n \) as \( n \to \infty \), while the left had side of (6) remains bounded. Thus, there are truly nonlinear phenomena that do not occur in the linear setting of Rademacher cotype which do not allow for the straightforward reversal of the inequality in (6).

In essence, the total mass of the measure that appears in the right-hand side of (6) is too large in comparison to the total mass of the measure...
that appears in the left-hand side of (6) for an inequality that is the reverse of (6) to make any sense even in Hilbert space (indeed, it fails in any metric space that isn’t a singleton; see [39]).

The solution to this problem comes by considering functions defined on $\mathbb{Z}_m^n$ rather than on $\{-1,1\}^n$, and scaling the argument of the function. Specifically, following [38] say that a metric space $(X,d_X)$ has metric cotype $s \in (0,\infty)$ if for every $n \in \mathbb{N}$ there exists $m \in \mathbb{N}$ such that

$$\forall f : \mathbb{Z}_m^n \to X, \quad \sum_{j=1}^n \mathbb{E} [d_X(f(x + me_j), f(x))^s] \leq_X \mathbb{E} [d_X(f(x + \varepsilon), f(x))^s],$$

where the expectation is with respect to $(x, \varepsilon) \in \mathbb{Z}_m^n \times \{-1,0,1\}^n$ chosen uniformly at random. It was proved in [38] that a Banach space $(X,\|\cdot\|_X)$ has Rademacher cotype $s$ if and only if it has metric cotype $s$, in particular $L_p$ has metric cotype max${p,2}$. “Scaling” refers to the fact that in (7) we consider displacements of the argument of $f$ by a multiple of $m$, i.e., we consider distances between $f(x + me_j)$ and $f(x)$ rather than distances between $f(x + e_j)$ and $f(x)$, and then we compensate for this by normalizing the distances appropriately. This idea makes its appearance also in the left-hand side of our metric $X_p$ inequality [3], but we shall see below that the need for scaling in the context of Theorem 1.1 is due to a more subtle reason than the above explanation of why scaling is needed in the context of metric cotype (compare the total masses of the measures that appear in both sides of [3] to see that it doesn’t cause the problem that we presented above).

1.3.1. The case $2 < q < p$. While Paley’s work [47] from 1936 established that $L_q$ is not isomorphic to a subspace for $L_p$ when $2 < q < p$, several decades later more structural approaches to this theorem were developed. In 1962, Kadec and Pełczyński [28] introduced an influential way to solve this problem through a structural study of basic sequences in $L_p$ spaces. In particular, it follows from [28] that for $p \in (2,\infty)$, any infinite symmetric basic sequence in $L_p$ is equivalent to either the standard basis of $\ell_p$ or the standard basis of $\ell_2$. Consequently, for $q \in (2,p)$ there does not exist a symmetric basic sequence in $L_q$ that is equivalent to the unit basis of $\ell_q$, and therefore $\ell_q$ cannot be isomorphic to a subspace of $\ell_p$. In 1979, Johnson, Maurey, Schechtman and Tzafriri [25] obtained a proof of Paley’s theorem via a classification of finite symmetric bases in function spaces, leading to a comprehensive theory of symmetric structures in Banach spaces to which the research monograph [25] is devoted. In particular, in [25] a “local” version of the above theorem of Kadec and Pełczyński is studied, leading to a classification of all finite symmetric bases in $L_p$. It turns out that in this finitary setting the classification involves more structures than those that are allowed (by the Kadec–Pełczyński theorem) for infinite symmetric sequences in $L_p$, namely, a one-parameter family of such sequences can occur, yet any finite symmetric sequence in $L_p$ is equivalent to a member of this one-parameter family. This theorem of [25] is the starting point of our work here.

Given a Banach space $(X,\|\cdot\|_X)$, $n \in \mathbb{N}$ and $K \in [1,\infty)$, recall that a linearly independent sequence of vectors $(x_1, \ldots, x_n) \in X^n$ is said to be $K$-symmetric if for every sequence of scalars $a_1, \ldots, a_n \in \mathbb{R}$ and every permutation $\pi \in S_n$ and every sequence of signs $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n) \in \{-1,1\}^n$, we have $\|\varepsilon_1 a_{\pi(1)} x_1 + \ldots + \varepsilon_n a_{\pi(n)} x_n\|_X \leq K\|a_1 x_1 + \ldots + a_n x_n\|_X$. The sequence $(x_1, \ldots, x_n) \in X^n$ is said to be normalized if $\|x_j\|_X = 1$ for all $j \in \{1, \ldots, n\}$. Given two Banach spaces $(X,\|\cdot\|_X)$ and $(Y,\|\cdot\|_Y)$, two sequences $(x_1, \ldots, x_n) \in X^n$ and $(y_1, \ldots, y_n) \in Y^n$ are said to be $K$-equivalent if there exists $s \in (0,\infty)$ such that $s\|a_1 x_1 + \ldots + a_n x_n\|_X \leq \|a_1 y_1 + \ldots + a_n y_n\|_Y \leq Ks\|a_1 x_1 + \ldots + a_n x_n\|_X$ for all choices of scalars $a_1, \ldots, a_n \in \mathbb{R}$.

Fixing $p \in (2,\infty)$, examples of symmetric sequences in $L_p$ are furnished by Rosenthal’s $X_p^n(\omega)$ symmetric bases [54], as $\omega$ ranges over $[0,\infty)$. The definition of these bases is very simple: let $u_1, \ldots, u_n$ be an orthonormal basis of $\ell_2$ and define $\{x_j(p,\omega)\}_{j=1}^n \subseteq (\ell_p \oplus \ell_2)_p$ by

$$x_j(p,\omega) \overset{\text{def}}{=} \frac{1}{(1+\omega^p)^\frac{1}{p}} \cdot e_j + \frac{\omega}{(1+\omega^p)^\frac{1}{p}} \cdot u_j.$$  

(8)
The 1-symmetric sequence \( \{x_j(p, \omega)\}_{j=1}^n \) is known in the literature as Rosenthal’s \( X_p^n(\omega) \) basis. Note that since \( \ell_2 \) is isometric to a subset of \( L_p \) (see e.g. \([59]\)), the sequence \( \{x_j(p, \omega)\}_{j=1}^n \) can be realized as elements of \( L_p \).

In \([25]\) it was proved that for every \( K \in [1, \infty) \) and \( p \in (2, \infty) \) there exists \( D(p, K) \in (0, \infty) \) such that every \( K \)-symmetric sequence \( \{x_1, \ldots, x_n\} \) in \( L_p \) is \( D(p, K) \)-equivalent to an \( X_p^n(\omega) \) basis for some \( \omega \in [0, \infty) \). This classification theorem has immediate relevance to linear embeddings of \( \ell_2^n \) into \( L_p \). Indeed, if \( T : \ell_2^n \to L_p \) is injective and linear then \( (Te_1, \ldots, Te_n) \) is a \( \text{dist}(T) \)-symmetric sequence in \( L_p \), and is therefore \( D(p, \text{dist}(T)) \)-equivalent to an \( X_p^n(\omega) \) basis for some \( \omega \in (0, \infty) \). Direct inspection now reveals that this is only possible if \( \text{dist}(T) \) tends to \( \infty \) as \( n \to \infty \). In fact, by computing the various bounds explicitly and optimizing over \( \omega \in [0, \infty) \), as done in \([19]\) (relying in part on a computation from \([22]\)), one can deduce that for every \( 2 < q < p < \infty \) there exists \( \sigma(p, q) \in (0, \infty) \) such that for every invertible linear mapping \( T : \ell_q^n \to L_p \) we have

\[
\text{dist}(T) \geq \sigma(p, q) \cdot n^{(p-q)(q-2)/p^2}. \tag{9}
\]

The lower bound in \([9]\) is asymptotically sharp (up to the implicit dependence on \( p, q \), as exhibited by the embedding \( J_{(q \to p; n)} : \ell_q^n \to (\ell_p^n \oplus \ell_2^n)_p \subseteq L_p \) given by \([7]\)

\[
\forall j \in \{1, \ldots, n\}, \quad J_{(q \to p; n)}(e_j) \overset{\text{def}}{=} n^{1/2} \cdot e_j + n^{1/2} \cdot u_j, \tag{10}
\]

where, as in \([8]\), \( u_1, \ldots, u_n \) is an orthonormal basis of \( \ell_2^n \). Indeed, by a straightforward Langrange multiplier argument (see Section \([2]\) below), for every \( 2 < q < p \) we have

\[
\text{dist}\left(J_{(q \to p; n)}^R\right) \prec n^{(p-q)(q-2)/p^2}. \tag{11}
\]

A sequence of random variables \( \{Y_j\}_{j=1}^n \) is said to be symmetrically exchangeable if for every \( \pi \in S_n \) and \( \varepsilon \in \{-1, 1\}^n \) the random vectors \( (\varepsilon_1 Y_{\pi(1)}, \ldots, \varepsilon_n Y_{\pi(n)}) \) and \( (Y_1, \ldots, Y_n) \) are identically distributed. The proof of the above classification of finite symmetric sequences in \( L_p \) relies on the following inequality \([25]\). Fix \( p \in [2, \infty) \) and suppose that \( \{Y_j\}_{j=1}^n \) are symmetrically exchangeable random variables with \( \mathbb{E}[|Y_j|^p] = 1 \) for all \( j \in \{1, \ldots, n\} \). Then for every \( t_1, \ldots, t_n \in \mathbb{R} \),

\[
\left( \frac{\log p}{p} \right)^p \cdot \mathbb{E}\left[ \left| \sum_{j=1}^n t_j Y_j \right|^p \right] \lesssim_p \sum_{j=1}^n |t_j|^p + \left( \frac{1}{n} \sum_{j=1}^n t_j^2 \right)^{p/2} \mathbb{E}\left[ \left( \sum_{j=1}^n Y_j^2 \right)^{p/2} \right]. \tag{12}
\]

The term \( (\log p)/p \) in the left-hand side of \([12]\) is sharp up to a universal constant factor: in this sharp form the inequality \([12]\) is due to \([26]\). Without a sharp dependence on \( p \), inequality \([12]\) was first proved in \([25]\). The proof of \([12]\) with sharp dependence on \( p \) is significantly more involved than the proof in \([25]\). The dependence on \( p \) is not of major importance for us here, but it is worthwhile to state the above sharp form of \([12]\) since it is available in the literature.

Fix \( p \in [2, \infty) \), \( n \in \mathbb{N} \), \( k \in \{1, \ldots, n\} \) and \( a_1, \ldots, a_n \in \mathbb{R} \). For \( (\varepsilon, \pi) \in \{-1, 1\}^n \times S_n \) chosen uniformly at random, define

\[
Y_j(\varepsilon, \pi) \overset{\text{def}}{=} \frac{\varepsilon_j a_{\pi(j)}}{\left( \frac{1}{n} \sum_{s=1}^n |a_s|^p \right)^{1/p}}.
\]

Then \( \{Y_j\}_{j=1}^n \) are symmetrically exchangeable random variables (the underlying probability space being the uniform measure on \( \{-1, 1\}^n \times S_n \)), with \( \mathbb{E}[|Y_j|^p] = 1 \). An application of \([12]\) with

---

1The superscript in the notation \( J_{(q \to p; n)}^R(\cdot) \) refers to Rosenthal.
Following inequality, which should hold true for every $S$ constant 1 (as shown by Enflo in [16]) implies that for every $k$, Inequality (16) holds true when $p = 2$. Indeed, the fact that the real line has Enflo type 2 with constant 1 (as shown by Enflo in [16]) implies that for every $S \subseteq \{1, \ldots, n\}$ we have

$$\mathbb{E} \left[ \left( f(\varepsilon) - f(\varepsilon_{\{1, \ldots, n\} \setminus S} - \varepsilon) \right)^2 \right] \leq \sum_{j \in S} \mathbb{E} \left[ \left( f(\varepsilon) - f(\varepsilon_{\{1, \ldots, n\} \setminus S} - \varepsilon) \right)^2 \right].$$

By averaging (17) over all of those $S \subseteq \{1, \ldots, n\}$ satisfying $|S| = k$ we see that (16) holds true when $p = 2$, with $\alpha(2) = 1$ and even without the final term in the right-hand side of (16).

The validity of (16) for $p = 2$ indicates that the reason why scaling is needed for the definition (7) of metric cotype does not arise in the context of (16). However, Proposition 1.3 below shows that scaling is nevertheless necessary in the context of metric $X_p$ inequalities, thus explaining our formulation of Theorem 1.1. Note that the conclusion of Theorem 1.1 implies the linear $X_p$ inequality (15). Roughly speaking, this follows by applying (3) to the linear function $f : \mathbb{Z}_m^n \to \mathbb{R}$ given by $f(x) = \sum_{j=1}^n x_j a_j$. However, this reasoning isn’t quite accurate because this $f$ isn’t well defined as a function on the discrete torus $\mathbb{Z}_m^n$; for a precise argument see Proposition 2.1 below.
Proposition 1.3 (Scaling is necessary). Fix \( p \in (2, \infty) \), \( \alpha \in (0, 1) \), \( m, n \in \mathbb{N} \) and \( k \in \{1, \ldots, m\} \). Suppose that for every \( f : \mathbb{Z}_m^n \to \mathbb{R} \) we have

\[
\frac{\alpha^p}{(\begin{smallmatrix} n \\ k \end{smallmatrix})} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{|S|=k} \mathbb{E}\left[|f(x + me_j) - f(x)|^p\right] \leq \frac{k}{n} \sum_{j=1}^n \mathbb{E}\left[|f(x + e_j) - f(x)|^p\right] + \left(\frac{k}{n}\right)^{\frac{p}{2}} \mathbb{E}\left[|f(x + \varepsilon) - f(x)|^p\right],
\]

where the expectation is with respect to \((x, \varepsilon) \in \mathbb{Z}_m^n \times \{-1, 1\}^n\) chosen uniformly at random. Then

\[
k \geq \left(\frac{5}{\alpha}\right)^{\frac{2p}{p-2}} \implies m \geq \frac{\alpha}{3} \sqrt{n/k}.
\]

The proof of Proposition 1.3 appears in Section 2. We conjecture that the dependence of \( m \) on \( n \) and \( k \) that appears in Proposition 1.3 is sharp, up to the (possibly \( p \)-dependent) constant. This is the content of Conjecture 1.4 below. It seems that in order to prove Conjecture 1.4 one would need to exploit cancelations that are more subtle than those that we used to prove Theorem 1.1.

Conjecture 1.4. For every \( p \in (2, \infty) \) there exist \( \alpha_p \in (0, 1) \) and \( C_p \in [1, \infty) \) such that if \( m, n \in \mathbb{N} \) and \( k \in \{1, \ldots, n\} \) satisfy \( m \geq C_p \sqrt{n/k} \) then for every \( f : \mathbb{Z}_m^n \to \mathbb{R} \) we have

\[
\frac{\alpha_p^p}{(\begin{smallmatrix} n \\ k \end{smallmatrix})} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{|S|=k} \mathbb{E}\left[|f(x + 2me_j) - f(x)|^p\right] \leq \frac{k}{n} \sum_{j=1}^n \mathbb{E}\left[|f(x + e_j) - f(x)|^p\right] + \left(\frac{k}{n}\right)^{\frac{p}{2}} \mathbb{E}\left[|f(x + \varepsilon) - f(x)|^p\right],
\]

where the expectation is with respect to \((x, \varepsilon) \in \mathbb{Z}_m^n \times \{-1, 1\}^n\) chosen uniformly at random.

We will see in Section 1.4.3 below that, in addition to its intrinsic interest, a positive resolution of Conjecture 1.4 would have striking consequences in the theory of metric embeddings. A conjectural convolution inequality (of independent interest) that we formulate in Question 6.1 below is shown in Proposition 6.2 below to imply a positive answer to Conjecture 1.4.

Before passing to a description of the geometric consequences of Theorem 1.1 we note that the linear \( X_p \) inequality (15) also has a (much easier) converse [25]. Specifically, for every \( p \in (2, \infty) \) there exists \( K(p) \in (0, \infty) \) such that for every \( a_1, \ldots, a_n \in \mathbb{R} \) we have

\[
\sum_{j=1}^n |a_j|^p \leq \left(\frac{k}{n}\right)^{\frac{p}{2}} \mathbb{E}\left[\sum_{j=1}^n |\varepsilon_j a_j|^p\right] \leq \left(\frac{k}{n}\right)^{\frac{p}{2}} \mathbb{E}\left[\sum_{S \subseteq \{1, \ldots, n\}} \sum_{|S|=k} \left|\sum_{j \in S} \varepsilon_j a_j\right|^p\right],
\]

where the expectation is over \( \varepsilon \in \{-1, 1\}^n \) chosen uniformly at random. An inspection of the proof of [21] in [25] (or in [26]) reveals that one can take \( K(p) \lesssim \sqrt{p} \) in [21]. Theorem 1.5 below is a nonlinear version of [21]. Although we do not have a new geometric application of the reverse metric \( X_p \) inequality that appears in Theorem 1.5 it is worthwhile to establish it so as to obtain a complete picture of the \( X_p \) phenomenon in the metric setting. As a side product, our proof of Theorem 1.5 yields some new information on metric cotype; see Theorem 5.2 below and the discussion immediately preceding it.
Theorem 1.5 (Reverse metric $X_p$ inequality). Fix $p \in [2, \infty)$ and $k, m \in \mathbb{N}$ with $m \geq \frac{k^{1/p}}{\sqrt{p}}$. Fix also an integer $n \geq k$. Then for every $f : \mathbb{Z}_{8m}^n \rightarrow L_p$ we have

$$ \frac{k}{n} \sum_{j=1}^{n} \frac{\mathbb{E} \left[ \| f(x + 4me_j) - f(x) \|_p^p \right]}{m^p} + \left( \frac{k}{n} \right)^{\frac{p}{q}} \mathbb{E} \left[ \| f(x + \varepsilon) - f(x - \varepsilon) \|_p^p \right] \lesssim_p \frac{p^2}{(m/k)^{p}} \sum_{S \subseteq \{1, \ldots, n\}} \mathbb{E} \left[ \| f(x + \varepsilon S) - f(x) \|_p^p \right], \quad (22)$$

where the expectation is with respect to $(x, \varepsilon) \in \mathbb{Z}_{8m}^n \times \{-1, 1\}^n$ chosen uniformly at random.

1.4. Metric $X_p$ inequalities as obstructions to embeddings. Theorem 1.1 yields a bi-Lipschitz invariant that can be used to obtain new nonembeddability results which we shall now describe.

1.4.1. Snowflakes. Fix $p, q \in [1, \infty)$. Sharp restrictions on those $\theta \in (0, 1]$ for which the $\theta$-snowflake of $L_q$ admits a bi-Lipschitz embedding into $L_p$ follow from the theory of metric type and cotype when either $q \in [1, 2]$ and $p \geq q$, or $q \in [2, \infty)$ and $p \leq q$; see [32, 38, 20]. Here we obtain, as a consequence of Theorem 1.1, the first such result when $2 < q < p$.

**Theorem 1.6 ($L_q$ snowflakes in $L_p$).** For every $2 < q < p$ there exists $\delta(p, q) > 0$ such that if $\theta \in (0, 1)$ is such that the metric space $(L_q, \| x - y \|_{q, \theta})$ admits a bi-Lipschitz embedding into $L_p$, then necessarily $\theta \leq 1 - \delta(p, q)$. Specifically, $\theta$ must satisfy

$$ \theta \leq \frac{2q(p - q) + q^2(p - 1)(p - 2)}{2p^2(q - 2)} \left( \sqrt{1 + \frac{4p(p - 2)(q - 2)}{(pq - 3q + 2)^2} - 1} \right) \leq 1 - \frac{(p - q)(q - 2)}{2p^3}. \quad (23)$$

It was shown in [37, Remark 5.10] that for $2 < q < p$ that $(q/p)$-snowflake of $L_q$ is isometric to a subset of $L_p$. We conjecture that this is sharp, i.e., that the upper bound on $\theta$ that appears in (23) can be improved to $\theta \leq q/p$.

**Conjecture 1.7.** Suppose that $2 < q < p$ and $\theta \in (0, 1)$ is such that the metric space $(L_q, \| x - y \|_{q, \theta})$ admits a bi-Lipschitz embedding into $L_p$. Then necessarily $\theta \leq q/p$.

In fact, when $2 < q \leq p$, we ask whether or not $L_q$ has a unique snowflake that admits a bi-Lipschitz embedding into $L_p$. If true, this would be manifestly different than the case $1 \leq q \leq p \leq 2$, where it is known [12] (see also [58]) that the metric space $(L_q, \| x - y \|_{q, \theta})$ admits an isometric embedding into $L_p$ for every $0 < \theta \leq q/p$.

**Question 1.8 (Uniqueness of snowflakes).** Suppose that $2 < q \leq p$ and $\theta \in (0, 1)$. Is it true that if the metric space $(L_q, \| x - y \|_{q, \theta})$ admits a bi-Lipschitz embedding into $L_p$ then necessarily $\theta = q/p$?

The case $q = p$ of Question 1.8 is a well-known problem that has been open for many years (though apparently not stated explicitly in the literature): is it true that if $p \in (2, \infty)$ then for no $\theta \in (0, 1)$ the metric space $(L_p, \| x - y \|_{p, \theta})$ admits a bi-Lipschitz embedding into $L_p$? Related results appear in [37, Section 5].

1.4.2. Grids. Recall that for $q \in [1, \infty)$ and $m, n \in \mathbb{N}$ the integer grid $\{1, \ldots, m\}^n$, equipped with the metric inherited from $\ell_q^n$, is denoted $[m]_q^n$. Theorem 1.9 below, which is a consequence of Theorem 1.1, contains the best-known lower bound on $c_p([m]_q^n)$ when $2 < q < p$, thus yielding another quantitative version of the fact that $L_q$ does not admit a bi-Lipschitz embedding into $L_p$. 


Theorem 1.9 ($L_p$ distortion of $L_q$ grids). For every $p \in (2, \infty)$ there exists $\alpha_p \in (0, \infty)$ such that for every $q \in (2, p)$ and $m, n \in \mathbb{N}$ we have

$$c_p([m]^n_q) \gtrsim \alpha_p \min \left\{ \frac{q^{(p-2)}}{m} n^{(p-2)+\frac{q}{p-q}}, n \right\}^{\frac{(p-q)(q-2)}{q^2(p-2)}}. \quad (24)$$

In particular,

$$m \geq n^{1 + \frac{p-q}{q(p-2)}} \implies c_p([m]^n_q) \geq \alpha_p n^{\frac{(p-q)(q-2)}{q^2(p-2)}} \gtrsim \alpha_p c_p(\ell^m_q). \quad (25)$$

The fact that the lower bound in (24) becomes weaker for smaller $m$ is necessary, as exhibited by the following embedding from [38]. First, let $G_{2,p} : L_2 \to L_p$ be an isometric embedding of $L_2$ into $L_p$. By a classical theorem of Schoenberg [55] (see also [58]) there exists an isometric embedding of the $(2/q)$-snowflake of $\ell^m_2$ into $L_2$, i.e., there exists $\psi_q^m : \ell^m_2 \to L_2$ such that

$$\forall x, y \in \ell^m_2, \quad \|\psi_q^m(x) - \psi_q^m(y)\|_2 = \|x - y\|_q^2.$$  

Finally, let $I_{q^{-2}}^m : \ell^m_2 \to \ell^m_2$ be the identity mapping, and define

$$J_{(q;p;n)}^S \overset{\text{def}}{=} G_{2,p} \circ \psi_q^m \circ I_{q^{-2}}^m : \ell^m_q \to L_p. \quad (26)$$

As argued in [38], the distortion of the restriction of $J_{(q;p;n)}^S$ to $[m]^n_q$ satisfies

$$\operatorname{dist} \left( J_{(q;p;n)}^S \big|_{[m]^n_q} \right) \leq m^{\frac{1}{q} - \frac{2}{q}}. \quad (27)$$

Recalling the definition of the embedding $J_{(q;p;n)}^R$ in (10), we therefore have

$$c_p([m]^n_q) \leq \min \left\{ \operatorname{dist} \left( J_{(q;p;n)}^R \big|_{[m]^n_q} \right), \operatorname{dist} \left( J_{(q;p;n)}^S \big|_{[m]^n_q} \right), m^{\frac{1}{q} - \frac{2}{q}} \right\} \leq \min \left\{ n^{\frac{(p-q)(q-2)}{q^2(p-2)}}, m^{\frac{1}{q} - \frac{2}{q}} \right\}. \quad (27)$$

We conjecture that (27) is asymptotically sharp up to constant factors that depend only on $p, q$.

