

MOMENTS OF THE DISTANCE BETWEEN INDEPENDENT RANDOM VECTORS

ASSAF NAOR AND KRZYSZTOF OLESZKIEWICZ

ABSTRACT. We derive various sharp bounds on moments of the distance between two independent random vectors taking values in a Banach space.

1. INTRODUCTION

Throughout what follows, all Banach spaces are tacitly assumed to be separable. This assumption removes the need to discuss measurability side-issues; alternatively one could consider throughout only the special case of finitely-supported random variables, which captures all of the key ideas. We will also tacitly assume that all Banach spaces are over the complex scalars \mathbb{C} . This assumption is convenient for the ensuing proofs, but the main statements (namely, those that do not mention complex scalars explicitly) hold over the real scalars as well, through a standard complexification procedure. All the notation and terminology from Banach space theory that occurs below is basic and standard, as in e.g. [15].

Our starting point is the following question. What is the smallest $C > 0$ such that for every Banach space $(F, \|\cdot\|_F)$ and every two independent F -valued integrable random vectors $X, Y \in L_1(F)$ we have

$$\inf_{z \in F} \mathbb{E} [\|X - z\|_F + \|Y - z\|_F] \leq C \mathbb{E} [\|X - Y\|_F]? \tag{1}$$

We will reason that (1) holds with $C = 3$, and that $C = 3$ is the sharp constant here. More generally, we have the following theorem.

Theorem 1.1. *Suppose that $p \geq 1$ and $(F, \|\cdot\|_F)$ is a Banach space. Let $X, Y \in L_p(F)$ be two independent F -valued p -integrable random vectors. Then*

$$\inf_{z \in F} \mathbb{E} [\|X - z\|_F^p + \|Y - z\|_F^p] \leq \frac{3^p}{2^{p-1}} \mathbb{E} [\|X - Y\|_F^p]. \tag{2}$$

The constant $\frac{3^p}{2^{p-1}}$ in (2) cannot be improved.

The Banach space F that exhibits this sharpness of (2) is, of course, a subspace of ℓ_∞ , but we do not know what is the optimal constant in (2) when $F = \ell_\infty$ itself. More generally, understanding the meaning of the optimal constant in (2) for specific Banach spaces is an interesting question, which we investigate in the rest of the present work for certain special classes of Banach spaces but do not fully resolve.

1.1. Geometric motivation. Our interest in (1) arose from investigations of [1] in the context of Riemannian/Alexandrov geometry. It is well established throughout an extensive geometric literature that a range of useful quadratic distance inequalities for a metric space $(\mathcal{M}, d_{\mathcal{M}})$ arise if one imposes bounds on its curvature in the sense of Alexandrov. The term “quadratic” here indicates that these inequalities involve squares of distances between finite point configurations in \mathcal{M} . A phenomenon that was established in [1] is that any such quadratic metric inequality that holds for every Alexandrov space of nonnegative curvature becomes valid in any metric space whatsoever if one removes the squaring of the distances, i.e., in essence upon “linearization” of the inequality; see [1] for a precise formulation. This led naturally to the question whether the same phenomenon holds for Hadamard spaces (complete simply

A.N. was supported by the Packard Foundation and the Simons Foundation. The research that is presented here was conducted under the auspices of the Simons Algorithms and Geometry (A&G) Think Tank. K.O. was partially supported by the National Science Centre, Poland, project number 2012/05/B/ST1/00412.

connected spaces whose Alexandrov curvature in nonpositive); see [1] for an extensive discussion as well as the recent negative resolution of this question in [11]. In the context of a Hadamard space $(\mathcal{M}, d_{\mathcal{M}})$, the analogue of (1) is that independent finitely-supported \mathcal{M} -valued random variables X, Y satisfy

$$\inf_{z \in \mathcal{M}} \mathbb{E} [d_{\mathcal{M}}(X, z)^2 + d_{\mathcal{M}}(Y, z)^2] \leq \mathbb{E} [d_{\mathcal{M}}(X, Y)^2]. \quad (3)$$

See [1] for a standard derivation of (3), where $z \in \mathcal{M}$ is an appropriate “geometric barycenter,” namely it is obtained as the minimizer of the expected squared distance from X to z . As explained in [1], by using (3) iteratively one can obtain quadratic metric inequalities that hold in any Hadamard space and serve as obstructions for certain geometric embeddings. The “linearized” version of (3), in the case of Banach spaces and allowing for a loss of a factor C , is precisely (1). So, in the spirit of [1] it is natural to ask what is the smallest C for which it holds. This is what we address here, leading to analytic questions about Banach spaces that are interesting in their own right from the probabilistic and geometric perspective. We note that there are questions along these lines that [1] raises and remain open; see e.g. [1, Question 32].

1.2. Probabilistic discussion. The inequality which reverses (1) holds trivially as a consequence of the triangle inequality, even when X and Y are not necessarily independent. Namely, any $X, Y \in L_1(F)$ satisfy

$$\mathbb{E} [\|X - Y\|_F] \leq \inf_{z \in F} \mathbb{E} [\|X - z\|_F + \|Y - z\|_F].$$

So, the above discussion is about the extent to which this use of the triangle inequality can be reversed.

Since the upper bound that we seek is in terms of the distance in $L_p(F)$ between independent copies of X and Y , this can be further used to control from above expressions such as $\mathbb{E}[\|X - Y\|_F^p]$ for X and Y not necessarily independent in terms of $\mathbb{E}[\|X' - Y'\|_F^p]$, where X' and Y' are independent, X' has the same distribution as X , and Y' has the same distribution as Y .

In order to analyse the inequality (2) in a specific Banach space $(F, \|\cdot\|_F)$, we consider the following geometric moduli. Given $p \geq 1$ let $\mathfrak{b}_p(F, \|\cdot\|_F)$, or simply $\mathfrak{b}_p(F)$ if the norm is clear from the context, be the infimum over those $m > 0$ such that every independent F -valued random variables $X, Y \in L_p(F)$ satisfy

$$\inf_{z \in F} \mathbb{E} [\|X - z\|_F^p + \mathbb{E} \|Y - z\|_F^p] \leq \mathfrak{b}_p \mathbb{E} [\|X - Y\|_F^p]. \quad (4)$$

Thus, $\mathfrak{b}_p(F)$ is precisely the best possible constant in the $L_p(F)$ -analogue of the aforementioned barycentric inequality (3). The use of the letter “ \mathfrak{b} ” in this notation is in reference to the word “barycentric.” Theorem 1.1 asserts that $\mathfrak{b}_p(F) \leq 3^p/2^{p-1}$, and that this bound cannot be improved in general.

Let $\mathfrak{m}_p(F, \|\cdot\|_F) > 0$, or simply $\mathfrak{m}_p(F)$ if the norm is clear from the context, be the infimum over those $\mathfrak{b} > 0$ such that every independent F -valued random variables $X, Y \in L_p(F)$ satisfy

$$\mathbb{E} \left[\left\| X - \frac{1}{2}\mathbb{E}[X] - \frac{1}{2}\mathbb{E}[Y] \right\|_F^p + \left\| Y - \frac{1}{2}\mathbb{E}[X] - \frac{1}{2}\mathbb{E}[Y] \right\|_F^p \right] \leq \mathfrak{m}_p \mathbb{E} [\|X - Y\|_F^p]. \quad (5)$$

The use of the letter “ \mathfrak{m} ” in this notation is in reference to the word “mixture,” since the left-hand side of (5) is equal to $2\mathbb{E}[\|Z - \mathbb{E}[Z]\|_F^p]$, where $Z \in L_p(F)$ distributed according to the mixture of the laws of X and Y , namely X is the F -valued random vector such that for every Borel set $A \subseteq F$,

$$\mathbb{P}[Z \in A] = \frac{1}{2}\mathbb{P}[X \in A] + \frac{1}{2}\mathbb{P}[Y \in A] \quad (6)$$

Obviously $\mathfrak{b}_p(F) \leq \mathfrak{m}_p(F)$, because (5) corresponds to choosing $z = \frac{1}{2}\mathbb{E}[X] + \frac{1}{2}\mathbb{E}[Y] \in F$ in (4).

While we sometimes bound $\mathfrak{m}_p(F)$ directly, it is beneficial to refine the considerations through the study of two further moduli that are natural in their own right and, as we shall see later, their use can lead to better bounds. Firstly, let $\mathfrak{z}_p(F, \|\cdot\|_F)$, or simply $\mathfrak{z}_p(F)$ if the norm is clear from the context, be the infimum over those $\mathfrak{z} > 0$ such that every independent F -valued random variables $X, Y \in L_p(F)$ satisfy

$$\mathbb{E} [\|X - X'\|_F^p] + \mathbb{E} [\|Y - Y'\|_F^p] \leq \mathfrak{z}_p \mathbb{E} [\|X - Y\|_F^p], \quad (7)$$

where X', Y' are independent copies of X and Y' , respectively. The use of the letter “ ε ” in this notation is in reference to the word “roundness,” as we shall next explain.

Observe also that (7) is a purely metric condition, i.e., it involves only distances between points. So, it makes sense to investigate (7) in any metric space $(\mathcal{M}, d_{\mathcal{M}})$, namely to study the inequality

$$\mathbb{E}[d_{\mathcal{M}}(X, X')^p] + \mathbb{E}[d_{\mathcal{M}}(Y, Y')^p] \leq \varepsilon \mathbb{E}[d_{\mathcal{M}}(X, Y)^p]. \quad (8)$$

One requires (8) to hold for \mathcal{M} -valued independent random variables X, X', Y, Y' (say, finitely-supported, to avoid measurability assumptions) such that each of the pairs X, X' and Y, Y' is identically distributed.

To the best of our knowledge, condition (8) was first studied systematically by Enflo [10], who defined a metric space $(\mathcal{M}, d_{\mathcal{M}})$ to have *generalized roundness* p if it satisfies (8) with $\varepsilon = 2$. He proved that L_p has generalized roundness p for $p \in [1, 2]$, and ingeniously used this notion to answer an old question of Smirnov. See [9] for a relatively recent example of substantial impact of Enflo’s approach. By combining [14] with [19], a metric space $(\mathcal{M}, d_{\mathcal{M}})$ has generalized roundness p if and only if $(\mathcal{M}, d_{\mathcal{M}}^{p/2})$ embeds isometrically into a Hilbert space. The case $\varepsilon > 1$ of (8) arose in [2] in the context of metric embeddings.

The final geometric modulus that we consider here is a quantity $j_p(F, \|\cdot\|_F)$, or simply $j_p(F)$ if the norm is clear from the context, that is defined to be the infimum over those $j \geq 1$ such that every independent and identically distributed F -valued random variables $Z, Z' \in L_p(F)$ satisfy

$$j \mathbb{E}[\|Z - \mathbb{E}[Z]\|_F^p] \leq \mathbb{E}[\|Z - Z'\|_F^p], \quad (9)$$

Note that (9) holds with $j = 1$ by Jensen’s inequality, so we are asking here for an improvement of (this use of) Jensen’s inequality by a definite factor; the letter “ j ” in this notation is in reference to “Jensen.”

