DISCRETE RIESZ TRANSFORMS AND SHARP METRIC X_p INEQUALITIES

ASSAF NAOR

Abstract. For $p \in [2,\infty)$ the metric X_p inequality with sharp scaling parameter is proven here to hold true in L_p. The geometric consequences of this result include the following sharp statements about embeddings of L_q into L_p when $2<q<p<\infty$: the maximal $\theta \in (0,1]$ for which L_q admits a bi-θ-Hölder embedding into L_p equals q/p, and for $m,n\in\mathbb{N}$ the smallest possible bi-Lipschitz distortion of any embedding into L_p of the grid $\{1,\ldots,m\}^n \subseteq \ell_q^n$ is bounded above and below by constant multiples (depending only on p,q) of the quantity $\min\{n^{(p-q)(q-2)/(q^2(p-2))}, m^{(q-2)/q}\}$.

1. Introduction

The purpose of the present article is to resolve positively three conjectures that were posed by the author in collaboration with G. Schechtman in [NS14]. Specifically, we shall prove here that Conjecture 1.5, Conjecture 1.8 and Conjecture 1.12 of [NS14] all have a positive answer. As we shall explain below, of these three conjectures, Conjecture 1.8 was a longstanding open problem in embedding theory, while Conjecture 1.12 asserts the validity of a quite subtle and perhaps unexpected phase transition phenomenon that was first formulated as conceivably holding true in [NS14]. Conjecture 1.5 relates to a bi-Lipschitz invariant that was introduced in [NS14], asking about finer properties of this invariant in terms of a certain auxiliary parameter.

It was proven in [NS14] that Conjecture 1.8 and Conjecture 1.12 follow from Conjecture 1.5. Thus Conjecture 1.5 is the heart of the matter and the main focus of the present article, but we shall first describe all of the above conjectures since, by proving their validity, we establish delicate geometric phenomena related to the metric structure of L_p spaces. In addition to these applications, a key contribution of the present article is the use of a deep result of Lust-Piquard [LP98] for geometric purposes. While [NS14] proposed an approach to resolve the above conjectures, formulated as Question 6.1 in [NS14] and discussed at length in [NS14, Section 6], where it was shown to imply the above conjectures, we do not pursue this approach here, and indeed Question 6.1 of [NS14] remains open. Below we take a different route, yielding a novel connection between purely geometric questions and investigations in modern harmonic analysis and operator algebras.

1.1. **Geometric statements.** Following standard notation in Banach space theory and embedding theory (as in, say, [LT77, Ost13]), for $n\in\mathbb{N}$ and $p\in[1,\infty)$ we let ℓ_p^n denote the space $\mathbb{R}^n$ equipped with the ℓ_p norm. When referring to the space L_p, we mean for concreteness the Lebesgue space $L_p(\mathbb{R})$, though all of our new geometric results apply equally well to any infinite dimensional $L_p(\mu)$ space. The L_p distortion of a metric space (X,d_X), denoted $c_p(X)\in[0,\infty]$, is the infimum over those $D\in[0,\infty]$ for which there exists a mapping $f:X\to L_p$ that satisfies

$$\forall\, x,y\in X,\qquad d_X(x,y)\leqslant \|f(x)-f(y)\|_{L_p}\leqslant Dd_X(x,y).$$

(X,d_X) is said to admit a bi-Lipschitz embedding into L_p if $c_p(X)<\infty$.

Given $m,n\in\mathbb{N}$ and $q\in[1,\infty)$, the metric space whose underlying set is $\{1,\ldots,m\}^n$ (the m-grid in $\mathbb{R}^n$), equipped with the metric inherited from ℓ_q^n, will be denoted below by $[m]_q^n$. It follows from the classical work [Pal36] of Paley, in combination with general principles related to differentiation of Lipschitz functions (see [BL00, Chapter 7]), that if $2<q<p<\infty$ then $\lim_{n\to\infty} c_p(\ell_q^n)=\infty$.

Supported in part by the BSF, the Packard Foundation and the Simons Foundation.

Since $[m]_q^n$ becomes "closer" to ℓ_q^n as $m \to \infty$, one can apply an ultrapower argument (see [Hei80]) to deduce from this that $\lim_{m,n\to\infty} c_p([m]_q^n) = \infty$, but such reasoning does not yield information on the rate of growth of $c_p([m]_q^n)$. Effective estimates here follow from an alternative approach of Bourgain [Bou87] (with an improvement in [GNS12]), as well as the approach of [NS14], but the resulting bounds are far from being sharp. Resolving Conjecture 1.12 of [NS14], Theorem 1 below computes the quantity $c_p([m]_q^n)$ up to constant factors that may depend on p, q but not on m, n.

Theorem 1 (Sharp evaluation of the L_p distortion of ℓ_q^n grids)**.** *Suppose that $p, q \in [2,\infty)$ satisfy $q < p$. Then for every $m, n \in \mathbb{N}$ we have*

$$c_p\big([m]_q^n\big) \asymp_{p,q} \min\left\{ n^{\frac{(p-q)(q-2)}{q^2(p-2)}}, m^{1-\frac{2}{q}} \right\} = \left(\min\left\{ n^{\frac{p-q}{q(p-2)}}, m \right\} \right)^{1-\frac{2}{q}}. \qquad (1)$$

In the statement of Theorem 1, as well as in what follows, we use standard asymptotic notation. Namely, 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 $(a \lesssim b) \wedge (b \lesssim a)$. When we allow for implicit constants to depend on parameters, we indicate this by subscripts. Thus $a \lesssim_{p,q} b$ (respectively $a \gtrsim_{p,q} b$) means that there exists $c(p,q) \in (0,\infty)$ that may depend only on p, q such that $a \leqslant c(p,q)b$ (respectively $a \geqslant c(p,q)b$). The notation $a \asymp_{p,q} b$ stands for $(a \lesssim_{p,q} b) \wedge (b \lesssim_{p,q} a)$.

Very few results at the level of precision of Theorem 1 are known, and analogous questions are open even for some values of p, q that are not covered by Theorem 1; see [NS14, Remark 1.13] for more on this interesting topic. The asymptotic formula (1) expresses the statement that there exist two specific embeddings of $[m]_q^n$ into L_p such that one of them is always the best possible embedding of $[m]_q^n$ into L_p, up to constant factors that do not depend on m, n. One of these embeddings arises from the work of Rosenthal [Ros70] (relying also on computations in [GPP80, FJS88]), and the other is due to Mendel and the author [MN06] (relying also on a construction from [Sch38]). These issues, including precise descriptions of the above two embeddings, are explained in detail in [NS14].

The following immediate corollary of Theorem 1 asserts that if $2 < q < p < \infty$ and $m, n \in \mathbb{N}$ then the L_p distortion of $[m]_q^n$ exhibits a phase transition at $m \asymp n^{(p-q)/(q(p-2))}$.

Corollary 2 (Sharp phase transition of the L_p distortion of ℓ_q^n grids)**.** *Suppose that $m, n \in \mathbb{N}$ and $p, q \in (2,\infty)$ satisfy $q < p$. Then*

$$m \gtrsim n^{\frac{p-q}{q(p-2)}} \implies c_p\big([m]_q^n\big) \asymp_{p,q} c_p\big(\ell_q^n\big),$$

while as $n \to \infty$ we have

$$m = o\left(n^{\frac{p-q}{q(p-2)}}\right) \implies c_p\big([m]_q^n\big) = o\left(c_p\big(\ell_q^n\big)\right).$$

Thus, to state one concrete example so as to illustrate the situation whose validity we establish here, when, say, $q = 3$ and $p = 4$, and one tries to embed the grid $[m]_3^n$ into L_4, one sees that there is a phase transition at $m \asymp \sqrt[6]{n}$. If $m \gtrsim \sqrt[6]{n}$ then any embedding of $[m]_3^n$ into L_4 incurs the same distortion (up to universal constant factors) as the distortion required to embed all of ℓ_3^n into L_4, which grows like $\sqrt[18]{n}$. However, if $m = o(\sqrt[6]{n})$ then one can embed $[m]_3^n$ into L_4 with distortion $o(\sqrt[18]{n})$, and in this case the L_4 distortion of $[m]_3^n$ is $\sqrt[3]{m}$, up to universal constant factors.

Our second geometric result is Theorem 3 below, which resolves Conjecture 1.8 of [NS14].

Theorem 3 (Evaluation of the critical L_p snowflake exponent of L_q)**.** *Suppose that $p, q \in (2,\infty)$ satisfy $q < p$. Then the maximal $\theta \in (0,1]$ for which the metric space $(L_q, \|x-y\|_{L_q}^\theta)$ admits a bi-Lipschitz embedding into L_p equals q/p.*

In the setting of Theorem 3, the fact that the metric space $(L_q, \|x-y\|_{L_q}^{q/p})$ does indeed admit a bi-Lipschitz (even isometric) embedding into L_p was established by Mendel and the author in [MN04].

Since then, it has been a well known conjecture that in this context the Hölder exponent q/p cannot be increased, but before [NS14] it wasn't even known that if $(L_q, \|x-y\|_{L_q}^\theta)$ admits a bi-Lipschitz embedding into L_p then necessarily $\theta < 1-\delta$ for some $\delta = \delta(p,q) > 0$. Note that the endpoint case $q = 2$ must be removed from Theorem 3 since L_2 embeds isometrically into L_p.

1.2. **Optimal scaling in the L_p-valued metric X_p inequality.** In what follows, given $n \in \mathbb{N}$ we shall denote the set $\{1,\ldots,n\}$ by $[n]$. The coordinate basis of $\mathbb{R}^n$ will be denoted by $e_1,\ldots,e_n$, and for a sign vector $\varepsilon = (\varepsilon_1,\ldots,\varepsilon_n) \in \{-1,1\}^n$ and a subset $\mathsf{S} \subseteq [n]$ we shall use the notation

$$\varepsilon_{\mathsf{S}} \stackrel{\mathrm{def}}{=} \sum_{j\in \mathsf{S}} \varepsilon_j e_j. \tag{2}$$

Fix $p \in (0,\infty)$. Following [NS14], a metric space (X, d_X) is said to be an X_p metric space if there exists $\mathfrak{X} \in (0,\infty)$ such that for every $n \in \mathbb{N}$ and $k \in [n]$ there exists $m \in \mathbb{N}$ such that, denoting as usual $\mathbb{Z}_{2m} = \mathbb{Z}/(2m\mathbb{Z})$, every function $f : \mathbb{Z}_{2m}^n \to X$ satisfies the following distance inequality.

$$\begin{aligned}\bigg(\frac{1}{\binom{n}{k}} \sum_{\substack{S\subseteq [n]\\ |S|=k}} \mathbb{E}\Big[d_X\big(f(x+m\varepsilon_S), f(x)\big)^p\Big]\bigg)^{\frac1p}\\ \leqslant \mathfrak{X} m \bigg(\frac{k}{n}\sum_{j=1}^n \mathbb{E}\Big[d_X\big(f(x+e_j), f(x)\big)^p\Big] + \Big(\frac{k}{n}\Big)^{\frac p2} \mathbb{E}\Big[d_X\big(f(x+\varepsilon), f(x)\big)^p\Big]\bigg)^{\frac1p}.\end{aligned} \tag{3}$$

The expectations in (3) are with respect to $(x,\varepsilon) \in \mathbb{Z}_{2m}^n \times \{-1,1\}^n$ chosen uniformly at random. We refer to [NS14] for a detailed discussion of the meaning of (3); see also Sections 1.2.1, 1.3 below.