Conjecture 1.10. For $2 < q < p$ and $m, n \in \mathbb{N}$, the better of the embeddings $J_{(q;p;n)}^R$ and $J_{(q;p;n)}^S$ appearing in (10) and (26), respectively, is the best possible bi-Lipschitz embedding of the $L_q$ integer grid $[m]^n_q$ into $L_p$. Equivalently, $c_p([m]^n_q)$ is bounded from above and from below by positive constants that may depend only on $p$ and $q$ times the quantity

$$\min \left\{ n^{\frac{(p-q)(q-2)}{q^2(p-2)}}, m^{\frac{1}{q} - \frac{2}{q}} \right\}. \quad (28)$$

In particular, there exists $\eta(p, q) \in (0, 1)$ such that

$$m \geq n^{\frac{p-q}{q(p-2)}} \implies c_p([m]^n_q) \gtrsim \eta(p, q) c_p(\ell^m_q),$$

yet

$$m = o \left( n^{\frac{p-q}{q(p-2)}} \right) \implies c_p([m]^n_q) = o(c_p(\ell^m_q)) \quad (\text{as } n \to \infty).$$

An affirmative answer to Conjecture 1.10 would imply that if the linear embedding $J_{(q;p;n)}^R$ of $\ell^m_q$ into an appropriate Rosenthal $X_p(\omega)$ space fails to yield the best possible bi-Lipschitz embedding of $[m]^n_q$ into $L_p$ then (up to constant factors that are independent of $m, n$), the best possible way to embed $[m]^n_q$ into $L_p$ would be to embed it into $L_q$ (ignoring the fact that we are seeking an embedding into the larger space $L_p$), via the (highly nonlinear) Schoenberg embedding $\psi_q^{\omega}$. Admittedly, if true, this phenomenon would be quite exotic, but we conjecture that it indeed occurs partially because it is a consequence of Conjecture 1.4 as we shall see in Section 1.4.3 below.
Remark 1.11. There are also interesting open problems related to embeddings of $[m]^n_p$ into $L_q$ when $p > q > 2$. Specifically, by combining the upper bound in [38] with the metric cotype-based lower bound in [39], we see that

$$\frac{1}{\sqrt{q}} \min \left\{ \left( n^{\frac{1}{q} - \frac{1}{p}}, m^{1 - \frac{2}{p}} \right) \right\} \lesssim c_q([m]^n_p) \lesssim \min \left\{ \left( n^{\frac{1}{q} - \frac{1}{p}}, m^{1 - \frac{2}{p}} \right) \right\}. \quad (29)$$

The bounds in (29) match only when $q = 2$, and it remains open to evaluate $c_q([m]^n_p)$ up to constant factors that are independent of $m, n$. An inspection of the argument in [38] reveals that the lower bound on $c_q([m]^n_p)$ in (29) would be sharp (up to constant factors that may depend only on $p, q$) if the $(q/p)$-snowflake of $L_q$ admitted a bi-Lipschitz embedding into $L_q$. When $q = 2$ this is indeed the case due to the theorem of Schoenberg that was quoted above, but for $q > 2$ a positive answer to Question 1.8 (see also the paragraph immediately following Question 1.8) would imply that no nontrivial snowflake of $L_q$ admits a bi-Lipschitz embedding into $L_q$. In the spirit of Conjecture (28), one is tempted to ask whether or not the upper bound on $c_q([m]^n_p)$ in (29) is asymptotically sharp, i.e., if also in this setting it is best to embed $[m]^n_p$ into $L_q$ via an appropriate embedding into the smaller space $L_2$. However, if this were true then one would need to find a better lower bound on $c_q([m]^n_p)$ than what follows from the fact that $L_q$ has metric cotype $q$. For this reason, at present we do not have a concrete conjecture as to the sharp asymptotics of $c_q([m]^n_p)$ when $p > q > 2$.

1.4.3. Consequences of Conjecture 1.4. The following theorem asserts that Conjecture 1.4 implies a positive solution of Conjecture 1.7 and Conjecture 1.10. Thus, obtaining the conjecturally sharp value of $m$ in the metric $X_p$ inequality of Theorem 1.1 in addition to its intrinsic analytic interest, would yield striking nonembeddability results. As we mentioned earlier, in Section 6 we present a concrete convolution inequality (that is interesting on its own right) and prove that it implies an affirmative answer to Conjecture 1.4 and hence also to Conjecture 1.7 and Conjecture 1.10.

Theorem 1.12. If Conjecture 1.4 holds true then for every $2 < q < p$ and $\theta \in (0, 1)$,

$$c_p \left( L_q, \|x - y\|_q^\theta \right) < \infty \implies \theta \leq \frac{q}{p}. \quad (30)$$

Moreover, for every $m, n \in \{1, \ldots, n\}$ the $L_p$ distortion $c_p([m]^n)$ is bounded from above and below by a constant that may depend only on $p$ times the quantity appearing in (28).

1.5. $X_p$ metric spaces? For $p \in (0, \infty)$, by pursuing the Ribe program in light of Theorem 1.1, one should define a metric space $(X, d_X)$ is an $X_p$ metric space if for every $n \in \mathbb{N}$ and $k \in \{1, \ldots, n\}$ there exists $m \in \mathbb{N}$ such that for every $f : \mathbb{Z}_{4m}^n \to X$ we have

$$\frac{1}{\binom{n}{k}} \sum_{S \subseteq \{1, \ldots, n\}} \frac{\mathbb{E} \left[ d_X \left( f(x + 2m \varepsilon_S), f(x) \right)^p \right]}{m^p} \lesssim X \sum_{j=1}^{n} \frac{\mathbb{E} \left[ d_X \left( f(x + e_j), f(x) \right)^p \right]}{n} + \frac{k^2}{n} \mathbb{E} \left[ d_X \left( f(x + \varepsilon), f(x) \right)^p \right], \quad (31)$$

where the expectation is with respect to $(x, \varepsilon) \in \mathbb{Z}_{4m}^n \times \{-1, 1\}^n$ chosen uniformly at random. One would then want to investigate the structure of such metric spaces, motivated in part by analogies from the linear theory. However, in contrast to previous successful steps in the Ribe program, in the present setting the linear theory of $X_p$ spaces hasn’t been studied yet, and it therefore seems to be more natural to first understand what makes a Banach space an $X_p$ Banach space. Specifically, say that a Banach space $(X, \| \cdot \|_X)$ is an $X_p$ Banach space if for every $n \in \mathbb{N}$ and $k \in \{1, \ldots, n\}$,
Every \( v_1, \ldots, v_n \in X \) satisfy
\[
\frac{1}{\binom{n}{k}} \sum_{S \subseteq \{1, \ldots, n\}} \mathbb{E} \left[ \left\| \sum_{j \in S} \varepsilon_j v_j \right\|_X^p \right] \lesssim_X \frac{k}{n} \sum_{j=1}^{n} \left\| v_j \right\|_X^p + \left( \frac{k}{n} \right)^{\frac{p}{2}} \mathbb{E} \left[ \left\| \sum_{j=1}^{n} \varepsilon_j v_j \right\|_X^p \right],
\]
where the expectation is over \( \varepsilon \in \{-1, 1\}^n \) chosen uniformly at random. Being an \( X_p \) Banach space is a local property. Our proof of Theorem 1.1 shows that a Banach space is an \( X_p \) metric space if and only if it is an \( X_p \) Banach space, thus completing the Ribe program in this setting.

For \( p > 2 \), it seems that the only Banach spaces that were previously known to be \( X_p \) Banach spaces were those that are isomorphic to subspaces of \( L_p \). However, there exist separable \( X_p \) Banach spaces that are not isomorphic to a subspace of \( L_p \). In Section 7 we prove that for \( p \in [2, \infty) \) the Schatten \( p \) trace class \( S_p \) is an \( X_p \) Banach space. The fact that \( S_p \) is not isomorphic to a subspace of \( L_p \) was proved in [36] (see also [49]). Obtaining a satisfactory understanding of those Banach spaces that are \( X_p \) spaces remains an interesting, though probably quite difficult, research challenge.

Since \( S_p \) is an \( X_p \) Banach space, our work here shows that it is also an \( X_p \) metric space. The nonembeddability results that were stated above for embeddings into \( L_p \) therefore hold true for embeddings into \( S_p \) as well. In the setting of \( S_p \), these nonembeddability results are new even in the linear category. It was known that for \( 2 < q < p \) the Banach–Mazur distance of \( \ell_q^n \) to any subspace of \( S_p \) must tend to \( \infty \) with \( n \); this follows from the non-commutative Kadec–Pełczyński result in [50]; see also Theorem 10.7 in [52]. The literature gives no information on the rate at which \( c_{S_p}(\ell_q^n) \) tends to infinity with \( n \) (extracting quantitative estimates from the proof in [56], if at all possible, would probably require significant effort and yield weak bounds). Here we see that \( c_{S_p}(\ell_q^n) \) is asymptotically \( n^{(p-q)(q-2)/(q^2(p-2))} \), up to constant factors that may depend only on \( p, q \).

### 2. Preliminaries

Here we establish some initial facts and prove some of the simpler statements that were presented in the Introduction. The results of the present section will not be used for the proofs of Theorem 1.1 and its consequences, so they could be skipped on a first reading.

We shall start with the proof of Proposition 1.3, i.e., that scaling is needed for the metric \( X_p \) inequality of Theorem 1.1 to hold true.

**Proof of Proposition 1.3** We shall use here the notation that was introduced in the statement of Proposition 1.3. Since \( \ell_2 \) embeds isometrically into \( L_p \), by [37, Lem. 5.2] there exists \( F : \mathbb{Z}_{2m}^n \to L_p \) such that for every distinct \( x, y \in \mathbb{Z}_{2m}^n \) we have
\[
1 \leq \frac{\| F(x) - F(y) \|_p}{\min \left\{ 2\sqrt{k}, \sqrt{\sum_{j=1}^{n} |e^{\pi i (x_j - y_j)}/m - 1|^2} \right\}} \leq 2. \tag{32}
\]

By integrating (18) we see that
\[
\frac{1}{\binom{n}{k}} \sum_{S \subseteq \{1, \ldots, n\}} \mathbb{E} \left[ \| F(x + m \varepsilon_S) - F(x) \|_p^p \right] \leq \frac{k}{n} \sum_{j=1}^{n} \mathbb{E} \left[ \| F(x + e_j) - F(x) \|_p^p \right] + \left( \frac{k}{n} \right)^{\frac{p}{2}} \mathbb{E} \left[ \| F(x + \varepsilon) - F(x) \|_p^p \right]. \tag{33}
\]

It follows from (32) that if \( S \subseteq \{1, \ldots, n\} \) satisfies \( |S| = k \) then \( \| F(x + m \varepsilon_S) - F(x) \|_p \geq 2\sqrt{k} \) for every \( x \in \mathbb{Z}_{2m}^n \). Also, the elementary inequality \( |e^{\pi i / m} - 1|^2 \leq \pi^2 / m^2 \) implies that for every
$j \in \{1, \ldots, n\}$ we have $\|F(x + e_j) - F(x)\|_p \leq 2\pi/m$, and for every $(x, \varepsilon) \in \mathbb{Z}^n_{2m} \times \{-1, 1\}^n$ we have $\|F(x + \varepsilon) - F(x)\|_p \leq 4\sqrt{k}$. In conjunction with (33), these estimates show that

$$\frac{2^p\alpha^p k^{\frac{p}{2}}}{m^p} \leq \frac{2^p \pi k^p}{m^p} + \frac{(4k)^p}{n^{2p}},$$

which yields the desired implication (19).

Next, we shall check the validity of (11), i.e., evaluate the distortion of the mapping $J^R_{(q \to p; n)}$ given in (10). This is a known (and easy) statement which is included here only because we could not locate a clean reference for it.

**Proof of (11).** The definition (10) implies that for every $x \in \ell^n_q$ we have

$$\left\| J^R_{(q \to p; n)}(x) \right\|^p_{(\ell^n_q \oplus \ell^n_2)_p} = n^p \|x\|^p_p + n^p \|x\|^p_2.

Consequently, it suffices to show that for every $x \in \ell^n_q$ we have

$$\frac{n^p \|x\|^p_q}{2^p n} \leq n^p \|x\|^p_p + n^p \|x\|^p_2 \leq 2n^p \|x\|^p_q. \quad (34)$$

The rightmost inequality in (34) is an immediate consequence of the estimates $\|x\|_2 \leq n^{\frac{1}{2} - \frac{1}{q}} \|x\|_q$ and $\|x\|_p \leq \|x\|_q$, which hold true because $2 < q < p$.

Let $x \in \ell^n_q$ with $\|x\|_q = 1$ be such that $n^p \|x\|^p_p + n^p \|x\|^p_2$ is minimal. We may also assume that the number of nonzero entries of $x$ is minimal, and that $x_1, \ldots, x_k > 0$ and $x_{k+1} = \ldots = x_n = 0$ for some $k \in \{1, \ldots, n\}$. Hence, there exists (a Lagrange multiplier) $\lambda \in \mathbb{R}$ such that

$$\forall j \in \{1, \ldots, k\}, \quad n^p x_j^{p-1} + n^p \|x\|_2^{-2} x_j = \lambda x_j^{q-1}. \quad (35)$$

For $s \in [0, \infty)$ write $\psi(s) \overset{\text{def}}{=} n^p s^{p-2} - \lambda s^{q-2} + n^p \|x\|_2^{-2}$. Since $p, q > 2$ we have $\psi(0) > 0$, and since $p > q$ we have $\lim_{s \to \infty} \psi(s) = \infty$. It follows from (35) that $\lambda > 0$, and therefore there is a unique $s_0 \in (0, \infty)$ for which $\psi(s_0) = 0$. This means that $\psi$ starts at a positive value, decreases on $(0, s_0)$, and then increases to $\infty$. Consequently, there exist $a, b \in (0, \infty)$ such that $\psi(s) = 0 \iff s \in \{a, b\}$ for every $s \in (0, \infty)$. Since by (35) we have $\psi(x_j) = 0$ for every $j \in \{1, \ldots, k\}$, it follows that there exists $S \subseteq \{1, \ldots, k\}$ such that $x_j = a 1_S(j) + b 1_{\{1, \ldots, k\}\setminus S}(j)$ for all $j \in \{1, \ldots, k\}$. Since $\|x\|_q = 1$, we may assume without loss of generality that $\alpha^q |S| \geq 1/2$, i.e., that $a \geq 1/(2|S|)^{1/q}$. Consequently,

$$n^p \|x\|^p_p + n^p \|x\|^p_2 \geq n^p |S| a^p + n^p |S| \|a\|^p \geq \frac{n^p}{2^p |S|^{\frac{p}{q}}} + \frac{n^p |S|^{\frac{2}{q} - \frac{p}{q}}}{2^p} \geq 2 \frac{p}{2^p} n^p \frac{p - q(p - 2)}{q^2(p - 2)},$$

where the last step follows by computing the minimum of $n^p s^{\frac{p}{q} - 1} + n^p s^{-\frac{p}{q}}$ over $s \in (0, \infty)$.

In the present work, Banach spaces are assumed to be over real scalars unless stated otherwise. However, it will sometimes be notationally convenient to work with complex Banach spaces, and in fact all the results presented below hold true for Banach spaces over the complex numbers as well. This follows from a straightforward complexification argument. Specifically, given a real Banach space $(Z, \|\cdot\|_Z)$ and $p \in [1, \infty)$ denote by $Z_p(\mathbb{C})$ the following $p$-complexification of $Z$. As a vector space, $Z_p(\mathbb{C}) = Z \times Z$. As usual, we consider $Z_p(\mathbb{C})$ as a vector space over $\mathbb{C}$ by setting $(a + bi)(u, v) = (au - bv, av + bu)$ for every $u, v \in Z$ and $a, b \in \mathbb{R}$. The norm on $Z_p(\mathbb{C})$ is given by

$$\forall (u, v) \in Z \times Z, \quad \|(u, v)\|_{Z_p(\mathbb{C})} \overset{\text{def}}{=} \left( \int_0^{2\pi} \|(\cos \theta)u - (\sin \theta)v\|^p_{Z^2} d\theta \right)^{\frac{1}{p}}. \quad (36)$$
This turns \((Z_p(\mathbb{C}), \|\cdot\|_{Z_p(\mathbb{C})})\) into a Banach space over the complex numbers, which is isometric as a real Banach space to a subspace of \(L_p([0, 2\pi], Z)\). For every \(z \in Z\) we have

\[
\|(z, 0)\|_{Z_p(\mathbb{C})}^p = \|z\|^p_Z \frac{1}{2\pi} \int_0^{2\pi} |\cos \theta|^p d\theta = \frac{\Gamma(\frac{p}{2} + 1/2)}{\Gamma(\frac{p}{2} + 1)} \|z\|^p_Z. \tag{37}
\]

Hence, by considering an appropriate rescaling of the first coordinate of elements of \(Z_p(\mathbb{C})\), we see that \(Z\) is isometric to a subspace of \(Z_p(\mathbb{C})\). Since \(Z_p(\mathbb{C})\) is a subspace of \(L_p([0, 2\pi], Z)\), all properties that are closed under \(\ell_p\) sums are inherited by \(Z_p(\mathbb{C})\) from \(Z\).

The final matter that will be treated in the present section is to show that the metric \(X_p\) inequality of Theorem 1.1 implies the linear \(X_p\) inequality \([15]\). We shall show this in the context of general Banach spaces, i.e., if a Banach space is an \(X_p\) metric space then it is also an \(X_p\) Banach space. The converse of this assertion, i.e., that an \(X_p\) Banach space is also an \(X_p\) metric space, will be proved in Section 4.

**Proposition 2.1** (Metric \(X_p\) inequalities imply linear \(X_p\) inequalities). Let \((Z, \|\cdot\|_Z)\) be a Banach space. Fix \(p \in [2, \infty)\) and \(\gamma \in (0, 1)\). Fix also \(m, n \in \mathbb{N}\) and \(k \in \{1, \ldots, n\}\). Suppose that for every \(f : Z_{2m}^n \rightarrow Z\) we have

\[
\frac{\gamma}{2^n(n!)} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{\varepsilon \in \{-1, 1\}} \sum_{x \in \mathbb{Z}_m^n} \left\|f(x + m\varepsilon) - f(x)\right\|_Z^p m^p \leq \frac{k}{n} \sum_{j=1}^n \|f(x + e_j) - f(x)\|_Z^p + \frac{(k/n)^{\frac{p}{2}}}{2^n} \sum_{\varepsilon \in \{-1, 1\}} \sum_{x \in \mathbb{Z}_m^n} \|f(x + \varepsilon) - f(x)\|_Z^p. \tag{38}
\]

Then for every \(z_1, \ldots, z_n \in Z\) we have

\[
\frac{(2/\pi)^{2p}}{2^n(n!)} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{\varepsilon \in \{-1, 1\}} \left\|\sum_{j \in S} \varepsilon z_j\right\|_Z^p \leq \frac{k}{n} \sum_{j=1}^n \|z_j\|_Z^p + \frac{(k/n)^{\frac{p}{2}}}{2^n} \sum_{\varepsilon \in \{-1, 1\}} \sum_{j=1}^n \|\varepsilon z_j\|_Z^p. \tag{39}
\]

**Proof.** Since (38) holds true in \(Z\), it also holds true in its \(p\)-complexification \(Z_p(\mathbb{C})\). Fixing \(z_1, \ldots, z_n \in Z\) and \(\delta \in \{-1, 1\}^n\), apply (38) to the function \(f_\delta : Z_{2m}^n \rightarrow Z_p(\mathbb{C})\) given by

\[
\forall x \in Z_{2m}^n, \quad f_\delta(x) \overset{\text{def}}{=} \sum_{j=1}^n \delta_j e^{\frac{\pi i x_j}{m}} (z_j, 0) \in Z \times Z.
\]

By averaging the resulting inequality over \(\delta \in \{-1, 1\}^n\), we deduce that

\[
\frac{2^p \gamma}{2^n(n!)} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{\delta \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_m^n} \left\|\sum_{j \in S} \delta_j e^{\frac{\pi i x_j}{m}} (z_j, 0)\right\|_{Z_p(\mathbb{C})}^p m^p \leq \frac{k(2m)^n}{n} \sum_{j=1}^n \left|1 - e^{\frac{\pi i}{m}}\right|^p \cdot \|z_j, 0\|_{Z_p(\mathbb{C})}^p \tag{40}
\]
where for the left-hand side of (40) we used the fact that $e^{x(x+\sigma \varepsilon)} - e^{x \varepsilon} = -2e^{x \varepsilon}$ for every $\sigma \in \{-1,1\}$ and $x \in \mathbb{Z}_{2m}$.

Recalling the definition of the norm of $Z_p(\mathbb{C})$, for every $S \subseteq \{1, \ldots, n\}$ we have

$$
\sum_{\delta \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_{2m}^n} \left\| \sum_{j \in S} \delta_j e^{\frac{\pi x_j}{m}} (z_j, 0) \right\|_Z^p = \sum_{x \in \mathbb{Z}_{2m}^n} \sum_{\delta \in \{-1,1\}^n} \left\| \sum_{j \in S} \delta_j \cos \left( \theta + \frac{\pi x_j}{m} \right) z_j \right\|_Z^p d\theta
$$

$$
= \sum_{x \in \mathbb{Z}_{2m}^n} \sum_{\delta \in \{-1,1\}^n} \left\| \sum_{j \in S} \delta_j \cos \left( \theta + \frac{\pi x_j}{m} \right) z_j \right\|_Z^p d\theta
$$

$$
\geq 2\pi (2m)^n \sum_{\delta \in \{-1,1\}^n} \left\| \sum_{j \in S} \frac{\pi}{2} \left( \int_0^{2\pi} \cos \theta \, d\theta \right) z_j \right\|_Z^p
$$

$$
= \frac{2^{p+1}(2m)^n}{\pi^{p-1}} \sum_{\delta \in \{-1,1\}^n} \left\| \sum_{j \in S} \delta_j z_j \right\|_Z^p,
$$

(42)

where in (41) we used Jensen’s inequality.

To bound the first term in the right-hand side of (40), use the fact that $|1 - e^{i\theta}| \leq \theta$ for every $\theta \in [0, \pi]$, and the identity (47) to get

$$
\sum_{j=1}^n \left| 1 - e^{\frac{\pi j}{m}} \right|^p \left\| (z_j, 0) \right\|_{Z_p(\mathbb{C})}^p \leq \frac{\pi^p}{m^p} \left( \int_0^{2\pi} |\cos \theta|^p d\theta \right) \sum_{j=1}^n \left\| z_j \right\|_Z^p \leq \frac{\pi^{p+1}^p}{m^p} \sum_{j=1}^n \left\| z_j \right\|_Z^p,
$$

(43)

where we used the fact that, since $p \geq 2$, we have $\int_0^{2\pi} |\cos \theta|^p d\theta \leq \int_0^{2\pi} \cos^2 \theta d\theta = \pi$. To bound the second term in the right-hand side of (40), recall the contraction principle (see [29 Thm. 4.4]), which asserts that for every $a_1, \ldots, a_n \in \mathbb{R}$ we have

$$
\sum_{\delta \in \{-1,1\}^n} \left\| \sum_{j=1}^n a_j \delta_j z_j \right\|_Z^p \leq \left( \max_{j \in \{1, \ldots, n\}} |a_j|^p \right) \sum_{\delta \in \{-1,1\}^n} \left\| \sum_{j=1}^n \delta_j z_j \right\|_Z^p.
$$

Hence, for every $x \in \mathbb{Z}_{2m}^n$ and $\varepsilon \in \{-1,1\}^n$ we have

$$
\sum_{\delta \in \{-1,1\}^n} \left\| \sum_{j=1}^n \delta_j \left( e^{\frac{\pi(x_j+\varepsilon_j)}{m}} - e^{\frac{\pi x_j}{m}} \right) (z_j, 0) \right\|_Z^p

= \int_0^{2\pi} \sum_{\delta \in \{-1,1\}^n} \left\| \sum_{j=1}^n \left( \cos \left( \theta + \frac{\pi x_j}{m} \right) + \frac{\pi \varepsilon_j}{m} \right) \right\|_Z^p \, d\theta

\leq 2\pi \left( \max_{\theta \in [0, \pi]} \max_{j \in \{1, \ldots, n\}} |\cos \left( \theta + \frac{\pi}{m} \right) - \cos \theta|^p \right) \sum_{\delta \in \{-1,1\}^n} \left\| \sum_{j=1}^n \delta_j z_j \right\|_Z^p

\leq \frac{2^{p+1}(2m)^n}{\pi^{p-1}} \sum_{\delta \in \{-1,1\}^n} \left\| \sum_{j=1}^n \delta_j z_j \right\|_Z^p,
$$

(44)

where we used the fact that $|\cos(\theta + \pi/m) - \cos \theta| = |\int_\theta^{\theta+\pi/m} \sin t \, dt| \leq \pi/m$. The desired inequality (39) now follows by combining (40), (42), (43) and (44). □
3. Nonembeddability

Here we assume for the moment the validity of Theorem 1.1 whose proof appears in Section 4 and proceed to deduce its geometric consequences that were stated in the Introduction. Namely, we will prove here Theorem 1.6, Theorem 1.9 and Theorem 1.12.