We have the following general bounds, which hold for every Banach space $(F, \|\cdot\|_F)$ and every $p \geq 1$.

$$\mathfrak{b}_p(F) \leq m_p(F) \leq \frac{2 + \varepsilon_p(F)}{2j_p(F)}. \quad (10)$$

Indeed, we already observed the first inequality in (10), and the second inequality in (10) is justified by taking independent random variables $X, Y \in L_p(F)$, considering their mixture $Z \in L_p(F)$ as defined in (6), letting X', Y', Z' be independent copies of X, Y, Z , respectively, and proceeding as follows.

$$\begin{aligned} & \mathbb{E} \left[\left\| X - \frac{1}{2}\mathbb{E}[X] - \frac{1}{2}\mathbb{E}[Y] \right\|_F^p + \left\| Y - \frac{1}{2}\mathbb{E}[X] - \frac{1}{2}\mathbb{E}[Y] \right\|_F^p \right] \stackrel{(6)}{=} 2\mathbb{E}[\|Z - \mathbb{E}[Z]\|_F^p] \stackrel{(9)}{\leq} \frac{2}{j_p(F)} \mathbb{E}[\|Z - Z'\|_F^p] \\ & \stackrel{(6)}{=} \frac{2}{j_p(F)} \left(\frac{1}{2}\mathbb{E}[\|X - Y\|_F^p] + \frac{1}{4}\mathbb{E}[\|X - X'\|_F^p] + \frac{1}{4}\mathbb{E}[\|Y - Y'\|_F^p] \right) \leq \frac{2}{j_p(F)} \left(\frac{1}{2} + \frac{1}{4}\varepsilon_p(F) \right) \mathbb{E}[\|X - Y\|_F^p]. \end{aligned}$$

Recalling the definition (5) of $m_p(F)$, this implies (10).

Here we prove the following bounds on $\mathfrak{b}_p(L_q), m_p(L_q), \varepsilon_p(L_q), j_p(L_q)$ for $p, q \in [1, \infty)$.

Theorem 1.2. *For every $p, q \in [1, \infty)$ we have $j_p(L_q) = 2^{c(p,q)}$, where*

$$c(p, q) \stackrel{\text{def}}{=} \min \left\{ 1, p-1, \frac{p}{q}, \frac{p(q-1)}{q} \right\} = \begin{cases} p-1 & \text{if } 1 \leq p \leq q \leq 2 \text{ or } 1 \leq p \leq \frac{q}{q-1} \leq 2, \\ \frac{p(q-1)}{q} & \text{if } q \leq p \leq \frac{q}{q-1}, \\ \frac{p}{q} & \text{if } \frac{q}{q-1} \leq p \leq q, \\ 1 & \text{if } p \geq \frac{q}{q-1} \geq 2 \text{ or } p \geq q \geq 2. \end{cases} \quad (11)$$

We also have $\varepsilon_p(L_q) \leq 2^{C(p,q)}$, where

$$C(p, q) \stackrel{\text{def}}{=} \begin{cases} p-1 & \text{if } \frac{p}{q-1} \leq q \leq p, \\ \frac{p(q-2)}{q} + 1 & \text{if } \frac{q}{q-1} \leq p \leq q, \\ 2 - \frac{p}{q} & \text{if } q \geq 2 \text{ and } 1 \leq p \leq \frac{q}{q-1}, \\ \frac{p}{q} & \text{if } q \leq 2 \text{ and } q \leq p \leq \frac{q}{q-1}, \\ 1 & \text{if } 1 \leq p \leq q \leq 2. \end{cases} \quad (12)$$

In fact, if $\frac{p}{p-1} \leq q \leq p$, then $\tau_p(L_q) = 2^{p-1}$, if $\frac{q}{q-1} \leq p \leq q$, then $\tau_p(L_q) = 2^{\frac{p(q-2)}{q}+1}$, and $\tau_p(L_q) = 2$ if $1 \leq p \leq q \leq 2$. Namely, the above bound on $\tau_p(L_q)$ is sharp in the first, second and fifth ranges in (12).

Furthermore, $\mathfrak{b}_p(L_q) = m_p(L_q) = 2^{2-p}$ if $p \leq q \leq 2$. More generally, we have the bound

$$\mathfrak{b}_p(L_q) \leq m_p(L_q) \leq \min \left\{ \frac{3^p}{2^{p-1}} \left(\frac{\sqrt{2}}{3} \right)^{2c(p,q)}, \frac{2^{C(p,q)} + 2}{2^{c(p,q)+1}} \right\}. \quad (13)$$

The upper bound on $\mathfrak{b}_p(L_q)$ in (13) improves over (2) when $F = L_q$ for all values of $p, q \in [1, \infty)$. It would be interesting to find the exact value of $\mathfrak{b}_p(L_q)$ in the entire range $p, q \in [1, \infty)$. Note that the second quantity in the minimum in the right hand side of (13) corresponds to using (10) together with the bounds on $j_p(L_q)$ and $\tau_p(L_q)$ that Theorem 1.2 provides; when, say, $p = q$, this quantity is smaller than the first quantity in the minimum in the right hand side of (13) if and only if $1 \leq p < 3$.

Theorem 1.2 states that the constant $C(p, q)$ is sharp in the first, second and fifth ranges in (12). The following conjecture formulates what we expect to be the sharp values of $\tau_p(L_q)$ for all $p, q \in [1, \infty)$.

Conjecture 1.3. For all $p, q \in [1, \infty)$ we have $\tau_p(L_q) = 2^{C_{\text{opt}}(p,q)}$, where

$$C_{\text{opt}}(p, q) \stackrel{\text{def}}{=} \max \left\{ 1, p-1, \frac{p(q-2)}{q} + 1 \right\} = \begin{cases} p-1 & \text{if } p \geq 2 \text{ and } 1 \leq q \leq p, \\ \frac{p(q-2)}{q} + 1 & \text{if } q \geq 2 \text{ and } 1 \leq p \leq q, \\ 1 & \text{if } p, q \in [1, 2]. \end{cases} \quad (14)$$

We will prove later that $\tau_p(L_q) \geq 2^{C_{\text{opt}}(p,q)}$, so Conjecture (1.3) is about improving our upper bounds on $\tau_p(L_q)$ in the remaining third and fourth ranges that appear in (12).

Question 1.4. Below we will obtain improvements over (2) for other spaces besides $\{L_q : q \in [1, \infty)\}$, including e.g. the Schatten–von Neumann trace classes (see e.g. [20]) $\{S_q : q \in (1, \infty)\}$. However, parts of Theorem 1.2 rely on “commutative” properties of L_q which are not valid for S_q , thus leading to even better bounds in the commutative setting. It would be especially interesting to obtain sharp bounds in noncommutative probabilistic inequalities such as the roundness inequality (7) when $F = S_q$. In particular, we ask what is the value of $\tau_1(S_1)$? At present, we know (as was already shown by Enflo [10]) that $\tau_1(L_1) = 2$ while the only bound that we have for S_1 is $\tau_1(S_1) \leq 4$. Note that 4 is a trivial upper bound here, which holds for every Banach space. Interestingly, it follows from [7] that $\tau_1(S_1) \geq 2\sqrt{2}$, as explained in Remark 3.1 below. So, there is a genuine difference between the commutative and noncommutative settings of L_1 and S_1 , respectively. As a more modest question, is $\tau_1(S_1)$ strictly less than 4?

1.3. Complex interpolation. We will use basic terminology, notation and results of complex interpolation of Banach spaces; the relevant background appears in [8, 4]. Theorem 1.2 is a special case of the following more general result about interpolation spaces. As such, it applies also to random variables that take values in certain spaces other than L_q , including, for examples, Schatten–von Neumann trace classes (see e.g. [20]) and, by an extrapolation theorem of Pisier [18], Banach lattices of nontrivial type.

Theorem 1.5. Fix $\theta \in [0, 1]$ and $\frac{2}{2-\theta} \leq p \leq \frac{2}{\theta}$. Let $(F, \|\cdot\|_F), (H, \|\cdot\|_H)$ be a compatible pair of Banach spaces such that $(H, \|\cdot\|_H)$ is a Hilbert space. Then the following estimates hold true.

$$\tau_p([F, H]_\theta) \leq 2^{1+(1-\theta)p} \quad \text{and} \quad j_p([F, H]_\theta) \geq 2^{\frac{\theta p}{2}}. \quad (15)$$

Additionally, we have

$$\mathfrak{b}_p([F, H]_\theta) \leq m_p([F, H]_\theta) \leq \min \left\{ \frac{3^p}{2^{p-1}} \left(\frac{\sqrt{2}}{3} \right)^{p\theta}, \frac{1 + 2^{(1-\theta)p}}{2^{\frac{\theta p}{2}}} \right\} = \begin{cases} \frac{3^p}{2^{p-1}} \left(\frac{\sqrt{2}}{3} \right)^{p\theta} & \text{if } \frac{1}{1-\theta} \leq p \leq \frac{2}{\theta}, \\ \frac{1 + 2^{(1-\theta)p}}{2^{\frac{\theta p}{2}}} & \text{if } \frac{2}{2-\theta} \leq p \leq \frac{1}{1-\theta}. \end{cases} \quad (16)$$

(Note that if the first range of values of p in the right hand side of (16) is nonempty, then necessarily $\theta \leq \frac{2}{3}$.)

The deduction of Theorem 1.2 from Theorem 1.5 appears in Section 3 below; in most cases this deduction is nothing more than a direct substitution into Theorem 1.5, but in some cases a further argument is needed. Theorem 1.5 itself is a special case of the following theorem.

Theorem 1.6. Fix $\theta \in [0, 1]$ and $p \in [1, \infty]$ that satisfy $\frac{2}{2-\theta} \leq p \leq \frac{2}{\theta}$. Let $(F, \|\cdot\|_F), (H, \|\cdot\|_H)$ be a compatible pair of Banach spaces such that $(H, \|\cdot\|_H)$ is a Hilbert space. Suppose that (\mathcal{X}, μ) and (\mathcal{Y}, ν) are probability spaces. Then, for every $f \in L_p(\mu \times \nu; [F, H]_\theta)$ we have

$$2^{1+(1-\theta)p} \iint_{\mathcal{X} \times \mathcal{Y}} \|f(x, y)\|_{[F, H]_\theta}^p d\mu(x) d\nu(y) \geq \iint_{\mathcal{X} \times \mathcal{X}} \left\| \int_{\mathcal{Y}} (f(x, y) - f(\chi, y)) d\nu(y) \right\|_{[F, H]_\theta}^p d\mu(x) d\mu(\chi) \\ + \iint_{\mathcal{Y} \times \mathcal{Y}} \left\| \int_{\mathcal{X}} (f(x, y) - f(x, \nu)) d\mu(x) \right\|_{[F, H]_\theta}^p d\nu(y) d\nu(\nu), \quad (17)$$

and

$$\frac{3^p}{2^{p-1}} \left(\frac{\sqrt{2}}{3} \right)^{p\theta} \iint_{\mathcal{X} \times \mathcal{Y}} \|f(x, y)\|_{[F, H]_\theta}^p d\mu(x) d\nu(y) \\ \geq \int_{\mathcal{X}} \left\| \int_{\mathcal{Y}} f(x, y) d\nu(y) - \frac{1}{2} \iint_{\mathcal{X} \times \mathcal{Y}} f(\chi, \nu) d\mu(\chi) d\nu(\nu) \right\|_{[F, H]_\theta}^p d\mu(x) \\ + \int_{\mathcal{Y}} \left\| \int_{\mathcal{X}} f(x, y) d\mu(x) - \frac{1}{2} \iint_{\mathcal{X} \times \mathcal{Y}} f(\chi, \nu) d\mu(\chi) d\nu(\nu) \right\|_{[F, H]_\theta}^p d\nu(y). \quad (18)$$

Furthermore, if $g \in L_p(\mu \times \mu; [F, H]_\theta)$, then

$$2^{(1-\frac{\theta}{2})p} \iint_{\mathcal{X} \times \mathcal{X}} \|g(x, \chi)\|_{[F, H]_\theta}^p d\mu(x) d\mu(\chi) \geq \int_{\mathcal{X}} \left\| \int_{\mathcal{X}} (g(x, \chi) - g(\chi, x)) d\mu(x) \right\|_{[F, H]_\theta}^p d\mu(\chi). \quad (19)$$

Proof of Theorem 1.5 assuming Theorem 1.6. Let X and Y be independent p -integrable $[F, H]_\theta$ -valued random vectors. Due to the independence assumption, without loss of generality there are probability spaces (\mathcal{X}, μ) and (\mathcal{Y}, ν) such that X and Y are elements of $L_p(\mu \times \nu; [F, H]_\theta)$ that depend only on the first variable and second variable, respectively. Then (17) and (18) applied to $f = X - Y$ become

$$\mathbb{E} \left[\|X - X'\|_{[F, H]_\theta}^p \right] + \mathbb{E} \left[\|Y - Y'\|_{[F, H]_\theta}^p \right] \leq 2^{1+(1-\theta)p} \mathbb{E} \left[\|X - Y\|_{[F, H]_\theta}^p \right],$$

and

$$\mathbb{E} \left[\left\| X - \frac{1}{2} \mathbb{E}[X] - \frac{1}{2} \mathbb{E}[Y] \right\|_{[F, H]_\theta}^p \right] + \mathbb{E} \left[\left\| Y - \frac{1}{2} \mathbb{E}[X] - \frac{1}{2} \mathbb{E}[Y] \right\|_{[F, H]_\theta}^p \right] \leq \frac{3^p}{2^{p-1}} \left(\frac{\sqrt{2}}{3} \right)^{p\theta} \mathbb{E} \left[\|X - Y\|_{[F, H]_\theta}^p \right].$$

We therefore established the first inequality in (15) as well as the upper bound on $m_p([F, H]_\theta)$ that corresponds to the first term in the minimum that appears in (16).