The above definition of X_p metric spaces introduces the auxiliary integer $m \in \mathbb{N}$, which we call the *scaling parameter* corresponding to n and k. For some purposes m can be allowed to be arbitrary, but for other purposes one needs to obtain good bounds on m (as a function of n,k). It can, however, be quite difficult to obtain sharp bounds on scaling parameters in metric inequalities (for example, an analogous question in the context of metric cotype [MN08] is longstanding and important). In [NS14] it was proven that if $p \in [2,\infty)$ then L_p is an X_p metric space. The proof in [NS14] yields the validity of (3) when $X = L_p$ whenever $m \gtrsim_p n^{3/2}/\sqrt{k}$. It was also shown in [NS14, Proposition 1.4] that if $p \in (2,\infty)$ and k is sufficiently large (as a function of p) then for (3) to hold true in L_p one must necessarily have $m \gtrsim_p \sqrt{n/k}$. Conjecture 1.5 of [NS14] asks whether for every $p \in (2,\infty)$ this lower bound on m actually expresses the asymptotic behavior of the best possible scaling parameter, i.e., whether the metric X_p inequality (3) holds true in L_p for every $m \gtrsim_p \sqrt{n/k}$. Theorem 4 below resolves this conjecture positively.

Theorem 4 (L_p is an X_p metric space with sharp scaling parameter)**.** *Suppose that $k,m,n \in \mathbb{N}$ satisfy $k \in [n]$ and $m \geqslant \sqrt{n/k}$. Suppose also that $p \in [2,\infty)$. Then every $f : \mathbb{Z}_{8m}^n \to L_p$ satisfies*

$$\begin{aligned}\bigg(\frac{1}{\binom{n}{k}} \sum_{\substack{\mathsf{S}\subseteq [n]\\ |\mathsf{S}|=k}} \mathbb{E}\Big[\,\|f(x+4m\varepsilon_{\mathsf{S}}) - f(x)\|_{L_p}^p\Big]\bigg)^{\frac1p}\\ \lesssim_p m \bigg(\frac{k}{n}\sum_{j=1}^n \mathbb{E}\Big[\|f(x+e_j) - f(x)\|_{L_p}^p\Big] + \Big(\frac{k}{n}\Big)^{\frac p2} \mathbb{E}\Big[\|f(x+\varepsilon) - f(x)\|_{L_p}^p\Big]\bigg)^{\frac1p},\end{aligned} \tag{4}$$

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

Remark 5. Our proof of Theorem 4 shows that the implicit constant in (4) is $O(p^4/\log p)$. As explained in [NS14], this constant must be at least a (universal) constant multiple of $p/\log p$. While it is conceivable that a more careful implementation of our approach could somewhat decrease the dependence on p that we obtain, it seems that a new idea is required in order to establish the sharp dependence of $O(p/\log p)$ in (4) (if true). We leave the question of determining the correct asymptotic dependence on p in (4) as an interesting (and perhaps quite challenging) open question.

1.2.1. *Applications of Theorem 4.* The usefulness of the metric X_p inequality for L_p stems in part from the fact that it allows one to rule out the existence of metric embeddings in situations where the classical differentiation techniques fail. Examples of such situations include the treatment of discrete sets as in Theorem 1, where it isn't clear how to interpret the notion of derivative, as well as the treatment of Hölder mappings as in Theorem 3, where, unlike the Lipschitz case, mappings need not have any point of differentiability. In fact, by [NS14, Theorem 1.14] both Theorem 1 and Theorem 3 follow from Theorem 4. For completeness, we shall now briefly sketch why this is so.

Suppose that $2 \leqslant q < p < \infty$ and $m,n \in \mathbb{N}$. It is simple to check, as done in [NS14, Lemma 3.1], that there exists $h : \mathbb{Z}_{8m}^n \to [32m]_q^n$ such that for $(x,\varepsilon) \in \mathbb{Z}_{8m}^n \times \{-1,1\}^n$, $\mathsf{S} \subseteq [n]$ and $j \in [n]$,

$$\|h(x+4m\varepsilon_{\mathsf{S}})-h(x)\|_{\ell_q^n} \asymp m|\mathsf{S}|^{\frac{1}{q}} \quad \text{and} \quad \|h(x+e_j)-h(x)\|_{\ell_q^n} \asymp 1 \quad \text{and} \quad \|h(x+\varepsilon)-h(x)\|_{\ell_q^n} \asymp n^{\frac{1}{q}}.$$

Fix $D \in [1,\infty)$ and suppose that $\phi : [32m]_q^n \to L_p$ satisfies $\|x-y\|_{\ell_q^n} \leqslant \|\phi(x)-\phi(y)\|_{L_p} \leqslant D\|x-y\|_{\ell_q^n}$ for every $x,y \in [32m]_q^n$. An application of Theorem 4 to $f = \phi \circ h$ (with m replaced by $4m$), which we are allowed to do only when $k \in [n]$ is such that $4m \geqslant \sqrt{n/k}$, yields the bound

$$D \gtrsim_p \max_{\substack{k\in[n]\\ k\geqslant n/(16m^2)}} \frac{k^{\frac{1}{q}}}{\left(k + k^{\frac{p}{2}} n^{\frac{p}{q}-\frac{p}{2}}\right)^{\frac{1}{p}}}. \tag{5}$$

By evaluating the maximum in (5), one arrives at the asymptotic lower bound on $c_p([32m]_q^n)$ that appears in (1). As we explained earlier, the matching upper bound in (1) corresponds to the better of two explicit embeddings that are described in equations (11) and (27) of [NS14]. This completes the deduction of Theorem 1. Next, fix $L \in [1,\infty)$ and $\theta \in (0,1]$. Suppose that $\psi : L_q \to L_p$ satisfies $\|x-y\|_{L_q}^{\theta} \leqslant \|\psi(x)-\psi(y)\|_{L_p} \leqslant L\|x-y\|_{L_q}^{\theta}$ for every $x,y \in L_q$. For $k,n \in \mathbb{N}$ with $k \in [n]$, fix $m = \lceil \sqrt{n/(2k)} \rceil$ and apply Theorem 4 to $f = \psi \circ h$. The estimate thus obtained is

$$\left(\frac{n}{k}\right)^{\frac{\theta}{2}} k^{\frac{\theta}{q}} \lesssim_p L\sqrt{\frac{n}{k}} \left(k + \left(\frac{k}{n}\right)^{\frac{p}{2}} n^{\frac{\theta p}{q}}\right)^{\frac{1}{p}}.$$

Hence, for every $n \in \mathbb{N}$ we have

$$1 \lesssim_p L n^{\frac{1-\theta}{2}} \cdot \min_{k\in[n]} \left(k + k^{\frac{p}{2}} n^{p\left(\frac{\theta}{q}-\frac{1}{2}\right)}\right)^{\frac{1}{p}} k^{\theta\left(\frac{1}{2}-\frac{1}{q}\right)-\frac{1}{2}}. \tag{6}$$

Theorem 3 now follows by choosing the optimal k in (6) and letting $n \to \infty$; complete details of this computation appear in the proof of Theorem 1.14 in [NS14].

1.3. **Hypercube Riesz transforms and an X_p inequality for Rademacher chaos.** Fixing $n \in \mathbb{N}$, for every $h : \{-1,1\}^n \to \mathbb{R}$ and $j \in [n]$ let $\partial_j h : \{-1,1\}^n \to \mathbb{R}$ be given by

$$\forall\, \varepsilon \in \{-1,1\}^n, \qquad \partial_j h(\varepsilon) \stackrel{\mathrm{def}}{=} h(\varepsilon) - h(\varepsilon_1,\ldots,\varepsilon_{j-1},-\varepsilon_j,\varepsilon_{j+1},\ldots,\varepsilon_n). \tag{7}$$

Also, given $\mathsf{S} \subseteq [n]$ we shall denote by $\mathsf{E}_{\mathsf{S}} f : \{-1,1\}^n \to \mathbb{R}$ the function that is obtained from h by averaging over the coordinates in S, i.e., recalling the notation (2), we define

$$\forall\, \varepsilon \in \{-1,1\}^n, \qquad \mathsf{E}_{\mathsf{S}} h(\varepsilon) \stackrel{\mathrm{def}}{=} \frac{1}{2^n} \sum_{\delta\in\{-1,1\}^n} h\left(\delta_{\mathsf{S}} + \varepsilon_{[n]\smallsetminus\mathsf{S}}\right). \tag{8}$$

In particular, $\mathsf{E}_{\mathsf{S}}h$ depends only on those entries of $\varepsilon\in\{-1,1\}^n$ that belong to $[n]\smallsetminus\mathsf{S}$. Given $p\in[1,\infty)$, we shall reserve from now on the notation $\|h\|_p$ exclusively for the L_p norm of h with respect to the *normalized* counting measure on the discrete hypercube $\{-1,1\}^n$, i.e.,

$$\|h\|_p\stackrel{\mathrm{def}}{=}\bigg(\frac{1}{2^n}\sum_{\varepsilon\in\{-1,1\}^n}|h(\varepsilon)|^p\bigg)^{\frac{1}{p}}=\big(\mathsf{E}_{[n]}|h|^p\big)^{\frac{1}{p}}.$$

In what follows, $L_p^0(\{-1,1\}^n)$ denotes the subspace of all those $h\in L_p(\{-1,1\}^n)$ with $\mathsf{E}_{[n]}h=0$.

We shall work with the usual Fourier–Walsh expansion of a function $h:\{-1,1\}^n\to\mathbb{R}$. Thus, for every $\mathsf{A}\subseteq[n]$ consider the corresponding Walsh function $W_{\mathsf{A}}:\{-1,1\}^n\to\mathbb{R}$ given by

$$\forall\,\varepsilon\in\{-1,1\}^n,\qquad W_{\mathsf{A}}(\varepsilon)\stackrel{\mathrm{def}}{=}\prod_{j\in\mathsf{A}}\varepsilon_j,$$

and denote

$$\widehat{h}(\mathsf{A})\stackrel{\mathrm{def}}{=}\frac{1}{2^n}\sum_{\varepsilon\in\{-1,1\}^n}h(\varepsilon)W_{\mathsf{A}}.$$

Then we have

$$\forall\,\varepsilon\in\{-1,1\}^n,\qquad h(\varepsilon)=\frac{1}{2^n}\sum_{\mathsf{A}\subseteq[n]}\widehat{h}(\mathsf{A})W_{\mathsf{A}}(\varepsilon).$$

In probabilistic terminology, the above representation of h as a multilinear polynomial in the variables $\varepsilon_1,\ldots,\varepsilon_n$ expresses it as Rademacher chaos. A useful inequality for Rademacher chaos of the first degree, i.e., for weighted sums of i.i.d. Bernoulli random variables, served as the inspiration for the metric X_p inequality (3). Specifically, (3) is a nonlinear extension of the following inequality, which holds true for every $p\in[2,\infty)$, $k,n\in\mathbb{N}$ with $k\in[n]$, and every $a_1,\ldots,a_n\in\mathbb{R}$.

$$\Bigg(\frac{1}{2^n\binom{n}{k}}\sum_{\substack{\mathsf{S}\subseteq[n]\\|\mathsf{S}|=k}}\sum_{\varepsilon\in\{-1,1\}^n}\Big|\sum_{j\in\mathsf{S}}\varepsilon_ja_j\Big|^p\Bigg)^{\frac{1}{p}}\lesssim\frac{p}{\log p}\Bigg(\frac{k}{n}\sum_{j=1}^n|a_j|^p+\frac{(k/n)^{\frac{p}{2}}}{2^n}\sum_{\varepsilon\in\{-1,1\}^n}\Big|\sum_{j=1}^n\varepsilon_ja_j\Big|^p\Bigg)^{\frac{1}{p}}.\tag{9}$$

This inequality is due to Johnson, Maurey, Schechtman and Tzafriri, who proved it in [JMST79] with a constant factor that grows to ∞ with p faster than the $p/\log p$ factor that appears in (9). The factor $p/\log p$ that is stated in (9) is best possible; in the above sharp form, (9) is due to Johnson, Schechtman and Zinn [JSZ85]. As a step towards Theorem 4, we shall prove the following theorem in Section 3 below, thus extending (9) to Rademacher chaos of arbitrary degree.