Proof of Theorem 1.6. We first make some preparatory elementary estimates that explain the origin of the quantities that appear in (23). Define \( \psi_{p,q} : \mathbb{R} \to \mathbb{R} \) by

\[
\psi_{p,q}(t) = \frac{3tp}{q} - 3 + \left( tp - \frac{2tp}{q} - p \right) \left( 1 + \frac{tp - q}{q(p-2)} \right) = \frac{p^2(q-2)}{q^2(p-2)} t^2 + \frac{p(pq - 3q + 2)}{q(p-2)} t - p.
\]

Then for every \( s \in (0,1) \) we have

\[
\frac{q^2(p-2)}{p} \cdot \psi_{p,q}(1-s) - (p-q)(q-2) = -(2pq + 2q - 4p + pq^2 - 3q^2)s + p(q-2)s^2
\]

\[
> -(2pq + 2q - 4p + pq^2)s > -(2p^2 - 2p + p^3)s > -2p^3s.
\]

Hence \( \psi_{p,q}(1-(p-q)(q-2)/(2p^3)) > 0 \). Note that \( \psi_{p,q}(0) = -p < 0 \) and \( \psi_{p,q}(q/p) = -(p-q) < 0 \). Since \( \psi \) is quadratic with \( \lim_{s \to \pm \infty} \psi_{p,q}(t) = \infty \), it follows that \( \psi_{p,q} \) has exactly one positive zero that lies in the interval \((q/p, 1-(p-q)(q-2)/(2p^3))\). One checks that \( \psi_{p,q}(\theta_{p,q}) = 0 \), where

\[
\theta_{p,q} = \frac{2q(p-q) + q^2(p-1)(p-2)}{2p^2(q-2)} \left( \sqrt{1 + \frac{4p(p-2)(q-2)}{(pq - 3q + 2)^2}} - 1 \right).
\]

Consequently, \( q/p < \theta_{p,q} < 1 - (p-q)(q-2)/(2p^3) \) (in particular, the rightmost inequality in (23) is valid), and

\[\forall \theta \in (0,1), \quad \psi_{p,q}(\theta) \leq 0 \implies \theta \leq \theta_{p,q}. \quad (45)\]

Now, suppose that \( (L_q, \|x-y\|_q^p) \) admits a bi-Lipschitz embedding into \( L_p \). If \( \theta \leq q/p < \theta_{p,q} \) then we are done, so we may assume below that \( \theta > q/p \). Since \( \ell_q(\mathbb{C}) \) embeds isometrically into \( L_q \), there exists \( \Lambda \in [1,\infty) \) such that for every \( m,n \in \mathbb{N} \) there is a mapping \( f_{m,n} : \mathbb{Z}_{4m}^n \to L_p \) that satisfies for every \( x,y \in \mathbb{Z}_{4m}^n \),

\[
\left( \sum_{j=1}^{n} e^{\frac{\pi i x_j}{2m}} - e^{\frac{\pi i y_j}{2m}} \right)^{\frac{\theta}{q}} \leq \| f_{m,n}(x) - f_{m,n}(y) \|_p \leq \Lambda \left( \sum_{j=1}^{n} e^{\frac{\pi i x_j}{2m}} - e^{\frac{\pi i y_j}{2m}} \right)^{\frac{\theta}{q}}. \quad (46)
\]

Suppose that \( m \geq n \) and \( m^2 \) divides \( n^3 \). Define \( k = n^3/m^2 \). By Theorem 1.1 and Remark 1.2 in conjunction with (46), we have

\[
\frac{\frac{\theta p}{q} - 3 + (\frac{\theta p}{q} - 2p)}{m^3} \leq (c(p)\Lambda)^p \left( \frac{k}{m^3} + \left( \frac{k}{n} \right)^{\frac{\theta}{p}} \cdot \frac{n^3}{m^3} \right) = (c(p)\Lambda)^p \left( \frac{n^3}{m^{3p+2}} + \frac{n^{3} + \frac{\theta p}{q}}{m^{(1+\theta)p}} \right), \quad (47)
\]

where \( c(p) \in (1,\infty) \) may depend only on \( p \).

Choose \( m \) to be the smallest integer that divides \( n^3 \) and

\[m \geq n^{\frac{p-3+\theta p/q}{p-2}} = n^{1 + \frac{\theta p - 2}{q(p-2)}}.\]

Observe that since \( \theta > q/p \) and \( p > 2 \) we have \( m \geq n \). The above choice of \( m \) ensures that

\[
\frac{n^3}{m^{3p+2}} + \frac{n^{3} + \frac{\theta p}{q}}{m^{(1+\theta)p}} \leq_p n^3 \frac{n^3}{m^{3p+2}},
\]

and therefore by (47) (and our choice of \( m \)) we have

\[
n^{\frac{3p}{q} - 3 + (\frac{\theta p}{q} + 2p - 2p)} \left( 1 + \frac{\theta p - 2}{q(p-2)} \right) \leq_p (c(p)\Lambda)^p. \quad (48)
\]
Lemma 3.1. For every two integers $m, n \geq 2$ there exists a mapping $h_m^n : \mathbb{Z}_m \to \{0, \ldots, 4m\}^{2n}$ such that for every $q \in [2, \infty)$ and $x, y \in \mathbb{Z}_m$ we have

$$m \left( \sum_{j=1}^n \left| e^{\frac{2\pi i x_j}{m}} - e^{\frac{2\pi i y_j}{m}} \right|^q \right)^{\frac{1}{q}} \leq \|h_m^n(x) - h_m^n(y)\|_q \leq 3m \left( \sum_{j=1}^n \left| e^{\frac{2\pi i x_j}{m}} - e^{\frac{2\pi i y_j}{m}} \right|^q \right)^{\frac{1}{q}}.$$

Proof. For every $u \in \mathbb{Z}_m$ choose $a_m(u), b_m(u) \in \{0, \ldots, 4m\}$ such that

$$2m + 2m \cos \left( \frac{2\pi u}{m} \right) - a_m(u) \leq \frac{1}{2} \quad \text{and} \quad 2m + 2m \sin \left( \frac{2\pi u}{m} \right) - b_m(u) \leq \frac{1}{2}.$$

Then, for every distinct $u, v \in \mathbb{Z}_m$ we have

$$ (|a_m(u) - a_m(v)|^q + |b_m(u) - b_m(v)|^q)^{\frac{1}{q}} \leq \sqrt{|a_m(u) - a_m(v)|^2 + |b_m(u) - b_m(v)|^2} \leq 2m \left| e^{\frac{2\pi i u}{m}} - e^{\frac{2\pi i v}{m}} \right| + \frac{2}{\sqrt{2}} \leq 3m \left| e^{\frac{2\pi i u}{m}} - e^{\frac{2\pi i v}{m}} \right|,$$

since for distinct $u, v \in \mathbb{Z}_m$ we have $\left| e^{\frac{2\pi i u}{m}} - e^{\frac{2\pi i v}{m}} \right| \geq e^{\frac{2\pi i}{m}} - 1 \geq \frac{4}{m}$. Similarly,

$$ (|a_m(u) - a_m(v)|^q + |b_m(u) - b_m(v)|^q)^{\frac{1}{q}} \geq \frac{1}{\sqrt{2}} \sqrt{|a_m(u) - a_m(v)|^2 + |b_m(u) - b_m(v)|^2} \geq \frac{2 - \frac{\sqrt{2}}{2}}{\sqrt{2}} m \left| e^{\frac{2\pi i u}{m}} - e^{\frac{2\pi i v}{m}} \right| \geq m \left| e^{\frac{2\pi i u}{m}} - e^{\frac{2\pi i v}{m}} \right|.$$

Hence $h_m^n(x) \overset{\text{def}}{=} (a_m(x_1), b_m(x_1), a_m(x_2), b_m(x_2), \ldots, a_m(x_n), b_m(x_n))$ has the desired property. □

Proof of Theorem 1.9 We shall show that for an appropriate choice of $\beta_p \in (0, \infty)$ we have

$$m \geq n^{1+\frac{p-q}{(p-q)q}} \implies c_p \left(16m^{2n}\right) \geq \beta_p m^{\frac{(p-q)(q-1)}{q(p-2)}}. \quad (49)$$

Since $[M]^n_q \succeq [n]^n_q$ for every integer $M \geq m$ and $[m]^n_N$ contains an isometric copy of $[m]^n_q$ for every integer $N \geq n$, the validity of (49) implies the desired estimate (24).

Fix $D \in [1, \infty)$ and suppose that $f : [16m]^m_q \to L_p$ satisfies

$$\forall x, y \in [16m]^m_q, \quad \|x - y\|_q \leq \|f(x) - f(y)\|_p \leq D\|x - y\|_q. \quad (50)$$

Our goal is to bound $D$ from below. Define $F : \mathbb{Z}_m^{4n} \to L_p$ by $F = f \circ h_m^n$, where $h_m^n$ is the mapping from Lemma 3.1. Then for every $x \in \mathbb{Z}_m^n$, every $j \in \{1, \ldots, n\}$, every $\varepsilon \in \{-1, 1\}^n$ and every $S \subseteq \{1, \ldots, n\}$ we have

$$\|F(x + e_j) - F(x)\|_p \leq 3mD \left| e^{\frac{\pi i x_j}{m}} - 1 \right| \lesssim D, \quad (51)$$

$$\|F(x + \varepsilon) - F(x)\|_p \leq 3mD \left( \sum_{j=1}^n \left| e^{\frac{\pi i x_j}{m}} - 1 \right|^q \right)^{\frac{1}{q}} \lesssim Dn^{\frac{1}{q}}, \quad (52)$$

Since (48) is supposed to hold true for $n$ that can be arbitrarily large, we necessarily have

$$\psi_{p,q}(\theta) = \frac{3\theta p}{q} - 3 + \left( \theta p + 2 - \frac{2\theta p}{q} - p \right) \left( 1 + \frac{\theta p - q}{q(p-2)} \right) \leq 0.$$

Recalling (49), this implies that $\theta \leq \theta_{p,q}$, as required. □
and

$$\|F(x + 2m\epsilon S) - F(x)\|_p \geq m \left( \sum_{j \in S} |e^{\pi i j} - 1|^q \right)^{\frac{1}{q}} \gtrsim m |S|^\frac{1}{q}. \quad (53)$$

Denote

$$k = \left\lceil \frac{\ell_p(q-2)}{n^{3/2}/\sqrt{k}} \right\rceil. \quad (54)$$

Then $k \leq n$ and the assumption on $m$ in (49) implies that $m \geq n^{3/2}/\sqrt{k}$. Hence, by Theorem 1.1 and Remark 1.2 combined with (51), (52) and (53), there exists $K_p \in (0, \infty)$ such that

$$\frac{\ell_p^2(q-2)}{n^{3/2}(p-2)} \leq k \frac{\ell_p}{n^{3/2}} \leq K_p D^p \left( k + \frac{k^{p/2}}{n^{3/2}} \right) \lesssim K_p D^p n^{3/2(p-2)}.$$  

Consequently,

$$D \gtrsim \frac{n^{(p-q)(q-2)}}{K_p n^{3/2(p-2)}}. \quad \square$$

**Remark 3.2.** Lower bounds on $c_p([m]^n_q)$ that are weaker than those of Theorem 1.9 can also be deduced from general discretization principles (combined with the asymptotic computation of $c_p(\ell^n_q)$ in (49)), namely from Bourgain’s discretization theorem [9] and its quantitative improvement for $L_p$ spaces in [21]. Specifically, let $B^n_q$ denote the unit ball of $\ell^n_q$. Observe that $\frac{1}{n} (-m, \ldots, m)^n$ contains a $\delta$-dense subset of $B^n_q$, with $\delta \leq n^{1/4}/m$. By Theorem 1.3 in [21] (and the discussion immediately following it) we see that there exists a universal constant $\gamma \in (0,1)$ such that if

$$\frac{n^{1/4}}{m} \leq \frac{\gamma}{\sigma(p,q)n^{2+\frac{(p-q)(q-2)}{q^2(q-2)}}} \leq \frac{\beta}{n^2 c_p(\ell^n_q)}$$

then

$$c_p([2m]^n_q) \geq \frac{c_p(\ell^n_q)}{2} \gtrsim \sigma(p,q)n^{\frac{(p-q)(q-2)}{q^2(q-2)}},$$

where $\sigma(p,q) \in (0,\infty)$ as in (9). Consequently,

$$m \geq \frac{\sigma(p)}{\gamma} \cdot n^{2+\frac{1}{q} + \frac{(p-q)(q-2)}{q^2(q-2)}} = \frac{\sigma(p)}{\gamma} \cdot n^{2+\frac{p-q}{q(p-2)} + \frac{(q-2)}{q^2(q-2)}} \implies c_p([2m]^n_q) \gtrsim \sigma(p)n^{\frac{(p-q)(q-2)}{q^2(q-2)}}. \quad (55)$$

We note that a direct application of Bourgain’s discretization theorem [9] (which holds true also for target spaces that need not be $L_p$ spaces) would imply the same bound on $c_p([2m]^n_q)$ as in (55), provided that $m$ is much larger than the requirement appearing in (55) (specifically, $m$ would have to be at least doubly exponential in $n \log n$).

**Proof of Theorem 1.12.** The proof follows the proofs of Theorem 1.6 and Theorem 1.9 with a different (optimal) setting of parameters that is made possible due to the assumed validity of Conjecture 1.4 Specifically, we are now assuming that (20) holds true provided $m \geq C_p \sqrt{n/k}$.

Dealing first with (30), fix $\theta \in (q/p, 1]$ and $n \in \mathbb{N}$. Choose $m, k \in \mathbb{N}$ as follows.

$$m \overset{\text{def}}{=} \left\lceil \frac{\sqrt{p-q}}{n^{3/2}/\sqrt{k}} \right\rceil \quad \text{and} \quad k \overset{\text{def}}{=} \left\lceil \frac{C_p^2 n}{m^2} \right\rceil. \quad (56)$$

Since $\theta > q/p$ we may assume that $n$ is large enough so that $m \geq C_p$, in which case we have $k \in \{1, \ldots, n\}$ and $m \geq C_p \sqrt{n/k}$. Suppose for the sake of obtaining a contradiction that there
exists \( f_{m,n} : \mathbb{Z}_{4m}^n \to L_p \) satisfying (46). An application of (20) then yields the following estimate.

\[
\alpha_p C_p^{2n/p} \frac{p(q^2 - 2\theta^2 p)}{n^{q/2}} \frac{p^q}{m^{q/p}} \leq \alpha_p k^{\frac{n}{q}} \frac{n^{q}}{m^{q/p}} \leq \Lambda^p \left( \frac{k}{m^{\theta p}} + \left( \frac{k}{n} \right)^{q/p} \right) \cdot \frac{n^{q}}{m^{q/p}} \leq \frac{50}{p} \left( C_p \Lambda \right)^p n^{1 - (2 + \theta p) \frac{q/2 - q}{q/2}}.
\]  

(57)

Since (57) holds true for arbitrarily large \( n \), we conclude that

\[
\frac{p(q^2 - 2\theta^2 p)}{q^2(p - 2)} \leq 1 - (2 + \theta p) \frac{q - \theta p}{q(p - 2)} = \frac{p(q^2 - 2\theta^2 p)}{q^2(p - 2)} - \frac{\theta(p - q)}{q(p - 2)}.
\]

Consequently \( \theta \leq q/p \), contradicting the initial assumption that \( \theta > q/p \). This proves (30).

Next, we have already seen in (27) that \( c_p([m]_q^n) \) is bounded from above by a constant multiple of the quantity appearing in (28). By arguing as in the beginning of the proof of Theorem 1.9, it therefore suffices to show that for every \( m, n \in \mathbb{N} \) we have

\[
m \geq \frac{n^{\frac{p-q}{4(p-2)}}}{\xi(p)n^{\frac{(p-q)(q-2)}{2(p-2)}}} \implies c_p([16m]_q^n) \geq \xi(p)n^{\frac{p(q-2)}{2(p-2)}}.
\]

(58)

for some \( \xi(p) \in (0, \infty) \). To this end, suppose that there exists \( f : [16m]_q^n \to L_p \) satisfying (50), our goal being to bound \( D \) from below. As explained in the proof of Theorem 1.9, this implies the existence of \( F : \mathbb{Z}_{4m}^n \to L_p \) that satisfies (51), (52) and (53). Similarly to (54), choose \( k \in \mathbb{N} \) to be

\[
k = \left( C_p^2 n^{\frac{p(q-2)}{2(p-2)}} \right).
\]

(59)

We may suppose that \( n \) is large enough so that \( k \in \{1, \ldots, n\} \), since otherwise (58) is vacuous. The lower bound on \( m \) that is assumed in (58) implies that \( m \geq C_p n^{1/k} \), so we may apply (20), yielding, in conjunction with (51), (52) and (53), that the following holds true.

\[
\alpha_p C_p^3 n^{\frac{p^2(q-2)}{q^2(p-2)}} \frac{p^q}{n^{q/2}} \leq \alpha_p k \frac{n^{q}}{n^{q/2}} \leq \Lambda^p \left( \frac{k}{m^{\theta p}} + \left( \frac{k}{n} \right)^{q/p} \right) \cdot \frac{n^{q}}{m^{q/p}} \leq \frac{50}{p} \left( C_p \Lambda \right)^p n^{1 - (2 + \theta p) \frac{q/2 - q}{q/2}}.
\]

(60)

\[\square\]

4. Proof of Theorem 1.1

Suppose from now on that \( m, n \in \mathbb{N} \) satisfy \( m \geq n \) and that \( R \in [n, 2m] \) is an odd integer. In what follows we shall use the canonical identification of \( \mathbb{Z}_{4m}^n \) with \([-2m + 1, 2m - 1]^n \cap \mathbb{Z}^n \). Fix \( S \subseteq \{1, \ldots, n\} \) and define \( U_S \subseteq \mathbb{Z}_{4m}^n \) by

\[
U_S \defeq \{ y \in [-R, R]^n : \forall (i,j) \in S \times \{(1, \ldots, n) \setminus S\}, \quad (y_i, y_j) \in (2\mathbb{Z}) \times (1 + 2\mathbb{Z}) \}.
\]

(61)

Thus \( U_S \) consists of those \( y \in \mathbb{Z}_{4m}^n \) satisfying \( |y_j| \leq R \) for every \( j \in \{1, \ldots, n\} \), and such that \( y_j \) is even for every \( j \in S \) and \( y_j \) is odd for every \( j \in \{1, \ldots, n\} \setminus S \). Observe that since \( R \) is odd, for every \( y \in U_S \) we actually have \( |y_j| < R \) if \( j \in S \). Hence \( |U_S| = R^{1-n} = |S| \). Given a Banach space \( (X, \| \cdot \|_X) \), the averaging operator corresponding to \( U_S \) will be denoted below by \( D_S : L_2(\mathbb{Z}_{4m}^n, X) \to L_2(\mathbb{Z}_{4m}^n, X) \), i.e., for every \( f : \mathbb{Z}_{4m}^n \to X \) and \( x \in \mathbb{Z}_{4m}^n \) we set

\[
D_S f(x) \defeq \frac{1}{|U_S|} \sum_{y \in U_S} f(x + y).
\]

The following lemma extends Lemma 5.1 in [39], which corresponds to the special case \( |S| = 1 \).
Lemma 4.1. Suppose that \((X, \| \cdot \|_X)\) is a Banach space and \(p \in [1, \infty)\). For every \(f : \mathbb{Z}_{4m}^n \rightarrow X\),
\[
\sum_{x \in \mathbb{Z}_{4m}^n} \|f(x) - D_s f(x)\|_X^p
\]
\[
\lesssim_p \frac{R^p}{2^n} \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_{4m}^n} \|f(x + \varepsilon) - f(x)\|_X^p + \frac{1}{2^n} \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_{4m}^n} \|f(x + \varepsilon S) - f(x)\|_X^p \quad (62)
\]

Proof. For every \(w \in \mathbb{Z}^n\) all of whose coordinates are odd fix \(\gamma_w : \mathbb{N} \cup \{0\} \rightarrow \mathbb{Z}^n\) that satisfies \(\gamma_w(0) = 0\), \(\gamma_w(\|w\|_\infty) = w\) and \(\gamma_w(t) - \gamma_w(t-1) \in \{-1, 1\}^n\) for every \(t \in \mathbb{N}\). The existence of such \(\gamma_w\) is explained in [39, Lem. 51], and we shall quickly recall now why this is so for the sake of completeness. We may assume without loss of generality that all the coordinates of \(w\) are positive, since for general \(w\) we could then define \(\gamma_w = \text{sign}(w) \cdot \gamma_{|w|}\), where the multiplication is coordinate-wise and we denote \(\text{sign}(w) = (\text{sign}(w_1), \ldots, \text{sign}(w_n))\) and \(|w| = (|w_1|, \ldots, |w_n|)\). Now, supposing that all the coordinates of \(w\) are positive, define \(\gamma_w(0) = 0\) and, inductively, for every \(t \in \mathbb{N}\) such that \(\gamma_w(2t - 2)\) has already been defined, set
\[
\gamma_w(2t - 1) = \gamma_w(2t - 2) + \sum_{j=1}^n e_j \quad \text{and} \quad \gamma_w(2t) = \gamma_w(2t - 1) + \sum_{j \in \{1, \ldots, n\}, \gamma_w(2t - 2) < w_j} e_j - \sum_{j \in \{1, \ldots, n\}, \gamma_w(2t - 2) = w_j} e_j.
\]

This explicit definition of \(\gamma_w\) is not used below; we shall only need to know that \(\gamma_w\) exists, and that, as our construction guarantees, we have \(\varepsilon \gamma_w = \gamma_{\varepsilon w}\) for every \(\varepsilon \in \{-1, 1\}^n\). Note that, since the restriction of \(\gamma_w\) to \(\{0, \ldots, \|w\|_\infty\}\) is an \(\ell_\infty\) geodesic joining 0 and \(w\), for every distinct \(s, t \in \{0, \ldots, \|w\|_\infty\}\) we have \(\gamma_w(s) \neq \gamma_w(t)\).

If \(y \in U_S\) and \(\eta \in \{-1, 1\}^n\) then all that coordinates of \(y - \eta S\) are odd, and we can therefore consider \(\gamma_{y - \eta S}\). For every \(x \in \mathbb{Z}_{4m}^n\) define \(\gamma_{x,y} : \mathbb{N} \rightarrow \mathbb{Z}^n\) by \(\gamma_{x,y}(t) = x + \eta S + \gamma_{y - \eta S}\). Thus \(\gamma_{x,y}(0) = x + \eta S\), \(\gamma_{x,y}(\|y - \eta S\|_\infty) = x + y\) and \(\gamma_{x,y}(t) - \gamma_{x,y}(t-1) \in \{-1, 1\}^n\) for all \(t \in \mathbb{N}\). Note that \(\gamma_{x,y}(t)\) depends only on those coordinates of \(\eta\) that belong to \(S\).

For every \(z \in \mathbb{Z}_{4m}^n\) and \(\varepsilon, \eta \in \{-1, 1\}^n\) define
\[
F_{\eta}(z, \varepsilon) \overset{\text{def}}{=} \{(x, y) \in \mathbb{Z}_{4m}^n \times U_S : \gamma_{x,y}(t-1) = z \quad \text{and} \quad \gamma_{x,y}(t) = z + \varepsilon \quad \text{for some} \quad t \in [1, \|y - \eta S\|_\infty]\}.
\]

Observe that for every \((x, y) \in \mathbb{Z}_{4m}^n \times U_S\) and \(\eta \in \{-1, 1\}^n\) there is at most one \(t \in \{1, \ldots, \|y - \eta S\|_\infty\}\) for which \(\gamma_{x,y}(t-1) = z\) and \(\gamma_{x,y}(t) = z + \varepsilon\).

