Extending Lipschitz functions via random metric partitions

James R. Lee[*]
U.C. Berkeley
jrl@cs.berkeley.edu

Assaf Naor
Microsoft Research
anaor@microsoft.com

1 Introduction

Many classical problems in geometry and analysis involve the gluing together of local information to produce a coherent global picture. Inevitably, the difficulty of such a procedure lies at the local boundary, where overlapping views of the same locality must somehow be merged. It is therefore desirable that the boundaries be "smooth," allowing a graceful transition from one viewpoint to the next. For instance, one may point to Whitney's use of partitions of unity in studying what is now known as the Whitney extension problem [36, 37].

In the present work, we consider what is perhaps the most basic Whitney-type extension problem, that of extending a Lipschitz function so that it remains Lipschitz. Often such a map is extended by first producing a cover of the new domain, extending the mapping locally, and then gluing together the individual pieces. Our main observation is that in many cases, if one chooses a *random* cover from the right distribution, the boundary can be made "smooth" on average, even when the local maps are individually quite coarse. This insight leads to the unification, generalization, and improvement of many known results, as well as to new results for many interesting spaces.

1.1 The Lipschitz extension problem

Let (Y, d_Y), (Z, d_Z) be metric spaces, and for every $X \subseteq Y$, denote by $e(X, Y, Z)$ the infimum over all constants K such that every Lipschitz function $f : X \to Z$ can be extended to a function $\tilde{f} : Y \to Z$ satisfying $\|\tilde{f}\|_{\mathrm{Lip}} \le K\|f\|_{\mathrm{Lip}}$. (If no such K exists, we set $e(X, Y, Z) = \infty$). We also define $e(Y, Z) = \sup\{e(X, Y, Z) : \ X \subseteq Y\}$ and for every integer n, $e_n(Y, Z) = \sup\{e(X, Y, Z) : \ X \subseteq Y, \ |X| \le n\}$.

Estimating $e(Y, Z)$ is a classical and fundamental problem that has attracted a lot of attention due to its intrinsic interest and applications to geometry and approximation theory. It is a classical fact that for every metric space Y, $e(Y, \ell_\infty) = 1$, and Kirszbraun's famous extension theorem [19] states that whenever H_1 and H_2 are Hilbert spaces, $e(H_1, H_2) = 1$. We refer to the books [3, 35] for a detailed account of the case $e(Y, Z) = 1$.

Typically, proofs of the fact that $e(Y, Z) = 1$ involve showing that it is possible to extend an arbitrary Lipschitz function to an additional point while preserving the Lipschitz constant. Once this is achieved, the existence of the required extension follows from Zorn's lemma. When $e(Y, Z) > 1$, this "one point at a time" argument cannot work, since the Lipschitz constant will deteriorate with each iteration. Hence, in proving "non-isometric" extension results, one

[*]Work partially supported by NSF grant CCR-0121555 and an NSF Graduate Research Fellowship. The majority of this work was completed while the author was an intern at Microsoft Research, Redmond, WA.

must argue that it is possible to extend Lipschitz functions to arbitrarily many points at once, and it is therefore not surprising that such results are more recent.

The following theorem was proved by Marcus and Pisier in [27], via a probabilistic argument.

Theorem 1.1 (Marcus-Pisier [27]). *For every $1 < p < 2$ there exists a constant $C(p)$ such that for every integer n,*

$$e_n(L_p, L_2) \le C(p)\,(\log n)^{1/p-1/2}.$$

The Marcus-Pisier theorem initiated the study of the parameter $e_n(\cdot,\cdot)$, and using a different probabilistic argument, Johnson and Lindenstrauss [17] subsequently proved the following theorem.

Theorem 1.2 (Johnson-Lindenstrauss [17]). *For every metric space Y and every integer n,*

$$e_n(Y, L_2) \le 2\sqrt{\log n}.$$

In [17], it was also shown that Theorem 1.1 and Theorem 1.2 are almost optimal, in the sense that, for every $1 \le p < 2$, $e_n(L_p, L_2) \ge C(p)\left(\frac{\log n}{\log\log n}\right)^{1/p-1/2}$.

In [18] Johnson, Lindenstrauss and Schechtman studied the case when the target space is an arbitrary Banach space, proving the following two theorems.

Theorem 1.3 (Johnson-Lindenstrauss-Schechtman [18]). *There exists a universal constant C such that for every metric space Y, every Banach space Z and every integer n,*

$$e_n(Y, Z) \le C \log n.$$

Theorem 1.4 (Johnson-Lindenstrauss-Schechtman [18]). *There exists a universal constant C such that for every d-dimensional normed space Y and every Banach space Z, $e(Y, Z) \le Cd$.*

Additionally, Matoušek has shown in [28] that

Theorem 1.5 (Matoušek [28]). *There exists a universal constant C such that for every metric tree T and every Banach space Z, $e(T, Z) \le C$.*

In the important paper [1] (which introduced, among other things, the notion of Markov type), K. Ball has shown that for every $1 < p < 2$, $e(L_2, L_p) \le \frac{6}{\sqrt{p-1}}$. More recently Tsar′kov [34] proved that for every $2 < p < \infty$, $e(L_p, L_2) \le C(p) < \infty$ and the second named author proved in [30] that in the same range of p, $e(L_2, L_p) = \infty$. A quantitative version of the last result is discussed in the next section.

The extension problem when the target space Z isn't linear (e.g. Hadamard manifolds) has also received a lot of attention. We deal with this problem in Section 1.3 below.

1.2 Absolute Lipschitz extendability

We return to these old problems with the new perspective made possible by recent advances in combinatorics and theoretical computer science. Often in theoretical computer science, one needs to analyze data with an inherent metric structure (e.g. graphs). A technique developed over the past decade to handle such problems can be referred to as the method of *stochastic metric decomposition*. The basic idea is that given a metric space X, one constructs a distribution over partitions of X with certain desirable properties. For example, one often requires that each set in the partition has small diameter and yet, in expectation, that every point is "far from the boundary" of the partition. Variants of this approach have appeared in numerous contexts; see for instance [26, 20, 2].

We offer a new approach to extension problems by showing that one can pass from a stochastic decomposition to a "well-behaved" partition of unity which, in turn, can be used to extend Lipschitz functions in such a way that we have control on the growth of the Lipschitz constant. This allows us to obtain simple proofs of many of the extension theorems stated above, and more importantly, to obtain new extension theorems which, in some cases, are significant generalizations of the above results.

Additionally, we observe a new phenomenon underlying some of the previous results which we refer to as *absolute extendability*—the notion that for some spaces X, Lipschitz functions f from X into *any* Banach space Z can be extended to *any* containing space $Y \supseteq X$, where the loss in the Lipschitz constant is independent of Y, Z, and f, and thus depends only on X. To this end, let us define, for a metric space X, the *absolute extendability constant* $ae(X)$ by

$$ae(X) = \sup\{e(X, Y, Z) : Y \supseteq X, Z \text{ a Banach space}\}.$$

If $ae(X) < \infty$, we say that X is *absolutely extendable*. Additionally, for a family of metric spaces $\mathcal{M}$, let us define $ae(\mathcal{M}) = \sup_{X \in \mathcal{M}} ae(X)$ to be a uniform bound on the extendability of metrics in $\mathcal{M}$. As far as we are aware, the only previously known families of absolutely extendable metrics had such a property for a "trivial" reason; these are the cases when X is an absolute Lipschitz retract or when the family $\mathcal{M}$ consists of finite metrics of uniformly bounded cardinality (it is not to difficult to see that Theorem 1.3 is true when $\log n$ is replaced by n). We now turn to our results, some of which have been announced in [24].

Recall that the *doubling constant* (see, e.g. [15]) of a metric space X, denoted $\lambda(X)$, is the infimum over all numbers λ such that every ball in X can be covered by λ balls of half the radius. When $\lambda(X) < \infty$, one says that X is *doubling*. Applying our approach to the stochastic decomposition of [14] yields the following result.

Theorem 1.6. *There exists a universal constant $C > 0$ such that*

$$ae(X) \leq C \log \lambda(X).$$

Since $\log \lambda(X) = O(\log n)$ for any n-point metric space X, and $\log \lambda(X) = O(d)$ whenever X is a subset of some d-dimensional normed space, Theorem 1.6 unifies and generalizes Theorems 1.3 and 1.4. Moreover, it is interesting to note that, unlike Theorem 1.4, Theorem 1.6 only assumes a bound on the dimension of some normed space containing X, while Y is allowed to be arbitrary. The proof of Theorem 1.4 in [18] cannot be used to obtain such a result, since their proof uses nets in the *complement* of X. Techniques used in the proof of the above theorem also allow one to extend Lipschitz functions to neighborhoods of subsets of manifolds of negative curvature using an estimate of Bishop (see Corollary 3.13 in Section 3.2).

Our next theorem provides a significant generalization of Theorem 1.5. Let $G = (V, E)$ be a countable graph with edge weights in $[0, \infty]$. Denote by $\Sigma(G)$ the one-dimensional simplicial complex that arises from G by replacing every edge e of G by an interval whose length is equal to that of e. Note that $\Sigma(G)$ has a natural Riemannian metric structure and that the shortest path metric on G occurs as a submetric. For a family of metric spaces $\mathcal{M}$, define $\overline{\mathcal{M}} = \{(X, d) : X \subseteq Y \text{ for some } (Y, d) \in \mathcal{M}\}$, i.e. the closure of $\mathcal{M}$ under taking submetrics.

We now define the set of metrics supported on G by

$$\langle G\rangle = \overline{\bigcup_{\text{all weights}} \Sigma(G)}$$

where the union is taken over all possible weights on edges of G. Finally, for a family of graphs $\mathcal{F}$, let $\langle\mathcal{F}\rangle = \bigcup_{G\in\mathcal{F}}\langle G\rangle$. Given such a family of graphs $\mathcal{F}$, one can ask when there exists a constant $K_{\mathcal{F}}$ such that, $ae(\langle\mathcal{F}\rangle) \le K_{\mathcal{F}}$. For instance, note that if $\mathcal{F}$ is the family of graph-theoretic trees, then $\langle\mathcal{F}\rangle$ is precisely the class of metrics which are a submetric of a separable metric tree, and such a result would strengthen Theorem 1.5 in the separable case. We will see momentarily that such a result does follow and thus implies that a Lipschitz function on any subset of a separable metric tree can be extended to *any superspace* with a universally bounded loss in the Lipschitz constant, while Matoušek's proof relies heavily on the fact that the superspace is a tree.

To state the next result, we need to recall the definition of a graph minor. Namely, consider the following two operations on a graph $G = (V, E)$.

1. Removal of an edge $e \in E$, i.e. moving to the graph $G' = (V, E \setminus \{e\})$.
2. Contraction of an edge $\{u, v\} \in E$, yielding the graph $G' = (V', E')$, where $V' = V \setminus \{u, v\} \cup \{v^*\}$ and $\{s, t\} \in E'$ precisely when either (i) $\{s, t\} \in E$ for $v^* \notin \{s, t\}$ or when (ii) $s = v^*, t \neq v^*$, and either $\{u, t\} \in E$ or $\{v, t\} \in E$.

When a graph G' is obtainable from G by a finite sequence of such operations, G' is called a *minor* of G. If a family of graphs $\mathcal{F}$ has the property that $G' \in \mathcal{F}$ whenever $G \in \mathcal{F}$ and G' is a minor of G, we say that $\mathcal{F}$ is *minor-closed*. Finally, we say that a graph G excludes K_r if it does not contain the complete graph on r vertices as a minor. Using the decomposition of [20], along with its stochastic formulation in [31] and a quantitative improvement due to [10] yields

Theorem 1.7. *If a graph G excludes K_r, then $ae(\langle G\rangle) \le C\, r^2$ for some universal constant C.*

In particular, it is not difficult to see that trees are precisely the class of graphs which exclude K_3 as a minor. But the above result provides an even more significant extension to Theorem 1.5 by showing, for instance, that the family of planar metrics is absolutely extendable (it is well-known that every planar graph excludes K_5 as a minor).

To see a striking consequence of this result, which is proved in Section 5 in conjunction with a deep theorem of Robertson and Seymour [32], let us restrict ourselves (for the moment) to families of finite graphs and ask the question, for such a family, when is it that $ae(\langle\mathcal{F}\rangle) \le K_{\mathcal{F}}$ and when do we have $ae(\langle\mathcal{F}\rangle) = \infty$? Clearly if $\langle\mathcal{F}\rangle$ contains *all* finite metrics, then $ae(\langle\mathcal{F}\rangle) = \infty$ (see, for instance, the remark after the statement of Theorem 1.2, which shows that there is no uniform bound for finite subsets of L_1). It happens that this is the only case.

Corollary 1.8. *For a family of finite graphs $\mathcal{F}$, $ae(\langle\mathcal{F}\rangle) < \infty$* if and only if *$\langle\mathcal{F}\rangle$ does not contain all finite metrics.*

The techniques used in proving Theorem 1.7 also yield the following result.

Corollary 1.9. *If M is a two-dimensional Riemannian manifold of genus g, then for every subset $X \subseteq M$, we have $ae(X) \le C \cdot (g+1)$ for some universal constant C.*

We note here that since any n-point metric space is isometrically embeddable in a two-dimensional Riemannian manifold of genus $O(n^3)$, in Corollary 1.9, $ae(M)$ must tend to infinity with the genus of M.

Using our approach, together with the stochastic decomposition of [6] and the improved analysis of [8, 9] (which we generalize to arbitrary measurable metric spaces in Theorem 3.17), it is possible to obtain an asymptotic improvement over Theorem 1.3.

Theorem 1.10. *There exists a universal constant C such that for every n-point metric space X,*

$$ae(X) \le C \frac{\log n}{\log \log n}.$$

Our techniques, in conjunction with the decomposition of [7], also give a different kind of improvement to Theorem 1.4 (of a geometric flavor).

Theorem 1.11. *For any Banach space Z, $e(\ell_2^d, Z) \le C\sqrt{d}$.*

It follows that for any normed space Y, we have $e(Y, Z) \le C\sqrt{d} \cdot dist(Y, \ell_2^d)$, where $dist$ is the Banach-Mazur distance. In particular, when Y is a d-dimensional normed space, John's theorem [16] implies that $e(Y, Z) \le C \cdot d$. When Y is closer to a Hilbert space, improved bounds are achieved.

Finally, in the case when Y is an L_p space and $1 < p \le 2$, we obtain the following estimate for finite subsets.

Theorem 1.12. *For every $1 < p \le 2$ there exits a constant $C(p)$ such that for every integer n and every Banach space Z,*

$$e_n(L_p, Z) \le C(p) (\log n)^{1/p}.$$

The case $p = 2$ in Theorem 1.12 may be viewed as a "dual" to Theorem 1.2.

It has been asked in [18] whether $e_n(L_2, Z)$ is bounded for every Banach space Z. Since in [30] it was shown by the second named author that for $2 < p < \infty$, $e(L_2, L_p) = \infty$, it follows that the answer to this question is negative (this fact was overlooked in [30]). Since $e_n(L_2, L_2) = e_n(L_2, L_\infty) = 1$, it is of interest to give quantitative lower bounds on $e_n(L_2, L_p)$ for $2 < p < \infty$. We therefore analyze the proof in [30] and obtain the following lower bound.

Lemma 1.13. *For every integer n and every $2 < p < \infty$,*

$$e_n(L_2, L_p) \ge C \left(\frac{\log n}{\log \log n} \right)^{\frac{p-2}{p^2}},$$

where C is a universal constant.

1.3 Lipschitz functions which take values in a barycentric metric space

In recent years there has been considerable effort to obtain extension theorems for Lipschitz functions which take values in spaces which are not Banach spaces. A generalization of Kirszbraun's extension theorem to metric spaces with certain curvature bounds was proved by Lang and Schroeder in [26]. The problem of estimating $e(X, Y, Z)$ when Z is the hyperbolic space $\mathbb{H}^n$ arose in the context of geometric group theory. This problem was posed by Gromov in [11] and has been subsequently studied by Lang, Pavlović and Schroeder in [23]. We refer to [5, 23, 26] and the references therein for a selection of related results.

