EUCLIDEAN EMBEDDING, RANDOMIZED CLUSTERING, AND LIPSCHITZ EXTENSION FOR FINITE AND DOUBLING SUBSETS OF L_p WHEN $p>2$

ASSAF NAOR AND KEVIN REN

ABSTRACT. Fix $p>2$. We prove that the Euclidean distortion of every n-point subset of L_p is $p^3(\log n)^{\frac12+o(1)}$, thus, in particular, demonstrating that all n-point subsets of L_p exhibit an asymptotic improvement over the $O(\log n)$ Euclidean distortion guarantee that Bourgain's embedding theorem provides for arbitrary n-point metric spaces. We also prove that the separation modulus of every n-point subset of L_p is $O(p^2\sqrt{\log n})$, which is sharp up to the dependence on p. We deduce from (a refinement of) this asymptotic evaluation of the finitary separation modulus of L_p that for any n-point subset $\mathcal{C}$ of L_p, any Banach space $\mathbf{Z}$, and any 1-Lipschitz function $f:\mathcal{C}\to\mathbf{Z}$, there exists a $O(p^2\sqrt{\log n})$-Lipschitz function $F:L_p\to\mathbf{Z}$ that extends f. We obtain analogous separation and extension statements for doubling subsets of L_p.

CONTENTS

A. N. was supported by NSF grant DMS-2453936, BSF grant 2018223, and a Simons Investigator award. K. R. was supported by an NSF GRFP fellowship.

1. INTRODUCTION

Prior to passing to detailed descriptions of concepts, results, methods and history, we open with the following quick list that describes in broad strokes the main three outcomes of the present article which are resolutions of longstanding open problems in metric embedding theory:

(1) For every $2<p<\infty$, every n-point subset of L_p is shown to embed into L_2 with bi-Lipschitz distortion that grows to ∞ as $n\to\infty$ asymptotically slower $\log n$. Whether or not this holds has been a well-known open problem ever since the 1985 work [16] proved that any n-point metric space whatsoever embeds into a Hilbert space with distortion $O(\log n)$. In other words, it was unknown (and resolved herein) if—in terms of their bi-Lipschitz embeddability into L_2— anything can be said about finite subsets of L_p beyond the fact that they are metric spaces.

(2) For every $2<p<\infty$, any n-point subset $\mathcal{C}$ of L_p, any Banach space $\mathbf{Z}$, and any 1-Lipschitz function $f:\mathcal{C}\to\mathbf{Z}$, there is an L-Lipschitz function $F:L_p\to\mathbf{Z}$ whose restriction to $\mathcal{C}$ coincides with f, where L grows to ∞ as $n\to\infty$ asymptotically slower than $(\log n)/\log\log n$. Whether or not this holds has been open ever since the 2004 work [52] proved that this extension statement holds with $L=O((\log n)/\log\log n)$ and with L_p replaced by any metric space whatsoever.

(3) For every $2<p<\infty$, answering a question that was posed in the 2017 work [77], it is proved that the largest possible separation modulus of an n-point subset of L_p is bounded from above and from below by positive multiples (which may depend only on p) of $\sqrt{\log n}$. This yields asymptotically optimal randomized clustering of finite subsets of L_p, where the quality of the clustering is measured by the ratio between the probability that it separates points and their distance, which is an influential method that was introduced [12] in the mid-1990s in the computer science literature and has since substantially impacted both algorithm design and pure mathematics.

In addition to the above listed answers to known questions, we obtain improved Lipschitz extension and randomized clustering results for doubling subsets of L_p when $2\leqslant p<\infty$, which are new even when $p=2$ but (to the best of our knowledge) they have not been previously posed as open problems. Specifically, setting $C_p=p^2\sqrt{\log p}$, we prove that if $\mathcal{D}\subseteq L_p$ is λ-doubling for some $\lambda\geqslant 2$, then for any Banach space $\mathbf{Z}$ and any 1-Lipschitz function $f:\mathcal{D}\to\mathbf{Z}$, there exists a $O(C_p\sqrt{\log\lambda})$-Lipschitz function $F:L_p\to\mathbf{Z}$ that extends f, and furthermore the separation modulus of $\mathcal{D}$ is $O(C_p\sqrt{\log\lambda})$.

All of the aforementioned results are proved through an induction on scales and localization argument that we develop herein, in combination with a novel property of the Mazur map that we introduce.

We will next turn to a more technical (but entirely self-contained) description of the above statements:

1.1. **Bi-Lipschitz embeddings.** The Euclidean distortion of a finite metric space $(\mathcal{M},d_{\mathcal{M}})$, which is commonly denoted $\mathsf{c}_2(\mathcal{M})$ following [58], is the smallest $D\geqslant 0$ such that there exists $f:\mathcal{M}\to L_2$ satisfying

$$\forall x,y\in\mathcal{M},\qquad d_{\mathcal{M}}(x,y)\leqslant\|f(x)-f(y)\|_{L_2}\leqslant Dd_{\mathcal{M}}(x,y).$$

The Euclidean distortion growth $\{\mathsf{c}_2^n(\mathcal{M})\}_{n=1}^\infty$ of an infinite metric space $(\mathcal{M},d_{\mathcal{M}})$ is defined by

$$\forall n\in\mathbb{N},\qquad \mathsf{c}_2^n(\mathcal{M})\stackrel{\mathrm{def}}{=}\sup_{\substack{\mathcal{C}\subseteq\mathcal{M}\\|\mathcal{C}|\leqslant n}}\mathsf{c}_2(\mathcal{C}). \tag{1}$$

One says that $(\mathcal{M},d_{\mathcal{M}})$ has nontrivial Euclidean distortion growth if

$$\lim_{n\to\infty}\frac{\mathsf{c}_2^n(\mathcal{M})}{\log n}=0. \tag{2}$$

The above use of the word "nontrivial" arises from the Bourgain embedding theorem [16], which asserts that $\mathsf{c}_2^n(\mathcal{M})=O(\log n)$ for *every* metric space $(\mathcal{M},d_{\mathcal{M}})$. A well-known open question (which has been prominently on researchers' minds ever since [16] appeared but to the best of our knowledge was not stated in print) is whether L_p has nontrivial Euclidean distortion growth for every $2 < p < \infty$.[1] In other words, in terms of their bi-Lipschitz embeddability into a Hilbert space, can anything be said about finite subsets of L_p beyond merely that they are metric spaces? Here we prove that the answer is affirmative:

Theorem 1. *If $2 < p < \infty$, then L_p has nontrivial Euclidean distortion growth. More precisely,*[2]

$$\forall n \in \{3,4,\ldots\}, \qquad \mathsf{c}_2^n(L_p) \lesssim p^3\sqrt{\log n}\log\log n. \tag{3}$$

If $1 \leqslant p \leqslant 2$, then $\mathsf{c}_2^n(L_p) \lesssim \sqrt{\log n}$ for every integer $n \geqslant 2$ by [22]. In fact, together with the matching lower bound from [31], it follows from [22] that $\mathsf{c}_2^n(L_1) \asymp \sqrt{\log n}$; at present, L_1 is the only classical Banach space for which the Euclidean distortion growth has been computed up to universal constant factors, other than the (trivial) case $\mathsf{c}_2^n(L_2) = 1$ and the (substantial) case $\mathsf{c}_2^n(L_\infty) \asymp \log n$ (the latter statement is a combination of [16] and [58, 8] for the upper and lower bounds on $\mathsf{c}_2^n(L_\infty)$, respectively).

If $p > 2$, then prior to Theorem 1 it was known that L_p has nontrivial Euclidean distortion growth only when $2 < p < 3\sqrt{e} \in [4.94, 4.95]$. Specifically, it was announced in Section 5.3 of the 2014 arXiv posting [13] that L_p has nontrivial Euclidean distortion growth when $2 < p < 4$. We warn that [13] was replaced by a version in which all discussion of distortion growth was removed (and, to the best of our knowledge what was removed from [13] did not appear elsewhere). We do not know why the authors of [13] chose to remove that material, as it seems to us that the sketch of proof in [13] is sound, albeit with missing steps. In our initial arXiv posting [83], which is superseded by the present article, we succeeded to prove that L_p has nontrivial Euclidean distortion growth only when $2 < p < 4$, while being unaware of [13] (the bound that we obtained in [83] is stronger than what is implied by [13]). The subsequent work [49] used a clever iterative application of our main embedding theorem in [83] to extend to $4 \leqslant p < 3\sqrt{e}$ the range for which L_p has nontrivial Euclidean distortion growth, but that method of [49] breaks down when $p \geqslant 3\sqrt{e}$. From the quantitative perspective, the bound (3) of Theorem 1 is the best-known such bound for all fixed $p > 2$.

Conjecture 2. *The bound* (3) *gives $\mathsf{c}_2^n(L_p) = o(\log n)$ when $p = o\big(\sqrt[6]{\log n}/\sqrt[3]{\log\log n}\big)$. If $p \gtrsim \log n$, then ℓ_∞^n embeds with $O(1)$ distortion into L_p, so $\mathsf{c}_2^n(L_p) \asymp \log n$ as all n-point metric spaces are isometric to a subset of ℓ_∞^n. We conjecture that $\mathsf{c}_2^n(L_p) = o(\log n)$ in the remaining range $\sqrt[6]{\log n}/\sqrt[3]{\log\log n} \lesssim p = o(\log n)$.*

1.1.1. *Toward a sharp Euclidean distortion growth rate for L_p.* Now that it is established that L_p has nontrivial Euclidean distortion growth for every $1 \leqslant p < \infty$, one can turn attention to determining the growth rate of $\mathsf{c}_2^n(L_p)$ as $n \to \infty$, which is currently known only for $p \in \{1,2\}$. This seems to be a highly nontrivial matter. The Lewis theorem [54] asserts that $\mathsf{c}_2(\mathbf{X}) \geqslant k^{|1/2-1/p|}$ for every k-dimensional subspace $\mathbf{X}$ of L_p (this lower bound holds as equality when $\mathbf{X} = \ell_p^k$; see e.g. [95, 43]). In accordance with the longstanding Ribe research program [17, 44, 74, 10, 90, 89, 34, 78]), this naturally leads to the following open question:

Question 3 (Nonlinear Lewis problem)**.** *Is it true that for every $1 \leqslant p \leqslant \infty$ we have*

$$\forall n \in \{2,3,\ldots\}, \qquad \mathsf{c}_2^n(L_p) \asymp_p (\log n)^{\left|\frac12 - \frac1p\right|}. \tag{4}$$

In fact, it is not even known whether (4) *holds for* any *fixed $p \in [1,\infty] \smallsetminus \{1,2,\infty\}$.*

---

[1]For concreteness, L_p will always be the space of (equivalence classes of) p-integrable functions on the interval $[0,1]$, equipped with Lebesgue measure, but all of our results hold for any $L_p(\mu)$ space; this follows formally as any separable $L_p(\mu)$ space embeds isometrically into L_p (see e.g. [97, Chapter III.A]), though the proofs extend effortlessly to arbitrary $L_p(\mu)$ spaces. Other such standard Banach space-theoretic notations and conventions that will be used herein are according to [56, 57].

[2]We will use throughout the ensuing text the following (standard) conventions for asymptotic notation, in addition to the usual $O(\cdot), o(\cdot), \Omega(\cdot), \Theta(\cdot)$ notation. Given $a,b > 0$, by writing $a \lesssim b$ or $b \gtrsim a$ we mean that $a \leqslant \kappa b$ for some universal constant $\kappa > 0$, and $a \asymp b$ stands for $(a \lesssim b) \wedge (b \lesssim a)$. When we will need to allow for dependence on parameters, we will indicate it by subscripts. For example, in the presence of auxiliary objects q, U, ϕ, the notation $a \lesssim_{q,U,\phi} b$ means that $a \leqslant \kappa(q,U,\phi) b$, where $\kappa(q,U,\phi) > 0$ may depend only on q, U, ϕ, and similarly for the notations $a \gtrsim_{q,U,\phi} b$ and $a \asymp_{q,U,\phi} b$. Also, in what follows when expressions like, say, $\sqrt{\log\log\log n}$ appear for some integer n, it will be assumed tacitly that n is greater than a sufficiently large universal constant, so that they make sense (thus, $n > 15$ in the above example). In all such occurrences, the corresponding statement will be self-evident for smaller values of $n \in \mathbb{N}$ (by suitably adjusting an implicit universal constant factor).

We expect that answering Question 3 would be very challenging (based in part on how the only known nontrivial cases $p \in \{1,\infty\}$ of Question 3 were resolved), especially if the answer is positive, which would likely be a major achievement that would entail introducing a significant new idea/embedding method.

The challenge within Question 3 is to prove the upper bound on $\mathsf{c}_2^n(L_p)$ in (4), as the lower bound

$$\forall\, n \in \{2,3,\ldots,\}, \qquad \mathsf{c}_2^n(L_p) \gtrsim (\log n)^{\left|\frac{1}{2}-\frac{1}{p}\right|}. \tag{5}$$

follows from known results (observe that thanks to (5) our new upper bound (3) is optimal up to lower order factors whenever $p = (\log n)^{o(1)} \to \infty$). To justify (5), if $1 \leqslant p < 2$, then for each $k \in \mathbb{N}$ consider the n-point hypercube $\{0,1\}^k$ as a subset of ℓ_p^k, thus $n = |\{0,1\}^n| = 2^k$. By [31] its Euclidean distortion is

$$\mathsf{c}_2\big(\{0,1\}^k, \|\cdot\|_{\ell_p^k}\big) = k^{\frac{1}{p}-\frac{1}{2}} \asymp (\log n)^{\frac{1}{p}-\frac{1}{2}},$$

which shows that (5) holds when n is a power of 2; it is straightforward to deduce from this that (5) holds for every integer $n \geqslant 2$ (e.g., by augmenting the hypercube with auxiliary very distant points that form an equilateral metric). If $p \geqslant 2$, then by [87] there is an n-vertex connected planar graph G_n, equipped with its shortest-path metric, whose Euclidean distortion satisfies $\mathsf{c}_2(\mathsf{G}_n) \gtrsim \sqrt{\log n}$. As $p \geqslant 2$, the proof of the main result of [91] (see also the exposition in [36, 47]) shows that there is an embedding f of G_n (indeed, of any connected planar graph) into L_p whose bi-Lipschitz distortion is at most a universal constant multiple of $(\log n)^{1/p}$; the image $f(\mathsf{G}_n) \subseteq L_p$ exhibits the validity of (5) in the remaining range $p \geqslant 2$.

The lower bound (5) is only part of the picture, as there *must* be some dependence on p in (4) that diverges as $p \to \infty$, which demonstrates that there is a qualitative difference between the nonlinear setting of Question 3 and its aforementioned linear counterpart [54]. Indeed, consider any n-point metric space $\mathcal{M}_n$ with $\mathsf{c}_2(\mathcal{M}_n) \gtrsim \log n$, which exists thanks to [58, 8]. Because $\mathcal{M}_n$ embeds isometrically into ℓ_∞^n, while ℓ_∞^n embeds with $O(1)$ distortion into $L_{\Omega(\log n)}$, it follows that for $p \asymp \log n$ we necessarily have

$$\mathsf{c}_2^n(L_p) \gtrsim \sqrt{\min\{p, \log n\}}(\log n)^{\frac{1}{2}-\frac{1}{p}}. \tag{6}$$

This suggests (but does not prove!) that perhaps (6) holds for all $p > 2$. If so, then it would follow that a power-type dependence on p in our new bound (3) is unavoidable if $\alpha(\log\log n)/\log\log\log n \leqslant p \lesssim \log n$ for some fixed $\alpha > 2$, e.g., combining (3) and (6) gives $\sqrt{(\log n)\log\log n} \lesssim \mathsf{c}_2^n(L_{\log\log n}) \lesssim \sqrt{\log n}(\log\log n)^3$.

1.2. **Lipschitz extension.** Given a (source) metric space $(\mathcal{M}, d_\mathcal{M})$, a subset $\mathcal{C}$ of $\mathcal{M}$ such that $|\mathcal{C}| \geqslant 2$, and a (target) metric space $(\mathcal{T}, d_\mathcal{T})$, one denotes (following [64]) by $\mathsf{e}(\mathcal{M}, \mathcal{C}; \mathcal{T}) \in [1,\infty]$ the infimum over those $K > 0$ such that for every function $f : \mathcal{C} \to \mathcal{T}$ there exists a function $F : \mathcal{M} \to \mathcal{T}$ whose restriction to $\mathcal{C}$ coincides with f, and the Lipschitz constant of F satisfies:

$$\|F\|_{\mathrm{Lip}(\mathcal{M};\mathcal{T})} \leqslant K\|f\|_{\mathrm{Lip}(\mathcal{C};\mathcal{T})}, \tag{7}$$

where (7) uses the following notation for Lipschitz constants, which will occur throughout what follows:

$$\|f\|_{\mathrm{Lip}(\mathcal{C};\mathcal{T})} \stackrel{\mathrm{def}}{=} \sup_{\substack{x,y\in\mathcal{C}\\ x\neq y}} \frac{d_\mathcal{T}\big(f(x), f(y)\big)}{d_\mathcal{M}(x,y)}.$$

The supremum of $\mathsf{e}(\mathcal{M}, \mathcal{C}; \mathcal{T})$ over all the subsets $\mathcal{C}$ of $\mathcal{M}$ containing at least two points is denoted $\mathsf{e}(\mathcal{M}; \mathcal{T})$.

For each integer $n \geqslant 2$ one defines $\mathsf{e}_n(\mathcal{M})$ to be the supremum of $\mathsf{e}(\mathcal{M}, \mathcal{C}; \mathbf{Z})$ over all $\mathcal{C} \subseteq \mathcal{M}$ satisfying $2 \leqslant |\mathcal{C}| \leqslant n$, and all Banach spaces $(\mathbf{Z}, \|\cdot\|_\mathbf{Z})$. By [53, Theorem 1.10], every metric space $(\mathcal{M}, d_\mathcal{M})$ satisfies:

$$\forall\, n \in \{3,4,\ldots\}, \qquad \mathsf{e}_n(\mathcal{M}) \lesssim \frac{\log n}{\log\log n}. \tag{8}$$

Remark 4. Given a metric space $(\mathcal{M}, d_\mathcal{M})$, its *Lipschitz extension modulus* $\mathsf{e}(\mathcal{M})$ is defines to be the supremum of $\mathsf{e}(\mathcal{M}, \mathcal{C}; \mathbf{Z})$ over all $\mathcal{C} \subseteq \mathcal{M}$ with $|\mathcal{C}| \geqslant 2$ and all Banach spaces $(\mathbf{Z}, \|\cdot\|_\mathbf{Z})$. When the (typical) situation $\mathsf{e}(\mathcal{M}) = \infty$ occurs, one can consider analogously to (1) the finitary invariant $\mathsf{e}^n(\mathcal{M})$ which is defined for each $n \in \{2,3,\ldots\}$ to be the supremum of $\mathsf{e}(\mathcal{C})$ over all subsets $\mathcal{C} \subseteq \mathcal{M}$ with $2 \leqslant |\mathcal{C}| \leqslant n$.[3] It is important to stress the difference between $\mathsf{e}_n(\mathcal{M})$ and $\mathsf{e}^n(\mathcal{M})$, namely, $\mathsf{e}_n(\mathcal{M})$ measures the extent to which for any $\mathcal{C} \subseteq \mathcal{M}$ with $2 \leqslant |\mathcal{C}| \leqslant n$, any Lipschitz function from $\mathcal{C}$ to a Banach space can be extended to a Lipschitz function on $\mathcal{M}$, while $\mathsf{e}^n(\mathcal{M})$ disregards how $\mathcal{C}$ is situated in the super-space $\mathcal{M}$, asking for such extensions to $\mathcal{C}$ of Banach space-valued functions from arbitrary subsets of $\mathcal{C}$. While $\mathsf{e}_n(\mathcal{M})$ was studied in the literature for a long time (starting with [42], inspired by [35, 55]), to the best of our knowledge $\mathsf{e}^n(\mathcal{M})$ was not considered before. Nevertheless, $\mathsf{e}^n(\mathcal{M})$ has a key role in our investigations herein; see Section 1.3.2.

By [53, Theorem 1.12] for every $1 < p \leqslant 2$ we have:

$$\forall n \in \{2,3,\ldots,\}, \qquad \mathsf{e}_n(L_p) \lesssim_p (\log n)^{\frac{1}{p}}. \tag{9}$$

The following Lipschitz extension theorem treats the analog of (9) in the previously unknown range $p > 2$:

Theorem 5. *If $2 < p < \infty$ and $n \in \{2,3,\ldots,\}$, then*

$$\mathsf{e}_n(L_p) \lesssim p^2 \sqrt{\log n}.$$

Theorem 5 answers the natural question that was left open by [53] (and stated elsewhere, e.g. [82]) whether for every (indeed, any) fixed $2 < p < \infty$ we have

$$\lim_{n\to\infty} \mathsf{e}_n(L_p)\frac{\log\log n}{\log n} = 0. \tag{10}$$

In other words, prior to Theorem 5 it was not even known if one could provide an upper bound on $\mathsf{e}_n(L_p)$ that is asymptotically better than merely using the fact that L_p is a metric space through (8).

Remark 6. We hesitate to call (10) "nontrivial extension growth" for L_p partly due to the distinction that is noted in Remark 4, and mainly because in contrast to the analogous setting in (2), it is unknown if (8) is asymptotically sharp, nor is it conjectured that this is the case (though, it could very well be so). Determining the largest possible growth rate as $n \to \infty$ of $\mathsf{e}_n(\mathcal{M})$ over all metric spaces $\mathcal{M}$ is a major open question for which the currently best-known lower bound [82] is that $\mathsf{e}_n(\mathcal{M})$ must sometimes be at least a positive universal constant multiple of $\sqrt{\log n}$. In other words, the currently best-known bounds on $\mathsf{e}_n(L_\infty)$ are $\sqrt{\log n} \lesssim \mathsf{e}_n(L_\infty) \lesssim (\log n)/\log\log n$. For finite p, at present there is insufficient information for there to be a widely accepted conjecture what the sharp asymptotic dependence as $n \to \infty$ should be in Theorem 5, as well as in the estimate (9) of [53]. The best-available bounds in this context are presented in Section 1.2.1 below, from which it follows that if p is allowed to depend on n so that it tends to ∞ at a $(\log n)^{o(1)}$ rate, then Theorem 5 is sharp up to lower order factors; e.g. we now know that:

$$\sqrt{\log n} \lesssim \mathsf{e}_n(L_{\log\log n}) \lesssim \sqrt{\log n}(\log\log n)^2.$$

Conceivably $\mathsf{e}_n(L_p) \asymp_p \sqrt{\log n}$ or even $\mathsf{e}_n(L_p) \asymp \sqrt{\log n}$ for all $p \geqslant 2$, but proving this would be a spectacular achievement; in particular, the former statement would settle the Hilbertian case $p = 2$ and the latter statement would also settle the aforementioned case $p = \infty$ of extension from finite subsets of general metric spaces, while both of these remain tantalizingly unknown despite major efforts over many years.

The implicit constant in (9) that the proof in [53] provides tends to ∞ as $p \to 1^+$; this seems inherent to the currently available approach due to its reliance on [63].[4] Thus the case $p = 1$ remains the last holdout for the question of improving (8) when $\mathcal{M}$ is L_p for some fixed $1 \leqslant p < \infty$.[5]

[3] More generally, we will maintain the following notational convention. Given an invariant $\mathsf{i}(\mathcal{M}) \in \mathbb{R} \cup \{\infty\}$ of metric spaces $(\mathcal{M}, d_{\mathcal{M}})$, for each $n \in \{2,3,\ldots\}$ the superscript notation $\mathsf{i}^n(\mathcal{M})$ is reserved for the supremum of $\mathsf{i}(\mathcal{C})$ over $\mathcal{C} \subseteq \mathcal{M}$ with $2 \leqslant |\mathcal{C}| \leqslant n$. In contrast, the subscript notation $\mathsf{i}_n(\mathcal{M})$ is used less consistently in the literature to denote more subtle finitary invariants that still consider arbitrary subsets $\mathcal{C}$ of $\mathcal{M}$ with $2 \leqslant |\mathcal{C}| \leqslant n$, but do not depend only on the intrinsic geometry of $(\mathcal{C}, d_{\mathcal{M}})$.

[4] The best available bound as $p \to 1^+$ on the implicit constant in (9) can be deduced from the proof in [80, Section 4.2], which implies that one can take it to be at most a universal constant multiple of $1/(p-1)$.

[5] The case $p = \infty$ is also open, as we recalled in Remark 6, since it corresponds to arbitrary metric spaces $(\mathcal{M}, d_{\mathcal{M}})$.

Problem 7. *Determine whether for $n \in \{3,4,\ldots\}$ we have* $\mathsf{e}_n(L_1) = o\left(\frac{\log n}{\log\log n}\right)$.

It is worthwhile to recall in the context of Problem 7 that by [69] we have $\mathsf{e}_n(\mathcal{M};\mathbf{Z}) \lesssim_{\mathbf{Z}} \sqrt{\log n}$ for every metric space $(\mathcal{M}, d_{\mathcal{M}})$ and every Banach space $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$ that has an equivalent norm whose modulus of uniform convexity has power type 2 (see e.g. [11] for background on this class of spaces); the Hilbertian special case $\mathbf{Z} = L_2$ of this statement is the famous Johnson–Lindenstrauss extension theorem [41].

A positive answer to Problem 7 would have significant algorithmic consequences by improving over the best-known upper bound on the existence of vertex cut sparsifiers of weighted graphs [72], which has stood for a long time despite substantial efforts to improve it in the computer science literature (see e.g. [24, 32, 62]). The connection between Lipschitz extension and graph sparsification that yields this potential application of Problem 7 was discovered in [62]; we omit the details and the definition of the relevant sparsification notion as they are covered thoroughly in [62], as well as in e.g. [22, Section 1.3.3].

Remark 8. Even though it is of secondary importance given the current state of knowledge, we note that if one allows $p \geqslant 2$ to depend on n, then Theorem 5 improves over the general estimate (8) if and only if $p = o(\sqrt[4]{\log n}/\sqrt{\log\log n})$. Understanding this for $p \gtrsim \sqrt[4]{\log n}/\sqrt{\log\log n}$ is open. For that matter, by examining its behavior as $p \to 1^+$, one sees that the bound (9) of [53] improves over (8) if and only if $(p-1)\log\log n = 2(\log\log\log n) - \log\log\log\log n + O(1)$; understanding this for other $1 < p < 2$ is open.

1.2.1. *Known Lipschitz extension impossibility results.* Unlike the discussion in Section 1.1.1, there is no widely accepted conjectural growth rate (as $n \to \infty$) for $\mathsf{e}_n(L_p)$ when $1 \leqslant p \leqslant \infty$ is fixed; this growth rate is not even known in the Hilbertian setting $p = 2$. A naïve appeal to [54] from the perspective of the Ribe program leads to the prediction $\mathsf{e}_n(L_p) \asymp_p (\log n)^{|1/2-1/p|}$. This was indeed a possibility for quite some time (see e.g. the questions posed in [42, 9] for $p = 2$), until it was shown in [73] to fail even when $p = 2$.

The best-known lower bound in the range $2 \leqslant p \leqslant 4$ is $\mathsf{e}_n(L_p) \gtrsim \sqrt[4]{(\log n)/\log\log n}$. In fact, for every infinite dimensional Banach space $\mathbf{X}$ we have $\mathsf{e}_n(\mathbf{X}) \gtrsim \sqrt[4]{(\log n)/\log\log n}$ because by [71] (see also [73]),

$$\forall m \in \mathbb{N}, \qquad \mathsf{e}_{m^{O(m)}}(\ell_2^m) \gtrsim \sqrt[4]{m}, \tag{11}$$

while for every $m \in \mathbb{N}$ by Dvoretzky's theorem [30] ℓ_2^m is $O(1)$-isomorphic to a subspace of $\mathbf{X}$.

For the remaining values of p, the best-known lower bounds are $\mathsf{e}_n(L_p) \gtrsim ((\log n)/\log\log n)^{|1/2-1/p|}$ when $p \in [1,2) \cup (2,\infty)$ and $\mathsf{e}_n(L_\infty) \gtrsim \sqrt{\log n}$. For $1 \leqslant p < 2$, this coincides with [41, Theorem 3]. If $p > 2$, then $\mathsf{e}_n(L_p) \gtrsim ((\log n)/\log\log n)^{1/2-1/p}$ also follows from the reasoning in [41], though it is not stated there explicitly. Indeed, by [94] for every $1 \leqslant p \leqslant \infty$ and every $m \in \mathbb{N}$ there is a linear subspace $\mathbf{Y} = \mathbf{Y}_p$ of ℓ_p^m such that $\|\mathsf{Proj}\|_{\ell_p^m \to \ell_p^m} \gtrsim m^{|1/2-1/p|}$ for every projection Proj from ℓ_p^m onto $\mathbf{Y}$. By [55], this implies:

$$\mathsf{e}(\ell_p^m, \mathbf{Y}; \mathbf{Y}) \gtrsim m^{\left|\frac{1}{2}-\frac{1}{p}\right|}. \tag{12}$$

The discretization method of [41] deduces from (12) that there is a subset $\mathcal{N} = \mathcal{N}_p$ of ℓ_p^m with $|\mathcal{N}| = m^{O(m)}$ (specifically, $\mathcal{N}$ is a $(1/m^{O(1)})$-net of the unit sphere of $\mathbf{Y}$) such that $\mathsf{e}(\ell_p^m, \mathcal{N}; \mathbf{Y}) \gtrsim m^{|1/2-1/p|}$.

1.2.2. *Lipschitz extension from doubling subsets.* Given $\lambda \in \{2,3,\ldots\}$, a metric space $(\mathcal{M}, d_{\mathcal{M}})$ is λ-doubling if for every $r \geqslant 0$ and every $x \in \mathcal{M}$ there exist $y_1, \ldots, y_\lambda \in \mathcal{M}$ with $B_{\mathcal{M}}(x, 2r) \subseteq B_{\mathcal{M}}(y_1, r) \cup \ldots \cup B_{\mathcal{M}}(y_\lambda, r)$.[6] Here, as well as throughout the ensuing discussion, $B_{\mathcal{M}}(u,\rho) = \{v \in \mathcal{M} :\ d_{\mathcal{M}}(u,v) \leqslant \rho\}$ denotes the *closed* $d_{\mathcal{M}}$-ball centered at a point $u \in \mathcal{M}$ of radius $\rho \geqslant 0$.

By [53, Theorem 1.6], for every metric space $(\mathcal{M}, d_{\mathcal{M}})$, for every λ-doubling subset $\mathcal{D}$ of $\mathcal{M}$, and for every Banach space $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$ we have

$$\mathsf{e}(\mathcal{M}, \mathcal{D}; \mathbf{Z}) \lesssim \log\lambda. \tag{13}$$

It is an important open problem to determine if in the above stated generality the right hand side of (13) can be reduced to $o(\log\lambda)$ as $\lambda \to \infty$. It is also natural to investigate if such an improvement could be

[6] The notion of λ-doubling when $\lambda \geqslant 2$ is not necessarily an integer coincides with $\lfloor\lambda\rfloor$-doubling.

achieved for specific metric spaces $\mathcal{M}$, though to the best of our knowledge this has not been previously posed as an open question. In particular, it was not known if (13) could be improved when $\mathcal{M}$ is a Hilbert space; this is answered as the special case $p=2$ of the following theorem:

Theorem 9. *If $2\leqslant p<\infty$, then the following estimate holds for every $\lambda\in\{2,3,\ldots\}$, for every subset $\mathcal{D}$ of L_p that is λ-doubling, and for every Banach space $(\mathbf{Z},\|\cdot\|_{\mathbf{Z}})$:*

$$\mathsf{e}(L_p,\mathcal{D};\mathbf{Z})\lesssim\left(p^2\sqrt{\log p}\right)\sqrt{\log\lambda}. \tag{14}$$

We expect that the right hand side of (13) could be improved to $o(\log\lambda)$ also in the remaining range $1\leqslant p<2$, i.e., for every $\lambda\in\{2,3,\ldots\}$, for every subset $\mathcal{D}$ of L_p that is λ-doubling, and for every Banach space $(\mathbf{Z},\|\cdot\|_{\mathbf{Z}})$, we conjecture that $\mathsf{e}(L_p,\mathcal{D};\mathbf{Z})=o(\log\lambda)$. The case $p=1$ here is most tenuous in terms of available methods and evidence. When $1<p<2$, an attempt to combine our approach herein with the proof of (9) in [53] leads to probabilistic issues that we currently do not know how to address but they could be quite tractable. In terms of the best-known lower bounds in the context of Theorem 9, as every finite dimensional normed space $\mathbf{Y}$ is λ-doubling for $\lambda=e^{O(\dim(\mathbf{Y}))}$, it follows from (11) and (12) that for every $\lambda\in\{2,3,\ldots\}$ and every $1\leqslant p\leqslant\infty$ there exists a λ-doubling subset $\mathcal{D}=\mathcal{D}_p$ of L_p and a Banach space $\mathbf{Z}=\mathbf{Z}_p$ for which $\mathsf{e}(L_p,\mathcal{D};\mathbf{Z})\gtrsim\sqrt[4]{\log\lambda}$ if $2\leqslant p\leqslant 4$ and $\mathsf{e}(L_p,\mathcal{D};\mathbf{Z})\gtrsim(\log\lambda)^{|1/2-1/p|}$ if $p\in[1,2)\cup(4,\infty]$.

1.3. **Randomized clustering.** Given a metric space $(\mathcal{M},d_{\mathcal{M}})$ and a partition $\mathcal{P}$ of $\mathcal{M}$, denote for each $x\in\mathcal{M}$ the element of $\mathcal{P}$ containing x by $\mathcal{P}(x)$. For $\Delta>0$, a partition $\mathcal{P}$ of $\mathcal{M}$ is said to be Δ-bounded if $\mathrm{diam}_{\mathcal{M}}(\mathcal{P}(x))\leqslant\Delta$ for every $x\in\mathcal{M}$. Here, as well as throughout the ensuing discussion, the $d_{\mathcal{M}}$-diameter of $\emptyset\neq\mathcal{C}\subseteq\mathcal{M}$ will be denoted $\mathrm{diam}_{\mathcal{M}}(\mathcal{C})=\sup_{x,y\in\mathcal{C}}d_{\mathcal{M}}(x,y)\in[0,\infty]$.