We claim that
\[
N \overset{\text{def}}{=} \sum_{\eta \in \{-1, 1\}^n} |F_{\eta}(z, \varepsilon)| \quad (63)
\]
is independent of \(z \in \mathbb{Z}_{4m}^n\) and \(\varepsilon, \eta \in \{-1, 1\}^n\). Indeed, for every \(\varepsilon, \delta \in \{-1, 1\}^n\) and \(z, w \in \mathbb{Z}_{4m}^n\) define a bijection \(\psi_{z,w}^\varepsilon \delta : \mathbb{Z}_{4m}^n \times U_S \rightarrow \mathbb{Z}_{4m}^n \times U_S\) by
\[
\psi_{z,w}^\varepsilon \delta(x, y) \overset{\text{def}}{=} (w - \varepsilon \delta z + \varepsilon \delta x, \varepsilon \delta y).
\]

Then for every \(\eta \in \{-1, 1\}^n\) we have \(\gamma_{z,w}^\varepsilon \delta = \gamma_{w,\delta}^\varepsilon \delta\). Consequently,
\[
(\gamma_{x,y}(t-1), \gamma_{x,y}(t)) = (z, z + \varepsilon) \iff (\gamma_{z,w}^\varepsilon \delta(x, y)(t-1), \gamma_{z,w}^\varepsilon \delta(x, y)(t)) = (w, w + \delta)
\]
for every \(t \in \{1, \ldots, \|y - \eta S\|_\infty\}\). This shows that for every \(\eta \in \{-1, 1\}^n\) the mapping \(\psi_{z,w}^\varepsilon \delta\) is a bijection between \(F_{\eta}(z, \varepsilon)\) and \(F_{\varepsilon \delta \eta}(w, \delta)\), whence \(|F_{\eta}(z, \varepsilon)| = |F_{\varepsilon \delta \eta}(w, \delta)|\). Consequently,
\[
\sum_{\eta \in \{-1, 1\}^n} |F_{\eta}(z, \varepsilon)| = \sum_{\eta \in \{-1, 1\}^n} |F_{\varepsilon \delta \eta}(w, \delta)| = \sum_{\eta \in \{-1, 1\}^n} |F_{\eta}(w, \delta)|,
\]
20.
implying that the integer $N$ defined in (63) is indeed independent of $(z, \varepsilon) \in \mathbb{Z}_4^m \times \{-1,1\}^n$.

We shall need an estimate on $N$, which is proved by double counting as follows.

$$N(8m)^n = \sum_{z \in \mathbb{Z}_4^m} \sum_{\varepsilon, \eta \in \{-1,1\}^n} |F_y(z, \varepsilon)|$$

$$\leq \sum_{z \in \mathbb{Z}_4^m} \sum_{\varepsilon, \eta \in \{-1,1\}^n} \sum_{(x,y) \in \mathbb{Z}_4^m \times U_S} \sum_{t=1}^{\|y-\eta S\|_{\infty}} \mathbb{1}_{\{\gamma_{x,y}^n(t-1) = z \wedge \gamma_{x,y}^n(t) = z+\varepsilon\}}$$

$$\leq \sum_{\varepsilon, \eta \in \{-1,1\}^n} \sum_{(x,y) \in \mathbb{Z}_4^m \times U_S} \|y - \eta S\|_{\infty}$$

$$\leq R(8m)^n |U_S|.$$

Consequently,

$$N \leq R|U_S|. \tag{64}$$

Now, fix $f : \mathbb{Z}_4^m \to X$. For every $x \in \mathbb{Z}_4^m$, $y \in U_S$ and $\eta \in \{-1,1\}^n$ we have

$$\|f(x) - f(x + y)\|_X^p$$

$$\leq_p \|f(x) - f(x + \eta S)\|_X^p + \|f(\gamma_{x,y}^n(0)) - f(\gamma_{x,y}^n(\|y-\eta S\|_{\infty}))\|_X^p$$

$$\leq_p \|f(x) - f(x + \eta S)\|_X^p + \|y - \eta S\|_{\infty}^{-1} \sum_{t=1}^{\|y-\eta S\|_{\infty}} \|f(\gamma_{x,y}^n(t-1)) - f(\gamma_{x,y}^n(t))\|_X^p$$

By averaging this inequality over $\eta \in \{-1,1\}^n$ we see that

$$\|f(x) - f(x + y)\|_X^p$$

$$\leq_p \frac{1}{2^n} \sum_{\eta \in \{-1,1\}^n} \|f(x) - f(x + \eta S)\|_X^p + \frac{R^{p-1}}{2^n} \sum_{\eta \in \{-1,1\}^n} \sum_{t=1}^{\|y-\eta S\|_{\infty}} \|f(\gamma_{x,y}^n(t-1)) - f(\gamma_{x,y}^n(t))\|_X^p.$$
Lemma 4.2. Suppose that \((X, d_X)\) is a metric space and \(p \in [1, \infty)\). Then for every \(f : \mathbb{Z}^n_{4m} \to X\) and \(\varepsilon \in \{-1, 1\}^n\) we have

\[
\sum_{x \in \mathbb{Z}^n_{4m}} d_X(f(x + \varepsilon), f(x))^p \leq |S|^{p-1} \sum_{j \in S} \sum_{z \in \mathbb{Z}^n_{4m}} d_X(f(z + e_j), f(z))^p. \tag{65}
\]

Proof. Write \(S = \{j(1), \ldots, j(|S|)\}\) and for every \(\ell \in \{0, \ldots, |S|\}\) denote \(\Sigma(\ell) = \{j(1), \ldots, j(\ell)\}\) (with the convention \(\Sigma(0) = \emptyset\)). Then by the triangle inequality and Hölder’s inequality, for every \(\varepsilon \in \{-1, 1\}^n\) we have

\[
d_X(f(x + \varepsilon), f(x))^p \leq |S|^{p-1} \sum_{\ell=1}^{|S|} d_X(f(x + \varepsilon_{\Sigma(\ell-1)} + \varepsilon_{j(\ell)} e_{j(\ell)}), f(x + \varepsilon_{\Sigma(\ell-1)}))^p.
\]

Hence,

\[
\sum_{x \in \mathbb{Z}^n_{4m}} d_X(f(x + \varepsilon), f(x))^p \leq |S|^{p-1} \sum_{\ell=1}^{|S|} \sum_{y \in \mathbb{Z}^n_{4m}} d_X(f(y + \varepsilon_{j(\ell)} e_{j(\ell)}), f(y))^p = |S|^{p-1} \sum_{j \in S} \sum_{z \in \mathbb{Z}^n_{4m}} d_X(f(z + e_j), f(z))^p. \tag{66}
\]

Lemma 4.3. Suppose that \((X, \| \cdot \|_X)\) is a Banach space, \(p \in [1, \infty)\) and \(k \in \{1, \ldots, n\}\). Then for every \(f : \mathbb{Z}^n_{4m} \to X\) and \(\delta \in \{-1, 1\}^n\) we have

\[
\frac{1}{\binom{n}{k}} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{x \in \mathbb{Z}^n_{4m}} \|f(x + 2m \delta) - f(x)\|_X^p \lesssim_p m^p \frac{R^p}{2^n} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{x \in \mathbb{Z}^n_{4m}} \|Dsf(x + 2 \delta) - Dsf(x)\|_X^p
\]

\[
+ \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}^n_{4m}} \|f(x + \varepsilon) - f(x)\|_X^p + \sum_{j=1}^n \sum_{x \in \mathbb{Z}^n_{4m}} \|f(x + e_j) - f(x)\|_X^p. \tag{66}
\]

Proof. For every \(S \subseteq \{1, \ldots, n\}\) with \(|S| = k\) we have

\[
\sum_{x \in \mathbb{Z}^n_{4m}} \|f(x + 2m \delta) - f(x)\|_X^p \lesssim_p \sum_{x \in \mathbb{Z}^n_{4m}} \|Dsf(x + 2m \delta) - Dsf(x)\|_X^p
\]

\[
+ \sum_{x \in \mathbb{Z}^n_{4m}} \|Dsf(x + 2m \delta) - f(x + 2m \delta)\|_X^p + \sum_{x \in \mathbb{Z}^n_{4m}} \|Dsf(x) - f(x)\|_X^p
\]

\[
= \sum_{x \in \mathbb{Z}^n_{4m}} \|Dsf(x + 2m \delta) - Dsf(x)\|_X^p + \sum_{x \in \mathbb{Z}^n_{4m}} \|Dsf(x) - f(x)\|_X^p. \tag{67}
\]
The first term in (67) can be bounded as follows.

\[ \sum_{x \in \mathbb{Z}_4^{4m}} \| D_S f(x + 2m\delta_S) - D_S f(x) \|_X^p \leq m^{p-1} \sum_{t=1}^{m} \sum_{x \in \mathbb{Z}_4^{4m}} \| D_S f(x + 2t\delta_S) - D_S f(x + (2t - 2)\delta_S) \|_X^p \]

\[ = m^p \sum_{x \in \mathbb{Z}_4^{4m}} \| D_S f(x + 2\delta_S) - D_S f(x) \|_X^p. \] (68)

The second term in (67) is bounded using Lemma 4.1 and Lemma 4.2 as follows.

\[ \sum_{x \in \mathbb{Z}_4^{4m}} \| D_S f(x) - f(x) \|_X^p \]

\[ \lesssim_p \frac{R^p}{2^n} \sum_{\varepsilon \in \{-1, 1\}^n} x \in \mathbb{Z}_4^{4m} \| f(x + \varepsilon) - f(x) \|_X^p + |S|^{p-1} \sum_{j \in S} \sum_{x \in \mathbb{Z}_4^{4m}} \| f(x + e_j) - f(x) \|_X^p. \] (69)

Note that for every \( x \in \mathbb{Z}_4^{4m} \),

\[ \frac{1}{\binom{n}{k}} \sum_{S \subseteq \{1, \ldots, n\}, |S| = k} \sum_{j \in S} \| f(x + e_j) - f(x) \|_X^p = \frac{k}{n} \sum_{j=1}^{n} \| f(x + e_j) - f(x) \|_X^p. \]

Hence, the desired inequality (66) follows by substituting (68) and (69) into (67) and averaging the resulting inequality over all \( S \subseteq \{1, \ldots, n\} \) with \( |S| = k \). \( \square \)

Our next goal is to bound the first term in the right-hand side of (66). To this end we first recall some results from [20].

Fixing a Banach space \( (X, \| \cdot \|_X) \), consider the averaging operator \( A : L_2(\mathbb{Z}_4^{4m}, X) \rightarrow L_2(\mathbb{Z}_4^{4m}, X) \) given, for every \( f : \mathbb{Z}_4^{4m} \rightarrow X \) and \( x \in \mathbb{Z}_4^{4m} \), by

\[ Af(x) \overset{\text{def}}{=} \frac{1}{R^n} \sum_{y \in (-R, R)^n \cap (2\mathbb{Z})^n} f(x + y). \] (70)

For \( j \in \{1, \ldots, n\} \) denote \( B_j = D_{(j)} \), i.e., \( B_j \) is the averaging operator corresponding to the set \( U_{(j)} \), which consists of those \( y \in [-R, R]^n \) such that \( y_j \) is even and \( y_{\ell} \) is odd for every \( \ell \in \{1, \ldots, n\} \setminus \{j\} \).

(In [20] the set \( U_{(j)} \) was denoted \( S_{(j, R)} \) and the operator \( B_j \) was denoted \( \mathcal{E}_j \).)

It follows from [20] that for every \( f : \mathbb{Z}_4^{4m} \rightarrow X \), every \( p \in [1, \infty) \) and every \( \varepsilon \in \{-1, 1\}^n \) we have

\[ \sum_{x \in \mathbb{Z}_4^{4m}} \left\| \left( \frac{R}{R + 1} \right)^{n-1} (Af(x + \varepsilon) - Af(x - \varepsilon)) - \sum_{j=1}^{n} \varepsilon_j [B_j f(x + e_j) - B_j f(x - e_j)] \right\|_X^p \]

\[ \lesssim_p \frac{R^p}{p^n} \sum_{s=0}^{n-1} \binom{n}{s} \frac{n(R)^{(n-s)p}}{(s)^p} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{x \in \mathbb{Z}_4^{4m}} \| f(x + 2\varepsilon S) - f(x) \|_X^p. \] (71)

Since (71) is only implicit in [20] (it follows from proofs in [20] rather than explicit statements in [20]), we shall now explain how to establish (71).

**Proof of (71).** For every \( T \subseteq \{1, \ldots, n\} \) define \( L_T \subseteq \mathbb{Z}_4^{4m} \) by

\[ L_T \overset{\text{def}}{=} \{ y \in (-R, R)^n : \forall (i, j) \in T \times (\{1, \ldots, n\} \setminus T), (y_i, y_j) \in 2\mathbb{Z} \times \{0\} \}. \]
Thus $L_T$ consists of those $y \in (-R,R)^n$ all of whose coordinates are even, and all of whose coordinates that lie outside $T$ vanish. As in [20, Def. 3.2], we let $\Delta_T : L_2(\mathbb{Z}_4^m, X) \to L_2(\mathbb{Z}_4^m, X)$ denote the averaging operator corresponding to $L_T$, i.e., for every $f : \mathbb{Z}_4^m \to X$ and $x \in \mathbb{Z}_4^m$,

$$\Delta_T f(x) \overset{\text{def}}{=} \frac{1}{|L_T|} \sum_{y \in L_T} f(x + y).$$

We note in passing that the operator $A$ given in (70) coincides with $\Delta\{1,\ldots,n\}$.

For every $\alpha \in \{0,\ldots,n\}$ and $\beta \in \{0,\ldots,\alpha\}$ define $V^\varepsilon_{\alpha,\beta} : L_2(\mathbb{Z}_4^m, X) \to L_2(\mathbb{Z}_4^m, X)$ by setting for every $f : \mathbb{Z}_4^m \to X$ and $x \in \mathbb{Z}_4^m$,

$$V^\varepsilon_{\alpha,\beta} f(x) \overset{\text{def}}{=} \sum_{T \subseteq \{1,\ldots,n\} \delta \in \{-1,1\}^{\{1,\ldots,n\} \setminus T}} \sum_{|T| = n - \alpha} \frac{|\Delta_T f(x + \varepsilon T) - \Delta_T f(x + R\delta - \varepsilon T)|}{2^\alpha}.$$

Here $\langle \cdot, \cdot \rangle$ denotes the standard scalar product on $\mathbb{R}^n$. It is worthwhile to compare the right-hand side of (72) to the right-hand side of equation (44) in [20] (however, note that there is a difference of a normalization factor. Our $R$ is the same as the parameter $k$ of [20]). By combining Lemma 3.8 of [20] with Lemma 3.5 of [20] and identity (44) of [20] we see that for every $\alpha \in \{0,\ldots,n\}$ and $\beta \in \{0,\ldots,\alpha\}$ there exists $h_{\alpha,\beta} \in \mathbb{R}$ (related to the bivariate Bernoulli numbers; see [20, Sec. 3.1]) such that $h_{0,0} = 1$,

$$\forall \alpha \in \{0,\ldots,n\}, \forall \beta \in \{0,\ldots,\alpha\}, \quad |h_{\alpha,\beta}| \lesssim \frac{(\alpha - \beta)!|\beta|!}{2^\alpha}, \quad (73)$$

and for every $f : \mathbb{Z}_4^m \to X$ and $x \in \mathbb{Z}_4^m$,

$$\sum_{j=1}^n \varepsilon_j [B_j f(x + e_j) - B_j f(x - e_j)] = \left( \frac{R}{R + 1} \right)^{n-1} \sum_{\alpha = 0}^n \sum_{\beta = 0}^\alpha \frac{h_{\alpha,\beta}}{R^\alpha} V^\varepsilon_{\alpha,\beta} f(x). \quad (74)$$

Observe that $V^\varepsilon_{0,0} f(x) = Af(x + \varepsilon) - Af(x - \varepsilon)$, so it follows from (74) that

$$\left\| \sum_{j=1}^n \varepsilon_j [B_j f(x + e_j) - B_j f(x - e_j)] - \left( \frac{R}{R + 1} \right)^{n-1} (Af(x + \varepsilon) - Af(x - \varepsilon)) \right\|_X \overset{\text{74}}{\lesssim} \sum_{\alpha = 1}^n \sum_{\beta = 0}^\alpha \frac{|h_{\alpha,\beta}|}{R^\alpha} \|V^\varepsilon_{\alpha,\beta} f(x)\|_X \lesssim \sum_{\alpha = 1}^n 2^{\alpha} \sum_{\beta = 0}^\alpha \frac{(\alpha - \beta)!|\beta|!}{R^\alpha} \|V^\varepsilon_{\alpha,\beta} f(x)\|_X. \quad (75)$$

By convexity, it follows from (75) that

$$\left\| \sum_{j=1}^n \varepsilon_j [B_j f(x + e_j) - B_j f(x - e_j)] - \left( \frac{R}{R + 1} \right)^{n-1} (Af(x + \varepsilon) - Af(x - \varepsilon)) \right\|_X \overset{\text{74}}{\lesssim} \sum_{\alpha = 1}^n 2^{\alpha} \left( \sum_{\beta = 0}^\alpha \frac{(\alpha - \beta)!|\beta|!}{R^\alpha} \|V^\varepsilon_{\alpha,\beta} f(x)\|_X \right)^p \overset{\text{75}}{\lesssim} \sum_{\alpha = 1}^n 2^{\alpha} \sum_{\beta = 0}^\alpha \frac{(\alpha + 1)^{p-1} ((\alpha - \beta)!|\beta|!)p}{2^\alpha R^\alpha} \|V^\varepsilon_{\alpha,\beta} f(x)\|_X^p.$$
We can therefore bound the left-hand side of (71) as follows.

\[
\sum_{x \in \mathbb{Z}_{4m}^n} \left\| \left( \frac{R}{R+1} \right)^{n-1} (Af(x + \varepsilon) - Af(x - \varepsilon)) - \sum_{j=1}^{n} \varepsilon_j [B_j f(x + e_j) - B_j f(x - e_j)] \right\|_X^p 
\lesssim_p \sum_{\alpha=1}^{n} \sum_{\beta=0}^{\alpha} \frac{(\alpha + 1)^{p-1} ((\alpha - \beta)! \beta!)^p}{2^\alpha R^{\alpha p}} \sum_{x \in \mathbb{Z}_{4m}^n} \| V_{\alpha,\beta} f(x) \|_X^p. \tag{76}
\]

Since the number of terms in the sums that appear in the definition (72) of \( V_{\alpha,\beta} \) is \( \binom{n}{\alpha} \binom{n}{\beta} \),

\[
\sum_{x \in \mathbb{Z}_{4m}^n} \| V_{\alpha,\beta} f(x) \|_X^p
\leq \left( \frac{n}{\alpha} \right)^{p-1} \left( \frac{\alpha}{\beta} \right)^{p-1} \sum_{T \subseteq \{1, \ldots, n\}} \sum_{|T|=n-\alpha} \sum_{x \in \mathbb{Z}_{4m}^n} \| \Delta_T f (x + R\delta + \varepsilon_T) - \Delta_T f (x + R\delta - \varepsilon_T) \|_X^p
\leq \left( \frac{n}{\alpha} \right)^{p-1} \left( \frac{\alpha}{\beta} \right)^p \sum_{T \subseteq \{1, \ldots, n\}} \sum_{|T|=n-\alpha} \| f (x + 2\varepsilon_T) - f (x) \|_X^p,
\tag{77}
\]

where (77) is valid since \( \Delta_T \) is an averaging operator.

By combining (76) with (77) we see that

\[
\sum_{x \in \mathbb{Z}_{4m}^n} \left\| \left( \frac{R}{R+1} \right)^{n-1} (Af(x + \varepsilon) - Af(x - \varepsilon)) - \sum_{j=1}^{n} \varepsilon_j [B_j f(x + e_j) - B_j f(x - e_j)] \right\|_X^p
\lesssim_p \sum_{\alpha=1}^{n} \sum_{\beta=0}^{\alpha} \frac{(\alpha + 1)^{p-1} ((\alpha - \beta)! \beta!)^p}{2^\alpha R^{\alpha p}} \left( \frac{n}{\alpha} \right)^{p-1} \left( \frac{\alpha}{\beta} \right)^p \sum_{x \in \mathbb{Z}_{4m}^n} \| f (x + 2\varepsilon_T) - f (x) \|_X^p
\leq \sum_{\alpha=1}^{n} \frac{(\alpha + 1)^{p-1} n!}{(n - \alpha)!} \left( \frac{n}{\alpha} \right)^p \sum_{T \subseteq \{1, \ldots, n\}} \sum_{|T|=n-\alpha} \| f (x + 2\varepsilon_T) - f (x) \|_X^p. \tag{78}
\]

The desired estimate (71) is a consequence of (78) via the change of variable \( s = n - \alpha \) and by using the bounds \( n!/(n - \alpha)! \leq n^\alpha \) and \( (\alpha + 1)^p/2^\alpha \leq (2p)^p \). \( \square \)

In what follows, we will use the following simple lemma several times.

**Lemma 4.4.** Suppose that \((X, d_X)\) is a metric space. Fix \( S \subseteq \{1, \ldots, n\} \) and \( p \in [1, \infty) \). Then for every \( f : \mathbb{Z}_{4m}^n \to X \) we have

\[
\sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_{4m}^n} d_X (f(x), f(x + 2\varepsilon S))^p \leq 2^p \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_{4m}^n} d_X (f(x), f(x + \varepsilon))^p. \tag{79}
\]
Proof. For every $\varepsilon, \delta \in \{-1, 1\}^n$ we have

$$d_X(f(x), f(x + 2\varepsilon S))^p \leq 2^{p-1}d_X((f(x), f(x + \varepsilon S + \delta_{\{1, \ldots, n\}\setminus S}))^p + 2^{p-1}d_X((f(x + \varepsilon S + \delta_{\{1, \ldots, n\}\setminus S}), f(x + 2\varepsilon S))^p.$$ 

Hence,

$$\sum_{x \in \mathbb{Z}_4^n} d_X(f(x), f(x + 2\varepsilon S))^p \leq 2^{p-1} \sum_{x \in \mathbb{Z}_4^n} (d_X(f(x), f(x + \varepsilon S + \delta_{\{1, \ldots, n\}\setminus S}))^p + d_X((f(x + \varepsilon S + \delta_{\{1, \ldots, n\}\setminus S}), f(x + 2\varepsilon S))^p) = 2^{p-1} \sum_{x \in \mathbb{Z}_4^n} (d_X(f(x), f(x + \varepsilon S + \delta_{\{1, \ldots, n\}\setminus S}))^p + d_X((f(x), f(x + \varepsilon S - \delta_{\{1, \ldots, n\}\setminus S}))^p).$$

(80)

By averaging [80] over $\delta \in \{-1, 1\}^n$ while using the fact that $\delta_{\{1, \ldots, n\}\setminus S}$ and $-\delta_{\{1, \ldots, n\}\setminus S}$ are identically distributed, we deduce that

$$\sum_{x \in \mathbb{Z}_4^n} d_X(f(x), f(x + 2\varepsilon S))^p \leq 2^p \sum_{\delta \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_4^n} d_X(f(x), f(x + \varepsilon S + \delta_{\{1, \ldots, n\}\setminus S}))^p. \quad (81)$$

If $\varepsilon$ and $\delta$ are i.i.d. and uniformly distributed over $\{-1, 1\}^n$ then the vector $\varepsilon S + \delta_{\{1, \ldots, n\}\setminus S}$ is also uniformly distributed over $\{-1, 1\}^n$. Consequently, the desired estimate (79) follows by averaging [81] over $\varepsilon \in \{-1, 1\}^n$.

The following two lemmas contain estimates that will be used crucially in the ensuing discussion.

**Lemma 4.5.** Let $(X, \| \cdot \|_X)$ be a Banach space. Suppose that $R \geq 2n - 1$ (in addition to the previous assumptions on $R$, i.e., that it is an odd integer with $R \leq 2m$). Then for every $p \in [1, \infty)$ and $f : \mathbb{Z}_4^n \rightarrow X$ we have

$$\sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_4^n} \left\| \frac{R}{R+1} (Af(x + \varepsilon) - Af(x - \varepsilon)) - \sum_{j=1}^n \varepsilon_j [B_j f(x + e_j) - B_j f(x - e_j)] \right\|_X^p \leq p \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_4^n} \| f(x) - f(x + \varepsilon) \|^p_X. \quad (82)$$

Proof. By summing (71) over $\varepsilon \in \{-1, 1\}^n$ and using Lemma 4.4 we see that the left-hand side of (82) is at most $(O(1)p)^p$ times the following quantity

$$\left( \sum_{s=0}^{n-1} \left( \frac{n}{R} \right)^{(n-s)p} \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_4^n} \| f(x) - f(x + \varepsilon) \|^p_X \right)^{\frac{1}{p}} \leq \left( \frac{n}{R} \right)^p \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_4^n} \| f(x) - f(x + \varepsilon) \|^p_X,$$

where in the last step we used the fact that $R \geq 2n - 1$. 

