Forum of Mathematics, Pi (2016), Vol. 4, e3, 81 pages
doi:10.1017/fmp.2016.1

METRIC X_p INEQUALITIES

ASSAF NAOR[1] and GIDEON SCHECHTMAN[2]

[1] Mathematics Department, Princeton University, Fine Hall, Washington Road, Princeton, NJ 08544-1000, USA;
email: naor@math.princeton.edu

[2] Department of Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel;
email: gideon@weizmann.ac.il

Received 8 September 2014; accepted 30 December 2015

Abstract

For every $p \in (0,\infty)$ we associate to every metric space (X, d_X) a numerical invariant $\mathfrak{X}_p(X) \in [0,\infty]$ such that if $\mathfrak{X}_p(X) < \infty$ and a metric space (Y, d_Y) admits a bi-Lipschitz embedding into X then also $\mathfrak{X}_p(Y) < \infty$. We prove that if $p, q \in (2,\infty)$ satisfy $q < p$ then $\mathfrak{X}_p(L_p) < \infty$ yet $\mathfrak{X}_p(L_q) = \infty$. Thus, our new bi-Lipschitz invariant certifies that L_q does not admit a bi-Lipschitz embedding into L_p when $2 < q < p < \infty$. This completes the long-standing search for bi-Lipschitz invariants that serve as an obstruction to the embeddability of L_p spaces into each other, the previously understood cases of which were metric notions of type and cotype, which however fail to certify the nonembeddability of L_q into L_p when $2 < q < p < \infty$. Among the consequences of our results are new quantitative restrictions on the bi-Lipschitz embeddability into L_p of snowflakes of L_q and integer grids in ℓ_q^n, for $2 < q < p < \infty$. 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.

2010 Mathematics Subject Classification: 46B80, 46B85 (primary); 46B25, 47B10 (secondary)

1. Introduction

1.1. Nontechnical overview. As a special case of the main contribution of the present article, for $p \in (0,\infty)$ we associate to every metric space (X, d_X) a numerical invariant $\mathfrak{X}_p(X) \in [0,\infty]$; a precise description of this quantity appears in Definition 1.1 below. Given $p \in (0,\infty)$ and two metric spaces (X, d_X) and (Y, d_Y), any $f : X \to Y$ incurs distortion at least $\mathfrak{X}_p(X)/\mathfrak{X}_p(Y)$.

© The Author(s) 2016. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

Thus, $\mathfrak{X}_p(\cdot)$ is a bi-Lipschitz invariant. We shall prove that for $2 < q < p < \infty$ we have $\mathfrak{X}_p(L_p) \asymp p/\log p$, while $\mathfrak{X}_p(L_q) = \infty$. Consequently, L_q does not admit a bi-Lipschitz embedding into L_p.

Qualitatively, the above nonembedding conclusion is well known. Namely, the fact that L_q fails to admit a bi-Lipschitz embedding into L_p when $2 < q < p < \infty$ follows from a differentiation argument that allows one to reduce the question to the linear theory. Specifically, every Lipschitz mapping $f : L_q \to L_p$ must have [5, 25, 54] a point of Gâteaux differentiability $x_0 \in L_q$. The derivative $f'(x_0) : X \to Y$ is a bounded linear operator, and if f were bi-Lipschitz then it would follow that $f'(x_0)$ is invertible with a bounded inverse, and therefore L_q would be isomorphic to the linear subspace $f'(x_0)L_q$ of L_p. However, a classical theorem of Paley [76] asserts that L_q is not isomorphic to any subspace of L_p, so it follows that L_q also fails to admit a bi-Lipschitz embedding into L_p. The above reasoning is due to Mankiewicz [54, Theorem 4]; Section 1.2 below contains a more detailed description of the relevant background.

Such differentiation arguments rely on an existential statement (a point of differentiability must exist), followed by a limiting procedure (differentiation itself) that uses the linear structure. As such, they do not apply in many settings, examples of which include understanding the L_p distortion of certain (often discrete) subsets of L_q, as well as treating non-Lipschitz (for example Hölder) mappings, a setting in which the mapping may be nondifferentiable at every point. (By [58, Remark 5.10], there does exist a bi-Hölder embedding of L_q into L_p when $2 < q < p < \infty$. Hence, the pertinent question is to determine which Hölder exponents are possible here. The non-Lipschitz setting therefore exhibits phenomena that are truly nonlinear and cannot be explained by a direct reduction to the linear theory.) Crucially, such arguments also fail to give any indication as to how to devise an invariant of metric spaces that certifies that the geometry of certain subsets of L_q is incompatible with the geometry of any subset of L_p.

The search for such metric invariants has been an important theme in modern metric geometry, underpinned by a classical rigidity theorem of Ribe [82] that laid the groundwork for what is known today as the Ribe program; for more on this research program, see its original formulation by Bourgain [17] as well as the recent (though by now not quite up-to-date) surveys [44, Section 3], [10, 66]. It suffices to say here that Ribe's theorem indicates that certain types of linear properties of Banach spaces (including those properties that are used in some, but not all, of the known proofs that L_q is not isomorphic to any linear subspace of L_p when $2 < q < p < \infty$), may in fact be metric properties in disguise, that is, they could be reformulated without making any reference to the linear structure whatsoever, so as to make sense in any metric space and thus provide a dictionary that allows one to apply linear intuitions in purely metric contexts. This paradigm

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

is very powerful, leading to solutions of questions in a wide variety of areas, ranging from the nonlinear geometry of Banach spaces themselves, to settings that a priori have seemingly nothing to do with Banach spaces, such as group theory, harmonic analysis, probability and combinatorial optimization.

Among the first questions that one would ask about bi-Lipschitz embeddings is to characterize those $p, q \in [1, \infty)$ such that L_q fails to admit a bi-Lipschitz embedding into L_p. Not surprisingly, efforts to understand this question influenced some of the most important developments in the Ribe program. By a reduction to the linear theory through differentiation in a manner that is similar to what we described above, the qualitative answer here is known: L_q does not admit a bi-Lipschitz embedding into L_p if and only if $p, q \in [1, \infty)$ satisfy one of the following three conditions.

$$q < \min\{p, 2\} \quad \text{or} \quad q > \max\{p, 2\} \quad \text{or} \quad 2 < q < p < \infty. \tag{1}$$

The search for metric invariants that explain the first range in (1) was an important impetus in the development of the theory of type of metric spaces, with notable contributions by Enflo [29–31], Bourgain–Milman–Wolfson [19], Pisier [79] and Ball [9]; see also [27, 35, 39, 60, 67, 69–72, 74]. The search for metric invariants that explain the second range in (1) was an important impetus in the development of the theory of cotype of metric spaces; see the work of Mendel and Naor [61] as well as [9, 34, 63, 64]. The second range in (1) could also be explained through a metric invariant called Markov convexity; see [17, 47, 62]. Over the years, many applications of the above invariants (metric type, metric cotype, Markov convexity) to a wide range of areas were discovered; the above- mentioned references contain examples of such results, and a variety of additional examples appears in [4, 7, 8, 13, 22, 48, 51, 59, 65, 87, 87]. Despite these developments, the question of formulating a metric invariant that explains the third range in (1) remained unresolved for many years. Here we settle this remaining case by introducing an invariant of metric spaces that serves as an obstruction to the embeddability of L_q into L_p when $2 < q < p < \infty$, thus completing the repertoire of metric invariants that classify those $p, q \in [1, \infty)$ for which L_q admits a bi-Lipschitz embedding into L_p.

Our new metric invariant is described in the following definition, in which (and in what follows) for every $n \in \mathbb{N}$ we let $e_1, \ldots, e_n$ denote the standard basis of $\mathbb{R}^n$, and for $S \subseteq \{1, \ldots, n\}$ and $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n) \in \{-1, 1\}^n$ we denote $\varepsilon_S = \sum_{j \in S} \varepsilon_j e_j$.

DEFINITION 1.1 (X_p metric space). Let (X, d_X) be a metric space and $p \in (0, \infty)$. Say that (X, d_X) is an X_p metric space if there exists $\mathfrak{X} \in (0, \infty)$ such that for every $n \in \mathbb{N}$ and $k \in \{1, \ldots, n\}$ there exists $m \in \mathbb{N}$ such that every

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

mapping $f : \mathbb{Z}_{2m}^n \to X$ satisfies

$$\left(\frac{1}{\binom{n}{k}} \sum_{\substack{S \subseteq \{1,\ldots,n\} \\ |S|=k}} \mathbb{E}[d_X(f(x+m\varepsilon_S), f(x))^p]\right)^{1/p}$$
$$\leqslant \mathfrak{X} m\left(\frac{k}{n}\sum_{j=1}^{n} \mathbb{E}[d_X(f(x+e_j), f(x))^p]\right.$$
$$\left. + \left(\frac{k}{n}\right)^{p/2} \mathbb{E}[d_X(f(x+\varepsilon), f(x))^p]\right)^{1/p}, \tag{2}$$

where the expectations in (2) are with respect to $(x, \varepsilon) \in \mathbb{Z}_{2m}^n \times \{-1, 1\}^n$ chosen uniformly at random. The infimum over those $\mathfrak{X} \in (0, \infty)$ for which (2) holds true is denoted $\mathfrak{X}_p(X, d_X)$, or simply $\mathfrak{X}_p(X)$ if the metric is clear from the context.

Theorem 1.2 below establishes that L_p is an X_p metric space when $p \geqslant 2$. We shall also check that L_q is not an X_p metric space when $q \in (2, p)$. Since for a metric space (X, d_X) the property of being an X_p metric space is obviously inherited by all the metric spaces that admit a bi-Lipschitz embedding into X, we thus obtain a new proof of the fact that L_q fails to admit a bi-Lipschitz embedding into L_p when $2 < q < p < \infty$. We shall show that the metric X_p invariant yields results that were beyond the reach of previous methods. For example, we shall obtain the first nontrivial upper bound on those $\theta \in (0, 1]$ for which L_q admits a bi-θ-Hölder embedding into L_p.

The above overview covered the context of our results without going into various technicalities, and as such it did not provide an explanation of how we arrived at Definition 1.1. There are also technical subtleties that partially explain (in hindsight) why understanding the third range in (1) remained open for so much longer than the same question for the first two ranges in (1). These matters will be clarified in the remainder of this introduction starting from Section 1.2 below, where we shall also describe consequences of our work, including new results even within the linear theory, as well as intriguing open questions that it raises.

1.2. Detailed statements and technical background. The ensuing discussion uses standard notation and terminology from Banach space theory, as in [**50**]. In particular, for $p \in [1, \infty]$ and $n \in \mathbb{N}$, the space ℓ_p^n (respectively $\ell_p^n(\mathbb{C})$) denotes the vector space $\mathbb{R}^n$ (respectively $\mathbb{C}^n$), equipped with the standard ℓ_p norm. Our results apply equally well to any infinite-dimensional Lebesgue function space $L_p(\mu)$, but for concreteness we fix (as usual) the space L_p to be

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

equal to $L_p([0,1],\mathscr{L})$, where $\mathscr{L}$ is the Lebesgue measure. 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 [55, 75]. 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 exist $s \in (0,\infty)$, $D \in [1,\infty)$ and a mapping $f : X \to Y$ such that

$$\forall x, y \in X, \quad s d_X(x,y) \leqslant d_Y(f(x), f(y)) \leqslant D s d_X(x,y). \tag{3}$$

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 (3) holds true is called the distortion of f and is denoted $\mathbf{dist}(f)$. If no such D exists set $\mathbf{dist}(f) = \infty$. 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 $c_{L_p}(X, d_X) = c_p(X, d_X)$.

As we discussed in Section 1.1, 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 L_p. This is well understood via a reduction to the linear theory, from which we deduce that L_q admits a bi-Lipschitz embedding into L_p if and only if either $q = 2$ or $1 \leqslant p \leqslant q \leqslant 2$ (moreover, in these cases we have $c_p(L_q) = 1$). Indeed, by general principles (see [15, Ch. 7] and the references therein), relying mainly on differentiation theorems for Lipschitz mappings between Banach spaces (the case $p = 1$ being somewhat different from the reflexive range), it suffices to understand when L_q is isomorphic to a subspace of L_p, a question that is perhaps among the first issues that one would investigate when studying linear embeddings of Banach spaces. Chapter 12 of Banach's book [12] is devoted to this topic. Banach proved there that if L_q is isomorphic to a subspace of L_p then necessarily either $p \leqslant q \leqslant 2$ or $2 \leqslant q \leqslant p$, and that L_2 is isomorphic to a subspace of L_p for all $p \in [1,\infty)$. Banach also conjectured [12, 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 [42], 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 [76], that is, L_q is not isomorphic to a subspace of L_p when $2 < q < p$.

As we explained above, our goal here is to obtain a nonlinear version of Paley's theorem, that is, the formulation of a bi-Lipschitz invariant that serves

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

as an obstruction to the embeddability of 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, which, using the notation and terminology of Definition 1.1, asserts that if $p \in (2, \infty)$ then L_p is an X_p metric space, with $\mathfrak{X}_p(L_p) \lesssim p/\log p$.

THEOREM 1.2 (Metric X_p inequality). *Fix $p \in [2, \infty)$. Suppose that $m, n \in \mathbb{N}$ and $k \in \{1, \ldots, n\}$ satisfy*

$$m \geqslant \frac{n^{3/2} \log p}{\sqrt{k}} + pn.$$

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

$$\begin{aligned} &\frac{(p/\log p)^{-p}}{\binom{n}{k}} \sum_{\substack{S \subseteq \{1,\ldots,n\} \\ |S|=k}} \frac{\mathbb{E}[\|f(x + 2m\varepsilon_S) - f(x)\|_p^p]}{m^p} \\ &\lesssim_p \frac{k}{n} \sum_{j=1}^{n} \mathbb{E}[\|f(x + e_j) - f(x)\|_p^p] \\ &\quad + \left(\frac{k}{n}\right)^{p/2} \mathbb{E}[\|f(x + \varepsilon) - f(x)\|_p^p], \end{aligned} \tag{4}$$

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 1.2, 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 $a \leqslant c^p b$ (respectively $a \geqslant c^p b$) for some universal constant $c \in (0, \infty)$. The notation $a \lesssim b$ (respectively $a \gtrsim b$) stands for $a \leqslant cb$ (respectively $a \geqslant cb$) for some universal constant $c \in (0, \infty)$. The notation $a \asymp b$ stands for $(a \lesssim b) \wedge (b \lesssim a)$. 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 $a \leqslant c(X)b$, where $c(X) \in (0, \infty)$ is allowed

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

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 (4) is sharp up to a universal constant factor. We defer the explanation of why (4) is called a metric X_p inequality to the ensuing discussion. Note that since (4) involves the p'th power of L_p norms, it suffices to prove its validity when f is real-valued, but we stated Theorem 1.2 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 1.2 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 (4) that we can imagine. However, it is straightforward to modify our proof of Theorem 1.2 so as to obtain variants of (4) for functions defined on discrete tori whose modulus is not necessarily divisible by 4.

REMARK 1.3. If one makes the weaker assumption $m \geqslant n^{3/2}/\sqrt{k}$ in Theorem 1.2 then (4) 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 1.2.

1.3. 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 θ-snowflake of (X, d_X) is defined (see for example [**26**]) to be the metric space (X, d_X^θ). A natural quantitative refinement of the assertion that L_q does not admit a bi-Lipschitz embedding into L_p is that if the θ-snowflake of L_q admits a bi-Lipschitz embedding into L_p then necessarily θ 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 $2 < q < p$, in Theorem 1.7 below we obtain, as a consequence of Theorem 1.2, an explicit $\delta(p,q) \in (0,1)$ such that if the θ-snowflake of L_q admits a bi-Lipschitz embedding into L_p then necessarily $\theta \leqslant 1-\delta(p,q)$. In Section 6, we formulate a conjectural convolution inequality that is shown to yield the sharp value $\delta(p,q)$ in this context. Since Hölder mappings need not be differentiable anywhere, and moreover, continuous

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

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 ℓ_q^n. While such an estimate can be obtained from general principles, namely Bourgain's discretization theorem [**18**, **36**] (see Remark 3.2 below), in Theorem 1.11 we obtain, as a consequence of Theorem 1.2, 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 1.14 below.

1.4. 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 ℓ_q^n into L_p tends to ∞ with n. Thus, there is a finite-dimensional linear obstruction (which will be stated explicitly in Section 1.5 below) showing that no n-dimensional subspace of L_p can be close to ℓ_q^n. As we discussed in Section 1.1, an important rigidity theorem of Ribe [**82**] 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 [**10**, **17**, **66**], 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 X_p metric spaces in their own right. However, unlike previous advances in the Ribe program, in the present setting it seems 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 1.7 below.

1.5. 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

$$\left(\mathbb{E}\left[\left\|\sum_{j=1}^n \varepsilon_j x_j\right\|_X^r\right]\right)^{1/r} \lesssim_X \left(\sum_{j=1}^n \|x_j\|_X^r\right)^{1/r} \quad \text{and}$$
$$\left(\sum_{j=1}^n \|x_j\|_X^s\right)^{1/s} \lesssim_X \left(\mathbb{E}\left[\left\|\sum_{j=1}^n \varepsilon_j x_j\right\|_X^s\right]\right)^{1/s}, \tag{5}$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

where the expectation is with respect to $\varepsilon \in \{-1,1\}^n$ chosen uniformly at random. The infimum over the implicit constants for which (5) holds true are denoted $T_r(X)$ and $C_s(X)$, respectively. See [56] and the references therein for more on these important notions. It suffices to say here that if $p \in [1,\infty)$ then L_p has type $\min\{p,2\}$ and cotype $\max\{q,2\}$, from which one deduces that there exists $\kappa(p) \in (0,\infty)$ such that if $T : \ell_q^n \to L_p$ is an invertible linear operator then necessarily

$$\mathbf{dist}(T) = \|T\| \cdot \|T^{-1}\| \geqslant \kappa(p) \cdot \begin{cases} n^{1/q-1/p} & \text{if } 1 \leqslant q \leqslant p \leqslant 2, \\ n^{1/q-1/2} & \text{if } 1 \leqslant q \leqslant 2 \leqslant p < \infty, \\ n^{1/p-1/q} & \text{if } 2 \leqslant p \leqslant q, \\ n^{1/2-1/q} & \text{if } 1 \leqslant p \leqslant 2 \leqslant q. \end{cases} \tag{6}$$

(6) follows from an application of (5) with $X = L_p$, $r = \min\{p,2\}$, $s = \max\{p,2\}$ and $x_j = Te_j$. The bounds in (6) 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 1.4, that is, they certify (in a sharp way) that if either $q \in [1,2)$ and $p > q$ or $q \in (2,\infty)$ and $q > p$, then any linear embedding of ℓ_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 (5), 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 [31], a metric space (X,d_X) is said to have Enflo type $r \in [1,\infty)$ if for every $n \in \mathbb{N}$ and $f : \{-1,1\}^n \to X$,

$$\mathbb{E}[d_X(f(\varepsilon), f(-\varepsilon))^r] \lesssim_X \sum_{j=1}^n \mathbb{E}[d_X(f(\varepsilon), f(\varepsilon_1,\ldots,\varepsilon_{j-1},-\varepsilon_j,\varepsilon_{j+1},\ldots,\varepsilon_n))^r], \tag{7}$$

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 (7) coincides with the leftmost inequality in (5) 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 $r = \min\{p,2\}$, that is, $X = L_p$ satisfies (7) 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 [29] and for $p \in (2,\infty)$ in [72].

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

One is tempted to define when a metric space (X, d_X) has cotype $s \in (0,\infty)$ by reversing the inequality in (7) (with r replaced by s). But, note that if $d_X(f(\varepsilon), f(\delta)) = 1$ for every distinct $\varepsilon, \delta \in \{-1,1\}^n$ (this can occur even if X is a Hilbert space), then the right-hand side of (7) grows linearly with n as $n \to \infty$, while the left-hand side of (7) 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 (7). In essence, the total mass of the measure that appears in the right-hand side of (7) is too large in comparison to the total mass of the measure that appears in the left-hand side of (7) for an inequality that is the reverse of (7) to make any sense even in Hilbert space (it actually fails in any nonsingleton metric space; see [**61**]).

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 [**61**] 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

$$\forall\, f : \mathbb{Z}_{2m}^n \to X, \quad \sum_{j=1}^n \frac{\mathbb{E}[d_X(f(x+me_j), f(x))^s]}{m^s} \lesssim_X \mathbb{E}[d_X(f(x+\varepsilon), f(x))^s], \tag{8}$$

where the expectation is with respect to $(x,\varepsilon) \in \mathbb{Z}_{2m}^n \times \{-1,0,1\}^n$ chosen uniformly at random. It was proved in [**61**] 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 (8) we consider displacements of the argument of f by a multiple of m, that is, 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 (4), but we shall see below that the need for scaling in the context of Theorem 1.2 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 (4) to see that it does not cause the problem that we presented above).

1.5.1. The case $2 < q < p$. While Paley's work [**76**] 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 [**43**] introduced an influential way to solve this problem through a structural study of basic sequences in L_p spaces. In particular,

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

it follows from [43] that for $p \in (2, \infty)$, any infinite symmetric basic sequence in L_p is equivalent to either the standard basis of ℓ_p or the standard basis of ℓ_2. Consequently, for $q \in (2, p)$ there does not exist a symmetric basic sequence in L_p that is equivalent to the unit basis of ℓ_q, and therefore ℓ_q cannot be isomorphic to a subspace of ℓ_p. In 1979, Johnson *et al.* [40] 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 [40] is devoted. In particular, in [40] 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 [40] 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 $(x_1, \dots, x_n) \in X^n$ is said to be K-symmetric if for every sequence of scalars $a_1, \dots, a_n \in \mathbb{R}$, every permutation $\pi \in S_n$ and every sequence of signs $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n) \in \{-1, 1\}^n$, we have

$$\|\varepsilon_1 a_{\pi(1)} x_1 + \cdots + \varepsilon_n a_{\pi(n)} x_n\|_X \leqslant K \|a_1 x_1 + \cdots + a_n x_n\|_X.$$

The sequence $(x_1, \dots, x_n) \in X^n$ is said to be normalized if $\|x_j\|_X = 1$ for all $j \in \{1, \dots, n\}$. Given two Banach spaces $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$, two sequences $(x_1, \dots, x_n) \in X^n$ and $(y_1, \dots, y_n) \in Y^n$ are said to be K-equivalent if there exists $s \in (0, \infty)$ such that

$$s\|a_1 x_1 + \cdots + a_n x_n\|_X \leqslant \|a_1 y_1 + \cdots + a_n y_n\|_Y \leqslant Ks\|a_1 x_1 + \cdots + a_n x_n\|_X$$

for all choices of scalars $a_1, \dots, a_n \in \mathbb{R}$.

Fixing $p \in (2, \infty)$, examples of symmetric sequences in L_p are furnished by Rosenthal's $X_p^n(\omega)$ symmetric bases [83], as ω ranges over $[0, \infty)$. The definition of these bases is very simple: let $u_1, \dots, u_n$ be an orthonormal basis of ℓ_2^n and define $\{x_j(p, \omega)\}_{j=1}^n \subseteq (\ell_p^n \oplus \ell_2^n)_p$ by

$$x_j(p, \omega) \stackrel{\text{def}}{=} \frac{1}{(1 + \omega^p)^{1/p}} \cdot e_j + \frac{\omega}{(1 + \omega^p)^{1/p}} \cdot u_j. \tag{9}$$

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 ℓ_2 is isometric to a subset of L_p (see for example [90]), the sequence $\{x_j(p, \omega)\}_{j=1}^n$ can be realized as elements of L_p.

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

In [40], 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 ℓ_q^n into L_p. Indeed, if $T : \ell_q^n \to L_p$ is injective and linear then $(Te_1,\ldots,Te_n)$ is a $\mathbf{dist}(T)$-symmetric sequence in L_p, and is therefore $D(p,\mathbf{dist}(T))$-equivalent to an $X_p^n(\omega)$ basis for some $\omega \in (0,\infty)$. Direct inspection now reveals that this is only possible if $\mathbf{dist}(T)$ tends to ∞ as $n \to \infty$. In fact, by computing the various bounds explicitly and optimizing over $\omega \in [0,\infty)$, as done in [33] (relying in part on a computation from [37]), 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

$$\mathbf{dist}(T) \geqslant \sigma(p,q) \cdot n^{(p-q)(q-2)/(q^2(p-2))}. \tag{10}$$

The lower bound in (10) is asymptotically sharp (up to the implicit dependence on p,q), as exhibited by the embedding $J^R_{(q\to p;n)} : \ell_q^n \to (\ell_p^n \oplus \ell_2^n)_p \subseteq L_p$ given by

$$\forall\, j \in \{1,\ldots,n\}, \qquad J^R_{(q\to p;n)}(e_j) \stackrel{\mathrm{def}}{=} n^{1/2}\cdot e_j + n^{1/q}\cdot u_j, \tag{11}$$

where, as in (9), $u_1,\ldots,u_n$ is an orthonormal basis of ℓ_2^n. (The superscript in the notation $J^R_{(q\to p;n)}(\cdot)$ refers to Rosenthal.) Indeed, by a straightforward Lagrange multiplier argument (see Section 2 below), for every $2 < q \leqslant p$ we have

$$\mathbf{dist}(J^R_{(q\to p;n)}) \asymp n^{(p-q)(q-2)/(q^2(p-2))}. \tag{12}$$

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 [40]. 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 $t_1,\ldots,t_n \in \mathbb{R}$,

$$\left(\frac{\log p}{p}\right)^p \cdot \mathbb{E}\left[\left|\sum_{j=1}^n t_j Y_j\right|^p\right] \lesssim_p \sum_{j=1}^n |t_j|^p + \left(\frac{1}{n}\sum_{j=1}^n t_j^2\right)^{p/2} \mathbb{E}\left[\left(\sum_{j=1}^n Y_j^2\right)^{p/2}\right]. \tag{13}$$

The term $(\log p)/p$ in the left-hand side of (13) is sharp up to a universal constant factor: in this sharp form the inequality (13) is due to [41]. Without a sharp dependence on p, inequality (13) was first proved in [40]. The proof of (13) with sharp dependence on p is significantly more involved than the proof in [40]. The dependence on p is not of major importance for us here, but it is worthwhile to state the above sharp form of (13) since it is available in the literature.