Since our approach to the extension problem is based on a random construction and an averaging argument, it is most natural to present it in the context of Banach spaces (as in this case taking expectation has a concrete meaning). However, our methods generalize to a wider class of target spaces which encompasses, for example, the spaces considered by Lang, Pavlović and Schroeder in [23]. The basic idea is that instead of taking expectations, we need to require that every compactly supported measure on the target space Z has a barycenter, and that the map which assigns to each measure its barycenter is Lipschitz continuous. A similar notion was introduced by Gromov in [13] (see also [33]). It will be instructive to begin by presenting the definition of Gromov codiffusion spaces from [13, 33]. For a metric space Z let $\mathcal{M}_Z$ be the set of all regular Borel probability measures on X, topologized with the total variation norm $\|\cdot\|_{TV}$. A *random walk*, or a *diffusion*, on Z is a continuous map $\mu : Z \to \mathcal{M}_Z$. A *codiffusion* on Z is a continuous map $c : M \to Z$ defined on a convex subset $M \subseteq \mathcal{M}_Z$ containing all the Dirac measures δ_z such that $c(\delta_z) = z$ for every $z \in Z$ and such that $c^{-1}(z)$ is convex for every $z \in Z$. The notion of a codiffusion is close to what we need, except that in the context of the Lipschitz extension problem it is natural to require that c is not only continuous, but also satisfies a certain Lipschitz condition, described in the following definition.

Definition 1.14 (Barycentric metric space). For a metric space Z let $\mathcal{M}_Z^{\text{bounded}}$ be the set of all regular Borel probability measures on Z with bounded support. We shall say that Z is barycentric if there exists a constant $\beta > 0$ and a map $c : \mathcal{M}_Z^{\text{bounded}} \to Z$ such that $c(\delta_z) = z$ for every $z \in Z$ and for every $\mu, \nu \in \mathcal{M}_Z^{\text{bounded}}$,

$$d_Z(c(\mu), c(\nu)) \le \beta \cdot \operatorname{diam}(\operatorname{supp}(\mu + \nu)) \cdot \|\mu - \nu\|_{TV}. \tag{1}$$

The least constant β for which there exists such a mapping c is denoted $\beta(Z)$. When $\beta(Z) \le \beta$ we say that Z is β-barycentric.

If Z is a Banach space then $\beta(Z) = 1$. Indeed, for $\mu \in \mathcal{M}_Z^{\text{bounded}}$ we define $c(\mu)$ to be the usual *"center of mass"* $\int_Z x d\mu(x)$. Barycentric metric spaces need not, however, be linear. Examples of barycentric spaces are *Hadamard spaces*, i.e. complete geodesic metric spaces satisfying the CAT(0) comparison inequality of Alexandrov-Toponogov (see [23]). This class of spaces includes Hadamard manifolds, i.e. complete, simply connected Riemannian manifolds of nonpositive sectional curvature and is closed under taking products and gluing along closed convex subsets. If Z is a Hadamard space and $\mu \in \mathcal{M}_Z^{\text{bounded}}$ then $c(\mu)$ is defined as the unique minimizer of the function $z \mapsto \int_Z d(y,z)^2 d\mu(y)$. The existence and uniqueness of $c(\mu)$, as well as the fact that this map satisfies the barycentric condition (1) with $\beta = 1$ are proved in Section 4 of [23].

Most of the results presented above (namely Theorem 1.6, Theorem 1.7, Corollary 1.8, Corollary 1.9, Theorem 1.11 and Theorem 1.12) transfer to the case when the target space Z is allowed to be any complete barycentric metric space, the only difference being that when Z is β-barycentric, the estimates for the Lipschitz constant of the extended function are multiplied by β. Our proof of Theorem 1.10 does not seem to admit such a generalization—this issue is briefly discussed in Appendix 6. We present our results first for the case of linear Z, and in Appendix 6 we explain the simple modifications required to transfer them to arbitrary barycentric target spaces.

1.4 Some open problems

We end this introduction by recalling some related problems which remain open. In [1], K. Ball asked whether $e(L_2, L_1)$ is finite or infinite. Similarly, it is not known whether for $2 < p < \infty$ and $1 < q < 2$, $e(L_p, L_q) < \infty$. The second problem will be answered affirmatively if Ball's Markov type 2 problem [1] is resolved positively. Finally, it is not known whether, for every metric space X and every normed space Z, $e_n(X, Z) = O\left(\sqrt{\log n}\right)$. All known lower bounds fail to beat the "$\sqrt{\log n}$ barrier."

The flexibility of our approach suggests that it could be applied to other extension problems. For example, the extendability of large scale Lipschitz maps (studied in [22]), Hölder maps (studied in [30]) and uniformly continuous maps (studied in [4]) are all of great interest. It is also interesting to study the applicability of the random method presented in this paper to the higher order Whitney extension problem.

Acknowledgements: We are grateful to Misha Gromov for his support and many helpful suggestions. Specifically, we thank him for suggesting that we look at codiffusion spaces. We also thank Yuval Peres for several useful discussions.

2 Gentle partitions of unity

Let (Y, d) be a metric space, X a subspace of Y and $(\Omega, \mathcal{F}, \mu)$ a measure space. Given $K > 0$ we shall say that a function $\Psi : \Omega \times Y \to [0, \infty)$ is a *K-gentle partition of unity with respect to X* if the following conditions hold true:

1. For every $x \in Y \setminus \overline{X}$ the function $\omega \mapsto \Psi(\omega, x)$ is measurable and $\int_\Omega \Psi(\omega, x) d\mu(\omega) = 1$.
2. For every $\omega \in \Omega$ and $x \in \overline{X}$, $\Psi(\omega, x) = 0$.
3. There exists a Borel measurable function $\gamma : \Omega \to \overline{X}$ such that for every $x, y \in Y$,

$$\int_\Omega d(\gamma(\omega), x) \cdot |\Psi(\omega, x) - \Psi(\omega, y)| \, d\mu(\omega) \le K d(x, y).$$

If in condition (1) above, we require that the set $\{\omega \in \Omega : \ \Psi(\omega, x) > 0\}$ is finite, we shall say the partition of unity Ψ is *locally finite*.

Lemma 2.1. *Let (Y, d) be a metric space and X a subspace of Y. Fix $K \ge 1$ and assume that Y admits a K-gentle partition of unity $\Psi : \Omega \times Y \to [0, \infty)$ with respect to X. Let C be a closed convex set in some Banach space Z. If Ψ is locally finite or if X is separable, then every Lipschitz function $f : X \to C$ can be extended to a function $\tilde{f} : Y \to C$ such that $||\tilde{f}||_{\mathrm{Lip}} \le 3K||f||_{\mathrm{Lip}}$. Furthermore, the extension depends linearly and continuously on f.*

Proof. The completeness of C implies that f can be extended to the closure of X with the same Lipschitz constant, so we may assume without loss of generality that X is closed. Let $\gamma : \Omega \to T$ be as in condition (2) above. If Ψ is locally finite, then all integrals appearing below reduce to finite sums, and we may ignore measurability issues. Thus we now assume that C is closed and X is separable.

In what follows we refer to [3] for the basic facts on Bochner integration which we use. Since X is separable we may assume that Z is separable. In this case, Pettis' measurability lemma implies that for every $y \in Y$ the function $\omega \mapsto f(\gamma(\omega))\Psi(\omega, y)$ is Bochner measurable. Moreover, for every $x \in X$,

$$\begin{aligned}
\int_\Omega \|f(\gamma(\omega))\|\Psi(\omega,y)\,d\mu(\omega) &\le \int_\Omega \big(\|f(\gamma(\omega)) - f(x)\| + \|f(x)\|\big)\Psi(\omega,y)\,d\mu(\omega) \\
&\le \|f\|_{\mathrm{Lip}} \cdot \int_\Omega d(\gamma(\omega),x)\cdot|\Psi(\omega,x) - \Psi(\omega,y)|\,d\mu(\omega) + \|f(x)\| \\
&\le \|f\|_{\mathrm{Lip}} \cdot K d(x,y) + \|f(x)\| < \infty.
\end{aligned}$$

Since C is closed, it follows that the Bochner integral $\int_\Omega d(\gamma(\omega),x)\Psi(\omega,y)\,d\mu(\omega)$ is a well-defined element of C. We can therefore define, for $x \in Y$,

$$\tilde{f}(x) = \begin{cases} \int_\Omega f(\gamma(\omega))\Psi(\omega,x)\,d\mu(\omega) & x \in Y \setminus X, \\ f(x) & x \in X. \end{cases} \tag{2}$$

Clearly $\tilde{f}$ is an extension of f. To check the Lipschitz condition, take $x \in Y$ and $y \in Y \setminus X$. Fix $z \in X$ such that $d(x,z) \le 2\,d(x,X)$, and observe that

$$\tilde{f}(y) - \tilde{f}(x) = \int_\Omega \big[f(\gamma(\omega)) - f(z)\big]\cdot\big[\Psi(\omega,y) - \Psi(\omega,x)\big]\,d\mu(\omega).$$

Indeed, this is obviously true if $x, y \in Y \setminus X$, while if $x \in X$ then necessarily $x = z$, and the required identity follows from (2) and the fact that $\Psi(\omega, z) \equiv 0$.

Now,

$$\begin{aligned}
||\tilde{f}(x) - \tilde{f}(y)|| &\le \|f\|_{\mathrm{Lip}} \int_\Omega d(\gamma(\omega),z)\cdot|\Psi(\omega,x) - \Psi(\omega,y)|\,d\mu(\omega) \\
&\le \|f\|_{\mathrm{Lip}} \int_\Omega \big[d(\gamma(\omega),x) + d(x,z)\big]\cdot|\Psi(\omega,x) - \Psi(\omega,y)|\,d\mu(\omega) \\
&\le 3\,\|f\|_{\mathrm{Lip}} \int_\Omega d(\gamma(\omega),x)\cdot|\Psi(\omega,x) - \Psi(\omega,y)|\,d\mu(\omega) \le 3\,K\cdot\|f\|_{\mathrm{Lip}}\,d(x,y).
\end{aligned}$$

□

3 Stochastic metric decomposition

In this section, we introduce various notions of "well-behaved" random (pointed) partitions of a metric space and show that they exist in several important cases. Section 4 establishes that such decompositions imply the existence of gentle partitions of unity.

3.1 Well-behaved decompositions

In what follows, we use the convention that the distance from a point in a metric space to the empty set is ∞.

Definition 3.1 (Stochastic decomposition). Let (Y, d) be a metric space and X a subspace of Y. We shall say that $(\Omega, \mu, \{\Gamma^i(\cdot), \gamma^i(\cdot)\}_{i\in I})$ is a *stochastic decomposition of* Y *with respect to* X if I is some index set, (Ω, μ) is a probability space, for every $\omega \in \Omega$, $\{\Gamma^i(\omega)\}_{i\in I}$ is a partition of Y into Borel subsets and for every $x \in Y$ the set $\{i \in I :\ \exists\,\omega \in \Omega \text{ such that } x \in \Gamma^i(\omega)\}$ is countable. We assume that for every $i \in I$, $\gamma^i : \Omega \to X$ is a Borel measurable function such that for every $\omega \in \Omega$, $d\left(\gamma^i(\omega), \Gamma^i(\omega)\right) < 2d\left(X, \Gamma^i(\omega)\right)$. Finally, we require that for every $x \in Y$ and $i \in I$ the set $\{\omega \in \Omega : \ x \in \Gamma^i(\omega)\}$ is measurable.

If $\{\{\Gamma^i(\omega)\}_{i\in I} : \ \omega \in \Omega\}$ ranges over only a finite number of partitions of Y we shall say that the decomposition $(\Omega, \mu, \{\Gamma^i(\cdot), \gamma^i(\cdot)\}_{i\in I})$ is *finitely supported*.

Definition 3.2 (Bounded decomposition). Let $(\Omega, \mu, \{\Gamma^i(\cdot), \gamma^i(\cdot)\}_{i\in I})$ be a stochastic decomposition of Y with respect to X and $\Delta > 0$. We say that it is Δ-bounded if for every $\omega \in \Omega$ and $i \in I$, $\operatorname{diam}(\Gamma^i(\omega)) \le \Delta$.

Definition 3.3 (Padded decomposition). Let $(\Omega, \mu, \{\Gamma^i(\cdot), \gamma^i(\cdot)\}_{i\in I})$ be a Δ-bounded stochastic decomposition of Y with respect to X and $\varepsilon, \delta > 0$. We say that it is (ε, δ)-*padded* if for every $x \in Y$ and $i \in I$ the function $\omega \mapsto d\left(x, Y \setminus \Gamma^i(\omega)\right)$ is measurable and if $d(x, X) \le \varepsilon\Delta$ then

$$\mu\left(\bigcup_{i\in I}\left\{\omega : \ d\left(x, X \setminus \Gamma^i(\omega)\right) \ge \varepsilon\Delta\right\}\right) \ge \delta.$$

Observe that the assumptions in Definition 3.1 imply that the above union is countable.

Remark 3.4. Note that when Y is countable, the requirement the function $\omega \mapsto d\left(x, Y \setminus \Gamma^i(\omega)\right)$ is measurable is redundant (i.e. it is enough to demand that $\{\omega \in \Omega : \ x \in \Gamma^i(\omega)\}$ is measurable). Indeed, for every $\rho > 0$,

$$\left\{\omega \in \Omega : \ d\left(x, Y \setminus \Gamma^i(\omega)\right) < \rho\right\} = \bigcup_{y \in B^\circ(x,\rho)} \{\omega \in \Omega : \ y \notin \Gamma^i(\omega)\}.$$

This observation is useful since if we are interested in extending Lipschitz functions from a closed subset X of a *separable* space Y then it is enough to restrict our attention to the case when Y is countable. Indeed, let $f : X \to Z$ be a Banach space-valued Lipschitz function and $S \subseteq X$, $T \subseteq Y$ be countable dense subsets in X and $Y \setminus X$, respectively. Clearly S is closed in $S \cup T$ and if we extend $f|_S$ to a Lipschitz function defined on $S \cup T$ then we may pass to the closure and obtain the required extension without further loss in the Lipschitz constant.

Definition 3.5 (Thick decomposition). Let $(\Omega, \mu, \{\Gamma^i(\cdot), \gamma^i(\cdot)\}_{i\in I})$ be a Δ-bounded stochastic decomposition of Y with respect to X and $\varepsilon, \delta > 0$. We say that it is (ε, δ)-*thick* if for every $x \in Y$ and $i \in I$ the function $\omega \mapsto d\left(x, Y \setminus \Gamma^i(\omega)\right)$ is measurable and if $d(x, X) \le \varepsilon\Delta$ then

$$\int_\Omega \sum_{i\in I} \min\left\{d\left(x, Y \setminus \Gamma^i(\omega)\right), \Delta\right\} \, d\mu(\omega) \ge \delta\Delta.$$

Observe that since $\{\Gamma^i(\omega)\}_{i=1}^\infty$ is a partition of X, the above sum contains only *one* element.

Remark 3.6. Observe that if $(\Omega, \mu, \{\Gamma^i(\cdot), \gamma^i(\cdot)\}_{i\in I})$ is (ε, δ)-padded then it is also $(\varepsilon, \varepsilon\delta)$-thick.