Following [80], one says that $\mathcal{P}=\{\Gamma^1,\Gamma^2,\ldots\}$ is a *random partition* of a metric space $(\mathcal{M},d_{\mathcal{M}})$ if there exists a probability space $(\Omega,\mathbb{P})$ such that $\Gamma^1:\Omega\to 2^{\mathcal{M}},\Gamma^2:\Omega\to 2^{\mathcal{M}},\ldots$ is a sequence of set-valued mappings that are strongly measurable,[7] namely, $\{\omega\in\Omega:\ E\cap\Gamma^i(\omega)\neq\emptyset\}$ is $\mathbb{P}$-measurable for every $i\in\mathbb{N}$ and every *closed* subset E of $\mathcal{M}$, and if we write $\mathcal{P}^\omega=\{\Gamma^k(\omega):\ k\in\mathbb{N}\}$ for each $\omega\in\Omega$, then the mapping $\omega\mapsto\mathcal{P}^\omega$ takes values in partitions of $\mathcal{M}$. Note that we are formally discussing here random *ordered* partitions of $\mathcal{M}$ into countably many clusters, but this nuance will not have a role in what follows (it is important only for some of the results from the literature that we will need to quote). Measurability is not relevant when $\mathcal{M}$ is finite, which is the setting that we will initially discuss below, but we will quickly need to also treat random partitions of infinite spaces, at which point we will verify measurability as required.

Given $\Delta>0$, one says that $\mathcal{P}$ is a Δ-bounded random partition of $(\mathcal{M},d_{\mathcal{M}})$ if $\mathcal{P}^\omega$ is a Δ-bounded partition of $(\mathcal{M},d_{\mathcal{M}})$ for every $\omega\in\Omega$. Given $\sigma\geqslant 0$, a random Δ-bounded partition $\mathcal{P}$ of a metric space $(\mathcal{M},d_{\mathcal{M}})$ is said to be σ-separating if the following requirement holds:

$$\forall\, x,y\in\mathcal{M},\qquad \mathbb{P}\left[\mathcal{P}(x)\neq\mathcal{P}(y)\right]\leqslant\frac{\sigma}{\Delta}d_{\mathcal{M}}(x,y). \tag{15}$$

The separation modulus of $(\mathcal{M},d_{\mathcal{M}})$, denoted $\mathsf{SEP}(\mathcal{M})$, is the infimum over $\sigma\geqslant 0$ such that for every $\Delta>0$ there exists a random Δ-bounded σ-separating partition $\mathcal{P}_\Delta$ of $\mathcal{M}$; if no such random partition exists, then set $\mathsf{SEP}(\mathcal{M})=\infty$. This important concept has been introduced by [12], see [80] for the history.

For an infinite metric space $(\mathcal{M},d_{\mathcal{M}})$, define its separation growth $\left\{\mathsf{SEP}^n(\mathcal{M})\right\}_{n=1}^\infty$ by

$$\forall\, n\in\mathbb{N},\qquad \mathsf{SEP}^n(\mathcal{M})\stackrel{\mathrm{def}}{=}\sup_{\substack{\mathcal{C}\subseteq\mathcal{M}\\ |\mathcal{C}|\leqslant n}}\mathsf{SEP}(\mathcal{C}).$$

We say that $(\mathcal{M},d_{\mathcal{M}})$ has nontrivial separation growth if

$$\lim_{n\to\infty}\frac{\mathsf{SEP}^n(\mathcal{M})}{\log n}=0. \tag{16}$$

[7]This notion of measurability of set-valued functions is called here "strongly measurable" even though parts of the literature calls it more simply "measurable" to distinguish it from the notion of a "weakly measurable" set-valued function which is also commonly used in the literature but is *not* what we need herein. See [39] for a treatment of these classical concepts.

The term "nontrivial" is used here since by [12] *every* metric space $(\mathcal{M},d_{\mathcal{M}})$ satisfies $\mathsf{SEP}^n(\mathcal{M})=O(\log n)$.

By [23, 51, 80], if $1\leqslant p\leqslant 2$, then $\mathsf{SEP}^n(L_p)=o(\log n)$ if $\lim_{n\to\infty}(p-1)(\log\log n)/\log\log\log n=\infty$, while if $\liminf_{n\to\infty}(p-1)\log\log n<\infty$, then $\mathsf{SEP}^n(L_p)\gtrsim\log n$. See [80, Section 1.7.6] for more on this (including sharp bounds for fixed $1<p\leqslant 2$), where it is conjectured that $\mathsf{SEP}^n(L_p)=o(\log n)$ if and only if $\lim_{n\to\infty}(p-1)\log\log n=\infty$.

It was asked in [77, Question 1] (reiterated in [80, Question 83]) if L_p has nontrivial separation growth when $2<p<\infty$ (see [80, Section 1.7.6] for the relation to metric dimension reduction). This question was answered affirmatively in [48] by showing that $\mathsf{SEP}^n(L_p)\lesssim(\log n)^{1-1/p}$ for every $(p,n)\in(2,\infty)\times\mathbb{N}$. It was also asked in [77, 80] if, in fact, for every $(p,n)\in(2,\infty)\times\mathbb{N}$ we have $\mathsf{SEP}^n(L_p)\lesssim_p\sqrt{\log n}$, which would be asymptotically optimal as $n\to\infty$ for every fixed $2<p<\infty$ (see below). Here we prove that this sharp evaluation of the largest possible separation modulus of an n-point subset of L_p indeed holds:

Theorem 10. *For every $2<p<\infty$ and every $n\in\{2,3,\ldots\}$ we have $\mathsf{SEP}^n(L_p)\asymp_p\sqrt{\log n}$. More precisely,*

$$\forall(p,n)\in(2,\infty)\times\mathbb{N},\qquad \sqrt{\log n}\lesssim\mathsf{SEP}^n(L_p)\lesssim p^2\sqrt{\log n}. \tag{17}$$

The new content of (17) is its upper bound on $\mathsf{SEP}^n(L_p)$; the lower bound on $\mathsf{SEP}^n(L_p)$ in (17) holds since $\mathsf{SEP}^n(L_2)\asymp\sqrt{\log n}$, by [23], while $\mathsf{SEP}^n(\mathbf{X})\geqslant\mathsf{SEP}^n(L_2)$ for every infinite dimensional Banach space $\mathbf{X}$, by Dvoretzky's theorem [30]. So, we now know that $\mathsf{SEP}^n(L_p)=o(\log n)$ as $n\to\infty$ if $2<p=o\big(\sqrt[4]{\log n}\big)$, and it fails if $p\gtrsim\log n$ because in that case ℓ_∞^n embeds with distortion $O(1)$ into L_p, any n-point metric space $\mathcal{M}$ embeds into ℓ_∞^n, and we already recalled that by [12] there exists such a space for which $\mathsf{SEP}(\mathcal{M})\gtrsim\log n$. It remains open (likely requiring a substantially new idea) to understand what happens in the remaining range $\sqrt[4]{\log n}\lesssim p=o(\log n)$.

Remark 11. The first arXiv posting [83] of our work (which the present article supersedes) suppressed the dependence on p in Theorem 10 because the main matter at hand is determining the asymptotic growth rate of $\mathsf{SEP}^n(L_p)$ as $n\to\infty$; this is what was asked in [77] and what Theorem 10 answers. Nevertheless, understanding the dependence on p is of value in its own right (partially because in applications sometimes p itself is allowed to depend on n), and an inspection of the proof in [83] reveals that it yields the estimate $\mathsf{SEP}^n(L_p)\leqslant e^{O(p)}\sqrt{\log n}$. This was achieved in [83] by combining the approach of [13, 48] with a bootstrapping argument, which, as explained in [83, Remark 8] can also be realized as an iterative procedure. The question of improving the dependence on p was broached in the subsequent work [49], which enhanced the recursion in a novel and interesting way to get the bound $\mathsf{SEP}^n(L_p)\lesssim p^4\sqrt{\log n}$. The proof herein of Theorem 10 incorporates a more geometric approach by examining radially bounded random partitions and relying on a new property of the Mazur map that could be of use elsewhere; an overview of the ideas and steps of that proof appears in Section 1.3.3 below.

We will also prove the following theorem for doubling subsets of L_p:

Theorem 12. *If $p,\lambda\geqslant 2$, then every λ-doubling subset $\mathcal{D}$ of L_p satisfies:*

$$\mathsf{SEP}(\mathcal{D})\lesssim\Big(p^2\sqrt{\log p}\Big)\sqrt{\log\lambda}.$$

1.3.1. *From separation of neighborhoods to Lipschitz extension.* The similarity between the conclusions of Theorem 10 and Theorem 5, as well as Theorem 12 and Theorem 9, are not coincidental. The link is provided by Theorem 13 below, which we will deduce quickly in Section 7 from [53] (with input from [80]).

In the formulation of Theorem 13, as well as throughout the ensuing discussions, we will use the following (natural but nonstandard) notation. Given a metric space $(\mathcal{M},d_{\mathcal{M}})$ and $\mathcal{C}\subseteq\mathcal{M}$, for every $r\geqslant 0$ we will denote the r-neighborhood $\mathcal{C}$ in $\mathcal{M}$ by:

$$B_{\mathcal{M}}(\mathcal{C},r)\stackrel{\mathrm{def}}{=}\bigcup_{x\in\mathcal{C}}B_{\mathcal{M}}(x,r). \tag{18}$$

Given $\Delta > 0$ we will denote by $\mathsf{SEP}_\Delta(\mathcal{M})$ the infimum over $\sigma \geqslant 0$ such that there is a random Δ-bounded σ-separating partition of $\mathcal{M}$ (again, if no such random partition exists, then write $\mathsf{SEP}(\mathcal{M}) = \infty$). Thus,

$$\mathsf{SEP}(\mathcal{M}) = \sup_{\Delta>0} \mathsf{SEP}_\Delta(\mathcal{M}).$$

Theorem 13. *Suppose that $(\mathcal{M}, d_\mathcal{M})$ is a metric space and that $\mathcal{C} \neq \emptyset$ is a locally compact subset of $\mathcal{M}$. Then, the following Lipschitz extension estimate holds for every Banach space $(\mathbf{Z}, \|\cdot\|_\mathbf{Z})$ and every $L > 0$:*

$$\mathsf{e}(\mathcal{M}, \mathcal{C}; \mathbf{Z}) \lesssim L + \sup_{\Delta>0} \mathsf{SEP}_\Delta\Big(B_\mathcal{M}\big(\mathcal{C}, \frac{1}{L}\Delta\big)\Big). \tag{19}$$

Our main contribution to randomized clustering of (neighborhoods of) subsets of L_p is:

Theorem 14. *There exists a universal constant $\gamma > 0$ with the following property. Suppose that $2 \leqslant p < \infty$. For every $n \in \{2, 3, \ldots\}$, if $\mathcal{C}$ is an n-point subset of L_p, then for every $\Delta > 0$ we have:*

$$\mathsf{SEP}_\Delta\Big(B_{L_p}\big(\mathcal{C}, \frac{\gamma}{p}\Delta\big)\Big) \lesssim p^2\sqrt{\log n}. \tag{20}$$

Furthermore, for every $\lambda \geqslant 2$, if $\mathcal{D}$ is a subset of L_p that is λ-doubling, then for every $\Delta > 0$ we have:

$$\mathsf{SEP}_\Delta\Big(B_{L_p}\big(\mathcal{D}, \frac{\gamma}{p}\Delta\big)\Big) \lesssim \Big(p^2\sqrt{\log p}\Big)\sqrt{\log \lambda}. \tag{21}$$

Because any neighborhood of a set contains the set itself, Theorem 14 strengthens (the upper bound on $\mathsf{SEP}^n(L_p)$ in) Theorem 10 and Theorem 12. Furthermore, Theorem 5 and Theorem 9 follow from Theorem 14 by invoking Theorem 13. Indeed, any Lipschitz function from subset of a metric space that take values in a complete metric space automatically extends to the closure of its domain, so in Theorem 14 we may assume that $\mathcal{D}$ is a closed subset of L_p. As $\mathcal{D}$ is also assumed in Theorem 14 to be doubling, it is locally compact (see e.g. [38, Lemma 4.1.14], or notice that any ball in $\mathcal{D}$ is compact because it is totally bounded and complete). The assumptions of Theorem 13 therefore hold with $L \asymp p$ by Theorem 14. Note here that even when $\mathcal{C} \subseteq L_p$ is finite, the above justification of Theorem 5 involves random partitions of infinite sets (neighborhoods of $\mathcal{C}$ in L_p), for which measurability considerations are pertinent.

1.3.2. *From separation and extension to Euclidean embedding.* The link between Theorem 5 and Theorem 1 is furnished by Theorem 15 below, which is an embedding statement of independent interest; one can view it as a variant of the "measured descent" embedding method [47] that incorporates separating partitions rather than the padded partitions that occur in [47]. The proof of Theorem 15, which appears in Section 2, combines multiple ingredients, many of which (but not all) refine the reasoning in [5].

For Theorem 15, recall that $\mathsf{e}^k(\mathcal{M}) = \sup\{\mathsf{e}(\mathcal{C}) : \mathcal{C} \subseteq \mathcal{M} \ \wedge\ |\mathcal{C}| \leqslant k\}$ was defined in Remark 4.

Theorem 15. *For every $n \in \{3, 4, \ldots\}$, every n-point metric space $(\mathcal{M}, d_\mathcal{M})$ satisfies*

$$\mathsf{c}_2(\mathcal{M}) \lesssim \sqrt{(\log n)\log\log n}\bigg(\sum_{k=2}^{n} \frac{\mathsf{SEP}^k(\mathcal{M})^2 \mathsf{e}^k(\mathcal{M}; L_2)^4}{k(\log k)^2}\bigg)^{\frac12}. \tag{22}$$

For every $k \in \{2, 3, \ldots\}$ we have $\mathsf{SEP}^k(L_p) \lesssim p^2\sqrt{\log k}$ by Theorem 10, and $\mathsf{e}^k(L_p) \leqslant \mathsf{e}(L_p) \lesssim \sqrt{p}$ by [81]. A substitution of these two estimates into Theorem 15 implies Theorem 1 as follows:

$$\begin{aligned}\mathsf{c}_2^n(L_p) &\lesssim \sqrt{(\log n)\log\log n}\bigg(\sum_{k=2}^{n} \frac{p^4(\log k)\cdot p^2}{k(\log k)^2}\bigg)^{\frac12} \\ &\asymp p^3\sqrt{(\log n)\log\log n}\bigg(\int_1^n \frac{\mathrm{d}s}{s\log s}\bigg)^{\frac12} = p^3\sqrt{\log n}\log\log n.\end{aligned}$$

1.3.3. *Localization and induction on scales, and a radial property of the Mazur map.* The purpose of this section is to provide an overview of the key ingredients of our proof of Theorem 14.

Suppose that $(\mathcal{M},d_{\mathcal{M}})$ is a metric space and $\mathcal{C}\subseteq\mathcal{M}$. The $d_{\mathcal{M}}$-(circum)radius of $\mathcal{C}$ is defined as follows:

$$\mathrm{rad}_{\mathcal{M}}(\mathcal{C})\stackrel{\mathrm{def}}{=}\inf\big\{r\in[0,\infty]:\ \exists x\in\mathcal{M}\text{ such that }B_{\mathcal{M}}(x,r)\supseteq\mathcal{C}\big\},\tag{23}$$

Given $\Delta>0$, we will say that a random partition $\mathcal{P}$ of $\mathcal{C}$ is radially Δ-bounded with respect to $\mathcal{M}$ if:

$$\forall x\in\mathcal{C},\qquad \mathrm{rad}_{\mathcal{M}}\big(\mathcal{P}(x)\big)\leqslant\Delta.\tag{24}$$

A random partition $(\omega\in\Omega)\mapsto\mathcal{P}^{\omega}$ of $\mathcal{C}$, defined on some probability space $(\Omega,\mathbb{P})$, is radially Δ-bounded with respect to $\mathcal{M}$ if $\mathrm{rad}_{\mathcal{M}}(\mathcal{P}^{\omega}(x))\leqslant\Delta$ for every $\omega\in\Omega$ and every $x\in\mathcal{C}$. Given a random partition $\mathcal{P}$ of $\mathcal{C}$ that is radially Δ-bounded with respect to $\mathcal{M}$, say that it is σ-separating for some $\sigma\geqslant 0$ if (15) holds.

Denote by $\widehat{\mathsf{SEP}}_{\Delta}(\mathcal{C};\mathcal{M})$ the infimum over $\sigma\geqslant0$ such that there exists a σ-separating random partition of $\mathcal{C}$ that is radially Δ-bounded with respect to $\mathcal{M}$ (as always, when no such partition exists we define this parameter to be ∞). Recalling that $\mathrm{diam}_{\mathcal{M}}(\cdot)$ denotes the $d_{\mathcal{M}}$-diameter, the following bounds hold:

$$\forall\emptyset\neq\mathcal{C}\subseteq\mathcal{M},\qquad \mathrm{rad}_{\mathcal{M}}(\mathcal{C})\leqslant\mathrm{diam}_{\mathcal{M}}(\mathcal{C})\leqslant2\mathrm{rad}_{\mathcal{M}}(\mathcal{C}).\tag{25}$$

Consequently, the following simple general relations between the separation and radial separation moduli are satisfied for every metric space $(\mathcal{M},d_{\mathcal{M}})$, every $\emptyset\neq\mathcal{C}\subseteq\mathcal{M}$, and every $\Delta>0$:

$$\widehat{\mathsf{SEP}}_{\Delta}(\mathcal{C};\mathcal{M})\leqslant\mathsf{SEP}_{\Delta}(\mathcal{C})\leqslant2\widehat{\mathsf{SEP}}_{\frac{\Delta}{2}}(\mathcal{C};\mathcal{M}),\tag{26}$$

Even though the above radial variants of the notions of random Δ-bounded and σ-separating partitions may seem to be nuanced minor tweaks of the standard (by now classical) definitions, we will see that they influence our results in a way that is more dramatic than what one might initially expect.

The following lemma is a localization and induction on scales principle for random radial separation:

Lemma 16. *Suppose that $(\mathcal{M},d_{\mathcal{M}})$ is a separable metric space and that $\emptyset\neq\mathcal{C}\subseteq\mathcal{M}$ satisfies*

$$\lim_{\Delta\to\infty}\frac{1}{\Delta}\mathsf{SEP}_{\Delta}(\mathcal{C})=0.\tag{27}$$

Then, the following estimate holds for every $\Delta>0$ and $K>1$:

$$\widehat{\mathsf{SEP}}_{\Delta}(\mathcal{C};\mathcal{M})\leqslant\sum_{s=0}^{\infty}\frac{1}{K^s}\lim_{\varepsilon\to0^+}\sup_{z\in\mathcal{M}}\widehat{\mathsf{SEP}}_{K^s\Delta}\big(\mathcal{C}\cap B_{\mathcal{M}}(z,K^{s+1}\Delta+\varepsilon);\mathcal{M}\big).\tag{28}$$

The limits in (28) exist as the summands are nondecreasing with ε. We used the term "localization" to describe Lemma 16 as the s-summand in (28) depends only on $\mathcal{C}\cap B_{\mathcal{M}}(z,K^{s+1}\Delta+\varepsilon)$, which is a "local snapshot" of $\mathcal{C}$. We used the term "induction on scales" to describe Lemma 16 as the left hand side of (28) treats partitions at scale Δ while the right hand side of (28) considers partitions at an increasing sequence of larger scales. Our proof of Lemma 16 is an iterative use of the following observation. One can partition $\mathcal{C}$ into clusters of radius at most Δ by first partitioning it into clusters of radius at most $K\Delta$; each of those clusters is thus contained in a ball of radius $K\Delta+\varepsilon$, so we can proceed to refine the aforementioned initial partition by partitioning each of the enclosing balls into pieces of radius at most Δ.

A standard tool for analysing the geometry L_p is the classical Mazur map [67] to L_2 (and shifts thereof), whose restriction to balls has well-understood (and widely-used) quite favorable uniform continuity properties, i.e., it is well-behaved on the local snapshots of $\mathcal{C}$ that appear in the right hand side of (28).

For $1\leqslant p,q<\infty$, the Mazur map $M_{p\to q}:L_p\to L_q$ is defined [67] by:

$$\forall\phi\in L_p,\ \forall t\in[0,1],\qquad M_{p\to q}(\phi)(t)\stackrel{\mathrm{def}}{=}|\phi(t)|^{\frac{p}{q}}\,\mathrm{sign}\big(\phi(t)\big).\tag{29}$$

If $p>2$, then $M_{p\to2}$ is Lipschitz by [67], so we may consider the following normalization of the Mazur map that makes it be a 1-Lipschitz function from L_p into the Hilbert space L_2:

$$\widetilde{M}_{p\to2}\stackrel{\mathrm{def}}{=}\frac{1}{\|M_{p\to2}\|_{\mathrm{Lip}(L_p;L_2)}}M_{p\to2}.\tag{30}$$

The normalization factor in (30) satisfies $\|M_{p\to 2}\|_{\mathrm{Lip}(L_p;L_2)} < p$, as computed in [75, equation (5.32)].

We will prove herein that there exists a universal constant $\beta > 0$ such that the following inclusion holds:

$$\forall \phi \in B_{L_p}, \qquad \big(\widetilde{M}_{p\to 2}\big)^{-1}\Big(B_{L_2}\Big(\widetilde{M}_{p\to 2}(\phi), \frac{\beta}{p}\Big)\Big) \subseteq B_{L_p}\Big(\frac{1}{p}\phi, 1-\frac{1}{4p}\Big), \tag{31}$$

where B_{L_p} is the ball in L_p of radius 1 centered at the 0-function; more generally, the unit ball centered at the origin of a Banach space $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ will always be denoted $B_{\mathbf{X}} = \{x \in \mathbf{X} : \|x\|_{\mathbf{X}} \leqslant 1\} = B_{\mathbf{X}}(0,1)$.

The upshot of (31) is that the restriction of $\widetilde{M}_{p\to 2}$ to B_{L_p} gives a 1-Lipschitz function f into a Hilbert space with the property that for any point $x \in B_{L_p}$, the pullback under f of the Hilbertian ball centered at $f(x)$ of radius β/p is contained in an L_p-ball whose radius is a definite number (independent of x) smaller than 1, albeit ever so slightly as $p \to \infty$.

Returning to the setting of Theorem 14, the above discussion sets the stage for the following procedure. First apply Lemma 16 with $K = 1/(1-1/(4p))$; the choice of this value of K will allow us to take advantage of the radius decrease in (31) from 1 (the radius of the domain of f) to $1-1/(4p)$.

Thus, suppressing (only for the purpose of the present proof sketch) the small additive $\varepsilon > 0$ correction of the radius that appears in (28) (the complete reasoning in Section 6 takes this correction into account), we may now fix $z \in \mathcal{M}$, $\Delta > 0$ and an integer $s \geqslant 0$ and focus our attention on obtaining a separating partition that is $K^s\Delta$-radially bounded with respect to L_p of the local snapshots $\mathcal{C} \cap B_{L_p}(z, K^{s+1}\Delta)$ and $\mathcal{D} \cap B_{L_p}(z, K^{s+1}\Delta)$. By translating by z and rescaling by $K^{s+1}\Delta$, we are thus interested in obtaining a separating partition that is $(1/K)$-radially bounded with respect to L_p of $\mathcal{C} \cap B_{L_p}$ and $\mathcal{D} \cap B_{L_p}$.

As these local snapshots are contained in the unit ball of L_p, we may apply the Mazur map to them, for which (31) holds. The resulting images are now subsets of L_2, on which we may us Euclidean geometric considerations to randomly partition them with good separation into clusters of L_2-diameter at most β/p (using the partitioning results of [23, 53, 80] as well as the Kirszbraun extension theorem [46] and the Johnson–Lindenstrauss dimension reduction lemma [41]). The pullback under the normalized Mazur map of the resulting partition will then have good separation, and by (31) it will have L_p-radius a most $1-1/(4p) = 1/K$, as desired. Note that the center $(1/p)\phi$ of the ball in the right hand side of (31) need not belong to the corresponding local snapshot of $\mathcal{C}$ and $\mathcal{D}$, but this is permitted by the definition of radially bounded random partitions with respect to the super-space L_p. By substituting the resulting bounds on the radial separation moduli of the snapshots and summing over s, Theorem 14 follows.

Remark 17. As we already mentioned, good bounds on the moduli of uniform continuity of the Mazur map are well known [75, equation (5.32)] and widely used. If one incorporates into the above reasoning those bounds in place of the new radial property (31) of the Mazur map that we obtain herein, one arrives at exponential dependence on p in the right had side of (20) and of (21). In Section 3 we will prove that such an exponential loss is inherent to the aforementioned alternative route, and it will be incurred by any uniform homeomorphism from the unit ball of $L_p(\mu)$ into Hilbert space, not only by the Mazur map. What (31) achieves is power-type dependence on p if one is willing to settle for a small gain in the radius of the pre-image. This forces us to work with K close to 1 in Lemma 16, so the geometric convergence in (28) is slow, but as $K-1$ is of order $1/p$, this leads to a multiplicative loss which is also just $O(p)$.

1.4. **Beyond L_p.** Here we will discuss generalizations of our results to Banach spaces that need not be L_p (or any $L_p(\mu)$ space, for which all the statements are identical). Those who are interested only the aforementioned statements for L_p can harmlessly skip this material, which assumes some (entirely standard, per e.g. [15]) background from Banach space theory that is not pertinent to the ensuing proofs.

An inspection of the reasoning herein reveals that if $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ is an infinite dimensional Banach space, then the conclusions of Theorem 1, Theorem 5, Theorem 9, Theorem 12, Theorem 10, and Theorem 14 hold with L_p replaced by $\mathbf{X}$ provided that there is an injective Lipschitz function from the unit ball of $\mathbf{X}$ into a Hilbert space whose inverse is uniformly continuous, with the only difference being that in this (much) more general setting the dependence on p should be replaced by a dependence on the Lipschitz constant and the modulus of uniform continuity of the inverse in the aforementioned assumption on $\mathbf{X}$. There is substantial literature [67, 20, 88, 21, 28, 15, 92, 2, 93, 26, 27] that obtains uniform homeomorphisms from the unit ball of certain Banach spaces into a Hilbert space, which we will next describe.

The important work [88] implies that if $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ is a Banach space with an unconditional basis whose modulus of uniform smoothness has power-type 2 and whose modulus of uniform convexity has power type p, then there is an injective Lipschitz function from the unit ball of $\mathbf{X}$ into a Hilbert space whose inverse is $(2/p)$-Hölder,[8] whence the theorems that we listed above hold for such $\mathbf{X}$. The work [21] generalizes the aforementioned result of [88] to Banach lattices, though in a manner that is lossy with respect to the Hölder estimates that are obtained; we expect, however, that with not much more care the reasoning in [21] could be adapted to show that if $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ is a Banach lattice whose modulus of uniform smoothness has power-type 2 and whose modulus of uniform convexity has power type p, there is an injective Lipschitz function from the unit ball of $\mathbf{X}$ into a Hilbert space whose inverse is $(2/p)$-Hölder.

By [93], if $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ is the Schatten–von Neumann trace class S_p, or more generally if it is a noncommutative L_p space over any von Neumann algebra, then there is an injective Lipschitz function from the unit ball of $\mathbf{X}$ into a Hilbert space whose inverse is $(2/p)$-Hölder, thus showing that the theorems that we listed above can be generalized to this noncommutative setting. We conjecture that the same holds for any unitary ideal $\mathsf{S}_{\mathbf{E}}$ over any Banach space $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$ with a 1-symmetric basis, provided that $\mathbf{E}$ has modulus of uniform smoothness of power-type 2 and modulus of uniform convexity has power type p; we leave this as an open question that seems quite accessible (perhaps even straightforward) with current technology (the work [40] could be especially helpful here).

There are valuable works that construct uniform homeomorphisms into Hilbert space of unit balls in interpolation spaces (see [28, 26, 27] and [15, Section 3 of Chapter 9]), but they do not yield the above Lipschitz estimate and it would be interesting to understand (perhaps requiring a substantial new idea) whether such a Lipschitz estimate could be derived in this context as well (under suitable assumptions).

Problem 18. The extent to which the inclusion (31) generalizes beyond L_p is uncharted (and currently unstudied) territory. In particular, it would be interesting to determine if (31) holds for the noncommutative Mazur map on the Schatten–von Neumann trace class S_p when $p > 2$. And, it would be worthwhile to understand how (31) should be changed when L_p is replaced by a Banach space $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ with an unconditional basis whose modulus of uniform smoothness has power-type 2, the pertinent question being how the radii of the balls that occur in (31) should depend on geometric characteristics of $\mathbf{X}$.

1.5. **Roadmap.** The rest of the ensuing text is organized as follows. We will start by proving Theorem 15 in Section 2. As we explained in Section 1.3.2, this will complete the reduction of Theorem 1 to Theorem 10. Thus, what will remain to be done after Section 2 is to prove Theorem 14, which we have already seen implies the rest of the new results that are obtained herein; the corresponding extension statements rely on Theorem 13, which is not stated in the literature but readily follows from the link between randomized separation and Lipschitz extension that was discovered in [53]. The justification of Theorem 13 using [53] appears in Section 7. In Section 3 we will set the stage for proving the inclusion (31), which we will do in Section 4. While the concepts that are discussed in Section 3 are not needed for the proof Theorem 14, those motivate their radial counterparts that are used for this purpose and are discussed in Section 4. Also, Section 3 proves an impossibility result that explains the need for the variants that Section 4 provides. Lemma 16 is proved in Section 5, and the proof of Theorem 14 is completed in Section 6.

2. MEASURED DESCENT FOR SEPARATED PARTITIONS (IN THE PRESENCE OF LIPSCHITZ EXTENSION)

The purpose of this section is to prove Theorem 15. As we explained in Section 1.3.2, this will reduce Theorem 1 to Theorem 14, which will be proven later, in Section 6 below.

[8] [88] obtains such a function into L_1 rather than into L_2, but it is straightforward to adapt the reasoning so as to obtain a function into L_2. Furthermore, these Lipschitz and Hölder assertions follow from an inspection of the proofs in [88], but they are not stated there explicitly. It is perhaps simplest to verify them by examining the exposition of [88] in the monograph [15].

2.1. **Weakly bi-Lipschitz embeddings.** Section 1.1 recalled (for brevity of the Introduction) only the notion of Euclidean distortion, but it is standard (and fruitful) to study the analogous concept for embeddings into an arbitrary metric space. A metric space $(\mathcal{M},d_\mathcal{M})$ has a bi-Lipschitz embedding of distortion $D\geqslant 1$ into a metric space $(\mathcal{Z},d_\mathcal{Z})$ if there is a nonconstant Lipschitz function $f:\mathcal{M}\to\mathcal{Z}$ satisfying

$$\forall x,y\in\mathcal{M},\qquad d_\mathcal{Z}\big(f(x),f(y)\big)\geqslant \frac{\|f\|_{\mathrm{Lip}(\mathcal{M};\mathcal{Z})}}{D}d_\mathcal{M}(x,y), \tag{32}$$

The $\mathcal{Z}$-distortion $\mathsf{c}_\mathcal{Z}(\mathcal{M})$ of $\mathcal{M}$ is the infimum over those $D\in[1,\infty]$ for which an embedding f as above exists. We also denote $\mathsf{c}_\mathcal{Z}^n(\mathcal{M})=\sup\{\mathsf{c}_\mathcal{Z}(\mathcal{C}):\ |\mathcal{C}|\subseteq\mathcal{M}\ \wedge\ 2\leqslant|\mathcal{C}|\leqslant n\}$ for every integer $n\geqslant 2$.

There is a "one scale at a time" variant of the above classical setup, which is an important and commonly used concept in the study of metric embeddings; the relevant terminology, which we next recall, was introduced in [81, p. 192] (the reason for the nomenclature is further articulated in [75, Section 7.2]).

Definition 19. *A metric space $(\mathcal{M},d_\mathcal{M})$ admits a weakly bi-Lipschitz embedding of distortion $D\geqslant 1$ into a metric space $(\mathcal{Z},d_\mathcal{Z})$ if for any $\Delta>0$ there is a nonconstant Lipschitz function $f=f_\Delta:\mathcal{M}\to\mathcal{Z}$ satisfying*

$$\forall x,y\in\mathcal{M},\qquad d_\mathcal{M}(x,y)\geqslant\Delta\implies d_\mathcal{Z}\big(f(x),f(y)\big)\geqslant\frac{\|f\|_{\mathrm{Lip}(\mathcal{M};\mathcal{Z})}}{D}\Delta.$$

Analogously to the above notation, in the setting of Definition 19 let $\mathsf{d}_\mathcal{Z}(\mathcal{M})$ be the infimum over those $D\in[1,\infty]$ for which $(\mathcal{M},d_\mathcal{M})$ has a weakly bi-Lipschitz embedding of distortion D into $(\mathcal{Z},d_\mathcal{Z})$, and write

$$\forall n\in\{2,3,\ldots\},\qquad \mathsf{d}_\mathcal{Z}^n(\mathcal{M})\stackrel{\mathrm{def}}{=}\sup_{\substack{\mathcal{C}\subseteq\mathcal{M}\\ 2\leqslant|\mathcal{C}|\leqslant n}}\mathsf{d}_\mathcal{Z}(\mathcal{C}).$$

We will also use the shorter notation $\mathsf{d}_{L_p}(\mathcal{M})=\mathsf{d}_p(\mathcal{M})$ and $\mathsf{d}_{L_p}^n(\mathcal{M})=\mathsf{d}_p^n(\mathcal{M})$ for $p\geqslant 1$ and $n\in\{2,3,\ldots\}$.

Suitably defined "spatially localized" variants of Definition 19 will be discussed in Section 3 below, as they are important for proving the results that obtained herein. In this section, though, it suffices to consider only the above "vanilla" notion of "one scale at a time" embedding, as well as the following ad hoc notation for an obvious (also standard) "two-sided scale-localized" version thereof.