Similarly, due to the fact that X and X' are i.i.d., without loss of generality there is a probability space (\mathcal{X}, μ) such that X and X' are elements of $L_p(\mu \times \mu; [F, H]_\theta)$ that depend only on the first variable and second variable, respectively. Then, (19) applied to $g = X - X'$ simplifies to give

$$\mathbb{E} \left[\|X - X'\|_{[F, H]_\theta}^p \right] \geq 2^{\frac{\theta p}{2}} \mathbb{E} \left[\|X - \mathbb{E}[X]\|_{[F, H]_\theta}^p \right].$$

This establishes the second inequality in (15), as well as the upper bound on $m_p([F, H]_\theta)$ that corresponds to the second term in the minimum that appears in (16), due to (10). \square

The first and third inequalities of Theorem 1.6 are generalizations of results that appeared in the literature. Specifically, (17) generalizes Lemma 6 of [2], and (19) generalizes Lemma 5 of [17], which is itself inspired by a step within the proof of Theorem 2 of [21]. The proof of Theorem 1.6, which appears in Section 3 below, differs from the proofs of [21, 17, 2], but relies on the same ideas.

2. PROOF OF THEOREM 1.1

Let $(F, \|\cdot\|_F)$ be a Banach space. Fix $p \geq 1$. Theorem 1.1 asserts that $\mathfrak{b}_p(F) \leq \frac{3^p}{2^{p-1}}$. In fact, $\mathfrak{m}_p(F) \leq \frac{3^p}{2^{p-1}}$, which is stronger by (10). To see this, let $X, Y \in L_p(F)$ be independent random vectors and observe that

$$\begin{aligned} \mathbb{E} \left[\left\| X - \frac{1}{2}\mathbb{E}[X] - \frac{1}{2}\mathbb{E}[Y] \right\|_F^p \right] &= \frac{3^p}{2^p} \mathbb{E} \left[\left\| \frac{2}{3}(X - \mathbb{E}[Y]) + \frac{1}{3}(\mathbb{E}[Y] - \mathbb{E}[X]) \right\|_F^p \right] \\ &\leq \frac{3^p}{2^p} \left(\frac{2}{3} \mathbb{E} [\|X - \mathbb{E}[Y]\|_F^p] + \frac{1}{3} \mathbb{E} [\|\mathbb{E}[Y] - \mathbb{E}[X]\|_F^p] \right) \leq \frac{3^p}{2^p} \mathbb{E} [\|X - Y\|_F^p], \end{aligned}$$

where the penultimate step holds due to the convexity of $\|\cdot\|_F^p$ and the final step holds because, by Jensen's inequality, both $\mathbb{E} [\|X - \mathbb{E}[Y]\|_F^p] = \mathbb{E} [\|\mathbb{E}_Y[X - Y]\|_F^p]$ and $\mathbb{E} [\|\mathbb{E}[Y] - \mathbb{E}[X]\|_F^p]$ are at most $\mathbb{E} [\|X - Y\|_F^p]$. The symmetric reasoning with X replaced by Y now gives

$$\begin{aligned} &\mathbb{E} \left[\left\| X - \frac{1}{2}\mathbb{E}[X] - \frac{1}{2}\mathbb{E}[Y] \right\|_F^p \right] + \mathbb{E} \left[\left\| Y - \frac{1}{2}\mathbb{E}[X] - \frac{1}{2}\mathbb{E}[Y] \right\|_F^p \right] \\ &\leq 2 \max \left\{ \mathbb{E} \left[\left\| X - \frac{1}{2}\mathbb{E}[X] - \frac{1}{2}\mathbb{E}[Y] \right\|_F^p \right], \mathbb{E} \left[\left\| Y - \frac{1}{2}\mathbb{E}[X] - \frac{1}{2}\mathbb{E}[Y] \right\|_F^p \right] \right\} \leq \frac{3^p}{2^{p-1}} \mathbb{E} [\|X - Y\|_F^p]. \end{aligned}$$

This shows that $\mathfrak{m}_p(F) \leq \frac{3^p}{2^{p-1}}$. It remains to prove that the bound $\mathfrak{b}_p(F) \leq \frac{3^p}{2^{p-1}}$ is optimal for general F .

Fix an integer $n \geq 2$ and consider

$$F_n \stackrel{\text{def}}{=} \left\{ x \in \mathbb{C}^{2n} : \sum_{k=1}^{2n} x_k = 0 \right\},$$

equipped with supremum norm inherited from ℓ_∞^{2n} . We will prove that

$$\mathfrak{b}_p(F_n) \geq 2 \left(\frac{3}{2} - \frac{1}{n} \right)^p \xrightarrow{n \rightarrow \infty} \frac{3^p}{2^{p-1}}. \quad (20)$$

Denote by $\{e_k\}_{k=1}^{2n}$ the standard coordinate basis of ℓ_∞^{2n} . Define two n -element sets $A_n, B_n \subseteq F_n$ by

$$A_n \stackrel{\text{def}}{=} \left\{ (3n-2)e_j - (n+2) \sum_{k \in \{1, \dots, n\} \setminus \{j\}} e_k + (n-2) \sum_{k=n+1}^{2n} e_k : j \in \{1, \dots, n\} \right\},$$

and

$$B_n \stackrel{\text{def}}{=} \left\{ (n-2) \sum_{k=1}^n e_k + (3n-2)e_j - (n+2) \sum_{k \in \{n+1, \dots, 2n\} \setminus \{j\}} e_k : j \in \{n+1, \dots, 2n\} \right\}.$$

Note that A_n and B_n are indeed subsets of F_n because $3n-2 - (n-1)(n+2) + n(n-2) = 0$. Let X, Y be independent and uniformly distributed on A_n, B_n , respectively. One checks that $\|a - b\|_\infty = 2n$ for any $a \in A_n$ and $b \in B_n$. So, $\mathbb{E} [\|X - Y\|_\infty^p] = (2n)^p$. The desired bound (20) will follow if we demonstrate that

$$\forall z \in F_n, \quad \mathbb{E} [\|X - z\|_\infty^p + \|Y - z\|_\infty^p] \geq 2(3n-2)^p. \quad (21)$$

The proof of (21) proceeds via symmetrization. For permutations $\sigma, \rho \in S_n$, define $T_{\sigma, \rho} : F_n \rightarrow F_n$ by

$$\forall x = (x_1, \dots, x_{2n}) \in F_n, \quad T_{\sigma, \rho}(x) \stackrel{\text{def}}{=} (x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)}, x_{n+\rho(1)}, x_{n+\rho(2)}, \dots, x_{n+\rho(n)}).$$

$T_{\sigma, \rho}$ is a linear isometry of F_n and the sets A_n and B_n are $T_{\sigma, \rho}$ -invariant. Hence, for any $z \in F_n$,

$$\begin{aligned} \mathbb{E} [\|X - z\|_\infty^p] &= \frac{1}{(n!)^2} \sum_{\sigma, \rho \in S_n} \mathbb{E} [\|T_{\sigma, \rho}(X) - z\|_\infty^p] = \frac{1}{(n!)^2} \sum_{\sigma, \rho \in S_n} \mathbb{E} [\|X - T_{\sigma^{-1}, \rho^{-1}}(z)\|_\infty^p] \\ &\geq \mathbb{E} \left[\left\| X - \frac{1}{(n!)^2} \sum_{\sigma, \rho \in S_n} T_{\sigma^{-1}, \rho^{-1}}(z) \right\|_\infty^p \right] = \mathbb{E} \left[\left\| X - \frac{z_1 + \dots + z_n}{n} \sum_{k=1}^n e_k + \frac{z_1 + \dots + z_n}{n} \sum_{k=n+1}^{2n} e_k \right\|_\infty^p \right]. \end{aligned} \quad (22)$$

Denoting $u = (z_1 + \dots + z_n)/n$, it follows from (22) that $\mathbb{E} [\|X - z\|_\infty^p] \geq |3n-2-u|^p$, because one of the first n coordinates of any member of the support of X equals $3n-2$. The same argument with X replaced

by Y gives that $\mathbb{E}[\|Y - z\|_\infty^p] \geq |3n - 2 + u|^p$, because now one of the last n coordinates of any member of the support of Y equals $3n - 2$. We conclude with the following application of the convexity of $|\cdot|^p$.

$$\mathbb{E}[\|X - z\|_\infty^p + \|Y - z\|_\infty^p] \geq |3n - 2 - u|^p + |3n - 2 + u|^p \geq 2(3n - 2)^p. \quad \square$$

Remark 2.1. It is worthwhile to examine what the above argument gives if we take the norm on F_n to be the norm inherited from ℓ_q^{2n} . One computes that $\|a - b\|_q = (2n)^{1+1/q}$ for every $a \in A_n$ and $b \in B_n$. So,

$$\mathbb{E}[\|X - Y\|_q^p] = (2n)^{\frac{p(q+1)}{q}}. \quad (23)$$

Also, it follows from the same reasoning that led to (22) that for every $z \in F_n$,

$$\mathbb{E}[\|X - z\|_q^p] \geq \mathbb{E}\left[\left\|X - u \sum_{k=1}^n e_k + u \sum_{k=n+1}^n e_k\right\|_q^p\right] = (|3n - 2 - u|^q + (n - 1)|n + 2 + u|^q + n|n - 2 + u|^q)^{\frac{p}{q}},$$

and

$$\mathbb{E}[\|Y - z\|_q^p] \geq \mathbb{E}\left[\left\|Y - u \sum_{k=1}^n e_k + u \sum_{k=n+1}^n e_k\right\|_q^p\right] = (|3n - 2 + u|^q + (n - 1)|n + 2 - u|^q + n|n - 2 - u|^q)^{\frac{p}{q}}.$$

Hence, using the convexity of the p 'th power of the ℓ_q norm on \mathbb{R}^3 , we see that

$$\mathbb{E}[\|X - z\|_q^p] + \mathbb{E}[\|Y - z\|_q^p] \geq 2((3n - 2)^q + (n - 1)(n + 2)^q + n(n - 2)^q)^{\frac{p}{q}}. \quad (24)$$

By contrasting (23) with (24) we conclude that

$$\mathfrak{b}_p(F_n, \|\cdot\|_q) \geq 2((3n - 2)^q + (n - 1)(n + 2)^q + n(n - 2)^q)^{\frac{p}{q}} (2n)^{-\frac{p(q+1)}{q}}.$$

In particular, if we take $p = q \geq 2$ and $n = \lceil q \rceil$, then we conclude that $\mathfrak{b}_q(F_n, \|\cdot\|_q) \geq \frac{c}{q} (\frac{3}{2})^q$ for some universal constant $c > 0$. So, there is very little potential asymptotic gain (as $q \rightarrow \infty$) if we know that the Banach space of Theorem 1.1 admits an isometric embedding into L_q .