The following lemma contains an estimate that will be used to control the average over all $\delta \in \{-1, 1\}^n$ of the first term in the right-hand side of (66).
Lemma 4.6. Let \((X, \| \cdot \|_X)\) be a Banach space and fix \(S \subseteq \{1, \ldots, n\}\). Suppose that \(R\) is an odd integer satisfying \(2|S| - 1 \leq R \leq 2m\). Then for every \(p \in [1, \infty)\) and \(f : \mathbb{Z}_4^m \to X\) we have

\[
\sum_{\delta \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_4^m} \|D_S f(x + 2\delta) - D_S f(x)\|_X^p \\
\lesssim_p \sum_{\epsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_4^m} \left\| \sum_{j \in S} \epsilon_j [B_j f(x + e_j) - B_j f(x - e_j)] \right\|_X^p \\
+ \left( \frac{p|S|}{R} \right)^p \sum_{\epsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_4^m} \|f(x) - f(x + \epsilon)\|_X^p.
\]  

(83)

Proof. Denote \(k \overset{\text{def}}{=} |S|\), \(T \overset{\text{def}}{=} \{1, \ldots, n\} \setminus S\) and consider \(Z_4^m\) as being equal to \(\mathbb{Z}_4^m \times \mathbb{Z}_4^m\). For every \(y \in \mathbb{Z}_4^m\) define \(f_y : \mathbb{Z}_4^m \to X\) by setting for every \(x \in \mathbb{Z}_4^m\),

\[
f_y(x) \overset{\text{def}}{=} \frac{1}{(R+1)^{n-k}} \sum_{z \in (1+2Z)^T \cap [-R,R]^T} f(x, y + z).
\]

Let \(A^{(S)}\) be the averaging operator corresponding to (70) with \(Z_4^m\) replaced by \(\mathbb{Z}_4^m\), i.e., for every \(h : \mathbb{Z}_4^m \to X\) and \(x \in \mathbb{Z}_4^m\),

\[
A^{(S)} h(x) \overset{\text{def}}{=} \frac{1}{R^k} \sum_{w \in (-R,R)^k \cap (2Z)^k} h(x + w).
\]

Similarly, for every \(j \in S\) let \(B_j^{(S)}\) be the averaging operator analogous to \(B_j\) but with \(\mathbb{Z}_4^m\) replaced by \(\mathbb{Z}_4^m\), i.e., for every \(h : \mathbb{Z}_4^m \to X\) and \(x \in \mathbb{Z}_4^m\),

\[
B_j^{(S)} h(x) \overset{\text{def}}{=} \frac{1}{R(R+1)^{k-1}} \sum_{\substack{a \in [-R,R] \cap (2Z) \cap [-R,R] \cap (1+2Z)^{S \setminus \{j\}}}} h \left( x + ae_j + \sum_{s \in S \setminus \{j\}} b_se_s \right)
\]

With these definitions, for every \((x, y) \in \mathbb{Z}_4^m \times \mathbb{Z}_4^m\) and \(j \in S\) we have

\[
D_S f(x, y) = A^{(S)} f_y(x) \quad \text{and} \quad B_j f(x, y) = B_j^{(S)} f_y(x).
\]  

(84)

Since \(R \geq 2k - 1\), an application of (71) to \(f_y\) yields the following estimate, which holds true for every fixed \(\delta \in \{-1, 1\}^n\) and \(y \in \mathbb{Z}_4^m\),

\[
\sum_{x \in \mathbb{Z}_4^m} \|A^{(S)} f_y(x + 2\delta) - A^{(S)} f_y(x)\|_X^p = \sum_{x \in \mathbb{Z}_4^m} \|A^{(S)} f_y(x + \delta) - A^{(S)} f_y(x - \delta)\|_X^p
\]

\[
\lesssim_p \sum_{x \in \mathbb{Z}_4^m} \sum_{j \in S} \|B_j^{(S)} f_y(x + e_j) - B_j^{(S)} f_y(x - e_j)\|_X^p
\]

\[
+ p^k \sum_{s=0}^{k-1} \frac{(k/R)^{(k-s)p}}{a^k} \sum_{W \subseteq S} \sum_{x \in \mathbb{Z}_4^m} \|f_y(x + 2\delta W) - f_y(x)\|_X^p.
\]

(85)
By summing (85) over \( \delta \in \{-1, 1\}^n \) and \( y \in \mathbb{Z}_{4m}^T \), while using the identities (84), we see that

\[
\sum_{\delta \in \{-1, 1\}^n} \sum_{z \in \mathbb{Z}_{4m}^n} \| D_S f(z + 2\delta) - D_S f(z) \|_X^p
\]

\[
\lesssim_p \sum_{\delta \in \{-1, 1\}^n} \sum_{z \in \mathbb{Z}_{4m}^n} \left\| \sum_{j \in S} \delta_j [B_j f(z + \epsilon) - B_j f(x - \epsilon)] \right\|_X^p
\]

\[
+ p^p \sum_{s=0}^{k-1} \frac{(k/R)^{(k-s)p}}{(k/s)} \sum_{W \subseteq S} \sum_{\delta \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_{4m}^n} \sum_{y \in \mathbb{Z}_{4m}^n} \| f_y(x + 2\delta W) - f_y(x) \|_X^p.
\]  

(86)

Recalling that \( f_y \) is obtained from \( f \) by averaging, it follows by convexity that for every \( W \subseteq S \) and \( \delta \in \{-1, 1\}^n \) we have

\[
\sum_{x \in \mathbb{Z}_{4m}^n} \sum_{y \in \mathbb{Z}_{4m}^n} \| f_y(x + 2\delta W) - f_y(x) \|_X^p \leq \sum_{z \in \mathbb{Z}_{4m}^n} \| f(z + 2\delta W) - f(z) \|_X^p.
\]

Consequently, using Lemma 4.4 and the assumption \( R \geq 2k - 1 \), the final term in (86) is at most \( (O(1)p)^p \) times the following quantity

\[
\left( \sum_{s=0}^{k-1} \left( \frac{k}{R} \right)^{(k-s)p} \right) \sum_{\epsilon \in \{-1, 1\}^n} \sum_{z \in \mathbb{Z}_{4m}^n} \| f(z + 2\epsilon) - f(z) \|_X^p
\]

\[
\lesssim \left( \frac{k}{R} \right)^p \sum_{\epsilon \in \{-1, 1\}^n} \sum_{z \in \mathbb{Z}_{4m}^n} \| f(z + \epsilon) - f(z) \|_X^p.
\]

Hence (86) implies the desired inequality (83). \( \square \)

**Proof of Theorem 1.1** From now on choose \( R \) to be the smallest odd integer that is greater than \( pn \), and suppose that

\[
m \geq \frac{n^{3/2} \log p}{\sqrt{k}} + pn.
\]

In particular we have \( 2n \leq R \leq 2m \). Fix \( x \in \mathbb{Z}_{4m}^n \) and apply inequality (15) to the scalars \( a_j = B_j f(x + \epsilon_j) - B_j f(x - \epsilon_j) \). The resulting estimate is

\[
\frac{(p/ \log p)^p}{2^n (n/k)} \sum_{S \subseteq \{1,...,n\}} \sum_{|S| = k} \left\| \sum_{j \in S} \epsilon_j [B_j f(x + \epsilon_j) - B_j f(x - \epsilon_j)] \right\|_X^p
\]

\[
\lesssim_p \frac{k}{n} \sum_{j=1}^{n} \| B_j f(x + \epsilon_j) - B_j f(x - \epsilon_j) \|_X^p
\]

\[
+ \frac{(k/n)^p}{2^n} \sum_{\epsilon \in \{-1, 1\}^n} \left\| \sum_{j=1}^{n} \epsilon_j [B_j f(x + \epsilon_j) - B_j f(x - \epsilon_j)] \right\|_X^p.
\]  

(87)

28
By summing (87) over \( x \in \mathbb{Z}_4^n \) we deduce that
\[
\frac{(p/ \log p)^{-p}}{2^n \binom{n}{k}^2} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{|S|=k} \sum_{x \in \mathbb{Z}_4^n} \sum_{j \in S} \varepsilon_j \left[ B_j f(x + e_j) - B_j f(x - e_j) \right]^p \leq \sum_{x \in \mathbb{Z}_4^n} |B_j f(x + 2e_j) - B_j f(x)|^p \leq \sum_{x \in \mathbb{Z}_4^n} |f(x + 2e_j) - f(x)|^p \leq \sum_{x \in \mathbb{Z}_4^n} |f(x + e_j) - f(x)|^p \leq \sum_{x \in \mathbb{Z}_4^n} |f(x + e) - f(x)|^p.,
\]
(88)

For every \( j \in \{1, \ldots, n\} \), since \( B_j \) is an averaging operator we have
\[
\sum_{x \in \mathbb{Z}_4^n} |B_j f(x + 2e_j) - B_j f(x)|^p \leq \sum_{x \in \mathbb{Z}_4^n} |f(x + 2e_j) - f(x)|^p \leq \sum_{x \in \mathbb{Z}_4^n} |f(x + e_j) - f(x)|^p. \quad (89)
\]

Recalling that \( R \geq pn \), by Lemma 4.5 we have
\[
\sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_4^n} \left| \sum_{j=1}^n \varepsilon_j [B_j f(x + e_j) - B_j f(x - e_j)] \right|^p \leq \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_4^n} |Af(x + 2\varepsilon) - Af(x)|^p + \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_4^n} |f(x + \varepsilon) - f(x)|^p \leq \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_4^n} |f(x + \varepsilon) - f(x)|^p \quad (90)
\]
\[
\sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_4^n} |f(x + \varepsilon) - f(x)|^p, \quad (91)
\]
where in (90) we used the fact that \( A \) is an averaging operator.

By substituting (89) and (91) into (88) we see that
\[
\frac{(p/ \log p)^{-p}}{2^n \binom{n}{k}^2} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{|S|=k} \sum_{x \in \mathbb{Z}_4^n} \sum_{j \in S} \varepsilon_j \left[ B_j f(x + e_j) - B_j f(x - e_j) \right]^p \leq \frac{k}{n} \sum_{j=1}^n \sum_{x \in \mathbb{Z}_4^n} |f(x + e_j) - f(x)|^p + \left( \frac{k}{n} \right)^2 \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_4^n} |f(x + \varepsilon) - f(x)|^p. \quad (92)
\]

By averaging (83) over all \( S \subseteq \{1, \ldots, n\} \) with \( |S| = k \) and substituting (92) into the resulting inequality, we obtain the following estimate.
\[
\frac{(p/ \log p)^{-p}}{2^n \binom{n}{k}^2} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{\delta \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_4^n} |D_S f(x + \delta) - D_S f(x - \delta)|^p \leq \frac{k}{n} \sum_{j=1}^n \sum_{x \in \mathbb{Z}_4^n} |f(x + e_j) - f(x)|^p + \left( \frac{k}{n} \right)^2 \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_4^n} |f(x + \varepsilon) - f(x)|^p. \quad (93)
\]

29
Next, average (66) over $\delta \in \{-1,1\}^n$ and substitute (93) into the resulting inequality, thus obtaining the following estimate (recall that in the present setting $R \lesssim p n$).

$$\frac{1}{2^n \binom{n}{k}} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_{4m}^n} \frac{|f(x + 2m \varepsilon y) - f(x)|^p}{m^p} \lesssim_p \left( \frac{p^p}{(\log p)^p} + \frac{k^{p-1}}{m^p} \right) \frac{k}{n} \sum_{j=1}^{n} \sum_{x \in \mathbb{Z}_{4m}^n} |f(x + ej) - f(x)|^p + \left( \frac{p^p}{(\log p)^p} + \frac{(pm)^p}{m^p} \left( \frac{n}{k} \right)^p \right) \frac{k}{n} \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_{4m}^n} |f(x + \varepsilon x) - f(x)|^p. \quad (94)$$

Since $m \geq \frac{n^{3/2} \log p}{\sqrt{k}}$, the desired inequality (3) is a consequence of (94). \(\square\)

5. PROOF OF THEOREM 1.5

The desired inequality (22) is equivalent to the conjunction of the following two inequalities.

$$\sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_{8m}^n} |f(x + 2\varepsilon) - f(x)|^p \lesssim_p \frac{(pm/k)^{p/2}}{(n)} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_{8m}^n} |f(x + \varepsilon S) - f(x)|^p, \quad (95)$$

and

$$\sum_{j=1}^{n} \sum_{x \in \mathbb{Z}_{8m}^n} |f(x + 4m ej) - f(x)|^p \lesssim_p \frac{p^{p/2}}{2^n \binom{n}{k-1}} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_{8m}^n} |f(x + \varepsilon S) - f(x)|^p. \quad (96)$$

The proofs of (95) and (96) are of a different nature: (95) is related to metric type and (96) is related to metric cotype. We therefore treat (95) and (96) in separate subsections.

5.1. Metric type and proof of (95).

For every $n \in \mathbb{N}$ and $j \in \{1, \ldots, n\}$ let $\sigma^j \in \{-1,1\}^n$ be given by

$$\sigma^j \overset{\text{def}}{=} -e_j + \sum_{s \in \{1, \ldots, n\} \setminus \{j\}} e_s.$$

Thus, for every $\varepsilon \in \{-1,1\}^n$, coordinate-wise multiplication by $\sigma^j$ yields

$$\sigma^j \varepsilon = (\varepsilon_1, \ldots, -\varepsilon_j, -\varepsilon_j, \varepsilon_{j+1}, \ldots, \varepsilon_n).$$

Suppose that $(X, \| \cdot \|_X)$ be a Banach space and that $p \in [1, \infty)$. Slightly abusing notation that was introduced in [24], let $\mathfrak{P}^p_n(X)$ be the infimum over those $\mathfrak{P} \in (0, \infty)$ such that for every $h : \{-1,1\}^n \to X$ we have

$$\frac{1}{2^n} \sum_{\varepsilon \in \{-1,1\}^n} \|h(\varepsilon) - h(-\varepsilon)\|_X^p \leq \frac{\mathfrak{P}^p_n}{4^n} \sum_{\varepsilon, \delta \in \{-1,1\}^n} \left\| \sum_{j=1}^{n} \delta_j [h(\sigma^j \varepsilon) - h(\varepsilon)] \right\|^p_X. \quad (97)$$

Note that in [24] the quantity $\mathfrak{P}^p_n(X)$ denotes the best constant in an inequality that is stronger than but closely related to (97). However, this distinction is not important for us here and we prefer to use the notation $\mathfrak{P}^p_n(X)$ rather than introducing ad hoc terminology.

The quantity $\mathfrak{P}^p_n(X)$ is called the Pisier constant of $(X, \| \cdot \|_X)$ (corresponding to dimension $n$ and exponent $p$). In the context of metric type, Pisier proved in [50] that $\mathfrak{P}^p_n(X) \lesssim \log n$ for every Banach space $(X, \| \cdot \|_X)$. In order to prove (95) we will deal with $X = \mathbb{R}$, in which case
it will be important that \( \sup_{n \in \mathbb{N}} \Psi^n_p(\mathbb{R}) < \infty \). This strengthening of Pisier's inequality for real-valued functions is due to Talagrand [57], who proved that \( \sup_{n \in \mathbb{N}} \Psi^n_p(\mathbb{R}) \leq K^p \) for some universal constant \( K \in (1, \infty) \), an estimate that was later improved in [43] to \( \sup_{n \in \mathbb{N}} \Psi^n_p(\mathbb{R}) \lesssim p \). The rate of growth of \( \sup_{n \in \mathbb{N}} \Psi^n_p(\mathbb{R}) \) as \( p \to \infty \) remains unknown, the best available lower bound, due to Talagrand [57], being that \( \sup_{n \in \mathbb{N}} \Psi^n_p(\mathbb{R}) \) is at least a constant multiple of \( \log p \). We refer to [43, 24] for additional classes of Banach space \((X, \|\cdot\|)_n \) for which \( \sup_{n \in \mathbb{N}} \Psi^n_p(X) < \infty \).

Given a metric space \((X, \|\cdot\|_X)\), for every \( n \in \mathbb{N} \) and \( q \in [1, \infty) \) define \( \text{BMW}^n_q(X; p) \) to be the infimum over those \( B \in [1, \infty) \) such that for every \( h : \{-1, 1\}^n \to X \) we have

\[
\sum_{\varepsilon \in \{-1, 1\}^n} d_X(h(\varepsilon), h(-\varepsilon))^p \leq B^n n^{p-1} \sum_{j=1}^n \sum_{\varepsilon \in \{-1, 1\}^n} d_X(h(\sigma^j \varepsilon), h(\varepsilon))^p . \tag{98}
\]

The quantity \( \text{BMW}^n_q(X; p) \) is called the Bourgain-Milman-Wolfson type \( q \) constant of \((X, \|\cdot\|_X)\) (corresponding to dimension \( n \) and exponent \( p \)). It was introduced and studied by Bourgain, Milman and Wolfson in [10], though, as we explained in the Introduction, the case \( p = q \) was previously introduced by Enflo [17] and Gromov [23] (Gromov only dealt with the case \( p = q = 2 \)). It follows from [97] and Hölder’s inequality that if \((X, \|\cdot\|_X)\) is a Banach space then

\[
\text{BMW}^n_q(X; p) \leq \Psi^n_p(X) \cdot T^n_q(X; p) \lesssim (\log n) \cdot T^n_q(X; p) , \tag{99}
\]

where the (Rademacher type \( q \)) constant \( T^n_q(X; p) \) is defined to be the infimum over those \( T \in (0, \infty) \) such that for every \( x_1, \ldots, x_n \in X \) we have

\[
\left( \frac{1}{2^n} \sum_{\delta \in \{-1, 1\}^n} \left\| \sum_{i=1}^n \delta_i x_i \right\|_X^p \right)^{\frac{1}{p}} \leq T \left( \sum_{j=1}^n \|x_j\|_X^q \right)^{\frac{1}{q}} .
\]

Since for many Banach spaces \((X, \|\cdot\|_X)\) good estimates on \( T^n_q(X; p) \) are known, in conjunction with the available bounds on \( \Psi^n_p(X) \), inequality (99) often yields a satisfactory estimate on \( \text{BMW}^n_q(X; p) \). Such an estimate will be relevant to the ensuing proof of a metric-space-valued extension of [45]. There are also several important classes of (non-Banach) metric spaces \((X, d_X)\) for which good bounds on \( \text{BMW}^n_q(X; p) \) have been obtained; see for example [43, 42, 45, 44, 41]. When \( X = \mathbb{R} \), a bound that is even better than what follows from (99) is known: see inequality (6.32) in [41], which yields the estimate

\[
\sup_{n \in \mathbb{N}} \text{BMW}^n_2(\mathbb{R}; p) \lesssim \sqrt{p} . \tag{100}
\]

The following lemma, in conjunction with (100), implies (95). Note that there is no requirement that \( m \) is sufficiently large here: the lower bound on \( m \) that is assumed in Theorem 1.5 will be needed only for the proof of (96).

**Lemma 5.1.** Suppose that \((X, d_X)\) is a metric space and \( p, q \in [1, \infty) \). Then for every \( n \in \mathbb{N} \), \( k \in \{1, \ldots, n\} \) and \( f : \mathbb{Z}_{8m} \to X \) we have

\[
\sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_{8m}} d_X(f(x + 2\varepsilon), f(x))^p \lesssim_p \left( \text{BMW}^{|n/k|+1}_q(X; p) \right)^p \frac{(n/k)^2}{(n)} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_{8m}} d_X(f(x + \varepsilon S), f(x))^p . \tag{101}
\]
Proof. Write \( n = ak + b \) where \( a = \lfloor n/k \rfloor \) and \( b \in \{0, \ldots, k - 1\} \). For every \( j \in \{1, \ldots, a\} \) define \( I_j = \{(j-1)k + 1, \ldots, jk\} \), and also define \( I_{a+1} = \{ak + 1, \ldots, ak + b\} \). Fix \( x \in \mathbb{Z}_8^n \) and \( \varepsilon \in \{-1, 1\}^n \). For every permutation \( \pi \in S_n \) define \( h_{x,\varepsilon}^\pi : \{-1, 1\}^{a+1} \to X \) by

\[
\forall \delta \in \{-1, 1\}^{a+1}, \quad h_{x,\varepsilon}^\pi(\delta) \overset{\text{def}}{=} f \left( x + \sum_{j=1}^{a+1} \delta_j \varepsilon_{\pi(I_j)} \right).
\]

Note that for every \( \pi \in S_n \), every \( x \in \mathbb{Z}_8^n \), and every \( \delta \in \{-1, 1\}^{a+1} \) we have

\[
\sum_{\varepsilon \in \{-1, 1\}^n} d_X (h_{x,\varepsilon}^\pi(\delta), h_{x,\varepsilon}^\pi(-\delta))^p = \sum_{\varepsilon \in \{-1, 1\}^n} d_X (f(x + \varepsilon), f(x - \varepsilon))^p = \sum_{\varepsilon \in \{-1, 1\}^n} d_X (f(x + 2\varepsilon), f(x))^p. \tag{102}
\]

Also, for every \( \pi \in S_n \) and \( j \in \{1, \ldots, n\} \) we have

\[
\frac{1}{2^{a+1}} \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{\delta \in \{-1, 1\}^{a+1}} \sum_{x \in \mathbb{Z}_8^n} d_X (h_{x,\varepsilon}^\pi(\delta), h_{x,\varepsilon}^\pi(-\delta))^p = \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_8^n} d_X (f(x + \varepsilon_{\pi(I_j)}), f(x - \varepsilon_{\pi(I_j)}))^p = \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_8^n} d_X (f(x + 2\varepsilon_{\pi(I_j)}), f(x))^p. \tag{103}
\]

Fix \( B > \text{BMW}^q_{a+1}(X;p) \), apply \( \square \) to \( h_{x,\varepsilon}^\pi \), and sum the resulting inequality over \( \varepsilon \in \{-1, 1\}^n \) and \( x \in \mathbb{Z}_8^n \), while using the identities \( \square \) and \( \square \). The resulting inequality is

\[
\sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_8^n} d_X (f(x + 2\varepsilon), f(x))^p
\leq B^p (a + 1)^\frac{\varepsilon}{q} - 1 \sum_{j=1}^{a+1} \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_8^n} d_X (f(x + 2\varepsilon_{\pi(I_j)}), f(x))^p. \tag{104}
\]

By averaging \( \square \) over \( \pi \in S_n \) we see that

\[
B^{-p} (a + 1)^{-\frac{\varepsilon}{q}} \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_8^n} d_X (f(x + 2\varepsilon), f(x))^p
\leq \sum_{j=1}^{a} \sum_{S \subseteq \{1, \ldots, n\} \atop |S| = k} \frac{|\{ \pi \in S_n : \pi(I_j) = S \}|}{n!} \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_8^n} d_X (f(x + 2\varepsilon_S), f(x))^p
\]

\[
+ \sum_{T \subseteq \{1, \ldots, n\} \atop |T| = b} \frac{|\{ \pi \in S_n : \pi(I_{a+1}) = T \}|}{n!} \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_8^n} d_X (f(x + 2\varepsilon_T), f(x))^p
\]

\[
= \binom{a}{k} \sum_{S \subseteq \{1, \ldots, n\} \atop |S| = k} \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_8^n} d_X (f(x + 2\varepsilon_S), f(x))^p
\]

\[
+ \binom{a}{b} \sum_{T \subseteq \{1, \ldots, n\} \atop |T| = b} \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_8^n} d_X (f(x + 2\varepsilon_T), f(x))^p. \tag{105}
\]
By Lemma 4.4 if $T \subseteq S \subseteq \{1, \ldots, n\}$ then
\[
\sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_m^n} d_X(f(x + 2\varepsilon T), f(x))^p \leq 2^p \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_m^n} d_X(f(x + \varepsilon S), f(x))^p. \tag{106}
\]

Fixing $T \subseteq \{1, \ldots, n\}$ with $|T| = b$, by averaging (106) over all $k$-point subsets $S \subseteq \{1, \ldots, n\}$ with $S \supseteq T$ we see that
\[
\sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_m^n} d_X(f(x + 2\varepsilon T), f(x))^p \leq \frac{2^p}{\binom{n}{k}} \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_m^n} d_X(f(x + \varepsilon S), f(x))^p.
\]

Consequently,
\[
\frac{1}{\binom{n}{k}} \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_m^n} d_X(f(x + 2\varepsilon T), f(x))^p \leq \frac{2^p}{\binom{n}{k}} \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_m^n} d_X(f(x + \varepsilon S), f(x))^p
\]
\[
= \frac{2^p}{\binom{n}{k}} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_m^n} d_X(f(x + \varepsilon S), f(x))^p
\]
\[
= \frac{2^p}{\binom{n}{k}} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_m^n} d_X(f(x + \varepsilon S), f(x))^p. \tag{107}
\]

Since by the triangle inequality we also have
\[
\sum_{S \subseteq \{1, \ldots, n\}} \sum_{x \in \mathbb{Z}_m^n} d_X(f(x + 2\varepsilon S), f(x))^p \leq 2^p \sum_{S \subseteq \{1, \ldots, n\}} \sum_{x \in \mathbb{Z}_m^n} d_X(f(x + \varepsilon S), f(x))^p,
\]
it follows from (105) and (107) that
\[
\sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_m^n} d_X(f(x + 2\varepsilon S), f(x))^p
\]
\[
\leq \frac{(2B)^p(a + 1)^{\frac{p}{q}}}{\binom{n}{k}} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_m^n} d_X(f(x + \varepsilon S), f(x))^p
\]
\[
\leq \frac{(2B)^p(2n/k)^{\frac{p}{q}}}{\binom{n}{k}} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_m^n} d_X(f(x + \varepsilon S), f(x))^p. \tag{108}
\]

5.2. **Metric cotype and proof of (96)**. Given a metric space $(X, d_X)$ and $m, n \in \mathbb{N}$, for $p \in (1, \infty)$ define $\Gamma_p(X; m, n)$ to be the infimum over those $\Gamma \in (0, \infty)$ such that for every $f : \mathbb{Z}_m^n \to X$,
\[
\sum_{j=1}^{n} \sum_{x \in \mathbb{Z}_m^n} d_X(f(x + me_j), f(x))^p \leq \frac{\Gamma^p}{3^n} \sum_{\varepsilon \in \{-1,0,1\}^n} \sum_{x \in \mathbb{Z}_m^n} d_X(f(x + \varepsilon), f(x))^p. \tag{108}
\]
As discussed in the Introduction, following [39], we say that \((X, d_X)\) has metric cotype \(q\) if

\[
\Gamma_p(X) \overset{\text{def}}{=} \sup_{n \in \mathbb{N}} \inf_{m \in \mathbb{N}} \Gamma_p(X; m, n) < \infty.
\]

We need to briefly recall some facts related to \(K\)-convexity of Banach spaces; see the survey [35] for much more on this topic. Given a Banach space \((X, \| \cdot \|_X), p \in (1, \infty)\) and \(n \in \mathbb{N}\), for every \(f : \{-1, 1\}^n \to X\) define its Rademacher projection \(\text{Rad}(f) : \{-1, 1\}^n \to X\) by

\[
\forall \varepsilon \in \{-1, 1\}^n, \quad \text{Rad}(f)(\varepsilon) \overset{\text{def}}{=} \sum_{j=1}^n \frac{n}{2^n} f(\delta_j) \varepsilon_j.
\]

For \(p \in (1, \infty)\) let \(K_p(X) \in [1, \infty]\) be the infimum over those \(K \in [1, \infty]\) such that for every \(n \in \mathbb{N}\) and every \(f : \{-1, 1\}^n \to X\) we have

\[
\sum_{\varepsilon \in \{-1, 1\}^n} \|\text{Rad}(f)(\varepsilon)\|_X^p \leq K_p \sum_{\varepsilon \in \{-1, 1\}^n} \|f(\varepsilon)\|_X^p.
\]

A simple application of Khinchine’s inequality (with asymptotically sharp constant, see [48, Lem. 2]) shows that \(K_p(\mathbb{R}) \lesssim \sqrt{p}\) for \(p \in [2, \infty)\). A Banach space \((X, \| \cdot \|_X)\) is said to be \(K\)-convex if \(K_p(X) < \infty\) for some (equivalently for all) \(p \in (1, \infty)\); see [35] and the references therein.

