Learning on mm spaces based on Gromov's reconstruction theorem

Claire Brécheteau

Nantes Université, Centrale Nantes, LMJL, France

Thomas Verdebout

Department of Mathematics, ECARES, ULB, Belgium

Abstract

We focus on the questions of testing and learning for datasets represented by a matrix of pairwise distances between datapoints. Such datasets can be considered as discrete versions of metric measure spaces (mm spaces). We recall the Gromov's mm spaces reconstruction theorem, that states that mm spaces can be represented by the distribution of pairwise distance matrices. We give an alternative and detailed proof of this theorem. Then we introduce a new metric between mm spaces, based on this theorem, as an alternative to the Gromov-Wasserstein distance, and prove stability results and in particular parametric rates. As the Gromov-Wasserstein distance, this metric allows to account for variations of both density and shape of datasets. We provide new goodness-of-fit and two-sample tests for mm spaces, but also new classification methods for data given by mm spaces (or pairwise distance matrices), based on this new metric.

2012 ACM Subject Classification Nonparametric statistics, Hypothesis testing and confidence interval computation, Computational geometry

Keywords and phrases metric measure spaces, classification, testing, characteristic function, bootstrap

Funding *Claire Brécheteau*: Pulsar, Région Pays de Loire, Centrale Nantes, ANR GeoDSiC
Thomas Verdebout: Tournesol, Hubert Curien, FNRS

Acknowledgements We want to thank Frédéric Chazal for pointing out the Gromov's mm spaces reconstruction theorem.

1 Introduction

Pairwise distance matrices between points in a dataset can often be used to provide learning methods. The only requirement is that the space containing datapoints has to be equipped with a metric. Examples of such spaces are graphs, Riemannian manifolds with in particular spaces for directional data (torus, circle, sphere, the Grassmannian, etc.). Concrete examples such as images also fit into this framework. Even for Euclidean spaces, considering pairwise distance matrices is relevant if we are not interested in the location nor in the principal directions of the datapoints, since two datasets equal up to a translation or a rotation have the same distance matrix.

Clustering methods have been widely used for such datasets, as for instance spectral clustering [32] or density modes and neighborhood graphs-based algorithms [13]. Such methods allow to extract different separated features from datasets. In this paper, we are more interested on infering the shape of these features or datasets, and their density variations. In particular, we focus on the question of testing that two samples are equal up to a rotation or a translation, or if a sample is uniform on a shape. We are also interested in defining learning algorithms (e.g. classification) where each datapoint is given by a pairwise distance matrix, and where we aim at assigning a class to each datapoint. To tackle these problematics, we consider the more general concept of metric measure space. We consider that the datapoints are a sample of n independent random variables generated according to the same probability distribution $\mathrm{P}_{\mathscr{X}}$ on a metric space $(\mathscr{X}, \mathrm{d}_{\mathscr{X}})$. If $(\mathscr{X}, \mathrm{d}_{\mathscr{X}})$ is complete and separable, and $\mathrm{P}_{\mathscr{X}}$ is a Borel probability measure with support $\mathrm{Supp}(\mathrm{P}_{\mathscr{X}})$ equal to $\mathscr{X}$ (defined as the smallest closed subset of $\mathscr{X}$ with $\mathrm{P}_{\mathscr{X}}$-probability 1), then, we say that $(\mathscr{X}, \mathrm{d}_{\mathscr{X}}, \mathrm{P}_{\mathscr{X}})$ is a metric measure space, or mm space for short. Such spaces have been studied by Gromov in [19] and Sturm in [28], and are still widely used, especially for data analysis. Two mm spaces $(\mathscr{X}, \mathrm{d}_{\mathscr{X}}, \mathrm{P}_{\mathscr{X}})$ and $(\mathscr{Y}, \mathrm{d}_{\mathscr{Y}}, \mathrm{P}_{\mathscr{Y}})$ are considered equal when they are isomorphic, in the sense that there exists some one-to-one and onto function $\phi : \mathscr{X} \mapsto \mathscr{Y}$ which is an isometry and so that the pushforward of the measure $\mathrm{P}_{\mathscr{X}}$ by ϕ, $\phi_\sharp(\mathrm{P}_{\mathscr{X}})$, is equal to the measure $\mathrm{P}_{\mathscr{Y}}$:

1. $\forall\, x, x' \in \mathscr{X}$, $\mathrm{d}_{\mathscr{X}}(x, x') = \mathrm{d}_{\mathscr{Y}}(\phi(x), \phi(x'))$,
2. $\forall\, B$, Borel set of $\mathscr{Y}$, $\mathrm{P}_{\mathscr{Y}}(B) = \mathrm{P}_{\mathscr{X}}(\phi^{-1}(B)) = \phi_\sharp(\mathrm{P}_{\mathscr{X}})(B)$.

The Gromov's mm spaces reconstruction theorem [19, Theorem 3 1/2.5.] states that two mm spaces are isomorphic if and only if the distribution of the sample pairwise distance matrices coincide, for every sample size. For a sample of size 2, this distribution coincides with the shape signature, [23]. Several other signatures induce pseudo-metrics to compare mm spaces. For instance, distance-to-measure signatures (DTM-signatures) based on distances to nearest neighbours have been used to test equality of two mm spaces up to an isomorphism [7] and to test uniformity of a sample on a compact homogeneous space such as a ball, a torus or a Grassmannian, [8]. Topological signatures such as persistence diagrams [11, 12] have been widely used to compare the topological features of datasets, including connected components, loops, voids, etc. If these signatures induce pseudo-metrics, a metric has been defined by Mémoli in [22] to compare mm spaces up to an isomorphism, the Gromov-Wasserstein distance. Since then, the litterature have been proficient on the subject. In [3], the Gromov-Wasserstein distance has been adapted to generalisations of mm spaces. Lots of references are available in this paper regarding recent use of Gromov-Wasserstein distance in machine learning and data science, for scalable and empirically accurate computational schemes for its estimation, as well as for other variants of Gromov-Wasserstein distances to compare mm spaces. Recently, [14] proved the existence of an optimal Monge transport plan for the Gromov-Wasserstein distance between two mm spaces. In [24], a notion of barycenter for mm spaces or for distance matrices has been considered. A notion of weak convergence for mm spaces has been studied in [18]. Comparing two measures defined on the same space may also been interesting. For instance, in [2], the Gromov-Wasserstein distance has been studied on spheres. Classical probability measures on the sphere $\mathscr{S}^{d-1}$ in the Euclidean space $\mathbb{R}^d$ are rotationally symmetric distributions $\mathscr{R}_{\theta,f}$ with density $x \mapsto f(\langle x, \theta\rangle)$ with respect to the uniform distribution on the sphere, for some monotonous function f and location parameter $\theta \in \mathscr{S}^{d-1}$, and more specifically von Mises distributions $\mathscr{M}_{\theta,\kappa}$ for the function $f : u \mapsto \exp(\kappa u)$. Two rotationally symmetric distributions are equal up to an isomorphism if they have the same function f or the same concentration parameter κ. In such a case, they are equal up to a rotation.

In this paper, we provide a detailed aternative proof of Gromov's mm spaces reconstruction theorem. Based on this theorem, we introduce a new metric between mm spaces, which is an alternative to the Gromov-Wasserstein distance. This metric is based on characteristic functions of measures defined on matrix spaces. Characteristic functions characterize measures and have been used recently for two-samples tests [21] and for goodness-of-fit tests [15]. We tackle these two types of tests for mm spaces. More precisely, we introduce a new uniformity test on spheres as well as a new two-sample test for mm spaces. For their practical relevance, we study discrete variants of our new metric. Some variant sends an mm space to an element of the Euclidean space $\mathbb{R}^D$ for some dimension D. From this representation, we investigate several classical classification methods on $\mathbb{R}^D$ to perform learning tasks on mm spaces. We show that these methods are able to infer shape and density from samples. Moreover, by turning a gray-level image into a sample of points, we show with numerical experiments that our method is also particularly performant for images classification tasks, with respect to convolution neural networks. Notice that a notable merit of our procedures is that the sample size does not have to be fixed for tests and learning tasks to compare mm spaces from which these samples have been generated.

The paper is organized as follows. In Section 2 we focus on the notion of mm space. First, we recall and prove the Gromov's mm spaces reconstruction theorem. Then, we define a new metric between mm spaces, based on the Gromov's theorem, and provide stability properties and statistical rates for this metric. In Section 3 we focus on testing problems based on this new metric: two-sample tests for mm spaces and goodness-of-fit tests, with the particular case of uniformity tests on spheres. The high performance of the new tests is confirmed with power computation on generated samples in several contexts. In Section 4, we focus on classification tasks for datapoints given by mm spaces, based on the new metric. In particular, we show that our method is a relevant competitor of classical Convolutional Neural Networks for images classification, in particular, in the context of digits recognition.

2 A new metric for metric measure spaces based on Gromov's mm spaces reconstruction theorem

2.1 On the Gromov's mm spaces reconstruction theorem

In this section, we recall the Gromov's mm spaces reconstruction theorem [19, Theorem 3 1/2.5.] and provide an alternative detailed proof of the result.