Given metric spaces $(\mathcal{M},d_\mathcal{M})$ and $(\mathcal{Z},d_\mathcal{Z})$, let $\widehat{\mathsf{d}}_\mathcal{Z}(\mathcal{M})$ denote the infimum over those $D\in[1,\infty]$ with the property that for any $\Delta>0$ there exists a nonconstant Lipschitz function $f=f_\Delta:\mathcal{M}\to\mathcal{Z}$ satisfying

$$\forall x,y\in\mathcal{M},\qquad \Delta\leqslant d_\mathcal{M}(x,y)\leqslant 2\Delta\implies d_\mathcal{Z}\big(f(x),f(y)\big)\geqslant\frac{\|f\|_{\mathrm{Lip}(\mathcal{M};\mathcal{Z})}}{D}\Delta. \tag{33}$$

We also use the shorter notation $\widehat{\mathsf{d}}_{L_p}(\mathcal{M})=\widehat{\mathsf{d}}_p(\mathcal{M})$ and $\widehat{\mathsf{d}}_{L_p}^n(\mathcal{M})=\widehat{\mathsf{d}}_p^n(\mathcal{M})$ for $p\geqslant 1$ and $n\in\{2,3,\ldots\}$, where

$$\forall n\in\{2,3,\ldots\},\qquad \widehat{\mathsf{d}}_\mathcal{Z}^n(\mathcal{M})\stackrel{\mathrm{def}}{=}\sup_{\substack{\mathcal{C}\subseteq\mathcal{M}\\ 2\leqslant|\mathcal{C}|\leqslant n}}\widehat{\mathsf{d}}_\mathcal{Z}(\mathcal{C}).$$

Remark 20. The factor 2 in (33) was fixed for concreteness, but it is an arbitrary choice of minor significance to the context of what we study herein. Specifically, suppose that $1<\alpha<\beta<\infty$ and for any $\Delta>0$ there exists a 1-Lipschitz function $f_\Delta:\mathcal{M}\to L_2$ satisfying

$$\forall x,y\in\mathcal{M},\qquad \Delta\leqslant d_\mathcal{M}(x,y)\leqslant\alpha\Delta\implies\|f_\Delta(x)-f_\Delta(y)\|_{L_2}\geqslant\frac{\Delta}{D}. \tag{34}$$

Set $k=\lceil(\log\beta)/\log\alpha\rceil-1$. If $x,y\in\mathcal{M}$ satisfy $\Delta\leqslant d_\mathcal{M}(x,y)\leqslant\beta\Delta$, then $\alpha^i\Delta\leqslant d_\mathcal{M}(x,y)\leqslant\alpha^{i+1}\Delta$ for some $i\in\{0,\ldots,k\}$, whence $\|f_{\alpha^i\Delta}(x)-f_{\alpha^i\Delta}(y)\|_{L_2}\geqslant\alpha^i\Delta/D\geqslant\Delta/D$. So, if we define $g_\Delta:\mathcal{M}\to L_2\otimes L_2\cong L_2$ by

$$g_\Delta(x)\stackrel{\mathrm{def}}{=}\frac{1}{\sqrt{k+1}}\sum_{i=0}^k f_{\alpha^i\Delta}(x)\otimes v_i,$$

where $v_0,\ldots,v_k$ is an arbitrary orthonormal system in L_2, then g_Δ is 1-Lipschitz and

$$\forall x,y\in\mathcal{M},\qquad \Delta\leqslant d_\mathcal{M}(x,y)\leqslant\beta\Delta\implies\|g_\Delta(x)-g_\Delta(y)\|_{L_2}\geqslant\frac{\Delta}{D\sqrt{k+1}}.$$

Therefore, the existence for every $\Delta > 0$ of a 1-Lipschitz function satisfying (34) implies the existence for every $\Delta > 0$ of such a function having the analogous property with α replaced by β and D replaced by

$$D\sqrt{\left\lceil \frac{\log \beta}{\log \alpha} \right\rceil} \asymp_{\alpha,\beta} D.$$

2.1.1. *From separation and extension to a weakly bi-Lipschitz embedding.* Lemma 21 below bounds $\mathsf{d}_2(\cdot)$ by a product of separation moduli and (squares of) Lipschitz extension moduli, as those that appear in the statement of Theorem 15. The statement of Lemma 21 uses the following common notation and terminology, as well as the nonstandard notation that we will introduce in (35).

Given a metric space $(\mathcal{M}, d_\mathcal{M})$, a subset $\mathcal{N}$ is called a net of $\mathcal{M}$ if there is $r > 0$ such that $\mathcal{N}$ is r-separated, i.e., $d_\mathcal{M}(a,b) \geqslant r$ for every distinct $x, y \in \mathcal{N}$, and $\mathcal{N}$ is also ε-dense in $\mathcal{M}$, i.e., for every $x \in \mathcal{M}$ there is $a \in \mathcal{N}$ such that $d_\mathcal{M}(x,a) \leqslant r$. For each target metric space $(\mathcal{T}, d_\mathcal{T})$, the Lipschitz extension modulus from nets $\mathsf{e}_{\mathrm{net}}(\mathcal{M};\mathcal{T})$ is defined to be the supremum of $\mathsf{e}(\mathcal{M},\mathcal{N};\mathcal{Z})$ over all possible nets $\mathcal{N}$ of $\mathcal{M}$, where we recall that the (standard) notation for Lipschitz extension moduli was already introduced in Section 1.2. Clearly $\mathsf{e}_{\mathrm{net}}(\mathcal{M};\mathcal{T}) \leqslant \mathsf{e}(\mathcal{M};\mathcal{T})$, but it is worthwhile to single out extension from nets because it is an especially important and useful instance of the Lipschitz extension problem, and there are situations in which bounds on $\mathsf{e}_{\mathrm{net}}(\mathcal{M};\mathcal{T})$ are known that are stronger than the available upper bounds on $\mathsf{e}(\mathcal{M};\mathcal{T})$; examples of works that either treat or rely on extension from nets include [18, 14, 71, 76, 82, 79, 3, 19, 85].

For Theorem 15, it is convenient to introduce the following notation for every metric space $(\mathcal{M}, d_\mathcal{M})$:

$$\Pi(\mathcal{M}) = \Pi(\mathcal{M}, d_\mathcal{M}) \stackrel{\mathrm{def}}{=} \max\Big\{\mathsf{SEP}(\mathcal{N})\mathsf{e}(\mathcal{M},\mathcal{N};L_2)^2 :\ \mathcal{N} \text{ is a net of } \mathcal{M}\Big\}. \tag{35}$$

Thus, $\Pi(\mathcal{M}) \leqslant \mathsf{SEP}(\mathcal{M})\mathsf{e}_{\mathrm{net}}(\mathcal{M};L_2)^2 \leqslant \mathsf{SEP}(\mathcal{M})\mathsf{e}(\mathcal{M};L_2)^2$. The following lemma relates the "separation-extension product" $\Pi(\mathcal{M})$ to the weakly bi-Lipschitz Euclidean distortion $\mathsf{d}_2(\mathcal{M})$:

Lemma 21. *Every finite metric space $(\mathcal{M}, d_\mathcal{M})$ satisfies $\mathsf{d}_2(\mathcal{M}) \leqslant 8\Pi(\mathcal{M}) + 1$.*

Proof. Our goal is to prove that for any $\Delta > 0$ there is a 1-Lipschitz function $f = f_\Delta : \mathcal{M} \to L_2$ such that

$$\forall x, y \in \mathcal{M}, \qquad d_\mathcal{M}(x,y) \geqslant \Delta \implies \|f(x) - f(y)\|_{L_2} \geqslant \frac{\Delta}{8\Pi(\mathcal{M}) + 1}. \tag{36}$$

Set $m = |\mathcal{M}|$ and fix $\Delta > 0$. Let $0 < \varepsilon < 1/2$ and $0 < \delta < 1 - 2\varepsilon$ be auxiliary parameters whose values will be specified later to optimize the ensuing reasoning. Fix from now an arbitrary $(\varepsilon\Delta)$-net $\mathcal{N}$ of $\mathcal{M}$.

Let $\{v_\mathcal{S}\}_{\mathcal{S} \subseteq \mathcal{N}} \subseteq L_2$ be an orthonormal system of $2^{|\mathcal{N}|}$ vectors in L_2, indexed by the subsets of $\mathcal{N}$. Denoting by Ω the collection of all the $(\delta\Delta)$-bounded partitions of $\mathcal{N}$, let μ be a probability measure Ω such that

$$\forall a, b \in \mathcal{N}, \qquad \mu\big[\mathcal{P} \in \Omega :\ \mathcal{P}(a) \neq \mathcal{P}(b)\big] \leqslant \frac{\mathsf{SEP}(\mathcal{N})}{\delta\Delta} d_\mathcal{M}(a,b). \tag{37}$$

We can now define a function $\psi : \mathcal{N} \to L_2(\mu; L_2)$ as follows:

$$\forall a \in \mathcal{N}, \forall \mathcal{P} \in \Omega, \qquad \psi(a)(\mathcal{P}) \stackrel{\mathrm{def}}{=} \frac{\Delta\sqrt{\varepsilon\delta}}{\sqrt{2\Pi(\mathcal{M})}} v_{\mathcal{P}(a)}. \tag{38}$$

As $\|v_\mathcal{S} - v_\mathcal{T}\|_{L_2}^2 = 2$ whenever $\mathcal{S}, \mathcal{T} \subseteq \mathcal{N}$ are distinct, every $a, b \in \mathcal{N}$ satisfy

$$\begin{aligned}\|\psi(a) - \psi(b)\|_{L_2(\mu;L_2)} &\stackrel{(38)}{=} \frac{\Delta\sqrt{\varepsilon\delta}}{\sqrt{2\Pi(\mathcal{M})}} \sqrt{2\mu\big[\mathcal{P} \in \Omega :\ \mathcal{P}(a) \neq \mathcal{P}(b)\big]} \\ &\stackrel{(37)}{\leqslant} \frac{\sqrt{\varepsilon\Delta\mathsf{SEP}(\mathcal{N}) d_\mathcal{M}(a,b)}}{\sqrt{\Pi(\mathcal{M})}} \stackrel{(35)}{\leqslant} \frac{\sqrt{\varepsilon\Delta d_\mathcal{M}(a,b)}}{\mathsf{e}(\mathcal{M},\mathcal{N};\ell_2)} \leqslant \frac{d_\mathcal{M}(a,b)}{\mathsf{e}(\mathcal{M},\mathcal{N};\ell_2)},\end{aligned} \tag{39}$$

where the last step of (39) holds as $d_\mathcal{M}(a,b) \geqslant \varepsilon\Delta$ for distinct $a, b \in \mathcal{N}$. By (39), the Lipschitz constant of ψ is at most $1/\mathsf{e}(\mathcal{M},\mathcal{N};\ell_2)$, so there is a 1-Lipschitz function $\Psi : \mathcal{M} \to L_2(\mu; L_2)$ that extends ψ.

Consider $x,y\in\mathcal{M}$ with $d_{\mathcal{M}}(x,y)\geqslant\Delta$. Take $a,b\in\mathcal{N}$ such that $d_{\mathcal{M}}(x,a),d_{\mathcal{M}}(y,b)\leqslant\varepsilon\Delta$. Then,

$$d_{\mathcal{M}}(a,b)\geqslant d_{\mathcal{M}}(x,y)-d_{\mathcal{M}}(x,a)-d_{\mathcal{M}}(y,b)\geqslant(1-2\varepsilon)\Delta>\delta\Delta,\tag{40}$$

where the last step of (40) is where the assumption $\delta<1-2\varepsilon$ is used. Since $\mathrm{diam}_{\mathcal{M}}(\mathcal{P}(a))\leqslant\delta\Delta$ for every $\mathcal{P}\in\Omega$, necessarily $b\notin\mathcal{P}(a)$. In other words, $\mathcal{P}(a)\neq\mathcal{P}(b)$ for every $\mathcal{P}\in\Omega$. Equivalently, every $\mathcal{P}\in\Omega$ satisfies $\|v_{\mathcal{P}(a)}-v_{\mathcal{P}(b)}\|_{L_2}=\sqrt{2}$. Recalling the definition (38) of ψ, because $a,b\in\mathcal{N}$ and Ψ extends ψ we consequently have $\|\Psi(a)-\Psi(b)\|_{L_2(\mu;L_2)}=\|\psi(a)-\psi(b)\|_{L_2(\mu;L_2)}=\Delta\sqrt{\varepsilon\delta/\Pi(\mathcal{M})}$. Therefore,

$$\begin{aligned}\|\Psi(x)-\Psi(y)\|_{L_2(\mu;L_2)}&\geqslant\|\Psi(a)-\Psi(b)\|_{L_2(\mu;L_2)}-\|\Psi(x)-\Psi(a)\|_{L_2(\mu;L_2)}-\|\Psi(y)-\Psi(b)\|_{L_2(\mu;L_2)}\\&\geqslant\frac{\Delta\sqrt{\varepsilon\delta}}{\sqrt{\Pi(\mathcal{M})}}-d_{\mathcal{M}}(x,a)-d_{\mathcal{M}}(y,b)\geqslant\left(\frac{\sqrt{\varepsilon\delta}}{\sqrt{\Pi(\mathcal{M})}}-2\varepsilon\right)\Delta,\end{aligned}\tag{41}$$

where the penultimate step of (41) uses the fact that Ψ is 1-Lipschitz.

The right hand side of (41) is maximal for $\varepsilon=\varepsilon_{\mathrm{opt}}\stackrel{\mathrm{def}}{=}\delta/(16\Pi(\mathcal{M}))$, which is a valid choice for the above reasoning provided that the requirement $\delta<1-2\varepsilon_{\mathrm{opt}}$ is satisfied, i.e., if $\delta<8\Pi(\mathcal{M})/(8\Pi(\mathcal{M})+1)$. Assuming from now on that this restriction on δ holds, substitute $\varepsilon_{\mathrm{opt}}$ into (41). Since Ψ takes values in an m-dimensional subspace of a Hilbert space, by composing Ψ with an isometry between that subspace and ℓ_2^m and translating the resulting function so that $0\in\ell_2^m$ is in its image, we conclude that for every $0<\delta<8\Pi(\mathcal{M})/(8\Pi(\mathcal{M})+1)$ there is a 1-Lipschitz function $\Phi_\delta:\mathcal{M}\to B_{\ell_2^m}\big(0,\mathrm{diam}(\mathcal{M})\big)$ that satisfies:

$$\forall\, x,y\in\mathcal{M},\qquad d_{\mathcal{M}}(x,y)\geqslant\Delta\implies\|\Phi_\delta(x)-\Phi_\delta(y)\|_{\ell_2^m}\geqslant\frac{\delta\Delta}{8D(\Pi(\mathcal{M})}.$$

Now (36) follows by taking a $\delta\to\big(8\Pi(\mathcal{M})/(8\Pi(\mathcal{M})+1)\big)^-$ limit of Φ_δ along a convergent subsequence. $\square$

2.2. **Strongly bi-Lipschitz from weakly bi-Lipschitz.** Here we will prove the following embedding result:

Theorem 22. *For every $n\in\{3,4,\ldots\}$, every n-point metric space $(\mathcal{M},d_{\mathcal{M}})$ satisfies*

$$\mathsf{c}_2(\mathcal{M})\lesssim\max\left\{\sqrt{\log n}\left(\sum_{k=2}^{n-1}\frac{\widehat{\mathsf{d}}_2^k(\mathcal{M})^2}{k(\log k)^2}\right)^{\frac12},\mathsf{d}_2(\mathcal{M})\right\}\sqrt{\log\log n}.\tag{42}$$

Thanks to Lemma 21, Theorem 22 implies Theorem 15. In fact, this yields the following estimate:

$$\begin{aligned}\mathsf{c}_2(\mathcal{M})&\lesssim\max\left\{\sqrt{\log n}\left(\sum_{k=2}^{n-1}\frac{\Pi^k(\mathcal{M})^2}{k(\log k)^2}\right)^{\frac12},\Pi(\mathcal{M})\right\}\sqrt{\log\log n}\\&\stackrel{(35)}{\leqslant}\max\left\{\sqrt{\log n}\left(\sum_{k=2}^{n-1}\frac{\mathsf{SEP}^k(\mathcal{M})^2\mathsf{e}_{\mathrm{net}}^k(\mathcal{M};L_2)^4}{k(\log k)^2}\right)^{\frac12},\mathsf{SEP}(\mathcal{M})^2\mathsf{e}_{\mathrm{net}}(\mathcal{M};L_2)^4\right\}\sqrt{\log\log n},\end{aligned}\tag{43}$$

which is a strengthening of (22), where (43) uses the following notation:

$$\forall\, k\in\mathbb{N},\qquad\Pi^k(\mathcal{M})\stackrel{\mathrm{def}}{=}\sup_{\substack{\emptyset\neq\mathcal{C}\subseteq\mathcal{M}\\|\mathcal{C}|\leqslant k}}\Pi(\mathcal{C})\qquad\text{and}\qquad\mathsf{e}_{\mathrm{net}}^k(\mathcal{M};L_2)\stackrel{\mathrm{def}}{=}\sup_{\substack{\emptyset\neq\mathcal{C}\subseteq\mathcal{M}\\|\mathcal{C}|\leqslant k}}\mathsf{e}_{\mathrm{net}}(\mathcal{C};L_2).$$

(This adheres to the notational convention that we have been maintaining throughout by which a superscript indicates the hereditary version of an invariant of metric spaces which is obtained by considering the subset of the metric space of a given size for which the invariant in question is maximal.)

For the sole purpose of proving Theorem 3 (as a consequence of our Theorem 10 and the Lipschitz extension theorem of [81]), one could use [5] as a "black box" rather than using Theorem 22. Specifically, Theorem 3 follows by fixing $p>2$, using Theorem 10 and the estimate $\mathsf{e}(L_p;L_2)\lesssim\sqrt{p}$ of [81] to deduce that $\Pi^k(L_p)\lesssim p^3\sqrt{\log k}$ for every integer $k\geqslant 2$, whence $\mathsf{d}_2^k(L_p)\lesssim p^3\sqrt{\log k}$ by Lemma 21, and now a substitution of this conclusion into [5, Theorem 4.1] yields Theorem 3.

Thus, those who wish to verify Theorem 3 while appealing to the published literature can stop reading Section 2 here, and continue reading the present article from Section 3 onwards; that material does not rely on the contents of the rest of Section 2.

Notwithstanding the above discussion, it is worthwhile to revisit the ideas of [5] so as to derive Theorem 22 because it is an independently interesting geometric statement that did not appear elsewhere. In comparison to the reasoning in [5], the ensuing proof of Theorem 22 is more modular and its steps are more general and flexible, so they could be of use for further investigations elsewhere.

Observe that Theorem 22 implies the following embedding result:

$$\forall \frac{1}{2}\leqslant \theta\leqslant 1,\ \forall C\geqslant 1,\qquad \max_{k\in\{2,\ldots,n\}}\frac{\widehat{\mathsf{d}}_2^k(\mathcal{M})}{(\log k)^\theta}\leqslant C \implies \mathsf{c}_2(\mathcal{M})\lesssim C(\log n)^\theta \min\left\{\frac{\sqrt{\log\log n}}{\sqrt{2\theta-1}},\log\log n\right\}. \tag{44}$$

The special case $\theta=1/2$ of (44) coincides with the case $\varepsilon=1/2$ of [5, Theorem 4.1]. However, if $1/2<\theta\leqslant 1$ is independent of n or more generally if $\lim_{n\to\infty}(\theta-1/2)\log\log n=\infty$, then (44) improves asymptotically over [5, Theorem 4.1], though this could also be deduced by a careful inspection of the proof in [5].

In fact, the ensuing justification of Theorem 22 is an involved but quite direct refinement and generalization of the reasoning in [5]; it is also not entirely self-contained because it uses [5, Theorem 4.5] as a "black box," since we do not have anything novel to add to [5, Theorem 4.5] (there are enhancements of [5, Theorem 4.5] that appeared in [4, 22], but those are not needed for the present purposes).

2.2.1. *Weakly bi-Lipschitz embeddings of controlled local growth centers.* A key part of the proof of Theorem 22 studies the geometry of the set of points in a metric space which are centers of balls whose size increases in a controlled manner for radii that belong to a prescribed range. The pertinent object is:

Notation 23 (controlled local growth centers)**.** *Suppose that $(\mathcal{M},d_{\mathcal{M}})$ is a locally finite metric space (thus, all balls in $\mathcal{M}$ are finite). For $K\geqslant 1$ and $R\geqslant r\geqslant 0$, let $\mathcal{G}_{\leqslant K}(r,R)$ denote the set of centers in $\mathcal{M}$ at which the growth rate of balls from radius r to radius R is at most K, i.e.,*

$$\mathcal{G}_{\leqslant K}(r,R)=\mathcal{G}_{\leqslant K}^{d_{\mathcal{M}}}(r,R)\stackrel{\mathrm{def}}{=}\left\{x\in\mathcal{M}:\ \frac{|B_{\mathcal{M}}(x,R)|}{|B_{\mathcal{M}}(x,r)|}\leqslant K\right\}. \tag{45}$$

Note in passing that the sets of Notation 23 obey the following immediate inclusions:

$$\forall 1\leqslant K\leqslant K',\ \forall 0\leqslant r\leqslant r'\leqslant R'\leqslant R,\qquad \mathcal{G}_{\leqslant K}(r,R)\subseteq \mathcal{G}_{\leqslant K'}(r',R'). \tag{46}$$

The following theorem derives favorable Euclidean embedding properties of the sets of controlled local growth centers from Definition 23, which refine and generalize results that were obtained in [5]:

Theorem 24. *Fix an integer $n\geqslant 3$ and supposed that $(\mathcal{M},d_{\mathcal{M}})$ is an n-point metric space. Then,*

- *There is a 1-Lipschitz function $\psi:\mathcal{M}\to L_2$ satisfying the following for every $0<r<R$ and $K\geqslant 1$:*

$$\forall (x,y)\in \mathcal{G}_{\leqslant K}(r,R)\times\mathcal{M},\qquad d_{\mathcal{M}}(x,y)>\frac{1}{2}r+\frac{3}{2}R \implies \|\psi(x)-\psi(y)\|_{L_2}\gtrsim \frac{R-r}{\sqrt{K\log n}}. \tag{47}$$

- *There is a universal constant $C\geqslant 1$ with the following property. Fix $K,D,\beta\geqslant 1$. Suppose that for any $\mathcal{C}\subseteq\mathcal{M}$ with $|\mathcal{C}|\leqslant K$ and any $\Delta>0$ there is a 1-Lipschitz function $f=f_{\mathcal{C},\Delta}:\mathcal{M}\to L_2$ satisfying:*

$$\forall x,y\in\mathcal{C},\qquad \Delta\leqslant d_{\mathcal{M}}(x,y)\leqslant 3\beta\Delta \implies \|f(x)-f(y)\|_{L_2}\geqslant \frac{\Delta}{D}. \tag{48}$$

  Then, for any $\Delta>0$ and any $R>r>0$ that satisfy the restrictions

$$\Delta\geqslant 9Dr\qquad\text{and}\qquad R-r\geqslant C\beta\big(\log|\mathcal{G}_{\leqslant K}(r,R)|\big)\Delta, \tag{49}$$

  there is a 1-Lipschitz function $\phi=\phi_{\Delta,R,r}:\mathcal{M}\to L_2$ for which the following property holds:

$$\forall (x,y)\in \mathcal{G}_{\leqslant K}(r,R)\times\mathcal{M},\qquad \Delta\leqslant d_{\mathcal{M}}(x,y)\leqslant \beta\Delta \implies \|\phi(x)-\phi(y)\|_{L_2}\gtrsim \frac{\Delta}{D}. \tag{50}$$

Prior to proving Theorem 24 we will next explain how it implies Theorem 22.

Proof of Theorem 22 assuming Theorem 24. Let $C \geqslant 1$ be the universal constant from (the second part of) Theorem 24. For the ensuing reasoning, it will be convenient to set the following notation:

$$\forall\, i\in\mathbb{Z},\ \forall\, k\in\{1,\ldots,n\},\qquad r_i^k \stackrel{\mathrm{def}}{=} \frac{2^i}{27\widehat{\mathsf{d}}_2^k(\mathcal{M})}\qquad\text{and}\qquad R_i^n \stackrel{\mathrm{def}}{=} 3C(\log n)2^i. \tag{51}$$

Fix $k\in\{1,\ldots,n\}$. By the definition of $\widehat{\mathsf{d}}_2^k(\mathcal{M})$ and Remark 20 for $\beta=6$ and $\alpha=2$, for every subset $\mathcal{C}\subseteq\mathcal{M}$ with $2\leqslant|\mathcal{C}|\leqslant k$ and every $\Delta>0$ there exists a 1-Lipschitz function $f=f_{\mathcal{C},\Delta}:\mathcal{M}\to L_2$ such that

$$\forall\, x,y\in\mathcal{C},\qquad \Delta\leqslant d_{\mathcal{M}}(x,y)\leqslant 6\Delta \implies \|f(x)-f(y)\|_{L_2}\geqslant \frac{\Delta}{3\widehat{\mathsf{d}}_2^k(\mathcal{M})}.$$

The assumption of the second part of Theorem 24 therefore holds with the following parameters:

$$\beta=2\qquad\text{and}\qquad K=k\qquad\text{and}\qquad D=3\widehat{\mathsf{d}}_2^k(\mathcal{M}). \tag{52}$$

Consequently, for every $i\in\mathbb{Z}$ there exists a 1-Lipschitz function $\phi_i^k:\mathcal{M}\to L_2$ such that

$$\forall (x,y)\in \mathcal{G}_{\leqslant k}\big(r_i^k,R_i^n\big)\times\mathcal{M},\qquad 2^i\leqslant d_{\mathcal{M}}(x,y)\leqslant 2^{i+1}\implies \|\phi_i^k(x)-\phi_i^k(y)\|_{L_2}\gtrsim r_i^k, \tag{53}$$

This is indeed a valid application of Theorem 24 for the parameter choices

$$\Delta=27\widehat{\mathsf{d}}_2^k(\mathcal{M})2^i\qquad\text{and}\qquad r=r_i^k\qquad\text{and}\qquad R=R_i^n, \tag{54}$$

because $\Delta=9Dr$ by (52) and (54), and furthermore since $C\geqslant 1$ and $|\mathcal{G}_{\leqslant K}(r,R)|\leqslant|\mathcal{M}|=n$ we have

$$R-r\stackrel{(51)}{\geqslant} 2C(\log n)2^i\geqslant 2C\big(\log|\mathcal{G}_{\leqslant K}(r,R)|\big)2^i \stackrel{(52)\wedge(54)}{=} \beta C\big(\log|\mathcal{G}_{\leqslant K}(r,R)|\big)\Delta.$$

Apply [5, Theorem 4.5] to $\{\phi_i^k\}_{k\in\mathbb{Z}}$ with the parameters $A=3C(\log n)/2\geqslant 1$ and $B=27\widehat{\mathsf{d}}_2^k(\mathcal{M})\lesssim\log n$, while using $B\lesssim\log n$, which holds by [16]. This yields $\varphi^k:\mathcal{M}\to L_2$ that has the following two properties:

$$\|\varphi^k\|_{\mathrm{Lip}(\mathcal{M};L_2)}\lesssim\sqrt{(\log n)\log\log n}, \tag{55}$$

and

$$\forall\, x,y\in\mathcal{M},\ \forall\, i\in\mathbb{Z},\qquad \|\varphi^k(x)-\varphi^k(y)\|_{L_2}^2\gtrsim\left\lfloor\log_2\frac{|B_{\mathcal{M}}(x,R_i^n)|}{|B_{\mathcal{M}}(x,r_i^k)|}\right\rfloor\min\left\{\frac{2^{2i}}{\widehat{\mathsf{d}}_2^k(\mathcal{M})^2},\|\phi_i^k(x)-\phi_i^k(y)\|_{L_2}^2\right\}. \tag{56}$$

For every distinct $x,y\in\mathcal{M}$ define $i(x,y)\in\mathbb{Z}$ by the requirement $2^{i(x,y)}\leqslant d_{\mathcal{M}}(x,y)<2^{i(x,y)+1}$, i.e.,

$$i(x,y)\stackrel{\mathrm{def}}{=}\big\lfloor\log_2 d_{\mathcal{M}}(x,y)\big\rfloor. \tag{57}$$

We will also set $i(x,x)=-\infty$ for every $x\in\mathcal{M}$. As $|\mathcal{M}|=n$ the definition (45) of $\mathcal{G}_{\leqslant n}(r,R)$ implies that $\mathcal{G}_{\leqslant n}(r,R)=\mathcal{M}$ for every $R\geqslant r\geqslant 0$. Hence, the following index $\mathscr{k}(x,y)\in\{1,\ldots,n\}$ is well-defined:

$$\mathscr{k}(x,y)\stackrel{\mathrm{def}}{=}\min\left\{k\in\{1,\ldots,n\}:\ x\in\bigcap_{\ell=k}^{n}\mathcal{G}_{\leqslant\ell}\big(r_{i(x,y)}^{\ell},R_{i(x,y)}^{n}\big)\right\}. \tag{58}$$

For every $x\in\mathcal{M}$, the above convention $i(x,x)=-\infty$ is consistent with setting $\mathscr{k}(x,x)=1$. We calim that

$$\forall\, x,y\in\mathcal{M},\ \forall\, k\in\{\mathscr{k}(x,y),\mathscr{k}(x,y)+1,\ldots,n\},\qquad \|\varphi^k(x)-\varphi^k(y)\|_{L_2}\gtrsim\frac{\sqrt{\lfloor\log\mathscr{k}(x,y)\rfloor}}{\widehat{\mathsf{d}}_2^k(\mathcal{M})}d_{\mathcal{M}}(x,y). \tag{59}$$

To justify (59), fix $x,y\in\mathcal{M}$ and $k\in\{\mathscr{k}(x,y),\mathscr{k}(x,y)+1,\ldots,n\}$. If $\mathscr{k}(x,y)\in\{1,2\}$, then the right hand side of the inequality in (59) vanishes, so (59) holds vacuously. We can therefore assume that $\mathscr{k}(x,y)\in\{3,4,\ldots,n\}$. As $k\geqslant\mathscr{k}(x,y)$, the definition (58) of $\mathscr{k}(x,y)$ and the definition (57) of $i(x,y)$ show that

$$x\in\mathcal{G}_{\leqslant k}\big(r_{i(x,y)}^k,R_{i(x,y)}^n\big)\qquad\text{and}\qquad 2^{i(x,y)}\leqslant d_{\mathcal{M}}(x,y)<2^{i(x,y)+1}. \tag{60}$$

Consequently,

$$\|\phi_{i(x,y)}^k(x)-\phi_{i(x,y)}^k(y)\|_{L_2}\stackrel{(60)\wedge(53)}{\gtrsim} r_{i(x,y)}^k\stackrel{(53)}{\asymp}\frac{d_{\mathcal{M}}(x,y)}{\widehat{\mathsf{d}}_2^k(\mathcal{M})}. \tag{61}$$

At the same time, by the minimality of $\mathscr{k}(x,y)$ per (58), we have

$$x \notin \mathcal{G}_{\leqslant \mathscr{k}(x,y)-1}\big(r_{i(x,y)}^{\mathscr{k}(x,y)-1}, R_{i(x,y)}^{n}\big),$$

which by (45) is equivalent to the following lower bound on the growth rate of balls centered at x:

$$\frac{\big|B_{\mathcal{M}}\big(x, R_{i(x,y)}^{n}\big)\big|}{\big|B_{\mathcal{M}}\big(x, r_{i(x,y)}^{\mathscr{k}(x,y)-1}\big)\big|} > \mathscr{k}(x,y)-1. \tag{62}$$

Since $\widehat{\mathsf{d}}_2^{k}(\mathcal{M}) \geqslant \widehat{\mathsf{d}}_2^{\mathscr{k}(x,y)-1}(\mathcal{M})$ we have $r_{i(x,y)}^{k} \leqslant r_{i(x,y)}^{\mathscr{k}(x,y)-1}$ by (51), and therefore

$$\big|B_{\mathcal{M}}\big(x, r_{i(x,y)}^{k}\big)\big| \leqslant \big|B_{\mathcal{M}}\big(x, r_{i(x,y)}^{\mathscr{k}(x,y)-1}\big)\big|. \tag{63}$$

Consequently,

$$\left\lfloor \log_2 \frac{\big|B_{\mathcal{M}}\big(x, R_{i(x,y)}^{n}\big)\big|}{\big|B_{\mathcal{M}}\big(x, r_{i(x,y)}^{k}\big)\big|}\right\rfloor \overset{(63)}{\geqslant} \left\lfloor \log_2 \frac{\big|B_{\mathcal{M}}\big(x, R_{i(x,y)}^{n}\big)\big|}{\big|B_{\mathcal{M}}\big(x, r_{i(x,y)}^{\mathscr{k}(x,y)-1}\big)\big|}\right\rfloor \overset{(62)}{\geqslant} \big\lfloor \log_2(\mathscr{k}(x,y)-1)\big\rfloor \asymp \log \mathscr{k}(x,y), \tag{64}$$

where the last step of (64) is valid because we are now working under the assumption $\mathscr{k}(x,y) \geqslant 3$. The desired assertion (59) now follows by substituting (61) and (64) into (56) while using $2^{i(x,y)} \asymp d_{\mathcal{M}}(x,y)$.

Next, observe that the function $\psi : \mathcal{M} \to L_2$ from the first part of Theorem 24 satisfies:

$$\|\psi\|_{\mathrm{Lip}(\mathcal{M};L_2)} \leqslant 1 \qquad \text{and} \qquad \forall\, x,y \in \mathcal{M}, \qquad \|\psi(x)-\psi(y)\|_{L_2} \gtrsim \frac{1}{\sqrt{\mathscr{k}(x,y)\log n}}\, d_{\mathcal{M}}(x,y). \tag{65}$$

Indeed, the first part of (65) is the upper bound on the Lipschitz constant of ψ from statement of Theorem 24. For the second part of (65), fix distinct $x,y \in \mathcal{M}$ and record the following (very) crude bounds:

$$R_{i(x,y)}^{n} \overset{(51)}{>} 2^{i(x,y)-1} \qquad \text{and} \qquad d_{\mathcal{M}}(x,y) \overset{(57)}{\geqslant} 2^{i(x,y)} \overset{(51)}{>} \frac{1}{2} r_{i(x,y)}^{\mathscr{k}(x,y)} + \frac{3}{2} 2^{i(x,y)-1}. \tag{66}$$

Thus,

$$x \in \mathcal{G}_{\leqslant \mathscr{k}(x,y)}\big(r_{i(x,y)}^{\mathscr{k}(x,y)}, R_{i(x,y)}^{n}\big) \subseteq \mathcal{G}_{\leqslant \mathscr{k}(x,y)}\big(r_{i(x,y)}^{\mathscr{k}(x,y)}, 2^{i(x,y)-1}\big), \tag{67}$$

where the first step of (67) follows from the case $k = \mathscr{k}(x,y)$ of (60) and the second step of (67) follows from the first part of (66) and the (trivial) monotonicity property (46) of the sets of controlled local growth centers. We can thereofre complete the justification of the second part of (65) as follows:

$$\|\psi(x)-\psi(y)\|_{L_2} \gtrsim \frac{2^{i(x,y)-1} - r_{i(x,y)}^{\mathscr{k}(x,y)}}{\sqrt{\mathscr{k}(x,y)\log n}} \asymp \frac{2^{i(x,y)-1}}{\sqrt{\mathscr{k}(x,y)\log n}} \asymp \frac{d_{\mathcal{M}}(x,y)}{\sqrt{\mathscr{k}(x,y)\log n}}, \tag{68}$$

where the first step of (68) uses the first part (47) of Theorem 24 with $r = r_{i(x,y)}^{\mathscr{k}(x,y)}$, $R = 2^{i(x,y)-1}$, $K = \mathscr{k}(x,y)$, which is valid thanks to (67) and the second part of (66).