Theorem 6 (X_p inequality for Rademacher chaos)**.** *Suppose that $p\in[2,\infty)$, $n\in\mathbb{N}$ and $k\in[n]$. Then every $h\in L_p^0(\{-1,1\}^n)$ satisfies*

$$\Bigg(\frac{1}{\binom{n}{k}}\sum_{\substack{\mathsf{S}\subseteq[n]\\|\mathsf{S}|=k}}\big\|\mathsf{E}_{[n]\smallsetminus\mathsf{S}}h\big\|_p^p\Bigg)^{\frac{1}{p}}\lesssim_p\Bigg(\frac{k}{n}\sum_{j=1}^n\|\partial_jh\|_p^p+\Big(\frac{k}{n}\Big)^{\frac{p}{2}}\|h\|_p^p\Bigg)^{\frac{1}{p}}.\tag{10}$$

The deduction of Theorem 4 from Theorem 6 appears in Section 2 below.

Remark 7. As in (4), the implicit constant that we obtain in (10) is $O(p^4/\log p)$. In fact, our proof yields the following slightly more refined estimate in the setting of Theorem 6.

$$\Bigg(\frac{1}{\binom{n}{k}}\sum_{\substack{\mathsf{S}\subseteq[n]\\|\mathsf{S}|=k}}\big\|\mathsf{E}_{[n]\smallsetminus\mathsf{S}}h\big\|_p^p\Bigg)^{\frac{1}{p}}\lesssim\frac{p^{\frac{5}{2}}}{\sqrt{\log p}}\Big(\frac{k}{n}\Big)^{\frac{1}{p}}\Bigg(\sum_{j=1}^n\|\partial_jh\|_p^p\Bigg)^{\frac{1}{p}}+\frac{p^4}{\log p}\sqrt{\frac{k}{n}}\cdot\|h\|_p.\tag{11}$$

It remains open to determine the growth rate as $p\to\infty$ of the implicit constant in (10).

1.3.1. *Lust-Piquard's work.* Our proof of Theorem 6 uses deep work [LP98] of Lust-Piquard on dimension-free bounds for discrete Riesz transforms. Specifically, for every $h:\{-1,1\}^n\to\mathbb{R}$ and $j\in[n]$ the jth (hypercube) Riesz transform of h, denoted $\mathsf{R}_j h:\{-1,1\}^n\to\mathbb{R}$, is defined as follows.

$$\forall\,\varepsilon\in\{-1,1\}^n,\qquad \mathsf{R}_j h(\varepsilon)\stackrel{\mathrm{def}}{=}\sum_{\substack{\mathsf{A}\subseteq[n]\\ j\in\mathsf{A}}}\frac{\widehat{h}(\mathsf{A})}{\sqrt{|\mathsf{A}|}}W_{\mathsf{A}}(\varepsilon). \tag{12}$$

Lust-Piquard proved the following inequalities, which hold true for $p\in[2,\infty)$ and $h\in L_p^0(\{-1,1\}^n)$.

$$\frac{1}{p^{3/2}}\|h\|_p\lesssim\bigg\|\Big(\sum_{j=1}^n(\mathsf{R}_j h)^2\Big)^{\frac12}\bigg\|_p\lesssim p\|h\|_p. \tag{13}$$

The inequalities in (13) were proved by Lust-Piquard in [LP98], though with a dependence on p that is worse than what we stated above. The dependence on p that appears in (13) follows from [BELP08] (specifically, the left hand inequality in (13) follows from [BELP08, Theorem 1.1] and the right hand inequality in (13) follows from [BELP08, Theorem 5.1]. Actually, the right hand inequality in (13) with the stated dependence on p follows implicitly from ingredients that are already present in [LP98]). Note that these estimates are stated in [BELP08] in terms of the strong (p,p) norm of the Hilbert transform with values in the Schatten–von Neumann trace class S_p, but this norm was shown to be $O(p)$ by Bourgain in [Bou86], and the bounds that we stated in (13) result from a direct substitution of Bourgain's bound into the statements in [BELP08].

The availability of dimension independent bounds for Riesz transforms is a well known paradigm in other (non-discrete) settings, originating from important classical work of Stein [Ste83] in the case of $\mathbb{R}^n$ equipped with Lebesgue measure (see also [GV79, DRdF85, Bañ86]). Most pertinent to the present context is the classical theorem of P. A. Meyer [Mey84] (see also [Gun86]) that obtained dimension independent bounds for the Riesz transforms that are associated to $\mathbb{R}^n$ equipped with the Gaussian measure (and the Ornstein–Uhlenbeck operator). Pisier discovered in [Pis88] an influential alternative proof of P. A. Meyer's theorem, based on a transference argument (see [CW76]) that allows one to reduce the question to the boundedness of the (one dimensional) Hilbert transform.

Lust-Piquard's work generally follows Pisier's strategy, but it also uncovers a phenomenon that is genuinely present in the hypercube setting and *not* in the Gaussian setting. Specifically, Lust-Piquard reduces the task of bounding the hypercube Riesz transforms to that of bounding the S_p norm of certain operators in a noncommutative $*$ algebra of (2^n by 2^n) matrices, and proceeds to do so using operator-theoretic methods, including her noncommutative Khinchine inequalities [LP86].

This indicates why the S_p-valued Hilbert transform makes its appearance in Lust-Piquard's inequality (recall the paragraph above, immediately following (13)), despite the fact that (13) deals with real-valued functions on the (commutative) hypercube. Significantly, while the classical results on Riesz transforms (with respect to either Lebesgue measure or the Gaussian measure) yield dimension independent bounds for every $p\in(1,\infty)$, it turns out that (13) actually fails to hold true when $p\in(1,2)$, as explained in [LP98] (where this observation is attributed to unpublished work of Lamberton); see also [BELP08, Section 5.5]. The reason for this disparity between the ranges $p\in(1,2)$ and $p\in[2,\infty)$ becomes clear when one transfers the question to the noncommutative setting, and this suggests a more complicated (but still dimension-free) replacement for (13) in the range $p\in(1,2)$, which Lust-Piquard also proved in [LP98]. So, while it is conceivable that a proof of (13) could be found that does not proceed along Lust-Piquard's noncommutative route, such a proof has not been found to date, and the qualitative divergence between the discrete situation and its continuous counterparts indicates that there may be an inherently different phenomenon at play here. Since its initial publication, Lust-Piquard's work influenced developments by herself and others that focused on proving related inequalities in other situations; we do not have anything new to add to this interesting body of work other than showing here that in addition to their intrinsic interest, such results can have a decisive role in understanding geometric embedding questions.

2. Deduction of Theorem 4 from Theorem 6

Assuming the validity of Theorem 6 for the moment, we shall now proceed to show how it implies Theorem 4. Note that since (4) involves only the pth powers of distances in L_p, by integration it suffices to prove Theorem 4 for real valued functions. So, from now on we shall assume that $m,n\in\mathbb{N}$ and we are given a function $f:\mathbb{Z}_{8m}^n\to\mathbb{R}$, the goal being to prove the validity of (4) for every $k\in[n]$ provided that $m\geqslant\sqrt{n/k}$, with the L_p norms replaced by absolute values in $\mathbb{R}$.

In what follows, given $\mathsf{S}\subseteq[n]$ and $f:\mathbb{Z}_{8m}^n\to\mathbb{R}$, define a function $\mathsf{T}_\mathsf{S}f:\mathbb{Z}_{8m}^n\to\mathbb{R}$ by

$$\forall\, x\in\mathbb{Z}_{8m}^n,\qquad \mathsf{T}_\mathsf{S}f(x)\stackrel{\mathrm{def}}{=}\frac{1}{2^n}\sum_{\delta\in\{-1,1\}^n}f(x+2\delta_\mathsf{S}).\tag{14}$$

We record for future use the following simple lemma.

Lemma 8. *For every $p\in[1,\infty)$, $m,n\in\mathbb{N}$, $\mathsf{S}\subseteq[n]$ and $f:\mathbb{Z}_{8m}^n\to\mathbb{R}$ we have*

$$\bigg(\frac{1}{(8m)^n}\sum_{x\in\mathbb{Z}_{8m}^n}|f(x)-\mathsf{T}_\mathsf{S}f(x)|^p\bigg)^{\frac1p}\leqslant 2\bigg(\frac{1}{(16m)^n}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}|f(x+\varepsilon)-f(x)|^p\bigg)^{\frac1p}.\tag{15}$$

Proof. By convexity, for every $x\in\mathbb{Z}_{8m}^n$ we have

$$\begin{aligned}|f(x)-\mathsf{T}_\mathsf{S}f(x)|^p&\leqslant\frac{1}{2^n}\sum_{\delta\in\{-1,1\}^n}|f(x)-f(x+2\delta_\mathsf{S})|^p\\&\leqslant\frac{2^{p-1}}{2^n}\sum_{\delta\in\{-1,1\}^n}\Big(\big|f(x)-f(x+\delta_\mathsf{S}+\delta_{[n]\smallsetminus\mathsf{S}})\big|^p+\big|f(x+\delta_\mathsf{S}+\delta_{[n]\smallsetminus\mathsf{S}})-f(x+2\delta_\mathsf{S})\big|^p\Big).\end{aligned}\tag{16}$$

The desired estimate (15) follows by averaging (16) over $x\in\mathbb{Z}_{8m}^n$ while using the translation invariance of the uniform measure on $\mathbb{Z}_{8m}^n$, and that if δ is uniformly distributed over $\{-1,1\}^n$ then the sign vectors $\delta_\mathsf{S}+\delta_{[n]\smallsetminus\mathsf{S}}$ and $-\delta_\mathsf{S}+\delta_{[n]\smallsetminus\mathsf{S}}$ are both also uniformly distributed over $\{-1,1\}^n$. □

Lemma 9. *Suppose that $m,n\in\mathbb{N}$ and $k\in[n]$. If $p\in[2,\infty)$ then every $f:\mathbb{Z}_{8m}^n\to\mathbb{R}$ satisfies*

$$\begin{aligned}&\frac{1}{(16m)^n\binom{n}{k}}\sum_{\substack{\mathsf{S}\subseteq[n]\\|\mathsf{S}|=k}}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}\frac{\big|\mathsf{T}_{[n]\smallsetminus\mathsf{S}}f(x+4m\varepsilon_\mathsf{S})-\mathsf{T}_{[n]\smallsetminus\mathsf{S}}f(x)\big|^p}{m^p}\\&\qquad\lesssim_p\frac{k/n}{(8m)^n}\sum_{j=1}^n\sum_{x\in\mathbb{Z}_{8m}^n}|f(x+e_j)-f(x)|^p+\frac{(k/n)^{\frac p2}}{(16m)^n}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}|f(x+\varepsilon)-f(x)|^p.\end{aligned}\tag{17}$$