2.1.1 The theorem

Given two mm spaces, $(\mathscr{X},\mathrm{d}_{\mathscr{X}},\mathrm{P}_{\mathscr{X}})$ and $(\mathscr{Y},\mathrm{d}_{\mathscr{Y}},\mathrm{P}_{\mathscr{Y}})$, the mm spaces reconstruction theorem recalled in Theorem 1 provides a criterion based on the distribution of samples pairwise distance matrices to determine if these two spaces are isomorphic or not. More precisely, the theorem states that the knowledge, for every sample size r, of the distribution of the pairwise distance matrix associated to an r-sample from a measure, suffices to reconstruct the associated metric measure space up to an isomorphism.

For a sample size $r\geq 2$, let $\mathrm{P}_{\mathscr{X},r}$ denote the distribution of the distance matrix $(\mathrm{d}_{\mathscr{X}}(X_i,X_j))_{1\leq i<j\leq r}\in\mathbb{R}_+^{\frac{r(r-1)}{2}}$ of an r-sample $X_1,\ldots,X_r$ (of r independent random variables with the same distribution $\mathrm{P}_{\mathscr{X}}$) from $\mathrm{P}_{\mathscr{X}}$. This distribution $\mathrm{P}_{\mathscr{X},r}=D_{\mathscr{X},r\sharp}(\mathrm{P}_{\mathscr{X}}^{\otimes r})$ is the pushforward of the measure $\mathrm{P}_{\mathscr{X}}^{\otimes r}$ of an r-sample from $\mathrm{P}_{\mathscr{X}}$, with the function $D_{\mathscr{X},r}:\mathscr{X}^r\to\mathbb{R}_+^{\frac{r(r-1)}{2}}$, $(x_1,x_2,\ldots,x_r)\mapsto(\mathrm{d}_{\mathscr{X}}(x_i,x_j))_{1\leq i<j\leq r}$. Within this framework, the theorem states as follows.

▶ **Theorem 1** (mm spaces reconstruction theorem [19]). *Let $(\mathscr{X},\mathrm{d}_{\mathscr{X}},\mathrm{P}_{\mathscr{X}})$ and $(\mathscr{Y},\mathrm{d}_{\mathscr{Y}},\mathrm{P}_{\mathscr{Y}})$ be two mm spaces. That is, each of them is a separable complete metric space equipped with a Borel probability measure whose support coincides with the whole space. Then, the following statements are equivalent:*

1. *The measures $\mathrm{P}_{\mathscr{X},r}$ and $\mathrm{P}_{\mathscr{Y},r}$ are equal for every $r\geq 2$.*
2. *The mm spaces $(\mathscr{X},\mathrm{d}_{\mathscr{X}},\mathrm{P}_{\mathscr{X}})$ and $(\mathscr{Y},\mathrm{d}_{\mathscr{Y}},\mathrm{P}_{\mathscr{Y}})$ are isomorphic.*

If point 1 is a direct consequence of point 2, the proof of the converse relies on the theorem of extension of isometries (Lemma 8) and is detailled in the following section. We illustrate this theorem with the following Example 2 that provides examples of isomorphic and non isomorphic mm spaces.

▶ **Example 2.** Let $\Xi_{\mathscr{X}} = (\mathscr{X}, \mathrm{d}_{\mathscr{X}}, \mathrm{P}_{\mathscr{X}})$, $\Xi_{\mathscr{Y}} = (\mathscr{Y}, \mathrm{d}_{\mathscr{Y}}, \mathrm{P}_{\mathscr{Y}})$ and $\Xi_{\mathscr{Z}} = (\mathscr{Z}, \mathrm{d}_{\mathscr{Z}}, \mathrm{P}_{\mathscr{Z}})$ be three mm spaces defined by $\mathscr{X} = \mathscr{Z} = \{0, 1\}$ and $\mathscr{Y} = \{1, 2\}$. These subsets of $\mathbb{R}$ are equipped with the absolute value $\mathrm{d}_{\mathscr{X}} = \mathrm{d}_{\mathscr{Y}} = \mathrm{d}_{\mathscr{Z}} = |\cdot - \cdot|$. We consider the following discrete probability measures on these spaces: $\mathrm{P}_{\mathscr{X}} = \frac{1}{4}\delta_0 + \frac{3}{4}\delta_1$ puts a mass $\frac{1}{4}$ to the point 0 and a mass $\frac{3}{4}$ to the point 1, $\mathrm{P}_{\mathscr{Y}} = \frac{3}{4}\delta_1 + \frac{1}{4}\delta_2$, and $\mathrm{P}_{\mathscr{Z}} = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$. Then, the spaces $\Xi_{\mathscr{X}}$ and $\Xi_{\mathscr{Y}}$ are isomorphic, but they are not isomorphic to $\Xi_{\mathscr{Z}}$. Indeed, the map $\phi : \mathscr{X} \mapsto \mathscr{Y}$ so that $\phi(0) = 2$ and $\phi(1) = 1$ is an isomorphism between $\Xi_{\mathscr{X}}$ and $\Xi_{\mathscr{Y}}$, but there is no possible isomorphism between $\Xi_{\mathscr{X}}$ and $\Xi_{\mathscr{Z}}$ since the mass repartition differs. The Gromov's theorem states that for every $r \geq 2$, $\mathrm{P}_{\mathscr{X},r} = \mathrm{P}_{\mathscr{Y},r}$, but for some $r > 2$, $\mathrm{P}_{\mathscr{X},r} \neq \mathrm{P}_{\mathscr{Z},r}$. Indeed, we get that $\mathrm{P}_{\mathscr{X},2} = \mathrm{P}_{\mathscr{Y},2} = \frac{10}{16}\delta_0 + \frac{6}{16}\delta_1$, whereas $\mathrm{P}_{\mathscr{Z},2} = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$. Moreover, $\mathrm{P}_{\mathscr{X},3} = \mathrm{P}_{\mathscr{Y},3} = \frac{28}{64}\delta_{(0,0,0)} + \frac{12}{64}\left(\delta_{(1,1,0)} + \delta_{(1,0,1)} + \delta_{(0,1,1)}\right)$, whereas $\mathrm{P}_{\mathscr{Z},3} = \frac{1}{4}\left(\delta_{(0,0,0)} + \delta_{(1,1,0)} + \delta_{(1,0,1)} + \delta_{(0,1,1)}\right)$.

2.1.2 Proof of the theorem

In this section, we prove that point 2 is a consequence of point 1 in Theorem 1. The idea is as follows. Let $(\mathscr{X}, \mathrm{d}_{\mathscr{X}}, \mathrm{P}_{\mathscr{X}})$ and $(\mathscr{Y}, \mathrm{d}_{\mathscr{Y}}, \mathrm{P}_{\mathscr{Y}})$ be two mm spaces for which the measures $\mathrm{P}_{\mathscr{X},r}$ and $\mathrm{P}_{\mathscr{Y},r}$ coincide for every $r \geq 2$. We build an isomorphism $\tilde{\psi}$ between these two mm spaces as follows. First, we define an isomorphism $\psi : x_i^* \mapsto y_i^*$ on a dense countable subset of the separable set $\mathscr{X}$: $\{x_i^*,\, i \in \mathbb{N}^*\}$. Then, we extend this isomorphism to an isomorphism $\tilde{\psi}$ from $\mathscr{X}$ to $\mathscr{Y}$, using Lemma 8. Finally, we prove that this isomorphism is one-to-one and onto, and sends the measure $\mathrm{P}_{\mathscr{X}}$ to $\mathrm{P}_{\mathscr{Y}}$, using the monotone class lemma, after proving equality of the measures $\mathrm{P}_{\mathscr{Y}}$ and $\tilde{\psi}_\sharp(\mathrm{P}_{\mathscr{X}})$ on a specific pi-system containing balls and intersections of balls.

The precise proof is as follows. Let $(\mathscr{X}, \mathrm{d}_{\mathscr{X}}, \mathrm{P}_{\mathscr{X}})$ and $(\mathscr{Y}, \mathrm{d}_{\mathscr{Y}}, \mathrm{P}_{\mathscr{Y}})$ be two mm spaces satisfying point 1 in Theorem 1. Let $\mathbf{x}^* = (x_i^*)_{i\in\mathbb{N}^*}$ be as above. Let R_1 be the set of sphere radii with center in $\mathbf{x}^*$ that are problematic in the sense that the $\mathrm{P}_{\mathscr{X}}$ mass of one such sphere does not vanish: $R_1 = \bigcup_{i\in\mathbb{N}^*}\{\rho \in \mathbb{R}_+,\, \mathrm{P}_{\mathscr{X}}(\partial\mathrm{B}(x_i^*, \rho)) > 0\}$. This set is a countable subset of $\mathbb{R}_+$, as a countable union of countable sets. Then we can define a sequence of non problematic radii $(\rho_n)_{n\in\mathbb{N}}$ in $\mathbb{R}_+^*\backslash R_1$, dense in $\mathbb{R}_+$, and set $R = \{\rho_n,\, n \in \mathbb{N}\} \cup \{\infty\}$. Then, $\bigcup_{n\in\mathbb{N}}\{\phi : \mathbb{N} \mapsto R \mid \forall\, i > n,\, \phi(i) = \infty\}$ is also countable. Thus, we can enumerate its elements in a sequence $(\phi_n)_{n\in\mathbb{N}^*}$. Now, we define the map $M_{\mathscr{X},n} : \mathscr{X}^n \to \mathbb{R}_+^{\frac{3n^2+n}{2}}$ for every $\mathbf{x}_n = (x_1, \ldots, x_n) \in \mathscr{X}^n$ by:

$$M_{\mathscr{X},n}(\mathbf{x}_n) = \left(B_{\mathscr{X},n}(\mathbf{x}_n), I_{\mathscr{X},n}(\mathbf{x}_n), D_{\mathscr{X},n}(\mathbf{x}_n)\right), \tag{1}$$

with:

- $B_{\mathscr{X},n}(\mathbf{x}_n) = (\mathrm{P}_{\mathscr{X}}(\mathrm{B}(x_i, \phi_j(i)))))_{1\leq i,j\leq n}$, the masses of balls,
- $I_{\mathscr{X},n}(\mathbf{x}_n) = \left(\mathrm{P}_{\mathscr{X}}\left(\bigcap_{i=1}^{\ell} \mathrm{B}(x_i, \phi_j(i))\right)\right)_{1\leq j,\ell\leq n}$, the masses of intersections of first balls,
- $D_{\mathscr{X},n}(\mathbf{x}_n) = (\mathrm{d}_{\mathscr{X}}(x_i, x_j))_{1\leq i<j\leq n}$, the pairwise distance matrix.