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

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

$$Y_j(\varepsilon, \pi) \stackrel{\text{def}}{=} \frac{\varepsilon_j a_{\pi(j)}}{((1/n)\sum_{s=1}^n |a_s|^p)^{1/p}}.$$

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 $\mathbb{E}[|Y_j|^p] = 1$. For $k \in \{1, \ldots, n\}$, an application of (13) with $t_1 = \cdots = t_k = 1$ and $t_{k+1} = \cdots = t_n = 0$ therefore yields the following inequality.

$$\frac{(p/\log p)^{-p}}{\binom{n}{k}} \sum_{\substack{S \subseteq \{1,\ldots,n\} \\ |S|=k}} \mathbb{E}\left[\left|\sum_{j \in S} \varepsilon_j a_j\right|^p\right] \lesssim_p \frac{k}{n}\sum_{j=1}^n |a_j|^p + \left(\frac{k}{n}\right)^{p/2}\left(\sum_{j=1}^n a_j^2\right)^{p/2}, \tag{14}$$

where in (14), as well as in (15)–(18) below, the expectation is with respect to $\varepsilon \in \{-1, 1\}^n$ chosen uniformly at random. Since, by Jensen's inequality,

$$\left(\sum_{j=1}^n a_j^2\right)^{p/2} = \left(\mathbb{E}\left[\left|\sum_{j=1}^n \varepsilon_j a_j\right|^2\right]\right)^{p/2} \leqslant \mathbb{E}\left[\left|\sum_{j=1}^n \varepsilon_j a_j\right|^p\right], \tag{15}$$

it follows from (14) that

$$\frac{(p/\log p)^{-p}}{\binom{n}{k}} \sum_{\substack{S \subseteq \{1,\ldots,n\} \\ |S|=k}} \mathbb{E}\left[\left|\sum_{j \in S} \varepsilon_j a_j\right|^p\right] \lesssim_p \frac{k}{n}\sum_{j=1}^n |a_j|^p + \left(\frac{k}{n}\right)^{p/2} \mathbb{E}\left[\left|\sum_{j=1}^n \varepsilon_j a_j\right|^p\right]. \tag{16}$$

An inspection of the argument in [41] reveals that the term $p/\log p$ in (16) is sharp up to a constant factor even in this special case of (13) (this is true if one requires the validity of (16) for all $k \in \{1, \ldots, n\}$, while for a fixed k there might be a better dependence as a function of k, n, p).

Our main result, namely Theorem 1.2, is a nonlinear version of (16). By following the reasoning that led to the definition (7) of Enflo type, one is tempted to try to establish the validity of the following inequality, which should hold true for every $f : \{-1, 1\}^n \to \mathbb{R}$ and for some $\alpha(p) \in (0, \infty)$.

$$\begin{aligned}\frac{\alpha(p)}{\binom{n}{k}} &\sum_{\substack{S \subseteq \{1,\ldots,n\} \\ |S|=k}} \mathbb{E}[|f(\varepsilon) - f(\varepsilon_{\{1,\ldots,n\}\setminus S} - \varepsilon_S)|^p] \\ &\leqslant \frac{k}{n}\sum_{j=1}^n \mathbb{E}[|f(\varepsilon) - f(\varepsilon_1, \ldots, \varepsilon_{j-1}, -\varepsilon_j, \varepsilon_{j+1}, \ldots, \varepsilon_n)|^p] \\ &\quad + \left(\frac{k}{n}\right)^{p/2} \mathbb{E}[|f(\varepsilon) - f(-\varepsilon)|^p]. \end{aligned} \tag{17}$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

Inequality (17) holds true when $p = 2$. Indeed, the fact that the real line has Enflo type 2 with constant 1 (as shown by Enflo in [29]) implies that for every $S \subseteq \{1, \ldots, n\}$ we have

$$\mathbb{E}[|f(\varepsilon) - f(\varepsilon_{\{1,\ldots,n\}\setminus S} - \varepsilon_S)|^2] \leqslant \sum_{j\in S} \mathbb{E}[|f(\varepsilon) - f(\varepsilon_1, \ldots, \varepsilon_{j-1}, -\varepsilon_j, \varepsilon_{j+1}, \ldots, \varepsilon_n)|^2]. \tag{18}$$

By averaging (18) over all of those $S \subseteq \{1, \ldots, n\}$ satisfying $|S| = k$ we see that (17) holds true when $p = 2$, with $\alpha(2) = 1$ and even without the final term in the right-hand side of (17).

The validity of (17) for $p = 2$ indicates that the reason why scaling is needed for the definition (8) of metric cotype does not arise in the context of (17). However, Proposition 1.4 below shows that scaling is nevertheless necessary in the context of metric X_p inequalities, thus explaining our formulation of Theorem 1.2. Note that the conclusion of Theorem 1.2 implies the linear X_p inequality (16). Roughly speaking, this follows by applying (4) to the linear function $f : \mathbb{Z}_{4m}^n \to \mathbb{R}$ given by $f(x) = \sum_{j=1}^n x_j a_j$. However, this reasoning is not quite accurate because this f is not well defined as a function on the discrete torus $\mathbb{Z}_{4m}^n$; for a precise argument, see Proposition 2.1 below.

PROPOSITION 1.4 (Scaling is necessary). *Fix $p \in (2, \infty)$, $\alpha \in (0, 1)$, $m, n \in \mathbb{N}$ and $k \in \{1, \ldots, n\}$. Suppose that for every $f : \mathbb{Z}_{2m}^n \to \mathbb{R}$ we have*

$$\frac{\alpha^p}{\binom{n}{k}} \sum_{\substack{S\subseteq\{1,\ldots,n\}\\ |S|=k}} \frac{\mathbb{E}[|f(x + m\varepsilon_S) - f(x)|^p]}{m^p} \leqslant \frac{k}{n} \sum_{j=1}^n \mathbb{E}[|f(x + e_j) - f(x)|^p] + \left(\frac{k}{n}\right)^{p/2} \mathbb{E}[|f(x + \varepsilon) - f(x)|^p], \tag{19}$$

where the expectation is with respect to $(x, \varepsilon) \in \mathbb{Z}_{2m}^n \times \{-1, 1\}^n$ chosen uniformly at random. Then

$$k \geqslant \left(\frac{5}{\alpha}\right)^{2p/(p-2)} \implies m \geqslant \frac{\alpha}{3}\sqrt{\frac{n}{k}}. \tag{20}$$

The proof of Proposition 1.4 appears in Section 2. We conjecture that the dependence of m on n and k that appears in Proposition 1.4 is sharp, up to the (possibly p-dependent) constant. This is the content of Conjecture 1.5

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

below. It seems that in order to prove Conjecture 1.5 one would need to exploit cancellations that are more subtle than those that we used to prove Theorem 1.2.

CONJECTURE 1.5. *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 $m \geqslant C_p\sqrt{n/k}$ then for every $f : \mathbb{Z}_{4m}^n \to \mathbb{R}$ we have*

$$\frac{\alpha_p}{\binom{n}{k}} \sum_{\substack{S\subseteq\{1,\ldots,n\}\\ |S|=k}} \frac{\mathbb{E}[|f(x+2m\varepsilon_S)-f(x)|^p]}{m^p} \lesssim_p \frac{k}{n}\sum_{j=1}^n \mathbb{E}[|f(x+e_j)-f(x)|^p] + \left(\frac{k}{n}\right)^{p/2} \mathbb{E}[|f(x+\varepsilon)-f(x)|^p], \quad (21)$$

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

We will see in Section 1.6.3 below that, in addition to its intrinsic interest, a positive resolution of Conjecture 1.5 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 1.5.

Before passing to a description of the geometric consequences of Theorem 1.2, we note that the linear X_p inequality (16) also has a (much easier) converse [**40**]. Specifically, for every $p \in (2,\infty)$ there exists $K(p) \in (0,\infty)$ such that for every $a_1,\ldots,a_n \in \mathbb{R}$ and $k \in \{1,\ldots,n\}$ we have

$$\frac{k}{n}\sum_{j=1}^n |a_j|^p + \left(\frac{k}{n}\right)^{p/2} \mathbb{E}\left[\left|\sum_{j=1}^n \varepsilon_j a_j\right|^p\right] \leqslant \frac{K(p)^p}{\binom{n}{k}} \sum_{\substack{S\subseteq\{1,\ldots,n\}\\ |S|=k}} \mathbb{E}\left[\left|\sum_{j\in S}\varepsilon_j a_j\right|^p\right], \quad (22)$$

where the expectation is over $\varepsilon \in \{-1,1\}^n$ chosen uniformly at random. An inspection of the proof of (22) in [**40**] (or in [**41**]) reveals that one can take $K(p) \lesssim \sqrt{p}$ in (22). Theorem 1.6 below is a nonlinear version of (22). Although we do not have a new geometric application of the reverse metric X_p inequality that appears in Theorem 1.6, 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 1.6 yields some new information on metric cotype; see Theorem 5.2 below and the discussion immediately preceding it.

THEOREM 1.6 (Reverse metric X_p inequality). *Fix $p \in [2,\infty)$ and $k,m \in \mathbb{N}$ with $m \geqslant k^{1/p}/\sqrt{p}$. Fix also an integer $n \geqslant k$. Then for every $f : \mathbb{Z}_{8m}^n \to L_p$*

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

we have

$$\begin{aligned}\frac{k}{n}\sum_{j=1}^{n}\frac{\mathbb{E}[\|f(x+4me_j)-f(x)\|_p^p]}{m^p}&\\ +\left(\frac{k}{n}\right)^{p/2}\mathbb{E}[\|f(x+\varepsilon)-f(x-\varepsilon)\|_p^p]&\\ \lesssim_p \frac{p^{p/2}}{\binom{n}{k}}\sum_{\substack{S\subseteq\{1,\ldots,n\}\\ |S|=k}}\mathbb{E}[\|f(x+\varepsilon_S)-f(x)\|_p^p],&\end{aligned}\tag{23}$$

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

1.6. Metric X_p inequalities as obstructions to embeddings. Theorem 1.2 yields a bi-Lipschitz invariant that can be used to obtain new nonembeddability results which we shall now describe.

1.6.1. Snowflakes. Fix $p,q\in[1,\infty)$. Sharp restrictions on those $\theta\in(0,1]$ for which the θ-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 $p\geqslant q$, or $q\in[2,\infty)$ and $p\leqslant q$; see [**34**, **52**, **59**]. Here we obtain, as a consequence of Theorem 1.2, the first such result when $2<q<p$.

THEOREM 1.7 (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 $\theta\leqslant 1-\delta(p,q)$. Specifically, θ must satisfy*

$$\begin{aligned}\theta&\leqslant\frac{2q(p-q)+q^2(p-1)(p-2)}{2p^2(q-2)}\left(\sqrt{1+\frac{4p(p-2)(q-2)}{(pq-3q+2)^2}}-1\right)\\ &\leqslant 1-\frac{(p-q)(q-2)}{2p^3}.\end{aligned}\tag{24}$$

It was shown in [**58**, Remark 5.10] that for $2<q<p$ that the (q/p)-snowflake of L_q is isometric to a subset of L_p. We conjecture that this is sharp, that is, that the upper bound on θ that appears in (24) can be improved to $\theta\leqslant q/p$.

CONJECTURE 1.8. *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 $\theta\leqslant q/p$.*

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

In fact, when $2 < q \leqslant p$, 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 $1 \leqslant q \leqslant p \leqslant 2$, where it is known [21] (see also [89]) that the metric space $(L_q, \|x-y\|_q^\theta)$ admits an isometric embedding into L_p for every $0 < \theta \leqslant q/p$.

QUESTION 1.9 (Uniqueness of snowflakes). Suppose that $2 < q \leqslant p$ 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 1.9 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 [58, Section 5].

REMARK 1.10. The analog of Conjecture 1.8 for sequence spaces has a positive answer. Indeed, a combination of [14, Corollary 2.19] and [14, Corollary 2.23] shows that for every $1 \leqslant q \leqslant p < \infty$, if $\theta \in (0,1]$ is such that the metric space $(\ell_q, \|x-y\|_q^\theta)$ admits a bi-Lipschitz embedding into ℓ_p then necessarily $\theta \leqslant q/p$. The proof of this result in [14] relies on an infinite-dimensional argument of [45] that is specific to sequence spaces (the above statement from [14] becomes false if $q \in [1,2]$, $p \in (2,\infty)$ and ℓ_p is replaced by L_p). Conversely, in [2] (see also [75, Exercise 1.61]) it was shown that for every $1 \leqslant q \leqslant p < \infty$ the (q/p)-snowflake of ℓ_q does admit a bi-Lipschitz embedding into ℓ_p.

1.6.2. Grids. Recall that for $q \in [1,\infty)$ and $m,n \in \mathbb{N}$ the grid $\{1,\ldots,m\}^n$, equipped with the metric inherited from ℓ_q^n, is denoted $[m]_q^n$. Theorem 1.11 below, which is a consequence of Theorem 1.2, 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 1.11 (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*

$$c_p([m]_q^n) \gtrsim \alpha_p \big(\min\{m^{(q(p-2))/(q(p-2)+p-q)}, n\}\big)^{(p-q)(q-2)/(q^2(p-2))}. \tag{25}$$

In particular,

$$m \geqslant n^{1+(p-q)/(q(p-2))} \implies c_p([m]_q^n) \geqslant \alpha_p n^{(p-q)(q-2)/(q^2(p-2))} \gtrsim \alpha_p c_p(\ell_q^n). \tag{26}$$

The fact that the lower bound in (25) becomes weaker for smaller m is necessary, as exhibited by the following embedding from [59]. First, let

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

$G_{2,p} : L_2 \to L_p$ be an isometric embedding of L_2 into L_p. By a classical theorem of Schoenberg [84] (see also [89]), there exists an isometric embedding of the $(2/q)$-snowflake of ℓ_2^n into L_2, that is, there exists $\psi_q^n : \ell_2^n \to L_2$ such that

$$\forall\, x, y \in \ell_2^n, \quad \|\psi_q^n(x) - \psi_q^n(y)\|_2 = \|x - y\|_2^{2/q}.$$

Finally, let $I_{q\to 2}^n : \ell_q^n \to \ell_2^n$ be the identity mapping, and define (the superscript in the notation $J_{(q\to p;n)}^S(\cdot)$ refers to Schoenberg)

$$J_{(q\to p;n)}^S \stackrel{\text{def}}{=} G_{2,p} \circ \psi_q^n \circ I_{q\to 2}^n : \ell_q^n \to L_p. \tag{27}$$

As argued in [59], the distortion of the restriction of $J_{(q\to p;n)}^S$ to $[m]_q^n$ satisfies

$$\mathbf{dist}(J_{(q\to p;n)}^S|_{[m]_q^n}) \leqslant m^{1-2/q}.$$

Recalling the definition of the embedding $J_{(q\to p;n)}^R$ in (11), we therefore have

$$\begin{aligned} c_p([m]_q^n) &\leqslant \min\{\mathbf{dist}(J_{(q\to p;n)}^R|_{[m]_q^n}), \mathbf{dist}(J_{(q\to p;n)}^S|_{[m]_q^n})\} \\ &\lesssim \min\{n^{(p-q)(q-2)/(q^2(p-2))}, m^{1-2/q}\}. \end{aligned} \tag{28}$$

We conjecture that (28) is asymptotically sharp up to constant factors that depend only on p, q.

CONJECTURE 1.12. *For $2 < q < p$ and $m, n \in \mathbb{N}$, the better of the embeddings $J_{(q\to p;n)}^R$ and $J_{(q\to p;n)}^S$ appearing in (11) and (27), respectively, is the best possible bi-Lipschitz embedding of the L_q integer grid $[m]_q^n$ into L_p. Equivalently, $c_p([m]_q^n)$ is bounded from above and from below by positive constants that may depend only on p and q times the quantity*

$$\min\{n^{(p-q)(q-2)/(q^2(p-2))}, m^{1-2/q}\}. \tag{29}$$

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

$$m \geqslant n^{(p-q)/q(p-2)} \implies c_p([m]_q^n) \geqslant \eta(p,q) c_p(\ell_q^n),$$

yet

$$m = o(n^{(p-q)/q(p-2)}) \implies c_p([m]_q^n) = o(c_p(\ell_q^n)) \quad (\text{as } n \to \infty).$$

An affirmative answer to Conjecture 1.12 would imply that if the linear embedding $J_{(q\to p;n)}^R$ of ℓ_q^n into an appropriate Rosenthal $X_p(\omega)$ space fails to yield the best possible bi-Lipschitz embedding of $[m]_q^n$ into L_p (up to constant factors that are independent of m, n), then the best possible way to embed $[m]_q^n$ into L_p would be to embed it into L_2 (ignoring the fact that we are seeking

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

an embedding into the larger space L_p), via the (highly nonlinear) Schoenberg embedding ψ_q^n. Admittedly, if true, this phenomenon would be quite exotic, but we conjecture that it indeed occurs partially because it is a consequence of Conjecture 1.5, as we shall see in Section 1.6.3 below.

REMARK 1.13. There are also interesting open problems related to embeddings of $[m]_p^n$ into L_q when $p > q > 2$. Specifically, by combining the upper bound in [59] with the metric cotype-based lower bound in [61], we see that

$$\frac{1}{\sqrt{q}} \cdot \min\{n^{1/q-1/p}, m^{1-q/p}\} \lesssim c_q([m]_p^n) \leqslant \min\{n^{1/q-1/p}, m^{1-2/p}\}. \tag{30}$$

The bounds in (30) match only when $q = 2$, and it remains open to evaluate $c_q([m]_p^n)$ up to constant factors that are independent of m, n. An inspection of the argument in [59] reveals that the lower bound on $c_q([m]_p^n)$ in (30) 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 1.9 (see also the paragraph immediately following Question 1.9) would imply that no nontrivial snowflake of L_q admits a bi-Lipschitz embedding into L_q. In the spirit of Conjecture 1.12, one is tempted to ask whether or not the upper bound on $c_q([m]_p^n)$ in (30) is asymptotically sharp, that is, if also in this setting it is best to embed $[m]_p^n$ 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 $c_q([m]_p^n)$ 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 $c_q([m]_p^n)$ when $p > q > 2$.

1.6.3. Consequences of Conjecture 1.5. The following theorem asserts that Conjecture 1.5 implies a positive solution of Conjecture 1.8 and Conjecture 1.12. Thus, obtaining the conjecturally sharp value of m in the metric X_p inequality of Theorem 1.2, 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 1.5, and hence also to Conjectures 1.8 and 1.12.

THEOREM 1.14. *If Conjecture 1.5 holds true then for every $2 < q < p$ and $\theta \in (0, 1)$,*

$$c_p(L_q, \|x-y\|_q^\theta) < \infty \implies \theta \leqslant \frac{q}{p}. \tag{31}$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

Moreover, for every $m, n \in \mathbb{N}$ the L_p distortion of the L_q grid $[m]_q^n$ is bounded from above and below by a constant that may depend only on p times the quantity appearing in (29).

1.7. X_p metric spaces? For $p \in (0, \infty)$, by pursuing the Ribe program in light of Theorem 1.2, one arrives at Definition 1.1 of when a metric space (X, d_X) is an X_p metric space. 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 has not 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

$$\frac{1}{\binom{n}{k}} \sum_{\substack{S \subseteq \{1,\ldots,n\} \\ |S|=k}} \mathbb{E}\left[\left\|\sum_{j\in S} \varepsilon_j v_j\right\|_X^p\right] \lesssim_X \frac{k}{n}\sum_{j=1}^n \|v_j\|_X^p + \left(\frac{k}{n}\right)^{p/2} \mathbb{E}\left[\left\|\sum_{j=1}^n \varepsilon_j v_j\right\|_X^p\right],$$

where the expectation is over $\varepsilon \in \{-1, 1\}^n$ chosen uniformly at random. Being an X_p Banach space is clearly a local property. Our proof of Theorem 1.2 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 [57] (see also [78]). 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 ℓ_q^n to any subspace of S_p must tend to ∞ with n: this follows from the noncommutative Kadec–Pełczyński result in [85]; see also [81, Theorem 10.7]. 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 [85], 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.

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

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 1.2 and its consequences, so they could be skipped on first reading.

We shall start with the proof of Proposition 1.4, that is, that scaling is needed for the metric X_p inequality of Theorem 1.2 to hold true.

Proof of Proposition 1.4. We shall use here the notation that was introduced in the statement of Proposition 1.4. Since ℓ_2 embeds isometrically into L_p, by [**58**, Lemma 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

$$1 \leqslant \frac{\|F(x) - F(y)\|_p}{\min\{2\sqrt{k}, \sqrt{\sum_{j=1}^n |e^{\pi i (x_j - y_j)/m} - 1|^2}\}} \leqslant 2. \tag{32}$$

By integrating (19) we see that

$$\frac{\alpha^p}{\binom{n}{k}} \sum_{\substack{S \subseteq \{1,\ldots,n\} \\ |S|=k}} \frac{\mathbb{E}[\|F(x + m\varepsilon_S) - F(x)\|_p^p]}{m^p}$$
$$\leqslant \frac{k}{n} \sum_{j=1}^n \mathbb{E}[\|F(x + e_j) - F(x)\|_p^p] + \left(\frac{k}{n}\right)^{p/2} \mathbb{E}[\|F(x + \varepsilon) - F(x)\|_p^p]. \tag{33}$$

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

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

which yields the desired implication (20). □

Next, we shall check the validity of (12), that is, evaluate the distortion of the mapping $J^R_{(q \to p;n)}$ given in (11). This is a known (and easy) statement which is included here only because we could not locate a clean reference for it.

Proof of (12). The definition (11) implies that for every $x \in \ell_q^n$ we have

$$\|J^R_{(q \to p;n)}(x)\|^p_{(\ell_p^n \oplus \ell_2^n)_p} = n^{p/2} \|x\|_p^p + n^{p/q} \|x\|_2^p.$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

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

$$\frac{n^{p/2}\|x\|_q^p}{2^{p/q} n^{(p(p-q)(q-2))/(q^2(p-2))}} \leqslant n^{p/2}\|x\|_p^p + n^{p/q}\|x\|_2^p \leqslant 2n^{p/2}\|x\|_q^p. \tag{34}$$

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

Let $x \in \ell_q^n$ with $\|x\|_q = 1$ be such that $n^{p/2}\|x\|_p^p + n^{p/q}\|x\|_2^p$ 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 $x_{k+1} = \cdots = x_n = 0$ for some $k \in \{1, \ldots, n\}$. Hence, there exists (a Lagrange multiplier) $\lambda \in \mathbb{R}$ such that

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

For $s \in [0, \infty)$ write $\psi(s) \stackrel{\text{def}}{=} n^{p/2}s^{p-2} - \lambda s^{q-2} + n^{p/q}\|x\|_2^{p-2}$. 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 ψ starts at a positive value, decreases on $(0, s_0)$, and then increases to ∞. Consequently, there exist $a, b \in (0, \infty)$ such that $\psi(s) = 0 \implies s \in \{a, b\}$ 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 $x_j = a\mathbf{1}_S(j) + b\mathbf{1}_{\{1,\ldots,k\}\setminus S}(j)$ for all $j \in \{1, \ldots, k\}$. Since $\|x\|_q = 1$, we may assume without loss of generality that $a^q|S| \geqslant 1/2$, that is, that $a \geqslant 1/(2|S|)^{1/q}$. Consequently,

$$\begin{aligned} n^{p/2}\|x\|_p^p + n^{p/q}\|x\|_2^p &\geqslant n^{p/2}|S|a^p + n^{p/q}|S|^{p/2}a^p \\ &\geqslant \frac{n^{p/2}}{2^{p/q}|S|^{p/q-1}} + \frac{n^{p/q}|S|^{p/2-p/q}}{2^{p/q}} \\ &\geqslant 2^{-p/q} n^{p/2-(p(p-q)(q-2))/(q^2(p-2))}, \end{aligned}$$

where the last step follows by computing the minimum of the quantity $n^{p/2}/s^{p/q-1} + n^{p/q}s^{p/2-p/q}$ 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 $(a + bi)(u, v) = (au - bv, av + bu)$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

for every $u, v \in Z$ and $a, b \in \mathbb{R}$. The norm on $Z_p(\mathbb{C})$ is given by

$$\forall\,(u,v) \in Z \times Z, \qquad \|(u,v)\|_{Z_p(\mathbb{C})} \stackrel{\text{def}}{=} \left(\int_0^{2\pi} \|(\cos\theta)u - (\sin\theta)v\|_Z^p \, d\theta \right)^{1/p}. \tag{36}$$

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

$$\|(z,0)\|_{Z_p(\mathbb{C})}^p = \|z\|_Z^p \int_0^{2\pi} |\cos\theta|^p \, d\theta = \frac{4\sqrt{\pi}\,\Gamma(p/2+1/2)}{\Gamma(p/2+1)} \|z\|_Z^p. \tag{37}$$

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 ℓ_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 1.2 implies the linear X_p inequality (16). We shall show this in the context of general Banach spaces, that is, if a Banach space is an X_p metric space then it is also an X_p Banach space. The converse of this assertion, that is, that an X_p Banach space is also an X_p metric space, follows from the proof of Theorem 1.2 that can be found 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 $k \in \{1,\dots,n\}$. Suppose that for every $f : \mathbb{Z}_{2m}^n \to Z$ we have*

$$\begin{aligned} &\frac{\gamma}{2^n \binom{n}{k}} \sum_{\substack{S \subseteq \{1,\dots,n\} \\ |S|=k}} \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_{2m}^n} \frac{\|f(x + m\varepsilon_S) - f(x)\|_Z^p}{m^p} \\ &\quad \leqslant \frac{k}{n} \sum_{j=1}^n \sum_{x \in \mathbb{Z}_{2m}^n} \|f(x+e_j) - f(x)\|_Z^p \\ &\qquad + \frac{(k/n)^{p/2}}{2^n} \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_{2m}^n} \|f(x+\varepsilon) - f(x)\|_Z^p. \end{aligned} \tag{38}$$

Then for every $z_1, \dots, z_n \in Z$ we have

$$\frac{(2/\pi)^{2p}\gamma}{2^n \binom{n}{k}} \sum_{\substack{S \subseteq \{1,\dots,n\} \\ |S|=k}} \sum_{\varepsilon \in \{-1,1\}^n} \left\| \sum_{j \in S} \varepsilon_j z_j \right\|_Z^p \leqslant \frac{k}{n} \sum_{j=1}^n \|z_j\|_Z^p + \frac{(k/n)^{p/2}}{2^n} \sum_{\varepsilon \in \{-1,1\}^n} \left\| \sum_{j=1}^n \varepsilon_j z_j \right\|_Z^p. \tag{39}$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

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 $f_\delta : \mathbb{Z}_{2m}^n \to Z_p(\mathbb{C})$ given by

$$\forall\, x \in \mathbb{Z}_{2m}^n, \qquad f_\delta(x) \stackrel{\text{def}}{=} \sum_{j=1}^n \delta_j e^{\pi i x_j/m}(z_j, 0) \in Z \times Z.$$

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

$$\begin{aligned}
&\frac{2^p\gamma}{2^n\binom{n}{k}} \sum_{\substack{S\subseteq\{1,\ldots,n\}\\ |S|=k}} \sum_{\delta\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{2m}^n} \frac{\left\|\sum_{j\in S}\delta_j e^{\pi i x_j/m}(z_j,0)\right\|_{Z_p(\mathbb{C})}^p}{m^p} \\
&\leqslant \frac{k(2m)^n}{n}\sum_{j=1}^n |1-e^{\pi i/m}|^p \cdot \|(z_j,0)\|_{Z_p(\mathbb{C})}^p + \frac{(k/n)^{p/2}}{4^n} \\
&\quad\times \sum_{x\in\mathbb{Z}_{2m}^n}\sum_{\varepsilon,\delta\in\{-1,1\}^n}\left\|\sum_{j=1}^n \delta_j(e^{(\pi i(x_j+\varepsilon_j))/m} - e^{\pi i x_j/m})(z_j,0)\right\|_{Z_p(\mathbb{C})}^p, \qquad (40)
\end{aligned}$$

where for the left-hand side of (40) we used the fact that $e^{(\pi i(x+m\sigma))/m} - e^{\pi i x/m} = -2e^{\pi i x/m}$ for every $\sigma \in \{-1,1\}$ and $x \in \mathbb{Z}_{2m}$.

