Chapter 5

Barycentric-valued Lipschitz extension

In this section, we will explain how separation profiles relate to Lipschitz extension. We cannot invoke [173] as a "black box" because we need a more general result and our definition of random partitions differs from that of [173]. But, the modifications that are required in order to apply the ideas of [173] in the present setting are of a secondary nature, and the main geometric content of the phenomenon that is explained below is the same as in [173].

In addition to making the present article self-contained, there are more advantages to including here complete proofs of Theorem 66 and Theorem 114. Firstly, the reasoning of [173] was designed to deal with a more general setting (treating multiple notions of random partitions at once), and it is illuminating to present a proof for separating decompositions in isolation, which leads to simplifications. Secondly, since [173] appeared, alternative viewpoints have been developed that relate it to optimal transport, as carried out by Kozdoba [158], Brudnyi and Brudnyi [62], Ohta [243], and culminating more recently with a comprehensive treatment by Ambrosio and Puglisi [11]. Here we will frame the construction using the optimal transport methodology, which has conceptual advantages that go beyond yielding a clearer restructuring of the argument. The optimal transport viewpoint had an important role in quantitative improvements that were obtained in [231, 233], as well as results that will appear in forthcoming works. As a byproduct, we will use this viewpoint to easily derive a stability statement for convex hull-valued Lipschitz extension under metric transforms.

5.1 Notational preliminaries

We will start by quickly setting notation and terminology for basic concepts in measure theory and optimal transport. Everything that we describe in this subsection is standard and is included here only in order to avoid any ambiguities in the subsequent discussions.

Given a signed measure μ on a measurable space $(\Omega, \mathcal{F})$, its Hahn–Jordan decomposition is denoted $\mu = \mu^+ - \mu^-$, i.e., μ^+, μ^- are disjointly supported nonnegative measures. The total variation measure of μ is $|\mu| = \mu^+ + \mu^-$. For $A \in \mathcal{F}$, the restriction of μ to A is denoted $\mu\lfloor_A$, i.e., $\mu\lfloor_A(E) = \mu(A \cap E)$ for $E \in \mathcal{F}$. If $(\Omega', \mathcal{F}')$ is another measurable space and $f : \Omega \to \Omega'$ is a measurable mapping, then the pushforward of μ under f is denoted $f_\#\mu$. Thus $f_\#\mu(E) = \mu(f^{-1}(E))$ for $E \in \mathcal{F}'$, or equivalently

$$\forall h \in L_1(f_\# \mu), \quad \int_{\Omega'} h(\omega')\, \mathrm{d} f_\# \mu(\omega') = \int_{\Omega} h(f(\omega))\, \mathrm{d}\mu(\omega).$$

Suppose from now on that $(\mathfrak{M}, d_{\mathfrak{M}})$ is a Polish metric space. A signed Borel measure μ on $\mathfrak{M}$ has finite first moment if $\int_{\mathfrak{M}} d_{\mathfrak{M}}(x,y)\, \mathrm{d}|\mu|(y) < \infty$ for all $x \in \mathfrak{M}$. Note that this implies in particular that $|\mu|(\mathfrak{M}) < \infty$, because if $x, x' \in \mathfrak{M}$ are distinct points, then the mapping $(y \in \mathfrak{M}) \mapsto [d_{\mathfrak{M}}(x,y) + d_{\mathfrak{M}}(x',y)]/d_{\mathfrak{M}}(x,x')$ belongs to $L_1(|\mu|)$ and takes values in $[1,\infty)$ by the triangle inequality.

The set of all of the signed Borel measures on $\mathfrak{M}$ of finite first moment is denoted $\mathsf{M}_1(\mathfrak{M}, d_{\mathfrak{M}})$ or simply $\mathsf{M}_1(\mathfrak{M})$ if the metric is clear from the context. The set of all nonnegative measures in $\mathsf{M}_1(\mathfrak{M})$ is denoted $\mathsf{M}_1^+(\mathfrak{M})$, the set of all $\mu \in \mathsf{M}_1(\mathfrak{M})$ with total mass 0, i.e., $\mu^+(\mathfrak{M}) = \mu^-(\mathfrak{M})$, is denoted $\mathsf{M}_1^0(\mathfrak{M})$, and the set of all probability measures in $\mathsf{M}_1(\mathfrak{M})$ is denoted $\mathsf{P}_1(\mathfrak{M})$.

Given $\mu, \nu \in \mathsf{M}_1^+(\mathfrak{M})$ with $\mu(\mathfrak{M}) = \nu(\mathfrak{M})$, a Borel measure π on $\mathfrak{M} \times \mathfrak{M}$ is a coupling of μ and ν if

$$\pi(E \times \mathfrak{M}) = \mu(A) \quad \text{and} \quad \pi(\mathfrak{M} \times E) = \nu(A)$$

for every Borel subset $E \subseteq \mathfrak{M}$. The set of couplings of μ and ν is denoted $\Pi(\mu,\nu) \subseteq \mathsf{M}_1^+(\mathfrak{M} \times \mathfrak{M})$. Note that $(\mu \times \nu)/\mu(\mathfrak{M}) = (\mu \times \nu)/\nu(\mathfrak{M}) \in \Pi(\mu,\nu)$, so $\Pi(\mu,\nu) \neq \varnothing$. The Wasserstein-1 distance between μ and ν that is induced by the metric $d_{\mathfrak{M}}$, denoted $\mathsf{W}_1^{d_{\mathfrak{M}}}(\mu,\nu)$ or simply $\mathsf{W}_1(\mu,\nu)$ if the metric is clear from the context, is the infimum of $\int_{\mathfrak{M}\times\mathfrak{M}} d_{\mathfrak{M}}(x,y)\, \mathrm{d}\pi(x,y)$ over all possible couplings $\pi \in \Pi(x,y)$. Since $(\mathfrak{M}, d_{\mathfrak{M}})$ is Polish, the metric space $(\mathsf{P}_1(\mathfrak{M}), \mathsf{W}_1)$ is also Polish; see, e.g., [42] or [10, Proposition 7.1.5]. Throughout what follows, $\mathsf{P}_1(\mathfrak{M})$ will be assumed to be equipped with the metric W_1. The Kantorovich–Rubinstein duality theorem (see, e.g., [307, Theorem 5.10]) asserts that

$$\mathsf{W}_1(\mu,\nu) = \sup_{\substack{\psi:\mathfrak{M}\to\mathbb{R} \\ \|\psi\|_{\mathrm{Lip}(\mathfrak{M})}=1}} \left(\int_{\mathfrak{M}} \psi\, \mathrm{d}\mu - \int_{\mathfrak{M}} \psi\, \mathrm{d}\nu \right). \tag{5.1}$$

Note that (5.1) implies that $\mathsf{W}_1(\mu+\tau, \nu+\tau) = \mathsf{W}_1(\mu,\nu)$ for every $\tau \in \mathsf{M}_1^+(\mathfrak{M})$.

For $\mu \in \mathsf{M}_1^0(\mathfrak{M})$ we have $\mu^+(\mathfrak{M}) = \mu^-(\mathfrak{M})$, so we can define[1]:

$$\|\mu\|_{\mathsf{W}_1(\mathfrak{M})} = \mathsf{W}_1(\mu^+, \mu^-).$$

[1]Note for later use that if $\mu, \nu \in \mathsf{M}_1^+(\mathfrak{M})$ satisfy $\mu(\mathfrak{M}) = \nu(\mathfrak{M})$, then $\mu - \nu \in \mathsf{M}_1^0(\mathfrak{M})$ and $\|\mu - \nu\|_{\mathsf{W}_1(\mathfrak{M})} = \mathsf{W}_1(\mu,\nu)$. For a standard justification of the latter assertion, see, e.g., the simple deduction of [236, equation (2.2)].

This turns $\mathsf{M}_1^0(\mathfrak{M})$ into a normed space whose completion is called the *free space* over $\mathfrak{M}$ (also known as the Arens–Eells space over $\mathfrak{M}$), and is denoted $\mathfrak{F}(\mathfrak{M})$; see [16, 113,310] for more on this topic, and note that while $\mathfrak{F}(\mathfrak{M})$ is commonly defined as the closure of the *finitely supported* measures in $\mathsf{M}_1^0(\mathfrak{M})$ with respect to the Wasserstein-1 norm, since the finitely supported measures are dense in $\mathsf{M}_1^0(\mathfrak{M})$ (see, e.g., [307, Theorem 6.18]), the definitions coincide. It follows from (5.1) that the dual of $\mathfrak{F}(\mathfrak{M})$ is canonically isometric to the space of all the real-valued Lipschitz functions on $\mathfrak{M}$ that vanish at some (arbitrary but fixed) point $x_0 \in \mathfrak{M}$, equipped with the norm $\|\cdot\|_{\mathrm{Lip}(\mathfrak{M})}$.

Suppose that $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$ is a separable Banach space and fix $\mu \in \mathsf{M}_1(\mathfrak{M})$. By the Pettis measurability criterion [249] (see also [36, Proposition 5.1]), any $f \in \mathrm{Lip}(\mathfrak{M}; \mathbf{Z})$ is $|\mu|$-measurable. Moreover, we have $\|f\|_{\mathbf{Z}} \in L_1(|\mu|)$ because if we fix $x \in \mathfrak{M}$, then for every $y \in \mathfrak{M}$,

$$\begin{aligned} \|f(y)\|_{\mathbf{Z}} &\leqslant \|f(y) - f(x)\|_{\mathbf{Z}} + \|f(x)\|_{\mathbf{X}} \\ &\leqslant \|f\|_{\mathrm{Lip}(\mathfrak{M};\mathbf{Z})} d_{\mathfrak{M}}(y, x) + \|f(x)\|_{\mathbf{X}} \in L_1(|\mu|), \end{aligned}$$

where the last step holds by the definition of $\mathsf{M}_1(\mathfrak{M})$ and the fact that it implies that $|\mu|(\mathfrak{M}) < \infty$. By Bochner's integrability criterion [40] (see also [36, Proposition 5.2]), it follows that the Bochner integrals $\int_{\mathfrak{M}} f \, \mathrm{d}\mu^+$ and $\int_{\mathfrak{M}} f \, \mathrm{d}\mu^-$ are well-defined elements of $\mathbf{Z}$, so we can consider the vector

$$\mathfrak{I}_f(\mu) \stackrel{\mathrm{def}}{=} \int_{\mathfrak{M}} f \, \mathrm{d}\mu = \int_{\mathfrak{M}} f \, \mathrm{d}\mu^+ - \int_{\mathfrak{M}} f \, \mathrm{d}\mu^- \in \mathbf{Z}. \tag{5.2}$$

If $\mu \in \mathsf{M}_1^0(\mathfrak{M})$, then $\mathfrak{I}_f(\mu) = \int_{\mathfrak{M}\times\mathfrak{M}} (f(x) - f(y)) \, \mathrm{d}\pi(x, y)$ for every coupling $\pi \in \Pi(\mu^+, \mu^-)$. Consequently, $\|\mathfrak{I}_f(\mu)\|_{\mathbf{Z}} \leqslant \|f\|_{\mathrm{Lip}(\mathfrak{M};\mathbf{Z})} \int_{\mathfrak{M}\times\mathfrak{M}} d_{\mathfrak{M}}(x, y) \, \mathrm{d}\pi(x, y)$, so by taking the infimum over all $\pi \in \Pi(\mu^+, \mu^-)$ we see that the norm of the linear operator $\mathfrak{I}_f$ from $(\mathsf{M}_1^0(\mathfrak{M}), \|\cdot\|_{\mathsf{W}_1})$ to $\mathbf{Z}$ satisfies

$$\|\mathfrak{I}_f\|_{(\mathsf{M}_1^0(\mathfrak{M}), \|\cdot\|_{\mathsf{W}_1}) \to \mathbf{Z}} \leqslant \|f\|_{\mathrm{Lip}(\mathfrak{M};\mathbf{Z})}. \tag{5.3}$$

Since $\mathsf{M}_1^0(\mathfrak{M})$ is dense in $\mathfrak{F}(\mathfrak{M})$, it follows that $\mathfrak{I}_f$ extends uniquely to a linear operator $\mathfrak{I}_f : \mathfrak{F}(\mathfrak{M}) \to \mathbf{Z}$ of norm at most $\|f\|_{\mathrm{Lip}(\mathfrak{M};\mathbf{Z})}$. So, even though elements of $\mathfrak{F}(\mathfrak{M})$ need not be measures, one can consider the "integral" $\mathfrak{I}_f(\phi) \in \mathbf{Z}$ of $f \in \mathrm{Lip}(\mathfrak{M}; \mathbf{Z})$ with respect to $\phi \in \mathfrak{F}(\mathfrak{M})$; see [114] for more on this topic.