According to Lemma 9, for every $n \in \mathbb{N}^*$, the pushforwards ${M_{\mathscr{X},n}}_\sharp(\mathrm{P}_{\mathscr{X}}^{\otimes n})$ and ${M_{\mathscr{Y},n}}_\sharp(\mathrm{P}_{\mathscr{Y}}^{\otimes n})$ coincide. Let $(\mathbf{y}^n)_{n\in\mathbb{N}^*}$ be a sequence of $\mathscr{Y}$-valued sequences $\mathbf{y}^n \in \mathscr{Y}^{\mathbb{N}}$, so that $\|M_{\mathscr{X},n}(\mathbf{x}_n^*) - M_{\mathscr{Y},n}(\mathbf{y}_n^n)\|_\infty \leq \frac{1}{n}$, as given by Lemma 10, where $\mathbf{y}_n^n = (y_1^n, \ldots, y_n^n)$ and $\mathbf{x}_n^* = (x_1^*, \ldots, x_n^*)$ are the vectors made of the first n coordinates of $\mathbf{y}^n$ and $\mathbf{x}^*$. A diagonal process, see Lemma 12, provides a subsequence of $(\mathbf{y}^n)_{n\in\mathbb{N}^*}$ whose ith coordinate converges to a point in $\mathscr{Y}$, that we denote by y_i^*. According to Lemma 14, $\mathbf{y}^* = (y_i^*)_{i\in\mathbb{N}}$ satisfies $M_{\mathscr{X},n}(\mathbf{x}_n^*) = M_{\mathscr{Y},n}(\mathbf{y}_n^*)$ for every $n \in \mathbb{N}^*$. In particular, it comes that the map ψ defined on $\mathbf{x}^*$ by $\psi(x_i^*) = y_i^*$ is an isometry since $D_{\mathscr{X},n}(\mathbf{x}_n^*) = D_{\mathscr{Y},n}(\mathbf{y}_n^*)$ for every $n \in \mathbb{N}^*$. This isometry ψ can be extended to an isometry $\tilde{\psi}$ on $\mathscr{X}$ that coincides with ψ on $\mathbf{x}^*$, according to Lemma 8, since $(\mathscr{Y}, \mathrm{d}_{\mathscr{Y}})$ is complete. Moreover, Lemma 15 yields that the sets $\tilde{\psi}(\mathscr{X})$ and $\mathscr{Y}$ are equal. Thus, $\tilde{\psi}$ is an isometry from $\mathscr{X}$ to $\mathscr{Y}$. In addition, we get that $\mathbf{y}^*$ is dense in $\mathscr{Y}$. It remains to prove that $\tilde{\psi}_\sharp(\mathrm{P}_{\mathscr{X}}) = \mathrm{P}_{\mathscr{Y}}$. Note that the collection $\{\bigcap_{i=1}^{\ell}(\mathrm{B}(y_i^*, \phi_j(i))) \mid j, \ell \in \mathbb{N}^*\}$ is a π-system, since it is stable by intersection. It generates the σ-algebra of Borel sets of $\mathscr{Y}$, since it contains all balls centered at points in $\mathbf{y}^*$, with radii in R. Moreover, the measures $\mathrm{P}_{\mathscr{Y}}$ and $\tilde{\psi}_\sharp(\mathrm{P}_{\mathscr{X}})$ coincide on this π-system since we proved that $I_{\mathscr{X},n}(\mathbf{x}_n^*) = I_{\mathscr{Y},n}(\mathbf{y}_n^*)$ for every $n \in \mathbb{N}$. Thus, according to the monotone class lemma, the measures $\mathrm{P}_{\mathscr{Y}}$ and $\tilde{\psi}_\sharp(\mathrm{P}_{\mathscr{X}})$ coincide. This concludes the proof of Theorem 1.

2.2 A new metric based on the distributions of pairwise distance matrices

2.2.1 Definition of the metric

Given two metric measure spaces $\Xi_{\mathscr{X}} = (\mathscr{X}, \mathrm{d}_{\mathscr{X}}, \mathrm{P}_{\mathscr{X}})$ and $\Xi_{\mathscr{Y}} = (\mathscr{Y}, \mathrm{d}_{\mathscr{Y}}, \mathrm{P}_{\mathscr{Y}})$, given $\boldsymbol{\alpha} = (\alpha_r)_{r\geq 2}$ some weights in $\mathbb{R}_+$, $\boldsymbol{\mu} = (\mu_r)_{r\geq 2}$ a sequence of probability measures μ_r on the sphere $\mathscr{S}^{\frac{r(r-1)}{2}-1}$, and ω a probability measure on $\mathbb{R}_+$ such that $\int_{\rho=0}^{+\infty} \rho \mathrm{d}\omega(\rho) < +\infty$, we define the "pseudo-metric" $\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}$ between $\Xi_{\mathscr{X}}$ and $\Xi_{\mathscr{Y}}$ by:

$$\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}}, \Xi_{\mathscr{Y}}) = \sum_{r=2}^{+\infty} \alpha_r \int_{\rho=0}^{+\infty} \int_{\theta\in\mathscr{S}^{\frac{r(r-1)}{2}-1}} |\phi_{\mathscr{X},r}(\rho\theta) - \phi_{\mathscr{Y},r}(\rho\theta)| \,\mathrm{d}\mu_r(\theta)\mathrm{d}\omega(\rho) \tag{2}$$

where $\phi_{\mathscr{X},r}(\rho\theta) = \mathbb{E}_{D\sim\mathrm{P}_{\mathscr{X},r}}[\exp(it\langle D, \rho\theta\rangle)]$ is the characteristic function of the distribution of r-samples pairwise distance matrices, $\mathrm{P}_{\mathscr{X},r}$, defined in Section 2.1.1, at the point $\rho\theta \in \mathbb{R}^{\frac{r(r-1)}{2}}$.

▶ **Proposition 3.** *Let $\boldsymbol{\alpha} = (\alpha_r)_{r\geq 2}$ be some weights in $\mathbb{R}_+$, $\boldsymbol{\mu} = (\mu_r)_{r\geq 2}$ be a sequence of probability measures μ_r on the sphere $\mathscr{S}^{\frac{r(r-1)}{2}-1}$, and ω be a probability measure on $\mathbb{R}_+$ such that $\int_{\rho=0}^{+\infty} \rho \mathrm{d}\omega(\rho) < +\infty$.*

If $\sum_{r=2}^{+\infty} r\alpha_r < +\infty$ and $p = 2$, or $\sum_{r=2}^{+\infty} r^3\alpha_r < +\infty$ and $p = 1$, then, $\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}$ is a pseudo-metric on the set of all mm spaces with finite pth moment (i.e. such that $\int_{\mathscr{X}} \mathrm{d}_{\mathscr{X}}^p(x, x_0)\mathrm{d}\mathrm{P}_{\mathscr{X}}(x) < +\infty$ for some $x_0 \in \mathscr{X}$), in the sense that, for every such mm spaces $\Xi_{\mathscr{X}}$, $\Xi_{\mathscr{Y}}$ and $\Xi_{\mathscr{Z}}$:

1. $\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}}, \Xi_{\mathscr{Y}}) \in \mathbb{R}_+$
2. $\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}}, \Xi_{\mathscr{Y}}) = \mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{Y}}, \Xi_{\mathscr{X}})$
3. $\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}}, \Xi_{\mathscr{Y}}) \leq \mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}}, \Xi_{\mathscr{Z}}) + \mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{Z}}, \Xi_{\mathscr{Y}})$.

Moreover, if $\Xi_{\mathscr{X}}$ and $\Xi_{\mathscr{Y}}$ are isomorphic, then, $\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}}, \Xi_{\mathscr{Y}}) = 0$.