Definition 3.7 (Separating decomposition). Let $(\Omega, \mu, \{\Gamma^i(\cdot), \gamma^i(\cdot)\}_{i\in I})$ be a Δ-bounded stochastic decomposition of Y with respect to X and $\varepsilon, \delta > 0$. We say that it is (ε, δ)-*separating* if for every $x, y \in Y$ such that $d(\{x, y\}, X) \le \varepsilon\Delta$,

$$\int_\Omega \sum_{i\in I} |\mathbf{1}_{\Gamma^i(\omega)}(x) - \mathbf{1}_{\Gamma^i(\omega)}(y)| \, d\mu(\omega) \le \frac{2d(x, y)}{\delta\Delta}.$$

The next lemma simply states that for a metric space (Y, d) and an arbitrary closed subspace X, an (ε, δ)-padded decomposition of X (with respect to itself) can be extended to a (roughly) (ε, δ)-padded decomposition of Y with respect to X. This will make it possible for us to place assumptions only on X, letting Y be arbitrary.

Lemma 3.8 (Partition extension). *Let (Y, d) be a metric space and X a closed subspace of Y. If X admits a finitely supported (ε, δ)-padded Δ-bounded stochastic decomposition (with respect to itself), then Y admits a finitely supported $(\frac{\varepsilon}{16+8\varepsilon}, \delta)$-padded $(1+\frac{\varepsilon}{2})\Delta$-bounded stochastic decomposition with respect to X.*

Proof. Let $(\Omega, \mu, \{\Gamma^i(\cdot), \gamma^i(\cdot)\}_{i\in I})$ be a finitely supported (ε, δ)-padded Δ-bounded stochastic decomposition of X. For every point $x \in Y$, let $t_x \in X$ be such that $d(x, t_x) \le 2d(x, X)$. Now, for every $\omega \in \Omega$ and $i \in I$, create a new set

$$\widehat{\Gamma}^i(\omega) = \Gamma^i(\omega) \bigcup \left\{ x \in Y : \ d\left(t_x, X \setminus \Gamma^i(\omega)\right) \ge \frac{\varepsilon\Delta}{2} \text{ and } d(x, t_x) \le \frac{\varepsilon\Delta}{4} \right\}.$$

Finally, for any point $x \in Y \setminus \bigcup_{i\in I} \widehat{\Gamma}^i(\omega)$, place x in a singleton cluster $\{x\}$. This constitutes a finitely supported distribution over partitions of Y. The selection function corresponding to $\widehat{\Gamma}^i(\cdot)$ is simply chosen to be $\hat{\gamma}^i = \gamma^i$, while the selection function corresponding to a singleton cluster $\{x\}$ is an arbitrary point $t \in X$ satisfying $d(t, x) \le 2d(x, X)$.

Let us now show that the above stochastic decomposition is $(\frac{\varepsilon}{16}, \delta)$-padded and $(1 + \frac{\varepsilon}{2})\Delta$-bounded. The $(1 + \frac{\varepsilon}{2})\Delta$-bounded condition is easy; singleton clusters have diameter zero. For points $x, y \in \widehat{\Gamma}^i(\omega)$, $i \in I$, we have

$$d(x, y) \le d(x, t_x) + d(t_x, t_y) + d(y, t_y) \le \frac{\varepsilon\Delta}{4} + \operatorname{diam}\left(\Gamma^i(\omega)\right) + \frac{\varepsilon\Delta}{4} \le \left(1 + \frac{\varepsilon}{2}\right)\Delta.$$

Fix some $x \in Y$ with $d(x, X) \le \frac{\varepsilon}{16+8\varepsilon}\left(1 + \frac{\varepsilon}{2}\right)\Delta = \frac{\varepsilon\Delta}{16}$. By the definition of padded decompositions, with probability at least δ, $d\left(t_x, X \setminus \Gamma^i(\omega)\right) \ge \varepsilon\Delta$ for some $i \in I$. Our goal will be to show that in this case we have $d\left(x, Y \setminus \widehat{\Gamma}^i(\omega)\right) \ge \frac{\varepsilon\Delta}{16}$, which will complete the proof. Assume to the contrary that there is some $y \in Y \setminus \widehat{\Gamma}^i(\omega)$ with $d(x, y) \le \frac{\varepsilon\Delta}{16}$. Observe that

$$d(t_x, t_y) \le d(t_x, x) + d(x, y) + d(y, t_y) \le 2d(x, X) + \frac{\varepsilon\Delta}{16} + 2\left(d(x, X) + d(x, y)\right) < \frac{\varepsilon\Delta}{2}.$$

Hence,

$$d\left(t_y, X \setminus \Gamma^i(\omega)\right) \ge d\left(t_x, X \setminus \Gamma^i(\omega)\right) - d(t_x, t_y) > \frac{\varepsilon\Delta}{2},$$

Since we also have that

$$d(y, t_y) \le 2d(y, X) \le 2d(x, X) + 2d(x, y) \le \frac{\varepsilon\Delta}{4},$$

we see that $y \in \widehat{\Gamma}^i(\omega)$, which contradicts the choice of y. $\square$

Remark 3.9. For later applications, we note that if X is $\frac{\varepsilon\Delta}{32}$-dense in Y and $\varepsilon < 1$, then the above proof shows that Y admits an $(\frac{\varepsilon}{32}, \delta)$-padded 2Δ-bounded stochastic decomposition with respect to itself.

Remark 3.10. Lemma 3.8 holds true when Y is countable and the decomposition in not necessarily finitely supported (in which case the extended decomposition is not necessarily finitely supported either). The proof is the same, and all one has to observe is that in this case the resulting decomposition satisfies the required measurability assumptions.

3.2 Constructions

In this section, we construct stochastic decompositions for various classes of metric spaces. Most of the constructions come directly from the theoretical computer science literature, but since we are dealing here with infinite spaces, we must be somewhat delicate in placing these finite random processes into the appropriate probability spaces and dealing with the measurability issues that arise. Because of this, some of the constructions are restated in a form which is different than that in which they originally appeared.

Let (X,d) be a metric space and $R > r > 0$. Denote by $\mathcal{C}(X;R,r)$ the largest cardinality of a set $N \subseteq X$ satisfying for every distinct $x, y \in N$, $r \le d(x,y) \le R$. The following result is essentially contained in [14], though we must deal with a technical issue due to our use of possibly infinite nets.

Lemma 3.11. *For every $\Delta > 0$, every metric space (X,d) admits an $\left(\varepsilon, \frac{1}{2}\right)$-padded Δ-bounded finitely supported stochastic decomposition of X with respect to itself, where $\varepsilon = \frac{1}{256\log[\mathcal{C}(X;2\Delta,\Delta/4)]}$.*

Proof. For ease of notation, we construct a 4Δ-bounded decomposition. We may assume that $\mathcal{C}(X;8\Delta,\Delta)$ is finite, since otherwise the result holds vacuously. Let N be a Δ-net of X. First, we need to introduce a distribution over partial orders $\prec$ on N such that for every ball $B \subset X$ of radius 3Δ, $\prec$ is a uniformly random total order on $B \cap N$ (note that $|B \cap N|$ is finite). To this end, consider the infinite graph $G = (N,E)$ where $\{x,y\} \in E$ if and only if $d(x,y) \le 3\Delta$. The degree of G is at most $\mathcal{C}(X;6\Delta,\Delta) < \infty$, and thus G admits a proper coloring using some finite number of color classes; call these classes $1, 2, \ldots, M$. Now let σ be a random permutation on $\{1,\ldots,M\}$ and let $\chi : N \to \{1,\ldots,M\}$ be a proper coloring. Finally, define $x \prec y$ if and only if $\sigma(\chi(x)) < \sigma(\chi(y))$. It is easy to see that for a ball B of radius 3Δ, every point in $B \cap N$ receives a unique color, and thus σ induces a uniformly random permutation on B. It follows that $\prec$ satisfies the desired properties.

Now choose a radius $R \in [\Delta, 2\Delta]$ uniformly at random. For each $y \in N$, define a cluster

$$C_y = \{x \in X : x \in B(y,R) \text{ and } y \prec z \text{ for all } z \in N \text{ with } x \in B(z,R)\}.$$

Since N is a Δ-net and $R \ge \Delta$, $P = \{C_y\}_{y \in N}$ constitutes a partition of X. Clearly all the clusters C_y have diameter at most 4Δ. Finally, we construct the required selectors as follows: for $y \in N$ let γ^y be the minimal element of N in C_y (with respect to $\prec$).

Now fix a value $t \in [0,\Delta]$ and some $x \in X$. Let $W = B(x, 2\Delta + t) \cap N$, and note that $m = |W| \le \mathcal{C}(X;6\Delta,\Delta)$. Arrange the points $w_1,\ldots,w_m \in W$ in order of increasing distance from x, and let I_k be the interval $[d(x,w_k) - t, d(x,w_k) + t]$. Let us say that $B(x,t)$ is *cut* if for some cluster C_{w_k}, $C_{w_k} \cap B(x,t) \neq \emptyset$, but $B(x,t) \nsubseteq C_{w_k}$. Finally, write $\mathcal{E}_k$ for the event that w_k is the minimal element in W (according to $\prec$) for which C_{w_k} cuts $B(x,t)$. Then,

$$\begin{aligned}
\Pr[B(x,t) \text{ is cut}] &\le \sum_{k=1}^{m} \Pr[\mathcal{E}_k] \\
&= \sum_{k=1}^{m} \Pr[R \in I_k] \cdot \Pr[\mathcal{E}_k \mid R \in I_k] \\
&\le \sum_{k=1}^{m} \frac{2t}{\Delta} \cdot \frac{1}{k} \le \frac{2t}{\Delta}(1 + \log m) \le \frac{8t}{\Delta} \log\left[\mathcal{C}(X;6\Delta,\Delta)\right].
\end{aligned}$$

Setting $t = \frac{\Delta}{64 \log[\mathcal{C}(X;6\Delta,\Delta)]}$ yields the required result. We remark that the stochastic decomposition can be made finitely supported by choosing R uniformly from a sufficiently fine mesh in $[\Delta, 2\Delta]$. $\square$

Corollary 3.12 (Doubling metrics). *Let (X, d) be a doubling metric space. Then for every $\Delta > 0$, there exists a $\left(\frac{1}{C \log \lambda(X)}, \frac{1}{2}\right)$-padded Δ-bounded finitely supported stochastic decomposition of X with respect to itself, where C is a universal constant.*

Proof. This is a direct consequence of Lemma 3.11, since it is evident that $\log[\mathcal{C}(X; 2\Delta, \Delta/4)] = O(\log \lambda(X))$. Indeed, let $N \subseteq X$ be a set such that, for every distinct $x, y \in N$, $\frac{\Delta}{4} \le d(x, y) \le 2\Delta$. Then clearly X is contained in a ball of radius 2Δ which can be covered by $\lambda(X)^4$ balls of radius $\frac{\Delta}{8}$. Since every such ball contains at most one point of N, we see that $|N| \le \lambda(X)^4$. $\square$

Corollary 3.13 (Negatively curved manifolds). *Fix $r > 0$ and let M be an n-dimensional Riemannian manifold satisfying $\operatorname{Ricci}(g) \ge -(n-1)rg$, where g is the Riemannian metric on M and $\operatorname{Ricci}(g)$ is the Ricci curvature of g. Then for any subset $A \subseteq M$ and every $\Delta > 0$, A admits an $\left(\varepsilon, \frac{1}{2}\right)$-padded Δ-bounded finitely supported stochastic decomposition (with respect to itself), where $\varepsilon = \frac{c}{n\left(1+\sqrt{r\Delta}\right)}$ and c is a universal constant.*

Proof. Fix $0 < R_1 < R_2$ and $x \in M$. The following inequality is a well known consequence of Bishop's inequality (see Lemma 5.3, pg. 275 in [12]):

$$\frac{\operatorname{Vol}(B_M(x, R_2))}{\operatorname{Vol}(B_M(x, R_1))} \le \frac{\int_0^{R_2} \left(e^{\sqrt{r}t} - e^{-\sqrt{r}t}\right)^{n-1} dt}{\int_0^{R_1} \left(e^{\sqrt{r}t} - e^{-\sqrt{r}t}\right)^{n-1} dt}.$$

By standard arguments it follows that:

$$\log[\mathcal{C}(A; 2\Delta, \Delta/4)] \le \log[\mathcal{C}(M; 2\Delta, \Delta/4)] \le O\left(1 + \sqrt{r\Delta}\right) \cdot n,$$

so that Lemma 3.11 implies the required result. $\square$

A strong decomposition for excluded-minor spaces follows from the (graph) decompositions of [20], and a similar notion was used in [31] to embed planar metrics into ℓ_2. We sketch the argument below.

Lemma 3.14 (Excluded minors). *Let G be a graph which does not possess K_r as a minor, and suppose that $X \in \langle G \rangle$. Then for every $\Delta > 0$, X admits a $\left(\frac{c}{r^2}, \frac{1}{2}\right)$-padded Δ-bounded finitely supported stochastic decomposition with respect to itself.*

Proof. Let G be a weighted graph without a K_r minor such that $X \subseteq \Sigma(G)$. Fix some $\beta \in (0, 1]$ and $k \in \mathbb{N}$. Consider the distribution μ on $\mathcal{P}$ which arises from the following random process. We decompose $\Sigma(G)$ recursively. Let x_0 be an arbitrary point of X. Choose now some $u \in [0, \beta\Delta)$ uniformly at random and consider, for each $n \in \mathbb{N} \cup \{0\}$, the annuli

$$A_n = \left\{x \in \Sigma(G) : (n-1)\beta\Delta + u \le d_{\Sigma(G)}(x, x_0) < n\beta\Delta + u\right\}.$$

In general, the sets A_n may be disconnected in the topology of $\Sigma(G)$. Let $\mathcal{C}$ denote the set of (disjoint) connected components of the $\{A_n\}$, and apply the above random process again to each component in $\mathcal{C}$. The process ends after k such steps, producing a partition P of $\Sigma(G)$.

In [20], it is shown that for some fixed $\beta = \beta(r)$ and $k = r$, the above process produces a partition P such that for every $C \in P$, $\mathrm{diam}(C) \le \Delta$. Improved quantitative bounds were obtained in [10], yielding $\beta(r) = \Omega\left(\frac{1}{r}\right)$.

For every partition P and $x \in X$ let $\pi_P(X)$ be the maximal $r \ge 0$ for which there exists $A \in P$ such that $B_{\Sigma(G)}(x,r) \subseteq A$. We claim that for any $x \in \Sigma(G)$ and every $t > 0$, $\Pr[\pi_P(x) \ge t] \ge \left(1 - \frac{2t}{\beta\Delta}\right)^k$. To see this, simply note that the probability of $B_{\Sigma(G)}(x,t)$ being separated into two different connected components at any step of the decomposition is at most $\frac{2t}{\beta\Delta}$. Choosing $t = \Omega\left(\frac{\Delta}{r^2}\right)$ gives $Pr[\pi_P(x) \ge t] \ge \frac{1}{2}$. This random partition can be modified to have finite support by choosing u uniformly from a sufficiently fine mesh in $[0, \beta\Delta)$, so that $\Sigma(G)$, and hence also X, admits the required padded stochastic decomposition. $\square$

Corollary 3.15 (Surfaces of bounded genus). *Let M be a two dimensional Riemannian manifold with genus g and $X \subseteq M$. Then for every $\Delta > 0$, X admits a $\left(\frac{c}{g+1}, \frac{1}{2}\right)$-padded Δ-bounded finitely supported stochastic decomposition with respect to itself.*

Proof. Let N be an η-net in X. For every $x, y \in N$ let $\ell_{x,y}$ be a minimal length geodesic joining x and y. Consider the set $N' \supseteq N$ obtained from adding all the points of intersection of the geodesics $\{\ell_{x,y}\}_{x,y \in N}$. Consider the graph $G = (N', E)$, where $\{u, v\} \in E$ if there is some $x, y \in N$ such that u and v are connected by a sub-geodesic $\ell \subseteq \ell_{x,y}$ for which $\ell \cap N' = \{u, v\}$. By construction, the graph G is embedded in M (in the graph-theoretic sense). It follows (see [29]) that G excludes a $K_{\Omega(\sqrt{g+1})}$ minor. Furthermore, N is isometric to a subset of $\Sigma(G)$ where the weight of an edge $\{u, v\}$ is equal to the length of the sub-geodesic connecting u and v. Hence N is a $K_{\Omega(\sqrt{g+1})}$-excluded metric, so that the required result follows from Lemma 3.14 and Remark 3.9 (for η small enough). $\square$

Optimal decompositions for finite subsets of ℓ_2^d were given in [7]. The following lemma is based on their techniques.