Theorem 5.2 below establishes a sharp metric cotype inequality for \(K\)-convex Banach spaces, with one difference: the averaging on the right-hand side is over \(\varepsilon \in \{-1, 1\}^n\) rather than \(\varepsilon \in \{-1, 0, 1\}^n\). The same result with averages over \(\varepsilon \in \{-1, 0, 1\}^n\) (and \(x \in \mathbb{Z}_4^n\), rather than \(x \in \mathbb{Z}_8^n\)) is the content of Theorem 4.1 in [39]. The proof here follows the argument in [39] with some technical modifications. It seems likely that a similar statement could be proved for the metric cotype \(p\) inequalities for Banach spaces of Rademacher cotype \(p\) (with no assumption of \(K\)-convexity) in [39, 20], though this may require changes to the arguments of [39, 20] that are more substantial than what we do here.

**Theorem 5.2.** Fix \(p \in [2, \infty)\) and \(\alpha \in [1, \infty)\). Let \((X, \| \cdot \|_X)\) be a \(K\)-convex Banach space of cotype \(p\). Suppose that \(m, n \in \mathbb{N}\) satisfy

\[
m \geq \frac{n^{1/p}}{\alpha K_p(X) C_p(X)}.
\]

Then for every \(f : \mathbb{Z}_8^n \to X\) we have

\[
\sum_{j=1}^n \sum_{x \in \mathbb{Z}_8^n} \frac{\|f(x + 4me_j) - f(x)\|_X^p}{m^p} \lesssim_p \left(\frac{\alpha K_p(X) C_p(X)}{2^n}\right)^p \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_8^n} \|f(x+\varepsilon) - f(x)\|_X^p.
\]

Before proving Theorem 5.2 we deduce the following simple corollary, which implies [96] because \(C_p(\mathbb{R}) = 1\) and \(K_p(\mathbb{R}) \lesssim \sqrt{p}\).

**Corollary 5.3.** Fix \(p \in [2, \infty)\) and \(\alpha \in [1, \infty)\). Let \((X, \| \cdot \|_X)\) be a \(K\)-convex Banach space of cotype \(p\). Suppose that \(m, n \in \mathbb{N}\) and \(k \in \{1, \ldots, n\}\) satisfy

\[
m \geq \frac{k^{1/p}}{\alpha K_p(X) C_p(X)}.
\]
Proof. By Theorem 5.2, for every $S \subseteq \{1, \ldots, n\}$ with $|S| = k$ we have
\[ \sum_{j \in S} \sum_{x \in \mathbb{Z}^n_{8m}} \frac{\|f(x + 4me_j) - f(x)\|^p_X}{m^p} \lesssim_p \frac{(\alpha K_p(X)C_p(X))^p}{2^n (\binom{n}{k-1})^p} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}^n_{8m}} \|f(x + \varepsilon S) - f(x)\|^p_X. \tag{111} \]

By averaging this inequality over all $S \subseteq \{1, \ldots, n\}$ with $|S| = k$ we obtain \eqref{111}. \hfill \Box

In order to prove Theorem 5.2 we first introduce a small amount of notation and prove an auxiliary lemma. For every $j \in \{1, \ldots, n\}$ define a linear operator $T_j : L_2(\mathbb{Z}^n_{8m}, X) \to L_2(\mathbb{Z}^n_{8m}, X)$ by setting for every $f : \mathbb{Z}^n_{8m} \to X$ and $x \in \mathbb{Z}^n_{8m}$,
\[ T_j f(x) \overset{\text{def}}{=} \frac{1}{2^n} \sum_{\varepsilon \in \{-1, 1\}^n} f(x + 2\varepsilon \{1, \ldots, n\} \setminus \{j\}). \tag{112} \]

Lemma 5.4. Let $(X, \|\cdot\|_X)$ be a Banach space and $p \in [1, \infty)$. Fix also $m, n \in \mathbb{N}$ and $j \in \{1, \ldots, n\}$. Then for every $f : \mathbb{Z}^n_{8m} \to X$ we have
\[ \sum_{x \in \mathbb{Z}^n_{8m}} \|f(x) - T_j f(x)\|^p_X \leq \frac{2^p}{2^m} \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}^n_{8m}} \|f(x + \varepsilon) - f(x)\|^p_X. \tag{113} \]

Moreover, for every $x \in \mathbb{Z}^n_{8m}$ we have
\[ \sum_{\varepsilon \in \{-1, 1\}^n} \left\| \sum_{j=1}^n \varepsilon_j [T_j f(x + 2e_j) - T_j f(x - 2e_j)] \right\|^p_X \leq (2K_p(X))^p \sum_{\varepsilon \in \{-1, 1\}^n} \|f(x + 2\varepsilon) - f(x)\|^p_X \leq (4K_p(X))^p \sum_{\varepsilon \in \{-1, 1\}^n} \|f(x + \varepsilon) - f(x)\|^p_X. \tag{114} \]

Proof. By the definition \eqref{112} and convexity,
\[ \sum_{x \in \mathbb{Z}^n_{8m}} \|f(x) - T_j f(x)\|^p_X \leq \frac{1}{2^n} \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}^n_{8m}} \|f(x) - f(x + 2\varepsilon \{1, \ldots, n\} \setminus \{j\})\|^p_X \leq \frac{2^p}{2^m} \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}^n_{8m}} \|f(x + \varepsilon) - f(x)\|^p_X, \]

where in the last step we used Lemma 4.4. This proves \eqref{113}.

To prove \eqref{114}, for every fixed $x \in \mathbb{Z}^n_{8m}$ define $h^x : \{-1, 1\}^n \to X$ by
\[ \forall \varepsilon \in \{-1, 1\}^n, \quad h^x(\varepsilon) \overset{\text{def}}{=} f(x + 2\varepsilon) - f(x). \]

We claim that the following identity holds true.
\[ \text{Rad}(h^x)(\varepsilon) = \frac{i}{2} \sum_{j=1}^n \varepsilon_j [T_j f(x + 2e_j) - T_j f(x - 2e_j)]. \tag{115} \]

Once \eqref{115} is proved, the desired inequality \eqref{114} would follow from the definition of $K_p(X)$.
By composing with linear functionals, it suffices to verify the validity of $\text{[115]}$ when $X = \mathbb{C}$. Moreover, for every $y \in \mathbb{Z}_{8m}^n$ define $W_y : \mathbb{Z}_{8m}^n \to \mathbb{C}$ by

$$
\forall x \in \mathbb{Z}_{8m}^n, \quad W_y(x) \overset{\text{def}}{=} \exp\left(\frac{\pi i}{4m} \sum_{j=1}^{n} x_j y_j\right).
$$

Then $\{W_y\}_{y \in \mathbb{Z}_{8m}^n}$ forms an orthonormal basis of $L_2(\mathbb{Z}_{8m}^n)$, and therefore it suffices to verify the validity of (115) when $f = W_y$ for some $y \in \mathbb{Z}_{8m}^n$. Now,

$$
W_y^x(\varepsilon) = -\left(1 - \prod_{j=1}^{n} \left(\cos\left(\frac{\pi \varepsilon_j y_j}{2m}\right) + i \sin\left(\frac{\pi \varepsilon_j y_j}{2m}\right)\right)\right) W_y(x)
= -\left(1 - \prod_{j=1}^{n} \left(\cos\left(\frac{\pi y_j}{2m}\right) + i \varepsilon_j \sin\left(\frac{\pi y_j}{2m}\right)\right)\right) W_y(x)
$$

Consequently,

$$
\text{Rad}(W_y^x)(\varepsilon) = i \left(\sum_{j=1}^{n} \varepsilon_j \sin\left(\frac{\pi y_j}{2m}\right) \prod_{s \in \{1, \ldots, n\} \setminus \{j\}} \cos\left(\frac{\pi y_s}{2m}\right)\right) W_y(x).
$$

At the same time, for every $j \in \{1, \ldots, n\}$ we have

$$
T_j W_y(x) = \left(\frac{1}{2^n} \sum_{\varepsilon \in \{-1,1\}^n} \prod_{s \in \{1, \ldots, n\} \setminus \{j\}} \exp\left(\frac{\pi i \varepsilon_s y_s}{2m}\right)\right) W_y(x)
= \left(\prod_{s \in \{1, \ldots, n\} \setminus \{j\}} \cos\left(\frac{\pi y_s}{2m}\right)\right) W_y(x).
$$

Therefore

$$
\sum_{j=1}^{n} \varepsilon_j [T_j f(x + 2e_j) - T_j f(x - 2e_j)]
= \left(\sum_{j=1}^{n} \varepsilon_j (W_y(2e_j) - W_y(-2e_j)) \prod_{s \in \{1, \ldots, n\} \setminus \{j\}} \cos\left(\frac{\pi y_s}{2m}\right)\right) W_y(x)
= 2 \left(\sum_{j=1}^{n} \varepsilon_j \sin\left(\frac{\pi y_j}{2m}\right) \prod_{s \in \{1, \ldots, n\} \setminus \{j\}} \cos\left(\frac{\pi y_s}{2m}\right)\right) W_y(x) = \frac{2}{i} \text{Rad}(W_y^x)(\varepsilon),
$$

where in the last step we used $\text{[116]}$. □

**Proof of Theorem 5.2.** By the triangle inequality, for every $j \in \{1, \ldots, n\}$ we have

$$
\|f(x + 4me_j) - f(x)\|_X \leq_p \|T_j f(x + 4me_j) - T_j f(x)\|_X + \|f(x + 4me_j) - T_j f(x + 4me_j)\|_X + \|f(x) - T_j f(x)\|_X.
$$
Hence, using (113) we see that
\[
\sum_{x \in \mathbb{Z}_8^n} \| f(x + 4me_j) - f(x) \|_X^p \\
\leq p \sum_{x \in \mathbb{Z}_8^n} \| T_j f(x + 4me_j) - T_j f(x) \|_X^p + \frac{1}{2^n} \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_8^n} \| f(x + \varepsilon) - f(x) \|_X^p. \quad (117)
\]
By the triangle inequality combined with Hölder’s inequality we have
\[
\sum_{x \in \mathbb{Z}_8^n} \| T_j f(x + 4me_j) - T_j f(x) \|_X^p \leq m^{p-1} \sum_{s=1}^m \sum_{x \in \mathbb{Z}_8^n} \| T_j f(x + 4se_j) - T_j f(x + 4(s-1)e_j) \|_X^p \\
= m^p \sum_{x \in \mathbb{Z}_8^n} \| T_j f(x + 2e_j) - T_j f(x - 2e_j) \|_X^p.
\]
In combination with (117), this implies that
\[
\sum_{j=1}^n \sum_{x \in \mathbb{Z}_8^n} \frac{\| f(x + 4me_j) - f(x) \|_X^p}{m^p} \\
\leq p \sum_{j=1}^n \sum_{x \in \mathbb{Z}_8^n} \| T_j f(x + 2e_j) - T_j f(x - 2e_j) \|_X^p + \frac{n}{m^p 2^n} \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_8^n} \| f(x + \varepsilon) - f(x) \|_X^p. \quad (118)
\]
By the definition of the cotype p constant $C_p(X)$, for every $x \in \mathbb{Z}_8^n$ we have
\[
\sum_{j=1}^n \| T_j f(x + 2e_j) - T_j f(x - 2e_j) \|_X^p \leq C_p(X)^p \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{j=1}^n \| T_j f(x + 2e_j) - T_j f(x - 2e_j) \|_X^p \\
\leq p (K_p(X)C_p(X))^p \sum_{\varepsilon \in \{-1, 1\}^n} \| f(x + \varepsilon) - f(x) \|_X^p, \quad (119)
\]
where in the last step of (119) we used (114). By substituting (119) into (118) we conclude that
\[
\sum_{j=1}^n \sum_{x \in \mathbb{Z}_8^n} \frac{\| f(x + 4me_j) - f(x) \|_X^p}{m^p} \\
\leq p (K_p(X)C_p(X))^p + \frac{n}{m^p 2^n} \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}_8^n} \| f(x + \varepsilon) - f(x) \|_X^p. \quad (120)
\]
Due to (109), (120) implies the desired inequality (110).

6. A CONJECTURAL CONVOLUTION INEQUALITY AS A WAY TO PROVE CONJECTURE 1.4

For every $j \in \{1, \ldots, n\}$ define an averaging operator $E_j : L_2(\mathbb{Z}_8^n) \rightarrow L_2(\mathbb{Z}_8^n)$ by setting for every $f : \mathbb{Z}_8^n \rightarrow \mathbb{R}$ and $x \in \mathbb{Z}_8^n$,
\[
E_j f(x) \overset{\text{def}}{=} \frac{f(x + e_j) + f(x - e_j)}{2}.
\]
We also set $E_j = \prod_{s \in \{1, \ldots, n\} \setminus \{j\}} E_s$ and $\mathcal{E} = \prod_{s=1}^n E_j$. Thus, for every $f : \mathbb{Z}_8^n \rightarrow \mathbb{R}$ and $x \in \mathbb{Z}_8^n$,
\[
E_j f(x) = \frac{1}{2^n} \sum_{\varepsilon \in \{-1, 1\}^n} f(x + \varepsilon_{\{1, \ldots, n\} \setminus \{j\}}) \quad \text{and} \quad \mathcal{E} f(x) = \frac{1}{2^n} \sum_{\varepsilon \in \{-1, 1\}^n} f(x + \varepsilon). \quad (121)
\]
Question 6.1. Is it true that for every $p \in (2, \infty)$ there exists $\beta_p \in (0, 1]$ such that for every $f : \mathbb{Z}^n_{4m} \to \mathbb{R}$ we have

$$
\frac{\beta_p}{2^n} \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}^n_m} |\mathcal{E}f(x + \varepsilon) - \mathcal{E}f(x - \varepsilon)|^p \leq \frac{1}{2^n} \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}^n_m} |f(x + \varepsilon) - f(x)|^p.
$$

(122)

It may very well be the case that (122) holds true without the second term that appears in the right-hand side, i.e., that

$$
\beta_p \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}^n_m} |\mathcal{E}f(x + \varepsilon) - \mathcal{E}f(x - \varepsilon)|^p \leq \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}^n_m} \varepsilon |\mathcal{E}f(x + \varepsilon) - \mathcal{E}f(x - \varepsilon)|^p.
$$

We formulated Question 6.1 in the above weaker form since it suffices for the following proposition.

Proposition 6.2. A positive answer to Question 6.1 implies that Conjecture 1.4 holds true, and hence also that all the conclusions of Theorem 1.12 hold true. Specifically, for every $\delta \in (0, \infty)$, if $m, n \in \mathbb{N}$ and $k \in \{1, \ldots, n\}$ satisfy $m \geq \delta \sqrt{n/k}$ then (20) holds true with

$$
\beta_p \lesssim_p \left( \frac{p^2}{\log p} + \frac{1}{\delta} \right)^p \alpha_p,
$$

where $\beta_p$ is as in (122).

Proof. Fix $f : \mathbb{Z}^n_{4m} \to \mathbb{R}$. By convexity, it follows from (121) that

$$
\sum_{x \in \mathbb{Z}^n_m} |f(x) - \mathcal{E}f(x)|^p \leq \frac{1}{2^n} \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}^n_m} |f(x + \varepsilon) - f(x)|^p.
$$

Hence, for every $S \subseteq \{1, \ldots, n\}$ we have

$$
\sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}^n_m} |f(x + 2m\varepsilon_S) - f(x)|^p \lesssim_p \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}^n_m} |\mathcal{E}f(x + 2m\varepsilon_S) - \mathcal{E}f(x)|^p + \frac{1}{2^n} \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}^n_m} |f(x + \varepsilon) - f(x)|^p.
$$

(123)

Arguing as in (68), it follows from the triangle inequality that

$$
\sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}^n_m} |\mathcal{E}f(x + 2m\varepsilon_S) - \mathcal{E}f(x)|^p \leq m^p \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}^n_m} |\mathcal{E}f(x + \varepsilon) - \mathcal{E}f(x - \varepsilon_S)|^p.
$$

(124)

For every $z \in \mathbb{Z}^{\{1, \ldots, n\}\setminus S}_{4m}$ apply (122) to the mapping $y \mapsto \prod_{s \in \{1, \ldots, n\}\setminus S} E_j f(y, z)$, and then average the resulting inequality over $z$. The estimate thus obtained is

$$
\frac{\beta_p}{2^n} \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}^n_m} |\mathcal{E}f(x + \varepsilon_S) - \mathcal{E}f(x - \varepsilon_S)|^p \leq \frac{1}{2^n} \sum_{\varepsilon \in \{-1, 1\}^n} \sum_{x \in \mathbb{Z}^n_m} \sum_{j \in S} \varepsilon_j |\mathcal{E}f(x + \varepsilon_j) - \mathcal{E}f(x - \varepsilon_j)|^p + \sum_{j \in S} \sum_{x \in \mathbb{Z}^n_m} |f(x + \varepsilon_j) - f(x)|^p.
$$

(125)
By averaging \([125]\) over those \(S \subseteq \{1, \ldots, n\}\) with \(|S| = k\) we see that
\[
\frac{\beta_p}{2^n \binom{n}{k}} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{\xi \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_4^k} |E f(x + \xi S) - E f(x - \xi S)|^p
\]
\[
\leq \frac{1}{2^n \binom{n}{k}} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{\xi \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_4^k} \left| \sum_{j \in S} \xi_j [E_j f(x + \xi_j) - E_j f(x - \xi_j)] \right|^p
\]
\[
+ \frac{k}{n} \sum_{j=1}^n \sum_{x \in \mathbb{Z}_4^k} |f(x + \xi_j) - f(x)|^p.
\] (126)

Note that since \(E_j\) is an averaging operator,
\[
\sum_{j=1}^n \sum_{x \in \mathbb{Z}_4^k} |E_j f(x + \xi_j) - E_j f(x - \xi_j)|^p \leq 2^p \sum_{j=1}^n \sum_{x \in \mathbb{Z}_4^k} |f(x + \xi_j) - f(x)|^p.
\]

Hence, using the linear \(X_p\) inequality \([15]\), we deduce that
\[
\frac{(p/\log p)^{-p}}{2^n \binom{n}{k}} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{\xi \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_4^k} \left| \sum_{j \in S} \xi_j [E_j f(x + \xi_j) - E_j f(x - \xi_j)] \right|^p
\]
\[
\leq_p \frac{k}{n} \sum_{j=1}^n \sum_{x \in \mathbb{Z}_4^k} |f(x + \xi_j) - f(x)|^p
\]
\[
+ \frac{(k/n)^{2/\pi}}{2^n} \sum_{\xi \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_4^k} \left| \sum_{j=1}^n \xi_j [E_j f(x + \xi_j) - E_j f(x - \xi_j)] \right|^p.
\] (127)

The same reasoning that leads to the identity \([115]\) (alternatively, by \([39, \text{Sec. 5}]\) shows that if for fixed \(x \in \mathbb{Z}_4^k\) we define \(g^x : \{-1,1\}^n \to \mathbb{R}\) by setting \(g^x(\xi) = f(x + \xi) - f(x)\) for every \(\xi \in \{-1,1\}^n\), then that Rademacher projection of \(g^x\) satisfies
\[
\text{Rad}(g^x)(\xi) = \frac{i}{2} \sum_{j=1}^n \xi_j [E_j f(x + \xi_j) - E_j f(x - \xi_j)].
\]

Hence, recalling that the \(K\)-convexity constant of \(\mathbb{R}\) satisfies \(K_p(\mathbb{R}) \lesssim \sqrt{p}\),
\[
\sum_{\xi \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_4^k} \left| \sum_{j=1}^n \xi_j [E_j f(x + \xi_j) - E_j f(x - \xi_j)] \right|^p \lesssim_p \frac{p}{2^n} \sum_{\xi \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_4^k} |f(x + \xi) - f(x)|^p.
\] (128)

By combining (126) with (127) and (128) we have
\[
\frac{(p^{3/2}/\log p)^{-p} \beta_p}{2^n \binom{n}{k}} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{\xi \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_4^k} |E f(x + \xi S) - E f(x - \xi S)|^p
\]
\[
\lesssim_p \frac{k}{n} \sum_{j=1}^n \sum_{x \in \mathbb{Z}_4^k} |f(x + \xi_j) - f(x)|^p + \frac{(k/n)^{2/\pi}}{2^n} \sum_{\xi \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_4^k} |f(x + \xi) - f(x)|^p.
\]
Recalling \((123)\) and \((124)\), we therefore have
\[
\frac{(p^{3/2}/\log p)^{-p} \beta_p}{2^n \binom{n}{k}} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{|S|=k} \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_{4m}^n} |f(x + 2m \varepsilon S) - f(x)|^p \]
\[
\lesssim_p \frac{k}{n} \sum_{j=1}^n \sum_{x \in \mathbb{Z}_{4m}^n} |f(x + \varepsilon_j) - f(x)|^p
\]
\[
+ \left(1 + \frac{(p^{3/2}/\log p)^{-p}}{m^p (k/n)^{3/2}}\right) \frac{(k/n)^{3/2}}{2^n} \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_{4m}^n} |f(x + \varepsilon) - f(x)|^p. \quad \Box
\]

7. The Schatten p trace class is an \(X_p\) Banach space

For \(p \in [1, \infty)\) and \(d \in \mathbb{N}\), the Schatten \(p\)-norm of a \(d\) by \(d\) matrix \(A \in M_d(\mathbb{R})\) is defined as
\[
\|A\|_{S_p} = \left(\text{Tr}\left( (A^*A)^{\frac{p}{2}} \right) \right)^{\frac{1}{p}} = \left(\text{Tr}\left( (AA^*)^{\frac{p}{2}} \right) \right)^{\frac{1}{p}}.
\]
See \([52]\) for relevant background. The following theorem asserts that \(S_p\) is an \(X_p\) Banach space.