Under the additional assumptions that $\alpha_r > 0$ and $\mathrm{Supp}(\mu_r) = \mathscr{S}^{\frac{r(r-1)}{2}-1}$ for every $r \geq 2$ and that $\mathrm{Supp}(\rho) = \mathbb{R}_+$, then, $\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}$ is a metric in the sense that for every mm spaces $\Xi_{\mathscr{X}}$ and $\Xi_{\mathscr{Y}}$: $\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}}, \Xi_{\mathscr{Y}}) = 0$ if and only if $\Xi_{\mathscr{X}}$ and $\Xi_{\mathscr{Y}}$ are isomorphic.

The proof of Proposition 3 is available in Section A.2.1.

2.2.2 Stability properties of the new metric

Given two metric measure spaces $\Xi_{\mathscr{X}} = (\mathscr{X}, \mathrm{d}_{\mathscr{X}}, \mathrm{P}_{\mathscr{X}})$ and $\Xi_{\mathscr{Y}} = (\mathscr{Y}, \mathrm{d}_{\mathscr{Y}}, \mathrm{P}_{\mathscr{Y}})$ with finite pth moment for some $p \in \{1, 2\}$, in this section, we show that the metric defined by (2) is stable with respect to Wasserstein and Gromov-Wasserstein distances. The p-Gromov-Wasserstein distance between $\Xi_{\mathscr{X}}$ and $\Xi_{\mathscr{Y}}$, $\mathscr{GW}_p(\Xi_{\mathscr{X}},\Xi_{\mathscr{Y}})$ is defined in [22] by

$$\mathscr{GW}_p^p(\Xi_{\mathscr{X}},\Xi_{\mathscr{Y}}) = \inf_{\pi\in\Pi(\mathrm{P}_{\mathscr{X}},\mathrm{P}_{\mathscr{Y}})} \int_{\mathscr{X}\times\mathscr{Y}}\int_{\mathscr{X}\times\mathscr{Y}} (\mathrm{d}_{\mathscr{X}}(x,x') - \mathrm{d}_{\mathscr{Y}}(y,y'))^p \,\mathrm{d}\pi(x,y)\mathrm{d}\pi(x',y'), \quad (3)$$

where $\Pi(\mathrm{P}_{\mathscr{X}},\mathrm{P}_{\mathscr{Y}})$ is the set of all transport plans (probability distributions) on $\mathscr{X}\times\mathscr{Y}$ with first marginal $\mathrm{P}_{\mathscr{X}}$ and second marginal $\mathrm{P}_{\mathscr{Y}}$. Roughly, the p-Gromov-Wasserstein distance corresponds to the minimal p-norm mean difference between the distance between two independent random variables in $\mathscr{X}$ and the distance between two independent random variables in $\mathscr{Y}$. When $(\mathscr{X},\mathrm{d}_{\mathscr{X}})$ and $(\mathscr{Y},\mathrm{d}_{\mathscr{Y}})$ are subspaces of a same space, then, the p-Wasserstein distance between the two probability measures $\mathrm{P}_{\mathscr{X}}$ and $\mathrm{P}_{\mathscr{Y}}$, $\mathscr{W}_p(\mathrm{P}_{\mathscr{X}},\mathrm{P}_{\mathscr{Y}})$, is defined in [31] by

$$\mathscr{W}_p^p(\mathrm{P}_{\mathscr{X}},\mathrm{P}_{\mathscr{Y}}) = \inf_{\pi\in\Pi(\mathrm{P}_{\mathscr{X}},\mathrm{P}_{\mathscr{Y}})} \int_{\mathscr{X}} \mathrm{d}_{\mathscr{X}}^p(x,y)\mathrm{d}\pi(x,y). \quad (4)$$

▶ **Proposition 4** (Stability). *Let $\boldsymbol{\alpha}$, $\boldsymbol{\mu}$, ω, be as in Proposition 3. Given two metric measure spaces $\Xi_{\mathscr{X}} = (\mathscr{X},\mathrm{d}_{\mathscr{X}},\mathrm{P}_{\mathscr{X}})$ and $\Xi_{\mathscr{Y}} = (\mathscr{Y},\mathrm{d}_{\mathscr{Y}},\mathrm{P}_{\mathscr{Y}})$ with finite pth moment, we get that:*

- *If $p=1$:*

$$\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}},\Xi_{\mathscr{Y}}) \leq \left(\sum_{r=2}^{+\infty}\alpha_r\left(\frac{r(r-1)}{2}\right)^{\frac{3}{2}}\right)\left(\int_{\rho=0}^{+\infty}\rho\mathrm{d}\omega(\rho)\right)\mathscr{GW}_1(\Xi_{\mathscr{X}},\Xi_{\mathscr{Y}}).$$

Moreover, when the spaces $(\mathscr{X},\mathrm{d}_{\mathscr{X}})$ and $(\mathscr{Y},\mathrm{d}_{\mathscr{Y}})$ are subspaces of a same space, we get that:

$$\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}},\Xi_{\mathscr{Y}}) \leq 2\left(\sum_{r=2}^{+\infty}\alpha_r\left(\frac{r(r-1)}{2}\right)^{\frac{3}{2}}\right)\left(\int_{\rho=0}^{+\infty}\rho\mathrm{d}\omega(\rho)\right)\mathscr{W}_1(\mathrm{P}_{\mathscr{X}},\mathrm{P}_{\mathscr{Y}}).$$

- *If $p=2$:*

$$\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}},\Xi_{\mathscr{Y}}) \leq \left(\sum_{r=2}^{+\infty}\alpha_r\sqrt{\frac{r(r-1)}{2}}\right)\left(\int_{\rho=0}^{+\infty}\rho\mathrm{d}\omega(\rho)\right)\mathscr{GW}_2(\Xi_{\mathscr{X}},\Xi_{\mathscr{Y}}).$$

Moreover, when the spaces $(\mathscr{X},\mathrm{d}_{\mathscr{X}})$ and $(\mathscr{Y},\mathrm{d}_{\mathscr{Y}})$ are subspaces of a same space, we get that:

$$\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}},\Xi_{\mathscr{Y}}) \leq 2\left(\sum_{r=2}^{+\infty}\alpha_r\sqrt{\frac{r(r-1)}{2}}\right)\left(\int_{\rho=0}^{+\infty}\rho\mathrm{d}\omega(\rho)\right)\mathscr{W}_2(\mathrm{P}_{\mathscr{X}},\mathrm{P}_{\mathscr{Y}}).$$

The proof of Proposition 4 is available in Section A.2.2. This proposition enhances that two mm spaces close in terms of Gromov-Wasserstein or Wasserstein distance will be close with respect to the new metric $\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}$. In particular, this occurs when we have access to some mm space $\Xi_{\mathscr{X}} = (\mathscr{X},\mathrm{d}_{\mathscr{X}},\mathrm{P}_{\mathscr{X}})$ through a sample of n points $(X_1\ldots,X_n)$. In this case, the space $\Xi_{\mathscr{X}}$ is approximated by the discrete mm space $\Xi_{\mathscr{X}_n} = (\mathscr{X}_n,\mathrm{d}_{\mathscr{X}},\mathrm{P}_{\mathscr{X}_n})$, also called empirical mm space, with $\mathscr{X}_n = \{X_1,\ldots,X_n\}$ and $\mathrm{P}_{\mathscr{X}_n} = \frac{1}{n}\sum_{i=1}^n\delta_{X_i}$. If $\mathrm{P}_{\mathscr{X}}$ satisfies that $\int_{\mathscr{X}}\mathrm{d}_{\mathscr{X}}(x,x_0)^p\mathrm{d}\mathrm{P}_{\mathscr{X}}(x) < +\infty$ for some $x_0\in\mathscr{X}$ and $p\geq 1$, then, $\mathscr{W}_p(\mathrm{P}_{\mathscr{X}},\mathrm{P}_{\mathscr{X}_n})$ converges to 0, as $n\to+\infty$, almost surely. This property recalled for instance in [5, Theorem 2.13] is a consequence of the Varadarajan theorem for separable metric spaces, that states the convergence in distribution of $\mathrm{P}_{\mathscr{X}_n}$ to $\mathrm{P}_{\mathscr{X}}$, for almost every sample $\mathscr{X}_n$ from $\mathrm{P}_{\mathscr{X}}$. As a consequence, we get that:

▶ **Corollary 5.** *Let $\boldsymbol{\alpha}$, $\boldsymbol{\mu}$, ω, be as in Proposition 3. If $\Xi_{\mathscr{X}}$ is a mm space so that $\int_{\mathscr{X}} \mathrm{d}_{\mathscr{X}}^p(x, x_0)\mathrm{dP}_{\mathscr{X}}(x) < \infty$ for some $x_0 \in \mathscr{X}$ (with $p = 1$ or $p = 2$), then, with probability 1, we get that:*

$$\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}}, \Xi_{\mathscr{X}_n}) \to 0, \text{ as } n \to +\infty.$$

According to Corollary 5, for large enough sample sizes, it will always be possible to distinguish between two different mm spaces (cf. Section 3). Specific rates of convergence have been derived in different contexts for the Wasserstein distance, see for instance [16, 33], just to name a few. In particular, for the Euclidean space $\mathbb{R}^d$, rates for the 1-Wasserstein distance are of order $n^{-\frac{1}{d}}$. This rate gets very slow when the dimension increases, this is the curse of dimensionality. The new distance that we propose to compare mm spaces does not suffer from this. Indeed, we prove in Proposition 6, that for any mm space, when we have access to this mm space only through a sample of n points, the difference between the unknown mm space and the empirical mm space based on this sample is of order $n^{-\frac{1}{2}}$, in terms of $\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}$. This corresponds to the classical parametric rate in statistics.