Lemma 3.16 (Finite dimensional Hilbert space.). *For any closed subset X of ℓ_2^d and for every $\Delta > 0$ there exists a stochastic decomposition of ℓ_2^d with respect to X which is (ε, δ)-separating and Δ-bounded for every $\varepsilon > 0$ and $\delta = \frac{1}{2\sqrt{d}}$.*

Proof. We construct a graph $G = (\mathbb{Z}^d, E)$ where for $u, v \in \mathbb{Z}^d$, $\{u, v\}$ is an edge if and only if $d\left(u + [0,1)^d, v + [0,1)^d\right) \le 8\Delta$. As before, the degree of G is uniformly bounded, and thus it admits an M-coloring $\chi : \mathbb{Z}^d \to \{1, 2, \ldots, M\}$ such that $\chi(u) \ne \chi(v)$ whenever $d\left(u + [0,1)^d, v + [0,1)^d\right) \le 8\Delta$.

Denote by m_d the Lebesgue measure on $[0,1]^d$ and by ν the uniform probability measure on $\{1, \ldots, M\}$. Consider the product space $\Omega = \prod_{i=1}^{\infty}\left([0,1)^d \times \{1, \ldots, M\}\right)$ equipped with the natural product measure μ. Given $\omega = (x_1, c_1, x_2, c_2, \ldots) \in \Omega$ we construct recursively a sequence of disjoint subsets of $\mathbb{R}^d$, $\{\Gamma^{i,v}(\omega) : \ i \in \mathbb{N}, \ v \in \mathbb{Z}^d\}$ as follows:

$$\Gamma^{k,v}(\omega) = \begin{cases} B_2\left(v + x_k, \frac{\Delta}{2}\right) \setminus \left(\bigcup_{j=1}^{k-1} \bigcup_{u \in \mathbb{Z}^d} \Gamma^{j,u}(\omega)\right) & \text{if } \chi(v) = c_k \\ \emptyset & \text{otherwise,} \end{cases}$$

where $B_2(y, \rho)$ denotes the closed Euclidean ball of radius ρ centered at y. The sets $\{\Gamma^{i,v}(\omega) : i \in \mathbb{N}, \ v \in \mathbb{Z}^d\}$ form a partition of $[0,1)^d$ with probability one. We claim that there exist measurable maps $\gamma^{i,v} : \Omega \to X$ such that for every $\omega \in \Omega$, $d\left(\gamma^{i,v}(\omega), \Gamma^{i,v}(\omega)\right) \le 2d\left(X, \Gamma^{i,v}(\omega)\right)$.

Indeed, by a classical measurable selection theorem of Kuratowski and Ryll-Nardzewski [21] it is enough to check that for every open ball $B \subseteq R^d$ the set

$$\left\{\omega \in \Omega : \; B \cap X \cap \left\{a \in \mathbb{R}^d ; \; d\left(a, \Gamma^{i,v}(\omega)\right) \le 2d\left(X, \Gamma^{i,v}(\omega)\right)\right\} \neq \emptyset\right\}$$

is measurable, and this fact follows directly from the construction.

The above decomposition is trivially Δ-bounded. Now, fix $x, y \in \mathbb{R}^d$. Since the (ε, δ)-separating condition is trivial for $d(x,y) > \Delta$, assume that $d(x,y) \le \Delta$. We now bound the probability that x and y end up in different clusters. It follows from the construction that x and y are separated in the partition induced by $\omega = (x_1, c_1, x_2, c_2, \ldots)$ precisely when there is an index j for which $\left[x_j + \chi^{-1}(c_j)\right] \cap \left[B_2\left(x, \frac{\Delta}{2}\right) \triangle B_2\left(y, \frac{\Delta}{2}\right)\right] \neq \emptyset$ while $\left[x_i + \chi^{-1}(c_i)\right] \cap \left[B_2\left(x, \frac{\Delta}{2}\right) \cap B_2\left(y, \frac{\Delta}{2}\right)\right] = \emptyset$ for every $i < j$. Notice that since $d(x,y) \le \Delta$, no two cubes from $\{v + [0,1)^d : \; v \in \mathbb{Z}^d\}$ which intersect $B_2\left(x, \frac{\Delta}{2}\right) \cup B_2\left(y, \frac{\Delta}{2}\right)$ have the same color. Denoting by $K \subseteq \mathbb{R}^n$ the union of all the cubes which intersect $B_2\left(x, \frac{\Delta}{2}\right) \cup B_2\left(y, \frac{\Delta}{2}\right)$, it follows that for every Borel measurable $A \subseteq K$, $\mu\left([x_k + \chi^{-1}(c_k)] \cap A \neq \emptyset\right) = \frac{\mathsf{Vol}(A)}{\mathsf{Vol}(K)}$. Hence

$$\begin{aligned}
\int_\Omega \sum_{i \in I} |\mathbf{1}_{\Gamma^i(\omega)}(x) - \mathbf{1}_{\Gamma^i(\omega)}(y)| \, d\mu(\omega) &= 2\mu(\{x \text{ and } y \text{ are not in the same } \Gamma^i(\omega)\}) \\
&\le \; 2\sum_{j=1}^{n} \frac{\mathsf{Vol}\left(B_2\left(x, \frac{\Delta}{2}\right) \triangle B_2\left(y, \frac{\Delta}{2}\right)\right)}{\mathsf{Vol}(K)} \left[1 - \frac{\mathsf{Vol}\left(B_2\left(x, \frac{\Delta}{2}\right) \cap B_2\left(y, \frac{\Delta}{2}\right)\right)}{\mathsf{Vol}(K)}\right]^{j-1} \\
&= \; 2\frac{\mathsf{Vol}\left(B_2\left(x, \frac{\Delta}{2}\right) \triangle B_2\left(y, \frac{\Delta}{2}\right)\right)}{\mathsf{Vol}\left(B_2\left(x, \frac{\Delta}{2}\right) \cap B_2\left(y, \frac{\Delta}{2}\right)\right)} \le \frac{4\sqrt{d}}{\Delta}\|x-y\|_2,
\end{aligned}$$

by straightforward volume estimates. □

Finally, we present a decomposition theorem for general metric spaces Y with respect to a compact, measurable submetric X. The analysis is based on ideas from [6] and [9] for finite metrics.

Theorem 3.17. *Let (Y,d) be a metric space and X a compact subspace of Y. If σ is any non-degenerate Borel measure on X (i.e. one which assigns non-zero measure to every ball in X), then for every $\Delta > 0$, there exists a Δ-bounded stochastic decomposition of Y with respect to X such that, for every $x, y \in Y$ with $d(\{x,y\}, X) < \frac{\Delta}{16}$,*

$$\int_\Omega \sum_{i \in I} |\mathbf{1}_{\Gamma^i(\omega)}(x) - \mathbf{1}_{\Gamma^i(\omega)}(y)| \, d\mu(\omega) \le \frac{2\,d(x,y)}{\Delta}\left[1 + \log\left(\frac{\sigma(B_X(x, 5\Delta))}{\sigma(B_X(x,\Delta))}\right)\right]. \tag{3}$$

Proof. Since X is compact, we may assume that σ is a probability measure on X. Let us then equip the product space $\Omega' = \prod_{i=1}^\infty X$ with the natural product measure μ'. Finally, let $\Omega = \Omega' \times [2\Delta, 4\Delta]$, equipped with the probability measure $\Pr = \mu' \times \lambda$, where λ is the normalized Lebesgue measure on $[2\Delta, 4\Delta]$. Given $\omega = (x_1, x_2, x_3, \ldots) \in \Omega'$ and some $R \in [2\Delta, 4\Delta]$, we construct recursively a sequence of disjoint subsets of Y, $\{\Gamma^i(\omega, R) : i \in \mathbb{N}\}$ as follows,

$$\Gamma^k(\omega, R) = B(x_k, R) \setminus \left(\bigcup_{j=1}^{k-1} \Gamma^j(\omega)\right).$$

Let $S = \{y \in Y : d(x,y) < \frac{\Delta}{2}\}$, then with probability 1, for any $R \in [2\Delta, 4\Delta]$, the sets $\{\Gamma^i(\omega, R) \cap S\}_{i=1}^\infty$ form a partition of S. To see this, let N be a $\frac{\Delta}{4}$-net in X, and consider the balls $B(x,\frac{\Delta}{8})$ for $x\in N$. Since σ is non-degenerate, it assigns any such ball positive measure, and hence with probability 1, we have $x_j\in B(x,\Delta/8)$ for some $j\in\mathbb{N}$. Now fix a point $y\in Y$ such that $d(y,X)<\frac{\Delta}{2}$, and let $z\in X$ be such that $d(y,z)<2\,d(y,X)$. Observe that z is within $\frac{\Delta}{4}$ of some point of N, and hence within $\frac{3\Delta}{8}$ of some x_j with probability 1. But now we see that $d(y,x_j)\le d(y,z)+d(z,x_j)<\Delta+\frac{3\Delta}{8}<2\Delta$, hence $y\in B(x_j,R)$. It follows that the proposed sets are indeed a partition of S with probability 1.

We define the final partition by

$$\{\Gamma^i(\omega,R)\}_{i=1}^{\infty}\bigcup\left\{\{y\}:y\in Y\setminus\bigcup_{i\in\mathbb{N}}\Gamma^i(\omega,R)\right\},$$

and note that the required selectors exist due to an application of the Kuratowski, Ryll-Nardzewski Theorem as in the proof of Lemma 3.16.

Fix $x\in Y$ such that $d(x,X)<\frac{\Delta}{2}$ and denote by ν the distribution of the random variable $z\mapsto d(z,x)$, i.e. for $0\le\alpha<\beta$, $\nu([\alpha,\beta))=\sigma(\{z\in X:\ \alpha\le d(z,x)<\beta\})$. Fix $t\le\Delta$ and for every $R\in[2\Delta,4\Delta]$ denote by D_R the set of all $z\in X$ for which $B(z,R)\cap B(x,t)\notin\{\emptyset,B(x,t)\}$ (when this happens we say that $B(x,t)$ is cut by $B(z,R)$). Finally, let

$$\Omega_{i,R}=\{\omega\in\Omega':\ \omega_i\in D_R\}\setminus\bigcup_{j=1}^{i-1}\{\omega\in\Omega':\ \omega_j\in D_R\}.$$

Observe that for $\omega=(\omega_1,\omega_2,\dots)\in\Omega'$, if $\omega\in\Omega_{i,R}$ then the triangle inequality implies that $R\in[d(x,\omega_i)-t,d(x,\omega_i)+t]$. Moreover, if $\omega_i\in B(x,\Delta)$ then $B(x,t)\subseteq B(\omega_i,R)$, since $t\le\Delta$ and $R\ge 2\Delta$, so that $B(\omega_i,R)$ can't cut $B(x,t)$. Additionally, $B(\omega_i,R)\cap B(x,t)\neq\emptyset$ implies $\omega_i\in B(x,5\Delta)$, since $R\le 4\Delta$. It follows that if $\omega\in\Omega_{i,R}$, then $\Delta<d(\omega_i,x)\le 5\Delta$. Finally, if $\omega\in\Omega_{i,R}$ then $d(\omega_j,x)>d(\omega_i,x)$ for $j<i$, since otherwise $B(\omega_i,R)$ will not be the first ball to cut $B(x,t)$. These observations imply that for every $i=1,2,\dots$, every $R\in[2\Delta,4\Delta]$ and every $\rho>0$,

$$\mu'(\Omega_{i,R}|d(x,\omega_i)=\rho)\le\mathbf{1}_{\{R\in[\rho-t,\rho+t]\}}\cdot\mathbf{1}_{\{\Delta<\rho\le 5\Delta\}}[1-\nu([0,\rho])]^{i-1}.$$

Hence,

$$\begin{aligned}
\Pr[B(x,t)\text{is cut}] &= \frac{1}{2\Delta}\int_{2\Delta}^{4\Delta}\left(\sum_{i=1}^{\infty}\mu'(\Omega_{i,R})\right)dR\\
&= \frac{1}{2\Delta}\int_{2\Delta}^{4\Delta}\left(\sum_{i=1}^{\infty}\int_0^{\infty}\mu'(\Omega_{i,R}|d(x,\omega_i)=\rho)d\nu(\rho)\right)dR\\
&\le \frac{1}{2\Delta}\int_{2\Delta}^{4\Delta}\left(\sum_{i=1}^{\infty}\int_{(\Delta,5\Delta]}\mathbf{1}_{\{R\in[\rho-t,\rho+t]\}}\cdot[1-\nu([0,\rho])]^{i-1}d\nu(\rho)\right)dR\\
&\le \frac{t}{\Delta}\int_{(\Delta,5\Delta]}\frac{d\nu(\rho)}{\nu([0,\rho])}\\
&\le \frac{t}{\Delta}\log\left(\frac{\nu([0,5\Delta])}{\nu([0,\Delta])}\right)=\frac{t}{\Delta}\log\left(\frac{\sigma(B_X(x,5\Delta))}{\sigma(B_X(x,\Delta))}\right). \qquad (4)
\end{aligned}$$

The last inequality above is a classical fact, which can be proved as follows: Approximate the integral by the sum $\sum_{i=1}^{k}\frac{\nu([0,\rho_i])-\nu([0,\rho_{i-1}])}{\nu([0,\rho_i])}$ for some $\Delta<\rho_0<\rho_1<\ldots<\rho_k=5\Delta$, and use the estimate $\frac{\nu([0,\rho_i])-\nu([0,\rho_{i-1}])}{\nu([0,\rho_i])}\le\int_{\nu([0,\rho_{i-1}])}^{\nu([0,\rho_i])}\frac{ds}{s}$.

Observe that (3) holds trivially if $d(x,y) \geq \Delta$. Otherwise, choosing $t = d(x,y)$ in (4) yields the required result. $\square$

4 Constructing gentle partitions

In this section we show that the various decompositions that were introduced in the previous section can be used to construct gentle partitions of unity.

Theorem 4.1 (Stochastic decompositions yield gentle partitions). *There exists a universal constant $C > 0$ such that for every metric space (Y,d) and every subspace $X \subseteq Y$ the following assertions hold true:*

1. *If for every $n \in \mathbb{Z}$, Y admits an (ε,δ)-thick 2^n-bounded stochastic decomposition with respect to X, then Y also admits a $\frac{C}{\varepsilon\delta}$-gentle partition of unity with respect to X.*
2. *If for every $n \in \mathbb{Z}$, Y admits an (ε,δ)-padded 2^n-bounded stochastic decomposition with respect to X, then Y also admits a $\frac{C}{\varepsilon\delta}$-gentle partition of unity with respect to X.*
3. *If for every $n \in \mathbb{Z}$, Y admits an (ε,δ)-separating 2^n-bounded stochastic decomposition with respect to X, then Y also admits a $C\left(\frac{1}{\varepsilon}+\frac{1}{\delta}\right)$-gentle partition of unity with respect to X.*

Proof. For every $n \in \mathbb{Z}$ let $(\Omega_n, \mu_n, \{\Gamma_n^i(\cdot), \gamma_n^i(\cdot)\}_{i\in I})$ be a 2^n-bounded stochastic decomposition of Y with respect to X. Let $\varphi : \mathbb{R} \to \mathbb{R}_+$ be any 2-Lipschitz map with $\operatorname{supp}(\varphi) \subset [\frac{1}{2}, 4]$ and $\varphi \equiv 1$ on $[1,2]$. Define $\varphi_n(x) = \varphi\left(\frac{d(x,X)}{\varepsilon 2^{n-3}}\right)$ and let (Ω,μ) be the disjoint union of $\{I \times \Omega_n\}_{n\in\mathbb{Z}}$ (where the measure on I is the counting measure). In all the cases of the theorem the partition of unity which we construct will have the following form: For every $n \in \mathbb{Z}$, $\omega \in \Omega_n$, $i \in I$ and $x \in Y$ denote:

$$\Psi(i,\omega,x) = \frac{1}{S(x)}\,\theta_\omega^n(x)\varphi_n(x)\mathbf{1}_{\Gamma_n^i(\omega)}(x), \tag{5}$$

where for every $n \in \mathbb{Z}$ and $x \in Y$ the function $\omega \mapsto \theta_\omega^n(x) \in [0,\infty)$ is μ_n-integrable and

$$S(x) = \sum_{n\in\mathbb{Z}}\sum_{i\in I}\int_{\Omega_n} \theta_\omega^n(x)\varphi_n(x)\mathbf{1}_{\Gamma_n^i(\omega)}(x)\,d\mu_n(\omega) = \sum_{n\in\mathbb{Z}}\varphi_n(x)\int_{\Omega_n}\theta_\omega^n(x)\,d\mu_n(\omega).$$

Additionally define $\gamma(i,\omega) = \gamma_n^i(\omega)$.