5.2 Refined extension moduli

Continuing with the notation that was introduced by Matoušek [199], we will consider the following parameters related to Lipschitz extension. Suppose that $(\mathfrak{M}, d_{\mathfrak{M}})$, $(\mathfrak{N}, d_{\mathfrak{N}})$ are metric spaces and that $\mathfrak{C} \subseteq \mathfrak{M}$. Denote by $\mathsf{e}(\mathfrak{M}, \mathfrak{C}; \mathfrak{N})$ the infimum over those $K \in [1,\infty]$ such that for every $f : \mathcal{C} \to \mathfrak{N}$ with $\|f\|_{\mathrm{Lip}(\mathcal{C};\mathfrak{N})} < \infty$ there is $F : \mathfrak{M} \to \mathfrak{N}$ that extends f and satisfies

$$\|F\|_{\mathrm{Lip}(\mathfrak{M};\mathfrak{N})} \leqslant K\|f\|_{\mathrm{Lip}(\mathcal{C};\mathfrak{N})}.$$

The supremum of $\mathsf{e}(\mathfrak{M},\mathcal{C};\mathfrak{N})$ over all subsets $\mathcal{C} \subseteq \mathfrak{M}$ will be denoted $\mathsf{e}(\mathfrak{M};\mathfrak{N})$. Note that when $\mathfrak{N}$ is complete, $\mathfrak{N}$-valued Lipschitz functions on $\mathcal{C}$ automatically extend to the closure of $\mathcal{C}$ while preserving the Lipschitz constant, so we may assume here that $\mathcal{C}$ is closed. The supremum of $\mathsf{e}(\mathfrak{M},\mathcal{C};\mathbf{Z})$ over all Banach spaces $(\mathbf{Z},\|\cdot\|_{\mathbf{Z}})$ will be denoted below by $\mathsf{e}(\mathfrak{M},\mathcal{C})$. Thus, the notation $\mathsf{e}(\mathfrak{M})$ of the Introduction coincides with the supremum of $\mathsf{e}(\mathfrak{M},\mathcal{C})$ over all subsets $\mathcal{C} \subseteq \mathfrak{M}$.

If $(\mathfrak{M},d_{\mathfrak{M}})$ is a metric space, $\mathcal{C} \subseteq \mathfrak{M}$, and $(\mathbf{Z},\|\cdot\|_{\mathbf{Z}})$ is a Banach space, then it is natural to consider variants of the above definitions with the additional restrictions that the extended mapping F is required to take values in either the closure of the linear span of $f(\mathcal{C})$ or the closure of the convex hull of $f(\mathcal{C})$. Namely, let $\mathsf{e}_{\mathrm{span}}(\mathfrak{M},\mathcal{C};\mathbf{Z})$ be the infimum over those $K \in [1,\infty]$ such that for every $f : \mathcal{C} \to \mathbf{Z}$ there exists

$$F : \mathfrak{M} \to \overline{\mathrm{span}}\big(f(\mathcal{C})\big)$$

that extends f and satisfies

$$\|F\|_{\mathrm{Lip}(\mathfrak{M};\mathbf{Z})} \leqslant K\|f\|_{\mathrm{Lip}(\mathcal{C};\mathbf{Z})}. \tag{5.4}$$

Analogously, let $\mathsf{e}_{\mathrm{conv}}(\mathfrak{M},\mathcal{C};\mathbf{Z})$ be the infimum over $K \in [1,\infty]$ such that for every $f : \mathcal{C} \to \mathbf{Z}$ there exists

$$F : \mathfrak{M} \to \overline{\mathrm{conv}}\big(f(\mathcal{C})\big)$$

that extends f and satisfies (5.4). We then define $\mathsf{e}_{\mathrm{conv}}(\mathfrak{M},\mathcal{C})$ to be the supremum of $\mathsf{e}_{\mathrm{conv}}(\mathfrak{M},\mathcal{C};\mathbf{Z})$ over all possible Banach spaces $(\mathbf{Z},\|\cdot\|_{\mathbf{Z}})$. Note that while one could attempt to define $\mathsf{e}_{\mathrm{span}}(\mathfrak{M},\mathcal{C})$ similarly, there is no point to do so because it would result in the previously defined quantity $\mathsf{e}(\mathfrak{M},\mathcal{C})$. By considering the supremum of $\mathsf{e}_{\mathrm{conv}}(\mathfrak{M},\mathcal{C})$ over all subsets $\mathcal{C} \subseteq \mathfrak{M}$, one defines the quantity $\mathsf{e}_{\mathrm{conv}}(\mathfrak{M})$.

Remark 133. By [179] one can have $\mathsf{e}(\mathfrak{M},\mathcal{C};\mathbf{Z}) = \mathsf{e}(\mathfrak{M};\mathbf{Z}) = 1$ yet $\mathsf{e}_{\mathrm{span}}(\mathfrak{M},\mathcal{C},\mathbf{Z}) = \infty$ for some metric space $(\mathfrak{M},d_{\mathfrak{M}})$, some $\mathcal{C} \subseteq \mathfrak{M}$ and some Banach space $(\mathbf{Z},\|\cdot\|_{\mathbf{Z}})$. Indeed, if $\mathbf{X}$ is a closed reflexive subspace of ℓ_∞ and $\mathbf{V} \subseteq \mathbf{X}$ is a closed uncomplemented subspace of $\mathbf{X}$, then by [179] (see also [36, Corollary 7.3]) there is no Lipschitz retraction from $\mathbf{X}$ onto $\mathbf{V}$. Equivalently, the identity mapping from $\mathbf{V}$ to $\mathbf{V}$ cannot be extended to a Lipschitz mapping from $\mathbf{X}$ to $\mathbf{V}$. Hence, since $\mathrm{span}(\mathbf{V}) = \mathbf{V} \subseteq \ell_\infty$, we have $\mathsf{e}_{\mathrm{span}}(\mathbf{X},\mathbf{V};\ell_\infty) = \infty$. In contrast, $\mathsf{e}(\mathbf{X};\ell_\infty) = 1$ by the nonlinear Hahn–Banach theorem (see [206] or, e.g., [36, Lemma 1.1]). By combining [290] with the discretization method of [138] (see also [195]), one can quantify the above example by showing that for arbitrarily large $n \in \mathbb{N}$ there are Banach spaces $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ and $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$, and a subset $\mathcal{C} \subseteq \mathbf{X}$ with $|\mathcal{C}| = n$ for which we have

$$\frac{\mathsf{e}_{\mathrm{span}}(\mathbf{X}, \mathcal{C}; \mathbf{Z})}{\mathsf{e}(\mathbf{X}, \mathcal{C}; \mathbf{Z})} \gtrsim \sqrt{\frac{\log n}{\log\log n}}. \tag{5.5}$$

(In fact, in (5.5) one can have $\mathsf{e}(\mathbf{X}, \mathcal{C}; \mathbf{Z}) = \mathsf{e}(\mathbf{X}; \mathbf{Z}) = 1$.) At present, the right-hand side of (5.5) is the largest asymptotic dependence on n that we are able to obtain for this question, and it remains an interesting open problem to determine the best possible asymptotics here.

Most, but not all, of the Lipschitz extension methods in the literature, including Kirszbraun's extension theorem [155], Ball's extension theorem [23] and methods that rely on (variants of) partitions of unity such as in [61, 140, 166, 173], yield convex hull-valued extensions, i.e., they actually provide bounds on the quantity $\mathsf{e}_{\mathrm{conv}}(\mathcal{M}, \mathcal{C}; \mathbf{Z})$. Nevertheless, it seems likely that there is no $\varphi : [1, \infty) \to [1, \infty)$ such that $\mathsf{e}_{\mathrm{conv}}(\mathcal{M}) \leqslant \varphi(\mathsf{e}(\mathcal{M}))$ for every Polish metric space $(\mathcal{M}, d_{\mathcal{M}})$, though if such an estimate were available, then it would be valuable; see, e.g., Remark 141. In fact, we propose the following conjecture.

Conjecture 134. There exists a Polish metric space $(\mathcal{M}, d_{\mathcal{M}})$ for which $\mathsf{e}(\mathcal{M}) < \infty$ yet $\mathsf{e}_{\mathrm{conv}}(\mathcal{M}) = \infty$.