Recalling the definition (36) of the norm of $Z_p(\mathbb{C})$, for every $S \subseteq \{1,\ldots,n\}$ we have

$$\begin{aligned}
&\sum_{\delta\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{2m}^n}\left\|\sum_{j\in S}\delta_j e^{\pi i x_j/m}(z_j,0)\right\|_{Z_p(\mathbb{C})}^p \\
&= \sum_{x\in\mathbb{Z}_{2m}^n}\int_0^{2\pi}\sum_{\delta\in\{-1,1\}^n}\left\|\sum_{j\in S}\delta_j\cos\left(\theta+\frac{\pi x_j}{m}\right)z_j\right\|_Z^p d\theta \\
&= \sum_{x\in\mathbb{Z}_{2m}^n}\int_0^{2\pi}\sum_{\delta\in\{-1,1\}^n}\left\|\sum_{j\in S}\delta_j\left|\cos\left(\theta+\frac{\pi x_j}{m}\right)\right|z_j\right\|_Z^p d\theta \\
&\geqslant 2\pi(2m)^n\sum_{\delta\in\{-1,1\}^n}\left\|\sum_{j\in S}\frac{\delta_j}{2\pi}\left(\int_0^{2\pi}|\cos\theta|\,d\theta\right)z_j\right\|_Z^p \qquad (41) \\
&= \frac{2^{p+1}(2m)^n}{\pi^{p-1}}\sum_{\delta\in\{-1,1\}^n}\left\|\sum_{j\in S}\delta_j z_j\right\|_Z^p, \qquad (42)
\end{aligned}$$

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

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

To bound the first term in the right-hand side of (40), use the fact that $|1-e^{i\theta}| \leqslant \theta$ for every $\theta \in [0,\pi]$, and the identity (37) to get

$$\begin{aligned}&\sum_{j=1}^{n} |1-e^{\pi i/m}|^p \cdot \|(z_j,0)\|^p_{Z_p(\mathbb{C})}\\ &\quad\leqslant \frac{\pi^p}{m^p}\left(\int_0^{2\pi} |\cos\theta|^p\,d\theta\right)\sum_{j=1}^{n}\|z_j\|_Z^p \leqslant \frac{\pi^{p+1}}{m^p}\sum_{j=1}^{n}\|z_j\|_Z^p, \end{aligned} \tag{43}$$

where we used that, since $p \geqslant 2$, we have $\int_0^{2\pi} |\cos\theta|^p\,d\theta \leqslant \int_0^{2\pi}\cos^2\theta\,d\theta = \pi$. To bound the second term in the right-hand side of (40), recall the contraction principle (see [46, Theorem 4.4]), which asserts that for every $a_1,\ldots,a_n \in \mathbb{R}$ we have

$$\sum_{\delta\in\{-1,1\}^n}\left\|\sum_{j=1}^{n} a_j\delta_j z_j\right\|_Z^p \leqslant \left(\max_{j\in\{1,\ldots,n\}}|a_j|^p\right)\sum_{\delta\in\{-1,1\}^n}\left\|\sum_{j=1}^{n}\delta_j z_j\right\|_Z^p. \tag{44}$$

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

$$\begin{aligned}&\sum_{\delta\in\{-1,1\}^n}\left\|\sum_{j=1}^{n}\delta_j(e^{(\pi i(x_j+\varepsilon_j))/m}-e^{\pi i x_j/m})(z_j,0)\right\|^p_{Z_p(\mathbb{C})}\\ &\quad=\int_0^{2\pi}\sum_{\delta\in\{-1,1\}^n}\left\|\sum_{j=1}^{n}\left(\cos\left(\theta+\frac{\pi x_j}{m}+\frac{\pi\varepsilon_j}{m}\right)-\cos\left(\theta+\frac{\pi x_j}{m}\right)\right)\delta_j z_j\right\|_Z^p d\theta\\ &\quad\leqslant 2\pi\left(\max_{\theta\in[0,2\pi]}\max_{j\in\{1,\ldots,n\}}\left|\cos\left(\theta+\frac{\pi x_j}{m}+\frac{\pi\varepsilon_j}{m}\right)\right.\right.\\ &\qquad\left.\left.-\cos\left(\theta+\frac{\pi x_j}{m}\right)\right|^p\right)\sum_{\delta\in\{-1,1\}^n}\left\|\sum_{j=1}^{n}\delta_j z_j\right\|_Z^p \qquad (45)\\ &\quad\leqslant \frac{2\pi^{p+1}}{m^p}\sum_{\delta\in\{-1,1\}^n}\left\|\sum_{j=1}^{n}\delta_j z_j\right\|_Z^p, \qquad (46)\end{aligned}$$

where (45) uses (44) and (46) uses $|\cos(\alpha\pm\pi/m)-\cos\alpha| = |\int_\alpha^{\alpha\pm\pi/m}\sin t\,dt| \leqslant \pi/m$, which holds true for every $\alpha \in [0,2\pi]$. The desired inequality (39) follows by combining (40), (42), (43) and (46). □

3. Nonembeddability

Here, we assume for the moment the validity of Theorem 1.2, whose proof appears in Section 4, and proceed to deduce its geometric consequences that

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

were stated in the Introduction. Namely, we will prove here Theorems 1.7, 1.11 and 1.14.

Proof of Theorem 1.7. We first make some preparatory elementary estimates that explain the origin of the quantities that appear in (24). Define $\psi_{p,q} : \mathbb{R} \to \mathbb{R}$ by

$$\psi_{p,q}(t) \stackrel{\text{def}}{=} \frac{3tp}{q} - 3 + \left(tp + 2 - \frac{2tp}{q} - p\right)\left(1 + \frac{tp - q}{q(p-2)}\right)$$
$$= \frac{p^2(q-2)}{q^2(p-2)}t^2 + \frac{p(pq - 3q + 2)}{q(p-2)}t - p.$$

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

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

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 ψ is quadratic with $\lim_{t\to\pm\infty} \psi_{p,q}(t) = \infty$, 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

$$\theta_{p,q} \stackrel{\text{def}}{=} \frac{2q(p-q) + q^2(p-1)(p-2)}{2p^2(q-2)}\left(\sqrt{1 + \frac{4p(p-2)(q-2)}{(pq - 3q + 2)^2}} - 1\right).$$

Consequently, $q/p < \theta_{p,q} < 1 - (p-q)(q-2)/(2p^3)$ (in particular, the rightmost inequality in (24) is valid), and

$$\forall\, \theta \in (0,1), \quad \psi_{p,q}(\theta) \leqslant 0 \implies \theta \leqslant \theta_{p,q}. \tag{47}$$

Now, suppose that $(L_q, \|x-y\|_q^\theta)$ admits a bi-Lipschitz embedding into L_p. If $\theta \leqslant q/p < \theta_{p,q}$ 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 $x, y \in \mathbb{Z}_{4m}^n$,

$$\left(\sum_{j=1}^n |e^{\pi i x_j/2m} - e^{\pi i y_j/2m}|^q\right)^{\theta/q} \leqslant \|f_{m,n}(x) - f_{m,n}(y)\|_p$$
$$\leqslant \Lambda\left(\sum_{j=1}^n |e^{\pi i x_j/2m} - e^{\pi i y_j/2m}|^q\right)^{\theta/q}. \tag{48}$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

Suppose that $m \geqslant n$ and define $k \in \{1, \dots, n\}$ by $k \stackrel{\text{def}}{=} \lceil n^3/m^2 \rceil$. By Theorem 1.2 and Remark 1.3, in conjunction with (48), we have

$$\frac{n^{3\theta p/q}}{m^{2\theta p/q+p}} = \frac{k^{\theta p/q}}{m^p} \leqslant (c(p)\Lambda)^p \left(\frac{k}{m^{\theta p}} + \left(\frac{k}{n}\right)^{p/2} \cdot \frac{n^{\theta p/q}}{m^{\theta p}} \right) = (2c(p)\Lambda)^p \left(\frac{n^3}{m^{\theta p+2}} + \frac{n^{p+\theta p/q}}{m^{(1+\theta)p}} \right), \tag{49}$$

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

Choose $m \in \mathbb{N}$ by setting

$$m \stackrel{\text{def}}{=} \lceil n^{(p-3+\theta p/q)/(p-2)} \rceil = \lceil n^{1+(\theta p-q)/(q(p-2))} \rceil.$$

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

$$\frac{n^3}{m^{\theta p+2}} + \frac{n^{p+\theta p/q}}{m^{(1+\theta)p}} \lesssim_p \frac{n^3}{m^{\theta p+2}},$$

and therefore by (49) (and our choice of m) we have

$$n^{3\theta p/q-3+(\theta p+2-2\theta p/q-p)(1+(\theta p-q)/(q(p-2)))} \lesssim_p (c(p)\Lambda)^p. \tag{50}$$

Since (50) is supposed to hold true for n that can be arbitrarily large, we necessarily have

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

Recalling (47), this implies that $\theta \leqslant \theta_{p,q}$, as required. □

Before proving Theorem 1.11 we record for future use the following very simple lemma.

LEMMA 3.1. *For every two integers $m, n \geqslant 2$ there exists a mapping $h_m^n : \mathbb{Z}_m^n \to \{0, \dots, 4m\}^{2n}$ such that for every $q \in [2, \infty)$ and $x, y \in \mathbb{Z}_m^n$ we have*

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

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

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

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

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

$$\begin{aligned}
&(|a_m(u) - a_m(v)|^q + |b_m(u) - b_m(v)|^q)^{1/q} \\
&\quad\leqslant \sqrt{|a_m(u) - a_m(v)|^2 + |b_m(u) - b_m(v)|^2} \\
&\quad\leqslant 2m|e^{2\pi iu/m} - e^{2\pi iv/m}| + \frac{2}{\sqrt{2}} \\
&\quad\leqslant 3m|e^{2\pi iu/m} - e^{2\pi iv/m}|,
\end{aligned}$$

since for distinct $u, v \in \mathbb{Z}_m$ we have $|e^{2\pi iu/m} - e^{2\pi iv/m}| \geqslant |e^{2\pi i/m} - 1| \geqslant 4/m$. Similarly,

$$\begin{aligned}
&(|a_m(u) - a_m(v)|^q + |b_m(u) - b_m(v)|^q)^{1/q} \\
&\quad\geqslant \frac{1}{\sqrt{2}}\sqrt{|a_m(u) - a_m(v)|^2 + |b_m(u) - b_m(v)|^2} \\
&\quad\geqslant \frac{2 - \sqrt{2}/4}{\sqrt{2}} m|e^{2\pi iu/m} - e^{2\pi iv/m}| \geqslant m|e^{2\pi iu/m} - e^{2\pi iv/m}|.
\end{aligned}$$

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

Proof of Theorem 1.11. We shall show that for an appropriate choice of $\beta_p \in (0, \infty)$ we have

$$m \geqslant n^{1+(p-q)/(q(p-2))} \implies c_p([16m]^{2n}_q) \geqslant \beta_p n^{(p-q)(q-2)/(q^2(p-2))}. \tag{51}$$

Since $[M]^n_q \supseteq [m]^n_q$ for every integer $M \geqslant m$ and $[m]^N_q$ contains an isometric copy of $[m]^n_q$ for every integer $N \geqslant n$, the validity of (51) implies the desired estimate (25).

Fix $D \in [1, \infty)$ and suppose that $f : [16m]^{2m}_q \to L_p$ satisfies

$$\forall\, x, y \in [16m]^{2n}_q, \quad \|x - y\|_q \leqslant \|f(x) - f(y)\|_p \leqslant D\|x - y\|_q. \tag{52}$$

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

$$\|F(x + e_j) - F(x)\|_p \leqslant 3mD|e^{\pi i/2m} - 1| \lesssim D, \tag{53}$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

$$\|F(x+\varepsilon)-F(x)\|_p \leqslant 3mD\left(\sum_{j=1}^{n}|e^{\pi i\varepsilon_j/2m}-1|^q\right)^{1/q} \lesssim Dn^{1/q}, \tag{54}$$

and

$$\|F(x+2m\varepsilon_S)-F(x)\|_p \geqslant m\left(\sum_{j\in S}|e^{\pi i}-1|^q\right)^{1/q} \gtrsim m|S|^{1/q}. \tag{55}$$

Denote

$$k = \lceil n^{(p(q-2))/(q(p-2))}\rceil. \tag{56}$$

Then $k \leqslant n$ and the assumption on m in (51) implies that $m \geqslant n^{3/2}/\sqrt{k}$. Hence, by Theorem 1.2 and Remark 1.3, combined with (53)–(55), there exists $K_p \in (0,\infty)$ such that

$$n^{(p^2(q-2))/(q^2(p-2))} \leqslant k^{p/q} \leqslant K_p^p D^p\left(k+\frac{k^{p/2}}{n^{p/2-p/q}}\right) \lesssim_p K_p^p D^p n^{(p(q-2))/(q(p-2))}.$$

Consequently,

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

REMARK 3.2. Lower bounds on $c_p([m]_q^n)$ that are weaker than those of Theorem 1.11 can also be deduced from general discretization principles (combined with the asymptotic computation of $c_p(\ell_q^n)$ in [33]), namely from Bourgain's discretization theorem [18] and its quantitative improvement for L_p spaces in [36]. Specifically, let B_q^n denote the unit ball of ℓ_q^n. Observe that $(1/m)\{-m,\ldots,m\}^n$ contains a δ-dense subset of B_q^n, with $\delta \leqslant n^{1/q}/m$. By [36, Theorem 1.3] (and the discussion immediately following it), we see that there exists a universal constant $\gamma \in (0,1)$ such that if

$$\frac{n^{1/q}}{m} \leqslant \frac{\gamma}{\sigma(p,q)n^{2+(p-q)(q-2)/(q^2(p-2))}} \leqslant \frac{\beta}{n^2 c_p(\ell_q^n)}$$

then

$$c_p([2m]_q^n) \geqslant \frac{c_p(\ell_q^n)}{2} \gtrsim \sigma(p,q)n^{(p-q)(q-2)/(q^2(p-2))},$$

where $\sigma(p,q) \in (0,\infty)$ is as in (10). Consequently,

$$\begin{aligned} m &\geqslant \frac{\sigma(p,q)}{\gamma}\cdot n^{2+1/q+(p-q)(q-2)/(q^2(p-2))} \\ &= \frac{\sigma(p,q)}{\gamma}\cdot n^{2+(p-q)/(q(p-2))+(p(q-2))/(q^2(p-2))} \\ &\Longrightarrow\ c_p([2m]_q^n) \gtrsim \sigma(p,q)n^{(p-q)(q-2)/(q^2(p-2))}. \end{aligned} \tag{57}$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

We note that a direct application of Bourgain's discretization theorem [18] (which holds true also for target spaces that need not be L_p spaces) would imply the same bound on $c_p([2m]_q^n)$ as in (57), provided that m is much larger than the requirement appearing in (57) (specifically, m would have to be at least doubly exponential in $n \log n$).

Proof of Theorem 1.14. The proof follows the proofs of Theorem 1.7 and Theorem 1.11 with a different (optimal) setting of parameters that is made possible due to the assumed validity of Conjecture 1.5. Specifically, we are now assuming that (21) holds true provided $m \geqslant C_p\sqrt{n/k}$.

Dealing first with (31), fix $\theta \in (q/p, 1]$ and $n \in \mathbb{N}$. Choose $m, k \in \mathbb{N}$ as follows.

$$m \stackrel{\text{def}}{=} \lfloor n^{(\theta p-q)/(q(p-2))} \rfloor \quad \text{and} \quad k \stackrel{\text{def}}{=} \left\lceil \frac{C_p^2 n}{m^2} \right\rceil. \tag{58}$$

Since $\theta > q/p$ we may assume that n is large enough so that $m \geqslant C_p$, in which case we have $k \in \{1, \ldots, n\}$ and $m \geqslant C_p\sqrt{n/k}$. Suppose for the sake of obtaining a contradiction that there exists $f_{m,n} : \mathbb{Z}_{4m}^n \to L_p$ satisfying (48). An application of (21) then yields the following estimate.

$$\begin{aligned}
\alpha_p C_p^{2\theta p/q} n^{(p(q^2-2\theta^2 p))/(q^2(p-2))} &\overset{(58)}{\leqslant} \frac{\alpha_p k^{\theta p/q}}{m^p} \\
&\overset{(21)\wedge(48)}{\leqslant} \Lambda^p \left(\frac{k}{m^{\theta p}} + \left(\frac{k}{n}\right)^{p/2} \cdot \frac{n^{\theta p/q}}{m^{\theta p}} \right) \\
&\overset{(58)}{\lesssim_p} (C_p \Lambda)^p n^{1-(2+\theta p)((\theta p-q)/(q(p-2)))}.
\end{aligned} \tag{59}$$

Since (59) holds true for arbitrarily large n, we conclude that

$$\frac{p(q^2-2\theta^2 p)}{q^2(p-2)} \leqslant 1-(2+\theta p)\frac{\theta p-q}{q(p-2)} = \frac{p(q^2-2\theta^2 p)}{q^2(p-2)} - \frac{\theta p(q-2)(\theta p-q)}{q^2(p-2)}.$$

Consequently $\theta \leqslant q/p$, contradicting the initial assumption that $\theta > q/p$. This proves (31).

Next, we have already seen in (28) that $c_p([m]_q^n)$ is bounded from above by a constant multiple of the quantity appearing in (29). By arguing as in the beginning of the proof of Theorem 1.11, it therefore suffices to show that for every $m, n \in \mathbb{N}$ we have

$$m \geqslant n^{(p-q)/(q(p-2))} \implies c_p([16m]_q^{2n}) \geqslant \xi(p) n^{(p-q)(q-2)/(q^2(p-2))} \tag{60}$$

for some $\xi(p) \in (0, \infty)$. To this end, suppose that there exists $f : [16m]_q^{2n} \to L_p$ satisfying (52), our goal being to bound D from below. As explained in the proof

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

of Theorem 1.11, this implies the existence of $F : \mathbb{Z}_{4m}^n \to L_p$ that satisfies (53), (54) and (55). Similarly, for (56), choose $k \in \mathbb{N}$ to be

$$k \stackrel{\text{def}}{=} \lceil C_p^2 n^{(p(q-2))/(q(p-2))} \rceil. \tag{61}$$

We may suppose that n is large enough so that $k \in \{1, \ldots, n\}$, since otherwise (60) is vacuous. The lower bound on m that is assumed in (60) implies that $m \geqslant C_p \sqrt{n/k}$, so we may apply (21), yielding, in conjunction with (53)–(55), that the following holds true.

$$\begin{aligned} \alpha_p C_p^{p/q} n^{(p^2(q-2))/(q^2(p-2))} &\stackrel{(61)}{\leqslant} \alpha_p k^{p/q} \lesssim_p D^p \left(k + \frac{k^{p/2}}{n^{p/2-p/q}} \right) \\ &\stackrel{(61)}{\lesssim_p} (C_p D)^p n^{(p(q-2))/(q(p-2))} \\ &\implies D \gtrsim \frac{\alpha_p^{1/p}}{C_p^{1-1/q}} \cdot n^{(p-q)(q-2)/(q^2(p-2))}. \end{aligned}$$

□

4. Proof of Theorem 1.2

Suppose from now on that $m, n \in \mathbb{N}$ satisfy $m \geqslant n$ 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

$$U_S \stackrel{\text{def}}{=} \{y \in [-R, R]^n : \forall (i, j) \in S \times (\{1, \ldots, n\} \setminus S), (y_i, y_j) \in (2\mathbb{Z}) \times (1 + 2\mathbb{Z})\}. \tag{62}$$

Thus U_S consists for those $y \in \mathbb{Z}_{4m}^n$ satisfying $|y_j| \leqslant R$ 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 $|U_S| = R^{|S|}(R+1)^{n-|S|}$. Given a Banach space $(X, \|\cdot\|_X)$, the averaging operator corresponding to U_S will be denoted below by $D_S : L_2(\mathbb{Z}_{4m}^n, X) \to L_2(\mathbb{Z}_{4m}^n, X)$, that is, for every $f : \mathbb{Z}_{4m}^n \to X$ and $x \in \mathbb{Z}_{4m}^n$ we set

$$D_S f(x) \stackrel{\text{def}}{=} \frac{1}{|U_S|} \sum_{y \in U_S} f(x + y). \tag{63}$$

The following lemma extends [61, Lemma 5.1], which corresponds to the special case $|S| = 1$.

LEMMA 4.1. *Suppose that $m, n \in \mathbb{N}$, and that $R \in \{1, \ldots, 2m-1\}$ is odd. Let $(X, \|\cdot\|_X)$ be a Banach space and $p \in [1, \infty)$. Then for every $f : \mathbb{Z}_{4m}^n \to X$ and*

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

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

$$\sum_{x\in\mathbb{Z}_{4m}^n} \|f(x)-D_Sf(x)\|_X^p \lesssim_p \frac{R^p}{2^n}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{4m}^n}\|f(x+\varepsilon)-f(x)\|_X^p + \frac{1}{2^n}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{4m}^n}\|f(x+\varepsilon_S)-f(x)\|_X^p. \quad (64)$$

Proof. For every $w\in\mathbb{Z}^n$ all of whose coordinates are odd fix $\gamma_w:\mathbb{N}\cup\{0\}\to\mathbb{Z}^n$ 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 γ_w is explained in [61, Lemma 5.1], 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 $\gamma_w=\mathrm{sign}(w)\cdot\gamma_{|w|}$, where the multiplication is coordinate-wise and we denote $|w|=(|w_1|,\ldots,|w_n|)$ and $\mathrm{sign}(w)=(\mathrm{sign}(w_1),\ldots,\mathrm{sign}(w_n))$. 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

$$\gamma_w(2t-1)\stackrel{\mathrm{def}}{=}\gamma_w(2t-2)+\sum_{j=1}^n e_j,$$

and

$$\gamma_w(2t)\stackrel{\mathrm{def}}{=}\gamma_w(2t-1)+\sum_{\substack{j\in\{1,\ldots,n\}\\ \gamma_w(2t-1)_j<w_j}} e_j-\sum_{\substack{j\in\{1,\ldots,n\}\\ \gamma_w(2t-1)_j=w_j}} e_j.$$

This explicit definition of γ_w is not used below; we shall only need to know that γ_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 γ_w to $\{0,\ldots,\|w\|_\infty\}$ is an ℓ_∞ 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 the 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 $\gamma_{x,y}^\eta:\mathbb{N}\to\mathbb{Z}^n$ by $\gamma_{x,y}^\eta=x+\eta_S+\gamma_{y-\eta_S}$. Thus, $\gamma_{x,y}^\eta(0)=x+\eta_S$, $\gamma_{x,y}^\eta(\|y-\eta_S\|_\infty)=x+y$ and $\gamma_{x,y}^\eta(t)-\gamma_{x,y}^\eta(t-1)\in\{-1,1\}^n$ for all $t\in\mathbb{N}$. Note that $\gamma_{x,y}^\eta$ depends only on those coordinates of η that belong to S.

For every $z\in\mathbb{Z}_{4m}^n$ and $\varepsilon,\eta\in\{-1,1\}^n$ define

$$F_\eta(z,\varepsilon)\stackrel{\mathrm{def}}{=}\{(x,y)\in\mathbb{Z}_{4m}^n\times U_S:\gamma_{x,y}^\eta(t-1)=z \text{ and } \gamma_{x,y}^\eta(t)=z+\varepsilon \text{ for some } t\in[1,\|y-\eta_S\|_\infty]\}.$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

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 $t \in \{1, \dots, \|y - \eta_S\|_\infty\}$ for which $\gamma_{x,y}^\eta(t-1) = z$ and $\gamma_{x,y}^\eta(t) = z + \varepsilon$.

We claim that

$$N \stackrel{\text{def}}{=} \sum_{\eta \in \{-1,1\}^n} |F_\eta(z, \varepsilon)| \tag{65}$$

is independent of $z \in \mathbb{Z}_{4m}^n$ and $\varepsilon \in \{-1, 1\}^n$. Indeed, for every $\varepsilon, \delta \in \{-1, 1\}^n$ and $z, w \in \mathbb{Z}_{4m}^n$ define a bijection $\psi_{z,w}^{\varepsilon,\delta} : \mathbb{Z}_{4m}^n \times U_S \to \mathbb{Z}_{4m}^n \times U_S$ by

$$\psi_{z,w}^{\varepsilon,\delta}(x, y) \stackrel{\text{def}}{=} (w - \varepsilon\delta z + \varepsilon\delta x, \varepsilon\delta y).$$

Then for every $\eta \in \{-1, 1\}^n$ we have $\gamma^{\varepsilon\delta\eta}_{\psi_{z,w}^{\varepsilon,\delta}(x,y)} = w - \varepsilon\delta z + \varepsilon\delta\gamma_{x,y}^\eta$. Consequently,

$$(\gamma_{x,y}^\eta(t-1), \gamma_{x,y}^\eta(t)) = (z, z+\varepsilon) \iff (\gamma^{\varepsilon\delta\eta}_{\psi_{z,w}^{\varepsilon,\delta}(x,y)}(t-1), \gamma^{\varepsilon\delta\eta}_{\psi_{z,w}^{\varepsilon,\delta}(x,y)}(t)) = (w, w+\delta)$$

for every $t \in \{1, \dots, \|y - \eta_S\|_\infty\}$. 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,

$$\sum_{\eta \in \{-1,1\}^n} |F_\eta(z, \varepsilon)| = \sum_{\eta \in \{-1,1\}^n} |F_{\varepsilon\delta\eta}(w, \delta)| = \sum_{\eta \in \{-1,1\}^n} |F_\eta(w, \delta)|,$$

implying that the integer N defined in (65) is indeed independent of $(z, \varepsilon) \in \mathbb{Z}_{4m}^n \times \{-1, 1\}^n$.

