Mathematische Annalen (2024) 390:3109–3131
https://doi.org/10.1007/s00208-024-02831-x

Mathematische Annalen

Bi-Lipschitz embeddings of the space of unordered m-tuples with a partial transportation metric

David Bate[1] · Ana Lucía Garcia Pulido[2]

Received: 19 May 2023 / Revised: 7 February 2024 / Accepted: 17 February 2024 /
Published online: 19 March 2024

Abstract

Let $\Omega \subset \mathbb{R}^n$ be non-empty, open and proper. This paper is concerned with $Wb_p(\Omega)$, the space of p-integrable Borel measures on Ω equipped with the *partial* transportation metric introduced by Figalli and Gigli that allows the creation and destruction of mass on $\partial\Omega$. Alternatively, we show that $Wb_p(\Omega)$ is isometric to a subset of Borel measures with the ordinary Wasserstein distance, on the one point completion of Ω equipped with the shortcut metric

$$\delta(x, y) = \min\{\|x - y\|, \operatorname{dist}(x, \partial\Omega) + \operatorname{dist}(y, \partial\Omega)\}.$$

In this article we construct bi-Lipschitz embeddings of the set of unordered m-tuples in $Wb_p(\Omega)$ into Hilbert space. This generalises Almgren's bi-Lipschitz embedding theorem to the setting of optimal partial transport.

1 Introduction

A striking variety of problems in geometry, analysis, combinatorics and a vast number of applications can be neatly formulated in terms of measures and their comparison using transportation metrics. The prototypical transportation metric is the *p-Wasserstein distance* [2]. This is defined between two Borel measures of the same

✉ David Bate
david.bate@warwick.ac.uk

Ana Lucía Garcia Pulido
analucia.garciapulido@stir.ac.uk

1 Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK

2 Division of Computing Science and Mathematics, University of Stirling, Stirling FK9 4LA, UK

total mass on a metric space (X, d) by

$$W_p(\mu, \nu) = \inf_{\gamma} \left(\int_{X \times X} d(x, y)^p \, \mathrm{d}\gamma(x, y) \right)^{\frac{1}{p}}, \tag{1}$$

where $p \geq 1$ and the infimum is taken over all measures γ on $X \times X$ with coordinate projections $\pi_1 \gamma = \mu$ and $\pi_2 \gamma = \nu$. The resulting metric space of p-integrable probability measures equipped with W_p is denoted by $W_p(X)$ (see Definition 2.1).

A drawback of the Wasserstein distance is the requirement that the compared measures must have the same total mass. Recently emerging theories of *optimal partial transport* pertain to the transportation of measures without a mass constraint [8, 16, 20]. This article concerns the following formulation due to Figalli and Gigli [11].

Let Ω be an open non-empty proper subset of X. For *measures* μ and ν on Ω, one defines $Wb_p(\mu, \nu)$ as in (1), but the infimum is taken over measures γ on $\overline{\Omega} \times \overline{\Omega}$ with

$$\pi_1 \gamma|_{\Omega} = \mu \quad \text{and} \quad \pi_2 \gamma|_{\Omega} = \nu.$$

The resulting metric space of p-integrable measures, equipped with Wb_p, will be denoted by $Wb_p(X)$ (see Definition 3.1).

The key property of Wb_p is that $\partial\Omega$ can be used to destroy or create mass, at a cost of transporting it to or from $\partial\Omega$. This allows measures of different total masses to be compared and hence one can construct a metric space consisting of *all* measures, instead of restricting to probability measures. Understanding the interplay between transportation metrics and $\partial\Omega$ is motivated by solving evolution equations with Dirichlet boundary conditions from gradient flows [11, 21]. The metric Wb_p has found further applications such as obtaining new comparison principles for viscosity solutions [13].

A natural approach to study a metric space is to embed it into a well known space, such as a Euclidean or Banach space, as this allows the metric space to inherit geometric properties of the ambient space. Recall that the *distortion* of an injective map f between two metric spaces is $\mathrm{Lip}(f) \cdot \mathrm{Lip}(f^{-1})$, where $\mathrm{Lip}(f)$ is the Lipschitz constant of f; f is *bi-Lipschitz* if it has finite distortion. Since bi-Lipschitz embeddings preserve relative distances, they are central to analysis and metric geometry [18] and have applications to algorithm design [14].

Due to the prominence of the Wasserstein spaces in various areas of mathematics, their embeddability has attracted much attention. The non-embeddability (into L^1) of W_1 over various discrete metric spaces [5] such as the planar grid [19] and Hamming cube [15] is known, as is the non-embeddability of $W_p(\mathbb{R}^3)$ for $p \geq 1$ [4]. The interest in bi-Lipschitz embeddings of the Wasserstein spaces dates back to the work of Almgren [1, 9], forming the foundations of his celebrated partial regularity theorem for area minimising currents. Almgren proved that, for any $m \in \mathbb{N}$, the set of unordered m-tuples of points in $\mathbb{R}^n$,

$$\mathcal{A}_m(\mathbb{R}^n) = \left\{ \sum_{i=1}^{m} [\![x_i]\!] : x_i \in \mathbb{R}^n \ \forall 1 \leq i \leq m \right\}$$

equipped with W_2, bi-Lipschitz embeds into some Euclidean space (see Theorem 2.3). Here and throughout, $[\![x]\!]$ will denote the Dirac mass at x.

In this article we generalise Almgren's embedding to $Wb_2(\Omega)$.

Theorem 1.1 *For $n \in \mathbb{N}$, let $\Omega \subset \mathbb{R}^n$ be non-empty, open and proper. The space $(\mathcal{B}_m(\Omega), Wb_2)$ of unordered tuples of at most m points bi-Lipschitz embeds into Hilbert space. The distortion of our embedding is at most $cm^{n+5/2}$, for some constant $c \geq 1$.*

In general, $Wb_p(\Omega)$ is not a doubling metric space and hence cannot be bi-Lipschitz embedded into any Euclidean space, see Lemma 3.8. Therefore Hilbert space[1] becomes the natural target for an embedding. Note that, since we are not constrained to comparing measures of the same total mass, in Theorem 1.1 we consider unordered tuples of *at most* m-points.

To prove Theorem 1.1, we first show, for $\Omega \subset X$, that $Wb_p(\Omega)$ isometrically embeds into the ordinary p-Wasserstein space of measures on (Ω^*, δ), where Ω^* is the one point completion of Ω equipped with the shortcut metric

$$\delta(x, y) = \min\{\|x - y\|, \operatorname{dist}(x, \partial\Omega) + \operatorname{dist}(y, \partial\Omega)\}$$

for every $x, y \in \Omega$ (see Lemma 3.3). This embedding maps $\mathcal{B}_m(\Omega)$ to $\mathcal{A}_m(\Omega^*)$ and so, in order to prove Theorem 1.1, it remains to construct a bi-Lipschitz embedding of $\mathcal{A}_m(\Omega^*)$ into Hilbert space.

We do this, for $\Omega \subset \mathbb{R}^n$, by considering a Whitney decomposition $\mathcal{C}$ of Ω into cubes. This decomposition is chosen such that, inside any cube $Q \in \mathcal{C}$, the shortcut metric equals the Euclidean metric and consequently

$$\mathcal{A}_m(Q, \delta) = \mathcal{A}_m(Q, \|\ \|). \tag{2}$$

In particular, Almgren's theorem gives an embedding of each $\mathcal{A}_m(Q)$ into some Euclidean space. Despite the fact that any measure can be written as a sum of measures supported on cubes in $\mathcal{C}$, the construction of the required bi-Lipschitz embedding of $\mathcal{A}_m(\Omega^*)$ cannot be obtained simply by restricting to cubes. Indeed, W_p may not even be defined between the restriction of two measures to a cube; even when it is, simple examples show that the optimal transport of the restricted measures may be incomparable to the optimal transport of the original measures.

Our approach uses (2) as the starting point to determine the optimal transport of measures between different cubes, see Sect. 4. From this analysis we construct a bi-Lipschitz embedding of $\mathcal{A}_m(\Omega^*)$ into the ℓ_2-sum of infinitely many copies of $\mathcal{A}_m(\mathbb{R}^{n+1})$, see Theorem 4.12. The proof of Theorem 1.1 is concluded in Sect. 5 by applying Almgren's embedding to each term of the ℓ_2-sum.