Proof. For every fixed $\mathsf{S}\subseteq[n]$ we have

$$\begin{aligned}&\bigg(\frac{1}{(16m)^n}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}\big|\mathsf{T}_{[n]\smallsetminus\mathsf{S}}f(x+4m\varepsilon_\mathsf{S})-\mathsf{T}_{[n]\smallsetminus\mathsf{S}}f(x)\big|^p\bigg)^{\frac1p}\\&\qquad\leqslant\sum_{k=1}^m\bigg(\frac{1}{(16m)^n}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}\big|\mathsf{T}_{[n]\smallsetminus\mathsf{S}}f(x+4k\varepsilon_\mathsf{S})-\mathsf{T}_{[n]\smallsetminus\mathsf{S}}f(x+4(k-1)\varepsilon_\mathsf{S})\big|^p\bigg)^{\frac1p}\qquad(18)\\&\qquad=m\bigg(\frac{1}{(16m)^n}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{y\in\mathbb{Z}_{8m}^n}\big|\mathsf{T}_{[n]\smallsetminus\mathsf{S}}f(y+2\varepsilon_\mathsf{S})-\mathsf{T}_{[n]\smallsetminus\mathsf{S}}f(y-2\varepsilon_\mathsf{S})\big|^p\bigg)^{\frac1p},\qquad(19)\end{aligned}$$

where for (19) make the change of variable $y = x + 2(2k+1)\varepsilon_S$ in each of the summands of (18).

For every $x \in \mathbb{Z}_{8m}^n$ define $h_x : \{-1,1\}^n \to \mathbb{R}$ by

$$\forall\, \varepsilon \in \{-1,1\}^n, \qquad h_x(\varepsilon) \stackrel{\mathrm{def}}{=} f(x+2\varepsilon) - f(x-2\varepsilon). \tag{20}$$

Recalling (8) and (14), observe that for every $(x,\varepsilon) \in \mathbb{Z}_{8m}^n \times \{-1,1\}^n$ and $\mathsf{S} \subseteq [n]$ we have

$$\mathsf{T}_{[n]\smallsetminus \mathsf{S}} f(x+2\varepsilon_{\mathsf{S}}) - \mathsf{T}_{[n]\smallsetminus \mathsf{S}} f(x-2\varepsilon_{\mathsf{S}}) = \mathsf{E}_{[n]\smallsetminus \mathsf{S}} h_x(\varepsilon).$$

It therefore follows from (19) that

$$\begin{aligned} &\frac{1}{(16m)^n\binom{n}{k}} \sum_{\substack{\mathsf{S}\subseteq[n]\\ |\mathsf{S}|=k}} \sum_{\varepsilon\in\{-1,1\}^n} \sum_{x\in\mathbb{Z}_{8m}^n} \frac{\left|\mathsf{T}_{[n]\smallsetminus \mathsf{S}} f(x+4m\varepsilon_{\mathsf{S}}) - \mathsf{T}_{[n]\smallsetminus \mathsf{S}} f(x)\right|^p}{m^p} \\ &\leqslant \frac{1}{(8m)^n\binom{n}{k}} \sum_{\substack{\mathsf{S}\subseteq[n]\\ |\mathsf{S}|=k}} \sum_{x\in\mathbb{Z}_{8m}^n} \left\|\mathsf{E}_{[n]\smallsetminus \mathsf{S}} h_x\right\|_p^p \lesssim_p \sum_{j=1}^n \frac{k/n}{(8m)^n} \sum_{x\in\mathbb{Z}_{8m}^n} \|\partial_j h_x\|_p^p + \frac{(k/n)^{\frac{p}{2}}}{(8m)^n} \sum_{x\in\mathbb{Z}_{8m}^n} \|h_x\|_p^p, \end{aligned} \tag{21}$$

where in the last step of (21) we applied Theorem 6 with h replaced by h_x, separately for each $x \in \mathbb{Z}_{8m}^n$, which we are allowed to do because the function h_x is odd, so $h_x \in L_p^0(\{-1,1\}^n)$.

Next, observe that for every $(x,\varepsilon) \in \mathbb{Z}_{8m}^n \times \{-1,1\}^n$ and $j \in [n]$ we have

$$\begin{aligned} |\partial_j h_x(\varepsilon)|^p &\stackrel{(20)}{=} |f(x+2\varepsilon) - f(x-2\varepsilon) - f(x+2\varepsilon-4\varepsilon_j e_j) + f(x-2\varepsilon+4\varepsilon_j e_j)|^p \\ &\leqslant 2^{p-1}|f(x+2\varepsilon) - f(x+2\varepsilon-4\varepsilon_j e_j)|^p + 2^{p-1}|f(x-2\varepsilon) - f(x-2\varepsilon+4\varepsilon_j e_j)|^p. \end{aligned} \tag{22}$$

By summing (22) over $(x,\varepsilon) \in \mathbb{Z}_{8m}^n \times \{-1,1\}^n$, we therefore see that

$$\forall\, j \in [n], \qquad \frac{1}{(8m)^n} \sum_{x\in\mathbb{Z}_{8m}^n} \|\partial_j h_x\|_p^p \leqslant \frac{2^p}{(8m)^n} \sum_{y\in\mathbb{Z}_{8m}^n} |f(y+4e_j) - f(y)|^p. \tag{23}$$

Since for every $y \in \mathbb{Z}_{8m}^n$ we have

$$|f(y+4e_j) - f(y)|^p \leqslant 4^{p-1} \sum_{k=1}^4 |f(y+ke_j) - f(y+(k-1)e_j)|^p,$$

it follows from (23) that

$$\frac{1}{(8m)^n} \sum_{j=1}^n \sum_{x\in\mathbb{Z}_{8m}^n} \|\partial_j h_x\|_p^p \leqslant \frac{8^p}{(8m)^n} \sum_{j=1}^n \sum_{z\in\mathbb{Z}_{8m}^n} |f(z+e_j) - f(z)|^p. \tag{24}$$

In the same vein to the above reasoning, for every $(x,\varepsilon) \in \mathbb{Z}_{8m}^n \times \{-1,1\}^n$ we have

$$|h_x(\varepsilon)|^p \stackrel{(20)}{\leqslant} 4^{p-1} \sum_{k=-1}^2 |f(x+k\varepsilon) - f(x+(k-1)\varepsilon)|^p.$$

Consequently,

$$\frac{1}{(8m)^n} \sum_{x\in\mathbb{Z}_{8m}^n} \|h_x\|_p^p \leqslant \frac{4^p}{(16m)^n} \sum_{\varepsilon\in\{-1,1\}^n} \sum_{z\in\mathbb{Z}_{8m}^n} |f(z+\varepsilon) - f(z)|^p. \tag{25}$$

The desired estimate (17) now follows from a substitution of (24) and (25) into (21). $\square$

Proof of Theorem 4. Fixing $(x,\varepsilon)\in\mathbb{Z}_{8m}^n\times\{-1,1\}^n$ and $\mathsf{S}\subseteq[n]$, observe that

$$|f(x+4m\varepsilon_{\mathsf{S}})-f(x)|^p\leqslant 3^{p-1}\Big(\big|\mathsf{T}_{[n]\smallsetminus\mathsf{S}}f(x+4m\varepsilon_{\mathsf{S}})-\mathsf{T}_{[n]\smallsetminus\mathsf{S}}f(x)\big|^p+\big|f(x)-\mathsf{T}_{[n]\smallsetminus\mathsf{S}}f(x)\big|^p \\ +|f(x+4m\varepsilon_{\mathsf{S}})-\mathsf{T}_{[n]\smallsetminus\mathsf{S}}f(x+4m\varepsilon_{\mathsf{S}})|^p\Big). \qquad (26)$$

By averaging (26) over $(x,\varepsilon)\in\mathbb{Z}_{8m}^n\times\{-1,1\}^n$ and all those $\mathsf{S}\subseteq[n]$ with $|\mathsf{S}|=k$, while using translation invariance in the variable x, we see that

$$\frac{1}{(16m)^n\binom{n}{k}}\sum_{\substack{\mathsf{S}\subseteq[n]\\|\mathsf{S}|=k}}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}\frac{|f(x+4m\varepsilon_{\mathsf{S}})-f(x)|^p}{m^p}$$

$$\lesssim_p\frac{1}{(16m)^n\binom{n}{k}}\sum_{\substack{\mathsf{S}\subseteq[n]\\|\mathsf{S}|=k}}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}\frac{\big|\mathsf{T}_{[n]\smallsetminus\mathsf{S}}f(x+4m\varepsilon_{\mathsf{S}})-\mathsf{T}_{[n]\smallsetminus\mathsf{S}}f(x)\big|^p}{m^p} \qquad (27)$$

$$+\frac{1}{m^p(8m)^n\binom{n}{k}}\sum_{\substack{\mathsf{S}\subseteq[n]\\|\mathsf{S}|=k}}\sum_{x\in\mathbb{Z}_{8m}^n}|f(x)-\mathsf{T}_{[n]\smallsetminus\mathsf{S}}f(x)|^p. \qquad (28)$$

The quantity that appears in (27) can be bounded from above using Lemma 9, and the quantity that appears in (28) can be bounded from above using Lemma 8. The resulting estimate is

$$\frac{1}{(16m)^n\binom{n}{k}}\sum_{\substack{\mathsf{S}\subseteq[n]\\|\mathsf{S}|=k}}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}\frac{|f(x+4m\varepsilon_{\mathsf{S}})-f(x)|^p}{m^p}$$

$$\lesssim_p\frac{k/n}{(8m)^n}\sum_{j=1}^n\sum_{x\in\mathbb{Z}_{8m}^n}|f(x+e_j)-f(x)|^p+\frac{\left(\frac{k}{n}\right)^{\frac{p}{2}}+\frac{1}{m^p}}{(16m)^n}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{Z}_{8m}^n}|f(x+\varepsilon)-f(x)|^p.$$

This implies the desired estimate (4), since we are assuming that $m\geqslant\sqrt{n/k}$. □

3. Proof of Theorem 6

Suppose that $n\in\mathbb{N}$ and $h:\{-1,1\}^n\to\mathbb{R}$. For every $k\in\{0,\ldots,n\}$ the kth Rademacher projection of h is the function $\mathbf{Rad}_k h:\{-1,1\}^n\to\mathbb{R}$ that is given by

$$\mathbf{Rad}_k h(\varepsilon)\stackrel{\mathrm{def}}{=}\sum_{\substack{\mathsf{A}\subseteq[n]\\|A|=k}}\widehat{h}(\mathsf{A})W_{\mathsf{A}}(\varepsilon).$$