Remark 135. By definition, for every metric space $(\mathcal{M}, d_{\mathcal{M}})$, every Banach space $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$ and every $\mathcal{C} \subseteq \mathcal{M}$,

$$\mathsf{e}_{\mathrm{conv}}(\mathcal{M}, \mathcal{C}; \mathbf{Z}) \geqslant \mathsf{e}_{\mathrm{span}}(\mathcal{M}, \mathcal{C}; \mathbf{Z}) \geqslant \mathsf{e}(\mathcal{M}, \mathcal{C}; \mathbf{Z}).$$

We explained in Remark 133 that the second of these inequalities can be strict (in a strong sense). However, as a complement to Conjecture 134, we state that to the best of our knowledge it is unknown whether this is so for the first of these inequalities, i.e., if it could happen that $\mathsf{e}_{\mathrm{span}}(\mathcal{M}, \mathcal{C}; \mathbf{Z}) < \infty$ yet $\mathsf{e}_{\mathrm{conv}}(\mathcal{M}, \mathcal{C}; \mathbf{Z}) = \infty$. We suspect that this is possible, but if not, then it would be interesting to investigate how one could bound $\mathsf{e}_{\mathrm{conv}}(\mathcal{M}, \mathcal{C}; \mathbf{Z})$ from above by a function of $\mathsf{e}_{\mathrm{span}}(\mathcal{M}, \mathcal{C}; \mathbf{Z})$. We do know that there are a metric space $(\mathcal{M}, d_{\mathcal{M}})$, a Banach space $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$, a subset $\mathcal{C} \subseteq \mathcal{M}$ and a Lipschitz mapping $f : \mathcal{C} \to \mathbf{Z}$ that can be extended to a Lipschitz mapping that takes values in $\overline{\mathrm{span}}(f(\mathcal{C}))$ but cannot be extended to a Lipschitz mapping that takes values in $\overline{\mathrm{conv}}(f(\mathcal{C}))$. To see this, let $\{e_j\}_{j=1}^{\infty}$ be the standard basis of ℓ_∞. For $n \in \mathbb{N}$ set $m(n) = n(n-1)/2$ and let $\mathbf{X}_n$ be the span of $\{e_{m(n)+1}, \ldots, e_{m(n+1)}\}$ in ℓ_∞. Thus, $\mathbf{X}_n$ is isometric to ℓ_∞^n and $\ell_\infty = (\bigoplus_{n=1}^{\infty} \mathbf{X}_n)_\infty$. By [290], there is a linear subspace $\mathbf{V}_n$ of $\mathbf{X}_n$ such that every linear projection $\mathsf{Q} : \mathbf{X}_n \to \mathbf{V}_n$ satisfies

$$\|\mathsf{Q}\|_{\mathbf{X}_n \to \mathbf{V}_n} \gtrsim \sqrt{n}.$$

By the method of [138], it follows that there exists[2] $\mathcal{A}_n \subseteq B_{V_n} = V_n \cap B_{\ell_\infty}$ with $|\mathcal{A}_n| \leqslant n^{O(n)}$ such that $\|F_n\|_{\mathrm{Lip}(X_n;V_n)} \gtrsim \sqrt{n}$ for any $F_n : \mathbf{X}_n \to \mathbf{V}_n$ that extends the formal identity $\mathsf{Id}_{\mathcal{A}_n\to\mathbf{V}_n} : \mathcal{A}_n \to \mathbf{V}_n$. By compactness, there exists $\delta_n \in (0,1)$ such that if we define

$$\mathcal{C}_n = \mathcal{A}_n \cup \left\{\delta_n e_{m(n)+1}, \ldots, \delta_n e_{m(n+1)}\right\} \cup \{0\},$$

then also $\|\Phi_n\|_{\mathrm{Lip}(\mathbf{X}_n;\mathbf{X}_n)} \gtrsim \sqrt{n}$ for any mapping Φ_n from $\mathbf{X}_n$ to the polytope $\overline{\mathrm{conv}}(\mathcal{C}_n)$ that extends the formal identity $\mathsf{Id}_{\mathcal{C}_n\to\mathbf{X}_n}$. Consider the subset

$$\mathcal{C} = \bigcup_{n=1}^{\infty}{}^{\cdot}\, \mathcal{C}_n \subseteq \ell_\infty.$$

Suppose that $\Phi : \ell_\infty \to \overline{\mathrm{conv}}(\mathcal{C})$ extends $\mathsf{Id}_{\mathcal{C}\to\ell_\infty}$. Then, for each $n \in \mathbb{N}$ the mapping $\mathsf{R}_n \circ (\Phi|_{\mathbf{X}_n}) : \mathbf{X}_n \to \mathbf{X}_n$ extends $\mathsf{Id}_{\mathcal{C}_n\to\ell_\infty}$ and takes values in $\overline{\mathrm{conv}}(\mathcal{C}_n)$, where we denote the canonical restriction operator from ℓ_∞ to $\mathbf{X}_n$ by $\mathsf{R}_n : \ell_\infty \to \mathbf{X}_n$. Hence,

$$\|\Phi\|_{\mathrm{Lip}(\ell_\infty;\mathbf{X}_n)} \geqslant \|\mathsf{R}_n \circ (\Phi|_{\mathbf{X}_n})\|_{\mathrm{Lip}(\mathbf{X}_n;\mathbf{X}_n)} \gtrsim \sqrt{n}.$$

Since this holds for every $n \in \mathbb{N}$, the mapping Φ is not Lipschitz. Consequently, we have $\mathsf{e}_{\mathrm{conv}}(\ell_\infty, \mathcal{C}; \ell_\infty) = \infty$. At the same time, by construction we have $\overline{\mathrm{span}}(\mathcal{C}) = \overline{\mathrm{span}}(\{e_j\}_{j=1}^{\infty}) = c_0$ (recall that c_0 commonly denotes the subspace of ℓ_∞ consisting of all those sequences that tend to 0). So, any 2-Lipschitz retraction ρ of ℓ_∞ onto c_0 extends $\mathsf{Id}_{\mathcal{C}\to\ell_\infty}$ and takes values in $\overline{\mathrm{span}}(\mathcal{C})$; the existence of such a retraction ρ is due to [179] (see also [36, Example 1.5]). If $\mathsf{e}_{\mathrm{span}}(\ell_\infty, \mathcal{C}; \ell_\infty)$ were finite, then this example would answer the above question,[3] but we suspect that in fact $\mathsf{e}_{\mathrm{span}}(\ell_\infty, \mathcal{C}; \ell_\infty) = \infty$.

Proposition 136 is a convenient characterization of the quantities $\mathsf{e}(\mathfrak{M}, \mathcal{C})$ and $\mathsf{e}_{\mathrm{conv}}(\mathfrak{M}, \mathcal{C})$; while it was not previously stated explicitly in this form, its proof is based on well-understood ideas.

Proposition 136. *Suppose that $(\mathfrak{M}, d_\mathfrak{M})$ is a metric space, $\mathcal{C}$ is a Polish subset of $\mathfrak{M}$ and $s_0 \in \mathcal{C}$. Fix two nonnegative functions $\mathfrak{d} : \mathfrak{M} \times \mathfrak{M} \to [0,\infty)$ and $\varepsilon : \mathcal{C} :\to [0,\infty)$. Then, the following two equivalences hold.*

---

[2]The subset $\mathcal{A}_n$ can be taken to be any ε_n-net of the unit sphere of $\mathbf{V}_n$, for any $\varepsilon_n \lesssim n^{-3/2}$. Note, however, that the bound that follows from [138] (and also [195, Appendix C]) is $\varepsilon_n \lesssim n^{-2}$, and this suffices for the present purposes; see [233, Theorem 23] for the above stated weaker requirement from ε_n.

[3]And, it would show that for arbitrarily large $k \in \mathbb{N}$ there exist a metric space $(\mathfrak{M}, d_\mathfrak{M})$, a Banach space $(\mathbf{Z}, \|\cdot\|_\mathbf{Z})$ and a subset $\mathcal{S} \subseteq \mathfrak{M}$ with $|\mathcal{S}| = k$ such that $\mathsf{e}_{\mathrm{conv}}(\mathfrak{M}, \mathcal{S}; \mathbf{Z})/\mathsf{e}_{\mathrm{span}}(\mathfrak{M}, \mathcal{S}; \mathbf{Z}) \gtrsim \sqrt{(\log k)/\log\log k}$. It would then remain an interesting open question to determine the largest possible asymptotic dependence on k here.

(1) *The following two statements are equivalent.*

- *For every Banach space* $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$ *and every mapping* $f : \mathcal{C} \to \mathbf{Z}$ *that is* 1*-Lipschitz with respect to the metric* $d_{\mathfrak{M}}$ *there exists* $F : \mathfrak{M} \to \mathbf{Z}$ *that satisfies the following two conditions.*
  - $\|F(s) - f(s)\|_{\mathbf{Z}} \leqslant \varepsilon(s)$ *for every* $s \in \mathcal{C}$.
  - $\|F(x) - F(y)\|_{\mathbf{Z}} \leqslant \mathfrak{d}(x, y)$ *for every* $x, y \in \mathfrak{M}$.
- *There exists a family* $\{\phi_x\}_{x\in\mathfrak{M}}$ *of elements of the free space* $\mathfrak{F}(\mathcal{C})$ *with the following properties.*
  - $\|\phi_s - \boldsymbol{\delta}_s + \boldsymbol{\delta}_{s_0}\|_{\mathfrak{F}(\mathcal{C})} \leqslant \varepsilon(s)$ *for every* $s \in \mathcal{C}$.
  - $\|\phi_x - \phi_y\|_{\mathfrak{F}(\mathcal{C})} \leqslant \mathfrak{d}(x, y)$ *for every* $x, y \in \mathfrak{M}$.

(2) *The following two statements are equivalent.*

- *For every Banach space* $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$ *and every mapping* $f : \mathcal{C} \to \mathbf{Z}$ *that is* 1*-Lipschitz with respect to the metric* $d_{\mathfrak{M}}$ *there exists* $F : \mathfrak{M} \to \overline{\mathrm{conv}}(f(\mathcal{C}))$ *that satisfies the following two conditions.*
  - $\|F(s) - f(s)\|_{\mathbf{Z}} \leqslant \varepsilon(s)$ *for every* $s \in \mathcal{C}$.
  - $\|F(x) - F(y)\|_{\mathbf{Z}} \leqslant \mathfrak{d}(x, y)$ *for every* $x, y \in \mathfrak{M}$.
- *There exists a family* $\{\mu_x\}_{x\in\mathfrak{M}}$ *of probability measures in* $\mathsf{P}_1(\mathcal{C})$ *with the following properties.*
  - $\mathsf{W}_1^{d_{\mathfrak{M}}}(\mu_s, \boldsymbol{\delta}_s) \leqslant \varepsilon(s)$ *for every* $s \in \mathcal{C}$.
  - $\mathsf{W}_1^{d_{\mathfrak{M}}}(\mu_x, \mu_y) \leqslant \mathfrak{d}(x, y)$ *for every* $x, y \in \mathfrak{M}$.

In the setting of Proposition 136, if $\varepsilon(s) = 0$ for every $s \in \mathcal{C}$ and also $\mathfrak{d} = Kd_{\mathfrak{M}}$ for some $K \geqslant 1$, then in [11, Definition 2.7] a family $\{\phi_x\}_{x\in\mathfrak{M}} \subseteq \mathfrak{F}(\mathcal{C})$ as in part (1) of Proposition 136 is called a K-random projection of $\mathfrak{M}$ onto $\mathcal{C}$, and in [243, Definition 3.1] a family $\{\mu_x\}_{x\in\mathfrak{M}} \subseteq \mathsf{P}_1(\mathcal{C})$ as in part (2) of Proposition 136 is called a stochastic K-Lipschitz retraction of $\mathfrak{M}$ onto $\mathcal{C}$ while in [11, Definition 2.7] it is called a strong K-random projection of $\mathfrak{M}$ onto $\mathcal{C}$.

Proof of Proposition 136. Suppose first that $\{\phi_x\}_{x\in\mathfrak{M}} \subseteq \mathfrak{F}(\mathcal{C})$ and $\{\mu_x\}_{x\in\mathfrak{M}} \subseteq \mathsf{P}_1(\mathcal{C})$ are as in the two parts of Proposition 136. Let $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$ be a Banach space and fix a 1-Lipschitz function $f : \mathcal{C} \to \mathbf{Z}$. Since $\mathcal{C}$ is Polish and hence separable, by replacing $\mathbf{Z}$ with the closure of the linear span of $f(\mathcal{C})$ we may assume that $\mathbf{Z}$ is separable. Recalling the notation (5.2) and the discussion immediately following it for the (integration) operator

$$\mathfrak{I}_f : \mathsf{M}_1(\mathfrak{M}) \cup \mathfrak{F}(\mathfrak{M}) \to \mathbf{Z},$$

define two (linear) mappings

$$\mathsf{Ext}_{\mathcal{C}}^{\phi} f, \mathsf{Ext}_{\mathcal{C}}^{\mu} f : \mathfrak{M} \to \mathbf{Z}$$

by setting for every $x \in \mathfrak{M}$,

$$\mathsf{Ext}_{\mathcal{C}}^{\phi} f(x) \stackrel{\text{def}}{=} f(s_0) + \mathfrak{I}_f(\phi_x) \quad \text{and} \quad \mathsf{Ext}_{\mathcal{C}}^{\mu} f(x) \stackrel{\text{def}}{=} \mathfrak{I}_f(\mu_x) \stackrel{(5.2)}{=} \int_{\mathcal{C}} f \, \mathrm{d}\mu_x. \tag{5.6}$$

Observe that since μ_x is a probability measure, $\mathsf{Ext}_{\mathcal{C}}^{\mu} f(x)$ belongs to the closure of the convex hull of $f(\mathcal{C})$.