We mention an application of Theorem 1.1 to persistence homology. The space of *persistence barcodes* can be viewed as $\cup_m \mathcal{B}_m(U)$ for

$$U = \{(x, y) \in \mathbb{R}^2 : y > x\},$$

[1] We adopt the standard convention that Hilbert space is the unique complete and separable infinite dimensional inner product space, up to isometric isomorphism.

see [10]. Theorem 1.1 shows that the space of persistence barcodes with at most m-points can be bi-Lipschitz embedded into Hilbert space. This answers questions raised by Carrière and Bauer [6]. Prior to our results, it was known that $\mathcal{B}_m(U)$ coarsely embeds into Hilbert space [17]. In fact, Theorem 1.1 applies to the generalised persistence barcodes introduced in [7] whenever the ambient space is Euclidean. Our theorem also holds when $\mathcal{B}_m(\Omega)$ is equipped with any W_p for $p \geq 1$; due to the equivalence of norms on $\mathbb{R}^m$, these metrics are all bi-Lipschitz equivalent.

Finally, we mention that the distortion of any embedding of $\mathcal{B}_m(\Omega)$ into Hilbert space, for $\Omega \subset \mathbb{R}^n$, must necessarily converge to ∞ as m does, see Remark 5.3.

2 Wasserstein distance and Almgren's embedding

Let (X, d) be a complete and separable metric space. We write $\mathcal{M}(X)$ for the set of Borel measures on X and $\mathcal{P}(X)$ for the set of Borel probability measures on X. The Wasserstein space is defined as follows [2, 3].

Definition 2.1 For $\mu, \nu \in \mathcal{M}(X)$ and $p \in [1, \infty)$ define

$$W_p(\mu, \nu) = \inf_{\gamma} \left(\int_{X \times X} d(x, y)^p \, \mathrm{d}\gamma(x, y) \right)^{\frac{1}{p}},$$

where the infimum is taken over all *couplings* $\gamma \in \mathcal{M}(X \times X)$ with coordinate projections $\pi_1 \gamma = \mu$ and $\pi_2 \gamma = \nu$. Note that $W_p(\mu, \nu) < \infty$ only if $\mu(X) = \nu(X)$ as otherwise there does not exist a γ as in Definition 2.1.

Let $\mathcal{P}_p(X)$ be those $\mu \in \mathcal{P}(X)$ with

$$\int_X d(x, x_0)^p \, \mathrm{d}\mu(x) < \infty$$

for some (equivalently all) $x_0 \in X$. Then W_p defines a metric on $\mathcal{P}_p(X)$. Analogous statements hold for the case $p = \infty$, where the L^p integral is replaced by an essential supremum. We write $W_p(X)$ for the set $\mathcal{P}_p(X)$ equipped with W_p.

Definition 2.2 For $m \in \mathbb{N}$, define the space of *unordered m-tuples*

$$\mathcal{A}_m(X) = \left\{ \sum_{i=1}^{m} [\![x_i]\!] : x_i \in X \; \forall 1 \leq i \leq m \right\},$$

equipped with W_2. Note that, on $\mathcal{A}_m(X)$, W_2 equals

$$W_2(p, q) = \min_{\sigma \in \Sigma_m} \sqrt{\sum_{i=1}^{m} d(p_i, q_{\sigma(i)})^2},$$

where $p = \sum_{i=1}^{m} [\![p_i]\!]$ and $q = \sum_{i=1}^{m} [\![q_i]\!]$.

A fundamental step in Almgren's study of area minimising currents was the following bi-Lipschitz embedding.

Theorem 2.3 (Almgren, Theorem 2.1 [9]) *For every $m \in \mathbb{N}$ there exists an $N \in \mathbb{N}$ and a bi-Lipschitz embedding $\xi : \mathcal{A}_m(\mathbb{R}^n) \to \mathbb{R}^N$. By inspecting the proof one sees that $\xi(0) = 0$ and, for all $p, q \in \mathcal{A}_m(\mathbb{R}^n)$,*

$$\frac{W_2(p,q)}{cm^{n+1}} \leq \|\xi(p) - \xi(q)\| \leq W_2(p,q)$$

for a constant $c \geq 1$.

3 Optimal partial transport and the shortcut metric

The transportation metric Wb introduced by Figalli and Gigli [11] is defined between two Borel measures. Originally defined for open and bounded $\Omega \subset \mathbb{R}^n$, we state the natural generalisation of Wb to complete and separable metric spaces (X, d) (the proof of the triangle inequality is identical).

Definition 3.1 Let $\Omega \subset X$ be proper and non-empty. For $\mu, \nu \in \mathcal{M}(\Omega)$ and $p \in [1, \infty)$ define

$$Wb_p(\mu, \nu) = \inf_\gamma \left(\int_{X \times X} d(x,y)^p \,\mathrm{d}\gamma(x,y) \right)^{\frac{1}{p}},$$

where the infimum is taken over all *couplings* $\gamma \in \mathcal{M}(X \times X)$ with $\pi_1\gamma|_\Omega = \mu$ and $\pi_2\gamma|_\Omega = \nu$. Then Wb_p defines a metric on

$$\mathcal{M}b_p(\Omega) := \{\mu \in \mathcal{M}(\Omega) : Wb_p(\mu, 0) < \infty\}.$$

Analogous statements hold for the case $p = \infty$, where the L^p integral is replaced by an essential supremum.

We write $Wb_p(\Omega)$ for the set $\mathcal{M}b_p(\Omega)$ equipped with Wb_p. We also write $Wb_p^1(\Omega)$ for the set of $\mu \in \mathcal{M}b_p(\Omega)$ with $\mu(\Omega) \leq 1$, equipped with Wb_p.

The first step in our proof of Theorem 1.1 is to show an equivalence between $Wb_p^1(\Omega)$ and $W_p(\Omega^*)$, for Ω^* the *shortcut* metric space, defined as the one point completion of Ω via its complement.

Definition 3.2 For $\Omega \subset X$ non-empty and proper, let $\Omega^* = \Omega \cup \{\partial\}$. For $x, y \in \Omega$ define

$$\delta(x,y) = \min\{\|x - y\|, \mathrm{dist}(x, X\backslash\Omega) + \mathrm{dist}(y, X\backslash\Omega)\}$$

and $\delta(x, \partial) = \mathrm{dist}(x, X\backslash\Omega)$. Then δ defines a metric on Ω^*.

Profeta and Sturm [21, Remark 1.9] mention that $Wb_1^1(\Omega)$ isometrically embeds into $W_1(\Omega^*)$, and give an example showing that their embedding is not an isometry for $p > 1$. We show that there exists an isometric embedding of $Wb_p^1(\Omega)$ into $2W_p(\Omega^*)$ for any $p \geq 1$. Here we write $2W_p(\Omega^*)$ for the space of measures with total mass equal to 2.

Lemma 3.3 *Let X be a separable metric space and $\Omega \subset X$ be non-empty and proper. For any $p \geq 1$, the map*

$$\begin{aligned} Wb_p^1(\Omega) &\to 2W_p(\Omega^*) \\ \iota(\mu) &= \mu + (2 - \mu(\Omega))[\![\partial]\!], \end{aligned}$$

is an isometric embedding.

Proof Given a coupling for μ, ν we use it to construct a coupling for $\iota(\mu), \iota(\nu)$ and vice versa.

First let $\mu, \nu \in Wb_p^1(\Omega)$ and suppose that $\gamma \in \mathcal{M}(X \times X)$ is a coupling for μ and ν in $Wb_p(\Omega)$. Let $\pi_\partial(x) = \partial$ for all $x \in X$ and define $\gamma' \in \mathcal{M}(\Omega^* \times \Omega^*)$ as

$$\begin{aligned} \gamma' = {} & \gamma|_{\Omega \times \Omega} + (\pi_\partial \times \mathrm{id})_\# \gamma|_{X \setminus \Omega \times \Omega} + (\mathrm{id} \times \pi_\partial)_\# \gamma|_{\Omega \times X \setminus \Omega} \\ & + (2 - [\gamma(\Omega \times \Omega) + \gamma(X \setminus \Omega \times \Omega) + \gamma(\Omega \times X \setminus \Omega)])[\![(\partial, \partial)]\!]. \end{aligned}$$

For notational convenience, we let κ denote the coefficient of $[\![(\partial, \partial)]\!]$ in this expression. Then

$$\begin{aligned} \pi_1 \gamma' &= \pi_1(\gamma|_{\Omega \times \Omega}) + \gamma(X \setminus \Omega \times \Omega)[\![\partial]\!] + \pi_1(\gamma|_{\Omega \times X \setminus \Omega}) + \kappa[\![\partial]\!] \\ &= \pi_1(\gamma|_{\Omega \times \Omega}) + \pi_1(\gamma|_{\Omega \times X \setminus \Omega}) + (2 - [\gamma(\Omega \times \Omega) + \gamma(\Omega \times X \setminus \Omega)])[\![\partial]\!] \\ &= \pi_1(\gamma|_{\Omega \times X}) + (2 - \gamma(\Omega \times X))[\![\partial]\!] \\ &= \mu + (2 - \mu(X))[\![\partial]\!] = \iota(\mu). \end{aligned}$$