Above, and in what follows, we stated that a normed space admits an isometric embedding into L_q without specifying whether the embedding is linear or not. Later we will need such embeddings to be linear, so we recall that for any $q \geq 1$, by a classical differentiation argument (see [3, Chapter 7] for a thorough treatment of such reductions to the linear setting), a normed space embeds isometrically into L_q as a metric space if and only if it admits a linear isometric embedding into L_q .

Note that the phenomenon of Remark 2.1 is special to random variables that have different expectations. Namely, if $\mathbb{E}[X] = \mathbb{E}[Y]$, then by Jensen's inequality the ratio that defines $\mathfrak{b}_q(F)$ is at most 2 rather than the aforementioned exponential growth as $q \rightarrow \infty$. The following proposition shows that if F is a subspace of L_q for $q \geq 3$, then when $\mathbb{E}[X] = \mathbb{E}[Y]$ this ratio is at most 1, which is easily seen to be best possible (consider any nontrivial symmetric random variable X , and take Y to be identically 0).

Proposition 2.2. *Let $(F, \|\cdot\|_F)$ be a Banach space that admits an isometric embedding into L_q for some $q \in [3, \infty)$. Then, for any pair of independent F -valued random vectors $X, Y \in L_q(F)$ with $\mathbb{E}[X] = \mathbb{E}[Y]$,*

$$\inf_{z \in F} \mathbb{E}[\|X - z\|_F^q + \|Y - z\|_F^q] \leq \mathbb{E}[\|X - \mathbb{E}[X]\|_F^q + \|Y - \mathbb{E}[X]\|_F^q] \leq \mathbb{E}[\|X - Y\|_F^q]. \quad (25)$$

Proof. L_q over \mathbb{C} embeds isometrically into L_q over \mathbb{R} (indeed, complex L_q is, as a real Banach space, the same as $L_q(\ell_2^2)$, so this follows from the fact that Hilbert space is isometric to a subspace of L_q). So, in Proposition 2.2 we may assume that F embeds isometrically into L_q over \mathbb{R} , and therefore by integration/Fubini it suffices to prove (25) for real-valued random variables. So, our goal is to show that if X, Y are independent mean-zero real random variables with $\mathbb{E}[|X|^q], \mathbb{E}[|Y|^q] < \infty$, then

$$\mathbb{E}[|X + Y|^q] \geq \mathbb{E}[|X|^q] + \mathbb{E}[|Y|^q]. \quad (26)$$

The bound (25) would then follow by applying (26) to the mean-zero variables $X - \mathbb{E}[X]$ and $\mathbb{E}[X] - Y$.

Note in passing that the assumption $q \geq 3$ is crucial here, i.e. (26) fails if $q \in (0, 3) \setminus \{2\}$. Indeed, if $\beta \in (0, \frac{1}{2})$ and $\mathbb{P}[X = 1 - \beta] = \mathbb{P}[Y = 1 - \beta] = \beta$ and $\mathbb{P}[X = -\beta] = \mathbb{P}[Y = -\beta] = 1 - \beta$, then $\mathbb{E}[X] = \mathbb{E}[Y] = 0$ but

$$\frac{\mathbb{E}[|X + Y|^q]}{\mathbb{E}[|X|^q] + \mathbb{E}[|Y|^q]} = \frac{\beta^2 2^q (1 - \beta)^q + (1 - \beta)^2 2^q \beta^q + 2\beta(1 - \beta)(1 - 2\beta)^q}{2\beta(1 - \beta)^q + 2(1 - \beta)\beta^q}. \quad (27)$$

If $q \in (0, 2)$, then the right hand side of (27) equals $2^{q-2} < 1$ for $\beta = \frac{1}{2}$. If $q \in (2, 3)$, then the right hand side of (27) equals $1 + (2^{q-1} - q - 1)\beta + o(\beta)$, which is less than 1 for small β since $2^{q-1} - q - 1 < 0$ for $q \in (2, 3)$.

To prove (26), for every $s > 0$ and $x \in \mathbb{R}$, denote $\phi_s(x) = \text{sign}(x) \cdot |x|^s$. Observe that

$$\forall x, y \in \mathbb{R}, \quad \alpha(x, y) \stackrel{\text{def}}{=} |x + y|^q - |x|^q - |y|^q - q\phi_{q-1}(x)y - qx\phi_{q-1}(y) \geq 0. \quad (28)$$

Once (28) is proved, (26) would follow because

$$\begin{aligned} 0 \leq \mathbb{E}[\alpha(X, Y)] &= \mathbb{E}[|X + Y|^q] - \mathbb{E}[|X|^q] - \mathbb{E}[|Y|^q] - q\mathbb{E}[Y]\mathbb{E}[\phi_{q-1}(X)] - q\mathbb{E}[X]\mathbb{E}[\phi_{q-1}(Y)] \\ &= \mathbb{E}[|X + Y|^q] - \mathbb{E}[|X|^q] - \mathbb{E}[|Y|^q], \end{aligned}$$

where the penultimate step uses the independence of X, Y and the last step uses $\mathbb{E}[X] = \mathbb{E}[Y] = 0$.

It suffices to prove (28) when $q > 3$; the case $q = 3$ follows by passing to the limit. One checks that

$$\frac{\partial^3 \alpha}{\partial x^2 \partial y} = p(p-1)(p-2)(\phi_{p-3}(x+y) - \phi_{p-3}(x)) \implies \text{sign}\left(\frac{\partial^3 \alpha}{\partial x^2 \partial y}(x, y)\right) = \text{sign}(y),$$

where the last step holds because ϕ_{p-3} is increasing. Hence, $y \mapsto \frac{\partial^2 \alpha}{\partial x^2}(x, y)$ is decreasing for $y < 0$ and increasing for $y > 0$. One checks that $\frac{\partial^2 \alpha}{\partial x^2}(x, 0) = 0$ for all $x \in \mathbb{R}$, so $\frac{\partial^2 \alpha}{\partial x^2}(x, y) \geq 0$. Thus $x \mapsto \alpha(x, y)$ is convex for every fixed $y \in \mathbb{R}$. But $\alpha(0, y) = \frac{\partial \alpha}{\partial x}(0, y) = 0$ for any $y \in \mathbb{R}$, i.e. the tangent to the graph of $x \mapsto \alpha(x, y)$ at $x = 0$ is the x -axis. Convexity implies that the graph of $x \mapsto \alpha(x, y)$ lies above the x -axis, as required. \square

We end this section with the following simpler metric space counterpart of Theorem 1.1.

Proposition 2.3. *Fix $p \geq 1$ and let X and Y be independent finitely supported random variables taking values in a metric space $(\mathcal{M}, d_{\mathcal{M}})$. Then*

$$\inf_{z \in \mathcal{M}} \mathbb{E}[d_{\mathcal{M}}(X, z)^p + d_{\mathcal{M}}(Y, z)^p] \leq (2^p + 1) \mathbb{E}[d_{\mathcal{M}}(X, Y)^p]. \quad (29)$$

The constant $2^p + 1$ in (29) is optimal.

Proof. Let X' have the same distribution as X and be independent of X and Y . The point-wise inequality

$$d_{\mathcal{M}}(X, X')^p \leq (d_{\mathcal{M}}(X, Y) + d_{\mathcal{M}}(Y, X'))^p \leq 2^{p-1} (d_{\mathcal{M}}(X, Y)^p + d_{\mathcal{M}}(X', Y)^p)$$

is a consequence of the triangle inequality and the convexity of $(u > 0) \mapsto u^p$. By taking expectations, we obtain $\mathbb{E}[d_{\mathcal{M}}(X, X')^p] \leq 2^p \mathbb{E}[d_{\mathcal{M}}(X, Y)^p]$, so that

$$\inf_{z \in \mathcal{M}} \mathbb{E}[d_{\mathcal{M}}(X, z)^p + d_{\mathcal{M}}(Y, z)^p] \leq \mathbb{E}[d_{\mathcal{M}}(X, X')^p + d_{\mathcal{M}}(Y, X')^p] \leq (2^p + 1) \mathbb{E}[d_{\mathcal{M}}(X, Y)^p].$$

To see that the constant $2^p + 1$ is optimal, fix $n \in \mathbb{N}$ and let \mathcal{M} be the complete bipartite graph $K_{n,n}$, equipped with its shortest-path metric. Equivalently, \mathcal{M} can be partitioned into two n -point subsets L, R , and for distinct $x, y \in \mathcal{M}$ we have $d_{\mathcal{M}}(x, y) = 2$ if $\{x, y\} \subseteq L$ or $\{x, y\} \subseteq R$, while $d_{\mathcal{M}}(x, y) = 1$ otherwise. Let X be uniformly distributed over L and Y be uniformly distributed over R . Then $d_{\mathcal{M}}(X, Y) = 1$ point-wise. If $z \in L$, then $d_{\mathcal{M}}(Y, z) = 1$ point-wise, while $\mathbb{P}[d_{\mathcal{M}}(X, z) = 2] = \frac{n-1}{n}$ and $\mathbb{P}[d_{\mathcal{M}}(X, z) = 0] = \frac{1}{n}$. Consequently,

$$\frac{\mathbb{E}[d_{\mathcal{M}}(X, z)^p + d_{\mathcal{M}}(Y, z)^p]}{\mathbb{E}[d_{\mathcal{M}}(X, Y)^p]} = \frac{n-1}{n} 2^p + 1 \xrightarrow{n \rightarrow \infty} 2^p + 1.$$

By symmetry, the same holds if $z \in R$. \square

3. PROOF OF THEOREM 1.6 AND ITS CONSEQUENCES

Here we prove Theorem 1.6 and deduce Theorem 1.2.

Proof of Theorem 1.6. The assumption $\frac{2}{2-\theta} \leq p \leq \frac{2}{\theta}$ implies that $\frac{1}{p} = \frac{1-\theta}{q} + \frac{\theta}{2}$ for some (unique) $q \in [1, \infty]$. We will fix this value of q for the rest of the proof of Theorem 1.6. All of the desired bounds (17), (18), (19) hold true when $\theta = 0$, namely for every Banach space $(F, \|\cdot\|_F)$ and every $f \in L_q(\mu \times \nu; F)$ we have

$$\begin{aligned} & 2^{q+1} \iint_{X \times Y} \|f(x, y)\|_F^q d\mu(x) d\nu(y) \\ & \geq \iint_{X \times X} \left\| \int_Y (f(x, y) - f(x, y)) d\nu(y) \right\|_F^q d\mu(x) d\mu(\chi) \\ & \quad + \iint_{Y \times Y} \left\| \int_X (f(x, y) - f(x, \nu)) d\mu(x) \right\|_F^q d\nu(y) d\nu(\nu), \end{aligned} \quad (30)$$

and

$$\begin{aligned} & \frac{3^q}{2^{q-1}} \iint_{X \times Y} \|f(x, y)\|_F^q d\mu(x) d\nu(y) \\ & \geq \int_X \left\| \int_Y f(x, y) d\nu(y) - \frac{1}{2} \iint_{X \times Y} f(\chi, \nu) d\mu(\chi) d\nu(\nu) \right\|_F^q d\mu(x) \\ & \quad + \int_Y \left\| \int_X f(x, y) d\mu(x) - \frac{1}{2} \iint_{X \times Y} f(\chi, \nu) d\mu(\chi) d\nu(\nu) \right\|_F^q d\nu(y). \end{aligned} \quad (31)$$

Furthermore, if $g \in L_q(\mu \times \mu; F)$, then

$$2^q \iint_{X \times X} \|g(x, \chi)\|_F^q d\mu(x) d\mu(\chi) \geq \int_X \left\| \int_X (g(x, \chi) - g(\chi, x)) d\mu(x) \right\|_F^q d\mu(\chi). \quad (32)$$

Indeed, (30), (31), (32) are direct consequences of the triangle inequality in $L_q(\mu \times \nu; F)$ and $L_q(\mu \times \mu; F)$ and Jensen's inequality, with the appropriate interpretation when $q = \infty$.