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

$$\begin{aligned} N(8m)^n &= \sum_{z \in \mathbb{Z}_{4m}^n} \sum_{\varepsilon, \eta \in \{-1,1\}^n} |F_\eta(z, \varepsilon)| \\ &= \sum_{z \in \mathbb{Z}_{4m}^n} \sum_{\varepsilon, \eta \in \{-1,1\}^n} \sum_{(x,y) \in \mathbb{Z}_{4m}^n \times U_S} \sum_{t=1}^{\|y-\eta_S\|_\infty} \mathbf{1}_{\{\gamma_{x,y}^\eta(t-1) = z \wedge \gamma_{x,y}^\eta(t) = z+\varepsilon\}} \\ &= \sum_{\eta \in \{-1,1\}^n} \sum_{(x,y) \in \mathbb{Z}_{4m}^n \times U_S} \|y - \eta_S\|_\infty \\ &\leqslant R(8m)^n |U_S|. \end{aligned}$$

Consequently,

$$N \leqslant R|U_S|. \tag{66}$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

Now, fix $f : \mathbb{Z}_{4m}^n \to X$. For every $x \in \mathbb{Z}_{4m}^n$, $y \in U_S$ and $\eta \in \{-1,1\}^n$ we have

$$\begin{aligned}
&\|f(x)-f(x+y)\|_X^p \\
&\quad\lesssim_p \|f(x)-f(x+\eta_S)\|_X^p + \|f(\gamma_{x,y}^\eta(0)) - f(\gamma_{x,y}^\eta(\|y-\eta_S\|_\infty))\|_X^p \\
&\quad\lesssim_p \|f(x)-f(x+\eta_S)\|_X^p + \|y-\eta_S\|_\infty^{p-1} \\
&\qquad\times \sum_{t=1}^{\|y-\eta_S\|_\infty} \|f(\gamma_{x,y}^\eta(t-1)) - f(\gamma_{x,y}^\eta(t))\|_X^p \\
&\quad\leqslant \|f(x)-f(x+\eta_S)\|_X^p + R^{p-1} \sum_{t=1}^{\|y-\eta_S\|_\infty} \|f(\gamma_{x,y}^\eta(t-1)) - f(\gamma_{x,y}^\eta(t))\|_X^p.
\end{aligned}$$

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

$$\begin{aligned}
\|f(x)-f(x+y)\|_X^p &\lesssim_p \frac{1}{2^n}\sum_{\eta\in\{-1,1\}^n} \|f(x)-f(x+\eta_S)\|_X^p \\
&\quad + \frac{R^{p-1}}{2^n}\sum_{\eta\in\{-1,1\}^n}\sum_{t=1}^{\|y-\eta_S\|_\infty} \|f(\gamma_{x,y}^\eta(t-1)) - f(\gamma_{x,y}^\eta(t))\|_X^p.
\end{aligned}$$

Consequently, using the definition of the operator D_S and convexity, we see that

$$\begin{aligned}
&\sum_{x\in Z_{4m}^n} \|f(x)-D_S f(x)\|_X^p \\
&\quad\leqslant \frac{1}{|U_S|}\sum_{x\in Z_{4m}^n}\sum_{y\in U_S} \|f(x)-f(x+y)\|_X^p \\
&\quad\lesssim_p \frac{1}{2^n}\sum_{\eta\in\{-1,1\}^n}\sum_{x\in Z_{4m}^n} \|f(x)-f(x+\eta_S)\|_X^p \\
&\qquad + \frac{R^{p-1}}{2^n|U_S|}\sum_{\eta\in\{-1,1\}^n}\sum_{t=1}^{\|y-\eta_S\|_\infty}\sum_{x\in Z_{4m}^n}\sum_{y\in U_S} \|f(\gamma_{x,y}^\eta(t-1)) - f(\gamma_{x,y}^\eta(t))\|_X^p \\
&\quad= \frac{1}{2^n}\sum_{\eta\in\{-1,1\}^n}\sum_{x\in Z_{4m}^n} \|f(x)-f(x+\eta_S)\|_X^p \\
&\qquad + \frac{R^{p-1}N}{2^n|U_S|}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{z\in\mathbb{Z}_{4m}^n} \|f(z+\varepsilon)-f(z)\|_X^p.
\end{aligned}$$

Recalling the upper bound on N appearing in (66), this implies the desired estimate (64). □

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

We record the following very simple lemma for future use.

LEMMA 4.2. *Suppose that (X, d_X) is a metric space and $p \in [1, \infty)$. Then for every $f : \mathbb{Z}_{4m}^n \to X$, $\varepsilon \in \{-1, 1\}^n$ and $S \subseteq \{1, \dots, n\}$, we have*

$$\sum_{x \in \mathbb{Z}_{4m}^n} d_X(f(x + \varepsilon_S), f(x))^p \leqslant |S|^{p-1} \sum_{j \in S} \sum_{x \in \mathbb{Z}_{4m}^n} d_X(f(x + e_j), f(x))^p. \quad (67)$$

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

$$\begin{aligned} &d_X(f(x + \varepsilon_S), f(x))^p \\ &\quad \leqslant |S|^{p-1} \sum_{\ell=1}^{|S|} d_X(f(x + \varepsilon_{S(\ell-1)} + \varepsilon_{j(\ell)} e_{j(\ell)}), f(x + \varepsilon_{S(\ell-1)}))^p. \end{aligned}$$

Hence,

$$\begin{aligned} \sum_{x \in \mathbb{Z}_{4m}^n} d_X(f(x + \varepsilon_S), f(x))^p &\leqslant |S|^{p-1} \sum_{\ell=1}^{|S|} \sum_{y \in \mathbb{Z}_{4m}^n} d_X(f(y + \varepsilon_{j(\ell)} e_{j(\ell)}), f(y))^p \\ &= |S|^{p-1} \sum_{\ell=1}^{|S|} \sum_{z \in \mathbb{Z}_{4m}^n} d_X(f(z + e_{j(\ell)}), f(z))^p \\ &= |S|^{p-1} \sum_{j \in S} \sum_{z \in \mathbb{Z}_{4m}^n} d_X(f(z + e_j), f(z))^p. \end{aligned}$$

□

LEMMA 4.3. *Suppose that $m, n \in \mathbb{N}$, and that $R \in \{1, \dots, 2m - 1\}$ is odd and $k \in \{1, \dots, n\}$. Let $(X, \|\cdot\|_X)$ be a Banach space and $p \in [1, \infty)$. Then for every $f : \mathbb{Z}_{4m}^n \to X$ and $\delta \in \{-1, 1\}^n$,*

$$\begin{aligned} &\frac{1}{\binom{n}{k}} \sum_{\substack{S \subseteq \{1, \dots, n\} \\ |S| = k}} \sum_{x \in \mathbb{Z}_{4m}^n} \|f(x + 2m\delta_S) - f(x)\|_X^p \\ &\lesssim_p \frac{m^p}{\binom{n}{k}} \sum_{\substack{S \subseteq \{1, \dots, n\} \\ |S| = k}} \sum_{x \in \mathbb{Z}_{4m}^n} \|D_S f(x + 2\delta_S) - D_S f(x)\|_X^p \\ &\quad + \frac{R^p}{2^n} \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_{4m}^n} \|f(x + \varepsilon) - f(x)\|_X^p \\ &\quad + \frac{k^p}{n} \sum_{j=1}^{n} \sum_{x \in \mathbb{Z}_{4m}^n} \|f(x + e_j) - f(x)\|_X^p. \end{aligned} \quad (68)$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

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

$$\begin{aligned}
&\sum_{x\in\mathbb{Z}_{4m}^n} \|f(x+2m\delta_S) - f(x)\|_X^p \\
&\lesssim_p \sum_{x\in\mathbb{Z}_{4m}^n} \|D_S f(x+2m\delta_S) - D_S f(x)\|_X^p \\
&\quad + \sum_{x\in\mathbb{Z}_{4m}^n} \|D_S f(x+2m\delta_S) - f(x+2m\delta_S)\|_X^p + \sum_{x\in\mathbb{Z}_{4m}^n} \|D_S f(x) - f(x)\|_X^p \\
&= \sum_{x\in\mathbb{Z}_{4m}^n} \|D_S f(x+2m\delta_S) - D_S f(x)\|_X^p + 2\sum_{x\in\mathbb{Z}_{4m}^n} \|D_S f(x) - f(x)\|_X^p. \quad (69)
\end{aligned}$$

The first term in (69) can be bounded as follows:

$$\begin{aligned}
&\sum_{x\in\mathbb{Z}_{4m}^n} \|D_S f(x+2m\delta_S) - D_S f(x)\|_X^p \\
&\leqslant m^{p-1} \sum_{t=1}^{m} \sum_{x\in\mathbb{Z}_{4m}^n} \|D_S f(x+2t\delta_S) - D_S f(x+(2t-2)\delta_S)\|_X^p \\
&= m^p \sum_{x\in\mathbb{Z}_{4m}^n} \|D_S f(x+2\delta_S) - D_S f(x)\|_X^p. \quad (70)
\end{aligned}$$

The second term in (69) is bounded using Lemmas 4.1 and 4.2 as follows:

$$\begin{aligned}
\sum_{x\in\mathbb{Z}_{4m}^n} \|D_S f(x) - f(x)\|_X^p \lesssim_p \frac{R^p}{2^n} &\sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^n} \|f(x+\varepsilon) - f(x)\|_X^p \\
&+ |S|^{p-1} \sum_{j\in S} \sum_{x\in\mathbb{Z}_{4m}^n} \|f(x+e_j) - f(x)\|_X^p. \quad (71)
\end{aligned}$$

Note that for every $x \in \mathbb{Z}_{4m}^n$,

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

Hence, the desired inequality (68) follows by substituting (70) and (71) into (69) and averaging the resulting inequality over all $S \subseteq \{1,\ldots,n\}$ with $|S| = k$. □

Our next goal is to bound the first term in the right-hand side of (68). To this end, we first recall some results from [**34**].

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

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

$$Af(x) \stackrel{\text{def}}{=} \frac{1}{R^n} \sum_{y \in (-R,R)^n \cap (2\mathbb{Z})^n} f(x+y). \tag{72}$$

For $j \in \{1, \dots, n\}$ denote $B_j = D_{\{j\}}$, that is, 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_ℓ is odd for every $\ell \in \{1, \dots, n\} \setminus \{j\}$. (In [34], the set $U_{\{j\}}$ was denoted $S(j, R)$ and the operator B_j was denoted $\mathcal{E}_j$.)

It follows from [34] that for every $f : \mathbb{Z}_{4m}^n \to X$, every $p \in [1, \infty)$ and every $\varepsilon \in \{-1, 1\}^n$ we have

$$\begin{aligned} &\sum_{x \in \mathbb{Z}_{4m}^n} \Bigg\| \left(\frac{R}{R+1}\right)^{n-1} (Af(x+\varepsilon) - Af(x-\varepsilon)) \\ &\qquad - \sum_{j=1}^n \varepsilon_j [B_j f(x+e_j) - B_j f(x-e_j)] \Bigg\|_X^p \\ &\lesssim_p p^p \sum_{s=0}^{n-1} \frac{(n/R)^{(n-s)p}}{\binom{n}{s}} \sum_{\substack{S \subseteq \{1,\dots,n\} \\ |S|=s}} \sum_{x \in \mathbb{Z}_{4m}^n} \|f(x + 2\varepsilon_S) - f(x)\|_X^p. \end{aligned} \tag{73}$$

Since (73) is only implicit in [34] (it follows from proofs in [34] rather than from explicit statements in [34]), we shall now explain how to establish (73).

Proof of (73). For every $T \subseteq \{1, \dots, n\}$ define $L_T \subseteq \mathbb{Z}_{4m}^n$ by

$$L_T \stackrel{\text{def}}{=} \{y \in (-R, R)^n : \forall (i, j) \in T \times (\{1, \dots, n\} \setminus T), (y_i, y_j) \in 2\mathbb{Z} \times \{0\}\}.$$

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 [34, Definition 3.2], we let $\Delta_T : L_2(\mathbb{Z}_{4m}^n, X) \to L_2(\mathbb{Z}_{4m}^n, X)$ denote the averaging operator corresponding to L_T, that is, for every $f : \mathbb{Z}_{4m}^n \to X$ and $x \in \mathbb{Z}_{4m}^n$,

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

We note in passing that the operator A given in (72) coincides with $\Delta_{\{1,\dots,n\}}$.

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

For $\varepsilon \in \{-1,1\}^n$, $\alpha \in \{0,\ldots,n\}$ and $\beta \in \{0,\ldots,\alpha\}$ define $V^{\varepsilon}_{\alpha,\beta} : L_2(\mathbb{Z}^n_{4m}, X) \to L_2(\mathbb{Z}^n_{4m}, X)$ by setting for every $f : \mathbb{Z}^n_{4m} \to X$ and $x \in \mathbb{Z}^n_{4m}$,

$$V^{\varepsilon}_{\alpha,\beta} f(x) \stackrel{\text{def}}{=} \sum_{\substack{T\subseteq\{1,\ldots,n\}\\ |T|=n-\alpha}} \sum_{\substack{\delta\in\{-1,1\}^{\{1,\ldots,n\}\setminus T}\\ \langle\delta,\varepsilon_{\{1,\ldots,n\}\setminus T}\rangle=\alpha-2\beta}} [\Delta_T f(x+R\delta+\varepsilon_T) - \Delta_T f(x+R\delta-\varepsilon_T)]. \tag{74}$$

Here $\langle\cdot,\cdot\rangle$ denotes the standard scalar product on $\mathbb{R}^n$. It is worthwhile to compare the right-hand side of (74) to the right-hand side of [**34**, equation (44)] (however, note that there is a difference of a normalization factor. Our R is the same as the parameter k of [**34**]). By combining [**34**, Lemma 3.8] with [**34**, Lemma 3.5] and identity (44) of [**34**], 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 [**34**, Section 3.1]) such that $h_{0,0} = 1$,

$$\forall\, \alpha \in \{0,\ldots,n\},\ \forall\, \beta \in \{0,\ldots,\alpha\}, \quad |h_{\alpha,\beta}| \lesssim \frac{(\alpha-\beta)!\beta!}{2^\alpha}, \tag{75}$$

and for every $f : \mathbb{Z}^n_{4m} \to X$ and $x \in \mathbb{Z}^n_{4m}$,

$$\sum_{j=1}^n \varepsilon_j [B_j f(x+e_j) - B_j f(x-e_j)] = \left(\frac{R}{R+1}\right)^{n-1} \sum_{\alpha=0}^n \sum_{\beta=0}^\alpha \frac{h_{\alpha,\beta}}{R^\alpha} V^{\varepsilon}_{\alpha,\beta} f(x). \tag{76}$$

Observe that $V^{\varepsilon}_{0,0} f(x) = Af(x+\varepsilon) - Af(x-\varepsilon)$, so it follows from (76) that

$$\begin{aligned}
&\Bigg\| \sum_{j=1}^n \varepsilon_j [B_j f(x+e_j) - B_j f(x-e_j)] \\
&\qquad - \left(\frac{R}{R+1}\right)^{n-1} (Af(x+\varepsilon) - Af(x-\varepsilon)) \Bigg\|_X \\
&\stackrel{(76)}{\leqslant} \sum_{\alpha=1}^n \sum_{\beta=0}^\alpha \frac{|h_{\alpha,\beta}|}{R^\alpha} \|V^{\varepsilon}_{\alpha,\beta} f(x)\|_X \stackrel{(75)}{\lesssim} \sum_{\alpha=1}^n \frac{1}{2^\alpha} \sum_{\beta=0}^\alpha \frac{(\alpha-\beta)!\beta!}{R^\alpha} \|V^{\varepsilon}_{\alpha,\beta} f(x)\|_X.
\end{aligned} \tag{77}$$

By convexity, it follows from (77) that

$$\begin{aligned}
&\Bigg\| \sum_{j=1}^n \varepsilon_j [B_j f(x+e_j) - B_j f(x-e_j)] - \left(\frac{R}{R+1}\right)^{n-1} (Af(x+\varepsilon) - Af(x-\varepsilon)) \Bigg\|^p_X \\
&\lesssim_p \sum_{\alpha=1}^n \frac{1}{2^\alpha} \left(\sum_{\beta=0}^\alpha \frac{(\alpha-\beta)!\beta!}{R^\alpha} \|V^{\varepsilon}_{\alpha,\beta} f(x)\|_X \right)^p \\
&\leqslant \sum_{\alpha=1}^n \sum_{\beta=0}^\alpha \frac{(\alpha+1)^{p-1}((\alpha-\beta)!\beta!)^p}{2^\alpha R^{\alpha p}} \|V^{\varepsilon}_{\alpha,\beta} f(x)\|^p_X.
\end{aligned}$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

We can therefore bound the left-hand side of (73) as follows:

$$\sum_{x\in\mathbb{Z}^n_{4m}}\Bigg\|\left(\frac{R}{R+1}\right)^{n-1}(Af(x+\varepsilon)-Af(x-\varepsilon)) - \sum_{j=1}^{n}\varepsilon_j[B_jf(x+e_j)-B_jf(x-e_j)]\Bigg\|_X^p \lesssim_p \sum_{\alpha=1}^{n}\sum_{\beta=0}^{\alpha}\frac{(\alpha+1)^{p-1}((\alpha-\beta)!\beta!)^p}{2^\alpha R^{\alpha p}}\sum_{x\in\mathbb{Z}^n_{4m}}\|V^\varepsilon_{\alpha,\beta}f(x)\|_X^p. \tag{78}$$

Since the number of terms in the sums that appear in the definition (74) of $V^\varepsilon_{\alpha,\beta}$ is $\binom{n}{\alpha}\binom{\alpha}{\beta}$,

$$\begin{aligned}
&\sum_{x\in\mathbb{Z}^n_{4m}}\|V^\varepsilon_{\alpha,\beta}f(x)\|_X^p\\
&\leqslant \binom{n}{\alpha}^{p-1}\binom{\alpha}{\beta}^{p-1}\sum_{\substack{T\subseteq\{1,\dots,n\}\\|T|=n-\alpha}}\sum_{\substack{\delta\in\{-1,1\}^{\{1,\dots,n\}\setminus T}\\ \langle\delta,\varepsilon_{\{1,\dots,n\}\setminus T}\rangle=\alpha-2\beta}}\\
&\quad\times\sum_{x\in\mathbb{Z}^n_{4m}}\|\Delta_Tf(x+R\delta+\varepsilon_T)-\Delta_Tf(x+R\delta-\varepsilon_T)\|_X^p \qquad (79)\\
&= \binom{n}{\alpha}^{p-1}\binom{\alpha}{\beta}^{p}\sum_{\substack{T\subseteq\{1,\dots,n\}\\|T|=n-\alpha}}\sum_{x\in\mathbb{Z}^n_{4m}}\|\Delta_Tf(x+2\varepsilon_T)-\Delta_Tf(x)\|_X^p\\
&\leqslant \binom{n}{\alpha}^{p-1}\binom{\alpha}{\beta}^{p}\sum_{\substack{T\subseteq\{1,\dots,n\}\\|T|=n-\alpha}}\sum_{x\in\mathbb{Z}^n_{4m}}\|f(x+2\varepsilon_T)-f(x)\|_X^p, \qquad (80)
\end{aligned}$$

where (80) is valid since Δ_T is an averaging operator.

By combining (78) with (80) we see that

$$\sum_{x\in\mathbb{Z}^n_{4m}}\Bigg\|\left(\frac{R}{R+1}\right)^{n-1}(Af(x+\varepsilon)-Af(x-\varepsilon)) - \sum_{j=1}^{n}\varepsilon_j[B_jf(x+e_j)-B_jf(x-e_j)]\Bigg\|_X^p \lesssim_p \sum_{\alpha=1}^{n}\sum_{\beta=0}^{\alpha}\frac{(\alpha+1)^{p-1}((\alpha-\beta)!\beta!)^p}{2^\alpha R^{\alpha p}}\binom{n}{\alpha}^{p-1}\binom{\alpha}{\beta}^{p}$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

$$\times \sum_{\substack{T\subseteq\{1,\ldots,n\}\\ |T|=n-\alpha}} \sum_{x\in\mathbb{Z}_{4m}^n} \|f(x+2\varepsilon_T)-f(x)\|_X^p \tag{81}$$

$$= \sum_{\alpha=1}^{n} \frac{(\alpha+1)^p}{2^\alpha R^{\alpha p}\binom{n}{\alpha}} \left(\frac{n!}{(n-\alpha)!}\right)^p \times \sum_{\substack{T\subseteq\{1,\ldots,n\}\\ |T|=n-\alpha}} \sum_{x\in\mathbb{Z}_{4m}^n} \|f(x+2\varepsilon_T)-f(x)\|_X^p. \tag{82}$$

The desired estimate (73) is a consequence of (81) via the change of variable $s=n-\alpha$ and by using the bounds $n!/(n-\alpha)!\leqslant n^\alpha$ and $(\alpha+1)^p/2^\alpha\leqslant(2p)^p$. □

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*

$$\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{4m}^n} d_X(f(x),f(x+2\varepsilon_S))^p \leqslant 2^p \sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{4m}^n} d_X(f(x),f(x+\varepsilon))^p. \tag{83}$$

Proof. For every $\varepsilon,\delta\in\{-1,1\}^n$, we have

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

Hence,

$$\sum_{x\in\mathbb{Z}_{4m}^n} d_X(f(x),f(x+2\varepsilon_S))^p \leqslant 2^{p-1}\sum_{x\in\mathbb{Z}_{4m}^n}(d_X(f(x),f(x+\varepsilon_S+\delta_{\{1,\ldots,n\}\setminus S}))^p + d_X(f(x+\varepsilon_S+\delta_{\{1,\ldots,n\}\setminus S}),f(x+2\varepsilon_S))^p) \tag{84}$$

$$= 2^{p-1}\sum_{x\in\mathbb{Z}_{4m}^n}(d_X(f(x),f(x+\varepsilon_S+\delta_{\{1,\ldots,n\}\setminus S}))^p + d_X(f(x),f(x+\varepsilon_S-\delta_{\{1,\ldots,n\}\setminus S}))^p). \tag{85}$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

By averaging (84) over $\delta \in \{-1,1\}^n$ while using the fact that $\delta_{\{1,\dots,n\}\setminus S}$ and $-\delta_{\{1,\dots,n\}\setminus S}$ are identically distributed, we deduce that

$$\sum_{x\in\mathbb{Z}_{4m}^n} d_X(f(x), f(x+2\varepsilon_S))^p \leqslant \frac{2^p}{2^n} \sum_{\delta\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^n} d_X(f(x), f(x+\varepsilon_S+\delta_{\{1,\dots,n\}\setminus S}))^p. \tag{86}$$

If ε and δ are i.i.d. and uniformly distributed over $\{-1,1\}^n$, then the vector $\varepsilon_S + \delta_{\{1,\dots,n\}\setminus S}$ is also uniformly distributed over $\{-1,1\}^n$. Consequently, the desired estimate (83) follows by averaging (86) 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 $R \geqslant 2n-1$ (in addition to the previous assumptions on R, that is, that it is an odd integer with $R \leqslant 2m$). Then for every $p \in [1,\infty)$ and $f : \mathbb{Z}_{4m}^n \to X$ we have*

$$\begin{aligned}\sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^n} \Bigg\| \left(\frac{R}{R+1}\right)^{n-1} (Af(x+\varepsilon) - Af(x-\varepsilon)) \\ - \sum_{j=1}^n \varepsilon_j [B_j f(x+e_j) - B_j f(x-e_j)] \Bigg\|_X^p \\ \lesssim_p \left(\frac{pn}{R}\right)^p \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^n} \|f(x) - f(x+\varepsilon)\|_X^p. \end{aligned} \tag{87}$$

Proof. By summing (73) over $\varepsilon \in \{-1,1\}^n$ and using Lemma 4.4 we see that the left-hand side of (87) is at most $(O(1)p)^p$ times the following quantity

$$\begin{aligned}\left(\sum_{s=0}^{n-1} \left(\frac{n}{R}\right)^{(n-s)p}\right) \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^n} \|f(x) - f(x+\varepsilon)\|_X^p \\ \lesssim \left(\frac{n}{R}\right)^p \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^n} \|f(x) - f(x+\varepsilon)\|_X^p, \end{aligned}$$

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

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 (68).

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

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 $2|S|-1 \leqslant R \leqslant 2m$. Then for every $p \in [1,\infty)$ and $f : \mathbb{Z}_{4m}^n \to X$ we have*

$$\begin{aligned}\sum_{\delta\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{4m}^n}\|D_S f(x+2\delta_S)-D_S f(x)\|_X^p &\lesssim_p \sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{4m}^n}\left\|\sum_{j\in S}\varepsilon_j[B_j f(x+e_j)-B_j f(x-e_j)]\right\|_X^p \\ &\quad+\left(\frac{p|S|}{R}\right)^p\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{4m}^n}\|f(x)-f(x+\varepsilon)\|_X^p. \end{aligned} \tag{88}$$

Proof. Denote $k \stackrel{\text{def}}{=} |S|$, $T \stackrel{\text{def}}{=} \{1,\ldots,n\}\setminus S$ and consider $\mathbb{Z}_{4m}^n$ as being equal to $\mathbb{Z}_{4m}^S \times \mathbb{Z}_{4m}^T$. For every $y \in \mathbb{Z}_{4m}^T$ define $f_y : \mathbb{Z}_{4m}^S \to X$ by setting for every $x \in \mathbb{Z}_{4m}^S$,

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

Let $A^{(S)}$ be the averaging operator corresponding to (72) with $\mathbb{Z}_{4m}^n$ replaced by $\mathbb{Z}_{4m}^S$, that is, for every $h : \mathbb{Z}_{4m}^S \to X$ and $x \in \mathbb{Z}_{4m}^S$,

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

Similarly, for every $j \in S$ let $B_j^{(S)}$ be the averaging operator analogous to B_j but with $\mathbb{Z}_{4m}^n$ replaced by $\mathbb{Z}_{4m}^S$, that is, for every $h : \mathbb{Z}_{4m}^S \to X$ and $x \in \mathbb{Z}_{4m}^S$,

$$B_j^{(S)}h(x) \stackrel{\text{def}}{=} \frac{1}{R(R+1)^{k-1}}\sum_{\substack{a\in[-R,R]\cap(2\mathbb{Z})\\ b\in([-R,R]\cap(1+2\mathbb{Z}))^{S\setminus\{j\}}}} h\left(x+ae_j+\sum_{s\in S\setminus\{j\}}b_s e_s\right).$$

With these definitions, for every $(x,y) \in \mathbb{Z}_{4m}^S \times \mathbb{Z}_{4m}^T$ and $j \in S$ we have

$$D_S f(x,y) = A^{(S)}f_y(x) \quad\text{and}\quad B_j f(x,y) = B_j^{(S)}f_y(x). \tag{89}$$

Since $R \geqslant 2k-1$, an application of (73) to f_y yields the following estimate, which holds true for every fixed $\delta \in \{-1,1\}^n$ and $y \in \mathbb{Z}_{4m}^T$.