▶ **Proposition 6.** *Let $\Xi_{\mathscr{X}} = (\mathscr{X}, \mathrm{d}_{\mathscr{X}}, \mathrm{P}_{\mathscr{X}})$ be some metric measure space, and $\Xi_{\mathscr{X}_n} = (\mathscr{X}_n, \mathrm{d}_{\mathscr{X}}, \mathrm{P}_{\mathscr{X}_n})$ be its empirical version, based on some n-sample. Then, for every $\ell > 0$, we get that, with probability at least $1 - n^{-\ell}$:*

$$\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}}, \Xi_{\mathscr{X}_n}) \leq \mathbb{E}\left[\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}}, \Xi_{\mathscr{X}_n})\right] + \sqrt{\frac{2\ell\log(n)}{n}} \sum_{r=2}^{+\infty} r\alpha_r,$$

with

$$\mathbb{E}\left[\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}}, \Xi_{\mathscr{X}_n})\right] \leq \frac{4}{\sqrt{n}} \sum_{r=2}^{+\infty} \sqrt{r}\alpha_r + \frac{1}{n}\sum_{r=2}^{n} r^2\alpha_r \leq \frac{5}{\sqrt{n}} \sum_{r=2}^{+\infty} r^{\frac{3}{2}}\alpha_r.$$

The proof of Proposition 6 is available in Section A.2.3. The rates obtained in Proposition 6 are tight in the sense that for some discrete metric space with $|\mathscr{X}| = V$, we get a lower bound of order $\frac{1}{\sqrt{n}}$, as noticed in the following Example 7.

▶ **Example 7.** Let $\Xi_{\mathscr{X}} = (\mathscr{X}, \mathrm{d}_{\mathscr{X}}, \mathrm{P}_{\mathscr{X}})$ be a mm space with $|\mathscr{X}| = V$ elements, so that $\mathrm{d}_{\mathscr{X}}(x, y) = \mathbb{1}_{x \neq y}$ for every $x, y \in \mathscr{X}$ and $\mathrm{P}_{\mathscr{X}} = \frac{1}{V}\sum_{x\in\mathscr{X}} \delta_x$ is the uniform measure on $\mathscr{X}$. For some absolute constant $C > 0$, we get that:

$$\lim \inf_{n\to+\infty} \mathbb{E}[\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}}, \Xi_{\mathscr{X}_n})] \geq C\frac{\alpha_2 \int_0^{+\infty} |1 - \exp(i\rho)| \,\mathrm{d}\omega(\rho)}{\sqrt{n}\sqrt{V}}.$$

The proof of the lower bound in Example 7 is available in Section A.2.4.

2.2.3 Computational considerations

In practice, we may not have access to an mm space $\Xi_{\mathscr{X}} = (\mathscr{X}, \mathrm{d}_{\mathscr{X}}, \mathrm{P}_{\mathscr{X}})$, but to an empirical version $\Xi_{\mathscr{X}_n} = (\mathscr{X}_n, \mathrm{d}_{\mathscr{X}}, \mathrm{P}_{\mathscr{X}_n})$, based on an n-sample $(X_1, \ldots, X_n)$, generated according to the distribution $\mathrm{P}_{\mathscr{X}}$, where $\mathrm{P}_{\mathscr{X}_n} = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$ is the uniform measure on $\mathscr{X}_n = \{X_1, \ldots, X_n\}$. For fixed $r \geq 2$, $\rho > 0$ and $\theta \in \mathscr{S}^{\frac{r(r-1)}{2}-1}$, the computation of the empirical characteristic function at the point $\rho\theta$:

$$\phi_{\mathscr{X}_n,r}(\rho\theta) = \frac{1}{n^r} \sum_{1\leq i_1,\ldots,i_r\leq n} \exp(i\langle\left(\mathrm{d}_{\mathscr{X}}(X_{i_\ell}, X_{i_j})\right)_{1\leq\ell<j\leq r}, \rho\theta\rangle), \tag{5}$$

requires $\frac{r(r-1)}{2}n^r$ operations. For large sample size n, this complexity is prohibitive. We may prefer to use a bootstrap strategy, that consists, given $B \in \mathbb{N}^*$, to generate B independent samples from $\mathrm{P}_{\mathscr{X}_n}$, $(\mathscr{X}_n^b = (X_1^b, \ldots, X_r^b))_{1 \le b \le B}$. It means that $X_1^b, \ldots, X_r^b$ are picked from $\mathscr{X}_n$ uniformly, independently and with replacement. Then, we estimate $\phi_{\mathscr{X}_n,r}(\rho\theta)$ with the bootstrap version:

$$\phi_{\mathscr{X}_n,r}^{\mathrm{boot},(B)}(\rho\theta) = \frac{1}{B}\sum_{b=1}^{B} \exp(i\langle(\mathrm{d}_{\mathscr{X}}(X_\ell^b, X_j^b))_{1\le \ell,j\le r}, \rho\theta\rangle), \tag{6}$$

or with its symmetrized version:

$$\phi_{\mathscr{X}_n,r}^{\mathrm{boot,sym},(B)}(\rho\theta) = \frac{1}{B}\sum_{b=1}^{B}\frac{1}{r!}\sum_{\sigma\in\mathfrak{S}_r} \exp(i\langle(\mathrm{d}_{\mathscr{X}}(X_{\sigma(\ell)}^b, X_{\sigma(j)}^b))_{1\le \ell,j\le r}, \rho\theta\rangle). \tag{7}$$

The summation over an infinite number of parameters $r \ge 2$ is impossible in practice, therefore, we may focus on truncated vectors $\boldsymbol{\alpha} = (\alpha_r)_{r\ge 2}$ with $\alpha_r = 0$ for $r \ge R$ for some fixed $R \ge 3$. As well, the measures μ_r and ω may be chosen as discrete and finitely supported: $\mu_r = \sum_{i=1}^{N_r}\beta_{r,i}\delta_{\theta_{r,i}}$ and $\omega = \sum_{i=1}^{M_r}\gamma_{r,i}\delta_{\rho_{r,i}}$ for some non negative weights $(\beta_{r,i})_{1\le i\le N_r}$ and $(\gamma_{r,i})_{1\le i\le M_r}$ with sum 1. The global complexity would then become $\sum_{r=2}^{R} N_r M_r \frac{r(r-1)}{2} n^r$, and respectively, $\sum_{r=2}^{R} N_r M_r \frac{r(r-1)}{2} B$ for the bootstrap and $\sum_{r=2}^{R} N_r M_r \frac{r(r-1)}{2} r! B$ for the symmetrized bootstrap versions.

According to the Gromov's theorem, given a finite number of mm spaces, it is always possible to find discrete measures such that the pairwise $\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}$-distances are all positive. However, when the discrete measures are fixed in advance, there could exist two mm spaces at 0 $\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}$-distance. A strategy to deal with this issue would be to consider an equivalent l_2 version of $\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}$:

$$\tilde{\mathrm{d}}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega} : (\Xi_{\mathscr{X}}, \Xi_{\mathscr{Y}}) \mapsto \sum_{r=2}^{+\infty}\alpha_r \int_{\rho=0}^{+\infty}\int_{\theta\in\mathscr{S}^{\frac{r(r-1)}{2}-1}} |\phi_{\mathscr{X},r}(\rho\theta) - \phi_{\mathscr{Y},r}(\rho\theta)|^2 \,\mathrm{d}\mu_r(\theta)\mathrm{d}\omega(\rho).$$

These two versions are equivalent in the sense that:

$$\frac{1}{2}\tilde{\mathrm{d}}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}}, \Xi_{\mathscr{Y}}) \le \mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}}, \Xi_{\mathscr{Y}}) \le \sqrt{\tilde{\mathrm{d}}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}}, \Xi_{\mathscr{Y}})}.$$

There exists a closed form for $\tilde{\mathrm{d}}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}$. This closed form is obtained by following the proof of [1, Lemma 1]:

$$\tilde{\mathrm{d}}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}}, \Xi_{\mathscr{Y}}) = \sum_{r=2}^{+\infty}\alpha_r\Big(\frac{1}{n^{2r}}\sum_{i_1,\ldots,i_r=1}^{n}\sum_{j_1,\ldots,j_r=1}^{n} u(D_{i_1,\ldots,i_r}^{\mathscr{X}} - D_{j_1,\ldots,j_r}^{\mathscr{X}}) \tag{8}$$

$$+\frac{1}{m^{2r}}\sum_{r_1,\ldots,r_r=1}^{m}\sum_{v_1,\ldots,v_r=1}^{m} u(D_{r_1,\ldots,r_r}^{\mathscr{Y}} - D_{v_1,\ldots,v_r}^{\mathscr{Y}}) \tag{9}$$

$$-\frac{2}{n^r m^r}\sum_{i_1,\ldots,i_r=1}^{n}\sum_{j_1,\ldots,j_r=1}^{m} u(D_{i_1,\ldots,i_r}^{\mathscr{X}} - D_{j_1,\ldots,j_r}^{\mathscr{Y}})\Big), \tag{10}$$

where $D_{i_1,\ldots,i_r}^{\mathscr{X}} = (\mathrm{d}_{\mathscr{X}}(X_{i_\ell}, X_{i_k}))_{1\le \ell<k\le r}$ and where $u(t)$ is the real part of the characteristic function of the distribution of $\rho\theta$, that is, $u(t) = \int_{\mathscr{S}^{\frac{r(r-1)}{2}-1}}\int_0^{+\infty}\cos(\langle\rho\theta, t\rangle)\mathrm{d}\omega(\rho)\mathrm{d}\mu_r(\theta)$. As noticed in [1, Example 1], when the distribution of $\rho\theta$ is the $\left(\frac{r(r-1)}{2}\right)$-variate normal distribution with mean 0 and variance-covariance matrix identity, then, $u(t) = \exp(-\frac{1}{2}\|t\|^2)$.