For every $x, y \in \mathfrak{M}$ we have

$$\left\|\mathsf{Ext}_{\mathcal{C}}^{\phi} f(x) - \mathsf{Ext}_{\mathcal{C}}^{\phi} f(y)\right\|_{\mathbf{Z}} = \left\|\mathfrak{I}_f(\phi_x - \phi_y)\right\|_{\mathbf{Z}} \stackrel{(5.3)}{\leqslant} \|\phi_x - \phi_y\|_{\mathfrak{F}(\mathcal{C})} \leqslant \mathfrak{d}(x, y),$$

and similarly (using Kantorovich–Rubinstein duality),

$$\left\|\mathsf{Ext}_{\mathcal{C}}^{\mu} f(x) - \mathsf{Ext}_{\mu}^{\phi} f(y)\right\|_{\mathbf{Z}} \leqslant \mathsf{W}_1^{d_{\mathfrak{M}}}(\mu_x, \mu_y) \leqslant \mathfrak{d}(x, y).$$

Also, for every $s \in \mathcal{C}$ we have

$$\left\|\mathsf{Ext}_{\mathcal{C}}^{\phi} f(s) - f(s)\right\|_{\mathbf{Z}} = \left\|\mathfrak{I}_f(\phi_s - \boldsymbol{\delta}_s + \boldsymbol{\delta}_{s_0})\right\|_{\mathbf{Z}} \leqslant \|\phi_s - \boldsymbol{\delta}_s + \boldsymbol{\delta}_{s_0}\|_{\mathfrak{F}(\mathcal{C})} \leqslant \varepsilon(s),$$

and similarly,

$$\left\|\mathsf{Ext}_{\mathcal{C}}^{\mu} f(s) - f(s)\right\|_{\mathbf{Z}} = \left\|\mathfrak{I}_f(\phi_s - \boldsymbol{\delta}_s)\right\|_{\mathbf{Z}} \leqslant \mathsf{W}_1^{d_{\mathfrak{M}}}(\mu_s, \boldsymbol{\delta}_s) \leqslant \varepsilon(s).$$

Conversely, define $f : \mathcal{C} \to \mathfrak{F}(\mathcal{C})$ by setting $f(s) = \boldsymbol{\delta}_s - \boldsymbol{\delta}_{s_0}$ for each $s \in \mathcal{C}$. Then f is 1-Lipschitz. Fix $F : \mathfrak{M} \to \mathfrak{F}(\mathcal{C})$. Writing $F(x) = \phi_x$ for each $x \in \mathfrak{M}$, the assumptions of the first half of part (1) of Proposition 136 coincide with the assertions of its second half. As $\mathcal{C}$ is Polish, $\mathsf{P}_1(\mathcal{C})$ is closed in $\mathfrak{F}(\mathcal{C})$. Therefore,

$$\overline{\operatorname{conv}}\big(f(\mathcal{C})\big) = \mathsf{P}_1(\mathcal{C}) - \boldsymbol{\delta}_{s_0},$$

where the closure is with respect to the topology of $\mathfrak{F}(\mathcal{C})$. Thus, if

$$F(\mathfrak{M}) \subseteq \overline{\operatorname{conv}}\big(f(\mathcal{C})\big),$$

then $\mu_x \stackrel{\text{def}}{=} F(x) + \boldsymbol{\delta}_{s_0} \in \mathsf{P}_1(\mathcal{C})$ and the assumptions of the first half of part (2) of Proposition 136 coincide with the assertions of its second half. ■

The proof of Proposition 136 shows that even though in the first parts of the two equivalences in Proposition 136 one assumes merely the existence of an F with the desired properties, it follows that such an F can in fact be chosen to depend *linearly* on the input f, per (5.6).

Due to Proposition 136, the following question is closely related to Conjecture 134, though we think that it is also of independent interest.

Question 137. Characterize those Polish metric spaces $(\mathfrak{M}, d_{\mathfrak{M}})$ for which there exists a Lipschitz mapping $\rho : \mathfrak{F}(\mathfrak{M}) \to \mathsf{P}_1(\mathfrak{M})$ (recall that by default $\mathsf{P}_1(\mathfrak{M})$ is equipped with the Wasserstein-1 metric) and $x_0 \in \mathfrak{M}$ such that $\rho(\delta_y - \delta_{x_0}) = \delta_y$ for every $y \in \mathfrak{M}$.

5.3 Barycentric targets

Following [210], say that a metric space $(\mathfrak{M}, d_{\mathfrak{M}})$ is W_1-*barycentric* with constant $\beta > 0$ if there is a mapping $\mathfrak{B} : \mathsf{P}_1(\mathfrak{M}) \to \mathfrak{M}$ that satisfies $\mathfrak{B}(\boldsymbol{\delta}_x) = x$ for every $x, \in \mathfrak{M}$, and also

$$\forall \mu, \nu \in \mathsf{P}_1(\mathfrak{M}), \quad d_{\mathfrak{M}}\big(\mathfrak{B}(\mu), \mathfrak{B}(\nu)\big) \leqslant \beta \mathsf{W}_1^{d_{\mathfrak{M}}}(\mu, \nu).$$

The infimal β for which this holds is denoted $\beta_1(\mathfrak{M})$. This notion (and variants thereof) were studied in various contexts; see, e.g., [17,33,94,119,165,173,178,210, 241,243,292]. Any normed space $\mathbf{X}$ is W_1-barycentric with constant 1, as seen by considering $\mathfrak{B}(\mu) = \int_{\mathbf{X}} x \,\mathrm{d}\mu(x)$. Other examples of spaces that are W_1-barycentric with constant 1 include Hadamard spaces and Busemann nonpositively curved spaces [57], or more generally spaces with a conical geodesic bicombing [86].

Thanks to Proposition 136, convex hull-valued (approximate) extension theorems automatically generalize to extension theorems for mappings that take value in W_1-barycentric metric spaces.

Proposition 138. *Let $(\mathfrak{M}, d_{\mathfrak{M}})$ be a metric space and let $\mathfrak{C} \subseteq \mathfrak{M}$ be a Polish subset of $\mathfrak{M}$. Fix $\mathfrak{d} : \mathfrak{M} \times \mathfrak{M} \to [0, \infty)$ and $\varepsilon : \mathfrak{C} \to [0, \infty)$. Assume that for every Banach space $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$ and every $f : \mathfrak{C} \to \mathbf{Z}$ that is 1-Lipschitz with respect to $d_{\mathfrak{M}}$ there is $F : \mathfrak{M} \to \overline{\mathrm{conv}}(f(\mathfrak{C}))$ that satisfies*

$$\forall s \in \mathfrak{C}, \quad \|F(s) - f(s)\|_{\mathbf{Z}} \leqslant \varepsilon(s)$$

and

$$\forall x, y \in \mathfrak{M}, \quad \|F(x) - F(y)\|_{\mathbf{Z}} \leqslant \mathfrak{d}(x, y).$$

Fix $\eta : \mathfrak{C} \to (1, \infty)$ and $\tau : \mathfrak{M} \times \mathfrak{M} \to (1, \infty)$, as well as $\beta > 0$ and a concave nondecreasing function $\omega : [0, \infty) \to [0, \infty)$ with $\omega(0) = 0$. If $(\mathfrak{N}, d_{\mathfrak{N}})$ is a W_1-barycentric metric space with constant β and $\phi : \mathfrak{C} \to \mathfrak{N}$ has modulus of uniform continuity ω with respect to $d_{\mathfrak{M}}$, namely $d_{\mathfrak{N}}(f(s), f(t)) \leqslant \omega(d_{\mathfrak{M}}(s,t))$ for every $s, t \in \mathfrak{C}$, then there is $\Phi : \mathfrak{M} \to \mathfrak{N}$ such that $d_{\mathfrak{N}}(\Phi(s), \phi(s)) \leqslant \omega(\eta(s)\varepsilon(s))$ for every $s \in \mathfrak{C}$ and $d_{\mathfrak{N}}(\Phi(x), \Phi(y)) \leqslant \omega(\tau(x,y)\mathfrak{d}(x,y))$ for every $x, y \in \mathfrak{M}$.

Proof. By Proposition 136, there is a collection of measures $\{\mu_x\}_{x \in \mathfrak{M}} \subseteq \mathsf{P}_1(\mathfrak{C})$ such that

$$\forall s \in \mathfrak{C}, \quad \mathsf{W}_1^{d_{\mathfrak{M}}}(\mu_s, \boldsymbol{\delta}_s) \leqslant \varepsilon(s) \quad \text{and} \quad \forall x, y \in \mathfrak{M} \quad \mathsf{W}_1^{d_{\mathfrak{M}}}(\mu_x, \mu_y) \leqslant \mathfrak{d}(x, y).$$

Hence, for every $s \in \mathfrak{C}$ and $x, y \in \mathfrak{M}$ there are couplings $\pi_s \in \Pi(\mu_s, \boldsymbol{\delta}_s)$ and $\pi_{x,y} \in \Pi(\mu_x, \mu_y)$ such that

$$\iint_{\mathfrak{C} \times \mathfrak{C}} d_{\mathfrak{M}}(u, v) \,\mathrm{d}\pi_s(u, v) \leqslant \eta(s)\varepsilon(s)$$

and

$$\iint_{\mathcal{C}\times\mathcal{C}} d_{\mathfrak{M}}(u,v)\,\mathrm{d}\pi_{x,y}(u,v) \leqslant \tau(x,y)\mathfrak{d}(x,y).$$

Since $(\phi\times\phi)_\#\pi_s \in \Pi(\phi_\#\mu_s, \phi_\#\boldsymbol{\delta}_s)$ and $(\phi\times\phi)_\#\pi_{x,y} \in \Pi(\phi_\#\mu_x, \phi_\#\mu_y)$, it follows that

$$\begin{aligned}
\mathsf{W}_1^{d_{\mathfrak{N}}}(\phi_\#\mu_s, \phi_\#\boldsymbol{\delta}_s) &\leqslant \iint_{\mathfrak{N}\times\mathfrak{N}} d_{\mathfrak{N}}(a,b)\,\mathrm{d}(\phi\times\phi)_\#\pi_s(a,b)\\
&= \iint_{\mathfrak{N}\times\mathfrak{N}} d_{\mathfrak{N}}\big(\phi(u),\phi(v)\big)\,\mathrm{d}\pi_s(u,v)\\
&\leqslant \iint_{\mathfrak{N}\times\mathfrak{N}} \omega\big(d_{\mathfrak{N}}(u,v)\big)\,\mathrm{d}\pi_s(u,v)\\
&\leqslant \omega\bigg(\iint_{\mathfrak{N}\times\mathfrak{N}} d_{\mathfrak{N}}(u,v)\,\mathrm{d}\pi_s(u,v)\bigg)\\
&\leqslant \omega\big(\eta(s)\varepsilon(s)\big),
\end{aligned}$$

where the penultimate step uses the concavity of ω. For the same reason, also

$$\mathsf{W}_1^{d_{\mathfrak{N}}}(\phi_\#\mu_x, \phi_\#\mu_y) \leqslant \omega\big(\tau(x,y)\mathfrak{d}(x,y)\big).$$