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

$$\begin{aligned}
&\sum_{x\in\mathbb{Z}_{4m}^S} \|A^{(S)}f_y(x+2\delta_S) - A^{(S)}f_y(x)\|_X^p \\
&\quad = \sum_{x\in\mathbb{Z}_{4m}^S} \|A^{(S)}f_y(x+\delta_S) - A^{(S)}f_y(x-\delta_S)\|_X^p \\
&\quad \lesssim_p \sum_{x\in\mathbb{Z}_{4m}^S} \left\| \sum_{j\in S} \delta_j [B_j^{(S)}f_y(x+e_j) - B_j^{(S)}f_y(x-e_j)] \right\|_X^p \\
&\qquad + p^p \sum_{s=0}^{k-1} \frac{(k/R)^{(k-s)p}}{\binom{k}{s}} \sum_{\substack{W\subseteq S\\ |W|=s}} \sum_{x\in\mathbb{Z}_{4m}^S} \|f_y(x+2\delta_W) - f_y(x)\|_X^p. \qquad (90)
\end{aligned}$$

By summing (90) over $\delta \in \{-1,1\}^n$ and $y \in \mathbb{Z}_{4m}^T$, while using the identities (89), we see that

$$\begin{aligned}
&\sum_{\delta\in\{-1,1\}^n} \sum_{z\in\mathbb{Z}_{4m}^n} \|D_S f(z+2\delta_S) - D_S f(z)\|_X^p \\
&\quad \lesssim_p \sum_{\delta\in\{-1,1\}^n} \sum_{z\in\mathbb{Z}_{4m}^n} \left\| \sum_{j\in S} \delta_j [B_j f(z+e_j) - B_j f(x-e_j)] \right\|_X^p \\
&\qquad + p^p \sum_{s=0}^{k-1} \frac{(k/R)^{(k-s)p}}{\binom{k}{s}} \\
&\qquad \times \sum_{\substack{W\subseteq S\\ |W|=s}} \sum_{\delta\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^S} \sum_{y\in\mathbb{Z}_{4m}^T} \|f_y(x+2\delta_W) - f_y(x)\|_X^p. \qquad (91)
\end{aligned}$$

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

$$\sum_{x\in\mathbb{Z}_{4m}^S} \sum_{y\in\mathbb{Z}_{4m}^T} \|f_y(x+2\delta_W) - f_y(x)\|_X^p \leqslant \sum_{z\in\mathbb{Z}_{4m}^n} \|f(z+2\delta_W) - f(z)\|_X^p.$$

Consequently, using Lemma 4.4 and the assumption $R \geqslant 2k-1$, the final term in (91) is at most $(O(1)p)^p$ times the following quantity

$$\begin{aligned}
&\left(\sum_{s=0}^{k-1} \left(\frac{k}{R}\right)^{(k-s)p} \right) \sum_{\varepsilon\in\{-1,1\}^n} \sum_{z\in\mathbb{Z}_{4m}^n} \|f(z+2\varepsilon) - f(z)\|_X^p \\
&\qquad \lesssim_p \left(\frac{k}{R}\right)^p \sum_{\varepsilon\in\{-1,1\}^n} \sum_{z\in\mathbb{Z}_{4m}^n} \|f(z+\varepsilon) - f(z)\|_X^p.
\end{aligned}$$

Hence (91) implies the desired inequality (88). □

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

Proof of Theorem 1.2. From now on, choose R to be the smallest odd integer that is greater than pn, and suppose that

$$m \geqslant \frac{n^{3/2}\log p}{\sqrt{k}} + pn.$$

In particular, we have $2n \leqslant R \leqslant 2m$. Fix $x \in \mathbb{Z}_{4m}^n$ and apply inequality (16) to the scalars $a_j = B_j f(x+e_j) - B_j f(x-e_j)$. The resulting estimate is

$$\begin{aligned}
&\frac{(p/\log p)^{-p}}{2^n\binom{n}{k}} \sum_{\substack{S\subseteq\{1,\ldots,n\}\\ |S|=k}} \sum_{\varepsilon\in\{-1,1\}^n} \left|\sum_{j\in S} \varepsilon_j [B_j f(x+e_j) - B_j f(x-e_j)]\right|^p \\
&\lesssim_p \frac{k}{n}\sum_{j=1}^n |B_j f(x+e_j) - B_j f(x-e_j)|^p \\
&\quad + \frac{(k/n)^{p/2}}{2^n} \sum_{\varepsilon\in\{-1,1\}^n} \left|\sum_{j=1}^n \varepsilon_j [B_j f(x+e_j) - B_j f(x-e_j)]\right|^p. \qquad (92)
\end{aligned}$$

By summing (92) over $x \in \mathbb{Z}_{4m}^n$ we deduce that

$$\begin{aligned}
&\frac{(p/\log p)^{-p}}{2^n\binom{n}{k}} \sum_{\substack{S\subseteq\{1,\ldots,n\}\\ |S|=k}} \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^n} \left|\sum_{j\in S} \varepsilon_j [B_j f(x+e_j) - B_j f(x-e_j)]\right|^p \\
&\lesssim_p \frac{k}{n}\sum_{x\in\mathbb{Z}_{4m}^n}\sum_{j=1}^n |B_j f(x+2e_j) - B_j f(x)|^p \\
&\quad + \frac{(k/n)^{p/2}}{2^n} \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^n} \left|\sum_{j=1}^n \varepsilon_j [B_j f(x+e_j) - B_j f(x-e_j)]\right|^p. \qquad (93)
\end{aligned}$$

For every $j \in \{1,\ldots,n\}$, since B_j is an averaging operator we have

$$\begin{aligned}
&\sum_{x\in\mathbb{Z}_{4m}^n} |B_j f(x+2e_j) - B_j f(x)|^p \\
&\quad \leqslant \sum_{x\in\mathbb{Z}_{4m}^n} |f(x+2e_j) - f(x)|^p \lesssim_p \sum_{x\in\mathbb{Z}_{4m}^n} |f(x+e_j) - f(x)|^p. \qquad (94)
\end{aligned}$$

Recalling that $R \geqslant pn$, by Lemma 4.5 we have

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

$$\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}^n_{4m}}\left|\sum_{j=1}^n \varepsilon_j[B_jf(x+e_j)-B_jf(x-e_j)]\right|^p$$
$$\lesssim_p \sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}^n_{4m}}|Af(x+2\varepsilon)-Af(x)|^p + \sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}^n_{4m}}|f(x+\varepsilon)-f(x)|^p$$
$$\leqslant \sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}^n_{4m}}|f(x+2\varepsilon)-f(x)|^p + \sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}^n_{4m}}|f(x+\varepsilon)-f(x)|^p \tag{95}$$
$$\lesssim_p \sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}^n_{4m}}|f(x+\varepsilon)-f(x)|^p, \tag{96}$$

where in (95) we used the fact that A is an averaging operator.

By substituting (94) and (96) into (93) we see that

$$\frac{(p/\log p)^{-p}}{2^n\binom{n}{k}}\sum_{\substack{S\subseteq\{1,\dots,n\}\\|S|=k}}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}^n_{4m}}\left|\sum_{j\in S}\varepsilon_j[B_jf(x+e_j)-B_jf(x-e_j)]\right|^p$$
$$\lesssim_p \frac{k}{n}\sum_{j=1}^n\sum_{x\in\mathbb{Z}^n_{4m}}|f(x+e_j)-f(x)|^p + \frac{1}{2^n}\left(\frac{k}{n}\right)^{p/2} \times \sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}^n_{4m}}|f(x+\varepsilon)-f(x)|^p. \tag{97}$$

By averaging (88) over all $S\subseteq\{1,\dots,n\}$ with $|S|=k$ and substituting (97) into the resulting inequality, we obtain the following estimate.

$$\frac{(p/\log p)^{-p}}{2^n\binom{n}{k}}\sum_{\substack{S\subseteq\{1,\dots,n\}\\|S|=k}}\sum_{\delta\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}^n_{4m}}|D_Sf(x+\delta_S)-D_Sf(x-\delta_S)|^p$$
$$\lesssim_p \frac{k}{n}\sum_{j=1}^n\sum_{x\in\mathbb{Z}^n_{4m}}|f(x+e_j)-f(x)|^p + \frac{1}{2^n}\left(\frac{k}{n}\right)^{p/2} \times \sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}^n_{4m}}|f(x+\varepsilon)-f(x)|^p. \tag{98}$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

Next, average (68) over $\delta \in \{-1,1\}^n$ and substitute (98) into the resulting inequality, thus obtaining the following estimate (recall that in the present setting $R \lesssim pn$).

$$\begin{aligned}
&\frac{1}{2^n\binom{n}{k}} \sum_{\substack{S\subseteq\{1,\ldots,n\}\\|S|=k}} \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^n} \frac{|f(x+2m\varepsilon_S)-f(x)|^p}{m^p} \\
&\lesssim_p \left(\frac{p^p}{(\log p)^p}+\frac{k^{p-1}}{m^p}\right)\frac{k}{n}\sum_{j=1}^n \sum_{x\in\mathbb{Z}_{4m}^n} |f(x+e_j)-f(x)|^p \\
&\quad+\left(\frac{p^p}{(\log p)^p}+\frac{(pn)^p}{m^p}\left(\frac{n}{k}\right)^{p/2}\right)\frac{1}{2^n}\left(\frac{k}{n}\right)^{p/2} \\
&\quad\times \sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{4m}^n} |f(x+\varepsilon)-f(x)|^p. \qquad (99)
\end{aligned}$$

Since $m \geqslant (n^{3/2}\log p)/\sqrt{k}$, the desired inequality (4) is a consequence of (99). □

5. Proof of Theorem 1.6

The desired inequality (23) is equivalent to the conjunction of the following two inequalities.

$$\begin{aligned}
&\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n} |f(x+2\varepsilon)-f(x)|^p \\
&\lesssim_p \frac{(pn/k)^{p/2}}{\binom{n}{k}} \sum_{\substack{S\subseteq\{1,\ldots,n\}\\|S|=k}} \sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n} |f(x+\varepsilon_S)-f(x)|^p, \qquad (100)
\end{aligned}$$

and

$$\begin{aligned}
&\sum_{j=1}^n \sum_{x\in\mathbb{Z}_{8m}^n} \frac{|f(x+4me_j)-f(x)|^p}{m^p} \\
&\lesssim_p \frac{p^{p/2}}{2^n\binom{n-1}{k-1}} \sum_{\substack{S\subseteq\{1,\ldots,n\}\\|S|=k}} \sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n} |f(x+\varepsilon_S)-f(x)|^p. \qquad (101)
\end{aligned}$$

The proofs of (100) and (101) are of a different nature: (100) is related to metric type and (101) is related to metric cotype. We therefore treat (100) and (101) in separate subsections.

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

5.1. Metric type and proof of (100). For every $n \in \mathbb{N}$ and $j \in \{1, \dots, n\}$ let $\sigma^j \in \{-1, 1\}^n$ be given by

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

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

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

Suppose that $(X, \|\cdot\|_X)$ is a Banach space and that $p \in [1, \infty]$. Slightly abusing notation that was introduced in [39], 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

$$\left(\frac{1}{2^n} \sum_{\varepsilon \in \{-1,1\}^n} \|h(\varepsilon) - h(-\varepsilon)\|_X^p\right)^{1/p} \leqslant \mathfrak{P}\left(\frac{1}{4^n} \sum_{\varepsilon,\delta \in \{-1,1\}^n} \left\|\sum_{j=1}^n \delta_j [h(\sigma^j \varepsilon) - h(\varepsilon)]\right\|_X^p\right)^{1/p}. \tag{102}$$

Note that in [39] the quantity $\mathfrak{P}_p^n(X)$ denotes the best constant in an inequality that is stronger than but closely related to (102). 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 his work on metric type, Pisier proved in [79] that $\mathfrak{P}_p^n(X) \lesssim \log n$ for every Banach space $(X, \|\cdot\|_X)$. In order to prove (100), we will deal with $X = \mathbb{R}$, in which case it will be important that $\sup_{n \in \mathbb{N}} \mathfrak{P}_p^n(\mathbb{R}) < \infty$. This strengthening of Pisier's inequality for real-valued functions is due to Talagrand [86], who proved that $\sup_{n \in \mathbb{N}} \mathfrak{P}_p^n(\mathbb{R}) \leqslant K^p$ for some universal constant $K \in (1, \infty)$, an estimate that was later improved in [72] to $\sup_{n \in \mathbb{N}} \mathfrak{P}_p^n(\mathbb{R}) \lesssim p$. The rate of growth of $\sup_{n \in \mathbb{N}} \mathfrak{P}_p^n(\mathbb{R})$ as $p \to \infty$ remains unknown, the best available lower bound, due to Talagrand [86], being that $\sup_{n \in \mathbb{N}} \mathfrak{P}_p^n(\mathbb{R})$ is at least a constant multiple of $\log p$. We refer to [39, 72, 88] for additional classes of Banach space $(X, \|\cdot\|_X)$ for which $\sup_{n \in \mathbb{N}} \mathfrak{P}_p^n(X) < \infty$.

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

$$\sum_{\varepsilon \in \{-1,1\}^n} d_X(h(\varepsilon), h(-\varepsilon))^p \leqslant B^p n^{p/q-1} \sum_{j=1}^n \sum_{\varepsilon \in \{-1,1\}^n} d_X(h(\sigma^j \varepsilon), h(\varepsilon))^p. \tag{103}$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

The quantity $\mathrm{BMW}_q^n(X;p)$ 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 *et al.* [19], though, as we explained in Section 1, the case $p=q$ was previously introduced by Enflo [31] and Gromov [38] (Gromov only dealt with the case $p=q=2$). It follows from (102) and Hölder's inequality that if $(X,\|\cdot\|_X)$ is a Banach space then

$$\mathrm{BMW}_q^n(X;p) \leqslant \mathfrak{P}_p^n(X)\cdot T_q^n(X;p) \lesssim (\log n)\cdot T_q^n(X;p), \tag{104}$$

where the (Rademacher type q) constant $T_q^n(X;p)$ 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

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

Since for many Banach spaces $(X,\|\cdot\|_X)$ good estimates on $T_q^n(X;p)$ are known, in conjunction with the available bounds on $\mathfrak{P}_p^n(X)$, inequality (104) often yields a satisfactory estimate on $\mathrm{BMW}_q^n(X;p)$. Such an estimate will be relevant to the ensuing proof of a metric-space-valued extension of (100). There are also several important classes of (non-Banach) metric spaces (X,d_X) for which good bounds on $\mathrm{BMW}_q^n(X;p)$ have been obtained; see for example [67, 71–74]. When $X=\mathbb{R}$, a bound that is even better than what follows from (104) is known: see inequality (6.32) in [67], which yields the estimate

$$\sup_{n\in\mathbb{N}} \mathrm{BMW}_2^n(\mathbb{R};p) \lesssim \sqrt{p}. \tag{105}$$

The following lemma, in conjunction with (105), implies (100). Note that there is no requirement that m is sufficiently large here: the lower bound on m that is assumed in Theorem 1.6 will be needed only for the proof of (101).

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 $f:\mathbb{Z}_{8m}^n\to X$ we have*

$$\begin{aligned}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n} d_X(f(x+2\varepsilon),f(x))^p &\lesssim_p (\mathrm{BMW}_q^{\lfloor n/k\rfloor+1}(X;p))^p \frac{(n/k)^{p/q}}{\binom{n}{k}}\\ &\quad\times \sum_{\substack{S\subseteq\{1,\ldots,n\}\\ |S|=k}}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n} d_X(f(x+\varepsilon_S),f(x))^p.\end{aligned} \tag{106}$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

Proof. Write $n = ak + b$ where $a = \lfloor n/k \rfloor$ and $b \in \{0, \dots, k-1\}$. For every $j \in \{1, \dots, a\}$ define $I_j = \{(j-1)k+1, \dots, jk\}$, and also define $I_{a+1} = \{ak+1, \dots, ak+b\}$. Fix $x \in \mathbb{Z}_{8m}^n$ 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

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

Note that for every $\pi \in S_n$, every $x \in \mathbb{Z}_{8m}^n$ and every $\delta \in \{-1,1\}^{a+1}$ we have

$$\begin{aligned}\sum_{\varepsilon\in\{-1,1\}^n} d_X(h_{x,\varepsilon}^\pi(\delta), h_{x,\varepsilon}^\pi(-\delta))^p &= \sum_{\varepsilon\in\{-1,1\}^n} d_X(f(x+\varepsilon), f(x-\varepsilon))^p \\ &= \sum_{\varepsilon\in\{-1,1\}^n} d_X(f(x+2\varepsilon), f(x))^p. \qquad (107)\end{aligned}$$

Also, for every $\pi \in S_n$ and $j \in \{1, \dots, n\}$ we have

$$\begin{aligned}&\frac{1}{2^{a+1}} \sum_{\varepsilon\in\{-1,1\}^n} \sum_{\delta\in\{-1,1\}^{a+1}} \sum_{x\in\mathbb{Z}_{8m}^n} d_X(h_{x,\varepsilon}^\pi(\sigma^j\delta), h_{x,\varepsilon}^\pi(\delta))^p \\ &= \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{8m}^n} d_X(f(x+\varepsilon_{\pi(I_j)}), f(x-\varepsilon_{\pi(I_j)}))^p \\ &= \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{8m}^n} d_X(f(x+2\varepsilon_{\pi(I_j)}), f(x))^p. \qquad (108)\end{aligned}$$

Fix $B > \mathrm{BMW}_q^{a+1}(X; p)$, apply (103) to $h_{x,\varepsilon}^\pi$, and sum the resulting inequality over $\varepsilon \in \{-1,1\}^n$ and $x \in \mathbb{Z}_{8m}^n$, while using the identities (107) and (108). The resulting inequality is

$$\begin{aligned}&\sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{8m}^n} d_X(f(x+2\varepsilon), f(x))^p \\ &\leqslant B^p (a+1)^{p/q-1} \sum_{j=1}^{a+1} \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{8m}^n} d_X(f(x+2\varepsilon_{\pi(I_j)}), f(x))^p. \qquad (109)\end{aligned}$$

By averaging (109) over $\pi \in S_n$ we see that

$$\begin{aligned}&B^{-p}(a+1)^{1-p/q} \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{8m}^n} d_X(f(x+2\varepsilon), f(x))^p \\ &\leqslant \sum_{j=1}^{a} \sum_{\substack{S\subseteq\{1,\dots,n\} \\ |S|=k}} \frac{|\{\pi \in S_n : \pi(I_j) = S\}|}{n!} \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{8m}^n} d_X(f(x+2\varepsilon_S), f(x))^p\end{aligned}$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

$$+\sum_{\substack{T\subseteq\{1,\ldots,n\}\\|T|=b}}\frac{|\{\pi\in S_n:\pi(I_{a+1})=T\}|}{n!}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}d_X(f(x+2\varepsilon_T),f(x))^p$$

$$=\frac{a}{\binom{n}{k}}\sum_{\substack{S\subseteq\{1,\ldots,n\}\\|S|=k}}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}d_X(f(x+2\varepsilon_S),f(x))^p$$

$$+\frac{1}{\binom{n}{b}}\sum_{\substack{T\subseteq\{1,\ldots,n\}\\|T|=b}}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}d_X(f(x+2\varepsilon_T),f(x))^p. \tag{110}$$

By Lemma 4.4, if $T\subseteq S\subseteq\{1,\ldots,n\}$ then

$$\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}d_X(f(x+2\varepsilon_T),f(x))^p\leqslant 2^p\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}d_X(f(x+\varepsilon_S),f(x))^p. \tag{111}$$

Fixing $T\subseteq\{1,\ldots,n\}$ with $|T|=b$, by averaging (111) over all k-point subsets $S\subseteq\{1,\ldots,n\}$ with $S\supseteq T$ we see that

$$\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}d_X(f(x+2\varepsilon_T),f(x))^p$$

$$\leqslant\frac{2^p}{\binom{n-b}{k-b}}\sum_{T\subseteq S\subseteq\{1,\ldots,n\}\atop|S|=k}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}d_X(f(x+\varepsilon_S),f(x))^p.$$

Consequently,

$$\frac{1}{\binom{n}{b}}\sum_{\substack{T\subseteq\{1,\ldots,n\}\\|T|=b}}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}d_X(f(x+2\varepsilon_T),f(x))^p$$

$$\leqslant\frac{2^p}{\binom{n}{b}\binom{n-b}{k-b}}\sum_{\substack{S\subseteq\{1,\ldots,n\}\\|S|=k}}|\{T\subseteq S:|T|=b\}|\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}d_X(f(x+\varepsilon_S),f(x))^p$$

$$=\frac{2^p\binom{k}{b}}{\binom{n}{b}\binom{n-b}{k-b}}\sum_{\substack{S\subseteq\{1,\ldots,n\}\\|S|=k}}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}d_X(f(x+\varepsilon_S),f(x))^p$$

$$=\frac{2^p}{\binom{n}{k}}\sum_{\substack{S\subseteq\{1,\ldots,n\}\\|S|=k}}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}d_X(f(x+\varepsilon_S),f(x))^p. \tag{112}$$

Since by the triangle inequality we also have

$$\sum_{\substack{S\subseteq\{1,\ldots,n\}\\|S|=k}}\sum_{x\in\mathbb{Z}_{8m}^n}d_X(f(x+2\varepsilon_S),f(x))^p\leqslant 2^p\sum_{\substack{S\subseteq\{1,\ldots,n\}\\|S|=k}}\sum_{x\in\mathbb{Z}_{8m}^n}d_X(f(x+\varepsilon_S),f(x))^p,$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

it follows from (110) and (112) that

$$\begin{aligned}&\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n} d_X(f(x+2\varepsilon),\,f(x))^p\\&\leqslant \frac{(2B)^p(a+1)^{p/q}}{\binom{n}{k}}\sum_{\substack{S\subseteq\{1,\ldots,n\}\\|S|=k}}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n} d_X(f(x+\varepsilon_S),\,f(x))^p\\&\leqslant \frac{(2B)^p(2n/k)^{p/q}}{\binom{n}{k}}\sum_{\substack{S\subseteq\{1,\ldots,n\}\\|S|=k}}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n} d_X(f(x+\varepsilon_S),\,f(x))^p. \qquad \square\end{aligned}$$

5.2. Metric cotype and proof of (101). 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 $f:\mathbb{Z}_{2m}^n\to X$,

$$\sum_{j=1}^n\sum_{x\in\mathbb{Z}_{2m}^n}\frac{d_X(f(x+me_j),\,f(x))^p}{m^p}\leqslant\frac{\Gamma^p}{3^n}\sum_{\varepsilon\in\{-1,0,1\}^n}\sum_{x\in\mathbb{Z}_{2m}^n} d_X(f(x+\varepsilon),\,f(x))^p. \tag{113}$$

As discussed in the Introduction, following [61], we say that (X,d_X) has metric cotype p if

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

We need to briefly recall some facts related to K-convexity of Banach spaces; see the survey [56] 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 $\mathbf{Rad}(f):\{-1,1\}^n\to X$ by

$$\forall\,\varepsilon\in\{-1,1\}^n,\qquad \mathbf{Rad}(f)(\varepsilon)\stackrel{\text{def}}{=}\sum_{j=1}^n\frac{\sum_{\delta\in\{-1,1\}^n}f(\delta)\delta_j}{2^n}\varepsilon_j.$$

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

$$\sum_{\varepsilon\in\{-1,1\}^n}\|\mathbf{Rad}(f)(\varepsilon)\|_X^p\leqslant K^p\sum_{\varepsilon\in\{-1,1\}^n}\|f(\varepsilon)\|_X^p.$$

A simple application of Khintchine's inequality (with asymptotically sharp constant, see [77, Lemma 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 [56] and the references therein.