If $v_1,\ldots,v_n$ is an orthonormal system in L_2, then define $f : \mathcal{M} \to L_2 \otimes L_2 \cong L_2$ by setting

$$\forall\, x \in \mathcal{M}, \qquad f(x) \stackrel{\mathrm{def}}{=} \sqrt{(\log n)\log\log n}\cdot \psi(x)\otimes v_1 + \sum_{k=2}^{n-1} \frac{\widehat{\mathsf{d}}_2^{k}(\mathcal{M})}{\sqrt{k}\log k}\varphi^{k}(x)\otimes v_k + \frac{\widehat{\mathsf{d}}_2(\mathcal{M})}{\sqrt{\log n}}\varphi^{n}(x)\otimes v_n. \tag{69}$$

The orthonormality of $v_1,\ldots,v_n$ and the the bounds on the Lipschitz constants in (55) and (65) give

$$\|f\|_{\mathrm{Lip}(\mathcal{M};L_2\otimes L_2)} \lesssim \max\left\{\sqrt{\log n}\bigg(\sum_{k=2}^{n-1} \frac{\widehat{\mathsf{d}}_2^{k}(\mathcal{M})^2}{k(\log k)^2}\bigg)^{\frac12}, \mathsf{d}_2(\mathcal{M})\right\}\sqrt{\log\log n}. \tag{70}$$

Furthermore, every distinct $x, y \in \mathcal{M}$ satisfy the following lower bound:

$$\begin{aligned}\frac{\|f(x)-f(y)\|_{L_2\otimes L_2}}{d_{\mathcal{M}}(x,y)} &\overset{(69)\wedge(59)\wedge(65)}{\gtrsim} \left(\frac{\log\log n}{\hslash(x,y)} + \lfloor \log \hslash(x,y)\rfloor \sum_{k=\hslash(x,y)}^{n-1} \frac{1}{k(\log k)^2} + \frac{\lfloor \log \hslash(x,y)\rfloor}{\log n}\right)^{\frac12} \\ &\geqslant \left(\frac{\log\log n}{\hslash(x,y)} + \lfloor \log \hslash(x,y)\rfloor \int_{\max\{\hslash(x,y),2\}}^{n} \frac{\mathrm{d}s}{s(\log s)^2} + \frac{\lfloor \log \hslash(x,y)\rfloor}{\log n}\right)^{\frac12} \\ &= \left(\frac{\log\log n}{\hslash(x,y)} + \lfloor \log \hslash(x,y)\rfloor \left(\frac{1}{\log\max\{\hslash(x,y),2\}} - \frac{1}{\log n}\right) + \frac{\lfloor \log \hslash(x,y)\rfloor}{\log n}\right)^{\frac12} \gtrsim 1,\end{aligned} \tag{71}$$

where the last step of (71) is verified by considering separately the cases $1 \leqslant \hslash(x,y) \leqslant \max\{\log\log n, 3\}$, $\max\{\log\log n, 3\} \leqslant \hslash(x,y) \leqslant \sqrt{n}$ and $\sqrt{n} \leqslant \hslash(x,y) \leqslant n$. The desired distortion bound (42) of Theorem 22 follows by combining (70) and (71) for the embedding f of $\mathcal{M}$ into Hilbert space. $\square$

Passing to the proof of Theorem 24, we will first justify its first part by following the reasoning in the proof of [5, Claim 4.6] as well as its predecessors [68, Lemma 3.4], [1, Theorem 2] and [25, Lemma 2.9], all of which are an elaboration of the important embedding technique that was introduced in [16].

Given a metric space $(\mathcal{M}, d_{\mathcal{M}})$, the distance of a point $x \in \mathcal{M}$ to a nonempty subset $\mathcal{C}$ of $\mathcal{M}$ is denoted as usual by $d_{\mathcal{M}}(x, \mathcal{C}) = \inf\{d_{\mathcal{M}}(x,y):\ y \in \mathcal{C}\}$. It is worthwhile to state separately the following observation:

Observation 25. *Let $(\mathcal{M}, d_{\mathcal{M}})$ be a finite metric space. Given $0 \leqslant \mathfrak{p} \leqslant 1$, denote by $\mathcal{Z}_{\mathfrak{p}}$ the ($\mathfrak{p}$-Bernoulli) random subset of $\mathcal{M}$ that is obtained by including independently each $x \in \mathcal{M}$ in $\mathcal{Z}$ with probability $\mathfrak{p}$, i.e.,*

$$\forall \mathcal{C} \subseteq \mathcal{M}, \qquad \mathbb{P}[\mathcal{Z}_{\mathfrak{p}} = \mathcal{C}] = \mathfrak{p}^{|\mathcal{C}|}(1-\mathfrak{p})^{n-|\mathcal{C}|}. \tag{72}$$

Then, for every $0 < r < R$ and every $x, y \in \mathcal{M}$ such that $d_{\mathcal{M}}(x,y) > \frac12 r + \frac32 R$, we have

$$\mathbb{P}\left[\mathcal{Z}_{\mathfrak{p}} \neq \emptyset \quad \text{and} \quad |d_{\mathcal{M}}(x,\mathcal{Z}_{\mathfrak{p}}) - d_{\mathcal{M}}(y,\mathcal{Z}_{\mathfrak{p}})| > \frac{R-r}{2}\right] \geqslant \min\left\{(1-\mathfrak{p})^{|B_{\mathcal{M}}(x,R)|}, 1-(1-\mathfrak{p})^{|B_{\mathcal{M}}(x,r)|}\right\}. \tag{73}$$

Proof. Consider the events E and F that are defined as follows:

$$\begin{cases} E \overset{\mathrm{def}}{=} \left\{d_{\mathcal{M}}(x,\mathcal{Z}_{\mathfrak{p}}) > R \quad \text{and} \quad d_{\mathcal{M}}(y,\mathcal{Z}_{\mathfrak{p}}) \leqslant \frac{r+R}{2}\right\}, \\ F \overset{\mathrm{def}}{=} \left\{d_{\mathcal{M}}(x,\mathcal{Z}_{\mathfrak{p}}) \leqslant r \quad \text{and} \quad d_{\mathcal{M}}(y,\mathcal{Z}_{\mathfrak{p}}) > \frac{r+R}{2}\right\}. \end{cases} \tag{74}$$

By definition, $E \cap F = \emptyset$ and $E \cup F \subseteq \{\mathcal{Z}_{\mathfrak{p}} \neq \emptyset \ \wedge\ |d_{\mathcal{M}}(x,\mathcal{Z}_{\mathfrak{p}}) - d_{\mathcal{M}}(y,\mathcal{Z}_{\mathfrak{p}})| > (R-r)/2\}$, so we have

$$\mathbb{P}\left[\mathcal{Z}_{\mathfrak{p}} \neq \emptyset \quad \text{and} \quad |d_{\mathcal{M}}(x,\mathcal{Z}_{\mathfrak{p}}) - d_{\mathcal{M}}(y,\mathcal{Z}_{\mathfrak{p}})| > \frac{R-r}{2}\right] \geqslant \mathbb{P}[E] + \mathbb{P}[F]. \tag{75}$$

Note that the event $\{d_{\mathcal{M}}(x,\mathcal{Z}_{\mathfrak{p}}) > R\}$ occurs if and only if $\mathcal{Z}_{\mathfrak{p}}$ contains no point from $B_{\mathcal{M}}(x,R)$, and the event $\{d_{\mathcal{M}}(y,\mathcal{Z}_{\mathfrak{p}}) \leqslant (r+R)/2\}$ occurs if and only if $\mathcal{Z}_{\mathfrak{p}}$ contains at least one point from $B_{\mathcal{M}}(y,(r+R)/2)$. By the triangle inequality for $d_{\mathcal{M}}$, the assumption $d_{\mathcal{M}}(x,y) > r/2+3R/2 = R+(r+R)/2$ of Observation 25 implies that $B_{\mathcal{M}}(x,R) \cap B_{\mathcal{M}}(y,(r+R)/2) = \emptyset$, so $\{d_{\mathcal{M}}(x,\mathcal{Z}_{\mathfrak{p}}) > R\}$ and $\{d_{\mathcal{M}}(y,\mathcal{Z}_{\mathfrak{p}}) \leqslant (r+R)/2\}$ are independent events, thanks to the definition of $\mathcal{Z}_{\mathfrak{p}}$. Consequently,

$$\mathbb{P}[E] \overset{(74)}{=} \mathbb{P}\left[\mathcal{Z}_{\mathfrak{p}} \cap B_{\mathcal{M}}(x,R) = \emptyset\right] \mathbb{P}\left[d_{\mathcal{M}}(y,\mathcal{Z}_{\mathfrak{p}}) \leqslant \frac{r+R}{2}\right] \overset{(72)}{=} (1-\mathfrak{p})^{|B_{\mathcal{M}}(x,R)|}\, \mathbb{P}\left[d_{\mathcal{M}}(y,\mathcal{Z}_{\mathfrak{p}}) \leqslant \frac{r+R}{2}\right]. \tag{76}$$

In the same vein, the events $\{d_{\mathcal{M}}(x,\mathcal{Z}_{\mathfrak{p}}) \leqslant r\}$ and $\{d_{\mathcal{M}}(y,\mathcal{Z}_{\mathfrak{p}}) > (r+R)/2\}$ coincide with, respectively, the events $\{\mathcal{Z}_{\mathfrak{p}} \cap B_{\mathcal{M}}(x,r) \neq \emptyset\}$ and $\{\mathcal{Z}_{\mathfrak{p}} \cap B_{\mathcal{M}}(y,(r+R)/2) = \emptyset\}$. The latter two events are independent because $B_{\mathcal{M}}(x,r) \cap B_{\mathcal{M}}(y,(r+R)/2) \subseteq B_{\mathcal{M}}(x,R) \cap B_{\mathcal{M}}(y,(r+R)/2) = \emptyset$, so we have

$$\mathbb{P}[F] \overset{(74)}{=} \mathbb{P}\left[\mathcal{Z}_{\mathfrak{p}} \cap B_{\mathcal{M}}(x,r) \neq \emptyset\right] \mathbb{P}\left[d_{\mathcal{M}}(y,\mathcal{Z}_{\mathfrak{p}}) > \frac{r+R}{2}\right] \overset{(72)}{=} \left(1-(1-\mathfrak{p})^{|B_{\mathcal{M}}(x,r)|}\right) \mathbb{P}\left[d_{\mathcal{M}}(y,\mathcal{Z}_{\mathfrak{p}}) > \frac{r+R}{2}\right]. \tag{77}$$

The desired bound (73) is now justified as follows:

$$
\begin{aligned}
&\mathbb{P}\left[\mathscr{Z}_{\mathfrak{p}}\neq\emptyset \quad \text{and} \quad |d_{\mathcal{M}}(x,\mathscr{Z}_{\mathfrak{p}})-d_{\mathcal{M}}(y,\mathscr{Z}_{\mathfrak{p}})|>\frac{R-r}{2}\right]\\
&\overset{(75)\wedge(76)\wedge(77)}{\geqslant} (1-\mathfrak{p})^{|B_{\mathcal{M}}(x,R)|}\,\mathbb{P}\left[d_{\mathcal{M}}(y,\mathscr{Z}_{\mathfrak{p}})\leqslant\frac{r+R}{2}\right]+\left(1-(1-\mathfrak{p})^{|B_{\mathcal{M}}(x,r)|}\right)\mathbb{P}\left[d_{\mathcal{M}}(y,\mathscr{Z}_{\mathfrak{p}})>\frac{r+R}{2}\right]\\
&\geqslant \min\left\{(1-\mathfrak{p})^{|B_{\mathcal{M}}(x,R)|},1-(1-\mathfrak{p})^{|B_{\mathcal{M}}(x,r)|}\right\}\left(\mathbb{P}\left[d_{\mathcal{M}}(y,\mathscr{Z}_{\mathfrak{p}})\leqslant\frac{r+R}{2}\right]+\mathbb{P}\left[d_{\mathcal{M}}(y,\mathscr{Z}_{\mathfrak{p}})>\frac{r+R}{2}\right]\right)\\
&= \min\left\{(1-\mathfrak{p})^{|B_{\mathcal{M}}(x,R)|},1-(1-\mathfrak{p})^{|B_{\mathcal{M}}(x,r)|}\right\}. \qquad \square
\end{aligned}
$$

For the of proof of the first part of Theorem 24, it will be notationally beneficial to use the convention that the distance of a point $x\in\mathcal{M}$ of a metric space $(\mathcal{M},d_{\mathcal{M}})$ to the empty set equals ∞ (in our setting $\mathcal{M}$ is finite, so the same purpose will be served by defining $d_{\mathcal{M}}(x,\emptyset)$ to be any fixed number that is strictly larger than the diameter $\mathrm{diam}(\mathcal{M})=\sup\{d_{\mathcal{M}}(y,z):\ y,z\in\mathcal{M}\}$ of $\mathcal{M}$, say, $d_{\mathcal{M}}(x,\emptyset)=\mathrm{diam}(\mathcal{M})+1$).

Proof of the first part of Theorem 24. For each $i\in\{1,\ldots,\lceil\log n\rceil\}$ let $\mathscr{Z}_{e^{-i}}$ be the random subset of $\mathcal{M}$ from Observation 25 with $\mathfrak{p}=2^{-i}$. Denote the probability space on which these sets are defined by $(\Omega,\mathbb{P})$; it can be the corresponding product space, though they need not be independent for the ensuing reasoning. Let $v_1,v_2,\ldots$ be an orthonormal basis of L_2 and define a 1-Lipschitz function $f:\mathcal{M}\to L_2(\mathbb{P};L_2)\cong L_2$ by

$$
\forall (x,\omega)\in\mathcal{M}\times\Omega,\qquad f(x)(\omega)\stackrel{\mathrm{def}}{=}\frac{1}{\sqrt{\lceil\log n\rceil}}\sum_{i=1}^{\lceil\log n\rceil}\min\left\{d_{\mathcal{M}}\big(x,\mathscr{Z}_{e^{-i}}(\omega)\big),\mathrm{diam}(\mathcal{M})\right\}v_i. \tag{78}
$$

For every $x\in\mathcal{M}$ and $R>0$ let $i_R(x)\in\{1,\ldots,\lceil\log n\rceil\}$ be such that $e^{i_R(x)-1}\leqslant|B_{\mathcal{M}}(x,R)|<e^{i_R(x)}$. Then,

$$
\forall K\geqslant1,\ \forall 0<r<R,\ \forall x\in\mathcal{G}_{\leqslant K}(r,R),\qquad e^{i_R(x)}>|B_{\mathcal{M}}(x,R)|\geqslant|B_{\mathcal{M}}(x,r)|\overset{(45)}{\geqslant}\frac{|B_{\mathcal{M}}(x,R)|}{K}\geqslant\frac{e^{i_R(x)-1}}{K}. \tag{79}
$$

Using Observation 25, for every $x\in\mathcal{G}_{\leqslant K}(r,R)$ and every $y\in\mathcal{M}$ with $d_{\mathcal{M}}(x,y)>r/2+3R/2$ we have

$$
\begin{aligned}
&\mathbb{P}\left[\left|\min\left\{d_{\mathcal{M}}\left(x,\mathscr{Z}_{e^{-i_R(x)}}\right),\mathrm{diam}(\mathcal{M})\right\}-\min\left\{d_{\mathcal{M}}\left(y,\mathscr{Z}_{e^{-i_R(x)}}\right),\mathrm{diam}(\mathcal{M})\right\}\right|>\frac{R-r}{2}\right]\\
&\qquad\qquad\overset{(73)\wedge(79)}{\geqslant}\min\left\{\left(1-e^{-i_R(x)}\right)^{e^{i_R(x)}},1-\left(1-e^{-i_R(x)}\right)^{\frac{e^{i_R(x)-1}}{K}}\right\}\asymp\frac{1}{K}.
\end{aligned}\tag{80}
$$

This directly implies the desired estimate (47) as follows:

$$
\begin{aligned}
&\|f(x)-f(y)\|_{L_2(\mathbb{P};L_2)}\\
&\overset{(78)}{=}\frac{1}{\sqrt{\lceil\log n\rceil}}\left(\sum_{j=1}^{\lceil\log n\rceil}\int_\Omega\left|\min\left\{d_{\mathcal{M}}\big(x,\mathscr{Z}_{e^{-j}}(\omega)\big),\mathrm{diam}(\mathcal{M})\right\}-\min\left\{d_{\mathcal{M}}\big(y,\mathscr{Z}_{e^{-j}}(\omega)\big),\mathrm{diam}(\mathcal{M})\right\}\right|^2\mathrm{d}\mathbb{P}(\omega)\right)^{\frac12}\\
&\geqslant\frac{1}{\sqrt{\lceil\log n\rceil}}\left(\int_\Omega\left|\min\left\{d_{\mathcal{M}}\big(x,\mathscr{Z}_{e^{-i_R(x)}}(\omega)\big),\mathrm{diam}(\mathcal{M})\right\}-\min\left\{d_{\mathcal{M}}\big(y,\mathscr{Z}_{e^{-i_R(x)}}(\omega)\big),\mathrm{diam}(\mathcal{M})\right\}\right|^2\mathrm{d}\mathbb{P}(\omega)\right)^{\frac12}\\
&\overset{(80)}{\gtrsim}\frac{R-r}{\sqrt{K\log n}}. \qquad \square
\end{aligned}
$$

It remains to prove the second part of Theorem 24. We will start with the following lemma:

Lemma 26. *Fix $K,D,\beta\geqslant1$. Let $(\mathcal{M},d_{\mathcal{M}})$ be a locally finite metric space such that for every $\mathcal{C}\subseteq\mathcal{M}$ with $|\mathcal{C}|\leqslant K$ and every $\Delta>0$ there exists a 1-Lipschitz function $f=f_{d_{\mathcal{M}},\mathcal{C},\Delta}:\mathcal{M}\to L_2$ that satisfies:*

$$
\forall x,y\in\mathcal{C},\qquad \Delta\leqslant d_{\mathcal{M}}(x,y)\leqslant3\beta\Delta\implies\|f(x)-f(y)\|_{L_2}\geqslant\frac{\Delta}{D}. \tag{81}
$$

Then, for every finite nonempty subset $\mathcal{U}$ *of* $\mathcal{M}$*, any* $\Delta > 0$ *and any* $R > r > 0$ *that satisfy the restrictions*

$$\Delta \geqslant 9Dr \qquad \text{and} \qquad R - r \geqslant \mathrm{diam}_{\mathcal{M}}(\mathcal{U}), \tag{82}$$

there exists a 1*-Lipschitz function* $\phi = \phi_{\mathcal{U},r,R,\Delta,d_{\mathcal{M}}} : \mathcal{M} \to L_2$ *such that*

$$\forall (x,y) \in \big(\mathcal{U} \cap \mathcal{G}_{\leqslant K}(r,R)\big) \times \mathcal{M}, \qquad \Delta \leqslant d_{\mathcal{M}}(x,y) \leqslant \beta\Delta \implies \|\phi(x)-\phi(y)\|_{L_2} \gtrsim \frac{\Delta}{D}. \tag{83}$$

Due to the similarity of Lemma 26 and the second part of Theorem 24, it is helpful to spell out how they differ. Assumption (81) of Lemma 26 coincides with assumption (48) of Theorem 24. Conclusion (83) of Lemma 26 is weaker than the desired conclusion (50) of Theorem 24: It imposes a further restriction that the point x belongs not only to the set $\mathcal{G}_{\leqslant K}(r,R)$ of centers whose local growth at the input radii $R > r > 0$ is at most K, but also to an arbitrary but fixed bounded nonempty subset $\mathcal{U}$ of $\mathcal{M}$; for this, Lemma 26 requires in (82) that those radii are separated by $\mathrm{diam}_{\mathcal{M}}(\mathcal{U})$, while Theorem 24 asks in (49) for them to be separated by a quantity that involves both the given scale $\Delta > 0$ and the size of $\mathcal{G}_{\leqslant K}(r,R)$.

Proof of Lemma 26. Let $\mathcal{N}$ be a subset of $\mathcal{U} \cap \mathcal{G}_{\leqslant K}(r,R)$ with the property that $d_{\mathcal{M}}(a,b) > 2r$ for all distinct $a,b \in \mathcal{N}$, and furthermore $\mathcal{N}$ is maximal with respect to inclusion among all the subsets of $\mathcal{U} \cap \mathcal{G}_{\leqslant K}(r,R)$ that have this property. By the triangle inequality for $d_{\mathcal{M}}$, the balls $\{B_{\mathcal{M}}(a,r)\}_{a\in\mathcal{N}}$ are disjoint, so we have

$$|\mathcal{N}| \min_{a\in\mathcal{N}} |B_{\mathcal{M}}(a,r)| \leqslant \sum_{a\in\mathcal{N}} |B_{\mathcal{M}}(a,r)| = \Big|\bigcup_{a\in\mathcal{N}} B_{\mathcal{M}}(a,r)\Big|. \tag{84}$$

Fix $a_{\min} \in \mathcal{N}$ such that $|B_{\mathcal{M}}(a_{\min},r)| = \min_{a\in\mathcal{N}} |B_{\mathcal{M}}(a,r)|$. By the triangle inequality for $d_{\mathcal{M}}$, since $\mathcal{N} \subseteq \mathcal{U}$ we have $B_{\mathcal{M}}(a,r) \subseteq B_{\mathcal{M}}(a_{\min}, \mathrm{diam}_{\mathcal{M}}(\mathcal{U}) + r) \subseteq B_{\mathcal{M}}(a_{\min}, R)$ for every $a \in \mathcal{N}$, as $R - r \geqslant \mathrm{diam}_{\mathcal{M}}(\mathcal{U})$. Thanks to (84), this implies $|\mathcal{N}| \leqslant |B_{\mathcal{M}}(a_{\min},R)|/|B_{\mathcal{M}}(a_{\min},r)|$. As $a_{\min} \in \mathcal{N} \subseteq \mathcal{G}_{\leqslant K}(r,R)$, we deduce that $|\mathcal{N}| \leqslant K$.

By the assumption of Lemma 26 applied to $\mathcal{C} = \mathcal{N}$, there is a 1-Lipschitz function $f : \mathcal{M} \to L_2$ for which

$$\forall a,b \in \mathcal{N}, \qquad \Delta_0 \leqslant d_{\mathcal{M}}(a,b) \leqslant 3\beta\Delta_0 \implies \|f(a)-f(b)\|_{L_2} \geqslant \frac{\Delta_0}{D}, \quad \text{where} \quad \Delta_0 \stackrel{\mathrm{def}}{=} \frac{19D-1}{19D}\Delta - 4r, \tag{85}$$

which is valid since $\Delta \geqslant 9Dr > 76Dr/(19D-1)$, as $D \geqslant 1$, so the parameter Δ_0 in (85) is indeed positive. Because $\mathcal{M}$ is locally finite, $\phi(\mathcal{M})$ is a countable subset of L_2, so by replacing L_2 by $L_2 \oplus \mathbb{R}$, we may assume that there exists a vector $v \in f(\mathcal{M})^{\perp} \subseteq L_2$ that satisfies $\|v\|_{L_2} = 1$ and v is orthogonal to $f(\mathcal{M})$. Since f is 1-Lipschitz, and by the triangle inequality for $d_{\mathcal{M}}$ also the function $(x \in \mathcal{M}) \mapsto d_{\mathcal{M}}(x,\mathcal{N})$ is 1-Lipschitz, the aforementioned orthogonality implies that the following function $\phi : \mathcal{M} \to L_2$ is 1-Lipschitz:

$$\forall x \in \mathcal{M}, \qquad \phi(x) \stackrel{\mathrm{def}}{=} \frac{1}{\sqrt{2}} f(x) + \frac{d_{\mathcal{M}}(x,\mathcal{N})}{\sqrt{2}} v.$$

We then have the following lower bound:

$$\forall x,y \in \mathcal{M}, \qquad \|\phi(x)-\phi(y)\|_{L_2} \gtrsim \max\big\{\|f(x)-f(y)\|_{L_2}, |d_{\mathcal{M}}(x,\mathcal{N}) - d_{\mathcal{M}}(y,\mathcal{N})|\big\}. \tag{86}$$

Fix $x \in \mathcal{U} \cap \mathcal{G}_{\leqslant K}(r,R)$. By the maximality of $\mathcal{N}$ with respect to inclusion, there must exist $a_x \in \mathcal{N}$ such that $d_{\mathcal{M}}(x,a_x) \leqslant 2r$, i.e., we have $d_{\mathcal{M}}(x,\mathcal{N}) \leqslant 2r$. Consequently, if $y \in \mathcal{M}$ satisfies

$$d_{\mathcal{M}}(y,\mathcal{N}) \geqslant 2r + \frac{\Delta}{19D}, \tag{87}$$

then it follows from (86) that $\|\phi(x)-\phi(y)\|_{L_2} \gtrsim \Delta/D$, which is the desired conclusion of (87).

The above simple preparatory reasoning demonstrates that the proof of Lemma 26 will be complete if we will prove that $\|\phi(x)-\phi(y)\|_{L_2} \gtrsim \Delta/D$ whenever $x,y \in \mathcal{M}$ have the following properties:

$$\Delta \leqslant d_{\mathcal{M}}(x,y) \leqslant \beta\Delta \qquad \text{and} \qquad \exists a_x, a_y \in \mathcal{N}, \qquad d_{\mathcal{M}}(a,a_x) \leqslant 2r \quad \text{and} \quad d_{\mathcal{M}}(y,a_y) < 2r + \frac{\Delta}{19D}. \tag{88}$$

This is indeed the case because by the triangle inequality for $d_{\mathcal{M}}$ we have

$$d_{\mathcal{M}}(a_x,a_y) \geqslant d_{\mathcal{M}}(x,y) - d_{\mathcal{M}}(x,a_x) - d_{\mathcal{M}}(y,b_y) \stackrel{(88)}{\geqslant} \Delta - 4r + \frac{\Delta}{19D} \stackrel{(85)}{=} \Delta_0, \tag{89}$$

and

$$d_{\mathcal{M}}(a_x,a_y)\leqslant d_{\mathcal{M}}(x,y)+d_{\mathcal{M}}(x,a_y)+d_{\mathcal{M}}(y,b_y)\overset{(88)}{\leqslant}\beta\Delta+4r+\frac{\Delta}{19D}\leqslant 3\beta\Delta_0, \tag{90}$$

where checking that the last step of (90) is valid is straightforward using the definition (85) of Δ_0 together with the assumptions $\beta\geqslant 1$ and $\Delta\geqslant 9Dr\geqslant 9r$. Thanks to (89) and (90) we may use (85) with $a=a_x$ and $b=a_y$ in combination with (86) to conclude that the following lower bound holds:

$$\begin{aligned}\|\phi(x)-\phi(y)\|_{L_2}&\overset{(86)}{\gtrsim}\|f(x)-f(y)\|_{L_2}\geqslant\|f(a_x)-f(a_y)\|_{L_2}-\|f(x)-f(a_x)\|_{L_2}-\|f(y)-f(a_y)\|_{L_2}\\&\overset{(85)}{\geqslant}\frac{\Delta_0}{D}-d_{\mathcal{M}}(x,a_x)-d_{\mathcal{M}}(y,a_y)\overset{(85)\wedge(88)}{\geqslant}\frac{\frac{19D-1}{19D}\Delta-4r}{D}-4r-\frac{\Delta}{19D}\gtrsim\frac{\Delta}{D},\end{aligned} \tag{91}$$

where the second step of (91) uses the triangle inequality in L_2, the third step of (91) uses the fact that f is 1-Lipschitz, and the final step of (91) is straightforward to verify using $r\leqslant\Delta/(9D)$ and $D\geqslant 1$. □

Lemma 27 below provides a link between Lemma 26 and Theorem 24; it is a spatial localization principle for Euclidean single scale embeddings that does not involve the controlled local growth centers. The term "spatial localization" is used here for the following reason: The assumption of Lemma 27 per (92) is that "one scale at a time" embeddings exist for arbitrary subsets of small diameter, and its conclusion per (93) is that such an embedding guarantee is possible provided one of the points is close to a subset whose cardinality is sufficiently small without any bound on the spacial size (diameter) of that subset.

Lemma 27. *There is a universal constant $C\geqslant 1$ with the following property. Suppose that $\beta,D\geqslant 1$ and $\mathscr{d},\Delta>0$. Let $(\mathcal{M},d_{\mathcal{M}})$ be a metric space and fix $\mathcal{X},\mathcal{Y}\subseteq\mathcal{M}$. Assume that for every $\emptyset\neq\mathcal{U}\subseteq\mathcal{M}$ with $\mathrm{diam}_{\mathcal{M}}(\mathcal{U})\leqslant\mathscr{d}$ there is a 1-Lipschitz function $f_{\mathcal{U}}:\mathcal{U}\to L_2$ satisfying*

$$\forall (x,y)\in(\mathcal{U}\cap\mathcal{X})\times(\mathcal{U}\cap\mathcal{Y}),\qquad \Delta\leqslant d_{\mathcal{M}}(x,y)\leqslant\beta\Delta\implies\|f_{\mathcal{U}}(x)-f_{\mathcal{U}}(y)\|_{L_2}\geqslant\frac{\Delta}{D}. \tag{92}$$

Then, for any $\mathcal{C}\subseteq\mathcal{M}$ with $2\leqslant|\mathcal{C}|\leqslant e^{\frac{\mathscr{d}}{C\beta\Delta}}$ there is a 1-Lipschitz function $\varphi_{\mathcal{C}}:\mathcal{M}\to L_2$ satisfying

$$\forall (x,y)\in\mathcal{X}\times\mathcal{Y},\qquad \big(\Delta\leqslant d_{\mathcal{M}}(x,y)\leqslant\beta\Delta\big)\wedge\big(d_{\mathcal{M}}(x,\mathcal{C})\leqslant(\beta+1)\Delta\big)\implies\|\varphi_{\mathcal{C}}(x)-\varphi_{\mathcal{C}}(y)\|_{L_2}\gtrsim\frac{\Delta}{D}. \tag{93}$$

Prior to proving Lemma 27, which is of independent interest and could be of value for other purposes elsewhere (in particular, another application of it will be worked out in Section 3), we will next demonstrate how to quickly deduce the second part of Theorem 24 assuming its validity:

Proof of the second part of Theorem 24 assuming Lemma 27. We may assume that $|\mathcal{G}_{\leqslant K}(r,R)|\geqslant 2$ since if $\mathcal{G}_{\leqslant K}(r,R)=\emptyset$, then the desired conclusion (50) is vacuous, and if $|\mathcal{G}_{\leqslant K}(r,R)|=1$, then let $x_0\in\mathcal{M}$ be such that $\mathcal{G}_{\leqslant K}(r,R)=\{x_0\}$ and the desired conclusion (50) holds for $\phi(x)=d_{\mathcal{M}}(x,x_0)\in\mathbb{R}$. Let $C\geqslant 1$ be the constant from Lemma 27. Assumption (48) of Theorem 24 coincides with assumption (81) of Lemma 26, and the second part of assumption (49) Theorem 24 implies that $R-r\geqslant\mathrm{diam}_{\mathcal{M}}(\mathcal{U})$ for every nonempty subset $\mathcal{U}$ of $\mathcal{M}$ for which $\mathrm{diam}_{\mathcal{M}}(\mathcal{U})\leqslant C\beta(\log|\mathcal{G}_{\leqslant K}(r,R)|)\Delta$. We may therefore apply Lemma 26 to deduce that for every such $\mathcal{U}$ there is a 1-Lipschitz function $f_{\mathcal{U}}:\mathcal{M}\to L_2$ for which assumption (92) of Lemma 27 holds for $\mathcal{X}=\mathcal{Y}=\mathcal{M}$ and $\mathscr{d}=C\beta(\log|\mathcal{G}_{\leqslant K}(r,R)|)\Delta$, and with D replaced by a positive universal constant multiple of D. The desired conclusion (50) of Theorem 24 is now a special case of the conclusion (93) of Lemma 27, applied to $\mathcal{C}=\mathcal{G}_{\leqslant K}(r,R)$, which is valid because $2\leqslant|\mathcal{G}_{\leqslant K}(r,R)|=e^{\mathscr{d}/(C\beta\Delta)}$. □

Thus, in order to complete the proof of Theorem 24 it remains to prove Lemma 27, which we do next:

Proof of Lemma 27. For each $\mathcal{S}\subseteq\mathcal{M}$ and $K>0$, let $\mathrm{Part}_{\mathcal{M}}(\mathcal{S};K)$ denote the set of all the partitions $\mathcal{P}$ of $\mathcal{S}$ that are K-bounded with respect to $d_{\mathcal{M}}$, i.e., $\mathrm{diam}_{\mathcal{M}}(\mathcal{P}(x))\leqslant K$ for every $x\in\mathcal{S}$.

By [12], there exist universal constants $0<\mathfrak{p}<1<\kappa$ such that for any $\mathcal{C}\subseteq\mathcal{M}$ with $|\mathcal{C}|>1$ and any $\Delta_0>0$ there exists a probability measure $\mu_{\mathcal{C},\Delta_0}$ on $\mathrm{Part}_{\mathcal{M}}(\mathcal{C};\Delta_0)$ satisfying the following condition:

$$\forall x\in\mathcal{C},\qquad \mu_{\mathcal{C},\Delta_0}\Big[\mathcal{P}_0\in \mathrm{Part}_{\mathcal{M}}(\mathcal{C};\Delta_0):\ B_{\mathcal{M}}\Big(x,\frac{1}{\kappa\log|\mathcal{C}|}\Delta_0\Big)\cap\mathcal{C}\subseteq\mathcal{P}_0(x)\Big]\geqslant\mathfrak{p}.$$

We will proceed to prove that Lemma 27 holds if we take the universal constant C equal to 60κ.