We also have the common notation $\mathbf{Rad}_1 h=\mathbf{Rad}h$. Note that $\mathbf{Rad}_0$ is the mean of h, i.e., recalling the notation (8), $\mathbf{Rad}_0 h=\mathsf{E}_{[n]}h$. By a classical theorem of Bonami [Bon68], if $\eta:\{-1,1\}^n\to\mathbb{R}$ is a Rademacher chaos of order at most k, i.e., $\widehat{\eta}(\mathsf{A})=0$ whenever $\mathsf{A}\subseteq[n]$ is such that $|\mathsf{A}|>k$, then for every $p\in[2,\infty)$ we have $\|\eta\|_p\leqslant(p-1)^{k/2}\|\eta\|_2\leqslant p^{k/2}\|\eta\|_2$. Consequently,

$$\|\mathbf{Rad}_k h\|_p\leqslant p^{\frac{k}{2}}\|\mathbf{Rad}_k h\|_2\leqslant p^{\frac{k}{2}}\|h\|_2\leqslant p^{\frac{k}{2}}\|h\|_p,$$

where we used the fact that (by Parseval's identity) $\|\mathbf{Rad}_k h\|_2\leqslant\|h\|_2$, and that $\|h\|_2\leqslant\|h\|_p$ since $p\geqslant 2$. This was a quick (and standard) derivation of the following well-known operator norm bound for $\mathbf{Rad}_k$, which we state here for ease of future reference.

$$\|\mathbf{Rad}_k\|_{p\to p}\leqslant p^{\frac{k}{2}}. \qquad (29)$$

Given $\mathsf{S} \subseteq [n]$ and $\alpha \in \mathbb{R}$, for every $h : \{-1,1\}^n \to \mathbb{R}$ define a function $\Delta_{\mathsf{S}}^{\alpha} h : \{-1,1\}^n \to \mathbb{R}$ by

$$\forall\, \varepsilon \in \{-1,1\}^n, \qquad \Delta_{\mathsf{S}}^{\alpha} h(\varepsilon) \stackrel{\mathrm{def}}{=} \sum_{\substack{\mathsf{A}\subseteq [n]\\ \mathsf{A}\cap \mathsf{S}\neq\varnothing}} |\mathsf{A}\cap\mathsf{S}|^{\alpha} \widehat{h}(\mathsf{A}) W_{\mathsf{A}}(\varepsilon).$$

Thus, recalling the notation (7) for the hypercube partial derivatives $\partial_1, \ldots, \partial_n$, as well the notation (12) for the hypercube Riesz transforms $\mathsf{R}_1, \ldots, \mathsf{R}_n$, we have the following standard identities.

$$\Delta_{\mathsf{S}} \stackrel{\mathrm{def}}{=} \Delta_{\mathsf{S}}^{1} = \frac{1}{2}\sum_{j\in S} \partial_j = \frac{1}{4}\sum_{j\in S}\partial_j^*\partial_j = \sum_{j\in S}\partial_j^2$$

(using that $\{\partial_j\}_{j=1}^n$ are self adjoint operators with $\partial_j^2 = 2\partial_j$ for all $j \in [n]$) and

$$\forall\, j \in [n], \qquad \mathsf{R}_j = \frac{1}{2}\partial_j \Delta_{[n]}^{-\frac{1}{2}}.$$

This means that Lust-Piquard's inequality (13) can we rewritten as follows.

$$\frac{1}{p^{3/2}}\left\|\Delta_{[n]}^{\frac{1}{2}} h\right\|_p \lesssim \left\|\Big(\sum_{j=1}^n (\partial_j h)^2\Big)^{\frac{1}{2}}\right\|_p \lesssim p\left\|\Delta_{[n]}^{\frac{1}{2}} h\right\|_p. \tag{30}$$

By Khinchine's inequality (with asymptotically sharp constant, see [PZ30, Lem. 2]), we have

$$\left\|\Big(\sum_{j=1}^n (\partial_j h)^2\Big)^{\frac{1}{2}}\right\|_p \leqslant \Bigg(\frac{1}{2^n}\sum_{\delta\in\{-1,1\}^n}\Big\|\sum_{j=1}^n \delta_j\partial_j h\Big\|_p^p\Bigg)^{\frac{1}{p}} \lesssim \sqrt{p}\left\|\Big(\sum_{j=1}^n (\partial_j h)^2\Big)^{\frac{1}{2}}\right\|_p.$$

In combination with (30), this implies that

$$\frac{1}{p^{\frac{3}{2}}}\left\|\Delta_{[n]}^{\frac{1}{2}} h\right\|_p \lesssim \Bigg(\frac{1}{2^n}\sum_{\delta\in\{-1,1\}^n}\Big\|\sum_{j=1}^n \delta_j\partial_j h\Big\|_p^p\Bigg)^{\frac{1}{p}} \lesssim p^{\frac{3}{2}}\left\|\Delta_{[n]}^{\frac{1}{2}} h\right\|_p. \tag{31}$$

For ease of future reference, we also record here the following formal consequence of (31).

$$\forall\, \mathsf{S}\subseteq[n], \qquad \frac{1}{p^{\frac{3}{2}}}\left\|\Delta_{\mathsf{S}}^{\frac{1}{2}} h\right\|_p \lesssim \Bigg(\frac{1}{2^n}\sum_{\delta\in\{-1,1\}^n}\Big\|\sum_{j\in\mathsf{S}} \delta_j\partial_j h\Big\|_p^p\Bigg)^{\frac{1}{p}} \lesssim p^{\frac{3}{2}}\left\|\Delta_{\mathsf{S}}^{\frac{1}{2}} h\right\|_p. \tag{32}$$

Indeed, by symmetry it suffices to establish the validity of (32) when $\mathsf{S} = \{1, \ldots, k\}$ for some $k \in [n]$, in which case (32) follows by first fixing the values of $\varepsilon_{k+1}, \ldots, \varepsilon_n \in \{-1,1\}$, applying (31) (with n replaced by k) to the mapping $(\varepsilon_1, \ldots, \varepsilon_k) \mapsto h(\varepsilon_1, \ldots, \varepsilon_k, \varepsilon_{k+1}, \ldots, \varepsilon_n)$, raising the resulting estimates to the power p, and then averaging over the remaining variables $(\varepsilon_{k+1}, \ldots, \varepsilon_n) \in \{-1,1\}^{n-k}$.

Lemma 10 below contains bounds on negative powers of the hypercube Laplacian $\Delta_{[n]}$ that will be used later, but are more general and precise than what we actually need for the proof of Theorem 6: we will only use the following operator norm estimate corresponding to the case $\alpha = 1/2$ of Lemma 10, and a worse dependence on p would have sufficed for our purposes as well.

$$\forall\, p \in [2,\infty), \qquad \sup_{n\in\mathbb{N}}\left\|\Delta_{[n]}^{-\frac{1}{2}}\right\|_{p\to p} \lesssim \sqrt{\log p}. \tag{33}$$

We include here the sharp estimates of Lemma 10 because they are interesting in their own right and our proof yields them without additional effort. The boundedness of negative powers of the hypercube Laplacian were studied in [NS02, Section 3] in the context of vector valued mappings. By specializing the bounds that are stated in [NS02] to the case of real valued mappings one obtains a variant of (33), but with a much worse dependence on p (the resulting bound grows exponentially with p). The (simple) proof below of Lemma 10 follows the strategy of [NS02] while using additional favorable properties of real valued mappings and taking care to obtain asymptotically sharp bounds.

Lemma 10. *Suppose that $p \in [2,\infty)$ and $\alpha \in (0,\infty)$ satisfy*

$$\alpha \leqslant \frac{5+\log p}{4}. \tag{34}$$

Then

$$\sup_{n\in\mathbb{N}} \left\|\Delta_{[n]}^{-\alpha}\right\|_{p\to p} \asymp \frac{(\log p)^\alpha}{2^\alpha\Gamma(1+\alpha)}. \tag{35}$$

Remark 11. Some restriction on α in the spirit of (34) is needed for (35) to hold true, since $\lim_{\alpha\to\infty}\Delta_{[n]}^{-\alpha} = \mathbf{Rad}$ and it is known that $\|\mathbf{Rad}\|_{p\to p} \gtrsim \sqrt{p}$ for n large enough (as a function of p).

Proof of Lemma 10. The lower estimate

$$\sup_{n\in\mathbb{N}} \left\|\Delta_{[n]}^{-\alpha}\right\|_{p\to p} \gtrsim \frac{(\log p)^\alpha}{2^\alpha\Gamma(1+\alpha)}. \tag{36}$$

holds true for every $\alpha \in (0,\infty)$, without the restriction (34). Indeed, denote $p^* \stackrel{\mathrm{def}}{=} p/(p-1)$ and observe that since $\Delta_{[n]}^{-\alpha}$ is self-adjoint it follows by duality that (36) is equivalent to the estimate

$$\sup_{n\in\mathbb{N}} \left\|\Delta_{[n]}^{-\alpha}\right\|_{p^*\to p^*} \gtrsim \frac{(\log p)^\alpha}{2^\alpha\Gamma(1+\alpha)}. \tag{37}$$

Fix an integer $n \geqslant 2$ and consider the following $f_p^n \in L_{p^*}(\{-1,1\}^n)$, for which $\|f_p^n\|_{p^*} = 1$.

$$\forall\, \varepsilon \in \{-1,1\}^n, \qquad f_p^n(\varepsilon) \stackrel{\mathrm{def}}{=} 2^{\frac{n}{p^*}}\delta_{(1,\ldots,1)}(\varepsilon) = 2^{-\frac{n}{p}}\prod_{j=1}^n(1+\varepsilon_j) = 2^{-\frac{n}{p}}\sum_{\mathsf{A}\subseteq[n]} W_{\mathsf{A}}.$$

For every $u \in (0,\infty)$ and $\alpha \in (0,\infty)$ the following identity holds true.

$$\frac{1}{u^\alpha} = \frac{1}{\Gamma(\alpha)}\int_0^\infty s^{\alpha-1}e^{-su}ds.$$

Consequently,

$$\Delta_{[n]}^{-\alpha} = \frac{1}{\Gamma(\alpha)}\int_0^\infty s^{\alpha-1}e^{-s\Delta_{[n]}}(I-\mathbf{Rad}_0)\mathrm{d}s. \tag{38}$$

Note that for every $s \in (0,\infty)$ and $\varepsilon \in \{-1,1\}^n$ we have

$$\begin{aligned} e^{-s\Delta_{[n]}}(I-\mathbf{Rad}_0)f_p^n(\varepsilon) &= 2^{-\frac{n}{p}}\sum_{\substack{\mathsf{A}\subseteq[n]\\ A\neq\varnothing}} e^{-s|\mathsf{A}|}W_{\mathsf{A}} \\ &= 2^{-\frac{n}{p}}\Big(\prod_{j=1}^n(1+e^{-s}\varepsilon_j)-1\Big) = -2^{-\frac{n}{p}}\left(1-(1+e^{-s})^{\kappa(\varepsilon)}(1-e^{-s})^{n-\kappa(\varepsilon)}\right), \end{aligned} \tag{39}$$

where we use the notation

$$\forall\, \varepsilon \in \{-1,1\}^n, \qquad \kappa(\varepsilon) \stackrel{\mathrm{def}}{=} \big|\{j\in[n]:\ \varepsilon_j = 1\}\big|.$$

Since the function $k \mapsto (1+e^{-s})^k(1-e^{-s})^{n-k}$ is increasing on $\{0,\ldots,n\}$, it follows from (39) that

$$\forall\, \varepsilon \in \{-1,1\}^n, \qquad \kappa(\varepsilon) \leqslant \frac{n}{2} \implies 2^{\frac{n}{p}}e^{-s\Delta_{[n]}}(I-\mathbf{Rad}_0)f_p^n(\varepsilon) \leqslant -\left(1-(1-e^{-2s})^{\frac{n}{2}}\right).$$

Recalling (38), it therefore follows that if $\varepsilon \in \{-1,1\}^n$ satisfies $\kappa(\varepsilon) \leqslant n/2$ then

$$2^{\frac{n}{p}}\left|\Delta_{[n]}^{-\alpha}f_p^n(\varepsilon)\right| \geqslant \frac{1}{\Gamma(\alpha)}\int_0^\infty s^{\alpha-1}\left(1-(1-e^{-2s})^{\frac{n}{2}}\right)\mathrm{d}s \geqslant \frac{1}{\Gamma(\alpha)}\int_0^{\frac{\log n}{2}}\frac{s^{\alpha-1}}{4}\mathrm{d}s = \frac{(\log n)^\alpha}{2^{2+\alpha}\Gamma(1+\alpha)}.$$

Hence, since $\left|\left\{\varepsilon \in \{-1,1\}^n :\ \kappa(\varepsilon) \leqslant \frac{n}{2}\right\}\right| \geqslant 2^{n-1}$, we have

$$\left\|\Delta_{[n]}^{-\alpha} f_p^n\right\|_{p^*} \gtrsim \frac{2^{-\frac{n}{p}}(\log n)^\alpha}{2^\alpha \Gamma(1+\alpha)}. \tag{40}$$

The desired estimate (37) now follows by choosing $n = \lceil p \rceil$ in (40).