**Theorem 7.1.** Fix \(p \in [2, \infty), d, n \in \mathbb{N}\) and \(k \in \{1, \ldots, n\}\). Then for every \(A_1, \ldots, A_n \in M_d(\mathbb{R})\),
\[
\frac{(p/\sqrt{\log p})^{-p}}{2^n \binom{n}{k}} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{|S|=k} \left\| \sum_{j \in S} \varepsilon_j A_j \right\|_{S_p}^p \lesssim_p \frac{k}{n} \sum_{j=1}^n \|A_j\|_{S_p}^p + \frac{(k/n)^{3/2}}{2^n} \sum_{\varepsilon \in \{-1,1\}^n} \left\| \sum_{j=1}^n \varepsilon_j A_j \right\|_{S_p}^p.
\]

**Question 7.2.** It remains an interesting open problem to determine whether or not the quantity \(p/\sqrt{\log p}\) in Theorem 7.1 can be replaced by the (sharp) quantity \(p/\log p\). This was proved in the scalar case in \([26]\), but additional ideas seem to be required in order to carry out the proof of \([26]\) in the above noncommutative setting.

The key step in the proof of Theorem 7.1 is the following proposition.

**Proposition 7.3.** Fix \(q \in [1, \infty), d, n \in \mathbb{N}\) and \(k \in \{1, \ldots, n\}\). Suppose that \(B_1, \ldots, B_n \in M_d(\mathbb{R})\) are symmetric and positive semidefinite. Then
\[
\frac{1}{\binom{n}{k}} \sum_{S \subseteq \{1, \ldots, n\}} \left( \sum_{j \in S} B_j \right)^q \lesssim_q \left( \frac{q}{\log(2q)} \right)^q \max \left\{ \frac{k}{n} \sum_{j=1}^n \text{Tr}\left( B_j^q \right), \left( \frac{k}{n} \right)^q \text{Tr}\left( \left( \sum_{j=1}^n B_j \right)^q \right) \right\}.
\]

Before proving Proposition 7.3, we assume its validity for the moment and proceed to show how it implies Theorem 7.1.

**Proof of Theorem 7.1** Lust-Piquard’s noncommutative Khinchine inequality \([33]\) asserts that for every \(S \subseteq \{1, \ldots, n\}\) we have
\[
\frac{p^{-\frac{3}{2}}}{2^n} \sum_{\varepsilon \in \{-1,1\}^n} \left\| \sum_{j \in S} \varepsilon_j A_j \right\|_{S_p}^p \lesssim \text{Tr}\left( \left( \sum_{j \in S} A_j^* A_j \right)^{\frac{p}{2}} \right) + \text{Tr}\left( \left( \sum_{j \in S} A_j A_j^* \right)^{\frac{p}{2}} \right). \tag{129}
\]

The (asymptotically optimal) dependence on \(p\) in the left-hand side of \((129)\) is not stated in Lust-Piquard’s original proof of \((129)\), but it can be found in \([51]\ page 106). By averaging \((129)\) over all
those $S \subseteq \{1, \ldots, n\}$ with $|S| = k$ we see that

$$\frac{p^{-p}}{2^n \binom{n}{k}} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{|S| = k} \left\| \sum_{j \in S} \varepsilon_j A_j \right\|_{S_p}^p \leq_p \frac{1}{\binom{n}{k}} \sum_{|S| = k} \left( \sum_{j \in S} A_j^* A_j \right)^{\frac{p}{2}} + \frac{1}{\binom{n}{k}} \sum_{|S| = k} \left( \sum_{j \in S} A_j A_j^* \right)^{\frac{p}{2}}. \quad (130)$$

Two applications of Proposition 7.3 with $q = p/2 \geq 1$, once with $B_j = A_j^* A_j$ and once with $B_j = A_j A_j^*$, so as to control the two terms that appear in the right-hand side of (130), yield

$$\frac{(p/\sqrt{\log p})^{-p}}{2^n \binom{n}{k}} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{|S| = k} \left\| \sum_{j \in S} \varepsilon_j A_j \right\|_{S_p}^p \leq_p \frac{k}{n} \sum_{j=1}^n \| A_j \|_{S_p}^p + \frac{k}{n} \left( \sum_{j=1}^n A_j \right)^{\frac{p}{2}} + \frac{k}{n} \left( \sum_{j=1}^n A_j^* \right)^{\frac{p}{2}}. \quad (131)$$

The other direction of Lust-Piquard’s noncommutative Khinchine inequality [33] asserts that

$$\left( \sum_{j=1}^n A_j^* A_j \right)^{\frac{p}{2}} + \left( \sum_{j=1}^n A_j A_j^* \right)^{\frac{p}{2}} \leq \frac{1}{2^n} \sum_{\varepsilon \in \{-1,1\}^n} \left\| \sum_{j=1}^n \varepsilon_j A_j \right\|_{S_p}^p. \quad (132)$$

Theorem 7.1 now follows by combining (131) and (132).

Lemma 7.4 below makes the same assertion as Proposition 7.3 but only for $k \leq n/2$ (and an explicit universal constant that arises from our proof; we do not claim that it is optimal). This is actually the main step in the proof of Proposition 7.3, which we will show below to easily follow from Lemma 7.4.

**Lemma 7.4.** Fix $q \in [1, \infty)$ and $d, k, n \in \mathbb{N}$ with $k \leq n/2$. Then for every $B_1, \ldots, B_n \in M_d(\mathbb{R})$ that are symmetric and positive semidefinite we have

$$\frac{1}{\binom{n}{k}} \sum_{S \subseteq \{1, \ldots, n\}} \left( \sum_{j \in S} B_j \right)^q \leq \left( \frac{4q}{\log(2q)} \right)^q \max \left\{ \frac{k}{n} \sum_{j=1}^n \text{Tr} \left( B_j^q \right), \left( \frac{k}{n} \right)^q \sum_{j=1}^n \text{Tr} \left( B_j^q \right) \right\}.$$

Assuming the validity of Lemma 7.4 for the moment, we proceed to deduce Proposition 7.3, which amounts to removing the restriction $k \leq n/2$ in Lemma 7.4.

**Proof of Proposition 7.3.** Write $k = u + v$ with $u, v \in \mathbb{N}$ satisfying $u, v \leq n/2$. By the triangle inequality in $S_q$, for every $S, T \subseteq \{1, \ldots, n\}$ with $T \subseteq S$ we have

$$\text{Tr} \left( \left( \sum_{j \in S} B_j \right)^q \right) \leq \left\| \sum_{s \in T} B_s + \sum_{s \in S \setminus T} B_s \right\|_{S_q}^q \leq 2^{q-1} \left\| \sum_{s \in T} B_s \right\|_{S_q}^q + 2^{q-1} \left\| \sum_{s \in S \setminus T} B_s \right\|_{S_q}^q.$$
Consequently,

\[
\frac{1}{(k)} \sum_{S \subseteq \{1, \ldots, n\}, |S| = k} \text{Tr} \left( \left( \sum_{j \in S} B_j \right)^q \right) \leq \frac{1}{(k)} \sum_{S \subseteq \{1, \ldots, n\}, |S| = k} \frac{2^{q-1}}{(u)} \sum_{T \subseteq S, |T| = u} \left( \left\| \sum_{s \in T} B_s \right\|_{S_q}^q + \left\| \sum_{s \in S \setminus T} B_s \right\|_{S_q}^q \right)
\]

\[
= \frac{2^{q-1}}{(u)} \sum_{U \subseteq \{1, \ldots, n\}, |U| = u} \text{Tr} \left( \left( \sum_{j \in U} B_j \right)^q \right) + \frac{2^{q-1}}{(v)} \sum_{V \subseteq \{1, \ldots, n\}, |V| = v} \text{Tr} \left( \left( \sum_{j \in V} B_j \right)^q \right). \tag{133}
\]

Proposition 7.3 now follows by applying Lemma 7.4 to each of the summands that appear in the right-hand side of (133).

Our proof of Lemma 7.4 relies on certain matrix inequalities of independent interest. These inequalities are established in the following section.

7.1. Auxiliary trace inequalities. Proposition 7.5 and Proposition 7.8 below will be used crucially in the proof of Lemma 7.4. Note that the same statements are trivial when matrices are replaced by scalars. See Section 7.1.1 for a discussion on the context of these results.

**Proposition 7.5.** Suppose that \( q \in [1, \infty) \) and \( d \in \mathbb{N} \). Then for every \( A, B \in M_d(\mathbb{R}) \) that are symmetric and positive semidefinite we have

\[
\left( \text{Tr} \left( (A + B)^q A \right) \right)^{\frac{1}{q}} \leq \left( \text{Tr} \left( A^{q+1} \right) \right)^{\frac{1}{q}} + \left( \text{Tr} \left( B^q A \right) \right)^{\frac{1}{q}}. \tag{134}
\]

Before proving Proposition 7.5, we record for future use the following H"older-type estimate.

**Lemma 7.6.** Fix \( d, k \in \mathbb{N} \) and \( q \in (0, \infty) \). Suppose that \( a_0, \ldots, a_{k-1}, b_1, \ldots, b_k \in (0, \infty) \) satisfy \( b_j + b_{j+1} \leq 2q a_j \) for every \( j \in \{0, \ldots, k-1\} \), where we set \( b_0 = b_k \). Suppose also that

\[
\sum_{j=0}^{k-1} a_j + \sum_{j=1}^{k} b_j = q + 1. \tag{135}
\]

Then for every \( A, B \in M_d(\mathbb{R}) \) that are symmetric and positive semidefinite we have

\[
\text{Tr} \left( A^{q+1} \left( \prod_{j=1}^{k-1} B_{p_j}^{b_j} A_{p_j}^{a_j} \right) B_{p_j}^{b_j} \right) \leq \left( \text{Tr} \left( A^{q+1} \right) \right)^{1 - \frac{1}{q}} \left( \text{Tr} \left( B^q A \right) \right)^{\frac{1}{q}} \sum_{j=1}^{k} b_j. \tag{136}
\]

**Proof.** By applying an arbitrarily small perturbation, we may assume that \( a_j - (b_j + b_{j+1})/(2q) > 0 \) for every \( j \in \{0, \ldots, k-1\} \). We can then define \( p_0, r_0, \ldots, p_{k-1}, r_{k-1} \in (0, \infty) \) by

\[
\forall j \in \{0, \ldots, k-1\}, \quad p_j \overset{\text{def}}{=} \frac{q + 1}{a_j - \frac{b_j + b_{j+1}}{2q}} \quad \text{and} \quad r_j \overset{\text{def}}{=} \frac{q}{b_j + 1}. \tag{136}
\]

Using the cyclicity of the trace, the choices in (136) imply that we have

\[
\text{Tr} \left( A^{q+1} \left( \prod_{j=1}^{k-1} B_{p_j}^{b_j} A_{p_j}^{a_j} \right) B_{p_j}^{b_j} \right) = \text{Tr} \left( \prod_{j=0}^{k-1} A_{p_j}^{q+1} \left( A^{\frac{1}{2q}} B_{r_j}^{q} A^{\frac{1}{2q}} \right) \right), \tag{137}
\]

Moreover,

\[
\sum_{j=0}^{k-1} \frac{1}{p_j} + \sum_{j=0}^{k-1} \frac{1}{r_j} = \frac{1}{q + 1} \sum_{j=0}^{k-1} a_j + \left( \frac{1}{q} - \frac{1}{q(q + 1)} \right) \sum_{j=1}^{k} b_j \overset{\text{135}}{=} 1.
\]
Therefore \( p_j, r_j \in (1, \infty) \) for all \( j \in \{0, \ldots, k-1\} \) and we may use Hölder’s inequality for traces (Théorème 6 in [13]) to deduce from (137) that
\[
\text{Tr} \left( A^{a_0} \left( \prod_{j=1}^{k-1} B^{b_j} A^{a_j} \right) B^{b_k} \right) \leq \prod_{j=0}^{k-1} \left( \text{Tr} \left( A^{a_j+1} \right) \right)^{\frac{1}{r_j}} \left( \text{Tr} \left( \left( A^{\frac{1}{r_j}} B^{\frac{q}{r_j}} A^{\frac{1}{r_j}} \right)^{r_j} \right) \right)^{\frac{1}{r_j}}. \tag{138}
\]

The Lieb-Thirring inequality [30] asserts that \( \text{Tr}((XYX)^r) \leq \text{Tr}(X^rY^rX^r) \) for every \( r \in [1, \infty) \) and for every symmetric and positive semidefinite matrices \( X, Y \in M_d(\mathbb{R}) \). Recalling the definition of \( r_0, \ldots, r_{k-1} \) in (136), for every \( j \in \{0, \ldots, k-1\} \) we therefore have
\[
\text{Tr} \left( \left( A^{\frac{1}{r_j}} B^{\frac{q}{r_j}} A^{\frac{1}{r_j}} \right)^{r_j} \right) \leq \text{Tr} \left( \sqrt{AB}^q \sqrt{A} \right) = \text{Tr} (B^q A).
\]

A substitution of this estimate into (138) gives
\[
\text{Tr} \left( A^{a_0} \left( \prod_{j=1}^{k-1} B^{b_j} A^{a_j} \right) B^{b_k} \right) \leq \left( \text{Tr} \left( A^{a_j+1} \right) \right)^{\frac{1}{r_j}} \left( \sum_{j=0}^{k-1} \frac{1}{r_j} \left( \text{Tr} \left( B^q A \right) \right)^{\frac{1}{r_j}} \right)^{\frac{k-1}{r_j}}.
\] (139)

where we used the fact that, due to (136), we have \( \sum_{j=0}^{k-1} \frac{1}{r_j} = \frac{1}{q} \sum_{j=1}^{k} b_j \).

Remark 7.7. For future use, note that if \( q, a_0, \ldots, a_{k-1}, b_1, \ldots, b_k \in (0, \infty) \) satisfy (135) and we also know that \( a_0, \ldots, a_{k-1} \geq 1 \) then the assumptions of Lemma 7.6 hold true, i.e., \( b_j + b_{j+1} \leq 2qa_j \) for every \( j \in \{0, \ldots, k-1\} \). Indeed, by (135) we have \( \max\{b_j, b_{j+1}\} \leq q + 1 - a_j \leq qa_j \), and consequently \( b_j + b_{j+1} \leq 2 \max\{b_j, b_{j+1}\} \leq 2qa_j \).

Proof of Proposition 7.4. Write \( q = 2m + \theta \), where \( m \in \mathbb{N} \cup \{0\} \) and \( \theta \in (0, 2] \). The proof of (134) treats the cases \( \theta \in (0, 1) \) and \( \theta \in [1, 2] \) differently.

Case 1: \( \theta \in [1, 2] \). In this range the mapping \( t \to t^\theta \) is operator-convex (see Theorem 2.6 in [13]). This means that for every \( s \in (0, 1) \) we have
\[
(A + B)^\theta = \left( \frac{A}{s} + \frac{B}{1-s} \right)^{\theta} \leq \frac{A^\theta}{s^{\theta-1}} + \frac{B^\theta}{(1-s)^{\theta-1}}, \tag{140}
\]
where, as usual, we interpret the inequality (140) in terms of the PSD order of matrices, i.e., that the right-hand side of (140) minus the left-hand side of (140) is a positive semidefinite matrix.

It follows from (140) that
\[
\sqrt{A} (A + B)^q \sqrt{A} \leq \sqrt{A} (A + B)^m A^\theta (A + B)^m \sqrt{A} + \sqrt{A} (A + B)^m B^\theta (A + B)^m \sqrt{A} \leq s^{\theta-1} \text{Tr} (A + B)^m A^\theta (A + B)^m A.
\]

So, by taking traces while making use of the cyclicity of the trace, we see that
\[
\text{Tr} ((A + B)^q A) \leq \frac{\text{Tr} ((A + B)^m A^\theta (A + B)^m A)}{s^{\theta-1}} + \frac{\text{Tr} ((A + B)^m B^\theta (A + B)^m A)}{(1-s)^{\theta-1}}. \tag{141}
\]

By choosing \( s \) so as to minimize the quantity appearing in the right-hand side of (141), we have
\[
\left( \text{Tr} ((A + B)^q A) \right)^{\frac{1}{q}} \leq \left( \text{Tr} ((A + B)^m A^\theta (A + B)^m A) \right)^{\frac{1}{q}} \cdot \left( \text{Tr} ((A + B)^m B^\theta (A + B)^m A) \right)^{\frac{1}{q}}. \tag{142}
\]

We shall now proceed to estimate each of the terms that appear in the right-hand side of (142) separately. By expanding the \( m \)th powers appearing in the matrix \( (A + B)^m A^\theta (A + B)^m A \), and
using the cyclicity of the trace, we see that $\text{Tr} \left( (A + B)^m A^\theta (A + B)^m A \right)$ equals the sum of $2^m$ terms, each of which is of the form

$$\text{Tr} \left( A^{a_0} \left( \prod_{j=1}^{k-1} B^{b_j} A^{a_j} \right) B^b \right), \quad (143)$$

for some $k \in \mathbb{N} \cup \{0\}$ and $a_0, \ldots, a_{k-1}, b_1, \ldots, b_k \in (0, \infty)$ that satisfy (135) (recall that $q = 2m + \theta$). Here we use the convention that when $k = 0$ the quantity appearing in (143) equals $\text{Tr} (A^{q+1})$. Note that $b_j$ is an integer for every $j \in \{1, \ldots, k\}$, and for every $r \in \{0, \ldots, 2m\}$ the number of terms of the form (143) that appear in the above expansion of $\text{Tr} \left( (A + B)^m A^\theta (A + B)^m A \right)$ with $\sum_{j=1}^k b_j \equiv r$ equals $\binom{2m}{r}$; this is because $\sum_{j=1}^k b_j$ is the total number of times that $B$ was chosen when one expands the two occurrences of $(A + B)^m$ in $(A + B)^m A^\theta (A + B)^m A$ as a product of matrices, each of which is either $A$ or $B$. Note also that $a_0, \ldots, a_{k-1} \geq 1$, since $\theta \geq 1$. Recalling Remark 7.7 we may therefore use Lemma 7.6 to deduce that

$$\text{Tr} \left( A^{a_0} \left( \prod_{j=1}^{k-1} B^{b_j} A^{a_j} \right) B^b \right) \leq (\text{Tr} (A^{q+1}))^{1 - \frac{1}{q} \sum_{j=1}^k b_j} \left( \text{Tr} (B^q A) \right)^{\frac{1}{q} \sum_{j=1}^k b_j}, \quad (144)$$

Hence,

$$\text{Tr} \left( (A + B)^m A^\theta (A + B)^m A \right) \leq \sum_{r=0}^{2m} \binom{2m}{r} \left( \text{Tr} (A^{q+1}) \right)^{1 - \frac{2m + \theta}{q}} \left( \text{Tr} (B^q A) \right)^{\frac{r + \theta}{q}} = \left( \text{Tr} (A^{q+1}) \right)^{\frac{2m + \theta}{q}} \left( \text{Tr} (B^q A) \right)^{\frac{1}{q}} \left( \text{Tr} (A^{q+1}) \right)^{\frac{1}{q}} \left( \text{Tr} (B^q A) \right)^{\frac{1}{q}}, \quad (145)$$

where in the final step we used the fact that $2m + \theta = q$.

The second term in the right-hand side of (142) is bounded using similar reasoning. As before, $\text{Tr} \left( (A + B)^m B^\theta (A + B)^m A \right)$ equals the sum of terms as in (143), for some $k \in \mathbb{N} \cup \{0\}$ and $a_0, \ldots, a_{k-1}, b_1, \ldots, b_k \in (0, \infty)$ that satisfy (135). However, now we know that $a_1, \ldots, a_{k-1} \in \mathbb{N}$ and $\sum_{j=0}^{k-1} b_j \equiv - \theta \in \{0, \ldots, 2m\}$. By Lemma 7.6 (and Remark 7.7), the estimate (144) holds true for the terms of the form (143) that appear in the expansion of $\text{Tr} \left( (A + B)^m B^\theta (A + B)^m A \right)$. For every $r \in \{0, \ldots, 2m\}$, the number of terms of the form (143) that appear in the expansion of $\text{Tr} \left( (A + B)^m B^\theta (A + B)^m A \right)$ with $\sum_{j=1}^k b_j = r + \theta$ equals $\binom{2m}{r}$, so by (144) we have

$$\text{Tr} \left( (A + B)^m B^\theta (A + B)^m A \right) \leq \sum_{r=0}^{2m} \binom{2m}{r} \left( \text{Tr} (A^{q+1}) \right)^{1 - \frac{2m + \theta}{q}} \left( \text{Tr} (B^q A) \right)^{\frac{r + \theta}{q}} = \left( \text{Tr} (B^q A) \right)^{\frac{2m + \theta}{q}} \left( \text{Tr} (A^{q+1}) \right)^{\frac{1}{q}} \left( \text{Tr} (B^q A) \right)^{\frac{1}{q}}, \quad (146)$$

where the last step uses the fact that $2m + \theta = q$.

By substituting (145) and (146) into (142) we see that

$$\left( \text{Tr} \left( (A + B)^q A \right) \right)^{\frac{1}{q}} \leq \left( \left( \text{Tr} (A^{q+1}) \right)^{\frac{1}{q}} + \left( \text{Tr} (B^q A) \right)^{\frac{1}{q}} \right)^{1 + \frac{2m}{q}} = \left( \left( \text{Tr} (A^{q+1}) \right)^{\frac{1}{q}} + \left( \text{Tr} (B^q A) \right)^{\frac{1}{q}} \right)^{\frac{1}{q}}, \quad 44$$
following from (147) that is an invertible matrix for every and .