By complex interpolation theory (specifically, by combining [4, Theorem 4.1.2] and [4, Theorem 5.1.2]), Theorem 1.6 will follow if we prove the $\theta = 1$ case of (17), (18), (19). To this end, as H is a Hilbert space and the inequalities in question are quadratic, it suffices to prove them coordinate-wise (with respect to any orthonormal basis of H), i.e., it suffices to show that for every (\mathbb{C} -valued) $f \in L_2(\mu \times \nu)$ and $g \in L_2(\mu \times \mu)$,

$$\begin{aligned} & 2 \iint_{X \times Y} |f(x, y)|^2 d\mu(x) d\nu(y) \\ & \geq \iint_{X \times X} \left| \int_Y (f(x, y) - f(x, y)) d\nu(y) \right|^2 d\mu(x) d\mu(\chi) \\ & \quad + \iint_{Y \times Y} \left| \int_X (f(x, y) - f(x, \nu)) d\mu(x) \right|^2 d\nu(y) d\nu(\nu), \end{aligned} \quad (33)$$

and

$$\begin{aligned} & \iint_{X \times Y} |f(x, y)|^2 d\mu(x) d\nu(y) \\ & \geq \int_X \left| \int_Y f(x, y) d\nu(y) - \frac{1}{2} \iint_{X \times Y} f(\chi, \nu) d\mu(\chi) d\nu(\nu) \right|^2 d\mu(x) \\ & \quad + \int_Y \left| \int_X f(x, y) d\mu(x) - \frac{1}{2} \iint_{X \times Y} f(\chi, \nu) d\mu(\chi) d\nu(\nu) \right|^2 d\nu(y), \end{aligned} \quad (34)$$

and

$$2 \iint_{X \times X} |g(x, \chi)|^2 d\mu(x) d\mu(\chi) \geq \int_X \left| \int_X (g(x, \chi) - g(\chi, x)) d\mu(x) \right|^2 d\mu(\chi). \quad (35)$$

The following derivation of the quadratic scalar inequalities (33), (34), (35) is an exercise in linear algebra.

Let $\{\varphi_j\}_{j=0}^\infty \subseteq L_2(\mu)$ and $\{\psi_k\}_{k=0}^\infty \subseteq L_2(\nu)$ be any orthonormal bases of $L_2(\mu)$ and $L_2(\nu)$, respectively, for which $\varphi_0 = \mathbf{1}_X$ and $\psi_0 = \mathbf{1}_Y$. Then $\{\varphi_j \otimes \psi_k\}_{j,k=0}^\infty$, $\{\varphi_j \otimes \varphi_k\}_{j,k=0}^\infty$ and $\{\psi_j \otimes \psi_k\}_{j,k=0}^\infty$ are orthonormal bases of $L_2(\mu \times \nu)$, $L_2(\mu \times \mu)$ and $L_2(\nu \times \nu)$, respectively, where for $\varphi \in L_2(\mu)$ and $\psi \in L_2(\nu)$ one defines (as usual) $\varphi \otimes \psi : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{C}$ by setting $\varphi \otimes \psi(x, y) = \varphi(x)\psi(y)$ for $(x, y) \in \mathcal{X} \times \mathcal{Y}$. We therefore have the following expansions, in the sense of convergence in $L_2(\mu \times \nu)$ and $L_2(\mu \times \mu)$, respectively.

$$f = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \langle \xi, \varphi_j \otimes \psi_k \rangle_{L_2(\mu \times \nu)} \varphi_j \otimes \psi_k \quad \text{and} \quad g = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \langle \zeta, \varphi_j \otimes \varphi_k \rangle_{L_2(\mu \times \mu)} \varphi_j \otimes \varphi_k.$$

In particular, by Parseval we have

$$\|f\|_{L_2(\mu \times \nu)}^2 = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} |\langle f, \varphi_j \otimes \psi_k \rangle_{L_2(\mu \times \nu)}|^2 \quad \text{and} \quad \|g\|_{L_2(\mu \times \mu)}^2 = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} |\langle g, \varphi_j \otimes \varphi_k \rangle_{L_2(\mu \times \mu)}|^2. \quad (36)$$

Define $R_X f \in L_2(\mu \times \mu)$ by

$$R_X f \stackrel{\text{def}}{=} \sum_{j=1}^{\infty} \langle f, \varphi_j \otimes \psi_0 \rangle_{L_2(\mu \times \nu)} \varphi_j \otimes \varphi_0 - \sum_{j=1}^{\infty} \langle f, \varphi_j \otimes \psi_0 \rangle_{L_2(\mu \times \nu)} \varphi_0 \otimes \varphi_j.$$

So, $(\mu \times \mu)$ -almost surely $R_X f(x, \chi) = \int_Y (f(x, y) - f(\chi, y)) d\nu(y)$. Also, define $R_Y f \in L_2(\nu \times \nu)$ by

$$R_Y f \stackrel{\text{def}}{=} \sum_{j=1}^{\infty} \langle f, \varphi_0 \otimes \psi_j \rangle_{L_2(\mu \times \nu)} \psi_j \otimes \psi_0 - \sum_{j=1}^{\infty} \langle f, \varphi_0 \otimes \psi_j \rangle_{L_2(\mu \times \nu)} \psi_0 \otimes \psi_j.$$

So, $(\nu \times \nu)$ -almost surely $R_Y f(y, \nu) = \int_X (f(x, y) - f(x, \nu)) d\mu(x)$. By Parseval in $L_2(\mu \times \mu)$, $L_2(\nu \times \nu)$, $L_2(\mu \times \nu)$,

$$\|R_X f\|_{L_2(\mu \times \mu)}^2 + \|R_Y f\|_{L_2(\nu \times \nu)}^2 = 2 \sum_{j=1}^{\infty} \left(|\langle f, \varphi_j \otimes \psi_0 \rangle_{L_2(\mu \times \nu)}|^2 + |\langle f, \varphi_0 \otimes \psi_j \rangle_{L_2(\mu \times \nu)}|^2 \right) \stackrel{(36)}{\leq} 2 \|f\|_{L_2(\mu \times \nu)}^2.$$

This is precisely (33).

Next, for every $\alpha, \beta \in \mathbb{C}$ define $S_X^\alpha f \in L_2(\mu)$ and $S_Y^\beta f \in L_2(\nu)$ by

$$S_X^\alpha f \stackrel{\text{def}}{=} (1 - \alpha) \langle f, \varphi_0 \otimes \psi_0 \rangle_{L_2(\mu \times \nu)} \varphi_0 + \sum_{j=1}^{\infty} \langle f, \varphi_j \otimes \psi_0 \rangle_{L_2(\mu \times \nu)} \varphi_j,$$

and

$$S_Y^\beta f \stackrel{\text{def}}{=} (1 - \beta) \langle f, \varphi_0 \otimes \psi_0 \rangle_{L_2(\mu \times \nu)} \psi_0 + \sum_{j=1}^{\infty} \langle f, \varphi_0 \otimes \psi_j \rangle_{L_2(\mu \times \nu)} \psi_j.$$

In other words, we have the following identities μ -almost surely and ν -almost surely, respectively.

$$S_X^\alpha f(x) = \int_Y f(x, y) d\nu(y) - \alpha \iint_{\mathcal{X} \times \mathcal{Y}} f(\chi, \nu) d\mu(\chi) d\nu(\nu),$$

and

$$S_Y^\beta f(y) = \int_X f(x, y) d\mu(x) - \beta \iint_{\mathcal{X} \times \mathcal{Y}} f(\chi, \nu) d\mu(\chi) d\nu(\nu),$$

By Parseval in $L_2(\mu)$, $L_2(\nu)$, $L_2(\mu \times \nu)$,

$$\begin{aligned} & \|S_X^\alpha f\|_{L_2(\mu)}^2 + \|S_Y^\beta f\|_{L_2(\nu)}^2 \\ &= (|1 - \alpha|^2 + |1 - \beta|^2) |\langle f, \varphi_0 \otimes \psi_0 \rangle_{L_2(\mu \times \nu)}|^2 + \sum_{j=1}^{\infty} \left(|\langle f, \varphi_j \otimes \psi_0 \rangle_{L_2(\mu \times \nu)}|^2 + |\langle f, \varphi_0 \otimes \psi_j \rangle_{L_2(\mu \times \nu)}|^2 \right) \\ &\stackrel{(36)}{\leq} \max\{|1 - \alpha|^2 + |1 - \beta|^2, 1\} \|f\|_{L_2(\mu \times \nu)}^2. \end{aligned}$$

The case $\alpha = \beta = \frac{1}{2}$ of this inequality is precisely (34). It is worthwhile to note in passing that this reasoning (substituted into the above interpolation argument) yields the following generalization of (18).

$$\begin{aligned} \max \left\{ (|1 - \alpha|^2 + |1 - \beta|^2)^{p\theta}, 1 \right\} & \left((1 + |\alpha|)^{\frac{2p(1-\theta)}{2-\theta p}} + (1 + |\beta|)^{\frac{2p(1-\theta)}{2-\theta p}} \right)^{1 - \frac{\theta p}{2}} \iint_{\mathcal{X} \times \mathcal{Y}} \|f(x, y)\|_{[F, H]_\theta}^p d\mu(x) d\nu(y) \\ & \geq \int_{\mathcal{X}} \left\| \int_{\mathcal{Y}} f(x, y) d\nu(y) - \alpha \iint_{\mathcal{X} \times \mathcal{Y}} f(\chi, \nu) d\mu(\chi) d\nu(\nu) \right\|_{[F, H]_\theta}^p d\mu(x) \\ & \quad + \int_{\mathcal{Y}} \left\| \int_{\mathcal{X}} f(x, y) d\mu(x) - \beta \iint_{\mathcal{X} \times \mathcal{Y}} f(\chi, \nu) d\mu(\chi) d\nu(\nu) \right\|_{[F, H]_\theta}^p d\nu(y). \end{aligned} \quad (37)$$

For the justification of the remaining inequality (35), define $Tg \in L_2(\mu)$ by

$$Tg \stackrel{\text{def}}{=} \sum_{j=1}^{\infty} (\langle g, \varphi_0 \otimes \varphi_j \rangle_{L_2(\mu \times \mu)} - \langle g, \varphi_j \otimes \varphi_0 \rangle_{L_2(\mu \times \mu)}) \varphi_j.$$

In other words, μ -almost surely $Tg(\chi) = \int_{\mathcal{X}} (g(x, \chi) - g(\chi, x)) d\mu(x)$. By Parseval in $L_2(\mu)$, $L_2(\mu \times \nu)$,

$$\begin{aligned} \|Tg\|_{L_2(\mu)}^2 &= \sum_{j=1}^{\infty} |\langle g, \varphi_0 \otimes \varphi_j \rangle_{L_2(\mu \times \mu)} - \langle g, \varphi_j \otimes \varphi_0 \rangle_{L_2(\mu \times \mu)}|^2 \\ &\leq \sum_{j=1}^{\infty} 2 \left(|\langle g, \varphi_0 \otimes \varphi_j \rangle_{L_2(\mu \times \mu)}|^2 + |\langle g, \varphi_j \otimes \varphi_0 \rangle_{L_2(\mu \times \mu)}|^2 \right) \stackrel{(36)}{\leq} 2 \|g\|_{L_2(\mu \times \mu)}^2, \end{aligned}$$

where in the penultimate step we used the convexity of $(\zeta \in \mathbb{C}) \mapsto |\zeta|^2$. This is precisely (35). \square

We will next deduce Theorem 1.2 from the special case of Theorem 1.6 that we stated as Theorem 1.5.