Observe that $\operatorname{supp}(\varphi_n) \subseteq \left\{x \in Y :\ d(x,X) \in \left[\varepsilon 2^{n-4}, \varepsilon 2^{n-1}\right]\right\}$, so the sum in the denominator of (5) contains at most 5 terms, and is therefore finite. Additionally, the definition of φ_n ensures that for every $x \in \overline{X}$ and $\omega \in \Omega$, $\Psi(\omega,x) = 0$.

The functions $\theta_\omega^n(x)$ will be different in each particular case, but we begin by making some general comments. Our goal is to show that for every $x,y \in Y$,

$$\sum_{n\in\mathbb{Z}}\sum_{i\in I}\int_{\Omega_n} d(\gamma(i,\omega),x)\cdot|\Psi(i,\omega,x)-\Psi(i,\omega,y)|\,d\mu_n(\omega) \leq \frac{C}{\varepsilon\delta}\cdot d(x,y). \tag{6}$$

In cases (1) and (2) above and

$$\sum_{n\in\mathbb{Z}}\sum_{i\in I}\int_{\Omega_n} d(\gamma(i,\omega),x)\cdot|\Psi(i,\omega,x)-\Psi(i,\omega,y)|\,d\mu_n(\omega) \leq C\left(\frac{1}{\varepsilon}+\frac{1}{\delta}\right)\cdot d(x,y). \tag{7}$$

in case (3).

Claim 4.2. *Fix $\omega \in \Omega_n$ and assume that $\Psi(i,\omega,x) \neq \Psi(i,\omega,y)$. Then*

$$d(\gamma(i,\omega),x) \leq d(x,y) + \frac{18}{\varepsilon} \cdot \max\{d(x,X), d(y,X)\}.$$

Proof. Our assumption implies that either $\Psi(i,\omega,x) > 0$ or $\Psi(i,\omega,y) > 0$. In the first case, $x \in \Gamma_n^i(\omega)$ and

$$d\left(\gamma(i,\omega),x\right) \leq d\left(\gamma(i,\omega),\Gamma_n^i(\omega)\right) + \operatorname{diam}(\Gamma_n^i(\omega)) \leq 2d\left(X,\Gamma_n^i(\omega)\right) + 2^n \leq 2d(x,X) + 2^n.$$

On the other hand, $\varphi_n(x) > 0$, so $d(x,X) \geq \varepsilon 2^{n-4}$, which implies the required estimate. In the second case, $\Psi(i,\omega,y) > 0$, so that

$$\begin{aligned} d(\gamma(i,\omega),x) &\leq d(\gamma(i,\omega),y) + d(x,y) \leq 2d(y,X) + \operatorname{diam}(\Gamma_n^i(\omega)) + d(x,y) \\ &\leq 2d(y,X) + 2^n + d(x,y) \leq \frac{18}{\varepsilon} \cdot d(y,X) + d(x,y). \end{aligned}$$

□

By Claim 4.2 we can estimate the left-hand side of (6) and (7) as follows:

$$\begin{aligned} &\sum_{n\in\mathbb{Z}} \sum_{i\in I} \int_{\Omega_n} d(\gamma(i,\omega),x) \cdot |\Psi(i,\omega,x) - \Psi(i,\omega,y)|\, d\mu_n(\omega) \\ &\leq \quad d(x,y) \cdot \sum_{n\in\mathbb{Z}} \sum_{i\in I} \int_{\Omega_n} \left[\Psi(i,\omega,x) + \Psi(i,\omega,y)\right] d\mu_n(\omega) \\ &\qquad + \frac{18}{\varepsilon} \cdot \max\{d(x,X), d(y,X)\} \sum_{n\in\mathbb{Z}} \sum_{i\in I} \int_{\Omega_n} |\Psi(i,\omega,x) - \Psi(i,\omega,y)|\, d\mu_n(\omega) \\ &= \quad 2d(x,y) + \frac{18}{\varepsilon} \cdot \max\{d(x,X), d(y,X)\} \sum_{n\in\mathbb{Z}} \sum_{i\in I} \int_{\Omega_n} |\Psi(i,\omega,x) - \Psi(i,\omega,y)|\, d\mu_n(\omega) \end{aligned}$$

It is therefore enough to show that:

$$\sum_{n\in\mathbb{Z}} \sum_{i\in I} \int_{\Omega_n} |\Psi(i,\omega,x) - \Psi(i,\omega,y)|\, d\mu_n(\omega) \leq \frac{C'}{\delta} \cdot \frac{d(x,y)}{\max\{d(x,X), d(y,X)\}}, \tag{8}$$

when the decompositions are either (ε,δ)-padded or (ε,δ)-thick and

$$\sum_{n\in\mathbb{Z}} \sum_{i\in I} \int_{\Omega_n} |\Psi(i,\omega,x) - \Psi(i,\omega,y)|\, d\mu_n(\omega) \leq C'\left(1 + \frac{\varepsilon}{\delta}\right) \cdot \frac{d(x,y)}{\max\{d(x,X), d(y,X)\}}, \tag{9}$$

when the decompositions are (ε,δ)-separating. Here C' is a universal constant.

We may assume that $C' > 4$, in which case inequality (8) (resp. inequality (9)) holds trivially when $d(x,y) \geq d(\{x,y\},X)$. Indeed, in this case $d(x,X) \leq d(x,y) + d(y,X) \leq 2d(x,y)$ and analogously $d(y,X) \leq 2d(x,y)$. Hence the right-hand side of (8) (resp. (9)) is greater than 2 while the left-hand side of (8) (resp. (9)) is at most 2 since by construction $\sum_{n\in\mathbb{Z}} \sum_{i\in I} \int_{\Omega_n} \Psi(i,\omega,z) d\mu_n(\omega) = 1$ for every $z \in Y$.

Our goal is therefore to prove (8) (resp. (9)) under the assumption $d(x,y) < d(\{x,y\},X)$. We may also assume without loss of generality that $d(x,X) \geq d(y,X)$. Now,

$$\begin{aligned}
&\sum_{n\in\mathbb{Z}}\sum_{i\in I}\int_{\Omega_n} |\Psi(i,\omega,x) - \Psi(i,\omega,y)|\, d\mu_n(\omega) \\
=&\ \sum_{n\in\mathbb{Z}}\int_{\Omega_n}\sum_{i\in I}\left|\frac{\theta^n_\omega(x)\varphi_n(x)\mathbf{1}_{\Gamma^i_n(\omega)}(x)S(y) - \theta^n_\omega(y)\varphi_n(y)\mathbf{1}_{\Gamma^i_n(\omega)}(y)S(x)}{S(x)S(y)}\right| d\mu_n(\omega) \\
\leq&\ \sum_{n\in\mathbb{Z}}\int_{\Omega_n}\sum_{i\in I}\frac{\left|\theta^n_\omega(x)\varphi_n(x)\mathbf{1}_{\Gamma^i_n(\omega)}(x) - \theta^n_\omega(y)\varphi_n(y)\mathbf{1}_{\Gamma^i_n(\omega)}(y)\right|}{S(x)}\, d\mu_n(\omega) + \\
&\left(\sum_{n\in\mathbb{Z}}\int_{\Omega_n}\theta^n_\omega(y)\varphi_n(y)\, d\mu_n(\omega)\right)\frac{|S(x)-S(y)|}{S(x)S(y)} \\
\leq&\ \sum_{n\in\mathbb{Z}}\int_{\Omega_n}\sum_{i\in I}\frac{\left|\theta^n_\omega(x)\varphi_n(x)\mathbf{1}_{\Gamma^i_n(\omega)}(x) - \theta^n_\omega(y)\varphi_n(y)\mathbf{1}_{\Gamma^i_n(\omega)}(y)\right|}{S(x)}\, d\mu_n(\omega) + \\
&\sum_{n\in\mathbb{Z}}\int_{\Omega_n}\theta^n_\omega(y)\varphi_n(y)\, d\mu_n(\omega)\sum_{k\in\mathbb{Z}}\int_{\Omega_k}\sum_{i\in I}\frac{\left|\theta^n_\tau(x)\varphi_k(x)\mathbf{1}_{\Gamma^i_k(\tau)}(x) - \theta^n_\tau(y)\varphi_k(y)\mathbf{1}_{\Gamma^i_k(\tau)}(y)\right|}{S(x)S(y)}\, d\mu_k(\tau) \\
=&\ \frac{2}{S(x)}\sum_{n\in\mathbb{Z}}\int_{\Omega_n}\sum_{i\in I}\left|\theta^n_\omega(x)\varphi_n(x)\mathbf{1}_{\Gamma^i_n(\omega)}(x) - \theta^n_\omega(y)\varphi_n(y)\mathbf{1}_{\Gamma^i_n(\omega)}(y)\right|\, d\mu_n(\omega). \qquad (10)
\end{aligned}$$

We now deal with each of the particular cases in the statement of the theorem:

Case 1. The stochastic decomposition $(\Omega_n, \mu_n, \{\Gamma^i_n(\cdot), \gamma^i_n(\cdot)\}_{i\in I})$ is (ε,δ)-thick and 2^n-bounded for all $n \in \mathbb{Z}$. In this case we take $\theta^n_\omega(x) = \pi^n_\omega(x)$, where for every $\omega \in \Omega_n$ and $x \in Y$,

$$\pi^n_\omega(x) = \sum_{i\in I}\min\left\{d\left(x, Y\setminus\Gamma^i_n(\omega)\right), 2^n\right\}. \qquad (11)$$

Observe that since $\{\Gamma^i_n(\omega)\}_{i\in I}$ is a partition of Y, the above sum consists of only *one* element.

Let n_0 be an integer such that $\frac{d(x,X)}{\varepsilon 2^{n_0-3}} \in [1,2]$. Then

$$\begin{aligned}
S(x) &\geq \sum_{n:\ d(x,X)\leq\varepsilon 2^n}\varphi_n(x)\int_{\Omega_n}\pi^n_\omega(x)\, d\mu_n(\omega) \\
&\geq \sum_{n:\ d(x,X)\leq\varepsilon 2^n}\delta 2^n\varphi_n(x) \geq \delta 2^{n_0}\varphi\left(\frac{d(x,X)}{\varepsilon 2^{n_0-1}}\right) \geq \delta 2^{n_0} \geq \frac{\delta}{\varepsilon}\, d(x,X). \qquad (12)
\end{aligned}$$

We now estimate the numerator in (10). Fix $n \in \mathbb{Z}$ and $\omega \in \Omega_n$. Assume first of all that $x, y \in \Gamma^j_n(\omega)$ for some $j \in I$. Hence

$$\begin{aligned}
&\sum_{i\in I}\left|\pi^n_\omega(x)\varphi_n(x)\mathbf{1}_{\Gamma^i_n(\omega)}(x) - \pi^n_\omega(y)\varphi_n(y)\mathbf{1}_{\Gamma^i_n(\omega)}(y)\right| = \left|\pi^n_\omega(x)\varphi\left(\frac{d(x,X)}{\varepsilon 2^{n-3}}\right) - \pi^n_\omega(y)\varphi\left(\frac{d(y,X)}{\varepsilon 2^{n-3}}\right)\right| \\
\leq&\ \varphi\left(\frac{d(x,X)}{\varepsilon 2^{n-3}}\right)|\pi^n_\omega(x) - \pi^n_\omega(y)| + \pi^n_\omega(y)\left|\varphi\left(\frac{d(x,X)}{\varepsilon 2^{n-3}}\right) - \varphi\left(\frac{d(y,X)}{\varepsilon 2^{n-3}}\right)\right| \\
\leq&\ d(x,y) + \pi^n_\omega(y)\frac{16 d(x,y)}{\varepsilon 2^n} \leq d(x,y) + 2^n\cdot\frac{16 d(x,y)}{\varepsilon 2^n} \leq \frac{17}{\varepsilon}\, d(x,y). \qquad (13)
\end{aligned}$$

If, on the other hand, there are distinct $i, j \in I$ such that $x \in \Gamma_n^i(\omega)$ and $y \in \Gamma_n^j(\omega)$ then clearly $\pi_\omega^n(x), \pi_\omega^n(y) \le d(x,y)$, so:

$$\begin{aligned}\sum_{\ell\in I}\left|\pi_\omega^n(x)\varphi_n(x)\mathbf{1}_{\Gamma_n^\ell(\omega)}(x)-\pi_\omega^n(y)\varphi_n(y)\mathbf{1}_{\Gamma_n^\ell(\omega)}(y)\right| &= \pi_\omega^n(x)\varphi\left(\frac{d(x,X)}{\varepsilon 2^{n-1}}\right)+\pi_\omega^n(y)\varphi\left(\frac{d(y,X)}{\varepsilon 2^{n-1}}\right)\\ &\le \pi_\omega^n(x)+\pi_\omega^n(y)\le 2\,d(x,y). \qquad (14)\end{aligned}$$

Plugging the estimates (12), (13), (14) into (10) we obtain

$$\begin{aligned}\sum_{n\in\mathbb{Z}}\sum_{i\in I}\int_{\Omega_n}|\Psi(i,\omega,x)-\Psi(i,\omega,y)|\,d\mu_n(\omega) &\le \frac{2\varepsilon}{\delta\,d(x,X)}\sum_{n:\ \{x,y\}\cap\mathrm{supp}(\varphi_n)\neq\emptyset}\int_{\Omega_n}\frac{17}{\varepsilon}\,d(x,y)d\mu_n\\ &\le \frac{340\,d(x,y)}{\delta\,d(x,X)}, \qquad (15)\end{aligned}$$

where we have used the fact that for every $z\in Y$, $|\{n:\ z\in\mathrm{supp}(\varphi_n)\}|\le 5$.