Having proven (36), it remains to show that under the assumption (34) we have

$$\sup_{n\in\mathbb{N}} \left\|\Delta_{[n]}^{-\alpha}\right\|_{p\to p} \lesssim \frac{(\log p)^\alpha}{2^\alpha \Gamma(1+\alpha)}. \tag{41}$$

To this end, observe first that the identity

$$e^{-s\Delta_{[n]}}(I - \mathbf{Rad}_0) = \sum_{k=1}^n \frac{1}{e^{sk}} \mathbf{Rad}_k$$

implies that

$$\left\|e^{-s\Delta_{[n]}}(I - \mathbf{Rad}_0)\right\|_{p\to p} \leqslant \sum_{k=1}^n \frac{1}{e^{sk}} \left\|\mathbf{Rad}_k\right\|_{p\to p} \stackrel{(29)}{\leqslant} \sum_{k=1}^n \left(\frac{\sqrt{p}}{e^s}\right)^k.$$

Hence,

$$e^s > \sqrt{p} \implies \left\|e^{-s\Delta_{[n]}}(I - \mathbf{Rad}_0)\right\|_{p\to p} \lesssim \frac{\sqrt{p}}{e^s - \sqrt{p}}. \tag{42}$$

Suppose that $M \in (0,\infty)$ satisfies

$$e^M \geqslant \sqrt{ep} \iff M \geqslant \frac{1+\log p}{2}. \tag{43}$$

Then, by (42) we have

$$\frac{1}{\Gamma(\alpha)} \int_M^\infty s^{\alpha-1} \left\|e^{-s\Delta_{[n]}}(I - \mathbf{Rad}_0)\right\|_{p\to p} \mathrm{d}s \lesssim \frac{\sqrt{p}}{\Gamma(\alpha)} \int_M^\infty \frac{s^{\alpha-1}}{e^s} \mathrm{d}s. \tag{44}$$

Due to (34) and (43) we have $M \geqslant 2(\alpha-1)$. Since the function $s \mapsto s^{\alpha-1}e^{-s/2}$ is decreasing on $[2(\alpha-1),\infty) \supseteq [M,\infty)$, it follows that for every $s \geqslant M$ we have $s^{\alpha-1}e^{-s} \leqslant M^{\alpha-1}e^{-M/2}e^{-s/2}$. So,

$$\int_M^\infty \frac{s^{\alpha-1}}{e^s} \mathrm{d}s \leqslant \frac{M^{\alpha-1}}{e^{M/2}} \int_M^\infty \frac{\mathrm{d}s}{e^{s/2}} \lesssim \frac{M^{\alpha-1}}{e^M}. \tag{45}$$

A substitution of (45) into (44) yields the estimate

$$\frac{1}{\Gamma(\alpha)} \int_M^\infty s^{\alpha-1} \left\|e^{-s\Delta_{[n]}}(I - \mathbf{Rad}_0)\right\|_{p\to p} \mathrm{d}s \lesssim \frac{M^{\alpha-1}\sqrt{p}}{e^M \Gamma(\alpha)}. \tag{46}$$

At the same time, since for every $s \in [0,\infty)$ we have $\|e^{-s\Delta_{[n]}}\|_{p\to p} \leqslant 1$ (because $e^{-s\Delta_{[n]}}$ is an averaging operator) and $\|I - \mathbf{Rad}_0\|_{p\to p} \leqslant 2$, we have

$$\frac{1}{\Gamma(\alpha)} \int_0^M s^{\alpha-1} \left\|e^{-s\Delta_{[n]}}(I - \mathbf{Rad}_0)\right\|_{p\to p} \mathrm{d}s \lesssim \frac{1}{\Gamma(\alpha)} \int_0^M s^{\alpha-1} \mathrm{d}s = \frac{M^\alpha}{\Gamma(1+\alpha)}. \tag{47}$$

Making the choice

$$M \stackrel{\mathrm{def}}{=} \frac{1+\log p}{2}, \tag{48}$$

we see that

$$\left\|\Delta_{[n]}^{-\alpha}\right\|_{p\to p} \overset{(38)}{\leqslant} \frac{1}{\Gamma(\alpha)}\int_0^\infty s^{\alpha-1}\left\|e^{-s\Delta_{[n]}}(I-\mathbf{Rad}_0)\right\|_{p\to p}\mathrm{d}s \overset{(46)\wedge(47)}{\lesssim} \frac{M^\alpha}{\Gamma(1+\alpha)}+\frac{M^{\alpha-1}\sqrt{p}}{e^M\Gamma(\alpha)}$$
$$\overset{(48)}{=} \frac{(1+\log p)^\alpha}{2^\alpha\Gamma(1+\alpha)}+\frac{(1+\log p)^{\alpha-1}}{2^{\alpha-1}\Gamma(\alpha)\sqrt{e}} \asymp \frac{(\log p)^\alpha}{2^\alpha\Gamma(\alpha)}\left(\frac{1}{\alpha}+\frac{1}{\log p}\right)\left(1+\frac{1}{\log p}\right)^\alpha \overset{(34)}{\asymp} \frac{(\log p)^\alpha}{2^\alpha\Gamma(1+\alpha)}.$$

This is precisely the desired estimate (41), thus completing the proof of Lemma 10. □

Lemma 12. *Fix $n\in\mathbb{N}$, $p\in[1,\infty)$ and $\alpha\in\mathbb{R}$. Then for $h\in L_p^0(\{-1,1\}^n)$ and $S\subseteq[n]$ we have,*

$$\left\|\mathsf{E}_{[n]\smallsetminus \mathsf{S}}h\right\|_p \leqslant \left\|\Delta_{\mathsf{S}}^{\alpha}\Delta_{[n]}^{-\alpha}h\right\|_p.$$

Proof. Observe that we have the following identity of operators on $L_p^0(\{-1,1\}^n)$.

$$\mathsf{E}_{[n]\smallsetminus \mathsf{S}}\Delta_{\mathsf{S}}^{\alpha}\Delta_{[n]}^{-\alpha} = \mathsf{E}_{[n]\smallsetminus \mathsf{S}}. \tag{49}$$

Indeed, if $\varnothing\neq\mathsf{A}\subseteq[n]$ then $\mathsf{E}_{[n]\smallsetminus \mathsf{S}}W_A=\mathbf{1}_{\{\mathsf{A}\subseteq\mathsf{S}\}}W_A$, and at the same time we have

$$\mathsf{E}_{[n]\smallsetminus \mathsf{S}}\Delta_{\mathsf{S}}^{\alpha}\Delta_{[n]}^{-\alpha}W_{\mathsf{A}} = \frac{|\mathsf{A}\cap\mathsf{S}|^\alpha}{|\mathsf{A}|^\alpha}\mathsf{E}_{[n]\smallsetminus \mathsf{S}}W_A = \frac{|\mathsf{A}\cap\mathsf{S}|^\alpha}{|\mathsf{A}|^\alpha}\mathbf{1}_{\{\mathsf{A}\subseteq\mathsf{S}\}}W_A = \mathbf{1}_{\{\mathsf{A}\subseteq\mathsf{S}\}}W_A.$$

Consequently,

$$\left\|\Delta_{\mathsf{S}}^{\alpha}\Delta_{[n]}^{-\alpha}h\right\|_p^p = \mathsf{E}_{\mathsf{S}}\mathsf{E}_{[n]\smallsetminus \mathsf{S}}\left|\Delta_{\mathsf{S}}^{\alpha}\Delta_{[n]}^{-\alpha}h\right|^p \geqslant \mathsf{E}_{\mathsf{S}}\left|\mathsf{E}_{[n]\smallsetminus \mathsf{S}}\Delta_{\mathsf{S}}^{\alpha}\Delta_{[n]}^{-\alpha}h\right|^p \overset{(49)}{=} \mathsf{E}_{\mathsf{S}}\left|\mathsf{E}_{[n]\smallsetminus \mathsf{S}}h\right|^p = \left\|\mathsf{E}_{[n]\smallsetminus \mathsf{S}}h\right\|_p^p, \tag{50}$$

where the inequality in (50) follows from Jensen's inequality ($\mathsf{E}_{[n]\smallsetminus \mathsf{S}}$ is an averaging operator). □

Proof of Theorem 6. By Lemma 12 we have

$$\left(\frac{1}{\binom{n}{k}}\sum_{\substack{\mathsf{S}\subseteq[n]\\|\mathsf{S}|=k}}\left\|\mathsf{E}_{[n]\smallsetminus \mathsf{S}}h\right\|_p^p\right)^{\frac1p} \leqslant \left(\frac{1}{\binom{n}{k}}\sum_{\substack{\mathsf{S}\subseteq[n]\\|\mathsf{S}|=k}}\left\|\Delta_{\mathsf{S}}^{\frac12}\Delta_{[n]}^{-\frac12}h\right\|_p^p\right)^{\frac1p}. \tag{51}$$

By Lust-Piquard's discrete Riesz transform inequality (32), for every fixed $\mathsf{S}\subseteq[n]$ we have

$$\left\|\Delta_{\mathsf{S}}^{\frac12}\Delta_{[n]}^{-\frac12}h\right\|_p \lesssim p^{\frac32}\left(\frac{1}{2^n}\sum_{\delta\in\{-1,1\}^n}\left\|\sum_{j\in\mathsf{S}}\delta_j\partial_j\Delta_{[n]}^{-\frac12}h\right\|_p^p\right)^{\frac1p}. \tag{52}$$