Similarly, by symmetry, $\pi_2 \gamma' = \iota(\nu)$. Thus γ' is a coupling of $\iota(\mu)$ and $\iota(\nu)$ in $W_p(\Omega^*)$. Moreover,

$$\begin{aligned} \int \delta(x, y)^p \, \mathrm{d}\gamma'(x, y) = {} & \int_{\Omega \times \Omega} \delta(x, y)^p \, \mathrm{d}\gamma(x, y) + \int_{X \setminus \Omega \times \Omega} \delta(\partial, y)^p \, \mathrm{d}\gamma(x, y) \\ & + \int_{\Omega \times X \setminus \Omega} \delta(x, \partial)^p d\gamma(x, y) + \kappa \delta(\partial, \partial)^p \\ \leq {} & \int_{\Omega \times \Omega} d(x, y)^p \, \mathrm{d}\gamma(x, y) + \int_{X \setminus \Omega \times \Omega} d(x, y)^p \, \mathrm{d}\gamma(x, y) \\ & + \int_{\Omega \times X \setminus \Omega} d(x, y)^p \, \mathrm{d}\gamma(x, y) \\ = {} & \int d(x, y)^p \, \mathrm{d}\gamma(x, y). \end{aligned} \tag{3}$$

Therefore,

$$W_p(\iota(\mu), \iota(\nu)) \le Wb_p(\mu, \nu).$$

Conversely, let γ be a coupling for $\iota(\mu)$ and $\iota(\nu)$ in $W_p(\Omega^*)$. Define the closed set

$$E = \{(x, y) \in \Omega \times \Omega : \delta(x, y) = d(x, y)\}.$$

Fix $\epsilon > 0$ and for each $x \in \Omega$, let $c(x) \in X \backslash \Omega$ with

$$d(x, c(x)) \le (1 + \epsilon)\, \mathrm{dist}(x, X \backslash \Omega).$$

Since X is separable, c may be chosen to be a Borel function with countable image. Let $c_1 = (\mathrm{id} \times c) \circ \pi_1$ and $c_2 = (c \times \mathrm{id}) \circ \pi_2$ and define

$$\gamma' = \gamma|_E + (c_1)_\# \gamma|_{(\Omega \times \Omega^*) \backslash E} + (c_2)_\# \gamma|_{(\Omega^* \times \Omega) \backslash E} \in \mathcal{M}(X \times X). \tag{4}$$

Note that, since $\pi_1((c_1)_\# \gamma)$ is supported on $X \backslash \Omega$, its restriction to Ω equals 0. Therefore,

$$(\pi_1 \gamma')|_\Omega = (\pi_1 \gamma|_E)|_\Omega + (\pi_1 \gamma|_{(\Omega \times \Omega^*) \backslash E})|_\Omega + 0 = (\pi_1 \gamma)|_\Omega = \mu.$$

Similarly, by symmetry, $(\pi_2 \gamma')|_\Omega = \nu$. Hence γ is a coupling for μ and ν in $Wb_p(\Omega)$.

Now, for any $(x, y) \in (\Omega \times \Omega) \backslash E$,

$$d(x, c(x))^p + d(c(y), y)^p \le (1 + \epsilon)^p (\delta(x, \partial)^p + \delta(\partial, y)^p) \le (1 + \epsilon)^p \delta(x, y)^p.$$

Therefore,

$$\begin{aligned}
\int_{X \times X} d(x, y)^p \,\mathrm{d}\gamma'(x, y) &= \int_E d(x, y)^p \,\mathrm{d}\gamma(x, y) + \int_{(\Omega \times \Omega^*) \backslash E} d(x, c(x))^p \,\mathrm{d}\gamma(x, y) \\
&\quad + \int_{(\Omega^* \times \Omega) \backslash E} d(c(y), y)^p \,\mathrm{d}\gamma(x, y) \\
&\le \int_E d(x, y)^p \,\mathrm{d}\gamma(x, y) \\
&\quad + \int_{(\Omega^* \times \Omega^*) \backslash E} (1 + \epsilon)^p \delta(x, y)^p \,\mathrm{d}\gamma(x, y) \\
&\le (1 + \epsilon)^p \int_{\Omega^* \times \Omega^*} \delta(x, y)^p \,\mathrm{d}\gamma(x, y).
\end{aligned} \tag{5}$$

Since $\epsilon > 0$ is arbitrary, this shows that

$$W_p(\iota(\mu), \iota(\nu)) \ge Wb_p(\mu, \nu).$$

$\square$

Remark 3.4 After the first version of this article appeared, we were made aware that the statement of Lemma 3.3, for the case $\Omega = U$ as defined in our introduction, appears in the work of Divol and Lacombe [10, Proposition 3.15]. Note that our proof does not rely on the existence of unique closest points in $\partial\Omega$, whilst the one in [10] does. However, a flaw in their argument makes the proof incorrect even for the case of $\Omega = U$.

Central to their proof is the definition of a measure $\tilde{\pi}'$ and the claim that it is a coupling of $\tilde{\mu}$ and $\tilde{\nu}$ in $W_p(\Omega^*)$ (using the variables of [10, Lemma 3.17]). Using this they derive [10, Equation (3.8)] from which the proof is concluded. However, examples such as [21, Remark 1.9] show this equation to be false. Moreover, this equation would imply that $\delta = d$ in Ω. These contradictions originate in the fact that $\tilde{\pi}'$ is not a coupling of $\tilde{\mu}$ and $\tilde{\nu}$, which can be verified by comparing the total measure of $\tilde{\pi}'$ to that of $\tilde{\mu}$, $\tilde{\nu}$ or $\tilde{\pi}$.

Since the map

$$\begin{aligned} 2W_p(\Omega^*) &\to W_p(\Omega^*) \\ \mu &\mapsto \mu/2 \end{aligned}$$

has distortion $2^{1/p}$, we obtain the following corollary.

Corollary 3.5 *Let X be a separable metric space and $\Omega \subset X$ be non-empty and proper. Then $Wb_p^1(\Omega)$ bi-Lipschitz embeds into $W_p(\Omega^*)$ with distortion 2.*

The same proof as the one for Lemma 3.3 shows that the full space $Wb_p(\Omega)$ isometrically embeds into $\mathcal{M}(\Omega^*)$.

Lemma 3.6 *Let X be a separable metric space and $\Omega \subset X$ non-empty and proper. For any $p \geq 1$,*

$$\begin{aligned} Wb_p(\Omega) &\to (\mathcal{M}(\Omega^*), W_p) \\ \iota'(\mu) &= \mu + \infty \cdot [\![\partial]\!] \end{aligned}$$

is an isometric embedding.

Remark 3.7 For any $\mu \in \mathcal{M}b_p(\Omega)$,

$$W_p(\iota'(\mu), \infty \cdot [\![\partial]\!]) = Wb_p(\mu, 0) < \infty.$$

Therefore, the triangle inequality for W_p implies that W_p is indeed a metric on the image of ι'.

Proof (Proof of Lemma 3.6) If $\mu, \nu \in \mathcal{M}b_p(\Omega)$ then

$$\gamma' = \gamma|_{\Omega\times\Omega} + (\pi_\partial \times \mathrm{id})_\# \gamma|_{X\setminus\Omega\times\Omega} + (\mathrm{id} \times \pi_\partial)_\# \gamma|_{\Omega\times X\setminus\Omega} + \infty \cdot [\![(\partial, \partial)]\!]$$

defines a coupling of $\iota'(\mu)$ and $\iota'(\nu)$. The calculation in (3) shows that

$$W_p(\iota'(\mu), \iota'(\nu)) \leq Wb_p(\mu, \nu).$$

Conversely, if $\mu, \nu \in \mathcal{M}b_p(\Omega)$, then γ' as defined in (4) is a coupling for μ, ν and (5) shows that

$$W_p(\iota'(\mu), \iota'(\nu)) \geq Wb_p(\mu, \nu).$$

□

3.1 The shortcut metric space is not doubling

A metric space X is *doubling* if there exists $N \in \mathbb{N}$ such that each ball $B \subset X$ is covered by N balls of half the radius of B.