3 Testing problems for mm spaces

3.1 Two-sample testing

Testing the equality of two measures from two samples using the characteristic function has already been considered in the litterature, see [1] for instance. In the context of mm spaces, we aim at testing $\mathcal{H}_0$:"$\Xi_{\mathscr{X}}$ and $\Xi_{\mathscr{Y}}$ are isomorphic", versus $\mathcal{H}_1$:"$\Xi_{\mathscr{X}}$ and $\Xi_{\mathscr{Y}}$ are not isomorphic". This question has been considered in [7], using distance-to-measure (dtm) signatures. In this paper, given a pseudo-metric d between mm spaces, we investigate tests that reject $\mathcal{H}_0$ when the test statistic $T = \mathrm{d}(\Xi_{\mathscr{X}_n}, \Xi_{\mathscr{Y}_m})$ is larger than some critical value $\hat{c}_\alpha$, that depends on the level $\alpha \in (0,1)$ of the test. We recall that the level of the test is defined by the probability to make the wrong decision (i.e. reject $\mathcal{H}_0$) based on the test statistic under $\mathcal{H}_0$. As mentioned in [1, Theorem 5], the problem is that under $\mathcal{H}_0$, the distribution of the test statistic T depends on the mm space $\mathscr{X}$. Solutions to this problem consists in using boostrap or permutation to mimic the behaviour of the test statistic T under $\mathcal{H}_0$. Both procedures are proved correct in the context of characteristic functions, [1, Theorems 3,4,5]. In the appendix, Section B, we focus on the permutation procedure, as described in [1]. We compute the level and the power of the test under several alternatives. The power is defined as the probability to make the right decision (i.e. reject $\mathcal{H}_0$) under an alternative to hypothesis $\mathcal{H}_0$. Notice that for alternatives close to hypothesis $\mathcal{H}_0$, the power should be close to the level α of the test. However, in the experiments in the appendix, the level is sometimes larger than the parameter α. Therefore, we propose in this section a slightly less powerful procedure that reaches the correct level, as follows:

1. Set $N = \min(n, m)/2$
2. Consider a subsample of size N of $\mathscr{X}_n$, $\mathscr{X}_N$, and a subsample of size N of $\mathscr{Y}_m$, $\mathscr{Y}_N$.
3. Calculate $T_{N,obs} = \mathrm{d}(\mathscr{X}_N, \mathscr{Y}_N)$ from $\mathscr{X}_N$ and $\mathscr{Y}_N$.
4. Generate B subsamples of size N, $\mathscr{X}_N^b$ from $\mathscr{X}_n$.
5. Calculate $T^b_{\mathscr{X},N} = \mathrm{d}(\mathscr{X}_N^b, \mathscr{X}_N^{b+(B/2)})$ for each $b \in \{1, \ldots, B/2\}$.
6. Generate B subsamples of size N, $\mathscr{Y}_N^b$ from $\mathscr{Y}_m$.
7. Calculate $T^b_{\mathscr{Y},N} = \mathrm{d}(\mathscr{Y}_N^b, \mathscr{Y}_N^{b+(B/2)})$ for each $b \in \{1, \ldots, B/2\}$.
8. Approximate the p-value of the test by:

$$\hat{p} = \frac{\mathrm{Card}\{b \in \{1, \ldots, B/2\},\, T^b_{\mathscr{X},N} \geq T_{N,obs}\} + \mathrm{Card}\{b \in \{1, \ldots, B/2\},\, T^b_{\mathscr{Y},N} \geq T_{N,obs}\}}{B}.$$

The power of the test with level $\alpha \in \{0.05, 0.1\}$ is estimated by $\hat{\mathrm{P}}_\alpha = \frac{\mathrm{Card}\{m \in \{1,\ldots,M\},\, \hat{p}_m \leq \alpha\}}{M}$, after $M = 1000$ independent replications of the experiment. Since mm spaces encode both an information about the shape and the measure, we illustrate the testing procedure in two different frameworks. First, we generate points on the l_p-ball in $\mathbb{R}^2$ with the following procedure:

$$X_i = \frac{Z_i}{\|Z_i\|_p} = \frac{Z_i}{(Z_{i,1}^p + Z_{i,2}^p)^{\frac{1}{p}}}$$

with $Z_1, \ldots, Z_n$ i.i.d. from the standard normal distribution on $\mathbb{R}^2$. We test $\mathcal{H}_0$:"$p_1 = p_2$", against $\mathcal{H}_1$:"$p_1 \neq p_2$". Then, we generate points on the circle in $\mathbb{R}^2$ according to a von Mises distribution $\mathscr{M}_{\theta,\kappa}$. Since two different parameters θ lead to isomorphic mm spaces, we test equality of the concentration parameter: $\mathcal{H}_0$:"$\kappa_1 = \kappa_2$", versus $\mathcal{H}_1$:"$\kappa_1 \neq \kappa_2$". This last problem has been tackled in [29, 30].

We propose two different methods based on characteristic functions. First, we use the closed form $\tilde{\mathrm{d}}^R$ defined by (8), for $\alpha_R = 1$ and $\alpha_r = 0$ for $r \neq R$, and $\tilde{\mathrm{d}}^{\mathrm{boot},R}$ a bootstrapped version where means are approximated through 1000 random selections of uplets (2-uplets for $R = 2$, 6-uplets for $R = 3$) of indices. Then, we consider $\mathrm{d}^{\mathrm{boot},R}$ the distance (2) with finitely supported measures μ_R and ω, for which the characteristic functions are estimated by bootstrap as in (6). For $R = 2$, we evaluate the characteristic function at $t \in \{0.1, 0.2, 0.5, 1.0, 5.0, 10.0, 100.0\}$, and for $R = 5$, at $t \in \{(0.5) \in \mathbb{R}^{\frac{2(2-1)}{2}}, (0.5, 0.5, 0.5) \in \mathbb{R}^{\frac{3(3-1)}{2}}, \dots, (0.5, \dots, 0.5) \in \mathbb{R}^{\frac{5(5-1)}{2}}\}$. As an alternative, we consider $\mathrm{d}^{\mathrm{shp}}$, the 1-Wasserstein distance (i.e. the L_1-norm between the quantile functions) between the shape signatures (i.e. $\mathrm{P}_{\mathscr{X}_n,r}$ and $\mathrm{P}_{\mathscr{Y}_m,r}$ for $r = 2$). The second alternative that we consider is $\mathrm{d}^{\mathrm{dtm},h}$, the 1-Wasserstein distance between the DTM-signatures of [7]. The DTM-signature is the distribution of the root mean square of the distances of points in $\mathscr{X}_n$ to their hn-nearest neighbours in $\mathscr{X}_n$, for some parameter $h \in (0, 1)$. The power of the tests are provided in Table 1 for the test on the parameter p of the l_p-ball, and in Table 2 for the test on the parameter κ of the von Mises distribution. Our new methods are good competitors for the von Mises distributions, and they outperform the other methods for the l_p-balls alternatives. The levels are correct, of order 5% (middle line).

p_2	$\mathrm{d}^{\mathrm{boot},2}$	$\mathrm{d}^{\mathrm{boot},5}$	$\mathrm{d}^{\mathrm{shp}}$	$\mathrm{d}^{\mathrm{dtm},0.05}$	$\mathrm{d}^{\mathrm{dtm},0.5}$	$\mathrm{d}^{\mathrm{dtm},1}$
1.4	0.285	0.425	0.191	0.052	0.055	0.098
1.45	0.113	0.157	0.088	0.062	0.068	0.08
1.5	0.055	0.061	0.066	0.054	0.07	0.076
1.55	0.102	0.167	0.089	0.073	0.072	0.072
1.6	0.241	0.435	0.148	0.064	0.071	0.085

Table 1 Two-sample test, l_p-balls, $p_1 = 1.5$, sample sizes $n = m = 100$, $\alpha = 0.05$.

p_2	$\mathrm{d}^{\mathrm{boot},2}$	$\mathrm{d}^{\mathrm{boot},5}$	$\mathrm{d}^{\mathrm{shp}}$	$\mathrm{d}^{\mathrm{dtm},0.05}$	$\mathrm{d}^{\mathrm{dtm},0.5}$	$\mathrm{d}^{\mathrm{dtm},1}$
1.0	0.668	0.637	0.704	0.331	0.666	0.696
1.5	0.228	0.199	0.236	0.112	0.220	0.236
2.0	0.066	0.067	0.060	0.053	0.067	0.062
2.5	0.144	0.129	0.151	0.095	0.143	0.146
3.0	0.395	0.343	0.424	0.184	0.381	0.437

Table 2 Two-sample test, von Mises distribution, $\kappa_1 = 2.0$, sample sizes $n = m = 100$, $\alpha = 0.05$.