Equivalently, we want to prove that it follows that for every

\[
\begin{align*}
\forall s \in (0, \infty), && (sA + B)^\theta = \sin(\pi \theta) \int_0^\infty t^\theta \left( \frac{1}{t} - \frac{1}{t + a} \right) dt,
\end{align*}
\]

it follows that for every \( s \in (0, \infty) \),

\[
(sA + B)^\theta = \sin(\pi \theta) \int_0^\infty t^\theta \left( \frac{1}{t} I - (tI + sA + B)^{-1} \right) dt. \tag{147}
\]

Since \( \frac{d}{dt} X(t)^{-1} = -X(t)^{-1} X'(t) X(t)^{-1} \) for every differentiable \( X : (0, \infty) \to M_d(\mathbb{R}) \) such that \( X(t) \) is an invertible matrix for every \( t \in (0, \infty) \) (simply differentiate the identity \( X(t)^{-1} X(t) = I \)), it follows from (147) that

\[
\frac{d}{ds} (sA + B)^\theta = \sin(\pi \theta) \int_0^\infty t^\theta \left( tI + sA + B \right)^{-1} A (tI + sA + B)^{-1} dt. \tag{148}
\]

By integrating over \( s \in [0, 1] \), in order to prove (134) it will suffice to show that

\[
\forall s \in (0, 1), \quad \frac{d}{ds} \left( (sA + B)^q A \right) \leq q \left( Tr ((sA + B)^q A) \right)^{\frac{1}{q}}. \tag{149}
\]

Equivalently, we want to prove that

\[
\forall s \in (0, 1), \quad \frac{d}{ds} \left( (sA + B)^q A \right) \leq q \left( Tr ((sA + B)^q A) \right)^{\frac{1}{q}} \left( (sA + B)^q A \right)^{1 - \frac{1}{q}}. \tag{149}
\]

Define for every \( s \in (0, 1) \),

\[
f(s) \overset{\text{def}}{=} Tr \left( \left( \frac{d}{ds} (sA + B)^m \right) (sA + B)^{m + \theta} A \right),
\]

\[
g(s) \overset{\text{def}}{=} Tr \left( (sA + B)^m \left( \frac{d}{ds} (sA + B)^\theta \right) (sA + B)^m A \right),
\]

and

\[
h(s) \overset{\text{def}}{=} Tr \left( (sA + B)^{m + \theta} \left( \frac{d}{ds} (sA + B)^m \right) A \right).
\]

Then, since \( (sA + B)^q = (sA + B)^m (sA + B)^\theta (sA + B)^m \), we have

\[
\frac{d}{ds} \left( (sA + B)^q A \right) = f(s) + g(s) + h(s).
\]

Hence, because \( q = 2m + \theta \), in order to establish the validity of (149) it suffice to show that for every \( s \in [0, 1] \) we have

\[
\max \{ f(s), h(s) \} \leq m \left( Tr ((sA + B)^{q+1}) \right)^{\frac{1}{q}} \left( Tr ((sA + B)^q A) \right)^{1 - \frac{1}{q}}, \tag{150}
\]

and

\[
g(s) \leq \theta \left( Tr ((sA + B)^{q+1}) \right)^{\frac{1}{q}} \left( Tr ((sA + B)^q A) \right)^{1 - \frac{1}{q}}. \tag{151}
\]

Observe that

\[
f(s) = \sum_{r=1}^m Tr \left( (sA + B)^{q-1} A (sA + B)^{m-r} (sA + B)^{m+\theta} A \right)
\]

\[
= \sum_{r=1}^m Tr \left( (sA + B)^{q-1} A (sA + B)^{q-r} A \right). \tag{152}
\]
Similarly, using the cyclicity of the trace, we have

\[ h(s) = \sum_{r=1}^{m} \text{Tr} \left( (sA + B)^{q-r} A (sA + B)^{r-1} A \right) = f(s). \] (153)

Finally, by the integral representation (148) we have

\[ g(s) = \frac{\sin(\pi \theta)}{\pi} \int_{0}^{\infty} t^\theta \text{Tr} \left( (sA + B)^m (tI + sA + B)^{-1}A(tI + sA + B)^{-1}(sA + B)^m A \right) dt. \] (154)

By denoting \( C = sA + B \), it follows from (152), (153) and (154) that the desired estimates (150) and (151) will be proven once we show that for every \( C \in M_d(\mathbb{R}) \) that is symmetric and positive semidefinite we have

\[ \forall r \in \{1, \ldots, m\}, \quad \text{Tr} \left( C^{r-1} A C^{q-r} A \right) \leq \left( \text{Tr} \left( A^{q+1} \right) \right)^{\frac{1}{q}} \left( \text{Tr} \left( C^q A \right) \right)^{1-\frac{1}{q}}, \] (155)

and

\[ \int_{0}^{\infty} t^\theta \text{Tr} \left( C^m (tI + C)^{-1}A(tI + C)^{-1}C^m A \right) dt \leq \frac{\pi \theta}{\sin(\pi \theta)} \left( \text{Tr} \left( A^{q+1} \right) \right)^{\frac{1}{q}} \left( \text{Tr} \left( C^q A \right) \right)^{1-\frac{1}{q}}. \] (156)

(155) is a consequence of Lemma 7.6 (with \( B \) replaced by \( C \)). It therefore remains to establish the validity of (156). To this end, note that for every \( t \in (0, \infty) \), since \((tI + C)^{-1}\) and \( C^m \) commute, by the cyclicity of the trace we have

\[ \text{Tr} \left( C^m(tI + C)^{-1}A(tI + C)^{-1}C^m A \right) = \text{Tr} \left( \left( \sqrt{A} C^m(tI + C)^{-1}\sqrt{A} \right)^2 \right) \]

\[ \leq \text{Tr} \left( A C^{2m}(tI + C)^{-2}A \right) = \text{Tr} \left( C^{2m}(tI + C)^{-2}A^2 \right), \] (157)

where for the inequality in (157) we used the Lieb-Thirring inequality. This upper bound on the integrand in the left-hand side of (156) yields the following estimate.

\[ \int_{0}^{\infty} t^\theta \text{Tr} \left( C^m(tI + C)^{-1}A(tI + C)^{-1}C^m A \right) dt \leq \text{Tr} \left( C^{2m} \left( \int_{0}^{\infty} t^\theta(tI + C)^{-2}dt \right) A^2 \right). \] (158)

Note that for every \( c \in (0, \infty) \) we have

\[ \int_{0}^{\infty} \frac{t^\theta}{(t + c)^2} dt = c^{\theta - 1} \int_{0}^{\infty} \frac{s^\theta}{(s + 1)^2} ds = \frac{\pi \theta}{\sin(\pi \theta)} c^{\theta - 1}. \]

Consequently,

\[ \text{Tr} \left( C^{2m} \left( \int_{0}^{\infty} t^\theta(tI + C)^{-2}dt \right) A^2 \right) = \frac{\pi \theta}{\sin(\pi \theta)} \text{Tr} \left( C^{2m+\theta-1}A^2 \right) = \frac{\pi \theta}{\sin(\pi \theta)} \text{Tr} \left( C^{q-1}A^2 \right) \]

\[ = \frac{\pi \theta}{\sin(\pi \theta)} \text{Tr} \left( A C^{q-1} A \right) \leq \frac{\pi \theta}{\sin(\pi \theta)} \left( \text{Tr} \left( A^{q+1} \right) \right)^{\frac{1}{q}} \left( \text{Tr} \left( C^q A \right) \right)^{1-\frac{1}{q}}. \] (159)

A substitution of (159) into (158) yields the desired inequality (156).

The following Proposition is a variant of Proposition 7.5 when \( q \in (0, 1) \).

**Proposition 7.8.** Suppose that \( q \in (0, 1) \) and \( d \in \mathbb{N} \). Then for every \( A, B \in M_d(\mathbb{R}) \) that are symmetric and positive semidefinite we have

\[ \text{Tr} \left( (A + B)^q A \right) \leq \text{Tr} \left( A^{q+1} \right) + \text{Tr} \left( B^q A \right). \]
Proof. By the integral identity (148), with \( \theta \) replaced by \( q \) (which is allowed since \( 0 < q < 1 \)), for every \( s \in (0, \infty) \) we have
\[
\frac{d}{ds} \text{Tr} ((sA + B)^q A) = \frac{\sin(\pi q)}{\pi} \int_0^\infty t^q \text{Tr} \left((tI + sA + B)^{-1} A (tI + sA + B)^{-1} A\right) dt. \tag{160}
\]

Fix \( s, t \in (0, \infty) \) and define \( F : [0, \infty) \rightarrow \mathbb{R} \) by
\[
\forall w \in [0, \infty), \quad F(w) \overset{\text{def}}{=} \text{Tr} \left((tI + sA + wB)^{-1} A (tI + sA + wB)^{-1} A\right).
\]
This mapping was investigated in Section III of [5], where it was shown to be convex. Here we need to know that it is non-increasing, which follows from the following computation.
\[
F'(w) = - \text{Tr} \left((tI + sA + wB)^{-1} B (tI + sA + wB)^{-1} A (tI + sA + wB)^{-1} A\right) - \text{Tr} \left((tI + sA + wB)^{-1} A (tI + sA + wB)^{-1} B (tI + sA + wB)^{-1} A\right)
\]
\[
= - \text{Tr}(CD) - \text{Tr}(DC) = -2 \text{Tr}(CD) \leq 0,
\]
where \( C, D \in M_d(\mathbb{R}) \) are the symmetric and positive semidefinite matrices given by
\[
C \overset{\text{def}}{=} (tI + sA + wB)^{-1} B (tI + sA + wB)^{-1} \quad \text{and} \quad D \overset{\text{def}}{=} A (tI + sA + wB)^{-1} A.
\]
It follows from these considerations that
\[
\text{Tr} \left((tI + sA + B)^{-1} A (tI + sA + B)^{-1} A\right) = F(1) \leq F(0) = \text{Tr} \left((tI + sA)^{-1} A (tI + sA)^{-1} A\right).
\]
A substitution of this estimate into (160) shows that
\[
\frac{d}{ds} \text{Tr} ((sA + B)^q A) \leq \frac{\sin(\pi q)}{\pi} \int_0^\infty t^q \text{Tr} \left((tI + sA)^{-1} A (tI + sA)^{-1} A\right) dt
\]
\[
\overset{\text{(137)}}{=} \frac{d}{ds} \text{Tr} ((sA)^q A) = qs^{q-1} \text{Tr} (A^{q+1}). \tag{161}
\]
By integrating (161) over \([0, 1]\) we therefore see that \( \text{Tr} ((A + B)^q A) - \text{Tr} (B^q A) \leq \text{Tr} (A^{q+1}) \). \( \square \)

We record for future use the following simple reformulation of Proposition 7.5 and Proposition 7.8. When \( q \in [1, 2] \) it follows from Proposition 7.8 (with \( q \) replaced by \( q - 1 \)), and when \( q > 2 \) it follows from Proposition 7.5 (with \( q \) replaced by \( q - 1 \)) and the convexity of \( t \rightarrow t^{q-1} \) on \([0, \infty)\).

**Corollary 7.9.** Suppose that \( q \in [1, \infty) \) and \( d \in \mathbb{N} \). Set \( r \overset{\text{def}}{=} \max\{q-2, 0\} \). For every \( A, B \in M_d(\mathbb{R}) \) that are symmetric and positive semidefinite we have
\[
\text{Tr} \left((A + B)^{q-1} A\right) \leq \min \left\{ \frac{\text{Tr} (A^q)}{\lambda^r} + \frac{\text{Tr} (B^{q-1} A)}{(1-\lambda)^r} : \lambda \in [0, 1] \right\}.
\]

7.1.1. **Discussion and counterexamples.** An inspection of our proof of Lemma 7.4 below shows that, for \( p \in (2, \infty) \), what we really need in order to show that \( S_p \) is an \( X_p \) Banach space is that there exists \( K = K_p \in (0, \infty) \) such that if \( A, B \in M_d(\mathbb{R}) \) are symmetric and positive semidefinite then
\[
\text{Tr} \left((A + B)^{\frac{p}{2} - 1} A\right) \leq K \left(\text{Tr} (A^{\frac{p}{2} - 1} A) + \text{Tr} (B^{\frac{p}{2} - 1} A)\right). \tag{162}
\]
Specifically, (102) implies Theorem 7.1 with the term \( p/\sqrt{\log p} \) replaced by a constant that depends only on \( K \) and \( p \). By Corollary 7.9, (162) holds true with \( K = 2^{\max\{0,(p-4)/4\}} \).

Setting \( q = (p-2)/2 > 0 \), it is natural to ask whether multiplication by \( A \) is crucial for (162) to hold true. Specifically, one would naturally investigate whether for every \( A, B, C \in M_d(\mathbb{R}) \) that are symmetric and positive semidefinite we have
\[
\text{Tr}((A + B)^q C) \leq K \left(\text{Tr} (A^q C) + \text{Tr} (B^q C)\right), \tag{163}
\]
with \( K \in (0, \infty) \) independent of \( A, B, C \). By a simple duality argument (e.g. Lemma 5.12 in [13]), the above requirement is equivalent to the matrix inequality
\[
(A + B)^q \leq K (A^q + B^q),
\]
where, as usual, we interpret the inequality \((164)\) in terms of the PSD order of matrices.

Since for \( q \in [1, 2] \) the mapping \( t \mapsto t^q \) is operator-convex, for such \( q \) the PSD inequality \((164)\) holds true with \( K = 2^{q-1} \) (recall \((140)\)). This yields a simple proof of \((162)\) when \( 4 \leq p \leq 6 \). However, as we shall explain below, if \( q \in (0, 1) \cup (2, \infty) \) then \((164)\) fails to hold true with any constant \( K \) that is independent of \( A, B \).

The failure of such PSD subadditivity inequalities prompted much work in search for substitutes (note, however, that the literature did not focus on inequalities that allow for an arbitrary constant \( K \)). Indeed, direct computation (via diagonalization) yields that
\[
(A + B)^q \leq K (A^q + B^q),
\]

where, as usual, we interpret the inequality \((164)\) in terms of the PSD order of matrices.

A similar, though more tedious, computation shows that \((164)\) also fails for every \( q \in (0, 1) \cup (2, \infty) \). This shows that \((164)\) fails to hold true for \( A \), \( B \), and consequently obtaining various nonembeddability results), but we believe that they are interesting in their own right and deserve further investigation. Possible extensions include understanding inequalities of the form \( \text{Tr}(f(A + B)) \leq K \text{Tr}(f(A)) + K \text{Tr}(f(B)) \).

We shall end this discussion by presenting the aforementioned example that exhibits the failure of \((164)\) for every \( q \in (0, 1) \cup (2, \infty) \) and \( K \in (0, \infty) \). Fix \( s \in (0, \infty) \) which we will eventually take to be sufficiently small. Define \( A_s, B_s \in M_2(\mathbb{R}) \) and \( w_s \in \mathbb{R}^2 \) by
\[
A_s \overset{\text{def}}{=} \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}, \quad B_s \overset{\text{def}}{=} \begin{pmatrix} 1 & s \\ s & s^2 \end{pmatrix} \quad \text{and} \quad w_s \overset{\text{def}}{=} \begin{pmatrix} -s \\ 1 \end{pmatrix},
\]

\( A_s \) and \( B_s \) are symmetric and positive semidefinite, yet by direct computation for every \( K \in (0, \infty) \),
\[
\langle (K(A_s^4 + B_s^4) - (A_s + B_s)^4) w_s, w_s \rangle = -s^6 - 3s^8 + (K - 1)s^{10}.
\]
The above quantity is negative for \( s < 1/\sqrt{K} \), in which case the matrix \( K(A_s^4 + B_s^4) - (A_s + B_s)^4 \) is not positive semidefinite. This shows that \((164)\) fails to hold true for \( q = 4 \) with any constant \( K \in (0, \infty) \) that is independent of \( A \) and \( B \) (this corresponds to the failure of \((162)\) when \( p = 10 \)). A similar, though more tedious, computation shows that \((164)\) also fails for every \( q \in (0, 1) \cup (2, \infty) \). Indeed, direct computation (via diagonalization) yields that \( A_s^q = s^{2q} A_s, B_s^q = (1 + s^2)^{q-1} B_s \) and
\[
(A_s + B_s)^q = \begin{pmatrix} \frac{\sqrt{1 + 4s^2} + b(s)}{2} & \frac{a(s)\sqrt{1 + 4s^2}}{2} \\ \frac{a(s)\sqrt{1 - b(s)^2}}{2} & \frac{\sqrt{1 + 4s^2}}{2} \end{pmatrix},
\]

where
\[
a(s) \overset{\text{def}}{=} s^2 + \frac{1}{2} + \frac{\sqrt{1 + 4s^2}}{2} \quad \text{and} \quad b(s) \overset{\text{def}}{=} s^2 + \frac{1}{2} - \frac{\sqrt{1 + 4s^2}}{2}.
\]
One then directly computes that as $s \to 0$,
\begin{equation}
((K(A^q_s + B^q_s) - (A_s + B_s)^q) \ w_s, w_s) = \left(K s^{2(q+1)} - s^6 - s^{4q}\right) \left(1 + O_{q,K}(s^2)\right).
\end{equation}
When $q \in (0, 1)$ we have $4q < \min\{2(q+1), 6\}$ and when $q \in (2, \infty)$ we have $6 < \min\{2(q+1), 4q\}$. Consequently, for $q \in (0, 1) \cup (2, \infty)$ the quantity appearing in (163) is negative for small enough $s$, which means that the matrix $K(A^q_s + B^q_s) - (A_s + B_s)^q$ is not positive semidefinite.

7.2. Proof of Lemma 7.4. For the sake of simplicity denote
\begin{equation}
U \defeq \frac{1}{n} \sum_{S \subseteq \{1, \ldots, n\}} |S| = k \quad \text{Tr} \left( \left( \sum_{j \in S} B_j \right)^q \right), \quad V \defeq \frac{k}{n} \sum_{j = 1}^n \text{Tr}(B_j^q), \quad W \defeq \text{Tr} \left( \left( \frac{k}{n} \sum_{j = 1}^n B_j \right)^q \right).
\end{equation}
Our goal is therefore to show that
\begin{equation}
U \leq \left( \frac{4q}{\log(2q)} \right)^q \max\{V, W\}.
\end{equation}
Fix $\lambda \in (0, 1)$ to be specified later. For every $S \subseteq \{1, \ldots, n\}$ and $j \in S$, by Corollary 7.9 with $A = B_j$ and $B = \sum_{s \in S \setminus \{j\}} B_s$, we have
\begin{align*}
\text{Tr} \left( \left( \sum_{s \in S} B_s \right)^{q-1} B_j \right) &\leq \frac{1}{\lambda^r} \text{Tr}(B_j^q) + \frac{1}{(1 - \lambda^r)^{q+1}} \sum_{j \in S} \text{Tr} \left( \left( \sum_{s \in S \setminus \{j\}} B_s \right)^{q-1} B_j \right),
\end{align*}
where, as denoted in Corollary 7.9, $r = \max\{q - 2, 0\}$. Hence,
\begin{equation}
\text{Tr} \left( \left( \sum_{j \in S} B_j \right)^q \right) = \sum_{j = 1}^n \text{Tr} \left( \left( \sum_{j \in S} B_j \right)^{q-1} B_j \right)
\end{equation}
\begin{equation}
\leq \frac{1}{\lambda^r} \sum_{j \in S} \text{Tr}(B_j^q) + \frac{1}{(1 - \lambda^r)^{q+1}} \sum_{j \in S} \text{Tr} \left( \left( \sum_{s \in S \setminus \{j\}} B_s \right)^{q-1} B_j \right).
\end{equation}
By averaging (168) over all of those $S \subseteq \{1, \ldots, n\}$ with $|S| = k$, and recalling (166), we see that
\begin{equation}
U \leq \frac{V}{\lambda^r} + \frac{1}{(1 - \lambda^r)^{q+1}} \sum_{S \subseteq \{1, \ldots, n\}} \sum_{|S| = k} \sum_{j \in S} \text{Tr} \left( \left( \sum_{s \in S \setminus \{j\}} B_s \right)^{q-1} B_j \right).
\end{equation}
Now,
\begin{align*}
\sum_{S \subseteq \{1, \ldots, n\}} \sum_{|S| = k} \text{Tr} \left( \left( \sum_{s \in S \setminus \{j\}} B_s \right)^{q-1} B_j \right) &= \sum_{T \subseteq \{1, \ldots, n\}} \sum_{|T| = k - 1} \text{Tr} \left( \left( \sum_{t \in T} B_t \right)^{q-1} B_j \right) \\
&= \sum_{T \subseteq \{1, \ldots, n\}} \sum_{|T| = k - 1} \text{Tr} \left( \left( \sum_{t \in T} B_t \right)^{q-1} \left( \sum_{j \in \{1, \ldots, n\} \setminus T} B_j \right) \right) \leq \sum_{T \subseteq \{1, \ldots, n\}} \sum_{|T| = k - 1} \text{Tr} \left( \left( \sum_{t \in T} B_t \right)^{q-1} \left( \sum_{j = 1}^n B_j \right) \right),
\end{align*}
where in the last step of (170) we used the fact that if $A, B, C \in M_n(\mathbb{R})$ are symmetric and positive semidefinite then $\text{Tr}(A B) \leq \text{Tr}(A(B + C))$. To bound the final term in (170), use Hölder’s inequality for traces to deduce that for every $T \subseteq \{1, \ldots, n\}$ we have
\[
\text{Tr}\left( \left( \sum_{t \in T} B_t \right)^{q-1} \left( \sum_{j=1}^n B_j \right) \right) \leq \left( \text{Tr}\left( \left( \sum_{j=1}^n B_j \right)^q \right) \right)^{\frac{1}{q}} \left( \text{Tr}\left( \left( \sum_{t \in T} B_t \right)^q \right) \right)^{1-\frac{1}{q}} = \frac{nW_1^q}{k} \left( \text{Tr}\left( \left( \sum_{t \in T} B_t \right)^q \right) \right)^{1-\frac{1}{q}},
\]

(171)

where we recall the definition of \(W\) in (166).

The function \(t \mapsto t^q\) is operator trace-increasing (see Theorem 2.10 in [13]), i.e., if \(C, D \in M_d(\mathbb{R})\) are symmetric and positive semidefinite with \(C \leq D\) then \(\text{Tr}(C^q) \leq \text{Tr}(D^q)\). Consequently, for every \(T \subseteq \{1, \ldots, n\}\) and \(u \in \{1, \ldots, n\}\) we have \(\text{Tr}\left( \left( \sum_{t \in T} B_t \right)^q \right) \leq \text{Tr}\left( \left( B_u + \sum_{t \in T} B_t \right)^q \right)\). By raising this inequality to the power \((q-1)/q\) and averaging over all \(u \in \{1, \ldots, n\} \setminus T\) we see that

\[
\left( \text{Tr}\left( \left( \sum_{t \in T} B_t \right)^q \right) \right)^{1-\frac{1}{q}} \leq \frac{1}{n - |T|} \sum_{u \in \{1, \ldots, n\} \setminus T} \left( \text{Tr}\left( \left( \sum_{t \in T \cup \{u\}} B_t \right)^q \right) \right)^{1-\frac{1}{q}}.
\]

(172)

Hence, by combining (171) and (172) with (170), we see that

\[
\sum_{S \subseteq \{1, \ldots, n\}} \sum_{j \in S} \sum_{k \leq |S|} \sum_{t \in T} \text{Tr}\left( \left( \sum_{s \in S \setminus \{j\}} B_s \right)^{q-1} B_j \right)
\]

\[
\leq \frac{nW_1^q}{k(n - k + 1)} \sum_{T \subseteq \{1, \ldots, n\}} \sum_{u \in \{1, \ldots, n\} \setminus T} \left( \text{Tr}\left( \left( \sum_{t \in T \cup \{u\}} B_t \right)^q \right) \right)^{1-\frac{1}{q}}
\]

(173)

\[
= \frac{nW_1^q}{n - k + 1} \sum_{S \subseteq \{1, \ldots, n\}} \left( \text{Tr}\left( \left( \sum_{s \in S} B_s \right)^q \right) \right)^{1-\frac{1}{q}}.
\]

(174)

where for (174) note that for every \(S \subseteq \{1, \ldots, n\}\) with \(|S| = k\) the term corresponding to \(\sum_{s \in S} B_s\) occurs in the sum that appears in (173) with multiplicity \(k\), once for each \(u \in S\).

Recalling the definition of \(U\) in (166), by Jensen’s inequality we see that

\[
\frac{1}{(n \choose k)} \sum_{S \subseteq \{1, \ldots, n\} \atop |S| = k} \left( \text{Tr}\left( \left( \sum_{s \in S} B_s \right)^q \right) \right)^{1-\frac{1}{q}} \leq U^{1-\frac{1}{q}}.
\]

(175)

By substituting (175) into (174) and using \(k \leq n/2\), we have

\[
\frac{1}{(n \choose k)} \sum_{S \subseteq \{1, \ldots, n\} \atop |S| = k} \sum_{j \in S} \sum_{k \leq |S|} \text{Tr}\left( \left( \sum_{s \in S \setminus \{j\}} B_s \right)^{q-1} B_j \right) \leq 2W_1^q U^{\frac{q-1}{q}}.
\]

(176)

In conjunction with (176), it follows from (169) that

\[
U \leq \min \left\{ \frac{V}{\lambda^l} + \frac{2W_1^q U^{q-1}}{(1-\lambda)^r} : \lambda \in [0, 1] \right\} \leq \left( \frac{V}{\lambda^l} + 2 \frac{1}{\lambda^r} W_1^q U^{q-1} \right)^{r+1},
\]

(177)
where the final inequality in (177) is seen by choosing $1/\lambda = 1 + \left(2W^{1/2}U_{V}^{q-1}/V^{1/2}\right)^{1/1}$. By (177),

$$U_{V}^{1/1} \leq V_{V}^{1/1} + 2r_{V}^{1/1} W_{V}^{1/4r(V+1)} U_{V}^{q-1/4r(V+1)}.$$  \hspace{1cm} (178)

The desired inequality (167) is a formal consequence of (178), as follows. If $U \leq (4q/\log(2q))^{r+1} V$ then (167) holds true because $r+1 \leq q$. We may therefore assume that $U > (4q/\log(2q))^{r+1} V$, in which case (178) implies that

$$\frac{U_{V}}{(2q)^{q/2}} \leq \left(1 - \frac{\log(2q)}{4q}\right) U_{V}^{1/1} \leq 2r_{V}^{1/1} W_{V}^{1/4r(V+1)} U_{V}^{q-1/4r(V+1)},$$  \hspace{1cm} (179)

where we used the fact that $(1 - t) \geq e^{-2t}$ for every $t \in [0, 1/2]$. The estimate (179) simplifies to

$$U \leq 2^{q} (2q)^{r+1} W \leq 2^{q} (2q)^{2} W \leq \left(\frac{4q}{\log(2q)}\right)^{q} W,$$

where we used the elementary inequality $\log t \leq \sqrt{t}$, which holds true for every $t \in (0, \infty)$. \hfill $\Box$

**Acknowledgements.** We are very grateful to Eric Carlen and Oded Regev for helpful pointers to the literature related to Section 7.1 as well as showing us counter-examples to (164) when $q = 4$.

**References**


Mathematics Department, Princeton University, Fine Hall, Washington Road, Princeton, NJ 08544-1000, USA
E-mail address: naor@math.princeton.edu

Department of Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel
E-mail address: gideon@weizmann.ac.il