Case 2. The stochastic decomposition $(\Omega_n,\mu_n,\{\Gamma_n^i(\cdot),\gamma_n^i(\cdot)\}_{i\in I})$ is (ε,δ)-padded and 2^n-bounded for all $n\in\mathbb{Z}$. In this case let $g:[0,\infty)\to[0,\infty)$ be given by:

$$g(x)=\begin{cases}1 & x\ge 2\\ x-1 & 1\le x\le 2\\ 0 & 0\le x\le 1,\end{cases}$$

and define $\theta_\omega^n(x)=g\left(\frac{\pi_\omega^n(x)}{\varepsilon 2^{n-1}}\right)$, where $\pi_\omega^n(x)$ is as in (11). Let n_0 be as in the proof of (12). Then:

$$\begin{aligned}S(x) &\ge \sum_{n:\ d(x,X)\le\varepsilon 2^n}\varphi_n(x)\int_{\{\omega\in\Omega_n:\ \pi_\omega^n(x)\ge\varepsilon 2^n\}} g\left(\frac{\pi_\omega^n(x)}{\varepsilon 2^{n-1}}\right)(x)\,d\mu_n(\omega)\\ &\ge \varphi\left(\frac{d(x,X)}{\varepsilon 2^{n_0-3}}\right)\mu_{n_0}\left(\bigcup_{i\in I}\left\{\omega\in\Omega_{n_0}:\ d\left(x,X\setminus\Gamma_n^i(\omega)\right)\ge\varepsilon 2^{n_0}\right\}\right)\ge\delta. \qquad (16)\end{aligned}$$

Fix $n\in\mathbb{Z}$ and $\omega\in\Omega_n$. Assume $\varphi_n(x)\cdot\varphi_n(y)>0$. In this case $\{d(x,X),d(y,X)\}\cap[\varepsilon 2^{n-4},\varepsilon 2^{n-1}]\neq\emptyset$, so that in particular $d(y,X)\le\varepsilon 2^{n-1}$. We are assuming that $d(x,y)<d(y,X)$, so $d(x,X)\le d(x,y)+d(y,X)\le 2d(y,X)\le\varepsilon 2^n$. If $x,y\in\Gamma_n^j(\omega)$ for some $j\in I$ then since g is Lipschitz with constant 1 and bounded by 1, the same reasoning as in (13) gives:

$$\sum_{i\in I}\left|\theta_\omega^n(x)\varphi_n(x)\mathbf{1}_{\Gamma_n^i(\omega)}(x)-\theta_\omega^n(y)\varphi_n(y)\mathbf{1}_{\Gamma_n^i(\omega)}(y)\right|\le\frac{6d(x,y)}{\varepsilon 2^n}\le\frac{6d(x,y)}{d(x,X)}. \qquad (17)$$

On the other hand, if there distinct $i,j\in I$ such that $x\in\Gamma_n^i(\omega)$ and $y\in\Gamma_n^j(\omega)$ then using the fact that $\pi_\omega^n(x),\pi_\omega^n(y)\le d(x,y)$ we get :

$$\begin{aligned}\sum_{\ell\in I}\left|\pi_\omega^n(x)\varphi_n(x)\mathbf{1}_{\Gamma_n^\ell(\omega)}(x)-\pi_\omega^n(y)\varphi_n(y)\mathbf{1}_{\Gamma_n^\ell(\omega)}(y)\right| &\le g\left(\frac{\pi_\omega^n(x)}{\varepsilon 2^{n-1}}\right)+g\left(\frac{\pi_\omega^n(y)}{\varepsilon 2^{n-1}}\right)\\ &\le \frac{\pi_\omega^n(x)}{\varepsilon 2^{n-1}}+\frac{\pi_\omega^n(y)}{\varepsilon 2^{n-1}}\le\frac{4d(x,y)}{d(x,X)}, \qquad (18)\end{aligned}$$

and we conclude as in (15).

Case 3. The stochastic decomposition $(\Omega_n, \mu_n, \{\Gamma_n^i(\cdot), \gamma_n^i(\cdot)\}_{i\in I})$ is (ε, δ)-separating 2^n-bounded for all $n \in \mathbb{Z}$. This case is simpler: we take $\theta_\omega^n(x) \equiv 1$. Arguing as in (12) we get that $S(x) \geq 1$. Fix $n \in \mathbb{Z}$ and $\omega \in \Omega_n$ and assume as before that $\varphi_n(x) \cdot \varphi_n(y) > 0$. Observe that

$$\begin{aligned}\sum_{i\in I} \left|\varphi_n(x)\mathbf{1}_{\Gamma_n^i(\omega)}(x) - \varphi_n(y)\mathbf{1}_{\Gamma_n^i(\omega)}(y)\right| &\leq |\varphi_n(x) - \varphi_n(y)| \\ &\quad + \frac{\varphi_n(x) + \varphi_n(y)}{2} \sum_{i\in I} \left|\mathbf{1}_{\Gamma_n^i(\omega)}(x) - \mathbf{1}_{\Gamma_n^i(\omega)}(y)\right| \\ &\leq \frac{8d(x,y)}{\varepsilon 2^n} + \sum_{i\in I} \left|\mathbf{1}_{\Gamma_n^i(\omega)}(x) - \mathbf{1}_{\Gamma_n^i(\omega)}(y)\right| \\ &\leq \frac{8d(x,y)}{d(x,X)} + \sum_{i\in I} \left|\mathbf{1}_{\Gamma_n^i(\omega)}(x) - \mathbf{1}_{\Gamma_n^i(\omega)}(y)\right|. \qquad (19)\end{aligned}$$

Plugging this estimate into (10) we get

$$\begin{aligned}\sum_{n\in\mathbb{Z}} \sum_{i\in I} \int_{\Omega_n} |\Psi(i,\omega,x) - \Psi(i,\omega,y)|\, d\mu_n(\omega) &\leq \frac{8d(x,y)}{d(x,X)} \cdot |\{n:\ \{x,y\} \cap \mathrm{supp}(\varphi_n) \neq \emptyset\}| \\ &\quad + \sum_{n:\ \{x,y\}\cap\mathrm{supp}(\varphi_n)\neq\emptyset} \int_{\Omega_n} \sum_{i\in I} \left|\mathbf{1}_{\Gamma_n^i(\omega)}(x) - \mathbf{1}_{\Gamma_n^i(\omega)}(y)\right| d\mu_n \\ &\leq \frac{80\, d(x,y)}{d(x,X)} + \sum_{n:\ \{x,y\}\cap\mathrm{supp}(\varphi_n)\neq\emptyset} \frac{2d(x,y)}{\delta 2^n} \\ &\leq \frac{80\, d(x,y)}{d(x,X)} + \frac{20\varepsilon d(x,y)}{\delta d(x,X)},\end{aligned}$$

where we have used the (ε, δ)-separating condition and the fact that in the above sum, $d(x,X) \leq \varepsilon 2^n$.

□

We proceed to construct gentle partitions of unity when X is finite. The proof is analogous to the proof of Theorem 4.1, but there are several subtle differences, so we deal with this important case separately. The analysis uses ideas from [8, 9], namely that a certain sum of logarithms collapses in the analysis of the decomposition of Lemma 3.17. This allows the smoothing function φ to have larger support, which will be the key to achieving an improved result.

Theorem 4.3. *There exists a universal constant $C > 0$ such that for all $m \in \mathbb{N}$, any metric space (Y,d) and any m-point subset $X \subseteq Y$, Y admits a $C\frac{\log m}{\log\log m}$-gentle partition of unity with respect to X.*

Proof. Let $M > 2$ be an integer which will be determined later. Let $\varphi : \mathbb{R} \to \mathbb{R}_+$ be a 5-Lipschitz map with $\mathrm{supp}(\varphi) \subseteq [\frac{1}{5}, 5^{M+1}]$ and $\varphi \equiv 1$ on $[1, 5^M]$. By Theorem 3.17, with σ the counting measure on X, for every $n \in \mathbb{Z}$ there exists a stochastic decomposition of Y with respect to X, $(\Omega_n, \mu_n, \{\Gamma_n^i(\cdot), \gamma_n^i(\cdot)\}_{i\in I})$, which is 5^n-bounded and for every $x, y \in Y$ satisfying $d(\{x,y\}, X) \leq 5^{n-2}$,

$$\int_{\Omega_n} \sum_{i\in I} \left|\mathbf{1}_{\Gamma_n^i(\omega)}(x) - \mathbf{1}_{\Gamma_n^i(\omega)}(y)\right| d\mu_n(\omega) \leq \frac{2d(x,y)}{5^n} \left[1 + \log\left(\frac{|B_X(x, 5^{n+1})|}{|B_X(x, 5^n)|}\right)\right]. \qquad (20)$$

Define $\varphi_n(x) = \varphi\left(\frac{d(x,X)}{5^{n-M-4}}\right)$, and let (Ω,μ), $\Psi(i,\omega,x)$, $\gamma(i,\omega)$, $S(x)$ be as in the proof of Theorem 4.1, with $\theta^n_\omega(x) \equiv 1$. Our goal is to show that for every $x, y \in Y$,

$$\sum_{n\in\mathbb{Z}}\sum_{i\in I}\int_{\Omega_n} d(\gamma(i,\omega),x)\cdot|\Psi(i,\omega,x)-\Psi(i,\omega,y)|\,d\mu(\omega) \le O\left(\frac{\log m}{\log\log m}\right)\cdot d(x,y).$$

In what follows we fix $x, y \in Y$ and we may assume without loss of generality that $d(x,X) \ge d(y,X)$. As in the proof of Claim 4.2, if $|\Psi(i,\omega,x)-\Psi(i,\omega,y)| > 0$ then $d(\gamma(i,\omega),x) \le 5^n + d(x,y)$. It is therefore enough to show that

$$\sum_{n\in\mathbb{Z}} 5^n \sum_{i\in I}\int_{\Omega_n} |\Psi(i,\omega,x)-\Psi(i,\omega,y)|\,d\mu(\omega) \le O\left(\frac{\log m}{\log\log m}\right)\cdot d(x,y).$$

Since $\varphi_n(x) = 1$ for at least $M-1$ values of n, $S(x) = \sum_{n\in\mathbb{Z}}\varphi_n(x) \ge M-1 \ge \frac{M}{2}$. Hence, arguing as in (10) we get that

$$\begin{aligned}
&\sum_{n\in\mathbb{Z}} 5^n \sum_{i\in I}\int_{\Omega_n} |\Psi(i,\omega,x)-\Psi(i,\omega,y)|\,d\mu(\omega)\\
&\le \frac{2}{M}\sum_{n\in\mathbb{Z}} 5^n \int_{\Omega_n}\sum_{i\in I}\left|\varphi_n(x)\mathbf{1}_{\Gamma^i_n(\omega)}(x) - \varphi_n(y)\mathbf{1}_{\Gamma^i_n(\omega)}(y)\right|\,d\mu_n(\omega) +\\
&\quad \frac{4}{M^2}\left(\sum_{n\in\mathbb{Z}} 5^n\varphi_n(y)\right)|S(x)-S(y)|.
\end{aligned} \tag{21}$$

Let j be the maximal integer for which $\varphi_j(y) > 0$. Then $5^j \le 5^{M+5} d(y,X)$ and

$$\sum_{n\in\mathbb{Z}} 5^n\varphi_n(y) \le \sum_{n\le j} 5^n \le 5^{j+1} \le 5^{M+6} d(y,X).$$

Analogously, $\sum_{n\in\mathbb{Z}} 5^n\varphi_n(x) \le 5^{M+6} d(x,X)$. Since in $|S(x)-S(y)|$ there are at most $4M$ non-zero summands, each of which is bounded by 2, we get the following estimate:

$$\begin{aligned}
\sum_{n\in\mathbb{Z}} 5^n \sum_{i\in I}\int_{\Omega_n} |\Psi(i,\omega,x)-\Psi(i,\omega,y)|\,d\mu(\omega) &\le \frac{4}{M}\sum_{n:\ \{x,y\}\cap\operatorname{supp}(\varphi_n)\neq\emptyset} 5^n + \frac{5^{M+9}d(y,X)}{M}\\
&\le \frac{5^{M+7}}{M}[d(x,X)+d(y,X)] + \frac{5^{M+9}d(y,X)}{M}\\
&\le \frac{5^{M+10}}{2M}[d(y,X)+d(x,y)].
\end{aligned}$$

Corollary 4.4. *If $d(y,X) \le d(x,y)$ then:*

$$\sum_{n\in\mathbb{Z}} 5^n \sum_{i\in I}\int_{\Omega_n} |\Psi(i,\omega,x)-\Psi(i,\omega,y)|\,d\mu(\omega) \le \frac{5^{M+10}}{M}d(x,y).$$

We now deal with the case $d(x,X) \ge d(y,X) > d(x,y)$. The reasoning above shows that the estimate (21) can be written as

$$\begin{aligned}
&\sum_{n\in\mathbb{Z}} 5^n \sum_{i\in I} \int_{\Omega_n} |\Psi(i,\omega,x)-\Psi(i,\omega,y)|\, d\mu(\omega)\\
&\le \frac{2}{M}\sum_{n\in\mathbb{Z}} 5^n \int_{\Omega_n} \sum_{i\in I} \left|\varphi_n(x)\mathbf{1}_{\Gamma_n^i(\omega)}(x)-\varphi_n(y)\mathbf{1}_{\Gamma_n^i(\omega)}(y)\right| d\mu_n(\omega) +\\
&\quad \frac{5^{M+7}}{M^2}\cdot d(x,X)\cdot |S(x)-S(y)|. \qquad (22)
\end{aligned}$$

For the sake of simplicity, denote

$$E_{x,y}=\{n\in\mathbb{Z}:\ \varphi_n(x)=\varphi_n(y)=1\} \quad\text{and}\quad F_{x,y}=\{n\in\mathbb{Z}:\ \varphi_n(x)+\varphi_n(y)>0\}\setminus E_{x,y}.$$

By the definition of φ_n, if $n\in F_{x,y}$ then 5^n must be in

$$[5^3 d(x,X), 5^{M+5}d(x,X)]\cup[5^3 d(y,X),5^{M+5}d(y,X)]$$

but not in

$$[5^4 d(x,X),5^{M+4}d(x,X)]\cap[5^4 d(y,X),5^{M+4}d(y,X)].$$

Since $d(y,X)\le d(x,X)\le d(y,X)+d(x,y)\le 2d(y,Y)$, it follows that n can take at most 10 values. We have shown that $|F_{x,y}|\le 10$. Hence, using the fact that φ is 5-Lipschitz,

$$|S(x)-S(y)|\le \sum_{n\in F_{x,y}}|\varphi_n(x)-\varphi_n(y)|\le 10\frac{d(x,y)}{5^{n-M-5}}\le 5^{M+7}\cdot\frac{d(x,y)}{d(x,X)}. \qquad (23)$$

Fix $n\in\mathbb{Z}$ and $\omega\in\Omega_n$ and argue as in (19) to get that

$$\sum_{i\in I}\left|\varphi_n(x)\mathbf{1}_{\Gamma_n^i(\omega)}(x)-\varphi_n(y)\mathbf{1}_{\Gamma_n^i(\omega)}(y)\right|\le \frac{d(x,y)}{5^{n-M-5}}\cdot\mathbf{1}_{F_{x,y}}(n)+\sum_{i\in I}\left|\mathbf{1}_{\Gamma_n^i(\omega)}(x)-\mathbf{1}_{\Gamma_n^i(\omega)}(y)\right|.$$

Since for $n\in E_{x,y}\cup F_{x,y}$, $d(x,X)\le 5^{n-2}$, we may use (20) to get that

$$\begin{aligned}
&\int_{\Omega_n}\sum_{i\in I}\left|\varphi_n(x)\mathbf{1}_{\Gamma_n^i(\omega)}(x)-\varphi_n(y)\mathbf{1}_{\Gamma_n^i(\omega)}(y)\right| d\mu_n(\omega)\\
&\le \frac{d(x,y)}{5^{n-M-5}}\cdot\mathbf{1}_{F_{x,y}}(n)+\frac{2d(x,y)}{5^n}\left[1+\log\left(\frac{|B_X(x,5^{n+1})|}{|B_X(x,5^n)|}\right)\right] \qquad (24)
\end{aligned}$$

Let j,J be the minimal and maximal integers i, respectively, for which $\varphi_i(x)+\varphi_i(y)>0$. Observe that in this case $J-j\le 4M$ and $5^{j-3}\ge\frac{d(x,X)}{2}$. Hence

$$\begin{aligned}
&\sum_{n\in\mathbb{Z}}5^n\int_{\Omega_n}\sum_{i\in I}\left|\varphi_n(x)\mathbf{1}_{\Gamma_n^i(\omega)}(x)-\varphi_n(y)\mathbf{1}_{\Gamma_n^i(\omega)}(y)\right| d\mu_n(\omega)\\
&\le \sum_{n\in F_{x,y}}5^n\cdot\frac{d(x,y)}{5^{n-M-5}}+\sum_{n\in E_{x,y}\cup F_{x,y}}5^n\int_{\Omega_n}\sum_{i\in I}\left|\mathbf{1}_{\Gamma_n^i(\omega)}(x)-\mathbf{1}_{\Gamma_n^i(\omega)}(y)\right| d\mu_n\\
&\le 10\cdot 5^{M+5}d(x,y)+\sum_{n=j}^{J}2d(x,y)\left[1+\log\left(\frac{|B_X(x,5^{n+1})|}{|B_X(x,5^n)|}\right)\right]\\
&\le 5^{M+7}d(x,y)+2d(x,y)\left[4M+\log\left(\frac{|B_X(x,5^{J+1})|}{|B_X(x,5^j)|}\right)\right]\\
&\le \left(5^{M+7}+8M+2\log m\right)\cdot d(x,y), \qquad (25)
\end{aligned}$$

Plugging (23) and (25) into (22) we arrive at the following corollary:

Corollary 4.5. *If $d(x,y) < d(y,X)$ then*

$$\sum_{n\in\mathbb{Z}} 5^n \sum_{i\in I} \int_{\Omega_n} |\Psi(i,\omega,x) - \Psi(i,\omega,y)|\, d\mu(\omega) \le O\left(\frac{5^{2M}}{M^2} + \frac{\log m}{M}\right) \cdot d(x,y).$$

Using Corollary 4.4 and Corollary 4.5 with $M \approx \log\log m$ yields the required result. $\square$

5 Extension theorems

The following theorem is a direct consequence of the results of Section 2, Section 3 and Section 4.