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

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 $x \in \mathbb{Z}_{4m}^n$ rather than $x \in \mathbb{Z}_{8m}^n$) is the content of [61, Theorem 4.1]. The proof here follows the argument in [61] 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 [34, 61], though this may require changes to the arguments of [34, 61] 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*

$$m \geqslant \frac{n^{1/p}}{\alpha K_p(X) C_p(X)}, \tag{114}$$

where, recalling (5), $C_p(X)$ is the cotype p constant of X. Then for every $f : \mathbb{Z}_{8m}^n \to X$ we have

$$\sum_{j=1}^n \sum_{x \in \mathbb{Z}_{8m}^n} \frac{\|f(x+4me_j) - f(x)\|_X^p}{m^p} \lesssim_p \frac{(\alpha K_p(X) C_p(X))^p}{2^n} \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_{8m}^n} \|f(x+\varepsilon) - f(x)\|_X^p. \tag{115}$$

Before proving Theorem 5.2 we deduce the following simple corollary, which implies (101) 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*

$$m \geqslant \frac{k^{1/p}}{\alpha K_p(X) C_p(X)}.$$

Then for every $f : \mathbb{Z}_{8m}^n \to X$ we have

$$\sum_{j=1}^n \sum_{x \in \mathbb{Z}_{8m}^n} \frac{\|f(x+4me_j) - f(x)\|_X^p}{m^p} \lesssim_p \frac{(\alpha K_p(X) C_p(X))^p}{2^n \binom{n-1}{k-1}} \sum_{\substack{S \subseteq \{1,\ldots,n\} \\ |S|=k}} \sum_{\varepsilon \in \{-1,1\}^n} \sum_{x \in \mathbb{Z}_{8m}^n} \|f(x+\varepsilon_S) - f(x)\|_X^p. \tag{116}$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

Proof. By Theorem 5.2, for every $S \subseteq \{1,\dots,n\}$ with $|S| = k$ we have

$$\sum_{j\in S}\sum_{x\in\mathbb{Z}_{8m}^n}\frac{\|f(x+4me_j)-f(x)\|_X^p}{m^p} \lesssim_p \frac{(\alpha K_p(X)C_p(X))^p}{2^n}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}\|f(x+\varepsilon_S)-f(x)\|_X^p.$$

By averaging this inequality over all $S \subseteq \{1,\dots,n\}$ with $|S| = k$ we obtain (116). □

In order to prove Theorem 5.2, we first introduce a small amount of notation and prove an auxiliary lemma. For every $j \in \{1,\dots,n\}$, define a linear operator $T_j : L_2(\mathbb{Z}_{8m}^n, X) \to L_2(\mathbb{Z}_{8m}^n, X)$ by setting for every $f : \mathbb{Z}_{8m}^n \to X$ and $x \in \mathbb{Z}_{8m}^n$,

$$T_j f(x) \stackrel{\text{def}}{=} \frac{1}{2^n}\sum_{\varepsilon\in\{-1,1\}^n} f(x+2\varepsilon_{\{1,\dots,n\}\setminus\{j\}}). \tag{117}$$

LEMMA 5.4. *Let $(X, \|\cdot\|_X)$ be a Banach space and $p \in [1,\infty)$. Fix also m, $n \in \mathbb{N}$ and $j \in \{1,\dots,n\}$. Then for every $f : \mathbb{Z}_{8m}^n \to X$ we have*

$$\sum_{x\in\mathbb{Z}_{8m}^n}\|f(x)-T_jf(x)\|_X^p \leqslant \frac{2^p}{2^n}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}\|f(x+\varepsilon)-f(x)\|_X^p. \tag{118}$$

Moreover, for every $x \in \mathbb{Z}_{8m}^n$ we have

$$\begin{aligned}\sum_{\varepsilon\in\{-1,1\}^n}\Big\|\sum_{j=1}^n \varepsilon_j[T_jf(x+2e_j)-T_jf(x-2e_j)]\Big\|_X^p &\leqslant (2K_p(X))^p\sum_{\varepsilon\in\{-1,1\}^n}\|f(x+2\varepsilon)-f(x)\|_X^p\\ &\leqslant (4K_p(X))^p\sum_{\varepsilon\in\{-1,1\}^n}\|f(x+\varepsilon)-f(x)\|_X^p.\end{aligned} \tag{119}$$

Proof. By the definition (117) and convexity,

$$\begin{aligned}\sum_{x\in\mathbb{Z}_{8m}^n}\|f(x)-T_jf(x)\|_X^p &\leqslant \frac{1}{2^n}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}\|f(x)-f(x+2\varepsilon_{\{1,\dots,n\}\setminus\{j\}})\|_X^p\\ &\leqslant \frac{2^p}{2^n}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}\|f(x+\varepsilon)-f(x)\|_X^p,\end{aligned}$$

where in the last step we used Lemma 4.4. This proves (118).

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

To prove (119), for every fixed $x \in \mathbb{Z}_{8m}^n$ define $h^x : \{-1,1\}^n \to X$ by

$$\forall\, \varepsilon \in \{-1,1\}^n, \quad h^x(\varepsilon) \stackrel{\text{def}}{=} f(x+2\varepsilon) - f(x).$$

We claim that the following identity holds true.

$$\mathbf{Rad}(h^x)(\varepsilon) = \frac{i}{2}\sum_{j=1}^n \varepsilon_j [T_j f(x+2e_j) - T_j f(x-2e_j)]. \tag{120}$$

Once (120) is proved, the desired inequality (119) would follow from the definition of $K_p(X)$.

By composing with linear functionals, it suffices to verify the validity of (120) when $X = \mathbb{C}$. Moreover, for every $y \in \mathbb{Z}_{8m}^n$ define $W_y : \mathbb{Z}_{8m}^n \to \mathbb{C}$ by

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

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 (120) when $f = W_y$ for some $y \in \mathbb{Z}_{8m}^n$. Now,

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

Consequently,

$$\mathbf{Rad}(W_y^x)(\varepsilon) = i\left(\sum_{j=1}^n \varepsilon_j \sin\left(\frac{\pi y_j}{2m}\right) \prod_{s\in\{1,\ldots,n\}\setminus\{j\}} \cos\left(\frac{\pi y_s}{2m}\right)\right) W_y(x). \tag{121}$$

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

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

Therefore,

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

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

where in the last step we used (121). □

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

$$\begin{aligned} &\|f(x+4me_j) - f(x)\|_X^p \\ &\lesssim_p \|T_j f(x+4me_j) - T_j f(x)\|_X^p \\ &\quad + \|f(x+4me_j) - T_j f(x+4me_j)\|_X^p + \|f(x) - T_j f(x)\|_X^p. \end{aligned}$$

Hence, using (118) we see that

$$\begin{aligned} &\sum_{x\in\mathbb{Z}_{8m}^n} \|f(x+4me_j) - f(x)\|_X^p \\ &\lesssim_p \sum_{x\in\mathbb{Z}_{8m}^n} \|T_j f(x+4me_j) - T_j f(x)\|_X^p \\ &\quad + \frac{1}{2^n} \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{8m}^n} \|f(x+\varepsilon) - f(x)\|_X^p. \end{aligned} \tag{122}$$

By the triangle inequality combined with Hölder's inequality we have

$$\begin{aligned} &\sum_{x\in\mathbb{Z}_{8m}^n} \|T_j f(x+4me_j) - T_j f(x)\|_X^p \\ &\leqslant m^{p-1} \sum_{s=1}^{m} \sum_{x\in\mathbb{Z}_{8m}^n} \|T_j f(x+4se_j) - T_j f(x+4(s-1)e_j)\|_X^p \\ &= m^p \sum_{x\in\mathbb{Z}_{8m}^n} \|T_j f(x+2e_j) - T_j f(x-2e_j)\|_X^p. \end{aligned}$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

In combination with (122), this implies that

$$\sum_{j=1}^{n}\sum_{x\in\mathbb{Z}_{8m}^n}\frac{\|f(x+4me_j)-f(x)\|_X^p}{m^p}$$
$$\lesssim_p \sum_{j=1}^{n}\sum_{x\in\mathbb{Z}_{8m}^n}\|T_jf(x+2e_j)-T_jf(x-2e_j)\|_X^p$$
$$+\frac{n}{m^p2^n}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}\|f(x+\varepsilon)-f(x)\|_X^p. \tag{123}$$

By the definition of the cotype p constant $C_p(X)$, for every $x\in\mathbb{Z}_{8m}^n$ we have

$$\sum_{j=1}^{n}\|T_jf(x+2e_j)-T_jf(x-2e_j)\|_X^p$$
$$\leqslant \frac{C_p(X)^p}{2^n}\sum_{\varepsilon\in\{-1,1\}^n}\left\|\sum_{j=1}^{n}\varepsilon_j[T_jf(x+2e_j)-T_jf(x-2e_j)]\right\|_X^p$$
$$\lesssim_p \frac{(K_p(X)C_p(X))^p}{2^n}\sum_{\varepsilon\in\{-1,1\}^n}\|f(x+\varepsilon)-f(x)\|_X^p, \tag{124}$$

where in the last step of (124) we used (119). By substituting (124) into (123) we conclude that

$$\sum_{j=1}^{n}\sum_{x\in\mathbb{Z}_{8m}^n}\frac{\|f(x+4me_j)-f(x)\|_X^p}{m^p}$$
$$\lesssim_p \frac{(K_p(X)C_p(X))^p+n/m^p}{2^n}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}\|f(x+\varepsilon)-f(x)\|_X^p. \tag{125}$$

Due to (114), (125) implies the desired inequality (115). □

6. A conjectural convolution inequality as a way to prove Conjecture 1.5

For every $j\in\{1,\ldots,n\}$, define an averaging operator $E_j:L_2(\mathbb{Z}_m^n)\to L_2(\mathbb{Z}_m^n)$ by setting for every $f:\mathbb{Z}_m^n\to\mathbb{R}$ and $x\in\mathbb{Z}_m^n$,

$$E_jf(x)\stackrel{\text{def}}{=}\frac{f(x+e_j)+f(x-e_j)}{2}.$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

We also set $\mathcal{E}_j \stackrel{\text{def}}{=} \prod_{s\in\{1,\ldots,n\}\setminus\{j\}} E_s$ and $\mathcal{E} \stackrel{\text{def}}{=} \prod_{s=1}^n E_j$. Thus, for every $f : \mathbb{Z}_m^n \to \mathbb{R}$ and $x \in \mathbb{Z}_m^n$,

$$\begin{aligned} \mathcal{E}_j f(x) &= \frac{1}{2^n} \sum_{\varepsilon\in\{-1,1\}^n} f(x + \varepsilon_{\{1,\ldots,n\}\setminus\{j\}}), \quad \text{and} \\ \mathcal{E} f(x) &= \frac{1}{2^n} \sum_{\varepsilon\in\{-1,1\}^n} f(x+\varepsilon). \end{aligned} \tag{126}$$

QUESTION 6.1. Is it true that for every $p \in (2,\infty)$ there exists $\beta_p \in (0,1]$ such that for every $m, n \in \mathbb{N}$, every $f : \mathbb{Z}_m^n \to \mathbb{R}$ satisfies

$$\begin{aligned} &\frac{\beta_p}{2^n} \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_m^n} |\mathcal{E} f(x+\varepsilon) - \mathcal{E} f(x-\varepsilon)|^p \\ &\quad\leqslant \frac{1}{2^n} \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_m^n} \left| \sum_{j=1}^n \varepsilon_j [\mathcal{E}_j f(x+e_j) - \mathcal{E}_j f(x-e_j)] \right|^p \\ &\qquad + \sum_{j=1}^n \sum_{x\in\mathbb{Z}_m^n} |f(x+e_j) - f(x)|^p. \end{aligned} \tag{127}$$

It may very well be the case that (127) holds true without the second term that appears in the right-hand side, that is, that

$$\begin{aligned} &\beta_p \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_m^n} |\mathcal{E} f(x+\varepsilon) - \mathcal{E} f(x-\varepsilon)|^p \\ &\quad\leqslant \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_m^n} \left| \sum_{j=1}^n \varepsilon_j [\mathcal{E}_j f(x+e_j) - \mathcal{E}_j f(x-e_j)] \right|^p. \end{aligned}$$

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 1.5 holds true, and hence also that all the conclusions of Theorem 1.14 hold true. Specifically, for every $\delta \in (0,\infty)$, if $m, n \in \mathbb{N}$ and $k \in \{1,\ldots,n\}$ satisfy $m \geqslant \delta\sqrt{n/k}$ then (21) holds true with*

$$\alpha_p \gtrsim_p \min\left\{ \beta_p \left(\frac{\log p}{p^{3/2}} \right)^p, \delta^p \right\},$$

where β_p is as in (127).

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

Proof. Fix $f : \mathbb{Z}_{4m}^n \to \mathbb{R}$. By convexity, it follows from (126) that

$$\sum_{x\in\mathbb{Z}_{4m}^n} |f(x) - \mathcal{E}f(x)|^p \leqslant \frac{1}{2^n} \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^n} |f(x+\varepsilon) - f(x)|^p.$$

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

$$\begin{aligned}
&\sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^n} |f(x+2m\varepsilon_S) - f(x)|^p \\
&\lesssim_p \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^n} |\mathcal{E}f(x+2m\varepsilon_S) - \mathcal{E}f(x)|^p \\
&\quad + \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^n} |f(x+\varepsilon) - f(x)|^p. \qquad (128)
\end{aligned}$$

Arguing as in (70), it follows from the triangle inequality that

$$\begin{aligned}
&\sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^n} |\mathcal{E}f(x+2m\varepsilon_S) - \mathcal{E}f(x)|^p \\
&\leqslant m^p \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^n} |\mathcal{E}f(x+\varepsilon_S) - \mathcal{E}f(x-\varepsilon_S)|^p. \qquad (129)
\end{aligned}$$

For every $z \in \mathbb{Z}_{4m}^{\{1,\ldots,n\}\setminus S}$ apply (127) to the mapping

$$(y \in \mathbb{Z}_{4m}^S) \mapsto \prod_{j\in\{1,\ldots,n\}\setminus S} E_j f(y,z),$$

and then average the resulting inequality over z. The estimate thus obtained is

$$\begin{aligned}
&\frac{\beta_p}{2^n} \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^n} |\mathcal{E}f(x+\varepsilon_S) - \mathcal{E}f(x-\varepsilon_S)|^p \\
&\leqslant \frac{1}{2^n} \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^n} \left| \sum_{j\in S} \varepsilon_j [\mathcal{E}_j f(x+e_j) - \mathcal{E}_j f(x-e_j)] \right|^p \\
&\quad + \sum_{j\in S} \sum_{x\in\mathbb{Z}_{4m}^n} |f(x+e_j) - f(x)|^p. \qquad (130)
\end{aligned}$$

By averaging (130) over those $S \subseteq \{1,\ldots,n\}$ with $|S| = k$ we see that

$$\frac{\beta_p}{2^n \binom{n}{k}} \sum_{\substack{S\subseteq\{1,\ldots,n\} \\ |S|=k}} \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^n} |\mathcal{E}f(x+\varepsilon_S) - \mathcal{E}f(x-\varepsilon_S)|^p$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

$$\leqslant \frac{1}{2^n\binom{n}{k}} \sum_{\substack{S\subseteq\{1,\dots,n\}\\|S|=k}} \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^n} \left|\sum_{j\in S} \varepsilon_j[\mathcal{E}_j f(x+e_j) - \mathcal{E}_j f(x-e_j)]\right|^p$$
$$+ \frac{k}{n}\sum_{j=1}^n \sum_{x\in\mathbb{Z}_{4m}^n} |f(x+e_j) - f(x)|^p. \tag{131}$$

Note that since $\mathcal{E}_j$ is an averaging operator,

$$\sum_{j=1}^n \sum_{x\in\mathbb{Z}_{4m}^n} |\mathcal{E}_j f(x+e_j) - \mathcal{E}_j f(x-e_j)|^p \leqslant 2^p \sum_{j=1}^n \sum_{x\in\mathbb{Z}_{4m}^n} |f(x+e_j) - f(x)|^p.$$

Hence, using the linear X_p inequality (16), we deduce that

$$\frac{(p/\log p)^{-p}}{2^n\binom{n}{k}} \sum_{\substack{S\subseteq\{1,\dots,n\}\\|S|=k}} \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^n} \left|\sum_{j\in S} \varepsilon_j[\mathcal{E}_j f(x+e_j) - \mathcal{E}_j f(x-e_j)]\right|^p$$
$$\lesssim_p \frac{k}{n}\sum_{j=1}^n \sum_{x\in\mathbb{Z}_{4m}^n} |f(x+e_j) - f(x)|^p$$
$$+ \frac{(k/n)^{p/2}}{2^n} \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^n} \left|\sum_{j=1}^n \varepsilon_j[\mathcal{E}_j f(x+e_j) - \mathcal{E}_j f(x-e_j)]\right|^p. \tag{132}$$

The same reasoning that leads to the identity (120) (alternatively, by [**61**, Section 5]) shows that if for fixed $x \in \mathbb{Z}_{4m}^n$ we define $g^x : \{-1,1\}^n \to \mathbb{R}$ by setting $g^x(\varepsilon) = f(x+\varepsilon) - f(x)$ for every $\varepsilon \in \{-1,1\}^n$, then that Rademacher projection of g^x satisfies

$$\mathbf{Rad}(g^x)(\varepsilon) = \frac{i}{2}\sum_{j=1}^n \varepsilon_j[\mathcal{E}_j f(x+e_j) - \mathcal{E}_j f(x-e_j)].$$

Hence, recalling that the K-convexity constant of $\mathbb{R}$ satisfies $K_p(\mathbb{R}) \lesssim \sqrt{p}$,

$$\sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^n} \left|\sum_{j=1}^n \varepsilon_j[\mathcal{E}_j f(x+e_j) - \mathcal{E}_j f(x-e_j)]\right|^p$$
$$\lesssim_p p^{p/2} \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^n} |f(x+\varepsilon) - f(x)|^p. \tag{133}$$

By combining (131) with (132) and (133) we have

$$\frac{(p^{3/2}/\log p)^{-p}\beta_p}{2^n\binom{n}{k}} \sum_{\substack{S\subseteq\{1,\dots,n\}\\|S|=k}} \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{4m}^n} |\mathcal{E} f(x+\varepsilon_S) - \mathcal{E} f(x-\varepsilon_S)|^p$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

$$\lesssim_p \frac{k}{n}\sum_{j=1}^{n}\sum_{x\in\mathbb{Z}_{4m}^n}|f(x+e_j)-f(x)|^p + \frac{(k/n)^{p/2}}{2^n}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{4m}^n}|f(x+\varepsilon)-f(x)|^p.$$

Recalling (128) and (129), we therefore have

$$\frac{(p^{3/2}/\log p)^{-p}\beta_p}{2^n\binom{n}{k}}\sum_{\substack{S\subseteq\{1,\ldots,n\}\\|S|=k}}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{4m}^n}\frac{|f(x+2m\varepsilon_S)-f(x)|^p}{m^p}$$

$$\lesssim_p \frac{k}{n}\sum_{j=1}^{n}\sum_{x\in\mathbb{Z}_{4m}^n}|f(x+e_j)-f(x)|^p + \left(1+\frac{(p^{3/2}/\log p)^{-p}}{m^p(k/n)^{p/2}}\beta_p\right)\frac{(k/n)^{p/2}}{2^n} \times \sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{4m}^n}|f(x+\varepsilon)-f(x)|^p. \qquad \square$$

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

$$\|A\|_{S_p}=(\mathbf{Tr}((A^*A)^{p/2}))^{1/p}=(\mathbf{Tr}((AA^*)^{p/2}))^{1/p}.$$

See [81] 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 $A_1,\ldots,A_n\in M_d(\mathbb{R})$,*

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

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)

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

quantity $p/\log p$. This was proved in the scalar case in [41], but additional ideas seem to be required in order to carry out the proof of [41] 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 $k \in \{1,\dots,n\}$. Suppose that $B_1,\dots,B_n \in M_d(\mathbb{R})$ are symmetric and positive semidefinite. Then*

$$\frac{1}{\binom{n}{k}} \sum_{\substack{S\subseteq\{1,\dots,n\}\\ |S|=k}} \mathbf{Tr}\left(\left(\sum_{j\in S} B_j\right)^q\right) \lesssim_q \left(\frac{q}{\log(2q)}\right)^q \max\left\{\frac{k}{n}\sum_{j=1}^n \mathbf{Tr}(B_j^q), \left(\frac{k}{n}\right)^q \mathbf{Tr}\left(\left(\sum_{j=1}^n B_j\right)^q\right)\right\}.$$

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 Khintchine inequality [53] asserts that for every $S \subseteq \{1,\dots,n\}$ we have

$$\frac{p^{-p/2}}{2^n} \sum_{\varepsilon\in\{-1,1\}^n} \left\|\sum_{j\in S} \varepsilon_j A_j\right\|_{S_p}^p \lesssim_p \mathbf{Tr}\left(\left(\sum_{j\in S} A_j^* A_j\right)^{p/2}\right) + \mathbf{Tr}\left(\left(\sum_{j\in S} A_j A_j^*\right)^{p/2}\right). \tag{134}$$

The (asymptotically optimal) dependence on p in the left-hand side of (134) is not stated in Lust-Piquard's original proof of (134), but it can be found in [80, page 106]. By averaging (134) over all those $S \subseteq \{1,\dots,n\}$ with $|S| = k$ we see that

$$\begin{aligned}\frac{p^{-p/2}}{2^n\binom{n}{k}} \sum_{\substack{S\subseteq\{1,\dots,n\}\\ |S|=k}} \sum_{\varepsilon\in\{-1,1\}^n} \left\|\sum_{j\in S} \varepsilon_j A_j\right\|_{S_p}^p &\lesssim_p \frac{1}{\binom{n}{k}} \sum_{\substack{S\subseteq\{1,\dots,n\}\\ |S|=k}} \mathbf{Tr}\left(\left(\sum_{j\in S} A_j^* A_j\right)^{p/2}\right)\\ &\quad + \frac{1}{\binom{n}{k}} \sum_{\substack{S\subseteq\{1,\dots,n\}\\ |S|=k}} \mathbf{Tr}\left(\left(\sum_{j\in S} A_j A_j^*\right)^{p/2}\right).\end{aligned} \tag{135}$$

Two applications of Proposition 7.3 with $q = p/2 \geqslant 1$, 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

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

right-hand side of (135), yield

$$\begin{aligned}&\frac{(p/\sqrt{\log p})^{-p}}{2^n\binom{n}{k}}\sum_{\substack{S\subseteq\{1,\ldots,n\}\\|S|=k}}\sum_{\varepsilon\in\{-1,1\}^n}\left\|\sum_{j\in S}\varepsilon_j A_j\right\|_{S_p}^p\\&\lesssim_p \frac{k}{n}\sum_{j=1}^n\|A_j\|_{S_p}^p+\left(\frac{k}{n}\right)^{p/2}\mathbf{Tr}\left(\left(\sum_{j=1}^n A_j^*A_j\right)^{p/2}\right)\\&\quad+\left(\frac{k}{n}\right)^{p/2}\mathbf{Tr}\left(\left(\sum_{j=1}^n A_jA_j^*\right)^{p/2}\right).\end{aligned}\tag{136}$$

The other direction of Lust-Piquard's noncommutative Khintchine inequality [**53**] asserts that

$$\mathbf{Tr}\left(\left(\sum_{j=1}^n A_j^*A_j\right)^{p/2}\right)+\mathbf{Tr}\left(\left(\sum_{j=1}^n A_jA_j^*\right)^{p/2}\right)\lesssim_p\frac{1}{2^n}\sum_{\varepsilon\in\{-1,1\}^n}\left\|\sum_{j=1}^n\varepsilon_jA_j\right\|_{S_p}^p.\tag{137}$$

Theorem 7.1 now follows by combining (136) and (137). □

Lemma 7.4 below makes the same assertion as Proposition 7.3, but only for $k\leqslant n/2$ (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 $k\leqslant n/2$. Then for every $B_1,\ldots,B_n\in M_d(\mathbb{R})$ that are symmetric and positive semidefinite we have*

$$\begin{aligned}&\frac{1}{\binom{n}{k}}\sum_{\substack{S\subseteq\{1,\ldots,n\}\\|S|=k}}\mathbf{Tr}\left(\left(\sum_{j\in S}B_j\right)^q\right)\\&\lesssim\left(\frac{4q}{\log(2q)}\right)^q\max\left\{\frac{k}{n}\sum_{j=1}^n\mathbf{Tr}(B_j^q),\left(\frac{k}{n}\right)^q\mathbf{Tr}\left(\left(\sum_{j=1}^n B_j\right)^q\right)\right\}.\end{aligned}$$

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

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

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

we have

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

Consequently,

$$\begin{aligned}
&\frac{1}{\binom{n}{k}}\sum_{\substack{S\subseteq\{1,\ldots,n\}\\|S|=k}} \mathbf{Tr}\left(\left(\sum_{j\in S} B_j\right)^q\right)\\
&\quad\leqslant \frac{1}{\binom{n}{k}}\sum_{\substack{S\subseteq\{1,\ldots,n\}\\|S|=k}} \frac{2^{q-1}}{\binom{k}{u}}\sum_{\substack{T\subseteq S\\|T|=u}}\left(\left\|\sum_{s\in T} B_s\right\|_{S_q}^q + \left\|\sum_{s\in S\setminus T} B_s\right\|_{S_q}^q\right)\\
&\quad= \frac{2^{q-1}}{\binom{n}{u}}\sum_{\substack{U\subseteq\{1,\ldots,n\}\\|U|=u}} \mathbf{Tr}\left(\left(\sum_{j\in U} B_j\right)^q\right) + \frac{2^{q-1}}{\binom{n}{v}}\sum_{\substack{V\subseteq\{1,\ldots,n\}\\|V|=v}} \mathbf{Tr}\left(\left(\sum_{j\in V} B_j\right)^q\right). \quad (138)
\end{aligned}$$

Proposition 7.3 now follows by applying Lemma 7.4 to each of the summands that appear in the right-hand side of (138). □

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. Propositions 7.5 and 7.8 below will be used crucially in the proof of Lemma 7.4. Note that the same statements are trivial when matrixes are replaced by scalars. See Section 7.1.1 for a discussion on the context of these results, where it is explained in particular that Proposition 7.5 was known when $q \in [1, 2]$ by either directly applying the work of Carlen and Lieb [**24**], or through a simple argument that relies on operator convexity. At the same time, it is explained in Section 7.1.1 that when $q \in (0, 1) \cup (2, \infty)$, a range of values of q that is used crucially in our proof of Lemma 7.4 below, Proposition 7.5 exhibits a phenomenon that is qualitatively different from the simpler case $q \in [1, 2]$.