The computational time (is seconds) of pseudo-distances are available in Table 3, after 1000 Monte Carlo replications (10 for method 2 with sample size 20, and method 1 with sample size 100).

n	$\tilde{\mathrm{d}}^2$	$\tilde{\mathrm{d}}^3$	$\tilde{\mathrm{d}}^{\mathrm{boot},2}$	$\tilde{\mathrm{d}}^{\mathrm{boot},3}$	$\mathrm{d}^{\mathrm{boot},2}$	$\mathrm{d}^{\mathrm{boot},5}$	$\mathrm{d}^{\mathrm{shp}}$	$\mathrm{d}^{\mathrm{dtm},0.05}$	$\mathrm{d}^{\mathrm{dtm},1}$
5	2.48e-5	1.31e-3	0.11e-3	4.08e-3	1.26e-3	3.12e-3	1.77e-6	1.34e-6	1.65e-6
10	0.28e-3	0.1157	0.11e-3	4.06e-3	1.55e-3	3.49e-3	1.14e-5	3.66e-6	3.91e-6
20	4.29e-3	8.3077	0.11e-3	4.11e-3	1.65e-3	4.00e-3	1.00e-5	1.08e-5	3.13e-5
50	0.2764	***	0.10e-3	4.10e-3	1.29e-3	3.18e-3	5.58e-5	8.61e-5	8.25e-5
100	3.5973	***	0.11e-3	4.07e-3	1.24e-3	3.00e-3	0.21e-3	0.38e-3	0.36e-3
1000	***	***	0.36e-3	4.72e-3	1.77e-3	4.36e-3	0.0257	0.0400	0.0426

Table 3 Pseudo-distances computational time

This computational time does not take into account the computational time of the distance matrix of the pooled sample, that should be done for all of the methods. Notice that the computational time does not change for $\mathrm{d}^{\mathrm{boot},2}$ and $\mathrm{d}^{\mathrm{boot},5}$ nor for $\tilde{\mathrm{d}}^{\mathrm{boot},2}$ and $\tilde{\mathrm{d}}^{\mathrm{boot},3}$ since the number of bootstrap replications has been fixed to 1000, whatever the sample size. In practice, the largest the sample size n, the largest this number of replications should be.

3.2 Goodness-of-fit tests

Goodness-of-fit tests with the characteristic function have been studied in the litterature to test if a sample has been generated from a fixed probability measure, for instance on the sphere in [15]. In this section, we consider the problem of testing that some mm space $\Xi_{\mathscr{X}}$ is isomorphic to some fixed mm space $\Xi_{\mathscr{X}_0}$, given an n-sample from $\Xi_{\mathscr{X}}$. The hypotheses are therefore: $\mathcal{H}_0$:"$\Xi_{\mathscr{X}}$ is isomorphic to $\Xi_{\mathscr{X}_0}$" and $\mathcal{H}_1$:"$\Xi_{\mathscr{X}}$ is not isomorphic to $\Xi_{\mathscr{X}_0}$". As for two-sample tests, if the test statistic properly renormalised converges to a fixed distribution as the sample size goes to ∞, then, the limit distribution depends on the mm space $\Xi_{\mathscr{X}_0}$. Therefore, we use a Monte Carlo procedure to approximate the distribution of the test statistic under $\mathcal{H}_0$. For this, we use B n-samples generated independently from $\Xi_{\mathscr{X}_0}$. The procedure is as follows:

1. Compute a discrete approximation of the mm space $\Xi_{\mathscr{X}_0}$, based on N points, N large, $\mathscr{X}_{0,N}$. (For instance, a regular grid on the circle if $\Xi_{\mathscr{X}_0}$ is the circle equipped with the uniform probability measure)
2. Calculate $T_{n,obs} = \mathrm{d}(\mathscr{X}_n, \mathscr{X}_{0,N})$ from the original sample $\mathscr{X}_n$ and $\mathscr{X}_{0,N}$.
3. Generate B samples $(\mathscr{X}_n^b)_{1\leq b\leq B}$ from $\Xi_{\mathscr{X}_0}$.
4. Calculate $T_n^b = \mathrm{d}(\mathscr{X}_n^b, \mathscr{X}_{0,N})$ for each $b \in \{1, \ldots, B\}$.
5. Approximate the p-value of the test by $\hat{p} = \frac{\mathrm{Card}\{b\in\{1,\ldots,B\},\, T_n^b \geq T_{n,obs}\}}{B}$.

The power of the test with level $\alpha \in \{0.05, 0.1\}$ is estimated by $\hat{\mathrm{P}}_\alpha = \frac{\mathrm{Card}\{m\in\{1,\ldots,M\},\, \hat{p}_m \leq \alpha\}}{M}$, after $M = 1000$ independent replications of the experiment.

The problem of testing uniformity on the circle or on the sphere has been widely considered [26, 4, 17, 8, 9]. In this paper, we use our goodness-of-fit test to test that a sample has been generated from the uniform distribution on the circle. We compare its performances to the classical Rayleigh and Bingham tests. This test is well suited for testing uniformity on the circle or on the sphere. Indeed, given that two measures equal up to a rotation on the sphere are isomorphic, the uniform distribution is the only distribution on the sphere, isomorphic to no other distribution. We compare performances of the procedure with the pseudo-metric $\mathrm{d}^{\mathrm{boot},R}$ (with characteristic function evaluation at $t \in \{0.1, 0.2, 0.5, 1.0, 5.0, 10.0, 100.0\}$ for $R = 2$, and at $t \in \{(0.5) \in \mathbb{R}^{\frac{2(2-1)}{2}}, (0.5, 0.5, 0.5) \in \mathbb{R}^{\frac{3(3-1)}{2}}, \ldots, (0.5, \ldots, 0.5) \in \mathbb{R}^{\frac{5(5-1)}{2}}\}$ for $R = 5$) against classical Rayleigh and Bingham tests of uniformity, and against the new test of [8], $\mathrm{d}^{\mathrm{dtm}}$, based on the DTM signature (we used the iidness version, with a selection parameter procedure, with varying parameter h in $[0.1, 0.2, 0.5, 1]$). The performance results are available in Table 4, 5 and 6. The alternatives we consider are: measures on l_p-balls and von Mises distributions, as in Section 3.1, but also mixtures of 4 von Mises distributions with the same concentration parameter κ, with centers given by the 4 vertices of a square, and with the same mass on each of the 4 components of the mixture.

If the Rayleigh and the Bingham tests have no power for mixtures alternatives on the circle [9], as also noticed in Table 6, our tests are performant, although not as much as the most powerful existing test $\mathrm{d}^{\mathrm{dtm}}$, [8], up to our knowledge. Our tests are almost as performant as the Rayleigh test that is optimal for von Mises alternatives. Moreover, unlike our new test, none of the Rayleigh, Bingham and DTM-signature-based test [8] show power under l_p-ball alternatives. This makes sense since the directions of the sample points are uniformly distributed on the circle for these alternatives. It should be noticed that the performance of our test depends on the choice of parameters as well as on the alternatives.

p	$\mathrm{d}^{\mathrm{boot},2}$	$\mathrm{d}^{\mathrm{boot},5}$	Rayleigh	Bingham	$\mathrm{d}^{\mathrm{dtm}}$
1.8	1.0	1.0	0.046	0.041	0.047
1.85	0.728	0.94	0.045	0.053	0.043
1.9	0.396	0.929	0.057	0.047	0.07
1.95	0.327	0.146	0.048	0.047	0.052
2.0	0.05	0.037	0.042	0.053	0.043

Table 4 Uniformity test, l_p-ball, sample size $n = 100$, $\alpha = 0.05$.

κ	$\mathrm{d}^{\mathrm{boot},2}$	$\mathrm{d}^{\mathrm{boot},5}$	Rayleigh	Bingham	$\mathrm{d}^{\mathrm{dtm}}$
0	0.046	0.054	0.049	0.041	0.047
0.1	0.067	0.073	0.083	0.049	0.072
0.2	0.069	0.179	0.216	0.051	0.190
0.5	0.600	0.788	0.892	0.078	0.840
1	0.999	1.0	1.0	0.247	1.0
2	1.0	1.0	1.0	0.983	1.0
5	1.0	1.0	1.0	1.0	1.0

Table 5 Uniformity test, von Mises distribution, sample size $n = 100$, $\alpha = 0.05$.

p	$\mathrm{d}^{\mathrm{boot},2}$	$\mathrm{d}^{\mathrm{boot},5}$	Rayleigh	Bingham	$\mathrm{d}^{\mathrm{dtm}}$
0	0.071	0.064	0.068	0.058	0.078
10	0.475	0.068	0.057	0.048	0.95
20	0.999	0.083	0.059	0.064	1.0
50	0.990	0.119	0.040	0.071	1.0
300	1.0	0.224	0.05	0.082	1.0

Table 6 Uniformity test, mixture of 4 von Mises distributions, sample size $n = 100$, $\alpha = 0.05$.