Fix any subset $\mathcal{C}$ of $\mathcal{M}$ whose size satisfies the assumptions in the statement of Lemma 27, i.e.,

$$2\leqslant|\mathcal{C}|\leqslant e^{\frac{\mathscr{d}}{C\beta\Delta}}=e^{\frac{\mathscr{d}}{60\kappa\beta\Delta}}. \tag{94}$$

It will be convenient to set the following notations:

$$\Delta_0\stackrel{\mathrm{def}}{=}16\kappa(\beta+1)\big(\log|\mathcal{C}|\big)\Delta\qquad\text{and}\qquad\Omega\stackrel{\mathrm{def}}{=}\mathrm{Part}_{\mathcal{M}}\big(\mathcal{M};8(\beta+1)(2\kappa\log|\mathcal{C}|+1)\Delta\big)\qquad\text{and}\qquad\mu\stackrel{\mathrm{def}}{=}\mu_{\mathcal{C},\Delta_0}.$$

Thus, Ω is the set of Δ'-bounded partitions of the entire metric space $\mathcal{M}$, where we introduce the notation

$$\Delta'\stackrel{\mathrm{def}}{=}8(\beta+1)(2\kappa\log|\mathcal{C}|+1)\Delta\stackrel{(94)}{\leqslant}8(\beta+1)\Big(\frac{1}{30\beta\Delta}\mathscr{d}+1\Big)\Delta\leqslant\mathscr{d}, \tag{95}$$

and the last step of (95) is valid because (94) implies in particular that $\mathscr{d}/\Delta\geqslant 60\kappa\beta\log 2$, and using this lower bound on $\mathscr{d}/\Delta$ the rightmost inequality in (94) is elementary to verify, since $\beta,\kappa\geqslant1$. Now, a substitution of μ into [53, Lemma 3.8] shows that there exists a probability measure ν on Ω such that

$$\forall x\in\mathcal{M},\qquad d_{\mathcal{M}}(x,\mathcal{C})\leqslant(\beta+1)\Delta\implies\nu\big[\mathcal{P}\in\Omega:\ B_{\mathcal{M}}(x,(\beta+1)\Delta)\subseteq\mathcal{P}(x)\big]\geqslant\mathfrak{p}. \tag{96}$$

By [68, Lemma 5.2] there exists a function $G=G_\Delta:L_2\to L_2$ such that for every $x,y\in L_2$,

$$\|G(x)\|_{L_2}=\|G(y)\|_{L_2}=\Delta\quad\text{and}\quad\frac12\min\big\{\Delta,\|x-y\|_{L_2}\big\}\leqslant\|G(x)-G(y)\|_{L_2}\leqslant\min\big\{\Delta,\|x-y\|_{L_2}\big\}. \tag{97}$$

For every $\mathcal{U}\subsetneq\mathcal{M}$ define $\mathfrak{d}_{\mathcal{U}}=\mathfrak{d}_{\mathcal{U},\Delta,d_{\mathcal{M}}}:\mathcal{M}\to[0,1]$ by

$$\forall z\in\mathcal{M},\qquad\mathfrak{d}_{\mathcal{U}}(z)\stackrel{\mathrm{def}}{=}\min\Big\{1,\frac{1}{\Delta}d_{\mathcal{M}}(z,\mathcal{M}\smallsetminus\mathcal{U})\Big\}. \tag{98}$$

As the function $(t\in\mathbb{R})\mapsto\min\{1,t/\Delta\}$ is $(1/\Delta)$-Lipschitz and nondecreasing, $\|\mathfrak{d}_{\mathcal{U}}\|_{\mathrm{Lip}(\mathcal{M};\mathbb{R})}\leqslant1/\Delta$. Finally, because by (95) we have $\mathrm{diam}_{\mathcal{M}}(\mathcal{P}(z))\leqslant\mathscr{d}$ for every $\mathcal{P}\in\Omega$ and $z\in\mathcal{M}$, so we can invoke the assumption of Lemma 27 to fix a 1-Lipschitz function $f_{\mathcal{P}(z)}:\mathcal{P}(z)\to L_2$ that satisfies (92) with $\mathcal{U}=\mathcal{P}(z)$, i.e.,

$$\forall(x,y)\in(\mathcal{P}(z)\cap\mathcal{X})\times(\mathcal{P}(z)\cap\mathcal{Y}),\qquad\Delta\leqslant d_{\mathcal{M}}(x,y)\leqslant\beta\Delta\implies\|f_{\mathcal{P}(z)}(x)-f_{\mathcal{P}(z)}(y)\|_{L_2}\geqslant\frac{\Delta}{D}. \tag{99}$$

In particular, $f_{\mathcal{P}(z)}(z)$ is well-defined since $z\in\mathcal{P}(z)$, so we can define $\varphi_{\mathcal{C}}:\mathcal{M}\to L_2(\nu;L_2)\cong L_2$ by setting

$$\forall z\in\mathcal{M},\ \forall\mathcal{P}\in\Omega,\qquad\varphi_{\mathcal{C}}(z)(\mathcal{P})\stackrel{\mathrm{def}}{=}\frac{\mathfrak{d}_{\mathcal{P}(z)}(z)}{2}G\big(f_{\mathcal{P}(z)}(z)\big). \tag{100}$$

We first claim that the following point-wise bound holds:

$$\forall x,y\in\mathcal{M},\ \forall\mathcal{P}\in\Omega,\qquad\|\varphi_{\mathcal{C}}(x)(\mathcal{P})-\varphi_{\mathcal{C}}(y)(\mathcal{P})\|_{L_2}\leqslant d_{\mathcal{M}}(x,y). \tag{101}$$

Observe that after (101) will be established, by squaring both of its sides and then integrating the resulting estimate $\mathrm{d}\nu(\mathcal{P})$, we will deduce that $\varphi_{\mathcal{C}}$ is 1-Lipschitz as a function $\mathcal{M}$ to $L_2(\nu;L_2)\cong L_2$.

To verify (101), fix $x,y\in\mathcal{M}$ and $\mathcal{P}\in\Omega$. Suppose first that $\mathcal{P}(x)=\mathcal{P}(y)\stackrel{\mathrm{def}}{=}\mathcal{U}$. Then

$$\begin{aligned}|\varphi_{\mathcal{C}}(x)(\mathcal{P})-\varphi_{\mathcal{C}}(x)(\mathcal{P})|&=\frac12\Big\|\big(\mathfrak{d}_{\mathcal{U}}(x)-\mathfrak{d}_{\mathcal{U}}(y)\big)G\big(f_{\mathcal{U}}(x)\big)+\mathfrak{d}_{\mathcal{U}}(y)\Big(G\big(f_{\mathcal{U}}(x)\big)-G\big(f_{\mathcal{U}}(y)\big)\Big)\Big\|_{L_2}\\&\leqslant\frac{|\mathfrak{d}_{\mathcal{U}}(x)-\mathfrak{d}_{\mathcal{U}}(y)|}{2}\big\|G\big(f_{\mathcal{U}}(x)\big)\big\|_{L_2}+\frac{\mathfrak{d}_{\mathcal{U}}(y)}{2}\big\|G\big(f_{\mathcal{U}}(x)\big)-G\big(f_{\mathcal{U}}(y)\big)\big\|_{L_2}\\&\leqslant\frac{\|\mathfrak{d}_{\mathcal{U}}\|_{\mathrm{Lip}(\mathcal{M};\mathbb{R})}}{2}\Delta d_{\mathcal{M}}(x,y)+\frac12\|f_{\mathcal{U}}\|_{\mathrm{Lip}(\mathcal{M};L_2)}d_{\mathcal{M}}(x,y)\leqslant d_{\mathcal{M}}(x,y),\end{aligned} \tag{102}$$

where the first step of (102) is a consequence of the definition (100) of $\varphi_{\mathcal{C}}$, using the current assumption $\mathcal{P}(x)=\mathcal{P}(y)=\mathcal{U}$, the penultimate step of (102) uses (97) and the fact that $0\leqslant\mathfrak{d}_{\mathcal{U}}(\cdot)\leqslant1$ by (98), and the final step of (102) holds as $f_{\mathcal{U}}$ is 1-Lipschitz and $\mathfrak{d}_{\mathcal{U}}$ is $(1/\Delta)$-Lipschitz. This establishes (101) if $\mathcal{P}(x)=\mathcal{P}(y)$.

It remains to verify (101) when $\mathcal{P}(x) \neq \mathcal{P}(y)$, i.e., $x \in \mathcal{M} \smallsetminus \mathcal{P}(y)$ and $y \in \mathcal{M} \smallsetminus \mathcal{P}(x)$, which implies that

$$\max\Big\{d_{\mathcal{M}}\big(x,\mathcal{M}\smallsetminus\mathcal{P}(x)\big), d_{\mathcal{M}}\big(y,\mathcal{M}\smallsetminus\mathcal{P}(y)\big)\Big\} \leqslant d_{\mathcal{M}}(x,y). \tag{103}$$

Consequently, when $\mathcal{P}(x) \neq \mathcal{P}(y)$ we can justify (101) as follows:

$$\begin{aligned}|\varphi_{\mathcal{C}}(x)(\mathcal{P}) - \varphi_{\mathcal{C}}(x)(\mathcal{P})| \leqslant |\varphi_{\mathcal{C}}(x)(\mathcal{P})| + |\varphi_{\mathcal{C}}(x)(\mathcal{P})| &\overset{(100)\wedge(97)}{=} \frac{1}{2}\mathfrak{d}_{\mathcal{P}(x)}(x)\Delta + \frac{1}{2}\mathfrak{d}_{\mathcal{P}(y)}(y)\Delta \\ &\overset{(98)}{\leqslant} \frac{1}{2}d_{\mathcal{M}}\big(x,\mathcal{M}\smallsetminus\mathcal{P}(x)\big) + \frac{1}{2}d_{\mathcal{M}}\big(y,\mathcal{M}\smallsetminus\mathcal{P}(y)\big) \overset{(103)}{\leqslant} d_{\mathcal{M}}(x,y).\end{aligned}$$

Next, fix $x, y \in \mathcal{M}$ that satisfy $d_{\mathcal{M}}(x,y) \leqslant \beta\Delta$. Using the triangle inequality for $d_{\mathcal{M}}$, it follows that $B_{\mathcal{M}}(y,\Delta) \subseteq B_{\mathcal{M}}(x,(\beta+1)\Delta)$. Hence, if $\mathcal{P} \in \Omega$ is such that $B_{\mathcal{M}}(x,(\beta+1)\Delta) \subseteq \mathcal{P}(x)$, then $\mathcal{P}(y) = \mathcal{P}(x)$ and also $\min\{d_{\mathcal{M}}(x,\mathcal{M}\smallsetminus\mathcal{P}(x)), d_{\mathcal{M}}(y,\mathcal{M}\smallsetminus\mathcal{P}(y))\} \geqslant \Delta$. Recalling (98), it follows that

$$\big(d_{\mathcal{M}}(x,y) \leqslant \beta\Delta\big) \,\wedge\, \big(B_{\mathcal{M}}(x,(\beta+1)\Delta) \subseteq \mathcal{P}(x)\big) \implies \big(\mathcal{P}(x) = \mathcal{P}(y)\big) \wedge \big(\mathfrak{d}_{\mathcal{P}(x)}(x) = \mathfrak{d}_{\mathcal{P}(y)}(y) = 1\big), \tag{104}$$

for every $x, y \in \mathcal{M}$ and every $\mathcal{P} \in \Omega_{\Delta}^{\mathcal{M}}$. Consequently, the following lower bound holds for all $(x,y) \in \mathcal{X}\times\mathcal{Y}$ that satisfy $d_{\mathcal{M}}(x,\mathcal{C}) \leqslant (\beta+1)\Delta$ and $\Delta \leqslant d_{\mathcal{M}}(x,y) \leqslant \beta\Delta$:

$$\begin{aligned}\|\varphi_{\mathcal{C}}(x) - \varphi_{\mathcal{C}}(y)\|_{L_2(\nu;L_2)} &\geqslant \bigg(\int_{\{\mathcal{P}\in\Omega:\ B_{\mathcal{M}}(x,(\beta+1)\Delta)\subseteq\mathcal{P}(x)\}} \|\varphi_{\mathcal{C}}(x) - \varphi_{\mathcal{C}}(y)\|_{L_2}^2 \,\mathrm{d}\nu(\mathcal{P})\bigg)^{\frac12} \\ &= \frac{1}{2}\bigg(\int_{\{\mathcal{P}\in\Omega:\ B_{\mathcal{M}}(x,(\beta+1)\Delta)\subseteq\mathcal{P}(x)\}} \big\|G\big(f_{\mathcal{P}(x)}(x)\big) - G\big(f_{\mathcal{P}(x)}(y)\big)\big\|_{L_2}^2 \,\mathrm{d}\nu(\mathcal{P})\bigg)^{\frac12} \\ &\geqslant \frac{1}{4}\bigg(\int_{\{\mathcal{P}\in\Omega:\ B_{\mathcal{M}}(x,(\beta+1)\Delta)\subseteq\mathcal{P}(x)\}} \min\big\{\Delta^2, \|f_{\mathcal{P}(x)}(x) - f_{\mathcal{P}(x)}(y)\|_{L_2}^2\big\} \,\mathrm{d}\nu(\mathcal{P})\bigg)^{\frac12} \\ &\geqslant \frac{\Delta}{4D}\sqrt{\nu\big(\{\mathcal{P}\in\Omega:\ B_{\mathcal{M}}(x,(\beta+1)\Delta)\subseteq\mathcal{P}(x)\}\big)} \geqslant \frac{\Delta\sqrt{\mathfrak{p}}}{4D} \asymp \frac{\Delta}{D},\end{aligned} \tag{105}$$

where the second step of (105) holds by (100) using (104), which is valid as $d_{\mathcal{M}}(x,y) \leqslant \beta\Delta$, the third step of (105) uses the first inequality in (97), the fourth step of (105) uses (99), which is valid as $\Delta \leqslant d_{\mathcal{M}}(x,y) \leqslant \beta\Delta$ and $(x,y) \in (\mathcal{P}(x)\cap\mathcal{X})\times(\mathcal{P}(x)\cap\mathcal{Y})$ by (104), and $D \geqslant 1$, and the penultimate step of (105) is where (96) is used, which is valid as $d_{\mathcal{M}}(x,\mathcal{C}) \leqslant (\beta+1)\Delta$. We have thus proved the remaining part (93) of Lemma 27. □

Remark 28. An inspection of the proof of Lemma 27 reveals that the restriction on the size of $\mathcal{C}$ appears in its statement only because by [12] the *padding modulus* of $\mathcal{C}$ is $O(\log|\mathcal{C}|)$. Specifically, using the notation of [80], given $0 < \mathfrak{p} < 1$ let $\mathsf{PAD}_{\mathfrak{p}}(\mathcal{C}) = \mathsf{PAD}_{\mathfrak{p}}(\mathcal{C}, d_{\mathcal{M}})$ be the smallest $K \geqslant 1$ such that for every $\Delta > 0$ there is a distribution over Δ-bounded random partitions $\mathcal{P}$ of $\mathcal{C}$ with the property that for every $x \in \mathcal{C}$ the probability that $B_{\mathcal{M}}(x,\Delta/K)\cap\mathcal{C}$ is contained in $\mathcal{P}(x)$ is at least $\mathfrak{p}$. A repetition of the reasoning of Lemma 27 gives mutatis mutandis that if one replaces its requirement $|\mathcal{C}| \leqslant e^{\ell/(C\beta\Delta)}$ by $\mathsf{PAD}_{\mathfrak{p}}(\mathcal{C}) \leqslant \ell/(C\beta\Delta)$, then its conclusion (93) holds with $\|\varphi_{\mathcal{C}}(x) - \varphi_{\mathcal{C}}(y)\|_{L_2} \gtrsim \Delta/D$ replaced by $\|\varphi_{\mathcal{C}}(x) - \varphi_{\mathcal{C}}(y)\|_{L_2} \gtrsim \Delta\sqrt{\mathfrak{p}}/D$.

3. Localized weakly bi-Lipschitz embeddings

The following "localized version" of Definition 19 is a slight generalization of a definition that appeared in [75, Section 7.2], which corresponds to the (arbitrary) choice $K = 32$ below:

Definition 29. *Given $K, D > 0$, a metric space $(\mathcal{M}, d_{\mathcal{M}})$ is said to admit a K-localized weakly bi-Lipschitz embedding into a metric space $(\mathcal{N}, d_{\mathcal{N}})$ with distortion D if for every $\Delta > 0$ and every $z \in \mathcal{M}$ there exists a non-constant Lipschitz function $f_{\Delta}^z : \mathcal{M} \to \mathcal{N}$ such that*

$$\forall x, y \in B_{\mathcal{M}}(z, K\Delta), \qquad d_{\mathcal{M}}(x,y) \geqslant \Delta \implies d_{\mathcal{N}}\big(f_{\Delta}^z(x), f_{\Delta}^z(y)\big) \geqslant \frac{\|f_{\Delta}^z\|_{\mathrm{Lip}}}{D}\Delta, \tag{106}$$

A key (well known) property of L_p that we will use herein is the (first part of) the following theorem:

Theorem 30. *If $K>1$ and $p>2$, then L_p admits a K-localized weakly bi-Lipschitz embedding into L_2 with distortion D, where $D\lesssim p2^{p/2}K^{p/2-1}$. Conversely, if L_p admits a K-localized weakly bi-Lipschitz embedding into L_2 with distortion D, then necessarily $D\gtrsim 2^{p/2}K^{p/2-1}$.*

Proof. The first part of Theorem (30) essentially coincides with Lemma 7.6 of [75], which notes the (special case $K=32$ of the) following general statement. Suppose that $(\mathbf{X},\|\cdot\|_{\mathbf{X}})$ and $(\mathbf{Y},\|\cdot\|_{\mathbf{Y}})$ are Banach spaces for which there exists a function $U:B_{\mathbf{X}}\to\mathbf{Y}$ from the unit ball $B_{\mathbf{X}}=\{x\in\mathbf{X}:\ \|x\|_{\mathbf{X}}\leqslant 1\}$ to $\mathbf{Y}$ satisfying

$$\forall x,y\in B_{\mathbf{X}},\qquad \omega\big(\|x-y\|_{\mathbf{X}}\big)\leqslant\|U(x)-U(y)\|_{\mathbf{Y}}\leqslant L\|x-y\|_{\mathbf{X}},\tag{107}$$

for some $L>0$ and an increasing modulus $\omega:[0,\infty)\to[0,\infty)$. Let $\rho_{\mathbf{X}}:\mathbf{X}\to B_{\mathbf{X}}$ be the standard retraction (e.g. [75, equation (5.2)]) from $\mathbf{X}$ onto $B_{\mathbf{X}}$, i.e., $\rho_{\mathbf{X}}(x)=x/\max\{\|x\|_{\mathbf{X}},1\}$ for every $x\in\mathbf{X}$. It is straightforward to check that $\|\rho_{\mathbf{X}}\|_{\mathrm{Lip}}\leqslant 2$, so if we define for $z\in\mathbf{X}$ and $K,\Delta>0$ a function $f_\Delta^z:\mathbf{X}\to\mathbf{Y}$ by:

$$\forall x\in\mathbf{X},\qquad f_\Delta^z(x)\stackrel{\mathrm{def}}{=}K\Delta U\bigg(\rho_{\mathbf{X}}\Big(\frac{1}{K\Delta}(x-z)\Big)\bigg),$$

then $\|f_\Delta^z\|_{\mathrm{Lip}}\leqslant 2L$ by the second inequality in (107). Using this together with the first inequality in (107) shows that for every $x,y\in B_{\mathbf{X}}(z,K\Delta)$, if $\|x-y\|_{\mathbf{X}}\geqslant\delta$, then $\|f_\Delta^z(x)-f_\Delta^z(y)\|_{\mathbf{Y}}\geqslant K\omega(1/K)\|f_\Delta^z\|_{\mathrm{Lip}}\Delta/(2L)$. Thus, $\mathbf{X}$ admits a K-localized weakly bi-Lipschitz embedding into $\mathbf{Y}$ with distortion $2L/(K\omega(1/K))$.

In particular, if in (107) we have $\omega(t)=t^\alpha/\beta$ for all $t\geqslant 0$ and some $\alpha,\beta\geqslant 1$, then $\mathbf{X}$ admits a K-localized weakly bi-Lipschitz embedding into $\mathbf{Y}$ with distortion $2L\beta K^{\alpha-1}$. When $\mathbf{X}=L_p$ for some $p\geqslant 2$ and $\mathbf{Y}=L_2$, one can take U to be the restriction to the unit ball of L_p of the classical Mazur map [67] $M_{p\to 2}:L_p\to L_2$, in which case this holds with $\alpha=p/2$ and $\beta\asymp 2^{p/2}$, and with L in (107) satisfying $L\asymp p$. A derivation of these values of α,β,L for the Mazur map appears in [75]; see specifically equation (5.32) there.

We will establish the reverse direction by adjusting the proof of [75, Lemma 52]. So, suppose that L_p admits a K-localized weakly bi-Lipschitz embedding into L_2 with distortion D, and our goal is to bound D from below. As the complex plane $\mathbb{C}=\ell_2^2$, embeds into L_p with distortion $1+\varepsilon$ for any $0<\varepsilon<K-1$ (e.g. by Dvoretzky's theorem [30], though using that theorem is overkill for this purpose), it follows that also $L_p(\mathbb{C})$ admits a $K/(1+\varepsilon)$-localized weakly bi-Lipschitz embedding into L_2 with distortion $(1+\varepsilon)D$. Therefore, it suffices to prove that if $L_p(\mathbb{C})$ admits a K-localized weakly bi-Lipschitz embedding into L_2 with distortion D, then necessarily $D\gtrsim 2^{p/2}K^{p/2-1}$ (considering $\varepsilon\asymp 1/p$ above suffices).

Fix $n\in\mathbb{N}$. Applying the assumption that $L_p(\mathbb{C})$ admits a K-localized weakly bi-Lipschitz embedding into L_2 of distortion D to $\Delta=n^{1/p}/K$ in Definition 29 , it follows that there is $f:n^{1/p}B_{\ell_p^n(\mathbb{C})}\to L_2$ satisfying

$$\|f\|_{\mathrm{Lip}}=1\qquad\text{and}\qquad\forall x,y\in n^{\frac1p}B_{\ell_p^n(\mathbb{C})},\qquad \|x-y\|_{\ell_p^n(\mathbb{C})}\geqslant\frac{n^{\frac1p}}{K}\implies\|f(x)-f(y)\|_{L_2}\geqslant\frac{n^{\frac1p}}{KD}.\tag{108}$$

For $m\in\mathbb{N}$ write $\mathfrak{u}_m\stackrel{\mathrm{def}}{=}e^{2\pi\mathrm{i}/m}\in\mathbb{C}$ and let $\mathbb{U}_m\stackrel{\mathrm{def}}{=}\{1,\mathfrak{u}_m,\mathfrak{u}_m^2,\ldots,\mathfrak{u}_m^{m-1}\}\subseteq\mathbb{C}$ be the cyclic group of the roots of unity of order m. By (the Hilbertian special case of) Theorem 5.2 in [84] (whose conclusion is a slightly simpler variant of [70, Theorem 4.1][9]), if $m\geqslant\sqrt{n}$ and m is divisible by 8, then every $g:\mathbb{U}_m^n\to L_2$ satisfies:

$$\begin{aligned}&\sum_{j=1}^n\sum_{x\in\mathbb{U}_m^n}\big\|g(x)-g(x_1,\ldots,x_{j-1},-x_j,x_{j+1},\ldots,x_n)\big\|_{L_2}^2\\&\qquad\lesssim\frac{m^2}{2^n}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{U}_m^n}\big\|g(x)-g(\mathfrak{u}_m^{\varepsilon_1}x_1,\ldots,\mathfrak{u}_m^{\varepsilon_n}x_n)\big\|_{L_2}^2.\end{aligned}\tag{109}$$

We will apply (109) to the restriction to $\mathbb{U}_m^n\subseteq n^{1/p}B_{\ell_p^n(\mathbb{C})}$ of the function $f:n^{1/p}B_{\ell_p^n(\mathbb{C})}\to L_2$ in (108). For that, observe that since f is 1-Lipschitz, every $\varepsilon=(\varepsilon_1,\ldots,\varepsilon_n)\in\{-1,1\}^n$ and $x=(x_1,\ldots,x_n)\in\mathbb{U}_m^n$ satisfy

$$\big\|f(x)-f(\mathfrak{u}_m^{\varepsilon_1}x_1,\ldots,\mathfrak{u}_m^{\varepsilon_n}x_n)\big\|_{L_2}\leqslant\big\|(1-\mathfrak{u}_m^{\varepsilon_1})x_1,\ldots,(1-\mathfrak{u}_m^{\varepsilon_n})x_n\big\|_{\ell_p^n(\mathbb{C})}=\sqrt{2}n^{\frac1p}\Big|1-\cos\Big(\frac{2\pi}{m}\Big)\Big|\asymp\frac{n^{\frac1p}}{m}.$$

[9] Alternatively one could combine here Theorem 4.1 in [70] with the discussion in Section 4 of [33].

Hence, the right hand side of (109) for $g = f|_{\mathbb{U}_m^n}$ can be bounded from above as follows:

$$\frac{m^2}{2^n}\sum_{\varepsilon\in\{-1,1\}^n}\sum_{x\in\mathbb{U}_m^n}\left\|f(x)-f(\mathrm{u}_m^{\varepsilon_1}x_1,\ldots,\mathrm{u}_m^{\varepsilon_n}x_n)\right\|_{L_2}^2 \lesssim n^{\frac{2}{p}}m^n. \tag{110}$$

To bound the right hand side of (109) from below, suppose that $n\leqslant(2K)^p$. Then, for every $x\in\mathbb{U}_m$ and every $j\in\{1,\ldots,n\}$ we have $\|x-(x_1,\ldots,x_{j-1},-x_j,x_{j+1},\ldots,x_n)\|_{L_p(\mathbb{C})}=2\geqslant n^{1/p}/K$, so using (108) we get:

$$\forall\, x\in\mathbb{U}_m^m,\qquad \left\|f(x)-f(x_1,\ldots,x_{j-1},-x_j,x_{j+1},\ldots,x_n)\right\|_{L_2}\geqslant\frac{n^{\frac{1}{p}}}{KD}.$$

Consequently,

$$\sum_{j=1}^n\sum_{x\in\mathbb{U}_m^n}\left\|f(x)-f(x_1,\ldots,x_{j-1},-x_j,x_{j+1},\ldots,x_n)\right\|_{L_2}^2\geqslant\frac{n^{1+\frac{2}{p}}m^n}{(KD)^2}. \tag{111}$$

By combining (110) and (111) with (109) and rearranging, we deduce that $D\gtrsim\sqrt{n}/K$. The requirements for this to hold were that $n\leqslant(2K)^p$, as well as that $m\geqslant\sqrt{n}$ and $m\equiv 0 \mod 8$, so choosing $n=\lfloor(2K)^p\rfloor$ and $m=8\lceil\sqrt{n}\rceil$ gives $D\gtrsim\sqrt{(2K)^p-1}/K\asymp 2^{p/2}K^{p/2-1}$, as required. $\square$

Remark 31. It is worthwhile to note (though not needed for the results herein) that by incorporating the Maurey–Pisier theorem [66] and the main result of [70] that Rademacher cotype and metric cotype coincide, the above proof of the lower bound on D in Theorem 30 yields mutatis mutandis the following statement. Let $(\mathbf{X},\|\cdot\|_{\mathbf{X}}),(\mathbf{Y},\|\cdot\|_{\mathbf{Y}})$ be Banach spaces. Suppose $\mathbf{Y}$ has Rademacher cotype q and that $p\geqslant 2$ equals the infimum over those $p'\geqslant 2$ such that $\mathbf{X}$ has Rademcaher cotype p'. Fix $K,D>1$ and $p>q$. If $\mathbf{X}$ admits a K-localized weakly bi-Lipschitz embedding into $\mathbf{Y}$ with distortion D, then $D\gtrsim 2^{p/q}K^{p/q-1}/C_q(\mathbf{Y})$, where $C_q(\mathbf{Y})$ is the Rademacher cotype q constant of $\mathbf{Y}$.

As done in (33) for the two-sided scale-localized variant of Definition 19, we will also say in the context of Definition 29 that a metric space $(\mathcal{M},d_{\mathcal{M}})$ admits a *two-sided scale-localized K-localized weakly bi-Lipschitz embedding into a metric space $(\mathcal{N},d_{\mathcal{N}})$ with distortion* $D>0$ if for every $\Delta>0$ and every $z\in\mathcal{M}$ there exists a non-constant Lipschitz function $g_\Delta^z:\mathcal{M}\to\mathcal{N}$ such that

$$\forall\, x,y\in B_{\mathcal{M}}(z,K\Delta),\qquad \Delta\leqslant d_{\mathcal{M}}(x,y)\leqslant 2\Delta\implies d_{\mathcal{N}}\big(g_\Delta^z(x),g_\Delta^z(y)\big)\geqslant\frac{\|g_\Delta^z\|_{\mathrm{Lip}}}{D}\Delta. \tag{112}$$

The localization principle of Lemma 27 allows us to relate as follows localized weakly bi-Lipschitz embeddings to the usual (global) weakly bi-Lipschitz embeddings that were discussed in Section 2.1:

Lemma 32. *There exists a universal constant $\kappa\geqslant 2$ with the following property. Fix a nondecreasing function $D:(1,\infty)\to(1,\infty)$. Suppose that $(\mathcal{M},d_{\mathcal{M}})$ is a metric space admitting a two-sided scale-localized K-localized weakly bi-Lipschitz embedding into L_2 with distortion less than $D(K)$ for every $K>1$. Then,*

$$\forall\, n\in\{2,3,\ldots\},\qquad \widehat{\mathrm{d}}_2^n(\mathcal{M})\leqslant D(\kappa\log n). \tag{113}$$

Proof. We will show that (113) holds for $\kappa=2C$, where $C\geqslant 1$ is the universal constant of Lemma 27. For every nonempty bounded subset $\mathcal{U}$ of $\mathcal{M}$, fix any $z_{\mathcal{U}}\in\mathcal{U}$. Then, $\mathcal{U}\subseteq B_{\mathcal{M}}(z_{\mathcal{U}},\mathrm{diam}_{\mathcal{M}}(\mathcal{U}))$, whence for every $\Delta>0$, by considering the restriction to $\mathcal{U}$ of the function in (112) with $K=\mathrm{diam}_{\mathcal{M}}(\mathcal{U})/\Delta$ and $z=z_{\mathcal{U}}$, the assumption of Lemma 32 implies that there exists a 1-Lipschitz function $f_{\mathcal{U},\Delta}:\mathcal{M}\to L_2$ that satisfies

$$\forall\, x,y\in\mathcal{U},\qquad \Delta\leqslant d_{\mathcal{M}}(x,y)\leqslant 2\Delta\implies\|f_{\mathcal{U},\Delta}(x)-f_{\mathcal{U},\Delta}(y)\|_{L_2}\geqslant\frac{\Delta}{D\left(\frac{\mathrm{diam}_{\mathcal{M}}(\mathcal{U})}{\Delta}\right)}.$$

So, for every integer $n\geqslant 2$ we may apply Lemma 27 with $\mathcal{X}=\mathcal{Y}=\mathcal{M}$ and $\beta=2$, as well as $\mathscr{d}=2C(\log n)\Delta$ and $D=D(\mathscr{d}/\Delta)$, to obtain for every $\mathcal{C}\subseteq\mathcal{M}$ with $|\mathcal{C}|=n$ a 1-Lipschitz function $\varphi_{\mathcal{C},\Delta}:\mathcal{M}\to L_2$ that satisfies

$$\forall\, x,y\in\mathcal{M},\qquad \big(\Delta\leqslant d_{\mathcal{M}}(x,y)\leqslant 2\Delta\big)\wedge\big(d_{\mathcal{M}}(x,\mathcal{C})\leqslant 3\Delta\big)\implies\|\varphi_{\mathcal{C},\Delta}(x)-\varphi_{\mathcal{C},\Delta}(y)\|_{L_2}\gtrsim\frac{\Delta}{D(2C\log n)}.$$

This is more than what is needed to deduce that $\widehat{\mathsf{d}}_2^n(\mathcal{M})\leqslant D(2C\log n)$, per (33). In particular, we obtained a function $\varphi_{\mathcal{C},\Delta}$ that is defined on all of $\mathcal{M}$ while we only need it to be a 1-Lipschitz function on $\mathcal{C}$. For that, in the above application Lemma 27 it would suffice that $f_{\mathcal{U},\Delta}$ is defined on $\mathcal{U}$, for which we could work here with the weakening of the notion of two-sided scale-localized K-localized weakly bi-Lipschitz embedding in which the function g_Δ^z in (112) is defined only on $B_{\mathcal{M}}(z,K\Delta)$. This observation could be relevant to future investigations, but not when $\mathcal{M}$ is a Banach space since in that case Lipschitz functions can be extended from any ball to the super-space while increasing their Lipschitz constant by a factor of at most 2, which is seen by composing with the radial retraction as we did in the proof of Theorem 30. $\square$

With Lemma 32 at hand, we can now summarize the best available upper bounds on $\widehat{\mathsf{d}}_2^n(L_p)$:

$$\forall n\in\{2,3\ldots,\},\qquad \widehat{\mathsf{d}}_2^n(L_p)\lesssim\begin{cases}\sqrt{\log n} & \text{if } 1\leqslant p\leqslant\sqrt{5}-1,\\ (\log n)^{\frac{2}{p}\left(\frac{1}{p}-\frac12\right)} & \text{if } \sqrt{5}-1\leqslant p\leqslant 2,\\ (\log n)^{\frac{p}{2}-1} & \text{if } 2\leqslant p\leqslant 3,\\ p^3\sqrt{\log n} & \text{if } 3\leqslant p\leqslant \sqrt[6]{\log n},\\ \log n & \text{if } p\geqslant\sqrt[6]{\log n}.\end{cases}\tag{114}$$

Our contribution here is the range $2\leqslant p=o(\sqrt[6]{\log n})$ of (114), in which the previously best known estimate was nothing more than the $O(\log n)$ upper bound that holds by [16] for any n-point metric space. The case $2\leqslant p\leqslant 3$ of (114) follows by substituting Theorem 30 into Lemma 32, and if $3\leqslant p\leqslant\sqrt[6]{\log n}$, then (114) follows by substituting Theorem 10 and the bound $\mathsf{e}(L_p;L_2)\lesssim\sqrt{p}$ of [81] into Lemma 21 (using Theorem 30 and Lemma 32 here would yield a weaker result). If $1\leqslant p\leqslant\sqrt{5}-1$, then (114) is due to [6], if $\sqrt{5}-1\leqslant p\leqslant 2$, then (114) is due to [50], and if $p\geqslant\sqrt[6]{\log n}$, then (114) follows from [16].