Lemma 3.8 *For $n \geq 2$, let $\Omega \subset \mathbb{R}^n$ be non-empty and open such that $\overline{\Omega}$ is a proper subset of $\mathbb{R}^n$. Then for any $N \in \mathbb{N}$ and any sufficiently small $\epsilon > 0$, there exist $y_1, \ldots, y_N \in \Omega$ with $\delta(y_i, y_j) = \epsilon$ for each $i \neq j$. In particular, Ω^* is not doubling.*

Proof Let $x \notin \overline{\Omega}$ and $y \in \Omega$. For $N \in \mathbb{N}$, let $y_1, \ldots, y_N \in \Omega$ lie on the circle centred on x of radius $\|x - y\|$ (such points exist since Ω is open). For each $1 \leq i \leq N$, let l_i' be the line segment connecting y_i to x and let l_i be the connected component of $l_i' \cap \Omega$ containing y_i. Since $x \notin \overline{\Omega}$, there exists $\eta > 0$ such that

$$\inf\{\|z - z'\| : z \in l_i,\ z' \in l_j,\ i \neq j\} > \eta.$$

Now, $\mathrm{dist}(\cdot, \partial\Omega)$ is continuous on each l_i and converges to 0 as one travels along l_i towards $\partial\Omega$. Therefore, for each sufficiently small $\epsilon > 0$ and each $1 \leq i \leq N$, there exists $z_i \in l_i$ with $\mathrm{dist}(z_i, \partial\Omega) = \epsilon/2$. In particular, if $\epsilon < \eta$, then $\delta(z_i, z_j) = \epsilon$ for each $1 \leq i \neq j \leq N$.

Finally, we see that $y_i \in B(y_1, \epsilon)$ for each $1 \leq j \leq N$, but we require at least N balls of radius $\epsilon/4$ to cover $B(y_1, \epsilon)$. Since $N \in \mathbb{N}$ is arbitrary, Ω^* cannot be doubling. □

Remark 3.9 Lemma 3.8 is sharp in the following sense. If $\Omega = (-1, 1) \subset \mathbb{R}$, then Ω^* is bi-Lipschitz equivalent to a Euclidean circle. For any $n \in \mathbb{N}$, if $\Omega = \mathbb{R}^n \backslash \{0\}$, then Ω^* is isometric to $\mathbb{R}^n$. In both of these cases, the conclusion of Lemma 3.8 fails.

Note that each Euclidean space is doubling and that the doubling property is preserved under taking subsets and bi-Lipschitz images. Therefore, if a metric space is bi-Lipschitz embeddable into some Euclidean space, it must necessarily be doubling.

Corollary 3.10 *For $n \geq 2$ let $\Omega \subset \mathbb{R}^n$ be non-empty and open such that $\overline{\Omega}$ is a proper subset of $\mathbb{R}^n$. Then Ω^* is not bi-Lipschitz embeddable into any Euclidean space.*

3.2 The space of unordered tuples of at most m points

Definition 3.11 Let X be a metric space, $\Omega \subset X$ non-empty and proper and $m \in \mathbb{N}$. Define the *space of unordered tuples of at most m points* as

$$\mathcal{B}_m(\Omega) = \bigcup_{k=1}^{m} \mathcal{A}_k(\Omega),$$

with the metric inherited from $Wb_2(\Omega)$.

This space is naturally identified with a subset of $\mathcal{A}_m(\Omega^*)$.

Corollary 3.12 *Let $m \in \mathbb{N}$. For any separable metric space X and non-empty and proper $\Omega \subset X$, $\mathcal{B}_m(\Omega)$ isometrically embeds into $\mathcal{A}_m(\Omega^*)$ via the map*

$$\sum_{i=1}^{k} [\![x_i]\!] \mapsto \sum_{i=1}^{k} [\![x_i]\!] + (2m-k)[\![\partial]\!].$$

Proof Embed $\mathcal{B}_m(X)$ into $Wb_p^1(X)$ by $\mu \mapsto \mu/m$, apply Lemma 3.3, and then embed into $\mathcal{A}_m(\Omega^*)$ by $\mu \mapsto m\mu$. □

4 A bi-Lipschitz description of $\mathcal{A}_m(\Omega^*)$ in terms of $\mathcal{A}_m(\mathbb{R}^{n+1})$

To construct the bi-Lipschitz embedding from Theorem 1.1, it would be natural to adapt the techniques from the proof of Theorem 2.3 to our setting. However, the proof of Theorem 2.3 strictly depends on both, the linear structure of $\mathbb{R}^n$ (in particular the existence of projections), and the compactness of the unit ball. Although $\Omega \subset \mathbb{R}^n$ as a set, δ bears no relationship to the linear structure of $\mathbb{R}^n$ and this fact prohibits the direct use of Almgren's techniques. On the other hand, whilst it is possible to find a bi-Lipschitz embedding of Ω^* into ℓ_2 to gain a linear structure, this comes at the expense of compactness of the unit ball. Thus it is not possible to modify Almgren's proof to our setting.

In order to prove Theorem 1.1 we will use a Whitney decomposition $\mathcal{C}$ of Ω into cubes

$$\Omega = \bigcup_{Q \in \mathcal{C}} Q$$

(see Proposition 4.2) such that, within each Q, δ is given by $\|\cdot\|$. Consequently, $\mathcal{A}_m(Q,\delta) = \mathcal{A}_m(Q,\|\cdot\|)$. Theorem 2.3 then gives a bi-Lipschitz embedding of each $\mathcal{A}_m(Q,\delta)$ into $\mathbb{R}^N$ and it would be favourable to use these embeddings as "coordinate projections" to construct a global embedding into Hilbert space. Of course, the union of the $\mathcal{A}_m(Q)$ does not cover $\mathcal{A}_m(\Omega^*)$ and therefore we cannot simply define coordinate projections by taking restrictions to each Q. Nevertheless, the fact that $\mathcal{A}_m(Q,\delta) =$

$\mathcal{A}_m(Q, \|\cdot\|)$ enables us to construct a map $\phi_Q^*\colon \mathcal{A}_m(\Omega^*) \to \mathcal{A}_m(\mathbb{R}^{n+1})$ which, roughly speaking, acts as a smooth projection to $\mathcal{A}_m(Q)$.

The main result of this section shows that the ϕ_Q^* can be combined to define a bi-Lipschitz embedding of $\mathcal{A}_m(\Omega^*)$ into the following metric space.

Definition 4.1 Let $\mathcal{C}$ be a countable set and define

$$\mathcal{T} := \sum_{Q\in\mathcal{C}} \mathcal{A}_m(\mathbb{R}^{n+1})$$

to be the ℓ_2-sum of copies of $\mathcal{A}_m(\mathbb{R}^{n+1})$. That is, $\mathcal{T}$ consists of sequences

$$\sum_{Q\in\mathcal{C}} a_Q$$

of elements of $\mathcal{A}_m(\mathbb{R}^{n+1})$ for which

$$\sum_{Q\in\mathcal{C}} W_2^2(a_Q, 0) < \infty,$$

where $0 = \sum_{i=1}^m [\![0]\!]$, equipped with the metric

$$\sqrt{\sum_{Q\in\mathcal{C}} W_2^2(a_Q, a'_Q)}.$$

Once we have an embedding into $\mathcal{T}$, we will show that it is possible to find an embedding into ℓ_2. Indeed, in Sect. 5, we apply Theorem 2.3 to each term in the definition of $\mathcal{T}$ to obtain a bi-Lipschitz embedding of $\mathcal{T}$ into ℓ_2.

4.1 A Whitney decomposition of Ω

To construct the embedding into $\mathcal{T}$, we will use a Whitney decomposition of Ω. For a cube $Q \subset \mathbb{R}^n$, let $l(Q)$ denote the side length of Q.

Proposition 4.2 (Appendix J [12]) *Let $\Omega \subset \mathbb{R}^n$ be non-empty, open and proper. There exists a family of closed cubes $\mathcal{C}$ such that*

1. $\cup\mathcal{C} = \Omega$ *and the elements of* $\mathcal{C}$ *have disjoint interiors.*
2. $\sqrt{n}l(Q) \le \operatorname{dist}(Q, \partial\Omega) \le 4\sqrt{n}l(Q)$ *for all* $Q \in \mathcal{C}$.
3. *If* $Q, Q' \in \mathcal{C}$ *and* $Q \cap Q' \neq \emptyset$ *then*

$$\frac{1}{4} \le \frac{l(Q)}{l(Q')} \le 4.$$

We say that Q, Q' *are* neighbours.