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*

$$(\mathbf{Tr}((A + B)^q A))^{1/q} \leqslant (\mathbf{Tr}(A^{q+1}))^{1/q} + (\mathbf{Tr}(B^q A))^{1/q}. \quad (139)$$

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

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

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 $b_j + b_{j+1} \leqslant 2q a_j$ for every $j \in \{0,\ldots,k-1\}$, where we set $b_0 = b_k$. Suppose also that*

$$\sum_{j=0}^{k-1} a_j + \sum_{j=1}^{k} b_j = q+1. \tag{140}$$

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

$$\mathbf{Tr}\left(A^{a_0}\left(\prod_{j=1}^{k-1} B^{b_j} A^{a_j}\right) B^{b_k}\right) \leqslant (\mathbf{Tr}(A^{q+1}))^{1-(1/q)\sum_{j=1}^k b_j} (\mathbf{Tr}(B^q A))^{(1/q)\sum_{j=1}^k b_j}.$$

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

$$\forall\, j \in \{0,\ldots,k-1\}, \quad p_j \stackrel{\text{def}}{=} \frac{q+1}{a_j - (b_j+b_{j+1})/2q} \quad \text{and} \quad r_j \stackrel{\text{def}}{=} \frac{q}{b_{j+1}}. \tag{141}$$

Using the cyclicity of the trace, the choices in (141) imply that we have

$$\mathbf{Tr}\left(A^{a_0}\left(\prod_{j=1}^{k-1} B^{b_j} A^{a_j}\right) B^{b_k}\right) = \mathbf{Tr}\left(\prod_{j=0}^{k-1} A^{(q+1)/p_j} (A^{1/2r_j} B^{q/r_j} A^{1/2r_j})\right). \tag{142}$$

Moreover,

$$\sum_{j=0}^{k-1} \frac{1}{p_j} + \sum_{j=0}^{k-1} \frac{1}{r_j} = \frac{1}{q+1}\sum_{j=0}^{k-1} a_j + \left(\frac{1}{q} - \frac{1}{q(q+1)}\right)\sum_{j=1}^{k} b_j \stackrel{(140)}{=} 1.$$

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 [**28**, Théorème 6] to deduce from (142) that

$$\mathbf{Tr}\left(A^{a_0}\left(\prod_{j=1}^{k-1} B^{b_j} A^{a_j}\right) B^{b_k}\right) \leqslant \prod_{j=0}^{k-1} (\mathbf{Tr}(A^{q+1}))^{1/p_j} (\mathbf{Tr}((A^{1/2r_j} B^{q/r_j} A^{1/2r_j})^{r_j}))^{1/r_j}. \tag{143}$$

The Lieb–Thirring inequality [**49**] asserts that $\mathbf{Tr}((XYX)^r) \leqslant \mathbf{Tr}(X^r Y^r X^r)$ for every $r \in [1,\infty)$ and for every symmetric and positive semidefinite matrixes $X, Y \in M_d(\mathbb{R})$. Recalling the definition of $r_0, \ldots, r_{k-1}$ in (141), for every $j \in \{0,\ldots,k-1\}$ we therefore have

$$\mathbf{Tr}((A^{1/2r_j} B^{q/r_j} A^{1/2r_j})^{r_j}) \leqslant \mathbf{Tr}(\sqrt{A} B^q \sqrt{A}) = \mathbf{Tr}(B^q A).$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

A substitution of this estimate into (143) gives

$$\begin{aligned}\mathbf{Tr}\left(A^{a_0}\left(\prod_{j=1}^{k-1}B^{b_j}A^{a_j}\right)B^{b_k}\right) &\leqslant (\mathbf{Tr}(A^{q+1}))^{\sum_{j=0}^{k-1}(1/p_j)}(\mathbf{Tr}(B^qA))^{\sum_{j=0}^{k-1}(1/r_j)}\\ &= (\mathbf{Tr}(A^{q+1}))^{1-(1/q)\sum_{j=1}^{k}b_j}(\mathbf{Tr}(B^qA))^{(1/q)\sum_{j=1}^{k}b_j},\end{aligned} \tag{144}$$

where we used the fact that, due to (141), we have $\sum_{j=0}^{k-1}(1/r_j) = (1/q)\sum_{j=1}^{k}b_j$. □

REMARK 7.7. For future use, note that if $q, a_0, \dots, a_{k-1}, b_1, \dots, b_k \in (0,\infty)$ satisfy (140) and we also know that $a_0, \dots, a_{k-1} \geqslant 1$, then the assumptions of Lemma 7.6 hold true, that is, $b_j + b_{j+1} \leqslant 2qa_j$ for every $j \in \{0, \dots, k-1\}$. Indeed, by (140) we have $\max\{b_j, b_{j+1}\} \leqslant q + 1 - a_j \leqslant qa_j$, and consequently $b_j + b_{j+1} \leqslant 2\max\{b_j, b_{j+1}\} \leqslant 2qa_j$.

Proof of Proposition 7.5. Write $q = 2m + \theta$, where $m \in \mathbb{N}\cup\{0\}$ and $\theta \in (0, 2]$. The proof of (139) 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 [23, Theorem 2.6]). This means that for every $s \in (0, 1)$ we have

$$(A+B)^\theta = \left(s\frac{A}{s} + (1-s)\frac{B}{1-s}\right)^\theta \leqslant \frac{A^\theta}{s^{\theta-1}} + \frac{B^\theta}{(1-s)^{\theta-1}}, \tag{145}$$

where, as usual, we interpret the inequality (145) in terms of the PSD order of matrixes, that is, that the right-hand side of (145) minus the left-hand side of (145) is a positive semidefinite matrix.

It follows from (145) that

$$\begin{aligned}&\sqrt{A}(A+B)^q\sqrt{A}\\ &\quad\leqslant \frac{\sqrt{A}(A+B)^mA^\theta(A+B)^m\sqrt{A}}{s^{\theta-1}} + \frac{\sqrt{A}(A+B)^mB^\theta(A+B)^m\sqrt{A}}{(1-s)^{\theta-1}}.\end{aligned}$$

So, by taking traces while making use of the cyclicity of the trace, we see that

$$\mathbf{Tr}((A+B)^qA) \leqslant \frac{\mathbf{Tr}((A+B)^mA^\theta(A+B)^mA)}{s^{\theta-1}} + \frac{\mathbf{Tr}((A+B)^mB^\theta(A+B)^mA)}{(1-s)^{\theta-1}}. \tag{146}$$

By choosing s so as to minimize the quantity appearing in the right-hand side of (146), we have

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

$$\begin{aligned}&(\mathbf{Tr}((A+B)^q A))^{1/\theta}\\ &\leqslant (\mathbf{Tr}((A+B)^m A^\theta (A+B)^m A))^{1/\theta}\\ &\quad + (\mathbf{Tr}((A+B)^m B^\theta (A+B)^m A))^{1/\theta}. \end{aligned} \tag{147}$$

We shall now proceed to estimate each of the terms that appear in the right-hand side of (147) separately. By expanding the mth powers appearing in the matrix $(A+B)^m A^\theta (A+B)^m A$, and using the cyclicity of the trace, we see that $\mathbf{Tr}((A+B)^m A^\theta (A+B)^m A)$ equals the sum of 2^{2m} terms, each of which is of the form

$$\mathbf{Tr}\left(A^{a_0}\left(\prod_{j=1}^{k-1} B^{b_j} A^{a_j}\right) B^{b_k}\right), \tag{148}$$

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 (140) (recall that $q = 2m + \theta$). Here we use the convention that when $k = 0$ the quantity appearing in (148) equals $\mathbf{Tr}(A^{q+1})$. 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 (148) that appear in the above expansion of $\mathbf{Tr}((A+B)^m A^\theta (A+B)^m A)$ with $\sum_{j=1}^k b_j = r$ 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 the matrix $(A+B)^m A^\theta (A+B)^m A$ as a product of matrixes, each of which is either A or B. Note also that $a_0, \ldots, a_{k-1} \geqslant 1$, since $\theta \geqslant 1$. Recalling Remark 7.7, we may therefore use Lemma 7.6 to deduce that

$$\mathbf{Tr}\left(A^{a_0}\left(\prod_{j=1}^{k-1} B^{b_j} A^{a_j}\right) B^{b_k}\right) \leqslant (\mathbf{Tr}(A^{q+1}))^{1-(1/q)\sum_{j=1}^k b_j} (\mathbf{Tr}(B^q A))^{(1/q)\sum_{j=1}^k b_j}. \tag{149}$$

Hence,

$$\begin{aligned}&\mathbf{Tr}((A+B)^m A^\theta (A+B)^m A)\\ &\leqslant \sum_{r=0}^{2m} \binom{2m}{r} (\mathbf{Tr}(A^{q+1}))^{1-r/q} (\mathbf{Tr}(B^q A))^{r/q}\\ &= (\mathbf{Tr}(A^{q+1}))^{1-2m/q} ((\mathbf{Tr}(A^{q+1}))^{1/q} + (\mathbf{Tr}(B^q A))^{1/q})^{2m}\\ &= (\mathbf{Tr}(A^{q+1}))^{\theta/q} ((\mathbf{Tr}(A^{q+1}))^{1/q} + (\mathbf{Tr}(B^q A))^{1/q})^{2m}, \end{aligned} \tag{150}$$

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

The second term in the right-hand side of (147) is bounded using similar reasoning. As before, $\mathbf{Tr}((A+B)^m B^\theta (A+B)^m A)$ equals the sum of terms as in (148), for some $k \in \mathbb{N} \cup \{0\}$ and $a_0, \ldots, a_{k-1}, b_1, \ldots, b_k \in (0, \infty)$ that

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

satisfy (140). However, now we know that $a_1, \dots, a_{k-1} \in \mathbb{N}$ and $\sum_{j=0}^{k-1} b_j - \theta \in \{0, \dots, 2m\}$. By Lemma 7.6 (and Remark 7.7), the estimate (149) holds true for the terms of the form (148) that appear in the expansion of the quantity $\mathbf{Tr}((A+B)^m B^\theta (A+B)^m A)$. For every $r \in \{0, \dots, 2m\}$, the number of terms of the form (148) that appear in the expansion of $\mathbf{Tr}((A+B)^m B^\theta (A+B)^m A)$ with $\sum_{j=1}^{k} b_j = r + \theta$ equals $\binom{2m}{r}$, so by (149) we have

$$\begin{aligned}
&\mathbf{Tr}((A+B)^m B^\theta (A+B)^m A) \\
&\quad \leqslant \sum_{r=0}^{2m} \binom{2m}{r} (\mathbf{Tr}(A^{q+1}))^{1-(r+\theta)/q} (\mathbf{Tr}(B^q A))^{(r+\theta)/q} \\
&\quad = (\mathbf{Tr}(A^{q+1}))^{1-(2m+\theta)/q} (\mathbf{Tr}(B^q A))^{\theta/q} ((\mathbf{Tr}(A^{q+1}))^{1/q} + (\mathbf{Tr}(B^q A))^{1/q})^{2m} \\
&\quad = (\mathbf{Tr}(B^q A))^{\theta/q} ((\mathbf{Tr}(A^{q+1}))^{1/q} + (\mathbf{Tr}(B^q A))^{1/q})^{2m}, \qquad (151)
\end{aligned}$$

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

By substituting (150) and (151) into (147) we see that

$$\begin{aligned}
&(\mathbf{Tr}((A+B)^q A))^{1/\theta} \\
&\quad \leqslant ((\mathbf{Tr}(A^{q+1}))^{1/q} + (\mathbf{Tr}(B^q A))^{1/q})^{1+2m/\theta} \\
&\quad = ((\mathbf{Tr}(A^{q+1}))^{1/q} + (\mathbf{Tr}(B^q A))^{1/q})^{q/\theta},
\end{aligned}$$

using $2m + \theta = q$ once more. This completes the proof of the desired estimate (139) in Case 1.

Case 2: $\theta \in (0, 1)$. Note that since the underlying assumption of Proposition 7.5 is that $q \geqslant 1$, the facts that $q = 2m + \theta$ and $\theta \in (0, 1)$ imply that the integer m is positive. Moreover, in the range $\theta \in (0, 1)$ the mapping $t \to t^\theta$ is no longer operator-convex but we have the following commonly used (see for example [**32**]) integral representation at our disposal. Since for every $a \in (0, \infty)$ we have

$$a^\theta = \frac{\sin(\pi\theta)}{\pi} \int_0^\infty t^\theta \left(\frac{1}{t} - \frac{1}{t+a} \right) dt,$$

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

$$(sA+B)^\theta = \frac{\sin(\pi\theta)}{\pi} \int_0^\infty t^\theta \left(\frac{1}{t} I - (tI + sA + B)^{-1} \right) dt. \qquad (152)$$

Since $(d/dt)X(t)^{-1} = -X(t)^{-1} X'(t) X(t)^{-1}$ 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 (152) that

$$\frac{d}{ds}(sA+B)^\theta = \frac{\sin(\pi\theta)}{\pi} \int_0^\infty t^\theta (tI + sA + B)^{-1} A (tI + sA + B)^{-1} dt. \qquad (153)$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

By integrating over $s \in [0,1]$, in order to prove (139) it will suffice to show that

$$\forall\, s \in (0,1), \qquad \frac{d}{ds}(\mathbf{Tr}((sA+B)^q A))^{1/q} \leqslant (\mathbf{Tr}(A^{q+1}))^{1/q}.$$

Equivalently, we want to prove that

$$\forall\, s \in (0,1), \qquad \frac{d}{ds}\mathbf{Tr}((sA+B)^q A) \leqslant q(\mathbf{Tr}(A^{q+1}))^{1/q}(\mathbf{Tr}((sA+B)^q A))^{1-1/q}. \tag{154}$$

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

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

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

and

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

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

$$\frac{d}{ds}\mathbf{Tr}((sA+B)^q A) = f(s)+g(s)+h(s).$$

Hence, because $q = 2m+\theta$, in order to establish the validity of (154) it suffice to show that for every $s \in [0,1]$ we have

$$\max\{f(s), h(s)\} \leqslant m(\mathbf{Tr}(A^{q+1}))^{1/q}(\mathbf{Tr}((sA+B)^q A))^{1-1/q}, \tag{155}$$

and

$$g(s) \leqslant \theta(\mathbf{Tr}(A^{q+1}))^{1/q}(\mathbf{Tr}((sA+B)^q A))^{1-1/q}. \tag{156}$$

Observe that

$$\begin{aligned} f(s) &= \sum_{r=1}^m \mathbf{Tr}((sA+B)^{r-1}A(sA+B)^{m-r}(sA+B)^{m+\theta}A) \\ &= \sum_{r=1}^m \mathbf{Tr}((sA+B)^{r-1}A(sA+B)^{q-r}A). \end{aligned} \tag{157}$$

Similarly, using the cyclicity of the trace, we have

$$h(s) = \sum_{r=1}^m \mathbf{Tr}((sA+B)^{q-r}A(sA+B)^{r-1}A) = f(s). \tag{158}$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

Finally, by the integral representation (153) we have

$$g(s) = \frac{\sin(\pi\theta)}{\pi} \times \int_0^\infty t^\theta \,\mathbf{Tr}((sA+B)^m(tI+sA+B)^{-1}A(tI+sA+B)^{-1} \times (sA+B)^m A)\,dt. \tag{159}$$

By denoting $C = sA + B$, it follows from (157), (158) and (159) that the desired estimates (155) and (156) will be proven once we show that for every $C \in M_d(\mathbb{R})$ that is symmetric and positive semidefinite we have

$$\forall\, r \in \{1,\dots,m\}, \quad \mathbf{Tr}(C^{r-1}AC^{q-r}A) \leqslant (\mathbf{Tr}(A^{q+1}))^{1/q}(\mathbf{Tr}(C^qA))^{1-1/q}, \tag{160}$$

and

$$\int_0^\infty t^\theta \,\mathbf{Tr}(C^m(tI+C)^{-1}A(tI+C)^{-1}C^mA)\,dt \leqslant \frac{\pi\theta}{\sin(\pi\theta)}(\mathbf{Tr}(A^{q+1}))^{1/q}(\mathbf{Tr}(C^qA))^{1-1/q}. \tag{161}$$

(160) is a consequence of Lemma 7.6 (with B replaced by C). It therefore remains to establish the validity of (161). 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

$$\begin{aligned}&\mathbf{Tr}(C^m(tI+C)^{-1}A(tI+C)^{-1}C^mA)\\ &\quad= \mathbf{Tr}((\sqrt{A}C^m(tI+C)^{-1}\sqrt{A})^2)\\ &\quad\leqslant \mathbf{Tr}(AC^{2m}(tI+C)^{-2}A) = \mathbf{Tr}(C^{2m}(tI+C)^{-2}A^2),\end{aligned} \tag{162}$$

where for the inequality in (162) we used the Lieb–Thirring inequality. This upper bound on the integrand in the left-hand side of (161) yields the following estimate.

$$\int_0^\infty t^\theta \,\mathbf{Tr}(C^m(tI+C)^{-1}A(tI+C)^{-1}C^mA)\,dt \leqslant \mathbf{Tr}\left(C^{2m}\left(\int_0^\infty t^\theta(tI+C)^{-2}\,dt\right)A^2\right). \tag{163}$$

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

$$\int_0^\infty \frac{t^\theta}{(t+c)^2}\,dt = c^{\theta-1}\int_0^\infty \frac{s^\theta}{(s+1)^2}\,ds = \frac{\pi\theta}{\sin(\pi\theta)}c^{\theta-1}.$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

Consequently,

$$\begin{aligned}
&\mathbf{Tr}\left(C^{2m}\left(\int_0^\infty t^\theta (tI+C)^{-2}\,dt\right)A^2\right)\\
&=\frac{\pi\theta}{\sin(\pi\theta)}\,\mathbf{Tr}(C^{2m+\theta-1}A^2)\\
&=\frac{\pi\theta}{\sin(\pi\theta)}\,\mathbf{Tr}(C^{q-1}A^2)\\
&=\frac{\pi\theta}{\sin(\pi\theta)}\,\mathbf{Tr}(AC^{q-1}A)\\
&\overset{(160)}{\leqslant}\frac{\pi\theta}{\sin(\pi\theta)}(\mathbf{Tr}(A^{q+1}))^{1/q}(\mathbf{Tr}(C^qA))^{1-1/q}. \qquad (164)
\end{aligned}$$

A substitution of (164) into (163) yields the desired inequality (161). □

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*

$$\mathbf{Tr}((A+B)^qA) \leqslant \mathbf{Tr}(A^{q+1}) + \mathbf{Tr}(B^qA).$$

Proof. By the integral identity (153), with θ replaced by q (which is allowed since $0<q<1$), for every $s \in (0,\infty)$ we have

$$\frac{d}{ds}\,\mathbf{Tr}((sA+B)^qA) = \frac{\sin(\pi q)}{\pi}\int_0^\infty t^q\,\mathbf{Tr}((tI+sA+B)^{-1}A(tI+sA+B)^{-1}A)\,dt. \qquad (165)$$

Fix $s,t \in (0,\infty)$ and define $F:[0,\infty)\to\mathbb{R}$ by

$$\forall\, w \in [0,\infty),\qquad F(w) \overset{\text{def}}{=} \mathbf{Tr}((tI+sA+wB)^{-1}A(tI+sA+wB)^{-1}A).$$

This mapping was investigated in [**11**, Section III], where it was shown to be convex. Here we need to know that it is nonincreasing, which follows from the following computation.

$$\begin{aligned}
F'(w) &= -\mathbf{Tr}((tI+sA+wB)^{-1}B(tI+sA+wB)^{-1}A(tI+sA+wB)^{-1}A)\\
&\quad -\mathbf{Tr}((tI+sA+wB)^{-1}A(tI+sA+wB)^{-1}B(tI+sA+wB)^{-1}A)\\
&= -\mathbf{Tr}(CD) - \mathbf{Tr}(DC) = -2\,\mathbf{Tr}(CD) \leqslant 0,
\end{aligned}$$

where $C, D \in M_d(\mathbb{R})$ are the symmetric and positive semidefinite matrixes given by

$$C \overset{\text{def}}{=} (tI+sA+wB)^{-1}B(tI+sA+wB)^{-1},$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

and

$$D \stackrel{\text{def}}{=} A(tI + sA + wB)^{-1}A.$$

It follows from these considerations that

$$\begin{aligned}&\mathbf{Tr}((tI + sA + B)^{-1}A(tI + sA + B)^{-1}A)\\ &\quad = F(1) \leqslant F(0)\\ &\quad = \mathbf{Tr}((tI + sA)^{-1}A(tI + sA)^{-1}A).\end{aligned}$$

A substitution of this estimate into (165) shows that

$$\begin{aligned}\frac{d}{ds}\,\mathbf{Tr}((sA + B)^q A) &\leqslant \frac{\sin(\pi q)}{\pi}\int_0^\infty t^q\,\mathbf{Tr}((tI + sA)^{-1}A(tI + sA)^{-1}A)\,dt\\ &\stackrel{(152)}{=} \frac{d}{ds}\,\mathbf{Tr}((sA)^q A) = qs^{q-1}\,\mathbf{Tr}(A^{q+1}).\end{aligned} \tag{166}$$

By integrating (166) over $[0, 1]$ we therefore see that

$$\mathbf{Tr}((A + B)^q A) - \mathbf{Tr}(B^q A) \leqslant \mathbf{Tr}(A^{q+1}). \qquad \square$$

We record for future use the following simple reformulation of Propositions 7.5 and 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 $t \mapsto t^{q-1}$ on $[0, \infty)$.

COROLLARY 7.9. *Suppose that $q \in [1, \infty)$ and $d \in \mathbb{N}$. Set $r \stackrel{\text{def}}{=} \max\{q - 2, 0\}$. For every $A, B \in M_d(\mathbb{R})$ that are symmetric and positive semidefinite we have*

$$\mathbf{Tr}((A + B)^{q-1}A) \leqslant \min\left\{\frac{\mathbf{Tr}(A^q)}{\lambda^r} + \frac{\mathbf{Tr}(B^{q-1}A)}{(1-\lambda)^r} : \lambda \in (0, 1)\right\}.$$

7.1.1. Discussion and counterexamples. A straightforward 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

$$\mathbf{Tr}((A + B)^{p/2-1}A) \leqslant K(\mathbf{Tr}(A^{p/2-1}A) + \mathbf{Tr}(B^{p/2-1}A)). \tag{167}$$

Specifically, (167) 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, (167) holds true with $K = 2^{\max\{0,(p-4)/2\}}$.

Setting $q = (p - 2)/2 > 0$, it is natural to ask whether multiplication by A is crucial for (167) to hold true. Specifically, one would naturally investigate

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

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

$$\mathbf{Tr}((A+B)^q C) \leqslant K(\mathbf{Tr}(A^q C) + \mathbf{Tr}(B^q C)), \tag{168}$$

with $K \in (0,\infty)$ independent of A, B, C. By a simple duality argument (for example [23, Lemma 5.12]), the above requirement is equivalent to the matrix inequality

$$(A+B)^q \leqslant K(A^q + B^q), \tag{169}$$

where, as usual, we interpret the inequality (169) in terms of the PSD order of matrixes.

Since for $q \in [1,2]$ the function $t \mapsto t^q$ is operator-convex (see for example [16]), for such q the PSD inequality (169) holds true with $K = 2^{q-1}$ (recall (145)). This yields a simple proof of (167) when $4 \leqslant p \leqslant 6$. Moreover, when $q \in [1,2]$ the operator convexity of the function $t \mapsto t^q$ shows that if $A, B, C \in M_d(\mathbb{R})$ are symmetric and positive semidefinite then for every $\lambda \in (0,1)$ we have

$$\mathbf{Tr}((A+B)^q C) \leqslant \frac{\mathbf{Tr}(A^q C)}{\lambda^{q-1}} + \frac{\mathbf{Tr}(B^q C)}{(1-\lambda)^{q-1}}. \tag{170}$$

By choosing λ so as to minimize the right-hand side of (170) we see that

$$(\mathbf{Tr}((A+B)^q C))^{1/q} \leqslant (\mathbf{Tr}(A^q C))^{1/q} + (\mathbf{Tr}(B^q C))^{1/q}. \tag{171}$$

The inequality (171) is a strengthening of Proposition 7.5 in the special case $q \in [1,2]$, showing that when q belongs to this range Proposition 7.5 is a simple consequence of the operator convexity of the function $t \mapsto t^q$ (alternatively, one can deduce Proposition 7.5 directly from the work of Carlen and Lieb [24]; see specifically [24, Theorem 1.1 and Remark 1.2]). However, the above argument is special to the range $q \in [1,2]$ since, as we shall explain below, if $q \in (0,1) \cup (2,\infty)$ then (169) 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 in (169), but was rather devoted to, for example, finding substitutes for (169) with $q \in (0,1)$ and $K = 1$). One such substitute allows for conjugation by unitary matrixes, as initiated in [1]. A satisfactory recent result [6] along these lines asserts that if $f : [0,\infty) \to \mathbb{R}$ is nondecreasing, concave, and $f(0) \geqslant 0$, then for every $A, B \in M_d(\mathbb{R})$ there exist unitary matrixes $U, V \in M_d(\mathbb{C})$ such that $f(A+B) \leqslant Uf(A)U^* + Vf(B)V^*$. Another substitute for PSD subadditivity is a subadditivity inequality for unitarily invariant norms. Recall that a norm $\|\cdot\|$ on $M_d(\mathbb{C})$ is unitarily invariant if $\|UXV\| = \|X\|$ for every $X, U, V \in M_d(\mathbb{C})$ such that U, V are unitary. The papers [3, 20] contain satisfactory results along

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

these lines, obtaining inequalities of the form $\|f(A+B)\| \leqslant \|f(A)+f(B)\|$. For $q \in (0,1)$, when $f(t) = t^q$ and $\|\cdot\|$ is the Schatten 1 norm, the resulting inequality goes back to [57] and it corresponds to (168) with $C = I$ (and $K = 1$).

Here we study a different type of substitute for (169). For example, when $A \in M_d(\mathbb{R})$ is symmetric and positive semidefinite define $F_A : M_d(\mathbb{R}) \to \mathbb{R}$ by $F_A(X) = (\mathbf{Tr}(|X|^q A))^{1/q}$ (F_A need not be unitarily invariant). Proposition 7.5 asserts that if $q > 1$ then $F_A(X+Y) \leqslant F_A(X)+F_A(Y)$ for symmetric and positive semidefinite $X, Y \in M_d(\mathbb{R})$, provided that either X or Y equals A. Weakenings of (168) (the special case $C = A$) suffice for our application (that is, proving the X_p inequality for S_p, 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 $\mathbf{Tr}(f(A+B)A) \leqslant K\,\mathbf{Tr}(f(A)A) + K\,\mathbf{Tr}(f(B)A)$.

We shall end this discussion by presenting the aforementioned example that exhibits the failure of (169) 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

$$A_s \stackrel{\text{def}}{=} \begin{pmatrix} s^2 & 0 \\ 0 & 0 \end{pmatrix}, \quad B_s \stackrel{\text{def}}{=} \begin{pmatrix} 1 & s \\ s & s^2 \end{pmatrix} \quad \text{and} \quad w_s \stackrel{\text{def}}{=} \begin{pmatrix} -s \\ 1 \end{pmatrix}.$$

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 $s < 1/\sqrt[4]{K}$, in which case the matrix $K(A_s^4+B_s^4)-(A_s+B_s)^4$ is not positive semidefinite. This shows that (169) 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 (167) when $p = 10$). A similar, though more tedious, computation shows that (169) 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

$$(A_s + B_s)^q = \begin{pmatrix} \dfrac{a(s)^q(\sqrt{1+4s^2}+1) + b(s)^q(\sqrt{1+4s^2}-1)}{2\sqrt{1+4s^2}} & \dfrac{s(a(s)^q - b(s)^q)}{\sqrt{1+4s^2}} \\ \dfrac{s(a(s)^q - b(s)^q)}{\sqrt{1+4s^2}} & \dfrac{a(s)^q(\sqrt{1+4s^2}-1) + b(s)^q(\sqrt{1+4s^2}+1)}{2\sqrt{1+4s^2}} \end{pmatrix},$$

where

$$a(s) \stackrel{\text{def}}{=} s^2 + \frac{1}{2} + \frac{\sqrt{1+4s^2}}{2} \quad \text{and} \quad b(s) \stackrel{\text{def}}{=} s^2 + \frac{1}{2} - \frac{\sqrt{1+4s^2}}{2}.$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

One then directly computes that as $s \to 0$,

$$\langle (K(A_s^q + B_s^q) - (A_s + B_s)^q) w_s, w_s \rangle = (K s^{2(q+1)} - s^6 - s^{4q})(1 + O_{q,K}(s^2)). \quad (172)$$

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 (172) is negative for small enough s, which means that the matrix $K(A_s^q + B_s^q) - (A_s + B_s)^q$ is not positive semidefinite.