Proof of Theorem 1.2. The largest $\theta \in [0, 1]$ for which $\frac{2}{2-\theta} \leq p \leq \frac{2}{\theta}$ and also $\frac{1}{q} = \frac{1-\theta}{r} + \frac{\theta}{2}$ for some $r \geq 1$ is

$$\theta_{\max} = \theta_{\max}(p, q) \stackrel{\text{def}}{=} 2 \min \left\{ \frac{1}{p}, 1 - \frac{1}{p}, \frac{1}{q}, 1 - \frac{1}{q} \right\}.$$

We then have $L_q = [L_r, L_2]_{\theta_{\max}}$. Note that the quantity $c(p, q)$ that is defined in (11) is equal to $\frac{p}{2} \theta_{\max}$.

By (15) with $\theta = \theta_{\max}$ and $F = L_r$ we have $j_p(L_q) \geq 2^{c(p, q)}$. The matching upper bound $j_p(L_q) \leq 2^{c(p, q)}$ holds due to the following quick examples. If X is uniformly distributed on $\{-1, 1\}$, then $\mathbb{E}[|X - \mathbb{E}[X]|^p]$ and $\mathbb{E}|X - X'|^p = 2^{p-1}$. So, $j_p(\mathbb{R}) \leq 2^{p-1}$. If $\varepsilon \in (0, 1)$ and $\mathbb{P}[X_\varepsilon = 0] = 1 - \varepsilon$ and $\mathbb{P}[X_\varepsilon = 1] = \varepsilon$, then for $p > 1$,

$$j_p(\mathbb{R}) \leq \frac{\mathbb{E} \left[\|X_\varepsilon - X'_\varepsilon\|_q^p \right]}{\mathbb{E} \left[\|X_\varepsilon - \mathbb{E}[X_\varepsilon]\|_q^p \right]} = \frac{2\varepsilon(1-\varepsilon)}{(1-\varepsilon)\varepsilon^p + \varepsilon(1-\varepsilon)^p} \xrightarrow{\varepsilon \rightarrow 0^+} 2.$$

If $n \in \mathbb{N}$ and X_n is uniformly distributed over $\{\pm e_1, \dots, \pm e_n\}$, where $\{e_j\}_{j=1}^\infty$ is the standard basis of ℓ_p , then

$$j_p(L_q) \leq j_p(\ell_q^n) \leq \frac{\mathbb{E} \left[\|X_n - X'_n\|_q^p \right]}{\mathbb{E} \left[\|X_n - \mathbb{E}[X_n]\|_q^p \right]} = \frac{n-1}{n} 2^{\frac{p}{q}} + \frac{1}{2n} 2^p \xrightarrow{n \rightarrow \infty} 2^{\frac{p}{q}}.$$

If r_1, \dots, r_n are i.i.d. symmetric Bernoulli random variables viewed as elements of L_q , e.g. they can be the coordinate functions in $L_q(\{-1, 1\}^n)$, then let R_n be uniformly distributed over $\{\pm r_1, \dots, \pm r_n\}$. Then,

$$j_p(L_q) \leq \frac{\mathbb{E} \left[\|R_n - R'_n\|_q^p \right]}{\mathbb{E} \left[\|R_n - \mathbb{E}[R_n]\|_q^p \right]} = \frac{n-1}{n} 2^{\frac{p(q-1)}{q}} + \frac{1}{2n} 2^p \xrightarrow{n \rightarrow \infty} 2^{\frac{p(q-1)}{q}}.$$

This completes the proof that $j_p(L_q) = 2^{c(p, q)}$.

Next, an application of (15) with $\theta = \theta_{\max}$ and $F = L_r$ gives $\tau_p(L_q) \leq 2^{1+(1-\theta_{\max})p}$. In other words,

$$\mathbb{E} \left[\|X - X'\|_q^p \right] + \mathbb{E} \left[\|Y - Y'\|_q^p \right] \leq 2^{\max\{p-1, 3-p, 1+\frac{p(q-2)}{q}, 1+\frac{p(2-q)}{q}\}} \mathbb{E} \left[\|X - Y\|_q^p \right], \quad (38)$$

for every p -integrable independent L_q -valued random variables X, X', Y, Y' such that (X, Y) and (X', Y') are identically distributed. The bound (38) coincides with (15), where $C(p, q)$ is as in (12), only in the first two ranges that appear in (12), namely when $\frac{p}{p-1} \leq q \leq p$ or when $\frac{q}{q-1} \leq p \leq q$. For the remaining ranges that appear in (12), the bound (38) is inferior to (15), so we reason as follows.

For every $q, Q \in [1, \infty]$ satisfying $Q \geq q$, by [16, Remark 5.10] (the case $Q \in [1, 2]$ is an older result [6]) there exists an embedding $\mathfrak{s} = \mathfrak{s}_{q,Q} : L_q \rightarrow L_Q$ (given by an explicit formula) such that

$$\forall x, y \in L_q, \quad \|\mathfrak{s}(x) - \mathfrak{s}(y)\|_Q = \|x - y\|_q^{\frac{q}{Q}}. \quad (39)$$

Apply (38) to the L_Q -valued random vectors $\mathfrak{s}(X), \mathfrak{s}(X'), \mathfrak{s}(Y), \mathfrak{s}(Y')$ with q replaced by Q and p replaced with $\frac{pQ}{q}$. The resulting estimate is

$$\begin{aligned} \mathbb{E} \left[\|X - X'\|_q^p \right] + \mathbb{E} \left[\|Y - Y'\|_q^p \right] &\stackrel{(39)}{=} \mathbb{E} \left[\|\mathfrak{s}(X) - \mathfrak{s}(X')\|_Q^{\frac{pQ}{q}} \right] + \mathbb{E} \left[\|\mathfrak{s}(Y) - \mathfrak{s}(Y')\|_Q^{\frac{pQ}{q}} \right] \\ &\stackrel{(38)}{\leq} 2^{\max\{\frac{pQ}{q}-1, 3-\frac{pQ}{q}, 1+\frac{p(Q-2)}{q}, 1+\frac{p(2-Q)}{q}\}} \mathbb{E} \left[\|\mathfrak{s}(X) - \mathfrak{s}(Y)\|_Q^{\frac{pQ}{q}} \right] \\ &\stackrel{(39)}{=} 2^{\max\{\frac{pQ}{q}-1, 3-\frac{pQ}{q}, 1+\frac{p(Q-2)}{q}, 1+\frac{p(2-Q)}{q}\}} \mathbb{E} \left[\|X - Y\|_q^p \right]. \end{aligned} \quad (40)$$

It is in our interest to choose $Q \geq q$ so as to minimize the right hand side of (40). If $\frac{1}{p} + \frac{1}{q} \leq 1$, then $Q = q$ is the optimal choice in (40), and therefore we return to (38). But, if $\frac{1}{p} + \frac{1}{q} \geq 1$, then $Q = 1 + \frac{q}{p} \geq q$ is the optimal choice in (40) and we arrive at the following estimate which is better than (38) in the stated range

$$\frac{1}{p} + \frac{1}{q} \geq 1 \implies \mathbb{E} \left[\|X - X'\|_q^p \right] + \mathbb{E} \left[\|Y - Y'\|_q^p \right] \leq 2^{\max\{\frac{p}{q}, 2-\frac{p}{q}\}} \mathbb{E} \left[\|X - Y\|_q^p \right]. \quad (41)$$

The bound (41) covers the third and fourth ranges that appear in (12), as well as the case $p = q \in [1, 2]$ of the fifth range that appears in (12). However, (41) is inferior to (12) when $1 \leq p < q \leq 2$. When this occurs, use the fact [12] that L_q is isometric to a subspace of L_p and apply the already established case $p = q$ to the L_p -valued random variables $i(X), i(X'), i(Y), i(Y')$, where $i : L_q \rightarrow L_p$ is any isometric embedding.

We will next prove that $\tau_p(L_q) \geq 2^{C_{\text{opt}}(p,q)}$, where $C_{\text{opt}}(p, q)$ is given in (14). In particular, this will justify the second sharpness assertion of Theorem 1.2, namely that (38) is sharp when p, q belong to the first, second or fifth ranges that appear in (12). Firstly, by considering the special case of (7) in which X, Y are i.i.d., we see that $\tau_p(F) \geq 1$ for any Banach space F . Next, fix $n \in \mathbb{N}$ and let $r_1, \dots, r_n, \rho_1, \dots, \rho_n \in L_q$ be such that r_1, \dots, r_n and ρ_1, \dots, ρ_n each form a sequence of i.i.d. symmetric Bernoulli random variables, and the supports of r_1, \dots, r_n are disjoint from the supports of ρ_1, \dots, ρ_n . For example, one could consider them as the elements of $L_q(\{-1, 1\}^n) \oplus_q L_q(\{-1, 1\}^n)$ that are given by $r_i = (\omega \mapsto \omega_i, 0)$ and $\rho_i = (0, \omega \mapsto \omega_i)$ for each $i \in \{1, \dots, n\}$. Let X be uniformly distributed over $\{r_1, \dots, r_n\}$ and Y be uniformly distributed over $\{\rho_1, \dots, \rho_n\}$. Due to the disjointness of the supports, we have $\|X - Y\|_q^p = (\|X\|_q^q + \|Y\|_q^q)^{p/q} = 2^{p/q}$ pointwise. At the same time, $\mathbb{E}[\|X - X'\|_q^p] + \mathbb{E}[\|Y - Y'\|_q^p] = 2(1 - 1/n)(2^q/2)^{p/q} = (1 - 1/n)2^{1+p(q-1)/q}$. By letting $n \rightarrow \infty$, this shows that necessarily $\tau_p(L_q) \leq 2^{1+p(q-2)/q}$. Finally, if (7) holds, then in particular it holds for scalar-valued random variables. By integrating, we see that $\tau_p(F) \geq \tau_p(L_p)$ for any Banach space F . But, the case $p = q$ of the above discussion gives $\tau_p(L_p) \geq 2^{1+p(p-2)/p} = 2^{p-1}$, as required.

The bound (13) of Theorem 1.2 coincides with (16). When $p \leq q \leq 2$, we have $C(p, q) = 1$, $c(p, q) = p-1$ and thus $m_p(L_q) \leq 2^{2-p}$. It therefore remains to check that $\mathfrak{b}_p(L_q) \geq 2^{2-p}$ when $p \leq q \leq 2$. In fact, $\mathfrak{b}_p(F) \geq 2^{2-p}$ for every $p \geq 1$ and every Banach space $(F, \|\cdot\|_F)$. Indeed, fix distinct $a, b \in F$. Let X, Y be

independent and uniformly distributed over $\{a, b\}$. Then

$$\mathfrak{b}_p(F) \geq \frac{\mathbb{E}[\|X - z\|_F^p + \|Y - z\|_F^p]}{\mathbb{E}[\|X - Y\|_F^p]} = \frac{\|a - z\|_F^p + \|b - z\|_F^p}{\frac{1}{2}\|a - b\|_F^p} \geq \frac{2\|\frac{1}{2}(a - z) - \frac{1}{2}(b - z)\|_F^p}{\frac{1}{2}\|a - b\|_F^p} = 2^{2-p},$$

where the penultimate step is an application of the convexity of $\|\cdot\|_F^p$. \square

Remark 3.1. Fix $n \in \mathbb{N}$. Following [7], for $a = (a_1, \dots, a_{2n}) \in \mathbb{C}^{2n}$ denote by $\Re(a) = (\Re(a_1), \dots, \Re(a_{2n})) \in \mathbb{R}^{2n}$ and $\Im(a) = (\Im(a_1), \dots, \Im(a_{2n})) \in \mathbb{R}^{2n}$ the vectors of real parts and imaginary parts of the entries of a , respectively. Let $\Lambda(a) \in [0, \infty)$ be the area of the parallelogram that is generated by $\Re(a)$ and $\Im(a)$, i.e.,

$$\Lambda(a) \stackrel{\text{def}}{=} \sqrt{\|\Re(a)\|_2^2 \|\Im(a)\|_2^2 - \langle \Re(a), \Im(a) \rangle}.$$