4. *Each* $Q \in \mathcal{C}$ *has at most* 12^n *neighbours.*

A Whitney decomposition of Ω estimates which quantity attains the minimum in the definition of δ.

Lemma 4.3 *Let $\mathcal{C}$ be a Whitney decomposition of $\Omega \subset \mathbb{R}^n$, $Q, Q' \in \mathcal{C}$ and $x \in Q$ and $y \in Q'$. Then*

$$\sqrt{n}l(Q) \leq \operatorname{dist}(x, \partial\Omega) \leq 5\sqrt{n}l(Q). \tag{6}$$

If Q, Q' are neighbours then

$$\delta(x, y) = \|x - y\|. \tag{7}$$

If Q, Q' are not neighbours then

$$\frac{l(Q) + l(Q')}{8} \leq \delta(x, y) \leq 5\sqrt{n}(l(Q) + l(Q')). \tag{8}$$

Proof The first inequality in (6) is implied by $\sqrt{n}l(Q) \leq \operatorname{dist}(Q, \partial\Omega)$. The second follows from the triangle inequality:

$$\operatorname{dist}(x, \partial\Omega) \leq \operatorname{dist}(Q, \partial\Omega) + \operatorname{diam}(Q) \leq 4\sqrt{n}l(Q) + \sqrt{n}l(Q).$$

Now suppose Q, Q' are neighbours and let $z \in Q \cap Q'$. Then by (6),

$$\begin{aligned}\operatorname{dist}(x, \partial\Omega) + \operatorname{dist}(y, \partial\Omega) &\geq \sqrt{n}(l(Q) + l(Q')) \\ &\geq \|x - z\| + \|z - y\| \geq \|x - y\|,\end{aligned}$$

giving (7). On the other hand, suppose that Q, Q' are not neighbours and $l(Q) \geq l(Q')$. Then $\|x - y\| \geq l(Q'')$ for Q'' a neighbour of Q. In particular

$$\|x - y\| \geq l(Q'') \geq \frac{l(Q)}{4} \geq \frac{l(Q) + l(Q')}{8},$$

giving the first inequality in (8). The second inequality follows from (6). □

For the remainder of the paper we fix $m \in \mathbb{N}$, $\Omega \subset \mathbb{R}^n$ non-empty, open and proper and $\mathcal{C}$ a Whitney decomposition of Ω as in Proposition 4.2. We also fix $\mathcal{T}$ as in Definition 4.1.

4.2 Constructing a coordinate system

To construct a bi-Lipschitz embedding of $\mathcal{A}_m(\Omega^*)$ into $\mathcal{T}$, we define projections

$$\phi_Q^* \colon \mathcal{A}_m(\Omega^*) \to \mathcal{A}_m(\mathbb{R}^{n+1})$$

that serve as a coordinate system for $\mathcal{A}_m(\Omega^*)$. The embedding into $\mathcal{T}$ will then be defined as the ℓ_2-sum of the ϕ_Q^* (see Definition 4.6).

We begin with the construction of a function ϕ_Q that approximates the identity within a given $Q \in \mathcal{C}$, is supported on the neighbours of Q, and maintains bi-Lipschitz bounds with δ. For $Q \in \mathcal{C}$ and $r > 0$, we write $B(Q, r)$ for the *closed* r-neighbourhood of Q.

Lemma 4.4 *For each $Q \in \mathcal{C}$ there exists a map*

$$\phi_Q \colon \Omega \to \mathbb{R}^{n+1}$$

such that

1. *ϕ_Q is $9\sqrt{n+1}$-Lipschitz;*
2. *$\phi_Q(x) = 0$ for all $x \notin B(Q, l(Q)/4)$. In particular, ϕ_Q is supported on the neighbours of Q;*
3. *$\|\phi_Q\|_\infty \leq \sqrt{n+1}\, l(Q)$;*
4. *For all $x, y \in B(Q, l(Q)/8)$,*

$$\|\phi_Q(x) - \phi_Q(y)\| = \|x - y\|;$$

5. *The extension of ϕ_Q to Ω^*, defined by $\phi_Q(\partial) = 0$, is $9\sqrt{n+1}$-Lipschitz with respect to δ;*
6. *If $x \in B(Q, l(Q)/8)$ and $y \in \Omega^*$, then*

$$\|\phi_Q(x) - \phi_Q(y)\| \geq \min\left\{\frac{\|x - y\|}{2\sqrt{n}}, l(Q)\right\}.$$

Proof Fix $Q \in \mathcal{C}$ and let c be the centre of Q. For each $x \in \Omega$, let

$$\eta(x) = \max\left\{1 - \operatorname{dist}\left(x, B\left(Q, \frac{l(Q)}{8}\right)\right)\frac{8}{l(Q)}, 0\right\}.$$

That is, η is an $8/l(Q)$-Lipschitz function with $\|\eta\|_\infty = 1$ that equals 1 on $B(Q, l(Q)/8)$ and 0 on $\Omega \backslash B(Q, l(Q)/4)$. We also set

$$\varphi(x) = (x - c, l(Q)) \in \mathbb{R}^{n+1},$$

a 1-Lipschitz function satisfying $\|\varphi(x)\| \leq \sqrt{n+1}\, l(Q)$ for all x in the support of η.

Define $\phi_Q = \eta\varphi$. Since ϕ_Q is a product of Lipschitz functions, the Lipschitz constant of ϕ_Q is bounded above by

$$\operatorname{Lip}\varphi\|\eta\|_\infty + \sup\{\|\varphi(x)\| : x \in \operatorname{spt}\eta\}\operatorname{Lip}\eta \leq 1 + \sqrt{n+1}\, l(Q)\frac{8}{l(Q)} \leq 9\sqrt{n+1}.$$

This demonstrates item 1. Items 2 to 4 are immediate.

To see item 5, first let $x \in \Omega^*$ be such that $\phi_Q(x) \neq 0$. Then by item 2, $x \in Q'$ for Q' a neighbour of Q, so that $l(Q') \geq l(Q)/4$. Therefore, by item 3,

$$\begin{aligned}\|\phi_Q(x)\| &\leq \sqrt{n+1}\, l(Q) \\ &\leq 4\sqrt{n+1}\, l(Q') \\ &\leq 8\,\mathrm{dist}(x, \partial\Omega),\end{aligned}$$

using Eq. (6) for the final inequality. Thus

$$\|\phi_Q(x)\| \leq 8\,\mathrm{dist}(x, \partial\Omega)$$

holds for any $x \in \Omega^*$ (including $x = \partial$). Therefore, by the triangle inequality, for any $x, y \in \Omega^*$,

$$\|\phi_Q(x) - \phi_Q(y)\| \leq 8\left(\mathrm{dist}(x, \partial\Omega) + \mathrm{dist}(y, \partial\Omega)\right).$$

Combining this inequality with item 1 shows that ϕ_Q is $9\sqrt{n+1}$-Lipschitz with respect to δ on Ω^*.

Finally, to see item 6, first suppose that $y \notin B(Q, l(Q)/4)$. Then by item 2,

$$\|\phi_Q(x) - \phi_Q(y)\| = \|\varphi(x)\| \geq l(Q),$$

so that item 6 holds in this case.

In the case $y \in B(Q, l(Q)/4)$ we will show that

$$\|\phi_Q(x) - \phi_Q(y)\| \geq \frac{\|x-y\|}{2\sqrt{n}}, \tag{9}$$

completing the proof of item 6. To this end, note that

$$\|y - c\| \leq \sqrt{n}\frac{l(Q)}{2} + \frac{l(Q)}{4} \leq \sqrt{n}l(Q).$$

Therefore, by considering the first component of ϕ_Q, we see that

$$\begin{aligned}\|\phi_Q(x) - \phi_Q(y)\| &\geq \|(x-c) - \eta(y)(y-c)\| \\ &\geq \|x-y\| - (1-\eta(y))\|y-c\| \\ &\geq \|x-y\| - \sqrt{n}(1-\eta(y))l(Q).\end{aligned}$$

Thus, if

$$\sqrt{n}(1-\eta(y))l(Q) \leq \frac{\|x-y\|}{2}, \tag{10}$$

then (9) holds. On the other hand, if (10) does not hold, then by considering the final component of ϕ, we have

$$\|\phi_Q(x) - \phi_Q(y)\| \geq (1 - \eta(y))l(Q) \geq \frac{\|x - y\|}{2\sqrt{n}},$$

giving (9). □

The pushforwards under each ϕ_Q define our coordinate projections on $\mathcal{A}_m(\Omega)$.