7.2. Proof of Lemma 7.4. For the sake of simplicity denote

$$U \stackrel{\text{def}}{=} \frac{1}{\binom{n}{k}} \sum_{\substack{S \subseteq \{1,\ldots,n\} \\ |S| = k}} \mathbf{Tr}\left(\left(\sum_{j \in S} B_j \right)^q \right),$$
$$V \stackrel{\text{def}}{=} \frac{k}{n} \sum_{j=1}^n \mathbf{Tr}(B_j^q), \quad W \stackrel{\text{def}}{=} \mathbf{Tr}\left(\left(\frac{k}{n} \sum_{j=1}^n B_j \right)^q \right). \quad (173)$$

Our goal is therefore to show that

$$U \leqslant \left(\frac{4q}{\log(2q)} \right)^q \max\{V, W\}. \quad (174)$$

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

$$\mathbf{Tr}\left(\left(\sum_{s \in S} B_s \right)^{q-1} B_j \right) \leqslant \frac{1}{\lambda^r} \mathbf{Tr}(B_j^q) + \frac{1}{(1-\lambda)^r} \mathbf{Tr}\left(\left(\sum_{s \in S \setminus \{j\}} B_s \right)^{q-1} B_j \right),$$

where, as denoted in Corollary 7.9, $r = \max\{q-2, 0\}$. Hence,

$$\mathbf{Tr}\left(\left(\sum_{j \in S} B_j \right)^q \right) = \sum_{j=1}^n \mathbf{Tr}\left(\left(\sum_{j \in S} B_j \right)^{q-1} B_j \right)$$
$$\leqslant \frac{1}{\lambda^r} \sum_{j \in S} \mathbf{Tr}(B_j^q) + \frac{1}{(1-\lambda)^r} \sum_{j \in S} \mathbf{Tr}\left(\left(\sum_{s \in S \setminus \{j\}} B_s \right)^{q-1} B_j \right). \quad (175)$$

By averaging (175) over all of those $S \subseteq \{1,\ldots,n\}$ with $|S| = k$, and recalling (173), we see that

$$U \leqslant \frac{V}{\lambda^r} + \frac{1}{(1-\lambda)^r \binom{n}{k}} \sum_{\substack{S \subseteq \{1,\ldots,n\} \\ |S| = k}} \sum_{j \in S} \mathbf{Tr}\left(\left(\sum_{s \in S \setminus \{j\}} B_s \right)^{q-1} B_j \right). \quad (176)$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

Now,

$$\sum_{\substack{S\subseteq\{1,\dots,n\}\\|S|=k}}\sum_{j\in S}\mathbf{Tr}\Bigg(\Bigg(\sum_{s\in S\setminus\{j\}}B_s\Bigg)^{q-1}B_j\Bigg)=\sum_{\substack{T\subseteq\{1,\dots,n\}\\|T|=k-1}}\sum_{j\in\{1,\dots,n\}\setminus T}\mathbf{Tr}\Bigg(\Bigg(\sum_{t\in T}B_t\Bigg)^{q-1}B_j\Bigg)$$

$$=\sum_{\substack{T\subseteq\{1,\dots,n\}\\|T|=k-1}}\mathbf{Tr}\Bigg(\Bigg(\sum_{t\in T}B_t\Bigg)^{q-1}\Bigg(\sum_{j\in\{1,\dots,n\}\setminus T}B_j\Bigg)\Bigg)$$

$$\leqslant\sum_{\substack{T\subseteq\{1,\dots,n\}\\|T|=k-1}}\mathbf{Tr}\Bigg(\Bigg(\sum_{t\in T}B_t\Bigg)^{q-1}\Bigg(\sum_{j=1}^{n}B_j\Bigg)\Bigg),\tag{177}$$

where in the last step of (177) we used the fact that if $A, B, C \in M_d(\mathbb{R})$ are symmetric and positive semidefinite then $\mathbf{Tr}(AB) \leqslant \mathbf{Tr}(A(B+C))$. To bound the final term in (177), use Hölder's inequality for traces to deduce that for every $T \subseteq \{1,\dots,n\}$ we have

$$\mathbf{Tr}\Bigg(\Bigg(\sum_{t\in T}B_t\Bigg)^{q-1}\Bigg(\sum_{j=1}^{n}B_j\Bigg)\Bigg)\leqslant\Bigg(\mathbf{Tr}\Bigg(\Bigg(\sum_{j=1}^{n}B_j\Bigg)^{q}\Bigg)\Bigg)^{1/q}\Bigg(\mathbf{Tr}\Bigg(\Bigg(\sum_{t\in T}B_t\Bigg)^{q}\Bigg)\Bigg)^{1-1/q}$$

$$=\frac{nW^{1/q}}{k}\Bigg(\mathbf{Tr}\Bigg(\Bigg(\sum_{t\in T}B_t\Bigg)^{q}\Bigg)\Bigg)^{1-1/q},\tag{178}$$

where we recall the definition of W in (173).

The function $t \mapsto t^q$ is operator trace-increasing (see [**23**, Theorem 2.10]), that is, if $C, D \in M_d(\mathbb{R})$ are symmetric and positive semidefinite with $C \leqslant D$ then $\mathbf{Tr}(C^q) \leqslant \mathbf{Tr}(D^q)$. Consequently, for every $T \subsetneq \{1,\dots,n\}$ and $u \in \{1,\dots,n\}$ we have $\mathbf{Tr}((\sum_{t\in T}B_t)^q) \leqslant \mathbf{Tr}((B_u+\sum_{t\in T}B_t)^q)$. By raising this inequality to the power $(q-1)/q$ and averaging over all $u \in \{1,\dots,n\}\setminus T$ we see that

$$\Bigg(\mathbf{Tr}\Bigg(\Bigg(\sum_{t\in T}B_t\Bigg)^{q}\Bigg)\Bigg)^{1-1/q}\leqslant\frac{1}{n-|T|}\sum_{u\in\{1,\dots,n\}\setminus T}\Bigg(\mathbf{Tr}\Bigg(\Bigg(\sum_{t\in T\cup\{u\}}B_t\Bigg)^{q}\Bigg)\Bigg)^{1-1/q}.\tag{179}$$

Hence, by combining (178) and (179) with (177), we see that

$$\sum_{\substack{S\subseteq\{1,\dots,n\}\\|S|=k}}\sum_{j\in S}\mathbf{Tr}\Bigg(\Bigg(\sum_{s\in S\setminus\{j\}}B_s\Bigg)^{q-1}B_j\Bigg)$$

$$\leqslant\frac{nW^{1/q}}{k(n-k+1)}\sum_{\substack{T\subseteq\{1,\dots,n\}\\|T|=k-1}}\sum_{u\in\{1,\dots,n\}\setminus T}\Bigg(\mathbf{Tr}\Bigg(\Bigg(\sum_{t\in T\cup\{u\}}B_t\Bigg)^{q}\Bigg)\Bigg)^{1-1/q}\tag{180}$$

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

$$= \frac{nW^{1/q}}{n-k+1} \sum_{\substack{S\subseteq\{1,\ldots,n\}\\ |S|=k}} \left(\mathbf{Tr}\left(\left(\sum_{s\in S} B_s\right)^q\right)\right)^{1-1/q}, \tag{181}$$

where for (181) 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 (180) with multiplicity k, once for each $u \in S$.

Recalling the definition of U in (173), by Jensen's inequality we see that

$$\frac{1}{\binom{n}{k}} \sum_{\substack{S\subseteq\{1,\ldots,n\}\\ |S|=k}} \left(\mathbf{Tr}\left(\left(\sum_{s\in S} B_s\right)^q\right)\right)^{1-1/q} \leqslant U^{1-1/q}. \tag{182}$$

By substituting (182) into (181) and using $k \leqslant n/2$, we have

$$\frac{1}{\binom{n}{k}} \sum_{\substack{S\subseteq\{1,\ldots,n\}\\ |S|=k}} \sum_{j\in S} \mathbf{Tr}\left(\left(\sum_{s\in S\setminus\{j\}} B_s\right)^{q-1} B_j\right) \leqslant 2W^{1/q}U^{(q-1)/q}. \tag{183}$$

In conjunction with (183), it follows from (176) that

$$\begin{aligned} U &\leqslant \min\left\{\frac{V}{\lambda^r} + \frac{2W^{1/q}U^{(q-1)/q}}{(1-\lambda)^r} : \lambda\in(0,1)\right\} \\ &\leqslant (V^{1/(r+1)} + 2^{1/(r+1)}W^{1/(q(r+1))}U^{(q-1)/(q(r+1))})^{r+1}, \end{aligned} \tag{184}$$

where the final inequality in (184) is seen by choosing $1/\lambda = 1 + (2W^{1/q}U^{(q-1)/q}/V)^{1/(r+1)}$. By (184),

$$U^{1/(r+1)} \leqslant V^{1/(r+1)} + 2^{1/(r+1)}W^{1/(q(r+1))}U^{(q-1)/(q(r+1))}. \tag{185}$$

The desired inequality (174) is a formal consequence of (185), as follows. If $U \leqslant (4q/\log(2q))^{r+1}V$ then (174) holds true because $r+1 \leqslant q$. We may therefore assume that $U > (4q/\log(2q))^{r+1}V$, in which case (185) implies that

$$\frac{U^{1/(r+1)}}{(2q)^{1/2q}} \leqslant \left(1 - \frac{\log(2q)}{4q}\right)U^{1/(r+1)} \leqslant 2^{1/(r+1)}W^{1/(q(r+1))}U^{(q-1)/(q(r+1))}, \tag{186}$$

where we used the fact that $(1-t) \geqslant e^{-2t}$ for every $t \in [0,1/2]$. The estimate (186) simplifies to

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

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

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

Acknowledgments

We are grateful to Eric Carlen and Oded Regev for helpful pointers to the literature related to Section 7.1, as well as showing us counterexamples to (169) when $q = 4$. We also thank the anonymous referee for carefully reading our manuscript and for the many helpful comments. A.N. was supported in part by the NSF, the BSF, the Packard Foundation and the Simons Foundation. G.S. was supported in part by the ISF and the BSF.

Added in proof

The recent preprint [**68**] resolves positively Conjecture 1.5, Conjecture 1.8 and Conjecture 1.12. This is achieved in [**68**] via a route that differs from the route that we proposed in Section 6; in particular Question 6.1 remains open.

References

[1] C. A. Akemann, J. Anderson and G. K. Pedersen, 'Triangle inequalities in operator algebras', *Linear Multilinear Algebra* **11**(2) (1982), 167–178.

[2] F. Albiac and F. Baudier, 'Embeddability of snowflaked metrics with applications to the nonlinear geometry of the spaces L_p and ℓ_p for $0 < p < \infty$', *J. Geom. Anal.* **25**(1) (2015), 1–24.

[3] T. Ando and X. Zhan, 'Norm inequalities related to operator monotone functions', *Math. Ann.* **315**(4) (1999), 771–780.

[4] A. Andoni, A. Naor and O. Neiman, 'Snowflake universality of Wasserstein spaces', Preprint, 2015, arXiv:1509.08677.

[5] N. Aronszajn, 'Differentiability of Lipschitzian mappings between Banach spaces', *Studia Math.* **57**(2) (1976), 147–190.

[6] J. S. Aujla and J.-C. Bourin, 'Eigenvalue inequalities for convex and log-convex functions', *Linear Algebra Appl.* **424**(1) (2007), 25–35.

[7] T. Austin and A. Naor, 'On the bi-Lipschitz structure of Wasserstein spaces', Preprint, 2015.

[8] T. Austin, A. Naor and Y. Peres, 'The wreath product of $\mathbb{Z}$ with $\mathbb{Z}$ has Hilbert compression exponent $2/3$', *Proc. Amer. Math. Soc.* **137**(1) (2009), 85–90.

[9] K. Ball, 'Markov chains, Riesz transforms and Lipschitz maps', *Geom. Funct. Anal.* **2**(2) (1992), 137–172.

[10] K. Ball, 'The Ribe programme', *Astérisque* (352): Exp. No. 1047, viii, 147, 159, 2013, Séminaire Bourbaki. Vol. 2011/2012. Exposés 1043–1058.

[11] K. Ball, E. A. Carlen and E. H. Lieb, 'Sharp uniform convexity and smoothness inequalities for trace norms', *Invent. Math.* **115**(3) (1994), 463–482.

[12] S. Banach, Théorie des opérations linéaires. Éditions Jacques Gabay, Sceaux, 1993. Reprint of the 1932 original.

[13] Y. Bartal, N. Linial, M. Mendel and A. Naor, 'On metric Ramsey-type phenomena', *Ann. of Math.* (2) **162**(2) (2005), 643–709.

[14] F. Baudier, 'Quantitative nonlinear embeddings into Lebesgue sequence spaces', *J. Topol. Anal.*, to appear, Preprint, 2012, arXiv:1210.0588.

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

[15] Y. Benyamini and J. Lindenstrauss, *Geometric Nonlinear Functional Analysis,* Vol. 1, American Mathematical Society Colloquium Publications, 48 (American Mathematical Society, Providence, RI, 2000).
[16] R. Bhatia, *Matrix Analysis*, Graduate Texts in Mathematics, 169 (Springer, New York, 1997).
[17] J. Bourgain, 'The metrical interpretation of superreflexivity in Banach spaces', *Israel J. Math.* **56**(2) (1986), 222–230.
[18] J. Bourgain, 'Remarks on the extension of Lipschitz maps defined on discrete sets and uniform homeomorphisms', in *Geometrical Aspects of Functional Analysis (1985/86)*, Lecture Notes in Mathematics, 1267 (Springer, Berlin, 1987), 157–167.
[19] J. Bourgain, V. Milman and H. Wolfson, 'On type of metric spaces', *Trans. Amer. Math. Soc.* **294**(1) (1986), 295–317.
[20] J.-C. Bourin and M. Uchiyama, 'A matrix subadditivity inequality for $f(A+B)$ and $f(A)+f(B)$', *Linear Algebra Appl.* **423**(2–3) (2007), 512–518.
[21] J. Bretagnolle, D. Dacunha-Castelle and J.-L. Krivine, 'Fonctions de type positif sur les espaces L^p', *C. R. Acad. Sci. Paris* **261** (1965), 2153–2156.
[22] J. Brieussel and T. Zheng, 'Speed of random walks, isoperimetry and compression of finitely generated groups', Preprint, 2015, arXiv:1510.08040.
[23] E. Carlen, 'Trace inequalities and quantum entropy: an introductory course', in *Entropy and the Quantum*, Contemporary Mathematics, 529 (American Mathematical Society, Providence, RI, 2010), 73–140.
[24] E. A. Carlen and E. H. Lieb, 'A Minkowski type trace inequality and strong subadditivity of quantum entropy. II. Convexity and concavity', *Lett. Math. Phys.* **83**(2) (2008), 107–126.
[25] J. P. R. Christensen, 'Measure theoretic zero sets in infinite dimensional spaces and applications to differentiability of Lipschitz mappings', *Publ. Dép. Math. (Lyon)* **10**(2) (1973), 29–39. Actes du Deuxième Colloque d'Analyse Fonctionnelle de Bordeaux (Univ. Bordeaux, 1973), I, pp. 29–39.
[26] G. David and S. Semmes, *Fractured Fractals and Broken Dreams*, Oxford Lecture Series in Mathematics and its Applications, 7 (The Clarendon Press, Oxford University Press, New York, 1997), Self-similar geometry through metric and measure.
[27] J. Ding, J. R. Lee and Y. Peres, 'Markov type and threshold embeddings', *Geom. Funct. Anal.* **23**(4) (2013), 1207–1229.
[28] J. Dixmier, 'Formes linéaires sur un anneau d'opérateurs', *Bull. Soc. Math. France* **81** (1953), 9–39.
[29] P. Enflo, 'On the nonexistence of uniform homeomorphisms between L_p-spaces', *Ark. Mat.* **8** (1969), 103–105.
[30] P. Enflo, 'Uniform structures and square roots in topological groups. I, II', *Israel J. Math.* **8** (1970), 230–252. ibid., **8** (1970), 253–272.
[31] P. Enflo, 'Uniform homeomorphisms between Banach spaces', in *Séminaire Maurey-Schwartz (1975–1976), Espaces, L^p, applications radonifiantes et géométrie des espaces de Banach, Exp. No. 18* (Centre Math., École Polytech., Palaiseau, 1976), 7.
[32] H. Epstein, 'Remarks on two theorems of E. Lieb', *Comm. Math. Phys.* **31** (1973), 317–325.
[33] T. Figiel, W. B. Johnson and G. Schechtman, 'Random sign embeddings from l_r^n, $2 < r < \infty$', *Proc. Amer. Math. Soc.* **102**(1) (1988), 102–106.
[34] O. Giladi, M. Mendel and A. Naor, 'Improved bounds in the metric cotype inequality for Banach spaces', *J. Funct. Anal.* **260**(1) (2011), 164–194.
[35] O. Giladi and A. Naor, 'Improved bounds in the scaled Enflo type inequality for Banach spaces', *Extracta Math.* **25**(2) (2010), 151–164.

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

[36] O. Giladi, A. Naor and G. Schechtman, 'Bourgain's discretization theorem', *Ann. Fac. Sci. Toulouse Math.* (6) **21**(4) (2012), 817–837.
[37] E. D. Gluskin, A. Pietsch and J. Puhl, 'A generalization of Khintchine's inequality and its application in the theory of operator ideals', *Studia Math.* **67**(2) (1980), 149–155.
[38] M. Gromov, 'Filling Riemannian manifolds', *J. Differential Geom.* **18**(1) (1983), 1–147.
[39] T. Hytönen and A. Naor, 'Pisier's inequality revisited', *Studia Math.* **215**(3) (2013), 221–235.
[40] W. B. Johnson, B. Maurey, G. Schechtman and L. Tzafriri, 'Symmetric structures in Banach spaces', *Mem. Amer. Math. Soc.* **19**(217) (1979), v+298.
[41] W. B. Johnson, G. Schechtman and J. Zinn, 'Best constants in moment inequalities for linear combinations of independent and exchangeable random variables', *Ann. Probab.* **13**(1) (1985), 234–253.
[42] M. Ĭ. Kadeć, 'Linear dimension of the spaces L_p and l_q', *Uspehi Mat. Nauk* **13**(6(84)) (1958), 95–98.
[43] M. I. Kadec and A. Pełczyński, 'Bases, lacunary sequences and complemented subspaces in the spaces L_p', *Studia Math.* **21** (1961/1962), 161–176.
[44] N. J. Kalton, 'The nonlinear geometry of Banach spaces', *Rev. Mat. Complut.* **21**(1) (2008), 7–60.
[45] N. J. Kalton and N. L. Randrianarivony, 'The coarse Lipschitz geometry of $l_p \oplus l_q$', *Math. Ann.* **341**(1) (2008), 223–237.
[46] M. Ledoux and M. Talagrand, *Probability in Banach Spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 23 (Springer, Berlin, 1991), Isoperimetry and processes.
[47] J. R. Lee, A. Naor and Y. Peres, 'Trees and Markov convexity', *Geom. Funct. Anal.* **18**(5) (2009), 1609–1659.
[48] S. Li, 'Markov convexity and nonembeddability of the Heisenberg group', *Ann. Inst. Fourier*, to appear, Preprint, 2014, arXiv:1404.6751.
[49] E. H. Lieb and W. E. Thirring, 'Inequalities for the moments of the eigenvalues of the Schrödinger Hamiltonian and their relation to Sobolev inequalities', in *Studies in Mathematical Physics* (Princeton University Press, Princeton, NJ, 1976), 269–303.
[50] J. Lindenstrauss and L. Tzafriri, *Classical Banach Spaces. I* (Springer, Berlin, New York, 1977), Sequence spaces, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 92.
[51] N. Linial, A. Magen and A. Naor, 'Girth and Euclidean distortion', *Geom. Funct. Anal.* **12**(2) (2002), 380–394.
[52] G.-M. Lövblom, 'Uniform homeomorphisms between unit balls in L_p-spaces', *Math. Scand.* **62**(2) (1988), 294–302.
[53] F. Lust-Piquard, 'Inégalités de Khintchine dans C_p $(1 < p < \infty)$', *C. R. Acad. Sci. Paris Sér. I Math.* **303**(7) (1986), 289–292.
[54] P. Mankiewicz, 'On Lipschitz mappings between Fréchet spaces', *Studia Math.* **41** (1972), 225–241.
[55] J. Matoušek, *Lectures on Discrete Geometry*, Graduate Texts in Mathematics, 212 (Springer, New York, 2002).
[56] B. Maurey, 'Type, cotype and K-convexity', in *Handbook of the Geometry of Banach Spaces,* Vol. 2 (North-Holland, Amsterdam, 2003), 1299–1332.
[57] C. A. McCarthy, 'c_p', *Israel J. Math.* **5** (1967), 249–271.
[58] M. Mendel and A. Naor, 'Euclidean quotients of finite metric spaces', *Adv. Math.* **189**(2) (2004), 451–494.

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

[59] M. Mendel and A. Naor, 'Some applications of Ball's extension theorem', *Proc. Amer. Math. Soc.* **134**(9) (2006), 2577–2584. (electronic).
[60] M. Mendel and A. Naor, 'Scaled Enflo type is equivalent to Rademacher type', *Bull. Lond. Math. Soc.* **39**(3) (2007), 493–498.
[61] M. Mendel and A. Naor, 'Metric cotype', *Ann. of Math.* (2) **168**(1) (2008), 247–298.
[62] M. Mendel and A. Naor, 'Markov convexity and local rigidity of distorted metrics', *J. Eur. Math. Soc. (JEMS)* **15**(1) (2013), 287–337.
[63] M. Mendel and A. Naor, 'Spectral calculus and Lipschitz extension for barycentric metric spaces', *Anal. Geom. Metr. Spaces* **1** (2013), 163–199.
[64] M. Mendel and A. Naor, 'Nonlinear spectral calculus and super-expanders', *Publ. Math. Inst. Hautes Études Sci.* **119** (2014), 1–95.
[65] M. Mendel and A. Naor, 'Expanders with respect to Hadamard spaces and random graphs', *Duke Math. J.* **164**(8) (2015), 1471–1548.
[66] A. Naor, 'An introduction to the Ribe program', *Jpn. J. Math.* **7**(2) (2012), 167–233.
[67] A. Naor, 'Comparison of metric spectral gaps', *Anal. Geom. Metr. Spaces* **2** (2014), Art. 1.
[68] A. Naor, 'Discrete Riesz transforms and sharp metric X_p inequalities', Preprint, 2016, arXiv:1601.03332.
[69] A. Naor and Y. Peres, 'Embeddings of discrete groups and the speed of random walks', *Int. Math. Res. Not. IMRN* (2008), pages Art. ID rnn 076, 34.
[70] A. Naor and Y. Peres, 'L_p compression, traveling salesmen, and stable walks', *Duke Math. J.* **157**(1) (2011), 53–108.
[71] A. Naor, Y. Peres, O. Schramm and S. Sheffield, 'Markov chains in smooth Banach spaces and Gromov-hyperbolic metric spaces', *Duke Math. J.* **134**(1) (2006), 165–197.
[72] A. Naor and G. Schechtman, 'Remarks on non linear type and Pisier's inequality', *J. Reine Angew. Math.* **552** (2002), 213–236.
[73] A. Naor and L. Silberman, 'Poincaré inequalities, embeddings, and wild groups', *Compos. Math.* **147**(5) (2011), 1546–1572.
[74] S.-i. Ohta, 'Markov type of Alexandrov spaces of non-negative curvature', *Mathematika* **55**(1–2) (2009), 177–189.
[75] M. I. Ostrovskii, *Metric Embeddings*, De Gruyter Studies in Mathematics, 49 (De Gruyter, Berlin, 2013), Bilipschitz and coarse embeddings into Banach spaces.
[76] R. E. A. C. Paley, 'Some theorems on abstract spaces', *Bull. Amer. Math. Soc.* **42**(4) (1936), 235–240.
[77] R. E. A. C. Paley and A. Zygmund, 'On some series of functions, (1)', *Math. Proc. Cambridge Philos. Soc.* **26**(3) (1930), 337–357.
[78] G. Pisier, 'Some results on Banach spaces without local unconditional structure', *Compos. Math.* **37**(1) (1978), 3–19.
[79] G. Pisier, 'Probabilistic methods in the geometry of Banach spaces', in *Probability and Analysis (Varenna, 1985)*, Lecture Notes in Mathematics, 1206 (Springer, Berlin, 1986), 167–241.
[80] G. Pisier, 'Non-commutative vector valued L_p-spaces and completely p-summing maps', *Astérisque* **247** (1998), vi+131.
[81] G. Pisier and Q. Xu, 'Non-commutative L^p-spaces', in *Handbook of the Geometry of Banach Spaces,* Vol. 2 (North-Holland, Amsterdam, 2003), 1459–1517.
[82] M. Ribe, 'On uniformly homeomorphic normed spaces', *Ark. Mat.* **14**(2) (1976), 237–244.
[83] H. P. Rosenthal, 'On the subspaces of $L^p(p > 2$ spanned by sequences of independent random variables', *Israel J. Math.* **8** (1970), 273–303.

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press

[84] I. J. Schoenberg, 'Metric spaces and positive definite functions', *Trans. Amer. Math. Soc.* **44**(3) (1938), 522–536.
[85] F. A. Sukochev, 'Non-isomorphism of L_p-spaces associated with finite and infinite von Neumann algebras', *Proc. Amer. Math. Soc.* **124**(5) (1996), 1517–1527.
[86] M. Talagrand, 'Isoperimetry, logarithmic Sobolev inequalities on the discrete cube, and Margulis' graph connectivity theorem', *Geom. Funct. Anal.* **3**(3) (1993), 295–314.
[87] E. Veomett and K. Wildrick, 'Spaces of small metric cotype', *J. Topol. Anal.* **2**(4) (2010), 581–597.
[88] R. Wagner, 'Notes on an inequality by Pisier for functions on the discrete cube', in *Geometric Aspects of Functional Analysis*, Lecture Notes in Mathematics, 1745 (Springer, Berlin, 2000), 263–268.
[89] J. H. Wells and L. R. Williams, *Embeddings and Extensions in Analysis* (Springer, New York, Heidelberg, 1975), Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 84.
[90] P. Wojtaszczyk, *Banach Spaces for Analysts*, Cambridge Studies in Advanced Mathematics, 25 (Cambridge University Press, Cambridge, 1991).

https://doi.org/10.1017/fmp.2016.1 Published online by Cambridge University Press