The best available lower bound on $\widehat{\mathsf{d}}_2^n(L_p)$ is a universal constant multiple of $(\log n)^{1/p-1/2}$ if $1\leqslant p\leqslant 2$ and $\max\{\min\{p,\log n\},((\log n)/\log\log n)^{1/2-1/p}\}$ if $p\geqslant 2$; when $1\leqslant p\leqslant 2$ this follows from [31] by considering as in Section 1.1.1 the discrete k-dimensional hypercube with the ℓ_p^k metric, and for $p\geqslant 2$ the $\Omega(\min\{p,\log n\})$ lower bound follows from [65] while the $\Omega(((\log n)/\log\log n)^{1/2-1/p}$ lower bound follows from [70] (one cannot consider for this purpose the planar graph example that was used in Section 1.1.1).

Thus, we know that (114) is optimal only when $(p-1)\log n=O(1)$ and $p\gtrsim\log n$ (and, trivially, when $p=2$), but it plausibly not optimal for the rest of the possible values of p. It would be worthwhile (and likely challenging) to obtain asymptotically sharp bounds here.

4. Localized radially weakly bi-Lipschitz embeddings

As we will see later, Definition 29 (correspondingly, the first part of Theorem 30) suffices for proving Theorem 1, Theorem 5, Theorem 9, Theorem 10, and Theorem 12 with a constant factor that has a much worse (exponential) dependence on p. The lower bound in the second part of Theorem 30 shows that such a loss is inherent to this approach. It is more delicate to get the stated dependence on p in the above theorems. For that purpose, we will next introduce a (quite subtle, but crucial) variant of Definition 29 which is interesting in its own right and likely useful for other purposes beyond its applications that we derive herein. We will prove that the Mazur map obeys the aforementioned variant with a much better dependence on p, and demonstrate that this new embedding notion preserves the separation modulus.

To motivate Definition 33 below, clarify its geometric meaning, and explain its nuanced difference from Definition 29, we will start by examining the following consequence of Definition 29.

Fix $K,\Delta>0$ and suppose that a metric space $(\mathcal{M},d_{\mathcal{M}})$ admits a K-localized weakly bi-Lipschitz embedding into a metric space $(\mathcal{N},d_{\mathcal{N}})$ with distortion D. Thus, for every $\Delta>0$ there is a nonconstant Lipschitz function $f_\Delta^z:\mathcal{M}\to\mathcal{N}$ for which (106) holds. If $x\in\mathcal{M}$ and $y,w\in B_{\mathcal{M}}(z,K\Delta)$ satisfy $d_{\mathcal{M}}(y,w)\geqslant\Delta$, then,

$$\frac{\|f_\Delta^z\|_{\mathrm{Lip}}}{D}\Delta\overset{(106)}{\leqslant} d_{\mathcal{N}}\big(f_\Delta^z(y),f_\Delta^z(w)\big)\leqslant 2\max\Big\{d_{\mathcal{N}}\big(f_\Delta^z(x),f_\Delta^z(y)\big),d_{\mathcal{N}}\big(f_\Delta^z(x),f_\Delta^z(w)\big)\Big\}.$$

Letting $B^\circ_{\mathcal{M}}(x,r)=\{y\in\mathcal{M}:\ d_{\mathcal{M}}(x,y)<r\}$ be the open $d_{\mathcal{M}}$-ball centered at $x\in\mathcal{M}$ of radius $r\geqslant 0$, we get:

$$\forall\, x\in\mathcal{M},\ \forall\, y,w\in B_{\mathcal{M}}(z,K\Delta),\qquad d_{\mathcal{M}}(y,w)\geqslant\Delta \implies \{y,w\}\not\subseteq \big(f^z_\Delta\big)^{-1}\Big(B^\circ_{\mathcal{N}}\big(f^z_\Delta(x),\frac{\|f^z_\Delta\|_{\mathrm{Lip}}}{2D}\Delta\big)\Big).$$

The contrapositive of this conclusion implies that

$$\forall\, x\in\mathcal{M},\qquad \mathrm{diam}_{\mathcal{M}}\bigg(B_{\mathcal{M}}(z,K\Delta)\cap\big(f^z_\Delta\big)^{-1}\Big(B^\circ_{\mathcal{N}}\big(f^z_\Delta(x),\frac{\|f^z_\Delta\|_{\mathrm{Lip}}}{2D}\Delta\big)\Big)\bigg)\leqslant\Delta. \tag{115}$$

Definition (33) below is important for our purposes. It modifies (115) in the following 5 ways. Firstly, it considers only closed balls and it replaces the weak inequality in (115) by a strict inequality; both of these modifications are essentially cosmetic as they have insignificant impact on the results that are obtained herein and they are introduced for ease of remembering the definition and to streamline its subsequent implementations. Definition (33) also changes the term $2D$ in (115) to D, which is merely a matter of notational convenience that impacts only the values of the (mostly implicit) universal constant factors in the results that are obtained herein. Next, it requires the nonconstant Lipschitz function f^z_Δ to be defined on $B_{\mathcal{M}}(z,K\Delta)$ rather than on all of $\mathcal{M}$; while this could be a genuine weakening when $\mathcal{M}$ is an arbitrary metric space, if $\mathcal{M}$ is a Banach space, then (as we recalled in the proof of Theorem 30) the standard normalization mapping yields a 2-Lipschitz retraction from $\mathcal{M}$ onto $B_{\mathcal{M}}(z,K\Delta)$, so by precomposing with this retraction we would obtain the same property for a function that is now defined on all of $\mathcal{M}$, at the cost of replacing D by $2D$; again, this would only impact the universal constant factors in the results that are obtained herein. Finally, Definition (33) replaces the $d_{\mathcal{M}}$-diameter in (115) by the $d_{\mathcal{M}}$-radius. While one might think that due to (25) this is also a minor modification, it is, in fact, substantial, as it leads to an exponential improvement of the dependence on p, as expressed in Lemma 34 below.

Definition 33. *Given $K,D>0$, we say that a metric space $(\mathcal{M},d_{\mathcal{M}})$ admits a K-localized* **radially** *weakly bi-Lipschitz embedding into a metric space $(\mathcal{N},d_{\mathcal{N}})$ with distortion D if for every $\Delta>0$ and every $z\in\mathcal{M}$ there exists a non-constant Lipschitz function $f^z_\Delta:B_{\mathcal{M}}(z,K\Delta)\to\mathcal{N}$ such that*

$$\forall\, x\in B_{\mathcal{M}}(z,K\Delta),\qquad \mathrm{rad}_{\mathcal{M}}\bigg(\big(f^z_\Delta\big)^{-1}\Big(B_{\mathcal{N}}\big(f^z_\Delta(x),\frac{\|f^z_\Delta\|_{\mathrm{Lip}}}{D}\Delta\big)\Big)\bigg)<\Delta. \tag{116}$$

By the above discussion, if a metric space $(\mathcal{M},d_{\mathcal{M}})$ admits a K-localized weakly bi-Lipschitz embedding into a metric space $(\mathcal{N},d_{\mathcal{N}})$ with distortion D, then for any $\varepsilon>0$ it also admits a $(1+\varepsilon)K$-localized radially weakly bi-Lipschitz embedding into $\mathcal{N}$ with distortion $2(1+\varepsilon)^2D$.

Conversely, suppose that there is $L\geqslant 1$ such that for every ball $B_{\mathcal{M}}(z,r)\subseteq\mathcal{M}$ there exists an L-Lipschitz retraction ρ^z_r from $\mathcal{M}$ onto $B_{\mathcal{M}}(z,r)$; if $\mathcal{M}$ is a Banach space, then we can take $L=2$. If $f:B_{\mathcal{M}}(z,K\Delta)\to\mathcal{N}$ is a nonconstant Lipschitz function satisfying (116), then the function $F^z_\Delta=f^z_\Delta\circ\rho^z_r:\mathcal{M}\to\mathcal{N}$ extends f^z_Δ and satisfies $\|F^z_\Delta\|_{\mathrm{Lip}}\leqslant L\|f^z_\Delta\|_{\mathrm{Lip}}$. Therefore, for every $x\in B_{\mathcal{M}}(z,K\Delta)$ we have:

$$\mathrm{diam}_{\mathcal{M}}\bigg(\big(f^z_\Delta\big)^{-1}\Big(B_{\mathcal{N}}\big(f^z_\Delta(x),\frac{\|F^z_\Delta\|_{\mathrm{Lip}}}{LD}\Delta\big)\Big)\bigg)\leqslant \mathrm{diam}_{\mathcal{M}}\bigg(\big(f^z_\Delta\big)^{-1}\Big(B_{\mathcal{N}}\big(f^z_\Delta(x),\frac{\|f^z_\Delta\|_{\mathrm{Lip}}}{D}\Delta\big)\Big)\bigg)\overset{(116)\wedge(25)}{<}2\Delta.$$

It follows that $y\notin(f^z_\Delta)^{-1}(B_{\mathcal{N}}\big(f^z_\Delta(x),\|F^z_\Delta\|_{\mathrm{Lip}}\Delta/(LD)\big)))$ whenever $x,y\in B_{\mathcal{M}}(z,K\Delta)$ satisfy $d_{\mathcal{M}}(x,y)\geqslant 2\Delta$. In particular, since F^z_Δ extends f^z_Δ, we have

$$\forall\, x,y\in B_{\mathcal{M}}(z,K\Delta)=B_{\mathcal{M}}\Big(z,\frac{K}{2}(2\Delta)\Big),\qquad d_{\mathcal{M}}(x,y)\geqslant 2\Delta\implies d_{\mathcal{N}}\big(F^z_\Delta(x),F^z_\Delta(y)\big)\geqslant\frac{\|F^z_\Delta\|_{\mathrm{Lip}}}{LD}\Delta=\frac{\|F^z_\Delta\|_{\mathrm{Lip}}}{2LD}2\Delta.$$

Recalling Definition 29, this observation shows that (under the assumption that an L-Lipschitz retraction from $\mathcal{M}$ onto any ball in $\mathcal{M}$ exists), if $\mathcal{M}$ admits a K-localized radially weakly bi-Lipschitz embedding into $\mathcal{N}$ with distortion D, then also $\mathcal{M}$ admits a $(K/2)$-localized weakly bi-Lipschitz embedding into $\mathcal{N}$ with distortion $2LD$. In particular, if $\mathcal{M}$ is a Banach space, then we deduce that it admits a $(K/2)$-localized weakly bi-Lipschitz embedding into $\mathcal{N}$ with distortion $4D$.

The above loss of a $2L$ factor in the distortion is of secondary importance if $L=O(1)$, as it only impacts the universal constant factors in the results that are obtained herein. However, the above reduction of K to $K/2$ is significant, as demonstrated by Proposition 34 below, which should be contrasted with the second part of Theorem 30, since it shows that if the word "radially" would be omitted from the statement of Proposition 34, then the distortion in its conclusion would have to be $e^{\Omega(p)}$.

Lemma 34. *If $p\geqslant 2$, then there are $K=K(p), D=D(p)>1$ satisfying $K-1\asymp 1/p$ and $D\asymp p$ such that L_p admits a K-localized* **radially** *weakly bi-Lipschitz embedding into L_2 with distortion D.*

Lemma 34 provides a novel property (potentially of use beyond its applications herein) of the Mazur map [67], whose classical definition was recalled in (29); the relevant property, which implies Lemma 34 and is, in fact, what Lemma 34 will prove, was already stated in the Introduction as the inclusion (31).

It is straightforward to check that $M_{p\to q}$ maps L_p bijectively onto L_q with $M_{p\to q}^{-1}=M_{q\to p}$. Furthermore, $\|M_{p\to q}(x)\|_{L_q}=\|x\|_{L_p}$ for every $x\in L_p$, and $M_{p\to q}(sx)=s^{p/q}\operatorname{sign}(s)M_{p\to q}(x)$ for every $x\in L_p$ and $s\in\mathbb{R}$. The property of the Mazur map that lies at the heart of Lemma 34 is the following inequality:

Lemma 35. *If $p\geqslant 2$, then for every $0<\alpha\leqslant \frac{1}{p}$ and every $0<\lambda<1$ we have*[10]

$$\forall x,y\in L_p,\qquad \|y-\alpha x\|_{L_p}^p\leqslant (1-\lambda\alpha)^p\|x\|_{L_p}^2+\frac{5}{(1-\lambda)p\alpha}\big\|M_{p\to 2}(y)-M_{p\to 2}(x)\big\|_{L_2}^2. \tag{117}$$

Note that since λ appears only in the right hand side of (117), the optimal way to use Lemma 35 is to apply (117) with the value $\lambda=\lambda(x,y,p,\alpha)$ that minimizes the right hand side of (117) over $0\leqslant\lambda\leqslant 1$.

Assuming the validity of Lemma 35 for the moment, we will next show how to deduce Lemma 34.

Proof of Lemma 34 assuming Lemma 35. We will start by demonstrating that there exists a universal constant $c>0$ such that the following inclusion holds for every $0<\alpha\leqslant 1/p$ and $0<\sigma,\lambda<1$:[11]

$$\forall x\in B_{L_p},\qquad M_{2\to p}\Big(B_{L_2}\Big(M_{p\to 2}(x), cp\alpha\sqrt{(1-\sigma)\lambda(1-\lambda)}\Big)\Big)\subseteq B_{L_p}\big(\alpha x,1-\sigma\lambda\alpha\big). \tag{118}$$

To verify (118), note first that by the mean value theorem there exists $\sigma\lambda\alpha\leqslant\theta\leqslant\lambda\alpha\leqslant\frac{\lambda}{p}$ such that

$$(1-\sigma\lambda\alpha)^p-(1-\lambda\alpha)^p=p(1-\theta)^{p-1}(1-\sigma)\lambda\alpha\geqslant(1-\sigma)\lambda p\alpha\left(1-\frac{\lambda}{p}\right)^{p-1}\geqslant\frac{\lambda}{e^{\lambda}}(1-\sigma)p\alpha\asymp\lambda(1-\sigma)p\alpha,$$

where we used the fact that $(p\geqslant 1)\mapsto(1-\lambda/p)^{p-1}$ is decreasing and tends to $e^{-\lambda}$ as $p\to\infty$. Therefore, if we define $r=r(\alpha,p,\sigma,\lambda)>0$ by

$$r\stackrel{\mathrm{def}}{=}\sqrt{\frac{(1-\lambda)p\alpha}{5}\big((1-\sigma\lambda\alpha)^p-(1-\lambda\alpha)^p\big)}, \tag{119}$$

then $r\gtrsim p\alpha\sqrt{(1-\sigma)\lambda(1-\lambda)}$. Consequently, the inclusion (118) would follow if we will show that

$$\forall x\in B_{L_p},\qquad M_{2\to p}\Big(B_{L_2}\big(M_{p\to 2}(x),r\big)\Big)\subseteq B_{L_p}\big(\alpha x,1-\sigma\lambda\alpha\big). \tag{120}$$

To justify (120) we need to prove that $\|M_{2\to p}(w)-\alpha x\|_{L_p}\leqslant 1-\sigma\lambda\alpha$ for every $w\in L_2$ satisfying

$$\|w-M_{p\to 2}(x)\|_{L_2}\leqslant r. \tag{121}$$

[10] The constant 5 in the right hand side of (117) is neither optimal nor does it have meaningful impact on the ensuing results.

[11] Stating (118) for general parameters α,λ,σ is beneficial because its proof is simpler to read without making choices that are arbitrary at this juncture, and also this could be relevant for future investigations. Below, we will use (118) for $\alpha=1/p$ and $\lambda=\sigma=1/2$, which suffices for the specific application herein even though these choices of λ,σ are not optimal for the ensuing reasoning; its optimization leads to a different settings of λ,σ (specifically, $\sigma=2/3$ and $\lambda=(5-\sqrt{13})/2=0.697...$ turn out to be best), but this only impacts the value of the (implicit) universal constant factors in our results.

This indeed holds thanks to the following application of Lemma 35 with $y = M_{2\to p}(w)$:

$$\|M_{2\to p}(w)-\alpha x\|_{L_p} \overset{(117)}{\leqslant} \left((1-\lambda\alpha)^p + \frac{5\|w-M_{p\to 2}(x)\|_{L_2}^2}{(1-\lambda)p\alpha}\right)^{\frac{1}{p}} \overset{(121)}{\leqslant} \left((1-\lambda\alpha)^p + \frac{5r^2}{(1-\lambda)p\alpha}\right)^{\frac{1}{p}} \overset{(119)}{=} 1-\sigma\lambda\alpha.$$

Having established (118), we will use it by noting that for $\alpha = \frac{1}{p}$ and $\lambda = \sigma = \frac{1}{2}$ it implies that

$$\forall x \in B_{L_p}, \qquad \mathrm{rad}_{L_p}\Big(M_{2\to p}\big(B_{L_2}(M_{p\to 2}(x),\gamma)\big)\Big) \leqslant 1-\frac{1}{4p}, \tag{122}$$

where $\gamma > 0$ is a universal constant. To deduce Lemma 34 from (122), fix $z \in L_p$ and $\Delta > 0$, and define

$$\forall x \in L_p, \qquad f_\Delta^z(x) \overset{\mathrm{def}}{=} K\Delta M_{p\to 2}\Big(\frac{1}{K\Delta}(x-z)\Big) \in L_2, \qquad \text{where} \qquad K = K(p) \overset{\mathrm{def}}{=} 1+\frac{1}{4p}. \tag{123}$$

Then, $f_\Delta^z : L_p \to L_2$ is a bijection whose inverse is given by

$$\forall w \in L_2, \qquad \big(f_\Delta^z\big)^{-1}(w) = z + K\Delta M_{2\to p}\Big(\frac{1}{K\Delta} w\Big). \tag{124}$$

Since $(K\Delta)^{-1}(x-z) \in B_{L_p}$ for every $x \in B_{L_p}(z, K\Delta)$, it follows that

$$\forall x \in B_{L_p}(z,K\Delta), \qquad \mathrm{rad}_{\mathcal{M}}\Big(\big(f_\Delta^z\big)^{-1}\big(B_{L_2}(f_\Delta^z(x),\gamma K\Delta)\big)\Big) \overset{(122)\wedge(123)\wedge(124)}{\leqslant} \Big(1-\frac{1}{4p}\Big)K\Delta \overset{(123)}{<} \Delta. \tag{125}$$

By [75, equation (5.32)] we have $\|M_{p\to 2}\|_{\mathrm{Lip}(L_p;L_2)} < p/\sqrt{2}$, so also $\|f_\Delta^z\|_{\mathrm{Lip}(L_p;L_2)} < p$ by (123). Consequently, for every $x \in L_p$ we have $B_{L_2}(f_\Delta^z(x),\gamma K\Delta) \supseteq B_{L_2}(f_\Delta^z(x),(\gamma K/p)\|f\|_{\mathrm{Lip}}\Delta))$, so, recalling Definition 33, we see that (125) implies that L_p admits a K-localized radially weakly bi-Lipschitz embedding into L_2 with distortion D, where $K = K(p)$ is given in (123), so $K-1 \asymp 1/p$, and $D = D(p) = p/(K\gamma) \asymp p$. □

Remark 36. The above proof of Lemma 34 yields (31) for $\beta = \sqrt{\sqrt[4]{e}-1}/\sqrt{5e} \in [0.14, 0.15]$, as seen from (120) while recalling (119) with our choices $\alpha = 1/p$ and $\lambda = \sigma = 1/2$, and using $\|M_{p\to 2}\|_{\mathrm{Lip}(L_p;L_2)} < p/\sqrt{2}$.

By a straightforward tensorization argument, Lemma 35 can be deduced from its one-dimensional counterpart, which amounts to the following numerical fact:

Lemma 37. *If $p \geqslant 2$, then the following estimate holds for every $0 < \alpha \leqslant \frac{1}{p}$ and every $0 < \lambda < 1$:*

$$\forall u, v \in \mathbb{R}, \qquad \Big||u+v|^{\frac{2}{p}}\,\mathrm{sign}(u+v) - \alpha|u|^{\frac{2}{p}}\,\mathrm{sign}(v)\Big|^p \leqslant (1-\lambda\alpha)^p u^2 + \frac{5v^2}{(1-\lambda)p\alpha}. \tag{126}$$

We will next explain how to quickly deduce Lemma 35 from Lemma 37:

Proof of Lemma 35 assuming Lemma 37. Fixing $x, y \in L_p$, define $u = u(x), v = v(x,y) \in L_2$ by

$$u \overset{\mathrm{def}}{=} M_{p\to 2}(x) \qquad \text{and} \qquad v \overset{\mathrm{def}}{=} M_{p\to 2}(y) - M_{p\to 2}(x). \tag{127}$$

Then,

$$\begin{aligned}
\|y-\alpha x\|_{L_p}^p &= \int_0^1 |y(t)-\alpha x(t)|^p \,\mathrm{d}t \\
&\overset{(29)\wedge(127)}{=} \int_0^1 \Big||u(t)+v(t)|^{\frac{2}{p}}\,\mathrm{sign}\big(u(t)+v(t)\big) - \alpha|u(t)|^{\frac{2}{p}}\,\mathrm{sign}\big(v(t)\big)\Big|^p \,\mathrm{d}t \\
&\overset{(126)}{\leqslant} (1-\lambda\alpha)^p\|u\|_{L_2}^2 + \frac{5}{(1-\lambda)p\alpha}\|v\|_{L_2}^2 \\
&\overset{(127)}{=} (1-\lambda\alpha)^p\|M_{p\to 2}(x)\|_{L_2}^2 + \frac{5}{(1-\lambda)p\alpha}\big\|M_{p\to 2}(y)-M_{p\to 2}(x)\big\|_{L_2}^2.
\end{aligned}$$

This coincides with (117) because $\|M_{p\to 2}(x)\|_{L_2} = \|x\|_{L_p}$. □

Proof of Lemma 37. If $u = 0$, then (126) holds (with room to spare), so assume $u \neq 0$. By normalization, it suffices to prove (126) when $u = 1$, i.e., our goal is equivalent to establishing the following estimate:

$$\forall v \in \mathbb{R}, \qquad \Big| |1+v|^{\frac{2}{p}} \operatorname{sign}(1+v) - \alpha \Big|^p \leqslant (1-\lambda\alpha)^p + \frac{5v^2}{(1-\lambda)p\alpha}. \tag{128}$$

Suppose first that $v \geqslant -1+\alpha^{p/2}$, in which case our goal (128) becomes the following inequality:

$$(1+v)^{\frac{2}{p}} \leqslant \alpha + \left((1-\lambda\alpha)^p + \frac{5v^2}{(1-\lambda)p\alpha}\right)^{\frac{1}{p}}. \tag{129}$$

If also $|v| \leqslant \sqrt{15(1-\lambda)p\alpha/20}$, which will see is when (129) is most meaningful, then we proceed as follows. The function $(t>0) \mapsto (1+t)^{1/t}$ is decreasing, so $(1+t)^{1/t} \geqslant (19/4)^{4/15} > e^{2/5}$ for $0 < t \leqslant 15/4$. Hence, $(1+t)^{1/p} \geqslant e^{2t/(5p)} \geqslant 1+2t/(5p)$ for $0 \leqslant t \leqslant 15/4$. Using this for $t = 5v^2/((1-\lambda)p\alpha) \leqslant 15/4$, we get

$$\left(1+\frac{5v^2}{(1-\lambda)p\alpha}\right)^{\frac{1}{p}} \geqslant 1+\frac{2v^2}{(1-\lambda)p^2\alpha}. \tag{130}$$

We can therefore bound the right hand side of (129) from below as follows:

$$\begin{aligned}\alpha + \left((1-\lambda\alpha)^p + \frac{5v^2}{(1-\lambda)p\alpha}\right)^{\frac{1}{p}} &\geqslant \alpha + (1-\lambda\alpha)\left(1+\frac{5v^2}{(1-\lambda)p\alpha}\right)^{\frac{1}{p}} \overset{(130)}{\geqslant} 1+(1-\lambda)\alpha + \frac{2(1-\lambda\alpha)v^2}{(1-\lambda)p^2\alpha}\\ &\geqslant 1+(1-\lambda)\alpha + \frac{v^2}{(1-\lambda)p^2\alpha} = 1+\frac{2}{p}v + \left(\frac{v}{p\sqrt{(1-\lambda)\alpha}} - \sqrt{(1-\lambda)\alpha}\right)^2 \geqslant 1+\frac{2}{p}v \geqslant (1+v)^{\frac{2}{p}},\end{aligned} \tag{131}$$

where the third step of (131) holds because $1-\lambda\alpha \geqslant 1-1/p \geqslant 1/2$, as $0<\lambda<1$, $0 \leqslant \alpha \leqslant 1/p$ and $p \geqslant 2$, and the final step of (131) holds because $0 < 2/p \leqslant 1$, so the function $(v > -1) \mapsto 1+2v/p-(1+y)^{2/p}$ attains its global minimum when $v = 0$, where it vanishes. We have thus completed the verification of (129) when $|v| \leqslant \sqrt{15(1-\lambda)p\alpha/20}$. If $\sqrt{15(1-\lambda)p\alpha/20} < |v| \leqslant 1$, then (129) holds because

$$(1+v)^{\frac{2}{p}} \leqslant 2^{\frac{2}{p}} = \left(\frac{1}{4}+\frac{15}{4}\right)^{\frac{1}{p}} \leqslant \left(\frac{1}{4}+\frac{5v^2}{(1-\lambda)p\alpha}\right)^{\frac{1}{p}} < \alpha + \left((1-\lambda\alpha)^p + \frac{5v^2}{(1-\lambda)p\alpha}\right)^{\frac{1}{p}}, \tag{132}$$

where in the last step of (132) we used that $(1-\lambda\alpha)^p \geqslant (1-1/p)^p \geqslant 1/4$, as $\lambda\alpha \leqslant \alpha \leqslant 1/p \leqslant 1/2$. The remaining case of (129) is when $v \geqslant 1$, which holds (with room to spare) because in this case we have

$$(1+v)^{\frac{2}{p}} \leqslant (2v)^{\frac{2}{p}} < \left(5v^2\right)^{\frac{1}{p}} < \alpha + \left((1-\lambda\alpha)^p + \frac{5v^2}{(1-\lambda)p\alpha}\right)^{\frac{1}{p}},$$

where the last step is valid because $(1-\lambda)p\alpha < p\alpha \leqslant 1$.

It remains to check (128) for $v < -1+\alpha^{p/2}$, in which case set $w = -v > 1-\alpha^{p/2} > 0$ and (128) becomes:

$$\alpha - |w-1|^{\frac{2}{p}} \operatorname{sign}(1-w) \leqslant \left((1-\lambda\alpha)^p + \frac{5w^2}{(1-\lambda)p\alpha}\right)^{\frac{1}{p}}. \tag{133}$$

But (133) is very crude since $-|w-1|^{2/p}\operatorname{sign}(1-w) < (5/((1-\lambda)p\alpha))^{1/p} w^{2/p}$, as $5/((1-\lambda)p\alpha) > 5/(p\alpha) > 5$, and also $\alpha \leqslant (1-\alpha) < (1-\lambda\alpha)$, as $\alpha \leqslant 1/p \leqslant 1/2$. This completes the proof of Lemma 37. □

5. Localization and induction on scales for separated random partitions

Here we will prove the general localization and induction on scales principle that was formulated as Lemma 16. All of the relevant definitions were provided in the Introduction. In particular, the notion of radially bounded separating random partitions was introduced in Section 1.3.3.

Proof of Lemma 16. We will prove that the following inequality holds for every $K>1$ and every $\Delta,\varepsilon>0$, even without the assumption (27) of Lemma 16, namely for any separable metric space $(\mathcal{M},d_{\mathcal{M}})$:

$$\widehat{\mathsf{SEP}}_{\Delta}(\mathcal{C};\mathcal{M})\leqslant\frac{1}{K}\widehat{\mathsf{SEP}}_{K\Delta}(\mathcal{C};\mathcal{M})+\sup_{z\in\mathcal{M}}\widehat{\mathsf{SEP}}_{\Delta}\big(\mathcal{C}\cap B_{\mathcal{M}}(z,K\Delta+\varepsilon);\mathcal{M}\big). \tag{134}$$

Accepting the validity of (134) for the moment (its justification appears below), we will next proceed to explain how to use it to quickly deduce Lemma 16.

Suppose first that $\widehat{\mathsf{SEP}}_{\Delta}(\mathcal{C};\mathcal{M})<\infty$, which is when (28) is most meaningful. Then, $\widehat{\mathsf{SEP}}_{\Delta'}(\mathcal{C};\mathcal{M})<\infty$ for every $\Delta'\geqslant\Delta$. We may therefore rearrange the limit as $\varepsilon\to 0^+$ of (134) with Δ replaced by $K^s\Delta$ for every integer $s\geqslant 0$ to obtain the following recursive estimate:

$$\frac{1}{K^s}\widehat{\mathsf{SEP}}_{K^s\Delta}(\mathcal{C};\mathcal{M})-\frac{1}{K^{s+1}}\widehat{\mathsf{SEP}}_{K^{s+1}\Delta}(\mathcal{C};\mathcal{M})\leqslant\frac{1}{K^s}\lim_{\varepsilon\to 0^+}\sup_{z\in\mathcal{M}}\widehat{\mathsf{SEP}}_{K^s\Delta}\big(\mathcal{C}\cap B_{\mathcal{M}}(z,K^{s+1}\Delta+\varepsilon);\mathcal{M}\big). \tag{135}$$

Thanks to the assumption (27), by summing (135) over $s\in\mathbb{N}\cup\{0\}$ and telescoping we get (28), as $K>1$.

If $\widehat{\mathsf{SEP}}_{\Delta}(\mathcal{C};\mathcal{M})=\infty$, then we need to demonstrate that the right hand side of (28) is also infinite. This is so because the assumption (27) implies in particular that there is $s\in\mathbb{N}$ for which $\mathsf{SEP}_{K^s\Delta}(\mathcal{C};\mathcal{M})<\infty$, whence also $\widehat{\mathsf{SEP}}_{K^s\Delta}(\mathcal{C};\mathcal{M})<\infty$ by (26). We can therefore consider the largest nonnegative integer s_0 for which $\widehat{\mathsf{SEP}}_{K^{s_0}\Delta}(\mathcal{C};\mathcal{M})=\infty$. By applying (134) with Δ replaced by $K^{s_0}\Delta$, we see that since the left hand side of that inequality is infinite while the first term in the right hand side of that inequality is finite by the maximality of s_0, necessarily $\sup_{z\in\mathcal{M}}\widehat{\mathsf{SEP}}_{K^{s_0}\Delta}(\mathcal{C}\cap B_{\mathcal{M}}(z,K^{s_0+1}\Delta+\varepsilon);\mathcal{M})=\infty$ for every $\varepsilon>0$. Therefore the s_0-summand in the right hand side of (28) is infinite, as required.

To prove (134), fix $\varepsilon>0$. Define $\sigma,\tau>0$ by

$$\sigma\stackrel{\mathrm{def}}{=}\widehat{\mathsf{SEP}}_{K\Delta}(\mathcal{C};\mathcal{M})\qquad\text{and}\qquad\tau\stackrel{\mathrm{def}}{=}\sup_{z\in\mathcal{M}}\widehat{\mathsf{SEP}}_{\Delta}\big(\mathcal{C}\cap B_{\mathcal{M}}(z,K\Delta+\varepsilon);\mathcal{M}\big). \tag{136}$$

We may assume from now that $\sigma,\tau<0$ because otherwise (134) is vacuous.