Theorem 5.1. *There exists a universal constant $C > 0$ such that*

1. *For every metric space X, $ae(X) \le C \log \lambda(X)$.*
2. *For every $r > 0$ and any K_r-excluded metric space X, $ae(X) \le C \cdot r^2$.*
3. *Let M be a two dimensional Riemannian manifold with genus g and $X \subseteq M$. Then $ae(X) \le C \cdot (g+1)$.*
4. *If X is an n-point metric space then $ae(X) \le C \frac{\log n}{\log\log n}$.*
5. *For every integer d and every Banach space Z, $e(\ell_2^d, Z) \le C\sqrt{d}$.*

We also have the following extension result for neighborhoods of subsets of negatively curved manifolds.

Proposition 5.2. *Fix $r > 0$ and let M be an n-dimensional Riemannian manifold satisfying* $\mathrm{Ricci}(g) \ge -(n-1)rg$*, where g is the Riemannian metric on M. Then for any subset $X \subseteq M$, any metric space $Y \supseteq X$ such that X is Δ-dense in Y, and any Banach space Z, $e(X,Y,Z) \le C \cdot n\left(1 + \sqrt{r\Delta}\right)$. Here C is a universal constant.*

Proof. Observe that in this case in the proof of 4.1 we only require the existence of 2^n-padded decomposition with $2^n \le \Delta$. Hence the required result follows from an application of Corollary 3.13 and Lemma 3.8. $\square$

Let us recall that for a family of finite graphs $\mathcal{F}$, we claim that $ae(\langle\mathcal{F}\rangle) \le K_{\mathcal{F}}$ if and only if $\langle\mathcal{F}\rangle$ does not contain all finite metrics. We now give the simple proof which is based on a deep result of Robertson and Seymour [32].

Proof of Corollary 1.8. For a family of graphs $\mathcal{F}$, let $mc(\mathcal{F})$ denote its closure under taking minors, i.e. the maximal minor-closed family containing $\mathcal{F}$. Robertson and Seymour proved that if $mc(\mathcal{F})$ is non-trivial, i.e. does not contain *all* finite graphs, then there is some finite list of graphs $H_1, \cdots, H_k$ such that $G \in mc(\mathcal{F})$ if and only if G does not contain any H_i as a minor.

Observe that contraction/deletion of an edge corresponds to weighting by $0/\infty$, respectively. It follows that $\langle\mathcal{F}\rangle = \langle mc(\mathcal{F})\rangle$, thus if $\langle\mathcal{F}\rangle$ does not contain all finite metrics, then certainly $mc(\mathcal{F})$ does not contain all finite graphs, hence if $X \in \langle\mathcal{F}\rangle$, it must be supported on some graph G which excludes a K_r minor with $r = \max_i |H_i|$, and in this case part (2) of Theorem 5.1 applies. $\square$

We are now in position to prove Theorem 1.12 which was stated in the introduction.

Proof of Theorem 1.12. We begin with the case $p = 2$. Let X be an n-point subset of L_2, Z a Banach space and $f : X \to Z$ a Lipschitz function. Let H be the linear span of X and let Q be the orthogonal projection from L_2 onto H. By the proof of the Johnson-Lindenstrauss dimension reduction lemma [17] there is a probability space (Ω, P) such that for every $\omega \in \Omega$ there is a rank d linear operator $T_\omega : H \to H$ such that for every $x \in H$, $P\left(\left|\|T_\omega(x)\|_2 - \|x\|_2\right| \geq \frac{1}{2}\right) \leq 2e^{-cd}$, where c is a universal constant. We can therefore find for $d \approx \frac{4}{c}\log n$ a subset $A \subset \Omega$ with $P(A) \geq \frac{1}{2}$ such that for every $x, y \in X$ and $\omega \in A$, $\|T_\omega(x) - T_\omega(y)\|_2 \geq \frac{1}{2}\|x-y\|_2$. The function $g_\omega = f \circ (T_\omega|_{T_\omega(X)})^{-1}$ is Lipschitz on $T_\omega(X)$ with constant $2\|f\|_{\mathrm{Lip}}$. By Theorem 5.1, there is a function $\tilde{g}_\omega : T_\omega(H) \to Z$ such that $\tilde{g}_\omega|_{T_\omega(X)} = g_\omega$ and $\|\tilde{g}_\omega\|_{\mathrm{Lip}} \leq \frac{4C}{\sqrt{c}}\sqrt{\log n} \cdot \|f\|_{\mathrm{Lip}}$. Define $\tilde{f} : L_2 \to Z$ by:

$$\tilde{f}(x) = \frac{1}{P(A)} \int_A \tilde{g}_\omega(T_\omega Q x) dP(\omega).$$

Clearly $\tilde{f}$ is an extension of f and since $P(A) \geq \frac{1}{2}$, for every $x, y \in L_p$,

$$\begin{aligned}
\|\tilde{f}(x) - \tilde{f}(y)\|_Z &\leq \frac{1}{P(A)} \int_A \|\tilde{g}_\omega(T_\omega Q x) - \tilde{g}_\omega(T_\omega Q y)\|_Z \, dP(\omega) \\
&\leq O\left(\sqrt{\log n}\right) \|f\|_{\mathrm{Lip}} \cdot \int_\Omega \|T_\omega(Qx - Qy)\|_2 \, dP(\omega) \\
&= O\left(\sqrt{\log n}\right) \|f\|_{\mathrm{Lip}} \cdot \|x-y\|_2,
\end{aligned}$$

where we have used the fact that the concentration inequality for $\|T_\omega z\|_2$, $z \in H$, implies that $\int_\Omega \|T_\omega z\| dP(\omega) = O(\|z\|_2)$.

We pass to general $1 < p \leq 2$ via a method due to Marcus and Pisier [27]. In [27] they show that for every $0 < p \leq 2$ there is a probability space (Ω', P') such that for every $\omega \in \Omega'$ there is a linear operator $S_\omega : L_p \to L_2$ such that for every $x \in L_p \setminus \{0\}$ the random variable $X = \frac{\|S_\omega(x)\|_2}{\|x\|_p}$ satisfies for every $a \in \mathbb{R}$, $\mathbb{E}e^{-aX^2} = e^{-a^{p/2}}$. Let T be an n-point subset of L_p, Z a Banach space and $f : T \to Z$ a Lipschitz function. A standard application of Markov's inequality shows that there is constant c_p and a subset $A' \subset \Omega$ with $P'(A') \geq \frac{1}{2}$ such that for every $x, y \in X$ and $\omega \in A'$,

$$\|x-y\|_p \leq c_p (\log n)^{\frac{1}{p}-\frac{1}{2}} \|S_\omega(x) - S_\omega(y)\|_2.$$

For every $\omega \in A'$ the function $g_\omega = f \circ (S_\omega|_{S_\omega(X)})^{-1}$ is $O\left((\log n)^{\frac{1}{p}-\frac{1}{2}}\right)\|f\|_{\mathrm{Lip}}$ Lipschitz. By the above reasoning for the case $p = 2$, g_ω can be extended to a function $\tilde{g}_\omega$ defined on all of L_2 which is Lipschitz with constant $O\left((\log n)^{1/p}\right)\|f\|_{\mathrm{Lip}}$. Define $\tilde{f} : L_p \to L_2$ by:

$$\tilde{f}(x) = \frac{1}{P'(A')} \int_{A'} \tilde{g}_\omega(S_\omega(x)) \, dP'(\omega).$$

Clearly $\tilde{f}$ is an extension of f and since $P'(A') \geq \frac{1}{2}$, for every $x, y \in L_p$,

$$\begin{aligned}
\|\tilde{f}(x) - \tilde{f}(y)\|_Z &\leq \frac{1}{P'(A')} \int_{A'} \|\tilde{g}_\omega(S_\omega(x)) - \tilde{g}_\omega(S_\omega(y))\|_Z \, dP'(\omega) \\
&\leq O\left((\log n)^{1/p}\right) \|f\|_{\mathrm{Lip}} \cdot \int_\Omega \|S_\omega(x-y)\|_2 \, dP'(\omega) \\
&= O\left((\log n)^{1/p}\right) \|f\|_{\mathrm{Lip}} \cdot \|x-y\|_p \cdot \mathbb{E}X,
\end{aligned}$$

and we conclude since for $p>1$, $\mathbb{E}X = C_p < \infty$.

$\square$

Remark 5.3. As stated in the introduction, it was asked in [17] and [18] whether for every Banach space X, $\sup_n e_n(L_2, X) < \infty$. This is false since in [30] it was shown that for $2 < p < \infty$, $e(L_2, L_p) = \infty$. We end by reproducing the argument from [30] (which is based on ideas from [25]) in such a way that we get quantitative lower bounds on $e_n(L_2, \ell_p)$, $2 < p < \infty$. Since we are essentially repeating the proof from [30], our argument will be somewhat sketchy.

We claim that, for every integer n and every $2 < p < \infty$,

$$e_n(L_2, L_p) \geq \Omega\left[\left(\frac{\log n}{\log\log n}\right)^{\frac{p-2}{p^2}}\right].$$

Proof (sketch). Fix an integer m and set $\varepsilon = \frac{1}{n^{1/2-1/p}}$. Let N be an ε net in the unit ball of ℓ_2^{2m}, denoted B. By standard (crude) volume estimates $|N| \leq m^{2m}$. Consider the Mazur map $f: \ell_2^{2m} \to \ell_p^{2m}$ given by $f(x)_i = |x_i|^{2/p}\mathrm{sign}(x_i)$. From the numerical inequality $|a^{2/p}\mathrm{sign}(a) - b^{2/p}\mathrm{sign}(b)| \leq 2^{1-2/p}|a-b|^{2/p}$ it follows that for every $x, y \in \ell_2^{2m}$, $\|f(x) - f(y)\|_p \leq 2\|x-y\|_2^{2/p}$. Since the elements of N are ε separated, the restriction of f to N is Lipschitz with constant $2/\varepsilon^{1-2/p} = 2m^{(1-2/p)(1/2-1/p)}$. Assume that it is possible to extend $f|_N$ to a function $g: \ell_2^{2m} \to \ell_p^{2m}$ which is K Lipschitz. Since N is ε-dense in B, for every $x \in B$, $\|f(x) - g(x)\|_p \leq K\varepsilon + 2\varepsilon^{2/p}$. As in [30], by averaging g over all permutations and sign changes we arrive at a function h satisfying $h(a\mathbf{1}_A) = b\mathbf{1}_A$ for all scalars a and $A \subset \{1, \ldots, 2m\}$, where b depends only on the a and the cardinality of A. Additionally, h is K Lipschitz, and since f is invariant under permutations and changes of sign of the coordinates, for every $x \in B$, $\|h(x) - f(x)\|_p \leq K\varepsilon + 2\varepsilon^{2/p}$. Setting $x_k = \frac{1}{\sqrt{2m}}\mathbf{1}_{\{k,\ldots,k+m-1\}}$ it follows that $\|h(x_{m+1}) - h(x_1)\|_p^p = \sum_{k=1}^m \|h(x_{k+1}) - h(x_k)\|_p^p$, and since h is K-Lipschitz we get the estimate $\|h(x_{m+1}) - h(x_1)\|_p \leq \frac{K}{m^{1/2-1/p}} = K\varepsilon$. On the other hand,

$$\begin{aligned} 2K\varepsilon + 2\varepsilon^{2/p} &\geq \|h(x_{m+1}) - f(x_{m+1})\|_p + \|h(x_1) - f(x_1)\|_p \\ &\geq \|f(x_{m+1}) - f(x_1)\|_p - \|h(x_{n+1}) - h(x_1)\|_p \geq 1 - K\varepsilon. \end{aligned}$$

This implies that the ratio between K and the Lipschitz constant of f is at least $K/(2\varepsilon^{2/p-1}) = \Omega(\varepsilon^{-2/p}) = \Omega\left\{[(\log n)/(\log\log n)]^{(2-p)/p^2}\right\}$, where $n = m^{2m} \geq |N|$. $\square$

6 Appendix: Passing to arbitrary barycentric target spaces

In order to deal with barycentric metric spaces it is convenient to introduce the following variant of the notion of a gentle partition of unity. Let (Y, d) be a metric space, X a subspace of Y and $(\Omega, \mathcal{F}, \mu)$ a measure space. Given $K, L > 0$ we shall say that a function $\Psi: \Omega \times Y \to [0, \infty)$ is a (K, L)*-gentle partition of unity with respect to* X if the following conditions hold true:

1. For every $x \in Y \setminus X$ the function $\omega \mapsto \Psi(\omega, x)$ is measurable and $\int_\Omega \Psi(\omega, x)d\mu(\omega) = 1$.

2. There exists a Borel measurable function $\gamma: \Omega \to \overline{X}$ such that for every $x, y \in Y$,

$$\mathrm{diam}(\{x,y\} \cup \{\gamma(\omega) :\ \Psi(\omega,x) + \Psi(\omega,y) > 0\}) \leq K \cdot [d(x,y) + \max\{d(x,X), d(y,X)\}],$$

3. For every $x, y \in Y$, $x \neq y$,

$$\int_\Omega |\Psi(\omega, x) - \Psi(\omega, y)|\, d\mu(\omega) \leq L \cdot \frac{d(x,y)}{d(x,y) + \max\{d(x,X), d(y,X)\}}.$$

The following lemma is a variant of Lemma 2.1

Lemma 6.1. *Let (Y, d_Y) be a metric space and X a subspace of Y. Fix $K, L > 0$ and assume that Y admits a (K, L)-gentle partition of unity $\Psi : \Omega \times Y \to [0, \infty)$ with respect to X. Let (Z, d_Z) be a complete barycentric metric space and $\beta > \beta(Z)$. Then every Lipschitz function $f : X \to Z$ can be extended to a function $\tilde{f} : Y \to Z$ such that $\|\tilde{f}\|_{\text{Lip}} \leq \beta \max\{KL, 2K + 2\}$.*

Proof. As in the proof of Lemma 2.1, we may assume that X is closed. Let $c : \mathcal{M}_Z^{\text{bounded}} \to Z$ be a mapping satisfying the conditions of Definition 1.14 and let $\gamma : \Omega \to X$ be as in condition 2 above.