Definition 4.5 For every $Q \in \mathcal{C}$, define ϕ_Q^* to be the pushforward under ϕ_Q. That is,

$$\begin{aligned} \phi_Q^* \colon \mathcal{A}_m(\Omega^*) &\to \mathcal{A}_m(\mathbb{R}^{n+1}) \\ \sum_{i=1}^m [\![p_i]\!] &\mapsto \sum_{i=1}^m [\![\phi_Q(p_i)]\!]. \end{aligned}$$

Recall the construction of $\mathcal{T}$ from Definition 4.1.

Definition 4.6 Define the embedding ϕ^* by

$$\begin{aligned} \mathcal{A}_m(\Omega^*) &\to \mathcal{T} \\ \phi^* &= \sum_{Q \in \mathcal{C}} \phi_Q^* \end{aligned}$$

This is well defined since each ϕ_Q is supported on the neighbours of Q, so that each $x \in \Omega$ is contained in the support of at most 12^n of the ϕ_Q.

4.3 ϕ^* is bi-Lipschitz

In this section we show that ϕ^* is a bi-Lipschitz embedding, beginning by showing that it is Lipschitz.

For $p \in (\Omega^*)^m$ and $S \subset \Omega^*$, let

$$p^{-1}(S) = \{1 \leq k \leq m : p_k \in S\}.$$

From now on we use the notation σq to denote the element of $(\mathbb{R}^n)^m$ arising from the natural action of the symmetric group Σ_m on $(\mathbb{R}^n)^m$: $(\sigma q)_i = q_{\sigma(i)}$ for each $1 \leq i \leq m$.

Lemma 4.7 *For any $p, q \in \mathcal{A}_m(\Omega^*)$,*

$$\sum_{Q \in \mathcal{C}} W_2(\phi_Q^*(p), \phi_Q^*(q))^2 \leq c_0 W_2^2(p, q),$$

where $c_0 \geq 1$ depends only upon n.

Proof Fix $p, q \in (\Omega^*)^m$ and let $Q \in \mathcal{C}$ and $\sigma \in \Sigma_m$. Set

$$J_Q^\sigma = p^{-1}(B(Q, l(Q)/4)) \cup (\sigma q)^{-1}(B(Q, l(Q)/4)),$$

so that, by Lemma 4.4 item 2,

$$\sum_{k=1}^m \|\phi_Q(p_k) - \phi_Q(q_{\sigma(k)})\|^2 = \sum_{k \in J_Q^\sigma} \|\phi_Q(p_k) - \phi_Q(q_{\sigma(k)})\|^2.$$

Applying Lemma 4.4 item 5 gives

$$\sum_{k=1}^m \|\phi_Q(p_k) - \phi_Q(q_{\sigma(k)})\|^2 \leq 9^2(n+1) \sum_{k \in J_Q^\sigma} \delta(p_k, q_{\sigma(k)})^2.$$

Therefore

$$\sum_{Q \in \mathcal{C}} \min_{\sigma \in \Sigma_m} \sum_{k=1}^m \|\phi_Q(p_k) - \phi_Q(q_{\sigma(k)})\|^2 \leq 9^2(n+1) \sum_{Q \in \mathcal{C}} \min_{\sigma \in \Sigma_m} \sum_{k \in J_Q^\sigma} \delta(p_k, q_{\sigma(k)})^2.$$

Further,

$$\begin{aligned}\sum_{Q \in \mathcal{C}} \min_{\sigma \in \Sigma_m} \sum_{k \in J_Q^\sigma} \delta(p_k, q_{\sigma(k)})^2 &\leq \min_{\sigma \in \Sigma_m} \sum_{Q \in \mathcal{C}} \sum_{k \in J_Q^\sigma} \delta(p_k, q_{\sigma(k)})^2 \\ &\leq \min_{\sigma \in \Sigma_m} 2 \cdot 12^n \sum_{k=1}^m \delta(p_k, q_{\sigma(k)})^2,\end{aligned}$$

since $B(Q, l(Q)/4)$ is contained within the union of the neighbours of Q. The result follows for $c_0 = 2 \cdot 9^2 \cdot 12^n(n+1)$. □

To prove the lower Lipschitz bound, we fix the following notation until the end of the section.

Notation 4.8 Fix $p, q \in (\Omega^*)^m$ and, for every $Q \in \mathcal{C}$, let $\sigma_Q \in \Sigma_m$ be such that

$$\sum_{k=1}^m \|\phi_Q(p_k) - \phi_Q(q_{\sigma_Q(k)})\|^2 = W_2(\phi_Q^*(p), \phi_Q^*(q))^2. \tag{11}$$

Let $Q \in \mathcal{C}$. For integer $0 \leq r \leq 2m$, the annuli

$$Q^r = B\left(Q, \frac{r+1}{3m}\frac{l(Q)}{8}\right) \setminus B\left(Q, \frac{r}{3m}\frac{l(Q)}{8}\right)$$

are disjoint and so there exists $0 \le r \le 2m$ such that

$$p^{-1}(Q^r) \cup (\sigma_Q q)^{-1}(Q^r) = \emptyset. \tag{12}$$

Set

$$\widehat{Q} = B\left(Q, \frac{r}{3m}\frac{l(Q)}{8}\right).$$

Note that $\widehat{Q}$ is contained within the union of the neighbours of Q.

Let $c_1 = (48\sqrt{n})^{-1}$ and define $\mathcal{C}'$ to be the set of $Q \in \mathcal{C}$ for which

$$W_2(\phi_Q^*(p), \phi_Q^*(q)) < c_1 \frac{l(Q)}{m}. \tag{13}$$

Set

$$E = \bigcup_{Q \in \mathcal{C}'} \widehat{Q}.$$

To obtain a lower bound of

$$\sum_{Q \in \mathcal{C}} W_2(\phi_Q^*(p), \phi_Q^*(q))^2 \tag{14}$$

in terms of $W_2^2(p,q)$, we will construct a $\tau \in \Sigma_m$ for which $\sum_{i=1}^m \delta(p_i, q_{\tau(i)})^2$ is comparable to (14). A first attempt to do this may be, for each $Q \in \mathcal{C}$ and each $i \in p^{-1}(Q)$, to define $\tau(i) = \sigma_Q(i)$. Of course, a τ defined in this way need not be injective, for example if there exist $Q \neq Q' \in \mathcal{C}$ and $i \neq j$ such that $q_{\sigma_Q(i)} = q_{\sigma_{Q'}(j)}$. Nonetheless, we will show that it is possible to construct a permutation for the cubes in $\mathcal{C}'$. Indeed, we now show that conditions (12) and (13) ensure that, for each $Q \in \mathcal{C}'$, $p_i \in \widehat{Q}$ if and only if $q_{\sigma_Q(i)} \in \widehat{Q}$: (12) provides a moat surrounding $\widehat{Q}$ and (13) ensures that the distance between p_i and $q_{\sigma_Q(i)}$ is less than the width of the moat.

Lemma 4.9 *For any $Q \in \mathcal{C}'$,*

$$p^{-1}(\widehat{Q}) = (\sigma_Q q)^{-1}(\widehat{Q}) \tag{15}$$

and

$$\|p_k - q_{\sigma_Q(k)}\| = \|\phi_Q(p_k) - \phi_Q(q_{\sigma_Q(k)})\| \quad \forall k \in p^{-1}(\widehat{Q}). \tag{16}$$

Moreover, if $R \in \mathcal{C}'$ with $l(R) \le l(Q)$,

$$p^{-1}(\widehat{Q} \cap \widehat{R}) = (\sigma_R q)^{-1}(\widehat{Q} \cap \widehat{R}) \tag{17}$$

Proof For any $k \in p^{-1}(\widehat{Q})$, (13) and Lemma 4.4 item 6 imply

$$\min\left\{\frac{\|p_k - q_{\sigma_Q(k)}\|}{2\sqrt{n}}, l(Q)\right\} < c_1 \frac{l(Q)}{m}.$$

In particular,

$$\|p_k - q_{\sigma_Q(k)}\| < \frac{l(Q)}{24m}. \tag{18}$$

Therefore (12) implies that $q_{\sigma_Q(k)} \in \widehat{Q}$. By symmetry, if $k \in (\sigma_Q q)^{-1}(\widehat{Q})$ then $k \in p^{-1}(\widehat{Q})$ and so (15) holds. Since $\widehat{Q} \subset B(Q, l(Q)/8)$, Lemma 4.4 item 4 implies (16).