Since $(\mathfrak{N}, d_{\mathfrak{N}})$ is β-barycentric there is $\mathfrak{B}:\mathsf{P}_1(\mathfrak{N})\to\mathfrak{N}$ satisfying $\mathfrak{B}(\boldsymbol{\delta}_z)=z$ for every $z,\in\mathfrak{N}$, and

$$\forall \nu_1,\nu_2\in\mathsf{P}_1(\mathfrak{N}),\quad d_{\mathfrak{N}}\big(\mathfrak{B}(\nu_1),\mathfrak{B}(\nu_2)\big)\leqslant \beta\mathsf{W}_1^{d_{\mathfrak{N}}}(\nu_1,\nu_2).$$

Define $\Phi:\mathfrak{M}\to\mathfrak{N}$ by

$$\forall x\in\mathfrak{M},\quad \Phi(x)\stackrel{\mathrm{def}}{=}\mathfrak{B}(\phi_\#\mu_x).$$

Then, for every $s\in\mathcal{C}$ we have

$$d_{\mathfrak{N}}\big(\Phi(s),\phi(s)\big)\leqslant \beta\mathsf{W}_1^{d_{\mathfrak{N}}}\big(\phi_\#\mu_s,\phi_\#\boldsymbol{\delta}_s\big)\leqslant\omega\big(\eta(s)\varepsilon(s)\big).$$

For the same reason also $d_{\mathfrak{N}}(\Phi(x),\phi(y))\leqslant\omega(\tau(x,y)\mathfrak{d}(x,y))$ for every $x,y\in\mathfrak{M}$. ■

Because (as we will soon see) all of our new Lipschitz extension theorems are in fact bounds on $\mathsf{e}_{\mathrm{conv}}(\cdot)$, the following immediate corollary of Proposition 138 (with $\mathfrak{d}$ a multiple of $d_{\mathfrak{M}}$ and ω linear) shows that they apply to barycentric targets and not only to Banach space targets.

Corollary 139. *Fix $\beta>0$. Suppose that $\mathfrak{M}$ is a Polish metric space and that $\mathfrak{N}$ is a complete W_1-barycentric metric space with constant β. Then,*

$$\mathsf{e}_{\mathrm{conv}}(\mathfrak{M},\mathfrak{N})\leqslant\beta\mathsf{e}_{\mathrm{conv}}(\mathfrak{M}).$$

Another noteworthy special case of Proposition 138 is when $\omega(s) = s^\theta$ for some $0 < \theta \leqslant 1$, i.e., in the setting of Hölder extension that we discussed in Remark 15 and Section 2.3. Analogously to (1.18), we denote the convex hull-valued θ-Hölder extend modulus of a metric space $(\mathfrak{M}, d_{\mathfrak{M}})$ by

$$\mathsf{e}^\theta_{\mathrm{conv}}(\mathfrak{M}) = \mathsf{e}_{\mathrm{conv}}\big(\mathfrak{M}, d^\theta_{\mathfrak{M}}\big).$$

Corollary 140. *Suppose that $\mathfrak{M}$ is a Polish metric space. Then, for every $0 < \theta \leqslant 1$ we have*

$$\mathsf{e}^\theta(\mathfrak{M}) \leqslant \mathsf{e}^\theta_{\mathrm{conv}}(\mathfrak{M}) \leqslant \mathsf{e}_{\mathrm{conv}}(\mathfrak{M})^\theta.$$

Because the upper bound on $\mathsf{e}(\ell_\infty^n)$ that we obtain in Theorem 14 is actually an upper bound on $\mathsf{e}_{\mathrm{conv}}(\ell_\infty^n)$, Corollary 140 implies (1.19). More generally, Proposition 138 implies that

$$\mathsf{e}_{\mathrm{conv}}\big(\mathfrak{M}, \omega \circ d_{\mathfrak{M}}\big) \leqslant \sup_{d>0} \frac{\omega\big(\mathsf{e}_{\mathrm{conv}}(\mathfrak{M})d\big)}{\omega(d)} \leqslant \mathsf{e}_{\mathrm{conv}}(\mathfrak{M})$$

for any concave nondecreasing function $\omega : [0,\infty) \to [0,\infty)$ with $\omega(0) = 0$.

Remark 141. The question of how Lipschitz extension results imply extension results for other moduli of uniform continuity was studied in [224] and treated definitively by Brudnyi and Shvartsman in [65] using an interesting connection to the Brudnyĭ–Krugljak K-divisibility theorem [66] (see also [82]) from the theory of real interpolation of Banach spaces. In particular, by [65] we have $\mathsf{e}^\theta(\mathfrak{M}) \lesssim \mathsf{e}(\mathfrak{M})^2$, which remains the best-known bound on $\mathsf{e}^\theta(\mathfrak{M})$ in terms of $\mathsf{e}(\mathfrak{M})$ and it would be interesting to determine if it could be improved. As Corollary 140 shows that a better bound is available in terms of $\mathsf{e}^\theta_{\mathrm{conv}}(\mathfrak{M})$, Conjecture 134 and Question 137 could be relevant for this purpose.

5.4 Gentle partitions of unity

The following definition describes a numerical parameter that underlies the extension method of [173].

Definition 142 (Modulus of gentle partition of unity). Suppose that $(\mathfrak{M}, d_{\mathfrak{M}})$ is a metric space and that $\mathcal{C} \subseteq \mathfrak{M}$ is nonempty and closed. Define the *modulus of gentle partition of unity* of $\mathfrak{M}$ relative to $\mathcal{C}$, denoted $\mathsf{GPU}(\mathfrak{M}, d_{\mathfrak{M}}; \mathcal{C})$ or simply $\mathsf{GPU}(\mathfrak{M}; \mathcal{C})$ when the metric is clear from the context, to be the infimum over those $\mathfrak{g} \in (0,\infty]$ such that for every $x \in \mathfrak{M}$ there is a Borel probability measure μ_x supported on $\mathcal{C}$ with the requirements that if $s \in \mathcal{C}$, then $\mu_s = \boldsymbol{\delta}_s$, and also for every $x, y \in \mathfrak{M}$ we have

$$\int_{\mathcal{C}} d_{\mathfrak{M}}(s,x)\, \mathrm{d}|\mu_x - \mu_y|(s) \leqslant \mathfrak{g}\, d_{\mathfrak{M}}(x,y).$$

The modulus of gentle partitions of unity of $\mathfrak{M}$, denoted $\mathsf{GPU}(\mathfrak{M}, d_{\mathfrak{M}})$ or simply $\mathsf{GPU}(\mathfrak{M})$ when the metric is clear from the context, is defined to be the supremum of $\mathsf{GPU}(\mathfrak{M}, d_{\mathfrak{M}}; \mathcal{C})$ over all nonempty closed subsets $\mathcal{C} \subseteq \mathfrak{M}$.

The nomenclature of Definition 142 is derived from [173], though we warn that Definition 142 considers objects that are not identical to those that were introduced in [173]. In [173] the measures $\{\mu_x\}_{x\in\mathfrak{M}\smallsetminus\mathcal{C}}$ were also required to have a Radon–Nikoým derivative with respect to some reference measure μ. This additional requirement arises automatically from the constructions of [173] but it is not needed for any of the known applications of gentle partitions of unity, so it is beneficial to remove it altogether. The formal connection between [173] and Definition 142 was clarified in [11].

In anticipation of the proof of Theorem 66, one can generalize Definition 142 to the case of general profiles, analogously to what we did in Definition 64.

Definition 143 (Gentle partition of unity profile). Suppose that $(\mathfrak{M}, d_{\mathfrak{M}})$ is a metric space and that $\mathcal{C} \subseteq \mathfrak{M}$ is nonempty and closed. A metric $\mathfrak{d} : \mathfrak{M} \times \mathfrak{M} \to [0, \infty)$ is called a *gentle partition of unity profile* for $(\mathfrak{M}, d_{\mathfrak{M}})$ relative to $\mathcal{C}$ if for every $x \in \mathfrak{M}$ there is a Borel probability measure μ_x supported on $\mathcal{C}$ with the requirements that if $s \in \mathcal{C}$, then $\mu_s = \boldsymbol{\delta}_s$, and also for every $x, y \in \mathfrak{M}$ we have

$$\int_{\mathcal{C}} d_{\mathfrak{M}}(s, x)\, \mathrm{d}|\mu_x - \mu_y|(s) \leqslant \mathfrak{d}(x, y).$$

If $\mathfrak{d}$ is a gentle partition of unity profile for $(\mathfrak{M}, d_{\mathfrak{M}})$ relative to every closed subset $\varnothing \neq \mathcal{C} \subseteq \mathfrak{M}$, then we say that $\mathfrak{d}$ is a gentle partition of unity profile for $(\mathfrak{M}, d_{\mathfrak{M}})$.

Note in passing that if $\mathfrak{d}$ is a gentle partition of unity profile for $(\mathfrak{M}, d_{\mathfrak{M}})$ relative to $\mathcal{C}$, then for every $x \in \mathfrak{M}$ the probability measure μ_x in Definition 143 has finite first moment. Indeed, for any $s_0 \in \mathcal{C}$,

$$\begin{aligned}\int_{\mathcal{C}} d_{\mathfrak{M}}(s_0, s)\, \mathrm{d}\mu_x(s) &= \int_{\mathcal{C}} d_{\mathfrak{M}}(s_0, s)\, \mathrm{d}\big(\mu_x - \delta_{s_0}\big)(s)\\ &\leqslant \int_{\mathcal{C}} d_{\mathfrak{M}}(s_0, s)\, \mathrm{d}\big|\mu_x - \mu_{s_0}\big|(s) \leqslant \mathfrak{d}(s_0, x) < \infty, \end{aligned} \tag{5.7}$$

where we used the fact that $\mu_{s_0} = \delta_{s_0}$, since $s_0 \in \mathcal{C}$.

Suppose that $(\mathfrak{M}, d_{\mathfrak{M}})$ is a Polish metric space. The following estimate is implicit in [173]:

$$\mathsf{e}_{\mathrm{conv}}(\mathfrak{M}) \leqslant 2\mathsf{GPU}(\mathfrak{M}).$$

In fact, the same reasoning as in [173] leads to the following more general lemma.