A substitution of (52) into (51) yields

$$\left(\frac{1}{\binom{n}{k}}\sum_{\substack{\mathsf{S}\subseteq[n]\\|\mathsf{S}|=k}}\left\|\mathsf{E}_{[n]\smallsetminus \mathsf{S}}h\right\|_p^p\right)^{\frac1p} \lesssim p^{\frac32}\left(\frac{1}{2^n\binom{n}{k}}\sum_{\substack{\mathsf{S}\subseteq[n]\\|\mathsf{S}|=k}}\sum_{\delta\in\{-1,1\}^n}\left\|\sum_{j\in\mathsf{S}}\delta_j\partial_j\Delta_{[n]}^{-\frac12}h\right\|_p^p\right)^{\frac1p}. \tag{53}$$

For fixed $\varepsilon\in\{-1,1\}^n$, the linear X_p inequality (9) with $\left\{a_j=\partial_j\Delta_{[n]}^{-\frac12}h(\varepsilon)\right\}_{j=1}^n$ yields the estimate

$$\left(\frac{1}{2^n\binom{n}{k}}\sum_{\substack{\mathsf{S}\subseteq[n]\\|\mathsf{S}|=k}}\sum_{\delta\in\{-1,1\}^n}\left|\sum_{j\in\mathsf{S}}\delta_j\partial_j\Delta_{[n]}^{-\frac12}h(\varepsilon)\right|^p\right)^{\frac1p}$$
$$\lesssim \frac{p}{\log p}\left(\frac{k}{n}\sum_{j=1}^n\left|\partial_j\Delta_{[n]}^{-\frac12}h(\varepsilon)\right|^p+\frac{(k/n)^{\frac p2}}{2^n}\sum_{\delta\in\{-1,1\}^n}\left|\sum_{j=1}^n\delta_j\partial_j\Delta_{[n]}^{-\frac12}h(\varepsilon)\right|^p\right)^{\frac1p}. \tag{54}$$

By taking L_p norms with respect to $\varepsilon \in \{-1,1\}^n$, it follows from (54) that

$$\Bigg(\frac{1}{2^n\binom{n}{k}} \sum_{\substack{\mathsf{S}\subseteq[n]\\ |\mathsf{S}|=k}} \sum_{\delta\in\{-1,1\}^n} \Big\|\sum_{j\in\mathsf{S}} \delta_j \partial_j \Delta_{[n]}^{-\frac12} h\Big\|_p^p\Bigg)^{\frac1p}$$

$$\lesssim \frac{p}{\log p}\Bigg(\frac{k}{n}\sum_{j=1}^n \Big\|\partial_j \Delta_{[n]}^{-\frac12} h\Big\|_p^p + \frac{(k/n)^{\frac p2}}{2^n} \sum_{\delta\in\{-1,1\}^n} \Big\|\sum_{j=1}^n \delta_j\partial_j\Delta_{[n]}^{-\frac12} h\Big\|_p^p\Bigg)^{\frac1p}. \qquad (55)$$

By Lemma 10 we have

$$\Bigg(\sum_{j=1}^n \Big\|\partial_j \Delta_{[n]}^{-\frac12} h\Big\|_p^p\Bigg)^{\frac1p} \lesssim \sqrt{\log p}\Bigg(\sum_{j=1}^n \|\partial_j h\|_p^p\Bigg)^{\frac1p}, \qquad (56)$$

and another application of Lust-Piquard's discrete Riesz transform inequality (31) shows that

$$\Bigg(\frac{1}{2^n} \sum_{\delta\in\{-1,1\}^n} \Big\|\sum_{j=1}^n \delta_j\partial_j\Delta_{[n]}^{-\frac12} h\Big\|_p^p\Bigg)^{\frac1p} \lesssim p^{\frac32}\|h\|_p. \qquad (57)$$

The desired estimate (10) (in its slightly more refined form (11)) now follows by substituting (56) and (57) into (55), and then substituting the resulting inequality into (53). □

4. Beyond L_p

Fix $p \in [2,\infty)$. Following [NS14], a Banach space $(X, \|\cdot\|_X)$ is said to be an X_p Banach space if for every $k, n \in \mathbb{N}$ with $k \in [n]$, every $\mathsf{v}_1, \ldots, \mathsf{v}_n \in X$ satisfy

$$\frac{1}{2^n\binom{n}{k}} \sum_{\substack{\mathsf{S}\subseteq[n]\\ |\mathsf{S}|=k}} \sum_{\varepsilon\in\{-1,1\}^n} \Big\|\sum_{j\in\mathsf{S}} \varepsilon_j \mathsf{v}_j\Big\|_X^p \lesssim_X \frac{k}{n}\sum_{j=1}^n \|\mathsf{v}_j\|_X^p + \frac{(k/n)^{\frac p2}}{2^n} \sum_{\varepsilon\in\{-1,1\}^n} \Big\|\sum_{j=1}^n \varepsilon_j \mathsf{v}_j\Big\|_X^p.$$

Inequality (9) implies that L_p is an X_p Banach space when $p \in [2,\infty)$, and in [NS14] it was proven that for p in this range also the Schatten–von Neumann trace class S_p is an X_p Banach space. Thus, due to [McC67], there exists an X_p Banach space that is not isomorphic to a subspace of L_p.

By [NS14] we know that any X_p Banach space is also an X_p metric space (see [Nao12, Bal13] for the significance of such results in the context of the Ribe program). Our proof of Theorem 4 does not imply this general statement, since it relies on additional properties of the target Banach space X, which in our case is L_p. An inspection of our proof reveals that it uses only two nontrivial properties of the target space Banach X. Firstly, we need the following operator norm bounds.

$$\sup_{n\in\mathbb{N}} \Big\|\Delta_{[n]}^{-\frac12}\otimes I_X\Big\|_{L_p(\{-1,1\}^n,X)\to L_p(\{-1,1\}^n,X)} < \infty. \qquad (58)$$

By [NS02, Theorem 5], the requirement (58) is equivalent to X being a K-convex Banach space (for background on K-convexity, see the survey [Mau03]). Secondly, we need X to satisfy the following vector valued version of Lust-Piquard's inequality (31) for every $n \in \mathbb{N}$ and $h : \{-1,1\}^n \to X$.

$$\frac{1}{2^n} \sum_{\delta\in\{-1,1\}^n} \Big\|\sum_{j=1}^n \delta_j\big(\partial_j\otimes I_X\big)h\Big\|_{L_p(\{-1,1\}^n,X)}^p \asymp_X \Big\|\Big(\Delta_{[n]}^{\frac12}\otimes I_X\Big)h\Big\|_{L_p(\{-1,1\}^n,X)}^p. \qquad (59)$$

So, the argument of the present article actually shows that any K-convex X_p Banach space X that satisfies (59) is also an X_p metric space, with the same scaling parameter as in the statement of Theorem 4. However, as we shall explain in Corollary 15 below, the validity of (59) already implies that X is K-convex. This means that (58) is a consequence of (59) and there is no need to stipulate the validity of (58) as a separate assumption. We therefore have the following theorem.

Theorem 13. *Suppose that $p\in[2,\infty)$ and that $(X,\|\cdot\|_X)$ is an X_p Banach space that satisfies* (59)*. Suppose also that $k,m,n\in\mathbb{N}$ satisfy $k\in[n]$ and $m\geqslant\sqrt{n/k}$. Then every $f:\mathbb{Z}_{8m}^n\to X$ satisfies*

$$\bigg(\frac{1}{\binom{n}{k}}\sum_{\substack{\mathsf{S}\subseteq[n]\\ |\mathsf{S}|=k}}\mathbb{E}\Big[\,\|f(x+4m\varepsilon_{\mathsf{S}})-f(x)\|_X^p\,\Big]\bigg)^{\frac1p}$$

$$\lesssim_X m\bigg(\frac{k}{n}\sum_{j=1}^n\mathbb{E}\Big[\|f(x+e_j)-f(x)\|_X^p\Big]+\Big(\frac{k}{n}\Big)^{\frac p2}\mathbb{E}\Big[\|f(x+\varepsilon)-f(x)\|_X^p\Big]\bigg)^{\frac1p}, \quad (60)$$

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

It seems to be quite challenging to obtain a clean and useful characterization of the class of Banach spaces that satisfy the dimension-independent vector valued discrete Riesz transform inequality (59). We did verify, in collaboration with A. Eskenazis, that the Schatten–von Neumann trace class S_p satisfies (59) when $p\in[2,\infty)$, but in order to see this one needs to reexamine Lust-Piquard's proof in [LP98] while checking in several instances that her argument could be adjusted so as to apply to S_p-valued functions as well. Since including such an argument here would be quite lengthy (and mostly a repetition of Lust-Piquard's work), we postpone the justification of (59) when $X=\mathsf{S}_p$ to forthcoming work that is devoted to vector valued Riesz transforms. Due to the fact that S_p was shown to be an X_p Banach space in [NS14], Theorem 13 holds true when $X=\mathsf{S}_p$, with our current proof showing that the implicit constant in (60) (with $X=\mathsf{S}_p$) is $O(p^4/\sqrt{\log p})$.

It remains to prove that if a Banach space X satisfies (59) then X is K-convex. In fact, the following stronger statement holds true (see [Mau03] for background on type of Banach spaces).

Proposition 14. *Suppose that $p\in[1,\infty)$ and $\alpha\in(0,1)$, and that $(X,\|\cdot\|_X)$ is a Banach space such that for every $n\in\mathbb{N}$ and every $h:\{-1,1\}^n\to X$ we have*

$$\frac{1}{2^n}\sum_{\delta\in\{-1,1\}^n}\Big\|\sum_{j=1}^n\delta_j(\partial_j\otimes I_X)h\Big\|_{L_p(\{-1,1\}^n,X)}^p\lesssim_X\Big\|\big(\Delta_{[n]}^\alpha\otimes I_X\big)h\Big\|_{L_p(\{-1,1\}^n,X)}^p. \quad (61)$$

Then X has type $\frac1\alpha-\tau$ for every $\tau\in(0,1]$. In particular, if (59) *holds true then X has type $2-\tau$ for every $\tau\in(0,1]$.*

By Pisier's K-convexity theorem [Pis82], a Banach space X has type strictly larger than 1 if and only if X is K-convex. We therefore have the following corollary of Proposition 14.