By [7, Lemma 5.2] there is a linear operator $\mathcal{C} : \mathbb{C}^{2n} \rightarrow \mathbb{M}_{2n}(\mathbb{C})$ from \mathbb{C}^{2n} to the space of 2^n by 2^n complex matrices, such that for any $a \in \mathbb{C}^{2n}$ the Schatten-1 norm of the matrix $\mathcal{C}(a)$ satisfies

$$\|\mathcal{C}(a)\|_{S_1} = \frac{1}{2}\sqrt{\|a\|_2^2 + 2\Lambda(a)} + \frac{1}{2}\sqrt{\|a\|_2^2 - 2\Lambda(a)}. \quad (42)$$

Let $e_1, \dots, e_{2n} \in \mathbb{C}^{2n}$ be the standard basis of \mathbb{C}^{2n} and define $2n$ matrices $x_1, \dots, x_n, y_1, \dots, y_n \in \mathbb{M}_{2n}(\mathbb{C})$ by $x_k = \mathcal{C}(e_k)$ and $y_k = \mathcal{C}(ie_{n+k})$ for $k \in \{1, \dots, n\}$. By (42) we have $\|x_j - x_k\|_{S_1} = \|y_j - y_k\|_{S_1} = \sqrt{2}$ for distinct $j, k \in \{1, \dots, n\}$, while $\|x_j - y_k\|_{S_1} = 1$ for all $j, k \in \{1, \dots, n\}$. Hence, if we let X and Y be independent and distributed uniformly over $\{x_1, \dots, x_n\}$ and $\{y_1, \dots, y_n\}$, respectively, and X', Y' are independent copies of X, Y , respectively, then for every $p \geq 1$ we have

$$\mathbb{E}[\|X - Y\|_{S_1}^p] = 1 \quad \text{and} \quad \mathbb{E}[\|X - X'\|_{S_1}^p] = \mathbb{E}[\|Y - Y'\|_{S_1}^p] = \frac{n-1}{n} 2^{\frac{p}{2}}.$$

By letting $n \rightarrow \infty$, this implies that $\tau_p(S_1) \geq 2^{\frac{p}{2}+1}$. In particular, $\tau_1(S_1) \geq 2\sqrt{2}$.

Remark 3.2. Fix $q \geq 1$. Let $(F, \|\cdot\|_F)$ be a Banach space. Assume that F has a linear subspace $G \subseteq F$ that is isometric to L_q (or the Schatten-von Neumann trace class S_q). If $X, Y \in L_p(G)$ are i.i.d. random variables taking values in G , then for $c(p, q)$ as in (11), by Theorem 1.2 we have

$$\mathbb{E}[\|X - Y\|_F^p] \geq 2^{c(p, q)} \cdot \inf_{z \in F} \mathbb{E}[\|X - z\|_F^p]. \quad (43)$$

We note that this inequality is optimal despite the fact that the infimum is now taken over z in the larger super-space F . Indeed, in the proof of Theorem 1.2 the random variables that established optimality of $c(p, q)$ were symmetric when p, q belong to the first three ranges that appear in (11). In these cases, by the convexity of $\|\cdot\|_F^p$, the infimum in the right hand side of (43) is attained at $z = 0 \in G$. The fact that the term $2^{c(p, q)}$ in the right hand side of (43) cannot be replaced by any value greater than 2 needs the following separate treatment. If $\varepsilon \in (0, 1)$ and $\mathbb{P}[X = \nu] = \varepsilon = 1 - \mathbb{P}[X = 0]$ for some $\nu \in G$ with $\|\nu\|_F = 1$, then $\mathbb{E}[\|X - Y\|_F^p] = 2\varepsilon(1 - \varepsilon)$. Next, for any $z \in F$ we have

$$\begin{aligned} \mathbb{E}[\|X - z\|_F^p] &= (1 - \varepsilon)\|z\|_F^p + \varepsilon\|\nu - z\|_F^p \geq (1 - \varepsilon)\|z\|_F^p + \varepsilon(\max\{0, 1 - \|z\|_F\})^p \\ &\geq \min_{r \geq 0} ((1 - \varepsilon)r^p + \varepsilon(\max\{0, 1 - r\})^p) = \frac{\varepsilon(1 - \varepsilon)}{\left(\varepsilon^{\frac{1}{p-1}} + (1 - \varepsilon)^{\frac{1}{p-1}}\right)^{p-1}}, \end{aligned}$$

where the final step follows by elementary calculus. Therefore,

$$\frac{\mathbb{E}[\|X - Y\|_F^p]}{\inf_{z \in F} \mathbb{E}[\|X - z\|_F^p]} \leq 2 \left(\varepsilon^{\frac{1}{p-1}} + (1 - \varepsilon)^{\frac{1}{p-1}}\right)^{p-1} \xrightarrow{\varepsilon \rightarrow 0^+} 2.$$

Remark 3.3. An extrapolation theorem of Pisier [18] asserts that if $(F, \|\cdot\|_F)$ is a Banach lattice that is both p -convex with constant 1 and q -concave with constant 1, where $\frac{1}{p} + \frac{1}{q} = 1$, then there exists a Banach lattice W , a Hilbert space H , and $\theta \in (0, 1]$ such that F is isometric to the complex interpolation space

$[W, H]_\theta$. Hence, Theorem 1.5 applies in this setting, implying in particular that there is $r \in [1, \infty)$, namely $r = \frac{2}{\theta}$, such that every i.i.d. F -valued random variables $X, Y \in L_r(F)$ satisfy

$$\mathbb{E}[\|X - Y\|_F^r] \geq 2\mathbb{E}[\|X - \mathbb{E}[X]\|_F^r].$$

We will conclude by discussing further bounds in the non-convex range $p < 1$, as well as their limit when $p \rightarrow 0^+$. When $p \in (0, 1)$, the topological vector space L_p is not a normed space. Despite this, when we say that a normed space $(F, \|\cdot\|_F)$ admits a linear isometric embedding into L_p we mean (as usual) that there exists a linear mapping $T: F \rightarrow L_p$ such that $\|Tx\|_p = \|x\|_F$ for all $x \in F$. This of course forces the L_p quasi-norm to induce a metric on the image of T , so the use of the term “isometric” is not out of place here, though note that it is inconsistent with the standard metric on L_p , which is given by $\|f - g\|_p^p$ for all $f, g \in L_p$. The following proposition treats the case $p \in (0, 2]$, though later we will mainly be interested in the non-convex range $p \in (0, 1)$. Note that the case $p = 1$ implies the stated inequalities for, say, any two-dimensional normed space, since any such space admits [5] an isometric embedding into L_1 .

Proposition 3.4. *Let $(F, \|\cdot\|_F)$ be a Banach space that admits an isometric linear embedding into L_p for some $p \in (0, 2]$. Let $X, X', Y, Y' \in L_p(F)$ be independent F -valued random vectors such that X' has the same distribution as X and Y' has the same distribution as Y . Then,*

$$\mathbb{E}[\|X - X'\|_F^p] + \mathbb{E}[\|Y - Y'\|_F^p] \leq 2\mathbb{E}[\|X - Y\|_F^p], \quad (44)$$

and

$$\inf_{z \in F} \mathbb{E}[\|X - z\|_F^p + \mathbb{E}\|Y - z\|_F^p] \leq \min\{2, 2^{2-p}\} \mathbb{E}[\|X - Y\|_F^p]. \quad (45)$$

The constants 2 and $\min\{2, 2^{2-p}\}$ in (44) and (45), respectively, cannot be improved.

Proof. By [19, 6] there is a mapping $\mathfrak{s}: F \rightarrow L_2$ such that $\|\mathfrak{s}(x) - \mathfrak{s}(y)\|_2 = \|x - y\|_F^{\frac{p}{2}}$ for all $x, y \in F$. By the (trivial) Hilbertian case $p = q = 2$ of Theorem 1.2 applied to the L_2 -valued random vectors $\mathfrak{s}(X), \mathfrak{s}(Y)$,

$$\begin{aligned} \mathbb{E}[\|X - X'\|_F^p] + \mathbb{E}[\|Y - Y'\|_F^p] &= \mathbb{E}[\|\mathfrak{s}(X) - \mathfrak{s}(X')\|_2^2] + \mathbb{E}[\|\mathfrak{s}(Y) - \mathfrak{s}(Y')\|_2^2] \\ &\leq 2\mathbb{E}[\|\mathfrak{s}(X) - \mathfrak{s}(Y)\|_2^2] = 2\mathbb{E}[\|X - Y\|_F^p]. \end{aligned}$$

This substantiates (44). When $p < 1$ we cannot proceed from here to prove (45) by considering the analogue of the mixture constant $m(\cdot)$, namely by bounding the left hand side of (5) as we did in the Introduction, since the present L_p integrability assumption on X, Y does not imply that $\mathbb{E}[X]$ and $\mathbb{E}[Y]$ are well-defined elements of F . Instead, let Z' be independent of X, Y and distributed according to the mixture of the laws of X and Y , as in (6). The point $z \in F$ will be chosen randomly according to Z' , i.e.,

$$\begin{aligned} \inf_{z \in F} \mathbb{E}[\|X - z\|_F^p + \mathbb{E}\|Y - z\|_F^p] &\leq \mathbb{E}[\|X - Z'\|_F^p + \mathbb{E}\|Y - Z'\|_F^p] \\ &= \frac{1}{2} \mathbb{E}[\|X - X'\|_F^p + \mathbb{E}\|Y - Y'\|_F^p] + \mathbb{E}[\|X - Y\|_F^p] \stackrel{(44)}{\leq} 2\mathbb{E}[\|X - Y\|_F^p]. \end{aligned} \quad (46)$$

For $p \geq 1$ we have $\tau_p(F) \leq 2$ by (44), and $j_p(F) \geq j_p(L_p) = p - 1$ by Theorem 1.2, so $\mathfrak{b}_p(F) \leq 2^{2-p}$, by (10).

The sharpness of (44) is seen by taking X and Y to be identically distributed. When $p \geq 1$, we already saw in the proof of Theorem 1.2 that $\mathfrak{b}_p(F) \geq 2^{2-p}$ for any Banach space F ; thus (45) is sharp in this range. The same reasoning as in the proof of Theorem 1.2 shows that the factor 2 in (45) cannot be improved in the non-convex range $p \in (0, 1)$ as well. Indeed, fix v with $\|v\|_F = 1$ and let X and Y be uniformly distributed over $\{0, v\}$. Then, $\mathbb{E}[\|X - z\|_F^p + \mathbb{E}\|Y - z\|_F^p] = \|z\|_F^p + \|v - z\|_F^p \geq (\|z\|_F + \|v - z\|_F)^p \geq \|v\|_F^p = 1$ for every $z \in F$, while $\mathbb{E}[\|X - Y\|_F^p] = \frac{1}{2} \|v\|_F^p = \frac{1}{2}$. \square

Proposition 3.5 below is the limit of Proposition 3.4 as $p \rightarrow 0^+$. While it is possible to deduce it formally from Proposition 3.4 by passing to the limit, a justification of this fact is quite complicated due to the singularity of the logarithm at zero. We will instead proceed via a shorter alternative approach.

Following [13], a real Banach space $(F, \|\cdot\|_F)$ is said to admit a linear isometric embedding into L_0 if there exists a probability space (Ω, μ) and a linear operator $T : F \rightarrow \text{Meas}(\Omega, \mu)$, where $\text{Meas}(\Omega, \mu)$ denotes the space of (equivalence classes of) real-valued μ -measurable functions on Ω , such that

$$\forall x \in F, \quad \|x\|_F = e^{\int_{\Omega} \log |Tx| d\mu}. \quad (47)$$

As shown in [13], every three-dimensional real normed space admits a linear isometric embedding into L_0 , so in particular the following proposition applies to any such space.