Lemma 144. *Suppose that $(\mathfrak{M}, d_{\mathfrak{M}})$ is a Polish metric space and that $\mathcal{C} \subseteq \mathfrak{M}$ is nonempty and closed. Assume that $\mathfrak{d} : \mathfrak{M} \times \mathfrak{M} \to [0, \infty)$ is a gentle partition of unity* *profile for* $(\mathfrak{M}, d_{\mathfrak{M}})$ *relative to* $\mathfrak{C}$. *Then, for every Banach space* $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$ *and every* 1*-Lipschitz mapping* $f : \mathfrak{C} \to \mathbf{Z}$ *there exists*

$$F : \mathfrak{M} \to \overline{\mathrm{conv}}\big(f(\mathfrak{C})\big)$$

that extends f *and satisfies* $\|F(x) - F(y)\|_{\mathbf{Z}} \leqslant 2\mathfrak{d}(x, y)$ *for every* $x, y \in \mathfrak{M}$.

Proof. Let $\{\mu_x\}_{x\in\mathfrak{M}}$ be probability measures as in Definition 143. Then, $\{\mu_x\}_{x\in\mathfrak{M}} \subseteq \mathsf{P}_1(\mathfrak{C})$ by (5.7). So, by Proposition 136 (with $\varepsilon \equiv 0$) it suffices to check that for every $x, y \in \mathfrak{M}$ we have $\mathsf{W}_1(\mu_x, \mu_y) \leqslant 2\mathfrak{d}(x, y)$. To this end, fix $\eta > 0$ and $s_0 \in \mathfrak{C}$ such that $d_{\mathfrak{M}}(x, s_0) \leqslant d_{\mathfrak{M}}(x, \mathfrak{C}) + \eta$. Then, for every $s \in \mathfrak{C}$ we have

$$d_{\mathfrak{M}}(s, s_0) \leqslant d_{\mathfrak{M}}(s, x) + d_{\mathfrak{M}}(x, s_0) \leqslant d_{\mathfrak{M}}(s, x) + d_{\mathfrak{M}}(x, \mathfrak{C}) + \eta \leqslant 2d_{\mathfrak{M}}(s, x) + \eta.$$

Consequently, every 1-Lipschitz function $\psi : \mathfrak{C} \to \mathbb{R}$ satisfies

$$\begin{aligned}
\int_{\mathfrak{C}} \psi \, \mathrm{d}\mu_x - \int_{\mathfrak{C}} \psi \, \mathrm{d}\mu_y &= \int_{\mathfrak{C}} \big(\psi(s) - \psi(s_0)\big) \, \mathrm{d}(\mu_x - \mu_y)(s) \\
&\leqslant \int_{\mathfrak{C}} |\psi(s) - \psi(s_0)| \, \mathrm{d}|\mu_x - \mu_y|(s) \\
&\leqslant \int_{\mathfrak{C}} d_{\mathfrak{M}}(s, s_0) \, \mathrm{d}|\mu_x - \mu_y|(s) \\
&\leqslant \int_{\mathfrak{C}} (2d_{\mathfrak{M}}(s, x) + \eta) \, \mathrm{d}|\mu_x - \mu_y|(s) \\
&\leqslant 2\mathfrak{d}(x, y) + 2\eta.
\end{aligned}$$

The desired conclusion follows by letting

$$\eta \to 0$$

and using the Kantorovich–Rubinstein duality (5.1). ∎

5.5 The multi-scale construction

Suppose that $(\mathfrak{M}, d_{\mathfrak{M}})$ is a Polish metric space and fix another metric $\mathfrak{d}$ on $\mathfrak{M}$. In this section we will show that there is a universal constant $\alpha \geqslant 1$ with the following property. Assume that either $(\mathfrak{M}, d_{\mathfrak{M}})$ is locally compact and $\mathfrak{d}$ is a separation modulus for $(\mathfrak{M}, d_{\mathfrak{M}})$ per Definition 64, or the assumptions of Theorem 114 are satisfied. We will prove that either of these assumptions implies that $\alpha\mathfrak{d}$ is a gentle partition of unity profile for $(\mathfrak{M}, d_{\mathfrak{M}})$. By Lemma 144 this gives Theorems 66 and 114, and will show that in fact these extension results are both convex hull-valued and via a linear extension operator. This also implies that every locally compact metric space $\mathfrak{M}$ satisfies

$$\mathsf{GPU}(\mathfrak{M}) \lesssim \mathsf{SEP}(\mathfrak{M}). \tag{5.8}$$

Remark 145. The bound (5.8) need not be sharp. Indeed, it was proved in [173] that if $\mathfrak{M}$ is finite, then

$$\mathsf{GPU}(\mathfrak{M}) \lesssim \frac{\log|\mathfrak{M}|}{\log\log|\mathfrak{M}|}. \tag{5.9}$$

However, by [29] sometimes $\mathsf{SEP}(\mathfrak{M}) \gtrsim \log|\mathfrak{M}|$ (and always $\mathsf{SEP}(\mathfrak{M}) \lesssim \log|\mathfrak{M}|$). A shorter presentation of the proof of (5.9) can be found in [226], and a different proof of (5.9) will appear in the forthcoming work [207]. Also, in the forthcoming work [212] it is proved that (5.9) is optimal.

The following theorem is a precise formulation of what we will prove in this section.

Theorem 146. *Let $(\mathfrak{M}, d_{\mathfrak{M}})$ be a Polish metric space and fix another metric $\mathfrak{d}$ on $\mathfrak{M}$. Suppose that for every $\Delta > 0$ there is a probability space $(\Omega_\Delta, \mathbf{Prob}_\Delta)$ and a sequence of set-valued mappings $\{\Gamma_\Delta^k : \Omega_\Delta \to 2^{\mathfrak{M}}\}_{k=1}^\infty$ such that one of the following two measurability assumptions hold.*

- *Either $(\mathfrak{M}, d_{\mathfrak{M}})$ is locally compact and Γ_Δ^k is strongly measurable for each fixed $k \in \mathbb{N}$ and $\Delta > 0$,*
- *or Ω_Δ is a Borel subset of some Polish metric space $\mathfrak{Z}_\Delta$ and $\mathbf{Prob}_\Delta$ is a Borel probability measure supported on Ω_Δ, and Γ_Δ^k is a standard set-valued mapping for each fixed $k \in \mathbb{N}$ and $\Delta > 0$.*

Suppose that the following three requirements hold.

(1) $\mathcal{P}_\Delta^\omega = \{\Gamma_\Delta^k(\omega)\}_{k=1}^\infty$ *is a partition of $\mathfrak{M}$ for every $\omega \in \Omega_\Delta$,*
(2) $\mathrm{diam}_{\mathfrak{M}}(\mathcal{P}_\Delta^\omega(x)) < \Delta$ *for every $x \in \mathfrak{M}$ and $\omega \in \Omega_\Delta$,*
(3) $\Delta\mathbf{Prob}_\Delta\big[\omega \in \Omega_\Delta : \mathcal{P}_\Delta^\omega(x) \neq \mathcal{P}_\Delta^\omega(y)\big] \leqslant \mathfrak{d}(x, y)$ *for every $x, y \in \mathfrak{M}$.*

Then, $\alpha\mathfrak{d}$ is a gentle partition of unity profile for $(\mathfrak{M}, d_{\mathfrak{M}})$ for some universal constant $\alpha \in [1, \infty)$.

Suppose from now on that $\mathfrak{C}$ is a nonempty *closed* subset of $\mathfrak{M}$. We will first set notation and record basic properties of a sequence of bump functions that will be used in the proof of Theorem 146; this part of the discussion is entirely standard and has nothing to do with random partitions.

Fix a 1-Lipschitz function $\psi : [0, \infty) \to [0, \infty)$ such that $\mathrm{supp}(\psi) \subseteq [1, 4]$ and $\psi(t) = 1$ for every $t \in [2, 3]$ (these requirements uniquely determine ψ, which is piecewise linear). Define for each $n \in \mathbb{Z}$,

$$\forall x \in \mathfrak{M}, \quad \phi_n(x) = \phi_n^{\mathfrak{C}}(x) \stackrel{\mathrm{def}}{=} \psi\big(2^{-n} d_{\mathfrak{M}}(x, \mathfrak{C})\big).$$

Then $\|\phi_n\|_{\mathrm{Lip}(\mathfrak{M})} \leqslant 2^{-n}$ and if $\phi_n(x) \neq 0$ then necessarily $2^n \leqslant d_{\mathfrak{M}}(x, \mathfrak{C}) \leqslant 2^{n+2}$. We also denote

$$\forall x \in \mathfrak{M}, \quad \Phi(x) = \Phi^{\mathfrak{C}}(x) \stackrel{\mathrm{def}}{=} \sum_{m \in \mathbb{Z}} \phi_n(x).$$

For each $x \in \mathfrak{M}$, at most two summands in the sum that defines $\Phi(x)$ do not vanish. If $x \in \mathfrak{M} \smallsetminus \mathcal{C}$, then since $\mathcal{C}$ is closed we have $d_{\mathfrak{M}}(x, \mathcal{C}) > 0$, and therefore there is $n \in \mathbb{Z}$ for which $2^n \leqslant d_{\mathfrak{M}}(x, \mathcal{C}) < 2^{n+1}$. For this value of n we have $\phi_n(x) = 1$, so $\Phi(x) \geqslant 1$ for every $x \in \mathfrak{M} \smallsetminus \mathcal{C}$. Finally, for each $n \in \mathbb{Z}$ define

$$\forall x \in \mathfrak{M}, \quad \lambda_n(x) = \lambda_n^{\mathcal{C}}(x) \stackrel{\mathrm{def}}{=} \begin{cases} \frac{\phi_n(x)}{\Phi(x)} & \text{if } x \in \mathfrak{M} \smallsetminus \mathcal{C}, \\ 0 & \text{if } x \in \mathcal{C}. \end{cases}$$

By design, $\sum_{n\in\mathbb{Z}} \lambda_n(x) = 1$ for every $x \in \mathfrak{M} \smallsetminus \mathcal{C}$. Further properties of these bump functions are recorded in the following basic lemma, for ease of later reference.

Lemma 147. *Suppose that $x, y \in \mathfrak{M}$ satisfy $d_{\mathfrak{M}}(x, \mathcal{C}) \geqslant d_{\mathfrak{M}}(y, \mathcal{C}) > d_{\mathfrak{M}}(x, y)$. Then for every $n \in \mathbb{Z}$,*

$$\frac{2^n}{d_{\mathfrak{M}}(y, \mathcal{C})} \notin \left(\frac{1}{4}, 2\right) \implies \phi_n(x) = \phi_n(y) = \lambda_n(x) = \lambda_n(y) = 0 \qquad (5.10)$$

and

$$2^{n-1} < d_{\mathfrak{M}}(y, \mathcal{C}) < 2^{n+2} \implies \left|\lambda_n(x) - \lambda_n(y)\right| \lesssim \frac{d_{\mathfrak{M}}(x, y)}{d_{\mathfrak{M}}(y, \mathcal{C})}. \qquad (5.11)$$

Proof. Our assumption implies that $d_{\mathfrak{M}}(x, \mathcal{C}), d_{\mathfrak{M}}(y, \mathcal{C}) > 0$, so $x, y \in \mathfrak{M} \smallsetminus \mathcal{C}$. To prove (5.10), suppose first that $2^n \geqslant 2d_{\mathfrak{M}}(y, \mathcal{C})$. Then, since $\mathrm{supp}(\psi) \subseteq [1, 4]$ and $2^{-n} d_{\mathfrak{M}}(y, \mathcal{C}) \leqslant 1$ we have $\phi_n(y) = \lambda_n(y) = 0$. Also,

$$d_{\mathfrak{M}}(x, \mathcal{C}) \leqslant d_{\mathfrak{M}}(x, y) + d_{\mathfrak{M}}(y, \mathcal{C}) < 2d_{\mathfrak{M}}(y, \mathcal{C}) \leqslant 2^n,$$

so $2^{-n} d_{\mathfrak{M}}(x, \mathcal{C}) \leqslant 1$ and hence $\phi_n(x) = \lambda_n(x) = 0$. The remaining case of (5.10) is when $d_{\mathfrak{M}}(y, \mathcal{C}) \geqslant 2^{n+2}$. When this holds we have $2^{-n} d_{\mathfrak{M}}(x, \mathcal{C}) \geqslant 2^{-n} d_{\mathfrak{M}}(y, \mathcal{C}) \geqslant 4$ and therefore $\{2^{-n} d_{\mathfrak{M}}(x, \mathcal{C}), 2^{-n} d_{\mathfrak{M}}(y, \mathcal{C})\} \cap \mathrm{supp}(\psi) = \emptyset$. Consequently, in this case we have $\phi_n(x) = \phi_n(y) = \lambda_n(x) = \lambda_n(y) = 0$.