Now let $R \in \mathcal{C}'$ with $l(R) \leq l(Q)$ and $k \in p^{-1}(\widehat{Q} \cap \widehat{R})$. Then (18) for R implies

$$\|p_k - q_{\sigma_R(k)}\| < \frac{l(R)}{24m} \leq \frac{l(Q)}{24m}$$

and so (12) implies $q_{\sigma_R(k)} \in \widehat{Q}$. The similar argument with p and $\sigma_R q$ exchanged gives (17). □

By carefully partitioning E using the $\widehat{Q}$, we use Lemma 4.9 to construct the desired permutation on $p^{-1}(E)$.

Proposition 4.10 *There exists a bijection* $\tau \colon p^{-1}(E) \to q^{-1}(E)$ *such that*

$$\sum_{k \in p^{-1}(E)} \|p_k - q_{\tau(k)}\|^2 \leq \sum_{Q \in \mathcal{C}'} W_2(\phi_Q^*(p), \phi_Q^*(q))^2.$$

Proof Let

$$\mathcal{C}'' = \{Q \in \mathcal{C}' : p^{-1}(\widehat{Q}) \neq \emptyset\}.$$

Note that, by (15), $\mathcal{C}''$ can equivalently be defined as the set of $Q \in \mathcal{C}$ with $q^{-1}(\widehat{Q}) \neq \emptyset$. Since $\mathcal{C}''$ is finite, we enumerate it as

$$\mathcal{C}'' = \{Q_1, Q_2, \ldots, Q_j\}$$

in such a way that

$$l(Q_1) \geq l(Q_2) \geq \cdots \geq l(Q_j).$$

Then, for $1 \leq i \leq k \leq j$, applying Lemma 4.9 with $Q = Q_k$ and $R = Q_i$ gives

$$p^{-1}\left(\widehat{Q}_i \cap \widehat{Q}_k\right) = (\sigma_{Q_k} q)^{-1}\left(\widehat{Q}_i \cap \widehat{Q}_k\right) \quad \forall 1 \leq i \leq k \leq j. \tag{19}$$

Let $B_1 = \widehat{Q}_1$ and for each $2 \leq k \leq j$ define

$$B_k := \widehat{Q}_k \setminus \bigcup_{i=1}^{k-1} \widehat{Q}_i = \widehat{Q}_k \setminus \bigcup_{i=1}^{k-1} \widehat{Q}_i \cap \widehat{Q}_k.$$

Then (19) implies that σ_{Q_k} is a permutation between $p^{-1}(B_k)$ and $\sigma_{Q_k} q^{-1}(B_k)$ for each $1 \leq k \leq j$. Therefore, we define a bijection

$$\tau\colon p^{-1}(E) \to q^{-1}(E)$$

by setting τ to equal σ_{Q_k} on $D_k := p^{-1}(B_k)$ for each $1 \leq k \leq j$. Then

$$\begin{aligned}\sum_{k \in p^{-1}(E)} \|p_k - q_{\tau(k)}\|^2 &= \sum_{i=1}^{j} \sum_{k \in D_i} \|p_k - q_{\tau(k)}\|^2 \\ &= \sum_{i=1}^{j} \sum_{k \in D_i} \|p_k - q_{\sigma_{Q_i}(k)}\|^2 \\ &= \sum_{i=1}^{j} \sum_{k \in D_i} \|\phi_{Q_i}(p_k) - \phi_{Q_i}(q_{\sigma_{Q_i}(k)})\|^2 \\ &\leq \sum_{Q \in \mathcal{C}''} \sum_{k=1}^{m} \|\phi_Q(p_k) - \phi_Q(q_{\sigma_Q(k)})\|^2,\end{aligned}$$

using (16) for the third equality. Finally (11) completes the proof. $\square$

Next we consider the points outside E for which we use the distance to $\partial\Omega$ to estimate δ.

Lemma 4.11 *For any bijection*

$$\sigma\colon p^{-1}(\Omega \setminus E) \to q^{-1}(\Omega \setminus E)$$

we have

$$\sum_{k \in p^{-1}(\Omega \setminus E)} (\mathrm{dist}(p_k, \partial\Omega) + \mathrm{dist}(q_{\sigma(k)}, \partial\Omega))^2 \leq m^3 c_2 \sum_{Q \in \mathcal{C} \setminus \mathcal{C}'} W_2(\phi_Q^*(p), \phi_Q^*(q))^2,$$

for $c_2 \geq 1$ that depends only upon n.

Proof For a moment fix $k \in p^{-1}(\Omega \setminus E)$ and let $Q \in \mathcal{C}$ contain p_k. Then necessarily $Q \notin \mathcal{C}'$. Therefore (13) and (6) imply

$$W_2(\phi_Q^*(p), \phi_Q^*(q)) \geq \frac{c_1}{m} l(Q) \geq \frac{c_1}{5\sqrt{n}m} \mathrm{dist}(p_k, \partial\Omega).$$

Since each $Q \in \mathcal{C}$ contains at most m such points p_k,

$$\sum_{k\in p^{-1}(\Omega\setminus E)} \operatorname{dist}(p_k, \partial\Omega)^2 \le \frac{25m^2 n}{c_1^2} m \sum_{Q\in\mathcal{C}\setminus\mathcal{C}'} W_2(\phi_Q^*(p), \phi_Q^*(q))^2.$$

The same estimate for σq gives the desired inequality for $c_2 = 4 \cdot 25n/c_1^2$. □

We combine our previous results to show that ϕ_Q^* is a bi-Lipschitz embedding.

Theorem 4.12 *For any $p, q \in \mathcal{A}_m(\Omega^*)$,*

$$\frac{W_2(p,q)^2}{c_3 m^3} \le \sum_{Q\in\mathcal{C}} W_2(\phi_Q^*(p), \phi_Q^*(q))^2 \le c_3 W_2(p,q)^2,$$

where $c_3 \ge 1$ depends only upon n.

Proof The right hand inequality is given by Lemma 4.7.

For the left hand inequality, let τ be the bijection obtained from Proposition 4.10 and arbitrarily extend it to a bijection of $\{1, \dots, m\}$. Then

$$\begin{aligned}
\sum_{Q\in\mathcal{C}} W_2(\phi_Q^*(p), \phi_Q^*(q))^2 &= \sum_{Q\in\mathcal{C}'} \sum_{k=1}^{m} \|\phi_Q(p_k) - \phi_Q(q_{\sigma_Q(k)})\|^2 \\
&\quad + \sum_{Q\notin\mathcal{C}'} \sum_{k=1}^{m} \|\phi_Q(p_k) - \phi_Q(q_{\sigma_Q(k)})\|^2 \\
&\ge \sum_{k\in p^{-1}(E)} \|p_k - q_{\tau(k)}\|^2 \\
&\quad + \frac{1}{c_2 m^3} \sum_{k\in p^{-1}(\Omega\setminus E)} (\operatorname{dist}(p_k, \partial\Omega) + \operatorname{dist}(q_{\tau(k)}, \partial\Omega))^2 \\
&\ge \frac{1}{c_2 m^3} \sum_{k=1}^{m} \delta(p_k, q_{\tau(k)})^2 \\
&\ge \frac{1}{c_2 m^3} W_2(p,q)^2,
\end{aligned}$$

using Proposition 4.10 and Lemma 4.11 for the first inequality. □

5 The embedding into Hilbert space

In this section we conclude the proof of Theorem 1.1. Let $\xi : \mathcal{A}_m(\mathbb{R}^{n+1}) \to \mathbb{R}^N$ be the embedding given by Theorem 2.3. We write

$$\ell_2 = \sum_{Q\in\mathcal{C}} \mathbb{R}^N$$

as a direct l_2-sum over $\mathcal{C}$. Recall the construction of $\mathcal{T}$ from Definition 4.1.

Lemma 5.1 *The function $\xi'\colon \mathcal{T} \to \ell_2$ defined by*

$$\sum_{Q\in\mathcal{C}} \mathcal{A}_m(\mathbb{R}^{n+1}) \to \sum_{Q\in\mathcal{C}} \mathbb{R}^N$$
$$\xi' = \sum_{Q\in\mathcal{C}} \xi$$

is well defined. Moreover, for any $a, b \in \mathcal{T}$,

$$\frac{1}{cm^{2n+2}} \sum_{Q\in\mathcal{C}} W_2(a_Q, b_Q)^2 \le \|\xi'(a) - \xi'(b)\|^2 \le \sum_{Q\in\mathcal{C}} W_2(a_Q, b_Q)^2,$$

for $c \ge 1$ depending only upon n.