Corollary 15. *If $p\in[1,\infty)$ and $(X,\|\cdot\|_X)$ is a Banach space that satisfies* (61) *then X is K-convex.*

Proof of Proposition 14. Let $r_X\in[1,2]$ be the supremum over those $r\in[1,2]$ such that X has type r. Our goal is to show that $r_X\geqslant 1/\alpha$. By the Maurey–Pisier theorem [MP76], for every $n\in\mathbb{N}$ there exists a linear operator $\mathsf{J}_n:L_{r_X}(\{-1,1\}^n)\to X$ such that

$$\forall\, g\in L_{r_X}(\{-1,1\}^n),\qquad \|g\|_{r_X}\leqslant\|\mathsf{J}_ng\|_X\leqslant 2\|g\|_{r_X}. \quad (62)$$

Fixing $n\in\mathbb{N}$ and $g\in L_{r_X}(\{-1,1\}^n)$, for every $\omega\in\{-1,1\}^n$ define $g_\omega\in L_{r_X}(\{-1,1\}^n)$ by

$$\forall\,\varepsilon\in\{-1,1\}^n,\qquad g_\omega(\varepsilon)\stackrel{\mathrm{def}}{=}g(\omega\varepsilon)=g(\omega_1\varepsilon_1,\ldots,\omega_n\varepsilon_n). \quad (63)$$

Next, define $h_g:\{-1,1\}^n\to X$ by setting

$$\forall\,\omega\in\{-1,1\}^n,\qquad h_g(\omega)\stackrel{\mathrm{def}}{=}\mathsf{J}_ng_\omega\in X. \quad (64)$$

It follows from (63) and (64) that

$$\forall\, \omega \in \{-1,1\}^n, \qquad h_g(\omega) = \mathsf{J}_n\bigg(\sum_{\mathsf{A}\subseteq[n]} \widehat{g}(\mathsf{A})W_{\mathsf{A}}(\omega)W_{\mathsf{A}}\bigg) = \sum_{\mathsf{A}\subseteq[n]} \widehat{g}(\mathsf{A})W_{\mathsf{A}}(\omega)\mathsf{J}_n(W_{\mathsf{A}}). \tag{65}$$

By (65), for every $\omega \in \{-1,1\}^n$ we have

$$\big(\Delta^{\alpha}_{[n]}\otimes I_X\big)h_g(\omega) = \sum_{\mathsf{A}\subseteq[n]} |\mathsf{A}|^{\alpha}\widehat{g}(\mathsf{A})W_{\mathsf{A}}(\omega)\mathsf{J}_n(W_{\mathsf{A}}) = \mathsf{J}_n\big(\Delta^{\alpha}_{[n]}g_{\omega}\big) = \mathsf{J}_n\Big(\big(\Delta^{\alpha}_{[n]}g\big)_{\omega}\Big). \tag{66}$$

Consequently,

$$\forall\, \omega \in \{-1,1\}^n, \qquad \Big\|\big(\Delta^{\alpha}_{[n]}\otimes I_X\big)h_g(\omega)\Big\|_X \overset{(66)\wedge(62)}{\leqslant} 2\Big\|\big(\Delta^{\alpha}_{[n]}g\big)_{\omega}\Big\|_{r_X} = 2\Big\|\Delta^{\alpha}_{[n]}g\Big\|_{r_X}. \tag{67}$$

In a similar vein, it follows from (65) that for every $\omega,\delta \in \{-1,1\}^n$ we have

$$\sum_{j=1}^n \delta_j\big(\partial_j\otimes I_X\big)h_g(\omega) = \mathsf{J}_n\bigg(\Big(\sum_{j=1}^n \delta_j\partial_j g\Big)_{\omega}\bigg). \tag{68}$$

Hence,

$$\forall\, \omega,\delta \in \{-1,1\}^n, \qquad \Big\|\sum_{j=1}^n \delta_j\big(\partial_j\otimes I_X\big)h_g(\omega)\Big\|_X \overset{(68)\wedge(62)}{\geqslant} \Big\|\sum_{j=1}^n \delta_j\partial_j g\Big\|_{r_X}. \tag{69}$$

By combining (67) and (69) with an application of (61) to $h = h_g$, it follows that

$$\bigg(\frac{1}{2^n}\sum_{\delta\in\{-1,1\}^n}\Big\|\sum_{j=1}^n \delta_j\partial_j g\Big\|_{r_X}^{r_X}\bigg)^{\frac{1}{r_X}} \lesssim_p \bigg(\frac{1}{2^n}\sum_{\delta\in\{-1,1\}^n}\Big\|\sum_{j=1}^n \delta_j\partial_j g\Big\|_{r_X}^{p}\bigg)^{\frac{1}{p}} \lesssim_{p,X} \Big\|\Delta^{\alpha}_{[n]}g\Big\|_{r_X}, \tag{70}$$

where the first step of (70) uses Kahane's inequality [Kah64]. When $\alpha = 1/2$, by a result of Lamberton [LP98, p. 283], and for general $\alpha \in (0,1)$ by a result of the author and Schechtman [BELP08, Section 5.5], the validity of (70) for every $n \in \mathbb{N}$ and $g \in L_{r_X}(\{-1,1\}^n)$ implies that $r_X \geqslant 1/\alpha$. $\square$

Acknowledgements. I am very grateful to Alexandros Eskenazis for many helpful discussions.

References

[Bal13] K. Ball. The Ribe programme. *Astérisque*, (352):Exp. No. 1047, viii, 147–159, 2013. Séminaire Bourbaki. Vol. 2011/2012. Exposés 1043–1058.

[Bañ86] R. Bañuelos. Martingale transforms and related singular integrals. *Trans. Amer. Math. Soc.*, 293(2):547–563, 1986.

[BELP08] L. Ben-Efraim and F. Lust-Piquard. Poincaré type inequalities on the discrete cube and in the CAR algebra. *Probab. Theory Related Fields*, 141(3-4):569–602, 2008.

[BL00] Y. Benyamini and J. Lindenstrauss. *Geometric nonlinear functional analysis. Vol. 1*, volume 48 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2000.

[Bon68] A. Bonami. Construction d'opérateurs de convolution sur le groupe D^{∞}. *C. R. Acad. Sci. Paris Sér. A-B*, 266:A864–A866, 1968.

[Bou86] J. Bourgain. Vector-valued singular integrals and the H^1-BMO duality. In *Probability theory and harmonic analysis (Cleveland, Ohio, 1983)*, volume 98 of *Monogr. Textbooks Pure Appl. Math.*, pages 1–19. Dekker, New York, 1986.

[Bou87] J. Bourgain. Remarks on the extension of Lipschitz maps defined on discrete sets and uniform homeomorphisms. In *Geometrical aspects of functional analysis (1985/86)*, volume 1267 of *Lecture Notes in Math.*, pages 157–167. Springer, Berlin, 1987.

[CW76] R. R. Coifman and G. Weiss. *Transference methods in analysis*. American Mathematical Society, Providence, R.I., 1976. Conference Board of the Mathematical Sciences Regional Conference Series in Mathematics, No. 31.

[DRdF85] J. Duoandikoetxea and J. L. Rubio de Francia. Estimations indépendantes de la dimension pour les transformées de Riesz. *C. R. Acad. Sci. Paris Sér. I Math.*, 300(7):193–196, 1985.

[FJS88] T. Figiel, W. B. Johnson, and G. Schechtman. Random sign embeddings from l_r^n, $2 < r < \infty$. *Proc. Amer. Math. Soc.*, 102(1):102–106, 1988.

[GNS12] O. Giladi, A. Naor, and G. Schechtman. Bourgain's discretization theorem. *Ann. Fac. Sci. Toulouse Math. (6)*, 21(4):817–837, 2012.

[GPP80] 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):149–155, 1980.

[Gun86] R. F. Gundy. Sur les transformations de Riesz pour le semi-groupe d'Ornstein-Uhlenbeck. *C. R. Acad. Sci. Paris Sér. I Math.*, 303(19):967–970, 1986.

[GV79] R. F. Gundy and N. T. Varopoulos. Les transformations de Riesz et les intégrales stochastiques. *C. R. Acad. Sci. Paris Sér. A-B*, 289(1):A13–A16, 1979.

[Hei80] S. Heinrich. Ultraproducts in Banach space theory. *J. Reine Angew. Math.*, 313:72–104, 1980.

[JMST79] W. B. Johnson, B. Maurey, G. Schechtman, and L. Tzafriri. Symmetric structures in Banach spaces. *Mem. Amer. Math. Soc.*, 19(217):v+298, 1979.

[JSZ85] 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):234–253, 1985.

[Kah64] J.-P. Kahane. Sur les sommes vectorielles $\sum \pm u_n$. *C. R. Acad. Sci. Paris*, 259:2577–2580, 1964.

[LP86] F. Lust-Piquard. Inégalités de Khintchine dans C_p $(1 < p < \infty)$. *C. R. Acad. Sci. Paris Sér. I Math.*, 303(7):289–292, 1986.

[LP98] F. Lust-Piquard. Riesz transforms associated with the number operator on the Walsh system and the fermions. *J. Funct. Anal.*, 155(1):263–285, 1998.

[LT77] J. Lindenstrauss and L. Tzafriri. *Classical Banach spaces. I.* Springer-Verlag, Berlin-New York, 1977. Sequence spaces, Ergebnisse der Mathematik und ihrer Grenzgebiete, Vol. 92.

[Mau03] B. Maurey. Type, cotype and K-convexity. In *Handbook of the geometry of Banach spaces, Vol. 2*, pages 1299–1332. North-Holland, Amsterdam, 2003.

[McC67] C. A. McCarthy. c_p. *Israel J. Math.*, 5:249–271, 1967.

[Mey84] P.-A. Meyer. Transformations de Riesz pour les lois gaussiennes. In *Seminar on probability, XVIII*, volume 1059 of *Lecture Notes in Math.*, pages 179–193. Springer, Berlin, 1984.

[MN04] M. Mendel and A. Naor. Euclidean quotients of finite metric spaces. *Adv. Math.*, 189(2):451–494, 2004.

[MN06] M. Mendel and A. Naor. Some applications of Ball's extension theorem. *Proc. Amer. Math. Soc.*, 134(9):2577–2584 (electronic), 2006.

[MN08] M. Mendel and A. Naor. Metric cotype. *Ann. of Math. (2)*, 168(1):247–298, 2008.

[MP76] B. Maurey and G. Pisier. Séries de variables aléatoires vectorielles indépendantes et propriétés géométriques des espaces de Banach. *Studia Math.*, 58(1):45–90, 1976.

[Nao12] A. Naor. An introduction to the Ribe program. *Jpn. J. Math.*, 7(2):167–233, 2012.

[NS02] A. Naor and G. Schechtman. Remarks on non linear type and Pisier's inequality. *J. Reine Angew. Math.*, 552:213–236, 2002.

[NS14] A. Naor and G. Schechtman. Metric X_p inequalities, 2014. Preprint available at `http://arxiv.org/abs/1408.5819`. To appear in Forum Math. Pi.

[Ost13] M. I. Ostrovskii. *Metric embeddings*, volume 49 of *De Gruyter Studies in Mathematics*. De Gruyter, Berlin, 2013. Bilipschitz and coarse embeddings into Banach spaces.

[Pal36] R. E. A. C. Paley. Some theorems on abstract spaces. *Bull. Amer. Math. Soc.*, 42(4):235–240, 1936.

[Pis82] G. Pisier. Holomorphic semigroups and the geometry of Banach spaces. *Ann. of Math. (2)*, 115(2):375–392, 1982.

[Pis88] G. Pisier. Riesz transforms: a simpler analytic proof of P.-A. Meyer's inequality. In *Séminaire de Probabilités, XXII*, volume 1321 of *Lecture Notes in Math.*, pages 485–501. Springer, Berlin, 1988.

[PZ30] R. E. A. C. Paley and A. Zygmund. On some series of functions, (1). *Math. Proc. Cambridge Philos. Soc.*, 26(3):337–357, 1930.

[Ros70] H. P. Rosenthal. On the subspaces of L^p $(p > 2)$ spanned by sequences of independent random variables. *Israel J. Math.*, 8:273–303, 1970.

[Sch38] I. J. Schoenberg. Metric spaces and positive definite functions. *Trans. Amer. Math. Soc.*, 44(3):522–536, 1938.

[Ste83] E. M. Stein. Some results in harmonic analysis in $\mathbf{R}^n$, for $n \to \infty$. *Bull. Amer. Math. Soc. (N.S.)*, 9(1):71–73, 1983.

Mathematics Department, Princeton University, Fine Hall, Washington Road, Princeton, NJ 08544-1000, USA

E-mail address: `naor@math.princeton.edu`