To prove (5.11), assume that $2^{n-1} < d_{\mathfrak{M}}(y, \mathcal{C}) < 2^{n+2}$. Recalling that (pointwise) on $\mathfrak{M} \smallsetminus \mathcal{C}$ we have $\lambda_n = \phi_n/\Phi$ for all $n \in \mathbb{Z}$ and $\Phi \geqslant 1$, and moreover $\|\phi_n\|_{\mathrm{Lip}(\mathfrak{M})} \leqslant 2^{-n}$, we conclude as follows:

$$\begin{aligned} \left|\lambda_n(x) - \lambda_n(y)\right| &\leqslant \frac{\left|\phi_n(x) - \phi_n(y)\right|}{\Phi(x)} + \frac{\phi_n(y)}{\Phi(x)\Phi(y)}\left|\Phi(y) - \Phi(x)\right| \\ &\leqslant 2^{-n} d_{\mathfrak{M}}(x, y) + \sum_{n\in\mathbb{Z}} \left|\phi_n(x) - \phi_n(y)\right| \\ &\stackrel{(5.10)}{\leqslant} 2^{-n} d_{\mathfrak{M}}(x, y) + \sum_{\substack{n\in\mathbb{Z} \\ 2^{n-1} < d_{\mathfrak{M}}(y,\mathcal{C}) < 2^{n+2}}} 2^{-n} d_{\mathfrak{M}}(x, y) \\ &\asymp \frac{d_{\mathfrak{M}}(x, y)}{d_{\mathfrak{M}}(y, \mathcal{C})}. \end{aligned}$$

■

The interaction between $\{\lambda_n\}_{n\in\mathbb{Z}}$ and the random partitions of Theorem 146 is the content of the following lemma. Note that by reasoning as in (1.94), the metric $\mathfrak{d}$ in Theorem 146 must satisfy

$$\forall x, y \in \mathfrak{M}, \quad \mathfrak{d}(x,y) \geqslant d_{\mathfrak{M}}(x,y).$$

Lemma 148. *In the setting of Theorem 146, if $x \in \mathfrak{M} \smallsetminus \mathcal{C}$ and $y \in \mathfrak{M} \smallsetminus \{x\}$ satisfy $d_{\mathfrak{M}}(x,\mathcal{C}) \geqslant d_{\mathfrak{M}}(y,\mathcal{C})$, then*

$$\sum_{n\in\mathbb{Z}}\sum_{k=1}^{\infty}\int_{\Omega_{2^n}} \left|\lambda_n(x)\mathbf{1}_{\Gamma^k_{2^n}(\omega)}(x) - \lambda_n(y)\mathbf{1}_{\Gamma^k_{2^n}(\omega)}(y)\right| \mathrm{d}\mathbf{Prob}_{2^n}(\omega) \lesssim \frac{\mathfrak{d}(x,y)}{d_{\mathfrak{M}}(y,\mathcal{C}) + d_{\mathfrak{M}}(x,y)}. \tag{5.12}$$

Proof. As $\sum_{n\in\mathbb{Z}}\lambda_n(x) = \sum_{n\in\mathbb{Z}}\lambda_n(y) = 1$ and

$$\sum_{k=1}^{\infty}\mathbf{1}_{\Gamma^k_{2^n}(\omega)}(x) = \sum_{k=1}^{\infty}\mathbf{1}_{\Gamma^k_{2^n}(\omega)}(y) = 1$$

for every $n \in \mathbb{Z}$ and $\omega \in \Omega_{2^n}$, the left-hand side of (5.12) is at most 2. Since $\mathfrak{d}(x,y) \geqslant d_{\mathfrak{M}}(x,y)$, it follows that (5.12) holds if $d_{\mathfrak{M}}(y,\mathcal{C}) \leqslant d_{\mathfrak{M}}(x,y)$. So, we will assume in the rest of the proof of Lemma 148 that $d_{\mathfrak{M}}(x,y) < d_{\mathfrak{M}}(y,\mathcal{C})$ (thus, in particular, $y \in \mathfrak{M} \smallsetminus \mathcal{C}$), in which case the right-hand side of (5.12) becomes at least a universal constant multiple of the quantity $\mathfrak{d}(x,y)/d_{\mathfrak{M}}(y,\mathcal{C})$.

We claim that for every $n \in \mathbb{Z}$ the following inequality holds for every $\omega \in \Omega_{2^n}$:

$$\sum_{k=1}^{\infty}\left|\lambda_n(x)\mathbf{1}_{\Gamma^k_{2^n}(\omega)}(x) - \lambda_n(y)\mathbf{1}_{\Gamma^k_{2^n}(\omega)}(y)\right| \lesssim \left(2^{-n}d_{\mathfrak{M}}(x,y) + \mathbf{1}_{\{\mathcal{P}^{\omega}_{2m}(x)\neq\mathcal{P}^{\omega}_{2^n}(y)\}}\right)\mathbf{1}_{\{\frac14<\frac{2^n}{d_{\mathfrak{M}}(y,\mathcal{C})}<2\}}. \tag{5.13}$$

Assuming (5.13) for the moment, we will conclude the proof of (5.12) in the remaining case $d_{\mathfrak{M}}(x,y) < d_{\mathfrak{M}}(y,\mathcal{C})$ as follows:

$$\begin{aligned}
&\sum_{n\in\mathbb{Z}}\sum_{k=1}^{\infty}\int_{\Omega_{2^n}} \left|\lambda_n(x)\mathbf{1}_{\Gamma^k_{2^n}(\omega)}(x) - \lambda_n(y)\mathbf{1}_{\Gamma^k_{2^n}(\omega)}(y)\right| \mathrm{d}\mathbf{Prob}_{2^n}(\omega)\\
&\lesssim \sum_{\substack{n\in\mathbb{Z}\\ 2^{n-1}<d_{\mathfrak{M}}(y,\mathcal{C})<2^{n+2}}} \left(2^{-n}d_{\mathfrak{M}}(x,y) + \mathbf{Prob}_{2^n}\left[\left\{\omega\in\Omega_{2^n} : \mathcal{P}^{\omega}_{2^n}(x) \neq \mathcal{P}^{\omega}_{2^n}(y)\right\}\right]\right)\\
&\lesssim \sum_{\substack{n\in\mathbb{Z}\\ 2^{mn-1}<d_{\mathfrak{M}}(y,\mathcal{C})<2^{n+2}}} 2^{-n}\left(d_{\mathfrak{M}}(x,y) + \mathfrak{d}(x,y)\right)\\
&\asymp \frac{\mathfrak{d}(x,y)}{d_{\mathfrak{M}}(y,\mathcal{C})} \asymp \frac{\mathfrak{d}(x,y)}{d_{\mathfrak{M}}(y,\mathcal{C}) + d_{\mathfrak{M}}(x,y)},
\end{aligned}$$

where the first step uses (5.13), the second step is where we used condition (3) of Theorem 146, the penultimate step uses $\mathfrak{d}(x,y) \geqslant d_{\mathfrak{M}}(x,y)$, and in the final step uses the assumption $d_{\mathfrak{M}}(x,y) < d_{\mathfrak{M}}(y,\mathcal{C})$.

It therefore remains to establish (5.13). By Lemma 147, if it is not the case that $2^{n-1} < d_{\mathfrak{M}}(y,\mathcal{C}) < 2^{n+2}$, then $\lambda_n(x) = \lambda_n(y) = 0$, so both sides of (5.13) vanish. Thus, we may assume from now that $2^{n-1} < d_{\mathfrak{M}}(y,\mathcal{C}) < 2^{n+2}$. Under this assumption, if $\mathcal{P}^{\omega}_{2^n}(x) \neq \mathcal{P}^{\omega}_{2^n}(y)$, then the right-hand side of (5.13) is at least 1, while the left-hand side of (5.13) consists of a sum of two numbers, each of which is at most 1. It therefore remains to establish (5.13) when $\mathcal{P}^{\omega}_{2^n}(x) = \mathcal{P}^{\omega}_{2^n}(y)$ (and still $2^{n-1} < d_{\mathfrak{M}}(y,\mathcal{C}) < 2^{n+2}$). In this case, (5.13) becomes the inequality $|\lambda_{2^n}(x) - \lambda_{2^n}(y)| \leqslant d_{\mathfrak{M}}(x,y)/d_{\mathfrak{M}}(y,\mathcal{C})$, which we proved in Lemma 147. ∎

Proof of Theorem 146. By Lemma 115 and Corollary 118, for every $\Delta > 0$ there exists a $\mathbf{Prob}_{\Delta}$-to-Borel measurable mapping $\gamma^k_{\Delta} : \Omega_m \to \mathcal{C}$ such that

$$\forall \omega \in \Omega_{\Delta}, \quad \Gamma^k_{\Delta}(\omega) \neq \varnothing \implies d_{\mathfrak{M}}\big(\gamma^k_{\Delta}(\omega), \Gamma^k_{\Delta}(\omega)\big) \leqslant d_{\mathfrak{M}}\big(\mathcal{C}, \Gamma^k_{\Delta}(\omega)\big) + \Delta. \quad (5.14)$$

(In fact, in the locally compact setting of Theorem 146, the use of Lemma 115 shows that the additive Δ term in the right-hand side of (5.14) can be removed).