Fix $\eta>0$. The definition of σ in (136) yields the existence of a probability space (Ω_0,μ_0) and a sequence of strongly measurable mappings

$$\big\{\Phi^i:\Omega_0\to 2^{\mathcal{C}}\big\}_{i=1}^{\infty}$$

satisfying

$$\forall(i,\omega_0)\in\mathbb{N}\times\Omega_0\qquad\mathrm{rad}_{\mathcal{M}}\big(\Phi^i(\omega_0)\big)\leqslant K\Delta, \tag{137}$$

and furthermore if we define

$$\forall\omega_0\in\Omega_0,\qquad\mathcal{Q}^{\omega_0}\stackrel{\mathrm{def}}{=}\big\{\Phi^i(\omega_0)\big\}_{i=1}^{\infty}, \tag{138}$$

then $\mathcal{Q}^{\omega_0}$ is a partition of $\mathcal{C}$ for each $\omega_0\in\Omega_0$, and the following requirement holds:

$$\forall x,y\in\mathcal{C},\qquad\mu_0\big[\omega_0\in\Omega_0:\ \mathcal{Q}^{\omega_0}(x)\neq\mathcal{Q}^{\omega_0}(y)\big]\leqslant\frac{\sigma+\eta}{K\Delta}d_{\mathcal{M}}(x,y). \tag{139}$$

Because $\mathcal{M}$ is separable, we can fix a sequence $\{u_n\}_{n=1}^{\infty}$ that is dense in $\mathcal{M}$. Thanks to (137) we can then define a sequence of random indices $\{n^i:\Omega_0\to\mathbb{N}\}_{i=1}^{\infty}\subseteq\mathbb{N}$ as follows:

$$\forall\omega_0\in\Omega,\qquad n^i(\omega_0)\stackrel{\mathrm{def}}{=}\min\big\{n\in\mathbb{N}:\ \Phi^i(\omega_0)\subseteq B_{\mathcal{M}}(u_n,K\Delta+\varepsilon)\big\}. \tag{140}$$

Note in passing (so that we could use it freely later) that the measurability of all of these random indices is a quick consequence of the strong measurability of each of $\{\Phi^i:\Omega_0\to 2^{\mathcal{C}}\}_{i=1}^{\infty}$. Indeed, for every $i,n\in\mathbb{N}$,

$$\begin{aligned}\big\{\omega_0\in\Omega_0:n^i(\omega_0)=n\big\}&\\&=\big\{\omega_0\in\Omega_0:\ \Phi^i(\omega_0)\cap\big(\mathcal{C}\smallsetminus B_{\mathcal{M}}(u_n,K\Delta+\varepsilon)\big)=\emptyset\big\}\bigcap\\&\qquad\left(\bigcap_{k=1}^{n-1}\big\{\omega_0\in\Omega_0:\ \Phi^i(\omega_0)\cap\big(\mathcal{C}\smallsetminus B_{\mathcal{M}}(u_k,K\Delta+\varepsilon)\big)\neq\emptyset\big\}\right).\end{aligned}$$

Next, by the definition of τ in (134) for every $n\in\mathbb{N}$ there is a probability space (Ω_n,μ_n) and a sequence of strongly measurable mappings

$$\left\{\Psi_n^j:\Omega_n\to 2^{\mathcal{C}\cap B_{\mathcal{M}}(u_n,K\Delta+\varepsilon)}\right\}_{j=1}^{\infty}$$

satisfying

$$\forall (j,n)\in\mathbb{N}\times\mathbb{N},\ \forall\omega_n\in\Omega_n\qquad \mathrm{rad}_{\mathcal{M}}\big(\Psi_n^j(\omega_n)\big)\leqslant\Delta, \tag{141}$$

and furthermore if we define

$$\forall n\in\mathbb{N},\ \forall \omega_n\in\Omega_n,\qquad \mathcal{R}_n^{\omega_n}\stackrel{\mathrm{def}}{=}\{\Psi_n^j(\omega_n)\}_{i=1}^{\infty}, \tag{142}$$

then $\mathcal{R}_n^{\omega_n}$ is a partition of $\mathcal{C}\cap B_{\mathcal{M}}(u_n,K\Delta+\varepsilon)$ for each $\omega_n\in\Omega_n$, and the following requirement holds:

$$\forall x,y\in\mathcal{C}\cap B_{\mathcal{M}}(u_n,K\Delta+\varepsilon),\qquad \mu_n\big[\omega_n\in\Omega_n:\ \mathcal{R}_n^{\omega_n}(x)\neq\mathcal{R}_n^{\omega_n}(y)\big]\leqslant\frac{\tau+\eta}{\Delta}d_{\mathcal{M}}(x,y). \tag{143}$$

We will henceforth work with the product space (Ω,μ) that is given by:

$$\Omega\stackrel{\mathrm{def}}{=}\prod_{n=0}^{\infty}\Omega_n\qquad\text{and}\qquad \mu\stackrel{\mathrm{def}}{=}\bigotimes_{n=0}^{\infty}\mu_n. \tag{144}$$

For every $i,j\in\mathbb{N}$ define $\Gamma^{i,j}:\Omega\to 2^{\mathcal{C}}$ by:

$$\forall (i,j)\in\mathbb{N}\times\mathbb{N},\ \forall\omega=(\omega_0,\omega_1,\ldots)\in\Omega,\qquad \Gamma^{i,j}(\omega)\stackrel{\mathrm{def}}{=}\Phi^i(\omega_0)\cap\Psi^j_{n^i(\omega_0)}\big(\omega_{n^i(\omega_0)}\big). \tag{145}$$

The strong measurability of $\Gamma^{i,j}$ follows from the assumed strong measurability of $\{\Psi_n^j\}_{j,n=1}^{\infty}$, together with the measurability of $\{n^i\}_{i=1}^{\infty}$ that we verified above, as a consequence of the assumed strong measurability of $\{\Phi^i\}_{i=1}^{\infty}$. Indeed, suppose that $E\subseteq\mathcal{C}$ is closed. For $\omega=(\omega_0,\omega_1,\ldots)\in\Omega$, definition (145) shows that $\Gamma^{i,j}(\omega)\cap E\neq\emptyset$ if and only if $\Psi^j_{n^i(\omega_0)}\big(\omega_{n^i(\omega_0)}\big)\cap E\neq\emptyset$, as $E\subseteq\mathcal{C}$ and $\{\Phi^i(\omega_0)\}_{i=1}^{\infty}$ is a partition of $\mathcal{C}$. Thus,

$$\begin{aligned}&\big\{\omega\in\Omega:\ \Gamma^{i,j}(\omega)\cap E\neq\emptyset\big\}\\ &\quad=\bigcup_{i=1}^{\infty}\bigcup_{j=1}^{\infty}\bigcup_{n=1}^{\infty}\big\{\omega_0\in\Omega_0:\ n^i(\omega_0)=n\big\}\times\Omega_1\times\ldots\times\Omega_{n-1}\times\big\{\omega_n\in\Omega_n:\ \Psi_n^j(\omega_n)\cap E\neq\emptyset\big\}\times\prod_{k=n+1}^{\infty}\Omega_k.\end{aligned}$$

Therefore, we can consider the random partition of $\mathcal{C}$ that is given by

$$\forall\omega=(\omega_0,\omega_1,\ldots)\in\Omega,\qquad \mathcal{P}^{\omega}\stackrel{\mathrm{def}}{=}\big\{\Gamma^{i,j}(\omega)\big\}_{i,j=1}^{\infty},$$

which satisfies $\mathrm{rad}_{\mathcal{M}}(\mathcal{P}^{\omega}(x))\leqslant\Delta$ for every $x\in\mathcal{C}$ thanks to (141) and (145). Finally, every $x,y\in\mathcal{C}$ satisfy:

$$\begin{aligned}&\mu\big[\omega\in\Omega:\ \mathcal{P}^{\omega}(x)\neq\mathcal{P}^{\omega}(y)\big]\\ &\stackrel{(145)}{=}1-\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\mu\big[\omega\in\Omega:\ \Phi^i(\omega_0)\supseteq\{x,y\}\quad\text{and}\quad\Psi^j_{n^i(\omega_0)}\big(\omega_{n^i(\omega_0)}\big)\supseteq\{x,y\}\big]\\ &=1-\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\sum_{n=1}^{\infty}\mu\big[\omega\in\Omega:\ \Phi^i(\omega_0)\supseteq\{x,y\}\quad\text{and}\quad n^i(\omega_0)=n\quad\text{and}\quad\Psi_n^j(\omega_n)\supseteq\{x,y\}\big]\\ &\stackrel{(144)}{=}1-\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\sum_{n=1}^{\infty}\mu_0\big[\omega\in\Omega_0:\ \Phi^i(\omega_0)\supseteq\{x,y\}\quad\text{and}\quad n^i(\omega_0)=n\big]\mu_n\big[\omega_n\in\Omega_n:\ \Psi_n^j(\omega_n)\supseteq\{x,y\}\big]\\ &\stackrel{(142)}{=}1-\sum_{n=1}^{\infty}\Big(1-\mu_n\big[\omega_n\in\Omega_n:\ \mathcal{R}_n^{\omega_n}(v)\neq\mathcal{R}_n^{\omega_n}(y)\big]\Big)\sum_{i=1}^{\infty}\mu_0\big[\omega\in\Omega_0:\ \Phi^i(\omega_0)\supseteq\{x,y\}\quad\text{and}\quad n^i(\omega_0)=n\big]\\ &\stackrel{(143)\wedge(138)}{\leqslant}1-\Big(1-\min\Big\{1,\frac{\tau+\eta}{\Delta}d_{\mathcal{M}}(x,y)\Big\}\Big)\mu_0\big[\omega_0\in\Omega_0:\ \mathcal{Q}^{\omega_0}(x)\neq\mathcal{Q}^{\omega_0}(y)\big]\\ &\stackrel{(139)}{\leqslant}1-\Big(1-\min\Big\{1,\frac{\tau+\eta}{\Delta}d_{\mathcal{M}}(x,y)\Big\}\Big)\Big(1-\min\Big\{1,\frac{\sigma+\eta}{K\Delta}d_{\mathcal{M}}(x,y)\Big\}\Big)\leqslant\frac{\frac{\sigma+\eta}{K}+\tau+\eta}{\Delta}d_{\mathcal{M}}(x,y).\end{aligned}$$

By taking the limit of this estimate as $\eta \to 0^+$ we conclude that (134) indeed holds. $\square$

6. Analytic separation is preserved under localized radially weakly bi-Lipschitz embeddings

We will need to impose a stronger measurability requirement from random partitions to be able to easily use localized radially weakly bi-Lipschitz embeddings to transfer separating random partitions from one metric space to another; this is the mechanism by which the proof of Theorem 14 will be completed.

Given a Polish metric space $(\mathcal{M}, d_{\mathcal{M}})$ and a probability space $(\Omega, \mathbb{P})$, call a sequence of set-valued mappings $\{\Gamma^i : \Omega \to 2^{\mathcal{M}}\}_{i=1}^{\infty}$ an analytic random partition of $\mathcal{M}$ if for every $i \in \mathbb{N}$ and every analytic subset A of $\mathcal{M}$ the set $\{\omega \in \Omega :\ A \cap \Gamma^i(\omega)\}$ is $\mathbb{P}$-measurable. For this, we recall the standard terminology that a metric space is called Polish if it is separable and complete and a subset of a Polish metric space is analytic if it is a continuous image of a Polish metric space; see [45] for a thorough treatment and [59] for the history.

Because closed subsets of a Polish metric space are analytic, any analytic random partition is in particular a random partition per the definition that we recalled in Section 1.3. The following theorem from [80] provides examples of analytic random partitions with good (optimal) separation properties:

Theorem 38. *For each $k \in \mathbb{N}$ and $1 \leqslant p \leqslant \infty$ there is $1 \leqslant \sigma = \sigma(p,k) \lesssim k^{\max\left\{\frac{1}{p}, \frac{1}{2}\right\}}$ such that for any $\Delta > 0$ there exists an analytic random partition of ℓ_p^k that is Δ-bounded and σ-separating.*

The Introduction of [80] states Theorem 38 without mentioning that the corresponding random partitions are analytic, but their analyticity is stated in [80, Lemma 119], which is what is applied in the proof of [80, Lemma 125] to derive the measurability of the random partition that Theorem 38uses.[12] We will use below only the case $p = 2$ of Theorem 38, for which the underlying construction for finite subsets of ℓ_2^k is due to [23], and its extension to random partitions of all of ℓ_2^k is due to [53, 80]. The short proof of following basic and useful lemma clarifies why it is beneficial to consider analytic random partitions:

Lemma 39. *Fix $\sigma, \Delta, R, L > 0$. Let $(\mathcal{N}, d_{\mathcal{N}})$ be a Polish metric space admitting an analytic random partition that is (LR)-bounded and σ-separating. Suppose that $(\mathcal{M}, d_{\mathcal{M}})$ is a Polish metric space, $\mathcal{S} \subseteq \mathcal{M}$ is a Borel subset of $\mathcal{M}$, and that $\varphi : \mathcal{S} \to \mathcal{N}$ is an L-Lipschitz function that satisfies the following property:*

$$\forall x \in \mathcal{S}, \qquad \operatorname{rad}_{\mathcal{M}}\Big(\varphi^{-1}\big(B_{\mathcal{N}}(\varphi(x), LR)\big)\Big) \leqslant \Delta. \tag{146}$$

Then, $(\mathcal{S}, d_{\mathcal{M}})$ admits an analytic random partition that is Δ-radially bounded and $\sigma\frac{\Delta}{R}$-separating. Hence,

$$\widehat{\mathsf{SEP}}_{\Delta}(\mathcal{S}; \mathcal{M}) \leqslant \sigma \frac{\Delta}{R}.$$

Proof. Fix a probability space (Ω, μ) and an analytic random partition

$$\mathcal{P} = \left\{\Gamma^i : \Omega \to 2^{\mathcal{N}}\right\}_{i=1}^{\infty} \tag{147}$$

that is (LR)-bounded and σ-separating. If $E \subseteq \mathcal{S}$ is analytic, then $\varphi(E) \subseteq \mathcal{N}$ is analytic as φ is continuous, whence by the assumed analyticity of the random partition (147) for every $i \in \mathbb{N}$ the set

$$\left\{\omega \in \Omega :\ \varphi^{-1}\big(\Gamma^i(\omega)\big) \cap E \neq \emptyset\right\} = \left\{\omega \in \Omega :\ \Gamma^i(\omega) \cap \varphi(E) \neq \emptyset\right\}$$

is μ-measurable. So, the following sequence of set-valued mappings is an analytic random partition of $\mathcal{S}$:

$$\mathcal{Q} = \left\{(\omega \in \Omega) \mapsto \mathcal{Q}^{\omega}\right\} \stackrel{\mathrm{def}}{=} \left\{(\omega \in \Omega) \mapsto \varphi^{-1}\big(\Gamma^i(\omega)\big) \subseteq \mathcal{S}\right\}_{i=1}^{\infty}, \tag{148}$$

A different way to write (148) is:

$$\forall \omega \in \Omega,\ \forall x \in \mathcal{S}, \qquad \mathcal{Q}^{\omega}(x) \stackrel{(147)\wedge(148)}{=} \varphi^{-1}\Big(\mathcal{P}^{\omega}\big(\varphi(x)\big)\Big). \tag{149}$$

---

[12] To notice that the statement of [80, Lemma 119] provides the measurability that we need, recall the important classical theorem of Luzin [61] (see also e.g. [45, Theorem 21.10]) that analytic sets are universally measurable, i.e., they are measurable with respect to every complete σ-finite Borel measure on the given Polish metric space.

Because $\mathcal{P}$ is (RL)-bounded by assumption, $\operatorname{diam}_{\mathcal{N}}(\mathcal{P}^\omega(\varphi(x))) \leqslant RL$ for every $\omega \in \Omega$ and $x \in \mathcal{M}$, whence $\mathcal{P}^\omega(\varphi(x)) \subseteq B_{\mathcal{N}}(\varphi(x), RL)$. Thanks to (149), this implies that $\mathcal{Q}^\omega(x)$ is contained in $\varphi^{-1}(B_{\mathcal{N}}(\varphi(x), RL))$. By invoking the assumption (146) we conclude that the random partition $\mathcal{Q}$ is Δ-radially bounded with respect to the super-space $(\mathcal{M}, d_{\mathcal{M}})$. Finally, for every $x, y \in \mathcal{S}$ we have

$$\begin{aligned} \mu\big[\omega \in \Omega : \mathcal{Q}^\omega(x) \neq \mathcal{Q}^\omega(y)\big] &\overset{(149)}{=} \mu\Big[\omega \in \Omega : \mathcal{P}^\omega\big(\varphi(x)\big) \neq \mathcal{P}^\omega\big(\varphi(y)\big)\Big] \\ &\leqslant \sigma \frac{d_{\mathcal{N}}\big(\varphi(x), \varphi(y)\big)}{LR} \leqslant \Big(\frac{\sigma\Delta}{R}\Big)\frac{d_{\mathcal{M}}(x,y)}{\Delta}, \end{aligned} \tag{150}$$

where the second step of (150) uses the assumption that $\mathcal{P}$ is σ-separating and (LR)-bounded, and the third step of (150) uses the assumption that φ is L-Lipschitz. We already checked that $\mathcal{Q}$ is Δ-radially bounded with respect to $(\mathcal{M}, d_{\mathcal{M}})$, so it follows from (150) that $\widehat{\mathsf{SEP}}_\Delta(\mathcal{S}; \mathcal{M}) \leqslant \sigma\Delta/R$, as required. $\square$

The following lemma sets the stage for our subsequent application of Lemma 39; its short proof proceed by combining two important Euclidean results, namely the Kirszbraun Lipschitz extension theorem [46] and the Johnson–Lindenstrauss dimension lemma [41].

Lemma 40. *For any $\mathcal{C} \subseteq L_2$ with $2 \leqslant |\mathcal{C}| < \infty$ there is an integer $1 \leqslant k \lesssim \log|\mathcal{C}|$ and $H : L_2 \to \ell_2^k$ satisfying:*

$$\forall x, y \in L_2, \qquad \frac{1}{2}\|x-y\|_{L_2} - \frac{3}{2}d_{L_2}(x, \mathcal{C}) - \frac{3}{2}d_{L_2}(y, \mathcal{C}) \leqslant \|H(x) - H(y)\|_{\ell_2^k} \leqslant \|x-y\|_{L_2}. \tag{151}$$

Therefore, the following inclusion holds for any $r > 0$ and any point x in the r-neighborhood $B_{L_2}(\mathcal{C}, r)$ of $\mathcal{C}$:

$$B_{L_2}(\mathcal{C}, r) \cap H^{-1}\big(B_{\ell_2^k}(H(x), r)\big) \subseteq B_{L_2}(x, 8r). \tag{152}$$

For (152), recall our (nonstandard) notation (18) for neighborhoods of subsets in a metric space.

Proof of Lemma 40. (152) follows from the first inequality in (151). Indeed, fix $r > 0$ and $x \in B_{L_2}(\mathcal{C}, r)$, i.e., $d_{L_2}(x, \mathcal{C}) \leqslant r$. Consider any point y that belongs to the set that appears in the left hand side of(152). Thus, $y \in B_{L_2}(\mathcal{C}, r)$, i.e., $d_{L_2}(y, \mathcal{C}) \leqslant r$, and $H(y) \in B_{\ell_2^k}(H(x), r)$, i.e. $\|H(x) - H(y)\|_{L_2} \leqslant r$. By substituting these 3 bounds into the first inequality in (151) and rearranging we arrive at $\|x-y\|_{L_2} \leqslant 8r$, i.e., $y \in B_{L_2}(x, 8r)$, as required. Lemma 40 will therefore be proven once we establish (40), which we will proceed to do next.

The Johnson–Lindenstrauss lemma [41] yields a positive integer $k \lesssim \log|\mathcal{C}|$ and $h : L_2 \to \ell_2^k$ such that

$$\forall x, y \in \mathcal{C}, \qquad \frac{1}{2}\|x-y\|_{L_2} \leqslant \|h(x) - h(y)\|_{\ell_2^k} \leqslant \|x-y\|_{L_2}. \tag{153}$$

Now, Kirszbraun's Lipschitz extension theorem [46] provides a function $H : L_2 \to \ell_2^k$ satisfying:

$$\forall a \in \mathcal{C}, \quad H(a) = h(a) \qquad \text{and} \qquad \forall x, y \in L_2, \quad \|H(x) - H(y)\|_{\ell_2^k} \leqslant \|x-y\|_{L_2}. \tag{154}$$

Given $x, y \in L_2$, the second inequality in (151), i.e., the fact that H is 1-Lipschitz, coincides with the second condition in (154). For the first inequality in (151) fix $a, b \in L_2$ such that

$$a, b \in \mathcal{C} \qquad \text{and} \qquad \|x-a\|_{L_2} = d_{L_2}(x, \mathcal{C}) \qquad \text{and} \qquad \|y-b\|_{L_2} = d_{L_2}(y, \mathcal{C}). \tag{155}$$

Then,

$$\begin{aligned} \|x-y\|_{L_2} &\leqslant \|x-a\|_{L_2} + \|a-b\|_{L_2} + \|b-y\|_{L_2} \\ &\overset{(153)\wedge(155)}{\leqslant} d_{L_2}(x, \mathcal{C}) + 2\|h(a) - h(b)\|_{\ell_2^k} + d_{L_2}(y, \mathcal{C}) \\ &\overset{(154)\wedge(155)}{=} d_{L_2}(x, \mathcal{C}) + 2\|H(a) - H(b)\|_{\ell_2^k} + d_{L_2}(y, \mathcal{C}) \\ &\leqslant d_{L_2}(x, \mathcal{C}) + 2\big(\|H(a) - H(x)\|_{\ell_2^k} + \|H(x) - H(y)\|_{\ell_2^k} + \|H(y) - H(b)\|_{\ell_2^k}\big) + d_{L_2}(y, \mathcal{C}) \\ &\overset{(154)}{\leqslant} d_{L_2}(x, \mathcal{C}) + 2\big(\|a-x\|_{L_2} + \|H(x) - H(y)\|_{\ell_2^k} + \|y-b\|_{L_2}\big) + d_{L_2}(y, \mathcal{C}) \\ &\overset{(155)}{=} d_{L_2}(x, \mathcal{C}) + 2\big(d_{L_2}(x, \mathcal{C}) + \|H(x) - H(y)\|_{\ell_2^k} + d_{L_2}(x, \mathcal{C})\big) + d_{L_2}(y, \mathcal{C}). \end{aligned}$$

which rearranges to give the desired lower bound on $\|H(x)-H(y)\|_{\ell_2^k}$ in the first inequality of (151). $\square$

Lemma 41. *Fix $K,D\geqslant 1$. Suppose that $(\mathcal{M},d_{\mathcal{M}})$ is a metric space that admits a K-localized radially weakly bi-Lipschitz embedding into L_2 with distortion D. If $\mathcal{C}\subseteq\mathcal{M}$ satisfies $2\leqslant|\mathcal{C}|<\infty$ and also*

$$\mathcal{C}\subseteq B_{\mathcal{M}}\Big(z,\Big(K-\frac{1}{8D}\Big)\Delta\Big) \tag{156}$$

for some $\Delta>0$ and $z\in\mathcal{M}$, then there is an integer $k=O(\log|\mathcal{C}|)$ a 1-Lipschitz function

$$\varphi: B_{\mathcal{M}}\Big(\mathcal{C},\frac{1}{8D}\Delta\Big)\to\ell_2^k$$

such that

$$\forall\, x\in B_{\mathcal{M}}\Big(\mathcal{C},\frac{1}{8D}\Delta\Big),\qquad \mathrm{rad}_{\mathcal{M}}\bigg(\varphi^{-1}\Big(B_{\ell_2^k}\Big(\varphi(x),\frac{1}{8D}\Delta\Big)\Big)\bigg)<\Delta. \tag{157}$$

Proof. Writing $r\stackrel{\mathrm{def}}{=}\Delta/(8D)$, we have $B_{\mathcal{M}}(\mathcal{C},r)\subseteq B_{\mathcal{M}}(z,K\Delta)$ thanks to (156) and the triangle inequality for $d_{\mathcal{M}}$. Hence, recalling Definition 33, the assumption of Lemma 41 implies that there is a 1-Lipschitz function $f:B_{\mathcal{M}}(\mathcal{C},r)\to L_2$ and $0<R<\Delta$ such that for any $x\in B_{\mathcal{M}}(\mathcal{C},r)$ there is $y_x\in\mathcal{M}$ satisfying:

$$f^{-1}\big(B_{L_2}(f(x),8r)\big)=f^{-1}\Big(B_{L_2}\Big(f(x),\frac{1}{D}\Delta\Big)\Big)\subseteq B_{\mathcal{M}}(y_x,R). \tag{158}$$

By applying Lemma 40 to $f(\mathcal{C})$ we get an integer $1\leqslant k\lesssim\log|\mathcal{C}|$ and a 1-Lipschitz function $H:L_2\to\ell_2^k$ that satisfies the following inclusion for any $v\in B_{L_2}(f(\mathcal{C}),r)$:

$$B_{L_2}(f(\mathcal{C}),r)\cap H^{-1}\big(B_{\ell_2^k}(H(v),r)\big)\subseteq B_{L_2}(v,8r).$$

Since f is 1-Lipschitz, $f(x)\in B_{L_2}(f(\mathcal{C}),r)$ for every $v\in B_{\mathcal{M}}(\mathcal{C},r)$, so the following holds a special case:

$$\forall\, x\in B_{\mathcal{M}}(\mathcal{C},r),\qquad B_{L_2}(f(\mathcal{C}),r)\cap H^{-1}\Big(B_{\ell_2^k}(\varphi(x),r)\Big)\subseteq B_{L_2}(f(x),8r), \tag{159}$$

where we define $\varphi\stackrel{\mathrm{def}}{=}H\circ f:B_{\mathcal{M}}(\mathcal{C},r)\to\ell_2^k$. Then, φ is 1-Lipschitz as both f and H are 1-Lipschitz. Observe furthermore that as f is 1-Lipschitz and its domain is $B_{\mathcal{M}}(\mathcal{C},r)$, all of its values belong to $B_{L_2}(f(\mathcal{C}),r)$. Therefore (159) can we rewritten as follows:

$$\forall\, x\in B_{\mathcal{M}}(\mathcal{C},r),\qquad f\big(B_{\mathcal{M}}(f(\mathcal{C}),r)\big)\cap H^{-1}\Big(B_{\ell_2^k}(\varphi(x),r)\Big)\subseteq B_{L_2}(f(x),8r), \tag{160}$$

By applying f^{-1} to both sides of (160) and then using (158), we conclude that

$$\forall\, x\in B_{\mathcal{M}}\Big(\mathcal{C},\frac{1}{8D}\Delta\Big)=B_{\mathcal{M}}(\mathcal{C},r),\qquad \varphi^{-1}\Big(B_{\ell_2^k}\Big(\varphi(x),\frac{1}{8D}\Delta\Big)\Big)=f^{-1}\Big(H^{-1}\Big(B_{\ell_2^k}\big(\varphi(x),r\big)\Big)\Big)\subseteq B_{\mathcal{M}}(y_x,R).$$

The $d_{\mathcal{M}}$-radius of $\varphi^{-1}(B_{\ell_2^k}(\varphi(x),\delta/(8D)))$ is therefore at most $R<\Delta$ for any $x\in B_{\mathcal{M}}(\mathcal{C},\Delta/(8D))$. In other words, the desired conclusion (157) of Lemma 41 indeed holds. $\square$

The following theorem implies Theorem 14 because if $\mathcal{M}=L_p$ for some $p>2$, then by Lemma 34 its assumptions hold with $D\asymp p$ and $K-1\asymp 1/p$.

Theorem 42. *Fix $K,D,\lambda>1$. Let $(\mathcal{M},d_{\mathcal{M}})$ be a Polish metric space admitting a K-localized radially weakly bi-Lipschitz embedding into L_2 with distortion D. Then, every λ-doubling Borel subset $\mathcal{D}$ of $\mathcal{M}$ satisfies:*

$$\forall\,\Delta>0,\qquad \widehat{\mathsf{SEP}}_\Delta\Big(B_{\mathcal{M}}\Big(\mathcal{D},\frac{1}{9D}\Delta\Big);\mathcal{M}\Big)\lesssim\frac{\sqrt{\log\lambda}}{K-1}\cdot\begin{cases}\frac{\sqrt{\log\frac{1}{K-1}}}{K-1} & \text{if } 1<K\leqslant 1+\frac{1}{D},\\ KD\sqrt{\log(KD)} & \text{if } K>1+\frac{1}{D}.\end{cases} \tag{161}$$

If furthermore $2\leqslant|\mathcal{D}|<\infty$, then

$$\forall\,\Delta>0,\qquad \widehat{\mathsf{SEP}}_\Delta\Big(B_{\mathcal{M}}\Big(\mathcal{D},\frac{1}{9D}\Delta\Big);\mathcal{M}\Big)\lesssim\frac{\sqrt{\log|\mathcal{D}|}}{K-1}\cdot\begin{cases}\frac{1}{K-1} & \text{if } 1<K\leqslant 1+\frac{1}{D},\\ KD & \text{if } K>1+\frac{1}{D}.\end{cases} \tag{162}$$

Note that if $|\mathcal{D}| = n < \infty$, then conclusion (163) of Theorem 42 is stronger than its conclusion (164) unless the doubling constant λ is very large; specifically, this occurs if and only if

$$\lambda = n^{o\left(\frac{1}{\log(eKD)}\right)}.$$

Proof of Theorem 42 . It suffices to prove Theorem 42 while imposing the further assumption $K \geqslant 1+1/D$ because if $1 < K < 1+1/D < 2$, then we can replace D by the larger quantity $1/(K-1)$. Thus, we will assume from now that $K \geqslant 1+1/D$ and our goal becomes to prove that

$$\widehat{\mathsf{SEP}}_\Delta\Big(B_{\mathcal{M}}(\mathcal{D}, \frac{1}{9D}\Delta); \mathcal{M}\Big) \lesssim \frac{KD\sqrt{\log(KD)}}{K-1}\sqrt{\log\lambda}, \tag{163}$$

and correspondingly if $2 \leqslant |\mathcal{D}| < \infty$, then

$$\widehat{\mathsf{SEP}}_\Delta\Big(B_{\mathcal{M}}(\mathcal{D}, \frac{1}{9D}\Delta); \mathcal{M}\Big) \lesssim \frac{KD}{K-1}\sqrt{\log|\mathcal{D}|}. \tag{164}$$

We will start by proving (163). Fix $\Delta > 0$. Our goal is to eventually apply Lemma 16 with $\mathcal{C}$ replaced by $B_{\mathcal{M}}(\mathcal{D}, \Delta/(9D))$, and with K replaced by K_*, where for convenience we set the following notation:

$$K_* \stackrel{\mathrm{def}}{=} K - \frac{1}{2D}. \tag{165}$$

Then, our assumption $K \geqslant 1+1/D$ ensures that

$$K_* - 1 \asymp K - 1. \tag{166}$$

In particular, $K_* > 1$, so this will be a valid instantiation of Lemma 16.

Fix $s \geqslant 0$ and define three auxiliary parameters $\alpha = \alpha(s,\Delta,K,D), \beta = \beta(s,\Delta,K,D), \varepsilon = \varepsilon(s,\Delta,K,D)$ by:

$$\alpha \stackrel{\mathrm{def}}{=} \Big(K - \frac{1}{8D}\Big)K_*^s\Delta \qquad \text{and} \qquad \beta \stackrel{\mathrm{def}}{=} \frac{1}{8D}K_*^s\Delta - \frac{1}{9D}\Delta \qquad \text{and} \qquad \varepsilon_0 \stackrel{\mathrm{def}}{=} \frac{\frac{3}{8}K_*^s - \frac{1}{9}}{D}\Delta. \tag{167}$$

Observe for later use that because $s \geqslant 0$, $K > K_* > 1$ and $D > 1$, we have:

$$\alpha \asymp KK_*^s\Delta \qquad \text{and} \qquad \alpha > \beta \asymp \frac{K_*^s}{D}\Delta \asymp \frac{\alpha}{KD} \qquad \text{and} \qquad \varepsilon_0 \asymp \frac{K_*^s}{D}\Delta. \tag{168}$$

Fix $z \in \mathcal{M}$. By a standard iteration of the assumed λ-doubling property of $\mathcal{D}$, there exists a finite subset $\mathcal{C}$ of $\mathcal{D} \cap B_{\mathcal{M}}(z,\alpha)$ which is β-dense in $\mathcal{D} \cap B_{\mathcal{M}}(z,\alpha)$, i.e., $B_{\mathcal{M}}(\mathcal{C},\beta) \supseteq \mathcal{D} \cap B_{\mathcal{M}}(z,\alpha)$, and whose size satisfies:

$$2 \leqslant |\mathcal{C}| \leqslant \lambda^{\left\lceil \log_2\left(\frac{2\alpha}{\beta}\right)\right\rceil}. \tag{169}$$

(Briefly, for completeness (see also [38, Lemma 4.1.12]): If $\ell \in \mathbb{N}$ satisfies $2^\ell \geqslant 2\alpha/\beta$, then iterate ℓ times the doubling condition for $\mathcal{D}$ to get the existence of a cover $\mathcal{D} \cap B_{\mathcal{M}}(z,\alpha)$ by λ^ℓ balls of radius $\alpha/2^\ell \leqslant \beta/2$, and then let $\mathcal{C}$ consist of one point from the intersection of each of these balls with $\mathcal{D} \cap B_{\mathcal{M}}(z,\alpha)$.) Observe that by the definition (167) of α, the aforementioned inclusion $\mathcal{C} \subseteq \mathcal{D} \cap B_{\mathcal{M}}(z,\alpha)$ can be rewritten as

$$\mathcal{C} \subseteq \mathcal{D} \cap B_{\mathcal{M}}\Big(z, \Big(K - \frac{1}{8D}\Big)K_*^s\Delta\Big). \tag{170}$$

Also, the parameters in (165) and (167) were chosen judiciously so that following inclusion holds:

$$B_{\mathcal{M}}\big(\mathcal{D}, \frac{1}{9D}\Delta\big) \cap B_{\mathcal{M}}(z, K_*^{s+1}\Delta + \varepsilon_0) \subseteq B_{\mathcal{M}}\big(\mathcal{C}, \frac{1}{8D}K_*^s\Delta\big). \tag{171}$$

Indeed, if $x \in \mathcal{M}$ satisfies $d_{\mathcal{M}}(x,y) \leqslant \Delta/(9D)$ for some $y \in \mathcal{D}$, and also $d_{\mathcal{M}}(x,z) \leqslant K_*^{s+1}\Delta + \varepsilon_0$, then

$$d_{\mathcal{M}}(y,z) \leqslant d_{\mathcal{M}}(x,y) + d_{\mathcal{M}}(x,z) \leqslant \frac{1}{9D}\Delta + K_*^{s+1}\Delta + \varepsilon_0 \stackrel{(165)\wedge(167)}{=} \alpha.$$

Therefore, $y \in \mathcal{D} \cap B_{\mathcal{M}}(z,\alpha)$. As $\mathcal{C}$ is β-dense in $\mathcal{D} \cap B_{\mathcal{M}}(z,\alpha)$, there exists $c \in \mathcal{C}$ with $d_{\mathcal{M}}(c,y) \leqslant \beta$, whence:

$$d_{\mathcal{M}}(x,c) \leqslant d_{\mathcal{M}}(x,y) + d_{\mathcal{M}}(y,c) \leqslant \frac{1}{9D}\Delta + \beta \stackrel{(167)}{=} \frac{1}{8D}K_*^s\Delta.$$

This implies that x belongs to the right hand side of (171), as required.