Proposition 3.5. *Let $(F, \|\cdot\|_F)$ be a real Banach space that admits a linear isometric embedding into L_0 . Let X, X', Y, Y' be independent F -valued random vectors such that X' has the same distribution as X and Y' has the same distribution as Y . Assume that $\mathbb{E}[\log(1 + \|X\|_F)] < \infty$ and $\mathbb{E}[\log(1 + \|Y\|_F)] < \infty$. Then,*

$$e^{\mathbb{E}[\log(\|X-X'\|_F \cdot \|Y-Y'\|_F)]} \leq e^{2\mathbb{E}[\log(\|X-Y\|_F)]}, \quad (48)$$

and

$$\inf_{z \in F} e^{\mathbb{E}[\log(\|X-z\|_F \cdot \|Y-z\|_F)]} \leq e^{2\mathbb{E}[\log(\|X-Y\|_F)]}. \quad (49)$$

The multiplicative constant 1 in both of these inequalities is optimal.

Proof. (49) is a consequence of (48) by reasoning analogously to (46). Due to the assumed representation (47), by Fubini's theorem it suffices to prove (48) for real-valued random variables.

So, suppose that X, Y are independent real-valued random variables such that $\mathbb{E}[\log(1 + |X|)] < \infty$ and $\mathbb{E}[\log(1 + |Y|)] < \infty$. Note that every nonnegative random variable W with $\mathbb{E}[\log(1 + W)] < \infty$ satisfies

$$\mathbb{E}[\log W] = \int_0^{\infty} \frac{e^{-s} - \mathbb{E}[e^{-sW}]}{s} ds. \quad (50)$$

Indeed, for every $a, b \in [0, \infty)$ with $a \leq b$ we have

$$\int_0^{\infty} \frac{e^{-as} - e^{-bs}}{s} ds = \int_0^{\infty} \left(\int_a^b s e^{-ts} dt \right) \frac{ds}{s} = \int_a^b \left(\int_0^{\infty} e^{-ts} ds \right) dt = \int_a^b \frac{dt}{t} = \log b - \log a,$$

so that (50) follows by applying this identity and the Fubini theorem separately on each of the events $\{W \geq 1\}$ and $\{W < 1\}$, taking advantage of the fact that $e^{-s} - e^{-sW}$ is of constant sign on both events.

Let Z, Z' be independent random variables whose law is the mixture of the laws of X, Y as in (6). the desired inequality (48) is equivalent to the assertion that $\mathbb{E}[\log(Z - Z')^2] \leq \mathbb{E}[\log(X - Y)^2]$. By two applications of (50), once with $W = (X - Y)^2$ and once with $W = (Z - Z')^2$, it suffices to prove that

$$\forall s \geq 0, \quad \mathbb{E}\left[e^{-s(Z-Z')^2}\right] \geq \mathbb{E}\left[e^{-s(X-Y)^2}\right].$$

This is so because, using the formula for the Fourier transform of the Gaussian density, we have

$$\begin{aligned} \mathbb{E}\left[e^{-s(Z-Z')^2}\right] &= \mathbb{E}\left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{it(Z-Z')\sqrt{2s} - \frac{t^2}{2}} dt\right] \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathbb{E}\left[e^{it\sqrt{2s}Z}\right] \cdot \mathbb{E}\left[e^{-it\sqrt{2s}Z'}\right] e^{-\frac{t^2}{2}} dt \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left|\mathbb{E}\left[e^{it\sqrt{2s}Z}\right]\right|^2 e^{-\frac{t^2}{2}} dt \end{aligned} \quad (51)$$

$$\begin{aligned} &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left|\frac{1}{2}\mathbb{E}\left[e^{it\sqrt{2s}X}\right] + \frac{1}{2}\mathbb{E}\left[e^{it\sqrt{2s}Y}\right]\right|^2 e^{-\frac{t^2}{2}} dt \\ &\geq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Re\left(\mathbb{E}\left[e^{it\sqrt{2s}X}\right] \cdot \overline{\mathbb{E}\left[e^{it\sqrt{2s}Y}\right]}\right) e^{-\frac{t^2}{2}} dt \end{aligned} \quad (52)$$

$$= \frac{1}{\sqrt{2\pi}} \Re\left(\int_{-\infty}^{\infty} \mathbb{E}\left[e^{it\sqrt{2s}(X-Y)}\right] e^{-\frac{t^2}{2}} dt\right) = \mathbb{E}\left[e^{-s(X-Y)^2}\right], \quad (53)$$

where (51) uses Fubini and the independence of Z and Z' , (52) uses the fact that for all $a, b \in \mathbb{C}$ we have $|(a+b)/2|^2 = |(a-b)/2|^2 + \Re(a\bar{b}) \geq \Re(a\bar{b})$, the first step of (53) uses the independence of X and Y , and the last step of (53) uses once more the formula for the Fourier transform of the Gaussian density.

The fact that (48) is sharp follows by considering the case when X, Y are i.i.d. and non-atomic. Note that when both X and Y have an atom at the same point, both sides of (49) equal 0. The example considered in the proof of Proposition 3.4 when $p > 0$ is therefore of no use for establishing the optimality of (49), due to the atomic nature of the distributions under consideration. Instead, for an arbitrary $v \in F$ such that $\|v\|_F = 1$, let us consider random vectors $X = (\cos\Theta)v$ and $Y = (\cos\Theta')v$, where Θ and Θ' are independent random variables uniformly distributed on $[0, 2\pi]$.

Observe that for every $\alpha \in \mathbb{R}$ we have

$$\begin{aligned} \mathbb{E}[\log|\cos\Theta - \cos\alpha|] &= \mathbb{E}\left[\log\left|2\sin\left(\frac{\Theta+\alpha}{2}\right)\sin\left(\frac{\Theta-\alpha}{2}\right)\right|\right] \\ &= \log 2 + \mathbb{E}\left[\log\left|\sin\left(\frac{\Theta+\alpha}{2}\right)\right|\right] + \mathbb{E}\left[\log\left|\sin\left(\frac{\Theta-\alpha}{2}\right)\right|\right] = \log 2 + 2\mathbb{E}[\log|\cos\Theta|], \end{aligned} \quad (54)$$

where the last step of (54) holds because, by periodicity, $|\sin(\frac{\Theta\pm\alpha}{2})|$ has the same distribution as $|\cos\Theta|$.

The case $\alpha = \frac{\pi}{2}$ of (54) simplifies to give $\mathbb{E}[\log|\cos\Theta|] = -\log 2$. Hence, (54) becomes

$$\forall \alpha \in \mathbb{R}, \quad \mathbb{E}[\log|\cos\Theta - \cos\alpha|] = -\log 2. \quad (55)$$

Consequently,

$$\forall t \in \mathbb{R}, \quad \mathbb{E}[\log|\cos\Theta - t|] \geq -\log 2. \quad (56)$$

Indeed, if $t \in [-1, 1]$, then one can write $t = \cos\alpha$ for some $\alpha \in \mathbb{R}$, so that by (55) the inequality in (56) holds as equality. If $|t| > 1$, then $|\cos\theta - t| \geq |\cos\theta - \text{sign}(t)|$ for all $\theta \in [0, 2\pi]$, thus implying (56). It also follows from (55) that

$$\mathbb{E}[\log(\|X - Y\|_F)] = \mathbb{E}[\log|\cos\Theta - \cos\Theta'|] \stackrel{(55)}{=} -\log 2.$$

Next, by the Hahn–Banach theorem, take $\varphi \in F^*$ such that $\|\varphi\|_{F^*} = 1$ and $\varphi(v) = \|v\|_F = 1$. For any $z \in F$,

$$\mathbb{E}[\log(\|X - z\|_F)] = \mathbb{E}[\log(\|Y - z\|_F)] \geq \mathbb{E}[\log|\varphi((\cos\Theta)v - z)|] = \mathbb{E}[\log|\cos\Theta - \varphi(z)|] \stackrel{(56)}{\geq} -\log 2.$$

This implies the asserted sharpness of (49). Note that the above argument that (49) cannot hold with a multiplicative constant less than 1 in the right hand side worked for any Banach space F whatsoever. \square

Acknowledgements. We are grateful to Oded Regev for pointing us to [7, Lemma 5.2] and for significantly simplifying our initial reasoning for the statement that is proved in Remark 3.1.

REFERENCES

- [1] A. Andoni, A. Naor, and O. Neiman. Snowflake universality of Wasserstein spaces. *Ann. Sci. Éc. Norm. Supér. (4)*, 51(3):657–700, 2018.
- [2] Y. Bartal, N. Linial, M. Mendel, and A. Naor. Some low distortion metric Ramsey problems. *Discrete Comput. Geom.*, 33(1):27–41, 2005.
- [3] 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.
- [4] J. Bergh and J. Löfström. *Interpolation spaces. An introduction*. Springer-Verlag, Berlin-New York, 1976. Grundlehren der Mathematischen Wissenschaften, No. 223.
- [5] E. D. Bolker. A class of convex bodies. *Trans. Amer. Math. Soc.*, 145:323–345, 1969.
- [6] J. Bretagnolle, D. Dacunha-Castelle, and J.-L. Krivine. Lois stables et espaces L^p . *Ann. Inst. H. Poincaré Sect. B (N.S.)*, 2:231–259, 1965/1966.
- [7] J. Briët, O. Regev, and R. Saker. Tight hardness of the non-commutative Grothendieck problem. *Theory Comput.*, 13:Paper No. 15, 24, 2017.
- [8] A.-P. Calderón. Intermediate spaces and interpolation, the complex method. *Studia Math.*, 24:113–190, 1964.
- [9] A. N. Dranishnikov, G. Gong, V. Lafforgue, and G. Yu. Uniform embeddings into Hilbert space and a question of Gromov. *Canad. Math. Bull.*, 45(1):60–70, 2002.

- [10] P. Enflo. On the nonexistence of uniform homeomorphisms between L_p -spaces. *Ark. Mat.*, 8:103–105, 1969.
- [11] A. Eskenazis, M. Mendel, and A. Naor. Nonpositive curvature is not coarsely universal. Preprint, available at <https://arxiv.org/abs/1808.02179>, 2018.
- [12] M. I. Kadec'. Linear dimension of the spaces L_p and l_q . *Uspehi Mat. Nauk*, 13(6 (84)):95–98, 1958.
- [13] N. Kalton, A. Koldobsky, V. Yaskin, and M. Yaskina. The geometry of L_0 . *Canadian J. Math.*, 59:1029–1049, 2007.
- [14] C. J. Lennard, A. M. Tonge, and A. Weston. Generalized roundness and negative type. *Michigan Math. J.*, 44(1):37–45, 1997.
- [15] 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.
- [16] M. Mendel and A. Naor. Euclidean quotients of finite metric spaces. *Adv. Math.*, 189(2):451–494, 2004.
- [17] A. Naor. A phase transition phenomenon between the isometric and isomorphic extension problems for Hölder functions between L_p spaces. *Mathematika*, 48(1-2):253–271 (2003), 2001.
- [18] G. Pisier. La méthode d'interpolation complexe: applications aux treillis de Banach. In *Séminaire d'Analyse Fonctionnelle (1978–1979)*, pages Exp. No. 17, 18. École Polytech., Palaiseau, 1979.
- [19] I. J. Schoenberg. Metric spaces and positive definite functions. *Trans. Amer. Math. Soc.*, 44(3):522–536, 1938.
- [20] B. Simon. *Trace ideals and their applications*, volume 35 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge-New York, 1979.
- [21] L. Williams, J. H. Wells, and T. L. Hayden. On the extension of Lipschitz-Hölder maps on L^P spaces. *Studia Math.*, 39:29–38, 1971.

ASSAF NAOR

Department of Mathematics, Princeton University, Fine Hall Washington Road, Princeton NJ, USA;
e-mail: naor@math.princeton.edu

KRZYSZTOF OLESZKIEWICZ

Institute of Mathematics, University of Warsaw, ul. Banacha 2, Warszawa, Poland;
Institute of Mathematics, Polish Academy of Sciences, ul. Śniadeckich 8, Warszawa, Poland;
e-mail: koles@mimuw.edu.pl