For every $x \in \mathfrak{M} \smallsetminus \mathcal{C}$ define a Borel measure μ_x supported on $\mathcal{C}$ by

$$\mu_x \stackrel{\text{def}}{=} \sum_{n\in\mathbb{Z}} \sum_{k=1}^{\infty} \lambda_n(x) \big(\gamma^k_{2^n}\big)_{\#}\big(\mathbf{Prob}_{2^n} \lfloor_{\{\omega\in\Omega_{2^n} : x \in \Gamma^k_{2^n}(\omega)\}}\big). \quad (5.15)$$

In other words, for every Borel-measurable mapping $h : \mathcal{C} \to [0,\infty)$ we have

$$\int_{\mathcal{C}} h(s)\,\mathrm{d}\mu_x(s) = \sum_{n\in\mathbb{Z}} \sum_{k=1}^{\infty} \lambda_n(x) \int_{\{\omega\in\Omega_{2^n} : x\in\Gamma^k_{2^n}(\omega)\}} h\big(\gamma^k_{2^n}(\omega)\big)\,\mathrm{d}\mathbf{Prob}_{2^n}(\omega). \quad (5.16)$$

Since $\mathcal{P}^{\omega}_{2^n}$ is a partition of X for every $n \in \mathbb{Z}$ and $\omega \in \Omega_{2^n}$, the special case $h = \mathbf{1}_{\mathcal{C}}$ of (5.16) implies that

$$\begin{aligned}\mu_x(\mathcal{C}) &= \sum_{n\in\mathbb{Z}} \sum_{k=1}^{\infty} \lambda_n(x) \mathbf{Prob}_{2^n}\big[\{\omega \in \Omega_{2^n} : x \in \Gamma^k_{2^n}(\omega)\}\big] \\ &= \sum_{n\in\mathbb{Z}} \lambda_n(x) \mathbf{Prob}_{2^n}\left[\left\{\omega \in \Omega_{2^n} : x \in \bigcup_{k=1}^{\infty} \Gamma^k_{2^n}(\omega)\right\}\right] = \sum_{n\in\mathbb{Z}} \lambda_n(x) = 1.\end{aligned}$$

Thus μ_x is a probability measure. Consequently, if we also denote $\mu_s = \boldsymbol{\delta}_s$ for every $s \in \mathcal{C}$, then the proof of Theorem 146 will be complete if we show that

$$\forall x, y \in \mathfrak{M}, \qquad \int_{\mathcal{C}} d_{\mathfrak{M}}(s,x)\,\mathrm{d}|\mu_x - \mu_y|(s) \lesssim \mathfrak{d}(x,y). \quad (5.17)$$

It suffices to prove (5.17) when $x, y \in \mathfrak{M}$ are distinct and $\{x, y\} \not\subseteq \mathcal{C}$. Indeed, if $\{x, y\} \subseteq \mathcal{C}$ then $\mu_x = \boldsymbol{\delta}_x$ and $\mu_y = \boldsymbol{\delta}_y$, so the left-hand side of (5.17) is equal to $d_{\mathfrak{M}}(x, y)$, which is at most $\mathfrak{d}(x, y)$. Hence, in the rest of the proof of Theorem 146 we will assume without loss of generality that $x \in \mathfrak{M} \smallsetminus \mathcal{C}$ and $d_{\mathfrak{M}}(x, \mathcal{C}) \geqslant d_{\mathfrak{M}}(y, \mathcal{C})$.

We claim that the left-hand side of (5.17) can be bounded from above as follows:

$$\int_{\mathcal{C}} d_{\mathfrak{M}}(s, x)\, \mathrm{d}|\mu_x - \mu_y|(s) \leqslant d_{\mathfrak{M}}(x, y)$$
$$+ \sum_{n\in\mathbb{Z}} \sum_{k=1}^{\infty} \int_{\Omega_{2^n}} d_{\mathfrak{M}}\big(\gamma_{2^n}^k(\omega), x\big)\big|\lambda_n(x)\mathbf{1}_{\Gamma_{2^n}^k(\omega)}(x) - \lambda_{2^n}(y)\mathbf{1}_{\Gamma_{2^n}^k(\omega)}(y)\big|\, \mathrm{d}\mathbf{Prob}_{2^n}(\omega). \tag{5.18}$$

Indeed, if $x, y \in \mathfrak{M} \smallsetminus \mathcal{C}$, then μ_x, μ_y are defined according to (5.15), so that

$$\int_{\mathcal{C}} d_{\mathfrak{M}}(s, x)\, \mathrm{d}|\mu_x - \mu_y|(s)$$
$$\leqslant \sum_{n\in\mathbb{Z}} \sum_{k=1}^{\infty} \int_{\mathcal{C}} d_{\mathfrak{M}}(s, x)$$
$$\mathrm{d}\Big(\big(\gamma_{2^n}^k\big)_{\#}\big|\lambda_n(x)\mathbf{Prob}_{2^n}\lfloor_{\{\omega\in\Omega_{2^n}:x\in\Gamma_{2^n}^k(\omega)\}} - \lambda_n(y)\mathbf{Prob}_{2^n}\lfloor_{\{\omega\in\Omega_{2^n}:y\in\Gamma_{2^n}^k(\omega)\}}\big|\Big)(s)$$
$$= \sum_{n\in\mathbb{Z}} \sum_{k=1}^{\infty} \int_{\Omega_{2^n}} d_{\mathfrak{M}}\big(\gamma_{2^n}^k(\omega), x\big)\big|\lambda_n(x)\mathbf{1}_{\Gamma_{2^n}^k(\omega)}(x) - \lambda_n(y)\mathbf{1}_{\Gamma_{2^n}^k(\omega)}(y)\big|\, \mathrm{d}\mathbf{Prob}_{2^n}(\omega),$$

thus establishing (5.18) in this case. The remaining case is when $x \in \mathfrak{M} \smallsetminus \mathcal{C}$ and $y \in \mathcal{C}$, so that μ_x is given in (5.15) and $\mu_y = \boldsymbol{\delta}_y$. We can then use the following (crude) estimate:

$$\int_{\mathcal{C}} d_{\mathfrak{M}}(s, x)\, \mathrm{d}|\mu_x - \mu_y|(s)$$
$$\leqslant \int_{\mathcal{C}} d_{\mathfrak{M}}(s, x)\, \mathrm{d}\mu_y(s) + \int_{\mathcal{C}} d_{\mathfrak{M}}(s, x)\, \mathrm{d}\mu_x(s)$$
$$= d_{\mathfrak{M}}(x, y) + \sum_{n\in\mathbb{Z}} \sum_{k=1}^{\infty} \int_{\Omega_{2^n}} d_{\mathfrak{M}}\big(\gamma_{2^n}^k(\omega), x\big)\lambda_n(x)\mathbf{1}_{\Gamma_{2^n}^k(\omega)}(x)\, \mathrm{d}\mathbf{Prob}_{2^n}(\omega). \tag{5.19}$$

It remains to observe that because $y \in \mathcal{C}$ we have $\lambda_n(y) = 0$ for all $n \in \mathbb{Z}$ and therefore the right-hand side of (5.19) coincides with the right-hand side of (5.18).

Next, we claim that for every $(n, k) \in \mathbb{Z} \times \mathbb{N}$ and every $\omega \in \Omega_{2^n}$ we have

$$d_{\mathfrak{M}}\big(\gamma_{2^n}^k(\omega), x\big)\big|\lambda_n(x)\mathbf{1}_{\Gamma_{2^n}^k(\omega)}(x) - \lambda_n(y)\mathbf{1}_{\Gamma_{2^n}^k(\omega)}(y)\big|$$
$$\lesssim \big(d_{\mathfrak{M}}(y, \mathcal{C}) + d_{\mathfrak{M}}(x, y)\big)\big|\lambda_n(x)\mathbf{1}_{\Gamma_{2^n}^k(\omega)}(x) - \lambda_n(y)\mathbf{1}_{\Gamma_{2^n}^k(\omega)}(y)\big|. \tag{5.20}$$

By substituting the point-wise estimate (5.20) into (5.18) and using $d_{\mathfrak{m}}(x, y) \leqslant \mathfrak{d}(x, y)$ the desired estimate (5.17) follows from Lemma 148, thus completing the proof of Theorem 146.

To verify (5.20), note first that both sides of (5.20) vanish unless $x \in \Gamma^k_{2^n}(\omega)$ or $y \in \Gamma^k_{2^n}(\omega)$ and also, due to Lemma 147, $2^{n-1} < d_{\mathfrak{m}}(y, \mathcal{C}) < 2^{n+2}$. So, assume from now on that

$$\{x, y\} \cap \Gamma^k_{2^n}(\omega) \neq \varnothing \quad \text{and} \quad 2^{n-1} < d_{\mathfrak{m}}(y, \mathcal{C}) < 2^{n+2}. \tag{5.21}$$

Our goal (5.20) then becomes to deduce that

$$d_{\mathfrak{m}}\big(\gamma^k_{2^n}(\omega), x\big) \lesssim d_{\mathfrak{m}}(y, \mathcal{C}) + d_{\mathfrak{m}}(x, y). \tag{5.22}$$

Choose a point $z \in \Gamma^k_m(\omega)$ such that

$$\begin{aligned} d_{\mathfrak{m}}\big(\gamma^k_{2^n}(\omega), z\big) &\leqslant d_{\mathfrak{m}}\big(\gamma^k_{2^n}(\omega), \Gamma^k_{2^n}(\omega)\big) + 2^n \\ &\overset{(5.14)}{=} d_{\mathfrak{m}}\big(\mathcal{C}, \Gamma^k_{2^n}(\omega)\big) + 2^{n+1} \\ &\overset{(5.21)}{\asymp} d_{\mathfrak{m}}\big(\mathcal{C}, \Gamma^k_{2^n}(\omega)\big) + d_{\mathfrak{m}}(y, \mathcal{C}). \end{aligned} \tag{5.23}$$

If $x \in \Gamma^k_{2^n}(\omega)$, then

$$d_{\mathfrak{m}}\big(\mathcal{C}, \Gamma^k_{2^n}(\omega)\big) \leqslant d_{\mathfrak{m}}(x, \mathcal{C}) \leqslant d_{\mathfrak{m}}(x, y) + d_{\mathfrak{m}}(y, \mathcal{C})$$

and

$$d_{\mathfrak{m}}(x, z) \leqslant \operatorname{diam}_{\mathfrak{m}}\big(\Gamma^k_{2^n}(\omega)\big) \leqslant 2^n \overset{(5.21)}{\asymp} d_{\mathfrak{m}}(y, \mathcal{C}).$$

By combining these two estimates with (5.23) and the triangle inequality, we see that

$$d_{\mathfrak{m}}\big(\gamma^k_{2^n}(\omega), x\big) \leqslant d_{\mathfrak{m}}\big(\gamma^k_{2^n}(\omega), z\big) + d_{\mathfrak{m}}(z, x) \lesssim d_{\mathfrak{m}}(x, y) + d_{\mathfrak{m}}(y, \mathcal{C}).$$

Hence, the desired estimate (5.22) holds when $x \in \Gamma^k_{2^n}(\omega)$.

It remains to check (5.22) when $y \in \Gamma^k_{2^n}(\omega)$, in which case we proceed similarly by noting that now

$$d_{\mathfrak{m}}\big(\mathcal{C}, \Gamma^k_{2^n}(\omega)\big) \leqslant d_{\mathfrak{m}}(y, \mathcal{C}),$$

and

$$d_{\mathfrak{m}}(y, z) \leqslant \operatorname{diam}_{\mathfrak{m}}\big(\Gamma^k_{2^n}(\omega)\big) \leqslant 2^n \overset{(5.21)}{\asymp} d_{\mathfrak{m}}(y, \mathcal{C}).$$

By combining these two estimates with (5.23) and the triangle inequality, we conclude that

$$\begin{aligned} d_{\mathfrak{m}}\big(\gamma^k_{2^n}(\omega), x\big) &\leqslant d_{\mathfrak{m}}\big(\gamma^k_{2^n}(\omega), z\big) + d_{\mathfrak{m}}(z, y) + d_{\mathfrak{m}}(y, x) \\ &\lesssim d_{\mathfrak{m}}(y, \mathcal{C}) + d_{\mathfrak{m}}(x, y). \end{aligned}$$

■