For every $x \in Y \setminus X$ define $\nu_x \in \mathcal{M}_Z^{\text{bounded}}$ by

$$\nu_x(A) = \int_{\gamma^{-1}(f^{-1}(A))} \Psi(\omega, x) d\mu(\omega).$$

In other words, ν_x is the pullback of the probability measure $\Psi(\cdot, x)d\mu$ under the mapping $f \circ \gamma$. For $x \in X$ we define $\nu_x = \delta_{f(x)}$. Finally, for $x \in Y$ set $\tilde{f}(x) = c(\nu_x)$. Clearly $\tilde{f}$ is an extension of f.

If $x, y \in Y$ then by (1),

$$d_Z(\tilde{f}(x), \tilde{f}(y)) \leq \beta \cdot \text{diam}(\text{supp}(\nu_x + \nu_y)) \cdot \|\nu_x - \nu_y\|_{TV}.$$

Now, if $x, y \in Y \setminus X$ then

$$\begin{aligned} \text{diam}(\text{supp}(\nu_x + \nu_y)) &\leq \|f\|_{\text{Lip}} \cdot \text{diam}(\{\gamma(\omega) : \ \Psi(\omega, x) + \Psi(\omega, y) > 0\}) \\ &\leq \|f\|_{\text{Lip}} \cdot K \cdot [d(x,y) + \max\{d(x,X), d(y,X)\}]. \end{aligned}$$

and

$$\|\nu_x - \nu_y\|_{TV} \leq \int_\Omega |\Psi(\omega, x) - \Psi(\omega, y)| d\mu(y) \leq L \cdot \frac{d(x,y)}{d(x,y) + \max\{d(x,X), d(y,X)\}}.$$

This implies that for $x, y \in Y \setminus X$,

$$d_Z(\tilde{f}(x), \tilde{f}(y)) \leq \beta KL \cdot \|f\|_{\text{Lip}}.$$

It remains to deal with the case $x \in X$ and $y \in Y \setminus X$. In this case set $\Gamma = \{\gamma(\omega) : \ \Psi(\omega, y) > 0\}$. Condition 2 (applied with $x = y$) implies that $\text{diam}(\{y\} \cup \Gamma) \leq K d(y, X)$. In particular, $d(x, \Gamma) \leq d(x,y) + d(y, \Gamma) \leq d(x,y) + K d(y,X) \leq (K+1) d(x,y)$. It follows that

$$\text{diam}(\text{supp}(\nu_x + \nu_y)) \leq \|f\|_{\text{Lip}} \cdot \text{diam}(\{x, y\} \cup \Gamma) \leq \|f\|_{\text{Lip}} (K+1) d(x,y).$$

On the other hand, $\|\nu_x - \nu_y\|_{TV} \leq 2$, and the required result follows. $\square$

The following theorem is a simple variant of Theorem 4.1, adapted to the case of (K, L) gentle partitions of unity.

Theorem 6.2. *For every $\varepsilon, \delta \in (0,1)$, every metric space (Y,d) and every subspace $X \subseteq Y$ the following assertions hold true:*

1. *If for every $n \in \mathbb{Z}$, Y admits an (ε,δ)-thick 2^n-bounded stochastic decomposition with respect to X, then Y also admits a $\left(\frac{36}{\varepsilon}, \frac{680}{\delta}\right)$-gentle partition of unity with respect to X.*

2. *If for every $n \in \mathbb{Z}$, Y admits an (ε,δ)-padded 2^n-bounded stochastic decomposition with respect to X, then Y also admits a $\left(\frac{36}{\varepsilon}, \frac{680}{\delta}\right)$-gentle partition of unity with respect to X.*

3. *If for every $n \in \mathbb{Z}$, Y admits an (ε,δ)-separating 2^n-bounded stochastic decomposition with respect to X, then Y also admits a $\left(\frac{36}{\varepsilon}, 160\left(1+\frac{\varepsilon}{\delta}\right)\right)$-gentle partition of unity with respect to X.*

Proof. Since the proof is a straightforward modification of the proof of Theorem 4.1, we will only sketch the necessary changes. As before, for every $n \in \mathbb{Z}$ let $(\Omega_n, \mu_n, \{\Gamma_n^i(\cdot), \gamma_n^i(\cdot)\}_{i\in I})$ be a 2^n-bounded stochastic decomposition of Y with respect to X. Let $\varphi : \mathbb{R} \to \mathbb{R}_+$ be any 2-Lipschitz map with $\mathrm{supp}(\varphi) \subset [\frac{1}{2}, 4]$ and $\varphi \equiv 1$ on $[1,2]$. Define $\varphi_n(x) = \varphi\left(\frac{d(x,X)}{\varepsilon 2^{n-3}}\right)$ and let (Ω, μ) be the disjoint union of $\{I \times \Omega_n\}_{n\in\mathbb{Z}}$ (where the measure on I is the counting measure). The partition of unity which we construct has the following form: For every $n \in \mathbb{Z}$, $\omega \in \Omega_n$, $i \in I$ and $x \in Y$ denote:

$$\Psi(i,\omega,x) = \frac{1}{S(x)}\, \theta_\omega^n(x)\varphi_n(x)\mathbf{1}_{\Gamma_n^i(\omega)}(x), \tag{26}$$

where for every $n \in \mathbb{Z}$ and $x \in Y$ the function $\omega \mapsto \theta_\omega^n(x) \in [0,\infty)$ is μ_n-integrable and

$$S(x) = \sum_{n\in\mathbb{Z}}\sum_{i\in I}\int_{\Omega_n} \theta_\omega^n(x)\varphi_n(x)\mathbf{1}_{\Gamma_n^i(\omega)}(x)\, d\mu_n(\omega) = \sum_{n\in\mathbb{Z}} \varphi_n(x)\int_{\Omega_n}\theta_\omega^n(x)\, d\mu_n(\omega).$$

Additionally define $\gamma(i,\omega) = \gamma_n^i(\omega)$. Fix $n \in \mathbb{Z}$, $\omega \in \Omega_n$, $i \in I$ and assume that $x \in Y$ is such that $\Psi(i,\omega,x) > 0$. Then $x \in \Gamma_n^i(\omega)$ and $d(x,X) \in \left[\varepsilon 2^{n-4}, \varepsilon 2^{n-1}\right]$. Since $d(\gamma_n^i(\omega), \Gamma_n^i(\omega)) < 2d(X, \Gamma_n^i(\omega)) \le 2d(x,X)$ there is $x' \in \Gamma_n^i(\omega)$ for which $d(\gamma_n^i(\omega), x') < 2d(x,X)$. Hence,

$$d(\gamma_n^i(\omega), x) \le 2d(x,X) + \mathrm{diam}(\Gamma_n^i(\omega)) \le 2d(x,X) + 2^n \le \left(2 + \frac{16}{\varepsilon}\right) d(x,X) \le \frac{18}{\varepsilon} d(x,X).$$

Similarly, if $m \in \mathbb{Z}$, $\omega' \in \Omega_m$, $j \in I$ are such that $\Psi(j,\omega',y) > 0$ for some $y \in Y$ then

$$d(\gamma_m^j(\omega'), y) \le \frac{18}{\varepsilon} d(y,X).$$

It follows that

$$d(\gamma_n^i(\omega), \gamma_m^j(\omega')) \le d(x,y) + \frac{18}{\varepsilon} d(x,X) + \frac{18}{\varepsilon} d(y,X) \le \frac{36}{\varepsilon}\left(d(x,y) + \max\{d(x,X), d(y,X)\}\right).$$

Hence,

$$\mathrm{diam}\left(\{x,y\} \cup \{\gamma(i,\omega) :\ \Psi(i,\omega,x) + \Psi(i,\omega,y) > 0\}\right) \le \frac{36}{\varepsilon}\left(d(x,y) + \max\{d(x,X), d(y,X)\}\right).$$

It is therefore enough to show that:

$$\sum_{n\in\mathbb{Z}}\sum_{i\in I}\int_{\Omega_n} |\Psi(i,\omega,x) - \Psi(i,\omega,y)|\, d\mu_n(\omega) \le \frac{680}{\delta} \cdot \frac{d(x,y)}{d(x,y) + \max\{d(x,X), d(y,X)\}},$$

when the decompositions are either (ε,δ)-padded or (ε,δ)-thick and

$$\sum_{n\in\mathbb{Z}}\sum_{i\in I}\int_{\Omega_n}|\Psi(i,\omega,x)-\Psi(i,\omega,y)|\,d\mu_n(\omega)\le 160\left(1+\frac{\varepsilon}{\delta}\right)\cdot\frac{d(x,y)}{d(x,y)+\max\{d(x,X),d(y,X)\}},$$

when the decompositions are (ε,δ)-separating. From here on the proof is exactly as in the proof of Theorem 4.1.

□

Lemma 6.1, Theorem 6.2 and the results of Section 3.2 imply the extension results stated in the introduction, in the case when the target space Z is a complete barycentric metric space. The above argument does not yield a similar generalization of Theorem 1.10, since the construction of Theorem 4.3 does not imply a satisfactory bound as in condition 2 in the definition of a (K,L) gentle partition of unity. An inspection of the proof of Lemma 2.1 shows that we have only used the following property of the Banach space Z: There exists a map $c:\mathcal{M}_Z^{\text{bounded}}\to Z$ such that $c(\delta_z)=z$ for every $z\in Z$ and for every $\mu,\nu\in\mathcal{M}_Z^{\text{bounded}}$ and $p\in Z$,

$$d_Z(c(\mu),c(\nu))\le C\int_Z d(p,z)d|\mu-\nu|(z), \tag{27}$$

where C is a constant (depending on Z). This inequality holds true in Banach spaces due to the identity

$$\int_Z zd\mu(z)-\int_Z zd\nu(z)=\int_Z(z-p)d(\mu-\nu)(z).$$

All the results presented in this paper extend to target spaces satisfying (27) (using K-gentle partitions of unity), and it would be of interest to find a wider class of spaces with this property.

References

[1] K. Ball. Markov chains, Riesz transforms and Lipschitz maps. *Geom. Funct. Anal.*, 2:137–172, 1992.

[2] Y. Bartal. Probabilistic approximation of metric spaces and its algorithmic applications. In *Proceedings of the 37th Annual Symposium Foundations of Computer Science*, 1998.

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

[4] Y. Brudnyi and P. Shvartsman. Stability of the Lipschitz extension property under metric transforms. *Geom. Funct. Anal.*, 12(1):73–79, 2002.

[5] S. Buyalo and V. Schroeder. Extension of Lipschitz maps into 3-manifolds. *Asian J. Math.*, 5(4):685–704, 2001.

[6] G. Calinescu, H. Karloff, and Y. Rabani. Approximation algorithms for the 0-extension problem. In *Proceedings of the 12th Annual ACM-SIAM Symposium on Discrete Algorithms*. ACM, 2001.

[7] M. Charikar, C. Chekuri, A. Goel, S. Guha, and S. A. Plotkin. Approximating a finite metric by a small number of tree metrics. In *Proceedings of the 39th Annual Symposium on Foundations of Computer Science*, 1998.

[8] J. Fakcharoenphol, C. Harrelson, S. Rao, and K. Talwar. An improved approximation algorithm for the 0-extension problem. In *Proceedings of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms*. ACM, 2003.

[9] J. Fakcharoenphol, S. Rao, and K. Talwar. A tight bound on approximating arbitrary metrics by tree metrics. In *35th Annual ACM Symposium on Theory of Computing*. ACM, 2003.

[10] J. Fakcharoenphol and K. Talwar. An improved decomposition theorem for graphs excluding a fixed minor. In *7th International Workshop on Randomization and Approximation Techniques in Computer Science*. Springer-Verlag, 2003.

[11] M. Gromov. Asymptotic invariants of infinite groups. In *Geometric group theory, Vol. 2 (Sussex, 1991)*, volume 182 of *London Math. Soc. Lecture Note Ser.*, pages 1–295. Cambridge Univ. Press, Cambridge, 1993.

[12] M. Gromov. *Metric structures for Riemannian and non-Riemannian spaces*. Birkhäuser, Boston, 1999.

[13] M. Gromov. Random walk in random groups. *Geom. Funct. Anal.*, 13:73–146, 2003.

[14] A. Gutpa, R. Krauthgamer, and J. R. Lee. Bounded geometries, fractals, and low-distortion embeddings. In *Proceedings of the 44th Annual Symposium on Foundations of Computer Science*, 2003.

[15] J. Heinonen. *Lectures on analysis on metric spaces*. Universitext. Springer-Verlag, New York, 2001.

[16] F. John. Extremum problems with inequlities as subsidiary conditions. *Courant Anniversary Volume. Interscience, New York*, pages 187–204–245, 1948.

[17] 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.

[18] W. B. Johnson, J. Lindenstrauss, and G. Schechtman. Extensions of Lipschitz maps into Banach spaces. *Israel J. Math.*, 54(2):129–138, 1986.

[19] M. D. Kirszbraun. Über die zusammenziehenden und Lipschitzchen Transformationen. *Fund. Math.*, (22):77–108, 1934.

[20] P. Klein, S. A. Plotkin, and S. Rao. Excluded minors, network decomposition, and multicommodity flow. In *25th Annual ACM Symposium on Theory of Computing*, pages 682–690, May 1993.

[21] K. Kuratowski and C. Ryll-Nardzewski. A general theorem on selectors. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astron. Phys.*, 13:397–403, 1965.

[22] U. Lang. Extendability of large-scale Lipschitz maps. *Trans. Amer. Math. Soc.*, 351(10):3975–3988, 1999.

[23] U. Lang, B. Pavlović, and V. Schroeder. Extensions of Lipschitz maps into Hadamard spaces. *Geom. Funct. Anal.*, 10(6):1527–1553, 2000.

[24] J. R. Lee and A. Naor. Absolute Lipschitz extendability. *C. R. Acad. Sci. Paris, to appear*, 2003.

[25] J. Lindenstrauss. On nonlinear projections in Banach spaces. *Michigan Math. J.*, 11:263–287, 1964.

[26] N. Linial and M. Saks. Low diameter graph decompositions. *Combinatorica*, 13(4):441–454, 1993.

[27] 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.

[28] J. Matoušek. Extension of Lipschitz mappings on metric trees. *Comment. Math. Univ. Carolin.*, 31(1):99–104, 1990.

[29] B. Mohar and C. Thomassen. *Graphs on surfaces*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, 2001.

[30] A. Naor. A phase transition phenomenon between the isometric and isomorphic extension problems for Hölder functions between L_p spaces. *Mathematika*, 48:253–271, 2001.

[31] S. Rao. Small distortion and volume preserving embeddings for planar and Euclidean metrics. In *Proceedings of the 15th Annual Symposium on Computational Geometry*, pages 300–306. ACM, 1999.

[32] N. Robertson and P. D. Seymour. Graph minors. VIII. A Kuratowski theorem for general surfaces. *J. Combin. Theory Ser. B*, 48(2):255–288, 1990.

[33] L. Silberman. Addendum to "Random walk in random groups" by M. Gromov. *Geom. Funct. Anal.*, 13:147–177, 2003.

[34] I. G. Tsar′kov. Extension of Hilbert-valued Lipschitz mappings. *Vestnik Moskov. Univ. Ser. I Mat. Mekh.*, (6):9–16, 72, 1999.

[35] J. H. Wells and L. R. Williams. *Embeddings and extensions in analysis*. Springer-Verlag, New York, 1975. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 84.

[36] H. Whitney. Analytic extension of differentiable functions defined in closed sets. *Trans. Amer. Math. Soc.*, 36:63–89, 1934.

[37] H. Whitney. Differentiable functions defined in closed sets I. *Trans. Amer. Math. Soc.*, 36:369–387, 1934.