Thanks to (170), we may invoke Lemma 41 with Δ replaced by $K_*^s\Delta$ to get an integer k satisfying

$$k \asymp \log|\mathcal{C}| \overset{(169)}{\lesssim} (\log\lambda)\log\left(\frac{\alpha}{\beta}\right) \overset{(168)}{\lesssim} (\log\lambda)\log(KD), \tag{172}$$

and a 1-Lipschitz function $\varphi = \varphi_{s,\varepsilon} : B_{\mathcal{M}}\big(\mathcal{D}, \frac{1}{9D}\Delta\big) \cap B_{\mathcal{M}}(z, K_*^{s+1}\Delta + \varepsilon_0) \to \ell_2^k$, such that

$$\forall x \in B_{\mathcal{M}}\big(\mathcal{D}, \frac{1}{9D}\Delta\big) \cap B_{\mathcal{M}}(z, K_*^{s+1}\Delta + \varepsilon_0), \qquad \mathrm{rad}_{\mathcal{M}}\left(\varphi^{-1}\left(B_{\ell_2^k}\Big(\varphi(x), \frac{1}{8D}K_*^s\Delta\Big)\right)\right) < K_*^s\Delta. \tag{173}$$

Now, thanks to (173) we may apply Lemma 39 to $\mathcal{S} = B_{\mathcal{M}}(\mathcal{D}, \Delta/(9D)) \cap B_{\mathcal{M}}(z, K_*^{s+1}\Delta + \varepsilon_0) \subseteq \mathcal{M}$ and the target space $\mathcal{N} = \ell_2^k$, for which the assumption of Lemma 39 holds for $\sigma \lesssim \sqrt{K}$ by Theorem 38, with the parameters $L = 1$, $R = K_*^s\Delta/(8D)$ and Δ replaced by $K_*^s\Delta$, to get that:

$$\forall z \in \mathcal{M}, \qquad \widehat{\mathsf{SEP}}_{K_*^s\Delta}\Big(B_{\mathcal{M}}\big(\mathcal{D}, \frac{1}{9D}\Delta\big) \cap B_{\mathcal{M}}(z, K_*^{s+1}\Delta + \varepsilon_0); \mathcal{M}\Big) \lesssim D\sqrt{k} \overset{(172)}{\lesssim} D\sqrt{(\log\lambda)\log(KD)}.$$

Consequently,

$$\begin{aligned}\lim_{\varepsilon \to 0^+} \sup_{z \in \mathcal{M}} \widehat{\mathsf{SEP}}_{K_*^s\Delta}\Big(B_{\mathcal{M}}\big(\mathcal{D}, \frac{1}{9D}\Delta\big) \cap B_{\mathcal{M}}(z, K_*^{s+1}\Delta + \varepsilon); \mathcal{M}\Big) \\ \leqslant \sup_{z \in \mathcal{M}} \widehat{\mathsf{SEP}}_{K_*^s\Delta}\Big(B_{\mathcal{M}}\big(\mathcal{D}, \frac{1}{9D}\Delta\big) \cap B_{\mathcal{M}}(z, K_*^{s+1}\Delta + \varepsilon_0); \mathcal{M}\Big) \lesssim D\sqrt{(\log\lambda)\log(KD)}.\end{aligned} \tag{174}$$

In order to be able to use Lemma 16 in conjunction with (174), we must first check that its assumption (27) holds with $\mathcal{C}$ replaced by $B_{\mathcal{M}}(\mathcal{D}, \Delta/(9D))$, i.e., we need to verify that:

$$\lim_{T \to \infty} \frac{1}{T} \mathsf{SEP}_T\Big(B_{\mathcal{M}}\big(\mathcal{D}, \frac{1}{9D}\Delta\big)\Big) = 0. \tag{175}$$

The ensuing justification of (175) is suboptimal from the quantitative perspective. We chose this route as it is quick and for our purposes a qualitative statement (namely, without providing a rate of convergence) such as (175) suffices. See Remark 43 below for an improved (optimal) statement.

A well-known classical result (contained in [7]; see also [37, Chapter 12] and the discussion in [29], or e.g. the proof of [86, Theorem 5.2]) asserts that there are $k = k(\lambda) \in \mathbb{N}$ and $0 < \eta = \eta(\lambda) \leqslant 1$ such that for any $T > 0$ there exists a function $f_T : \mathcal{M} \to \ell_2^k$ which satisfies:

$$\|f_T\|_{\mathrm{Lip}(\mathcal{M};\ell_2^k)} \leqslant 1 \qquad \text{and} \qquad \forall a, b \in \mathcal{D}, \qquad d_{\mathcal{M}}(a,b) \geqslant \frac{1}{4}T \implies d_{\mathcal{M}}\big(f_T(a), f_T(b)\big) \geqslant \eta T. \tag{176}$$

We will next proceed to demonstrate that the following inclusion holds:

$$\forall T \geqslant \frac{\Delta}{2\eta D},\ \forall x \in B_{\mathcal{M}}\big(\mathcal{D}, \frac{1}{9D}\Delta\big), \qquad B_{\mathcal{M}}\big(\mathcal{D}, \frac{1}{9D}\Delta\big) \cap f_T^{-1}\Big(B_{\ell_2^k}\big(f_T(x), \frac{\eta}{2}T\big)\Big) \subseteq B_{\mathcal{M}}\big(x, \frac{1}{2}T\big). \tag{177}$$

After (177) will be established, we will proceed by fixing $T \geqslant \Delta/(2\eta D)$ and applying Lemma 39 with both $\mathcal{M}$ and $\mathcal{S}$ replaced by $B_{\mathcal{M}}(\mathcal{D}, \Delta/(9D))$, the function φ replaced by the restriction of f_T to $B_{\mathcal{M}}(\mathcal{D}, \Delta/(9D))$, the target space $\mathcal{N} = \ell_2^k$, for which the assumption of Lemma 39 holds for $\sigma \lesssim \sqrt{K}$ by Theorem 38, and with the parameters $L = 1 \geqslant \|f_T\|_{\mathrm{Lip}(\mathcal{M};\ell_2^k)}$, $R = \eta T/2$ and Δ replaced by $T/2$, to get that:

$$\forall T \geqslant \frac{\Delta}{\eta D}, \qquad \mathsf{SEP}_T\Big(B_{\mathcal{M}}\big(\mathcal{D}, \frac{1}{9D}\Delta\big)\Big) \overset{(26)}{\leqslant} 2\widehat{\mathsf{SEP}}_{\frac{1}{2}T}\Big(B_{\mathcal{M}}\big(\mathcal{D}, \frac{1}{9D}\Delta\big)\Big) \lesssim \frac{\sqrt{k}}{\eta} \lesssim_\lambda 1, \tag{178}$$

which handily implies (175).

Thus, it remains to prove (177), which we will next do by contradiction. The contrapositive of (177) is that there exist $T > 0$, as well as $x, y \in \mathcal{M}$ and $a, b \in \mathcal{D}$ that satisfy:

$$T \geqslant \frac{\Delta}{\eta D} \quad \text{and} \quad \|f_T(x) - f_T(y)\|_{\ell_2^k} \leqslant \frac{\eta T}{2} \quad \text{and} \quad d_{\mathcal{M}}(x,a), d_{\mathcal{M}}(y,b) \leqslant \frac{\Delta}{9D} \quad \text{and} \quad d_{\mathcal{M}}(x,y) > \frac{T}{2}. \tag{179}$$

It follows in particular from (179) that $d_{\mathcal{M}}(a,b)$ is sufficiently large so that the second part of (176) applies:

$$d_{\mathcal{M}}(a,b) \geqslant d_{\mathcal{M}}(x,y) - d_{\mathcal{M}}(x,a) - d_{\mathcal{M}}(y,b) \overset{(179)}{>} \frac{T}{2} - \frac{2\Delta}{9D} \overset{(179)}{\geqslant} \frac{T}{2} - \frac{2\eta T}{9} > \frac{T}{4}. \tag{180}$$

We may therefore use (176) to deduce the following contradictory chain of inequalities:

$$\begin{aligned}\frac{\eta T}{2} &\overset{(179)}{\geqslant} \|f_T(x)-f_T(y)\|_{\ell_2^k} \geqslant \|f_T(a)-f_T(b)\|_{\ell_2^k} - \|f_T(x)-f_T(a)\|_{\ell_2^k} - \|f_T(y)-f_T(b)\|_{\ell_2^k} \\ &\geqslant \|f_T(a)-f_T(b)\|_{\ell_2^k} - \|f_T\|_{\mathrm{Lip}(\mathcal{M};\ell_2^k)}\big(d_{\mathcal{M}}(x,a)+d_{\mathcal{M}}(x,b)\big) \overset{(180)\wedge(176)}{\geqslant} \eta T - \frac{2\Delta}{9D} \overset{(179)}{\geqslant} \eta T - \frac{4\eta T}{9}.\end{aligned}$$

Having checked the assumption (27) of Lemma 16 with $\mathcal{C}$ replaced by $B_{\mathcal{M}}(\mathcal{D},\Delta/(9D))$ is indeed satisfied, substituting (174) into the conclusion (28) of Lemma 16 and K replaced by K_* gives:

$$\widehat{\mathsf{SEP}}_\Delta\Big(B_{\mathcal{M}}\big(\mathcal{D},\frac{1}{9D}\Delta\big);\mathcal{M}\Big) \lesssim D\sqrt{(\log\lambda)\log(KD)}\sum_{s=0}^{\infty}\frac{1}{K_*^s} = \frac{K_*D\sqrt{\log(KD)}}{K_*-1}\log\lambda,$$

which implies the desired estimate (163) by the definition (165) of K_* and (166).

The justification of the finitary variant (164) is identical to the above reasoning, except that in this case one can simply work with $\mathcal{C}=\mathcal{D}$, and then replace the bound on k in (172) by $k \lesssim \log|\mathcal{D}|$. □

Remark 43. An optimal version of (178) is the following statement. There is a universal constant $C>0$ such that if $(\mathcal{M},d_{\mathcal{M}})$ is a metric space and $\mathcal{D}\subseteq\mathcal{M}$ is complete and λ-doubling for some $\lambda\geqslant 2$, then:

$$\forall\, r>0,\ \forall\, T\geqslant Cr,\qquad \mathsf{SEP}_T\big(B_{\mathcal{M}}(\mathcal{D},r)\big)\lesssim\log\lambda. \tag{181}$$

If $\mathcal{D}$ is compact, then (181) follows from [53, Theorem 3.17] applied to any $\lambda^{O(1)}$-doubling nondegenerate measure σ on $\mathcal{D}$; the existence of such a measure is due to [96], though the aforementioned dependence of its doubling constant on λ is not stated there but instead follows from an inspection of its proof (see also [37, Chapter 13]). If $\mathcal{D}$ is complete but not necessarily compact, then (181) is not proved in the literature but we will next indicate two ways to justify it. One such route is to use [53, Corollary 3.12] to obtain a $O(\log\lambda)$-padded partition of $\mathcal{D}$, then use [53, Lemma 3.8] to extend it to a $O(\log\lambda)$-padded random partition of $B_{\mathcal{M}}(\mathcal{D},r)$, and finally use [51, Theorem 2.2] to transform the extended random partition to a random partition of $B_{\mathcal{M}}(\mathcal{D},r)$ that is $O(\log\lambda)$-separating. However, [51] is an unpublished manuscript and one needs to justify the measurability that we need of each of the above steps (we checked that this is indeed the case, but it is somewhat tedious and does not appear in the literature). An alternative way to prove (181) is to follow the approach in the proof of [53, Theorem 3.17] when σ is now the $\lambda^{O(1)}$-doubling nondegenerate measure σ on $\mathcal{D}$ whose existence is proved in [60] (since $\mathcal{D}$ is complete). As this σ can have infinite mass (unlike the compact case), one needs to incorporate a (somewhat involved) reasoning that is a suitable adaptation of what was done in [80, Chapter 4]; an upshot of this route is that the measurability that we need is already worked out in [80]. Both of the above ways to demonstrate (181) result in digressions that are lengthier than how we proceeded above to establish (178), albeit suboptimally.

7. Deduction of Theorem 13 from [53]

Here will explain how to quickly derive Theorem 13 from one of the main results of [53].

Proof of Theorem 13. Fix $L>0$ and denote

$$\sigma \overset{\mathrm{def}}{=} \sup_{\Delta>0}\mathsf{SEP}_\Delta\Big(B_{\mathcal{M}}\big(\mathcal{C},\frac{1}{L}\Delta\big)\Big).$$

We may assume that $\sigma<\infty$ as otherwise (19) is vacuous. Fix $\Delta>0$. By the definition of σ, for any $\varepsilon>0$ there exists a probability space $(\Omega,\mathbb{P})=(\Omega_{\varepsilon,\Delta,L},\mathbb{P}_{\varepsilon,\Delta,L})$ and a random partition

$$\mathcal{P}\overset{\mathrm{def}}{=}\mathcal{P}_{\varepsilon,\Delta,L}=\Big\{\Gamma^i:\Omega\to 2^{B_{\mathcal{M}}(\mathcal{C},\frac{1}{L}\Delta)}\Big\}_{i=1}^{\infty}$$

of $B_{\mathcal{M}}(\mathcal{C},\Delta/L)$ that is Δ-bounded and $(\sigma+\varepsilon)$-separating. Because $\mathcal{C}$ is nonempty and locally compact, by [80, Lemma 115] for each $i\in\mathbb{N}$ there exists a $\mathbb{P}$-to-Borel measurable function $\gamma^i:\Omega\to\mathcal{C}$ that satisfies

$$\forall\omega\in\Omega,\qquad \Gamma^i(\omega)\neq\emptyset \implies d_{\mathcal{M}}\big(\gamma^i(\omega),\Gamma^i(\omega)\big)=d_{\mathcal{M}}\big(\mathcal{C},\Gamma^i(\omega)\big). \tag{182}$$

For each $x\in\mathcal{M}\smallsetminus B_{\mathcal{M}}(\mathcal{C},\Delta/L)$ fix any $c_x\in\mathcal{C}$ with $d_{\mathcal{M}}(x,c_x)<2d_{\mathcal{M}}(x,\mathcal{C})$. Denote by $\Gamma^x:\Omega\to 2^{\mathcal{M}}$ and $\gamma^x:\Omega\to\mathcal{C}$ the constant (set-valued and point-valued, respectively) functions that are given by setting $\Gamma^x(\omega)=\{x\}$ and $\gamma^x(\omega)=c_x$ for every $\omega\in\Omega$. Then, the following forms a stochastic decomposition[13] of $\mathcal{M}$ with respect to $\mathcal{C}$ in the sense of [53, Definition 3.1]:

$$\Big(\Omega,\mathbb{P},\{\Gamma^i(\cdot),\gamma^i(\cdot)\}_{i=1}^\infty\cup\{\Gamma^x(\cdot),\gamma^x(\cdot)\}_{x\in\mathcal{M}\smallsetminus B_{\mathcal{M}}(\mathcal{C},\frac{1}{L}\Delta)}\Big). \tag{183}$$

By combining [53, Lemma 2.1] and [53, Theorem 4.1, part 3.] it suffices to check that the stochastic decomposition (183) is Δ-bounded and $(1/(2L),1/\max\{2L,\sigma+\varepsilon\})$-separating with respect to $\mathcal{C}$, in the sense of [53, Definition 3.2] and [53, Definition 3.7], respectively. The former requirement is immediate as (183) adds singleton clusters to the random partition $\mathcal{P}$. The latter requirements mean that if we write

$$\forall\omega\in\Omega,\qquad \mathcal{Q}^\omega\stackrel{\mathrm{def}}{=}\mathcal{P}^\omega\cup\big\{\{x\}\big\}_{x\in\mathcal{M}\smallsetminus B_{\mathcal{M}}(\mathcal{C},\frac{1}{L}\Delta)}, \tag{184}$$

then

$$\forall x,y\in\mathcal{M},\qquad d_{\mathcal{M}}(\{x,y\},\mathcal{C})\leqslant\frac{1}{2L}\Delta\implies\mathbb{P}\big[\mathcal{Q}^\omega(x)\neq\mathcal{Q}^\omega(y)\big]\leqslant\frac{\max\{2L,\sigma+\varepsilon\}}{\Delta}d_{\mathcal{M}}(x,y). \tag{185}$$

To verify (185), fix $x,y\in\mathcal{M}$ with $d_{\mathcal{M}}(\{x,y\},\mathcal{C})\leqslant\Delta/(2L)$. Assume that also $d_{\mathcal{M}}(x,y)<\Delta/(2L)$, as otherwise the right hand side of (185) is at least 1. So, $\max\{d_{\mathcal{M}}(x,\mathcal{C}),d_{\mathcal{M}}(y,\mathcal{C})\}\leqslant d_{\mathcal{M}}(\{x,y\},\mathcal{C})+d_{\mathcal{M}}(x,y)\leqslant\Delta/L$, i.e., $x,y\in B_{\mathcal{M}}(\mathcal{C},\Delta/L)$, whence (185) follows from the definition (184) of $\mathcal{Q}$ and the fact that $\mathcal{P}$ is assumed to be a $(\sigma+\varepsilon)$-separating random partition of $B_{\mathcal{M}}(\mathcal{C},\Delta/L)$. □

Remark 44. We assumed in Theorem 13 that $\mathcal{C}$ is locally compact only for invoking [80, Lemma 115] to get measurable functions $\{\gamma^i:\Omega\to\mathcal{C}\}_{i=1}^\infty$ that satisfy (182) (an inspection of the proof of [80, Lemma 115] reveals that all that is required for this is that $\mathcal{C}$ is σ-compact). One can alternatively stipulate the existence of such functions (even with the weaker requirement $d_{\mathcal{M}}(\gamma^i(\omega),\Gamma^i(\omega))<2d_{\mathcal{M}}(\mathcal{C},\Gamma^i(\omega))$ in (182)) as part of the assumption of Theorem 13. This would be analogous to the route that was pursued in [53], which axiomatizes the minimal assumptions that are needed for its proofs. However, it is hard to maintain such an assumption under various natural operations, including those that are applied herein, such as preimages under Lipschitz functions and intersections. The foundational reworking of [80] circumvents such issues by imposing stronger measurability requirements that are readily seen to be perserved under natural geometric operations; adopting this approach facilitates the constructions herein.

References

[1] I. Abraham, Y. Bartal, and O. Neiman. Local embeddings of metric spaces. *Algorithmica*, 72(2):539–606, 2015.

[2] A. B. Aleksandrov and V. V. Peller. Functions of operators under perturbations of class $\mathbf{S}_p$. *J. Funct. Anal.*, 258(11):3675–3724, 2010.

[3] A. Andoni, A. Naor, and O. Neiman. On isomorphic dimension reduction in ℓ_1. Preprint, 2025.

[4] S. Arora, J. R. Lee, and A. Naor. Fréchet embeddings of negative type metrics. *Discrete Comput. Geom.*, 38(4):726–739, 2007.

[5] S. Arora, J. R. Lee, and A. Naor. Euclidean distortion and the sparsest cut. *J. Amer. Math. Soc.*, 21(1):1–21, 2008.

[13] Formally, Definition 3.1 of [53] also requires that $\Gamma^i(\omega)$ is a Borel subset of $\mathcal{M}$ for every $i\in\mathbb{N}$ and every $\omega\in\Omega$. However, this assumption is never used in [53]. To see this, note that such measurability is not required in the definition of gentle partition of unity in [53, Section 2]. The gentle partition of unity that is constructed in [53] from the stochastic decomposition is given in [53, equation (6)]. For it to be obey the measurability requirements of [53, Section 2] one needs that for each fixed $x\in\mathcal{M}\smallsetminus\mathcal{C}$ the function $(\omega\in\Omega)\mapsto\mathbf{1}_{\Gamma^i(\omega)}(x)\in\{0,1\}$ is $\mathbb{P}$-measurable, which is equivalent to the set $\{\omega\in\Omega:\ x\in\Gamma^i(\omega)\}$ being $\mathbb{P}$-measurable, and this is a special case of the strong measurability which is imposed by the definition of random partition that is used in the present work (following [80]). That said, all of the random partitions that we use herein are into Borel subsets of the given metric space, as is evident by their constructions (including those that are cited, specifically from [80], for Theorem 38).

[6] S. Arora, S. Rao, and U. Vazirani. Expander flows, geometric embeddings and graph partitioning. *J. ACM*, 56(2):Art. 5, 37, 2009.
[7] P. Assouad. Plongements lipschitziens dans $\mathbf{R}^n$. *Bull. Soc. Math. France*, 111(4):429–448, 1983.
[8] Y. Aumann and Y. Rabani. An $O(\log k)$ approximate min-cut max-flow theorem and approximation algorithm. *SIAM J. Comput.*, 27(1):291–301, 1998.
[9] K. Ball. Markov chains, Riesz transforms and Lipschitz maps. *Geom. Funct. Anal.*, 2(2):137–172, 1992.
[10] K. Ball. The Ribe programme. Number 352, pages Exp. No. 1047, viii, 147–159. 2013. Séminaire Bourbaki. Vol. 2011/2012. Exposés 1043–1058.
[11] K. Ball, E. A. Carlen, and E. H. Lieb. Sharp uniform convexity and smoothness inequalities for trace norms. *Invent. Math.*, 115(3):463–482, 1994.
[12] Y. Bartal. Probabilistic approximation of metric spaces and its algorithmic applications. In *Proceedings of 37th Conference on Foundations of Computer Science*, pages 184–193. IEEE, 1996.
[13] Y. Bartal and L.-A. Gottlieb. Dimension reduction techniques for ℓ_p, $1 \leqslant p < \infty$, with applications. Preprint available at `https://arxiv.org/pdf/1408.1789v2` (observe that this is version v2), 2014.
[14] B. Begun. A remark on almost extensions of Lipschitz functions. *Israel J. Math.*, 109:151–155, 1999.
[15] 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.
[16] J. Bourgain. On Lipschitz embedding of finite metric spaces in Hilbert space. *Israel Journal of Mathematics*, 52:46–52, 1985.
[17] J. Bourgain. The metrical interpretation of superreflexivity in Banach spaces. *Israel J. Math.*, 56(2):222–230, 1986.
[18] J. Bourgain. Remarks on the extension of Lipschitz maps defined on discrete sets and uniform homeomorphisms. In *Geometrical aspects of functional analysis (1985/86)*, volume 1267 of *Lecture Notes in Math.*, pages 157–167. Springer, Berlin, 1987.
[19] M. Braverman and A. Naor. Quantitative Wasserstein rounding. Preprint, 2025.
[20] E. A. Carlen and E. H. Lieb. Optimal hypercontractivity for Fermi fields and related noncommutative integration inequalities. *Comm. Math. Phys.*, 155(1):27–46, 1993.
[21] F. Chaatit. On uniform homeomorphisms of the unit spheres of certain Banach lattices. *Pacific J. Math.*, 168(1):11–31, 1995.
[22] A. Chang, A. Naor, and K. Ren. Random zero sets with local growth guarantees. Preprint available at `https://arxiv.org/pdf/2410.21931`, 2024.
[23] M. Charikar, C. Chekuri, A. Goel, S. Guha, and S. Plotkin. Approximating a finite metric by a small number of tree metrics. In *Proceedings 39th Annual Symposium on Foundations of Computer Science (Cat. No.98CB36280)*, pages 379–388, 1998.
[24] M. Charikar, T. Leighton, S. Li, and A. Moitra. Vertex sparsifiers and abstract rounding algorithms. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science—FOCS 2010*, pages 265–274. IEEE Computer Soc., Los Alamitos, CA, 2010.
[25] M. Charikar, K. Makarychev, and Y. Makarychev. Local global tradeoffs in metric embeddings. *SIAM J. Comput.*, 39(6):2487–2512, 2010.
[26] W. Corrêa. Uniform homeomorphisms between spheres induced by interpolation methods. *Proc. Amer. Math. Soc.*, 152(5):2157–2167, 2024.
[27] W. Corrêa, V. Ferenczi, R. Gesing, and P. Tradacete. Extremes of interpolation scales of Banach spaces. *J. Funct. Anal.*, 289(2):Paper No. 110924, 41, 2025.
[28] M. Daher. Homéomorphismes uniformes entre les sphères unité des espaces d'interpolation. *Canad. Math. Bull.*, 38(3):286–294, 1995.
[29] J. Ding, J. R. Lee, and Y. Peres. Markov type and threshold embeddings. *Geom. Funct. Anal.*, 23(4):1207–1229, 2013.
[30] A. Dvoretzky. Some results on convex bodies and Banach spaces. In *Proc. Internat. Sympos. Linear Spaces (Jerusalem, 1960)*, pages 123–160. Jerusalem Academic Press, Jerusalem, 1961.
[31] P. Enflo. On the nonexistence of uniform homeomorphisms between L_p-spaces. *Ark. Mat.*, 8:103–105, 1969.
[32] M. Englert, A. Gupta, R. Krauthgamer, H. Räcke, I. Talgam-Cohen, and K. Talwar. Vertex sparsifiers: new results from old techniques. *SIAM J. Comput.*, 43(4):1239–1262, 2014.
[33] A. Eskenazis, M. Mendel, and A. Naor. Nonpositive curvature is not coarsely universal. *Invent. Math.*, 217(3):833–886, 2019.
[34] G. Godefroy. De Grothendieck à Naor: une promenade dans l'analyse métrique des espaces de Banach. *Gaz. Math.*, (151):13–24, 2017.
[35] B. Grünbaum. Some applications of expansion constants. *Pacific J. Math.*, 10:193–201, 1960.
[36] A. Gupta, R. Krauthgamer, and J. R. Lee. Bounded geometries, fractals, and low-distortion embeddings. In *44th Symposium on Foundations of Computer Science (FOCS 2003), 11-14 October 2003, Cambridge, MA, USA, Proceedings*, pages 534–543. IEEE Computer Society, 2003.
[37] J. Heinonen. *Lectures on analysis on metric spaces*. Universitext. Springer-Verlag, New York, 2001.
[38] J. Heinonen, P. Koskela, N. Shanmugalingam, and J. T. Tyson. *Sobolev spaces on metric measure spaces*, volume 27 of *New Mathematical Monographs*. Cambridge University Press, Cambridge, 2015. An approach based on upper gradients.
[39] C. J. Himmelberg. Measurable relations. *Fund. Math.*, 87:53–72, 1975.

[40] J. Huang, F. Sukochev, and D. Zanin. Operator θ-Hölder functions with respect to $\|\cdot\|_p$, $0 < p \leqslant \infty$. *J. Lond. Math. Soc. (2)*, 105(4):2436–2477, 2022.
[41] W. B. Johnson and J. Lindenstrauss. Extensions of Lipschitz mappings into a Hilbert space. In *Conference in modern analysis and probability (New Haven, Conn., 1982)*, volume 26 of *Contemp. Math.*, pages 189–206. Amer. Math. Soc., Providence, RI, 1984.
[42] W. B. Johnson, J. Lindenstrauss, and G. Schechtman. Extensions of Lipschitz maps into Banach spaces. *Israel J. Math.*, 54(2):129–138, 1986.
[43] W. B. Johnson and G. Schechtman. Finite dimensional subspaces of L_p. In *Handbook of the geometry of Banach spaces, Vol. I*, pages 837–870. North-Holland, Amsterdam, 2001.
[44] N. J. Kalton. The nonlinear geometry of Banach spaces. *Rev. Mat. Complut.*, 21(1):7–60, 2008.
[45] A. S. Kechris. *Classical descriptive set theory*, volume 156 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995.
[46] M. Kirszbraun. Über die zusammenziehende und lipschitzsche transformationen. *Fundamenta Mathematicae*, 22(1):77–108, 1934.
[47] R. Krauthgamer, J. R. Lee, M. Mendel, and A. Naor. Measured descent: a new embedding method for finite metrics. *Geom. Funct. Anal.*, 15(4):839–858, 2005.
[48] R. Krauthgamer and N. Petruschka. Lipschitz decompositions of finite ℓ_p metrics. *Preprint available at* `https://arxiv.org/pdf/2502.01120`, 2025.
[49] R. Krauthgamer, N. Petruschka, and S. Sapir. The Power of Recursive Embeddings for ℓ_p Metrics. Preprint available at `https://arxiv.org/pdf/2503.18508`, 2025.
[50] J. R. Lee. On distance scales, embeddings, and efficient relaxations of the cut cone. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 92–101. ACM, New York, 2005.
[51] J. R. Lee and A. Naor. Metric decomposition, smooth measures, and clustering. Unpublished manuscript, available on request, 2003.
[52] J. R. Lee and A. Naor. Absolute Lipschitz extendability. *C. R. Math. Acad. Sci. Paris*, 338(11):859–862, 2004.
[53] J. R. Lee and A. Naor. Extending Lipschitz functions via random metric partitions. *Invent. Math.*, 160(1):59–95, 2005.
[54] D. R. Lewis. Finite dimensional subspaces of L_p. *Studia Math.*, 63(2):207–212, 1978.
[55] J. Lindenstrauss. On nonlinear projections in Banach spaces. *Michigan Math. J.*, 11:263–287, 1964.
[56] J. Lindenstrauss and L. Tzafriri. *Classical Banach spaces. I*, volume Band 92 of *Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas]*. Springer-Verlag, Berlin-New York, 1977. Sequence spaces.
[57] J. Lindenstrauss and L. Tzafriri. *Classical Banach spaces. II*, volume 97 of *Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas]*. Springer-Verlag, Berlin-New York, 1979. Function spaces.
[58] N. Linial, E. London, and Y. Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15:215–245, 1995.
[59] G. G. Lorentz. Who discovered analytic sets? *Math. Intelligencer*, 23(4):28–32, 2001.
[60] J. Luukkainen and E. Saksman. Every complete doubling metric space carries a doubling measure. *Proc. Amer. Math. Soc.*, 126(2):531–534, 1998.
[61] N. N. Luzin. Sur la classification de M. Baire. *C. R. Math. Acad. Sci. Paris*, 164:91–94, 1917.
[62] K. Makarychev and Y. Makarychev. Metric extension operators, vertex sparsifiers and Lipschitz extendability. *Israel J. Math.*, 212(2):913–959, 2016.
[63] M. B. Marcus and G. Pisier. Characterizations of almost surely continuous p-stable random Fourier series and strongly stationary processes. *Acta Math.*, 152(3-4):245–301, 1984.
[64] J. Matoušek. Extension of Lipschitz mappings on metric trees. *Comment. Math. Univ. Carolin.*, 31(1):99–104, 1990.
[65] J. Matoušek. On embedding expanders into l_p spaces. *Israel J. Math.*, 102:189–197, 1997.
[66] 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.
[67] S. Mazur. Une remarque sur l'homéomorphie des champs fonctionnels. *Studia Mathematica*, 1(1):83–85, 1929.
[68] M. Mendel and A. Naor. Euclidean quotients of finite metric spaces. *Advances in Mathematics*, 189(2):451–494, 2004.
[69] M. Mendel and A. Naor. Some applications of Ball's extension theorem. *Proc. Amer. Math. Soc.*, 134(9):2577–2584, 2006.
[70] M. Mendel and A. Naor. Metric cotype. *Ann. of Math. (2)*, 168(1):247–298, 2008.
[71] M. Mendel and A. Naor. Spectral calculus and Lipschitz extension for barycentric metric spaces. *Anal. Geom. Metr. Spaces*, 1:163–199, 2013.
[72] A. Moitra. Approximation algorithms for multicommodity-type problems with guarantees independent of the graph size. In *2009 50th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2009*, pages 3–12. IEEE Computer Soc., Los Alamitos, CA, 2009.
[73] 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, 2001.
[74] A. Naor. An introduction to the Ribe program. *Japanese Journal of Mathematics*, 7(2):167–233, 2012.

[75] A. Naor. Comparison of metric spectral gaps. *Analysis and Geometry in Metric Spaces*, 2(1):20141001, 2014.

[76] A. Naor. Uniform nonextendability from nets. *C. R. Math. Acad. Sci. Paris*, 353(11):991–994, 2015.

[77] A. Naor. Probabilistic clustering of high dimensional norms. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 690–709. SIAM, 2017.

[78] A. Naor. Metric dimension reduction: a snapshot of the Ribe program. In *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. I. Plenary lectures*, pages 759–837. World Sci. Publ., Hackensack, NJ, 2018.

[79] A. Naor. Impossibility of almost extension. *Adv. Math.*, 384:Paper No. 107761, 34, 2021.

[80] A. Naor. *Extension, separation and isomorphic reverse isoperimetry*, volume 11 of *Memoirs of the European Mathematical Society*. European Mathematical Society (EMS), Berlin, 2024.

[81] A. Naor, Y. Peres, O. Schramm, and S. Sheffield. Markov chains in smooth Banach spaces and Gromov-hyperbolic metric spaces. *Duke Math. J.*, 134(1):165–197, 2006.

[82] A. Naor and Y. Rabani. On Lipschitz extension from finite subsets. *Israel J. Math.*, 219(1):115–161, 2017.

[83] A. Naor and K. Ren. ℓ_p has nontrivial Euclidean distortion growth when $2 < p < 4$. Preprint available at `https://arxiv.org/pdf/2410.21931v1` (observe that this is version v1), 2025.

[84] A. Naor and G. Schechtman. Metric X_p inequalities. *Forum Math. Pi*, 4:e3, 81, 2016.

[85] A. Naor and G. Schechtman. Lipschitz almost-extension and nonexistence of uniform embeddings of balls in Schatten classes. Preprint, 2025.

[86] A. Naor and L. Silberman. Poincaré inequalities, embeddings, and wild groups. *Compos. Math.*, 147(5):1546–1572, 2011.

[87] I. Newman and Y. Rabinovich. A lower bound on the distortion of embedding planar metrics into Euclidean space. *Discrete Comput. Geom.*, 29(1):77–81, 2003.

[88] E. Odell and T. Schlumprecht. The distortion problem. *Acta Math.*, 173(2):259–281, 1994.

[89] M. Ostrovskii. Metric characterizations of some classes of Banach spaces. In *Harmonic analysis, partial differential equations, complex analysis, Banach spaces, and operator theory. Vol. 1*, volume 4 of *Assoc. Women Math. Ser.*, pages 307–347. Springer, [Cham], 2016.

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

[91] S. Rao. Small distortion and volume preserving embeddings for planar and Euclidean metrics. In *Proceedings of the Fifteenth Annual Symposium on Computational Geometry (Miami Beach, FL, 1999)*, pages 300–306. ACM, New York, 1999.

[92] Y. Raynaud. On ultrapowers of non commutative L_p spaces. *J. Operator Theory*, 48(1):41–68, 2002.

[93] E. Ricard. Hölder estimates for the noncommutative Mazur maps. *Arch. Math. (Basel)*, 104(1):37–45, 2015.

[94] A. Sobczyk. Projections in Minkowski and Banach spaces. *Duke Math. J.*, 8:78–106, 1941.

[95] N. Tomczak-Jaegermann. *Banach-Mazur distances and finite-dimensional operator ideals*, volume 38 of *Pitman Monographs and Surveys in Pure and Applied Mathematics*. Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc., New York, 1989.

[96] A. L. Vol'berg and S. V. Konyagin. A homogeneous measure exists on any compactum in $\mathbf{R}^n$. *Dokl. Akad. Nauk SSSR*, 278(4):783–786, 1984.

[97] P. Wojtaszczyk. *Banach spaces for analysts*, volume 25 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1991.

DEPARTMENT OF MATHEMATICS, PRINCETON NJ 08544-1000
Email address: `naor@math.princeton.edu`

DEPARTMENT OF MATHEMATICS, PRINCETON NJ 08544-1000
Email address: `kr5621@princeton.edu`