Proof Let $a \in \mathcal{T}$, so that

$$\sum_{Q\in\mathcal{C}} W_2(p_Q, 0)^2 < \infty.$$

Since ξ is 1-Lipschitz this implies that

$$\sum_{Q\in\mathcal{C}} \|\xi(p_Q)\|^2 = \sum_{Q\in\mathcal{C}} \|\xi(p_Q) - \xi(0)\|^2 \le \sum_{Q\in\mathcal{C}} W_2(p_Q, 0)^2 < \infty.$$

Hence, ξ' is well defined. Moreover, using that ξ is 1-Lipschitz again, we have, for any $b \in \mathcal{T}$,

$$\sum_{Q\in\mathcal{C}} \|\xi(a_Q) - \xi(b_Q)\|^2 \le \sum_{Q\in\mathcal{C}} W_2(a_Q, b_Q)^2,$$

so that ξ' is also 1-Lipschitz. Finally, Theorem 2.3 gives

$$\sum_{Q\in\mathcal{C}} \|\xi(a_Q) - \xi(b_Q)\|^2 \ge \frac{1}{cm^{2n+2}} \sum_{Q\in\mathcal{C}} W_2(a_Q, b_Q)^2.$$

□

Theorem 5.2 *There exists a bi-Lipschitz embedding $\zeta\colon \mathcal{B}_m(\Omega) \to \ell_2$ with distortion at most $cm^{n+5/2}$, for $c \ge 1$ depending only upon n. That is, for any $p, q \in \mathcal{B}_m(\Omega)$,*

$$\frac{W_2(p, q)}{cm^{n+5/2}} \le \|\zeta(p) - \zeta(q)\| \le cW_2(p, q).$$

Proof First isometrically embed $\mathcal{B}_m(\Omega)$ into $\mathcal{A}_m(\Omega^*)$ via Corollary 3.12. One then applies Theorem 4.12 to bi-Lipschitz embed $\mathcal{A}_m(\Omega^*)$ into $\mathcal{T}$. Finally, Lemma 5.1 bi-Lipschitz embeds $\mathcal{T}$ into ℓ_2, as required. □

Remark 5.3 For $n \geq 3$, the distortion of any embedding of $\mathcal{A}_m(\Omega^*)$ into ℓ_2 converges to ∞ as m increases. In particular, $Wb_2(\Omega)$ does not bi-Lipschitz embed into ℓ_2.

Indeed, by Eq. (7) we see that $\mathcal{A}_m(\Omega^*)$ contains an isometric copy of $A_m(Q)$ for some cube Q. Thus, the distortion of any embedding into ℓ_2 is at least that of $\mathcal{A}_m(Q)$. For $n \geq 3$, Andoni, Naor and Nieman [4, Theorem 7] prove that $W_2(\mathbb{R}^n)$ does not coarsely, in particular bi-Lipschitz, embed into any Banach space of non-trivial type, namely Hilbert space. Since the set of discrete measures is dense in $W_2(\mathbb{R}^n)$, a scaling argument shows that the distortion of any bi-Lipschitz embedding of $\mathcal{A}_m(Q)$ must converge to ∞ as m does.

The same conclusion can be made for $n = 2$ using an unpublished result of Austin and Naor announced in [4, Remark 8], which states that $W_2(\mathbb{R}^2)$ does not bi-Lipschitz embed into L_1 and, hence, does not bi-Lipschitz embed into ℓ_2.

Acknowledgements D.B. was supported by the European Union's Horizon 2020 research and innovation programme grant number 948021. A.L.G.P. was supported by the Engineering and Physical Sciences Research Council grant number EP/R018472/1. We would like to thank Andrea Marchese for useful discussions regarding Almgren's m-valued functions. We would also like to thank the referee for carefully reading this article and providing valuable suggestions that improved the exposition of this work.

Data availability There is no data.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there are no conflicts of interest.

References

1. Almgren, F.J.: Almgren's Big Regularity Paper. In: Scheffer, V., Taylor, Jean E. (eds.) Q-valued Functions Minimizing Dirichlet's Integral and the Regularity of Area-minimizing Rectifiable Currents up to Codimension 2. Springer, New York (2000)
2. Ambrosio, L., Brué, E., Semola, D.: Lectures on Optimal Transport. Unitext, vol. 130, p. 250. Springer, Berlin (2021). https://doi.org/10.1007/978-3-030-72162-6
3. Ambrosio, L., Gigli, N., Savaré, G.: Gradient Flows in Metric Spaces and in the Space of Probability Measures, Lectures in Mathematics ETH Zürich, p. 334 (2008)
4. Andoni, A., Naor, A., Neiman, O.: Snowflake universality of Wasserstein spaces. Ann. Sci. Éc. Norm. Supér. (4) **51**(3), 657–700 (2018). https://doi.org/10.24033/asens.2363

5. Baudier, F., Gartland, C., Schlumprecht, T.: L_1-distortion of Wasserstein metrics: a tale of two dimensions. Trans. Am. Math. Soc. Ser. B **10**, 1077–1118 (2023). https://doi.org/10.1090/btran/143
6. Carrière, M., Bauer, U.: On the metric distortion of embedding persistence diagrams into separable Hilbert spaces. In: 35th International Symposium on Computational Geometry, pp. 21–15 (2019)
7. Che, M., Galaz-García, F., Guijarro, L., Solis, I.M.: Metric Geometry of Spaces of Persistence Diagrams (2021)
8. Chizat, L., Peyré, G., Schmitzer, B., Vialard, F.-X.: Unbalanced optimal transport: dynamic and Kantorovich formulations. J. Funct. Anal. **274**(11), 3090–3123 (2018). https://doi.org/10.1016/j.jfa.2018.03.008
9. De Lellis, C., Spadaro, E.N.: Q-valued functions revisited. Mem. Am. Math. Soc. **211**(991), 79 (2011). https://doi.org/10.1090/S0065-9266-10-00607-1
10. Divol, V., Lacombe, T.: Understanding the topology and the geometry of the space of persistence diagrams via optimal partial transport. J. Appl. Comput. Topol. **5**(1), 1–53 (2021). https://doi.org/10.1007/s41468-020-00061-z
11. Figalli, A., Gigli, N.: A new transportation distance between non-negative measures, with applications to gradients flows with Dirichlet boundary conditions. J. Math. Pures Appl. (9) **94**(2), 107–130 (2010). https://doi.org/10.1016/j.matpur.2009.11.005
12. Grafakos, L.: Classical Fourier Analysis, 3rd edn. Grad. Texts Math., vol. 249 (2014). https://doi.org/10.1007/978-1-4939-1194-3
13. Guillen, N., Mou, C., Święch, A.: Coupling Lévy measures and comparison principles for viscosity solutions. Trans. Am. Math. Soc. **372**(10), 7327–7370 (2019). https://doi.org/10.1090/tran/7877
14. Indyk, P.: Algorithmic applications of low-distortion geometric embeddings. In: 42nd IEEE Symposium on Foundations of Computer Science, pp. 10–33 (2001)
15. Khot, S., Naor, A.: Nonembeddability theorems via Fourier analysis. Math. Ann. **334**(4), 821–852 (2006). https://doi.org/10.1007/s00208-005-0745-0
16. Liero, M., Mielke, A., Savaré, G.: Optimal entropy-transport problems and a new Hellinger–Kantorovich distance between positive measures. Invent. Math. **211**(3), 969–1117 (2018). https://doi.org/10.1007/s00222-017-0759-8
17. Mitra, A., Virk, Ž: The space of persistence diagrams on n points coarsely embeds into Hilbert space. Proc. Am. Math. Soc. **149**(6), 2693–2703 (2021). https://doi.org/10.1090/proc/15363
18. Naor, A.: Metric dimension reduction: a snapshot of the Ribe program. In: Proceedings of the International Congress of Mathematicians 2018, pp. 759–837 (2018). https://doi.org/10.1142/9789813272880_0029
19. Naor, A., Schechtman, G.: Planar earthmover is not in L_1. SIAM J. Comput. **37**(3), 804–826 (2007). https://doi.org/10.1137/05064206X
20. Piccoli, B., Rossi, F.: Generalized Wasserstein distance and its application to transport equations with source. Arch. Ration. Mech. Anal. **211**(1), 335–358 (2014). https://doi.org/10.1007/s00205-013-0669-x
21. Profeta, A., Sturm, K.-T.: Heat flow with Dirichlet boundary conditions via optimal transport and gluing of metric measure spaces. Calc. Var. Partial Differ. Equ. **59**(4), 34 (2020). https://doi.org/10.1007/s00526-020-01774-w