4 Learning for metric measure spaces

Let $(\Xi_{\mathscr{X}^{\ell}_{n_\ell}}, Y_\ell)_{\ell\in L}$ be a training dataset of Card(L) labelled mm spaces, with Y_ℓ a random variable that is $\{0,1\}$-valued (or $\mathfrak{D}$-valued for some finite set $\mathfrak{D}$) for the problem of classification. In this section, we propose new learning methods. Roughly, we develop a new classifier f that assigns to any given new mm space $\Xi_{\mathscr{X}_n}$, a label $f(\Xi_{\mathscr{X}_n})$. This label is supposed to approximate well the unknown variable Y associated to $\Xi_{\mathscr{X}_n}$. The performance of any function f is measured in terms of missclassification error rate $\frac{1}{\mathrm{Card}T}\sum_{t\in T}\mathbb{1}_{Y_t\neq f(\Xi_{\mathscr{X}^t_{n_t}})}$, on a testing dataset $(\Xi_{\mathscr{X}^t_{n_t}}, Y_t)_{t\in T}$. This problem has been tackled in [22], where a k-nearest-neighbours classifier is proposed. Roughly, to a new mm space $\Xi_{\mathscr{X}_n}$, we associate the label that appears the most often among the k nearest neighbours of $\Xi_{\mathscr{X}_n}$ in the training set. In [22], the Gromov-Wasserstein distance is used to compute nearest neighbours. In this paper, we investigate the use of the statistic defined in (6) combined with classical learning methods (convolutional neural networks) for the problem of digits recognition and shape detection.

4.1 Digits classification

In this section, we complete classification task on the famous MNIST digits dataset. We compare our method to a classical convolutional neural network (CNN) on the 8×8-sized digits images, from the `load_digits` function of the `sklearn.datasets` Python library. Each image has been turned into a dataset by assigning an integer weight to the pixel centers, proportional to the gray level of the pixel. The network structures used are as follows:

- CNN images: `Conv((3,3),(1,6),relu,MaxPool(2,2)); Flatten; Dense(54,20,relu); Dense(20,10); softmax`,
- NN images: `Flatten; Dense(64,128,relu); Dense(128,10); softmax`,
- Characteristic function: `Dense(4,15,relu); Dense(15,10,sigmoid); softmax`.

For the error computations, we made 50 replications of the experiment with 80% of data for learning and 20% for testing. We made `epoch`=20 complete pass through the training dataset for learning. For the characteristic functions based method, we used parameters $\mathfrak{R} \times \Theta = \{(0.5), (0.5, 0.5), (0.5, 0.5, 0.5), (0.5, 0.5, 0.5, 0.5), (0.5, 0.5, 0.5, 0.5, 0.5)\}$. As noticed in Figure 1, our method is a serious competitor to classical CNN. The median error represented by a line is in favor of our method. Moreover, the computational time is much smaller.

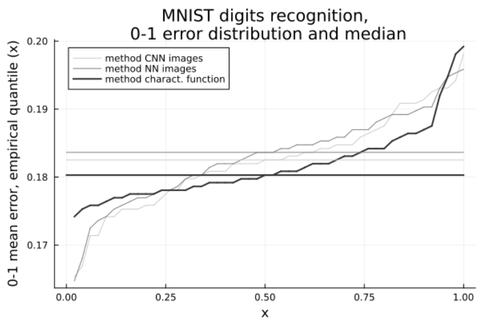

Figure 1 Misclassification rate for MNIST dataset classification.

4.2 Application to shape detection

In this section, we focus on the problem of classification, where a shape is labeled either as a segment or as a circle. Our method has been trained on a dataset of 11 circle-labeled 50-samples (from l_p-balls as in Section 3.1, with p in $\{1, 1.1, \ldots, 2\}$), and 11 segment-labeled 50-samples (uniform from the segment $[0, 1]$).

We consider a dataset represented in Figure 2. It has been generated from 5 l_p-balls, with the following parameters `(sample size, p, radius, center)`: $(3000, 1.7, 4.5, [0, 0])$, $(1500, 2, 2.5, [4, 0])$, $(500, 1.2, 1, [-2, 2])$, $(1000, 1.8, 2, [1, -1])$, $(500, 1.3, 0.8, [-3, -3])$. For parameters p close to 1, the l_p-balls look like squares and can be considered as a union of 4 segments, whereas for parameters p close to 2, the l_p-balls look like regular circles. We aim at extracting points from l_p-balls with small parameter p (in purple in Figure 2) and points from l_p-balls with large parameter p (in yellow), from the knowledge of the positions of the 6500 points only. For this, we proceed as follows:

- We turn each datapoint in $\mathbb{R}^2$ into a datapoint in the Euclidean space $\mathbb{R}^6$ by associating to the position of the point, the coordinates of the projection matrix onto its normal vector. This projection matrix is approximated using the Python notebook from [10].
- We use the ToMATo clustering algorithm [13], from the Python Gudhi library [25], to recover 11 components, as guided by the 11 points far from the diagonal of the persistence diagram in Figure 2. We used the parameters `graph_type='knn'`, `density_type='logDTM'`, `k = 30`, `k_DTM = 50`.
- Finally, we apply our classification algorithm (the same as the one in previous section) to each of the 11 components, trained on the 22-sized dataset. The result is represented in Figure 2 (right).

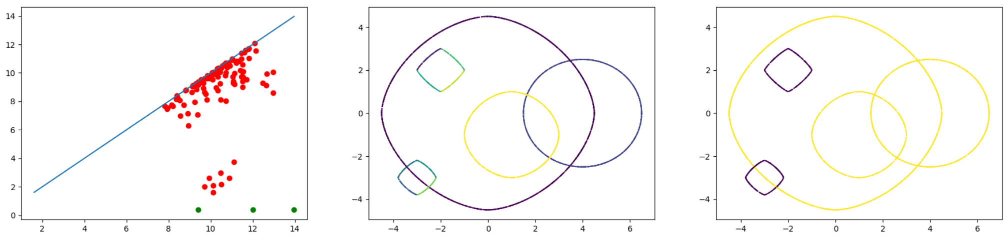

Figure 2 Persistence diagram for ToMATo algorithm (left), resulting clustering in 11 components (middle), (right) components are colored in purple if considered as segments by our learning algorithm and in yellow if considered as circles.

The result looks very satisfactory. We recover circle-shaped l_p-balls in yellow and squared-shaped l_p-balls in purple.

5 Conclusion and Perspectives

In this paper, we have introduced and studied a new metric to compare mm spaces up to an isomorphism. Unlike the Gromov-Wasserstein distance, this new metric does not suffer from the curse of dimensionality, in the sense that an mm space can be approximated at a parametric rate $n^{-\frac{1}{2}}$ from an n-sample. Moreover, we have proposed fast computationally tractable alternatives to this new metric. We have used this new metric and these alternatives to derive two-sample and goodness-of-fit tests to compare mm spaces from samples and in particular a new test of uniformity on the sphere. We have also developped a new learning procedure for learning on mm spaces. If this new method consists in assigning to any mm space, a value in the Euclidean space, and then to apply learning methods on the Euclidean space, as for instance neural networks, it should be interesting to develop a new neural networks structure specific to data like pairwise distance matrices. Such structures should be able to learn on matrices with potentially different size, as the new procedure introduced in this paper.

References

1 V. Alba Fernández, M. D. Jiménez Gamero, and J. Muñoz García. A test for the two-sample problem based on empirical characteristic functions. *Comput. Statist. Data Anal.*, 52(7):3730–3748, 2008. doi:10.1016/j.csda.2007.12.013.

2 Shreya Arya, Arnab Auddy, Ranthony A Clark, Sunhyuk Lim, Facundo Memoli, and Daniel Packer. The gromov–wasserstein distance between spheres. *Foundations of Computational Mathematics*, pages 1–56, 2024.

3 Martin Bauer, Facundo Mémoli, Tom Needham, and Mao Nishino. The z-gromov-wasserstein distance. *arXiv preprint arXiv:2408.08233*, 2024.

4 Christopher Bingham. An antipodally symmetric distribution on the sphere. *Ann. Statist.*, 2:1201–1225, 1974. URL: http://links.jstor.org/sici?sici=0090-5364(197411)2:6<1201:AASDOT>2.0.CO;2-A&origin=MSN.

5 Sergey Bobkov and Michel Ledoux. One-dimensional empirical measures, order statistics, and Kantorovich transport distances. *Mem. Amer. Math. Soc.*, 261(1259):v+126, 2019. doi:10.1090/memo/1259.

6 Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities.* Oxford University Press, Oxford, 2013. A nonasymptotic theory of independence, With a foreword by Michel Ledoux. doi:10.1093/acprof:oso/9780199535255.001.0001.

7 Claire Brécheteau. A statistical test of isomorphism between metric-measure spaces using the distance-to-a-measure signature. *Electron. J. Stat.*, 13(1):795–849, 2019. doi:10.1214/19-ejs1539.

8 Claire Brécheteau. Two distance-based families of statistical tests of uniformity for probability measures on homogeneous compact polish spaces. 2025. submitted.

9 Claire Brécheteau, Davy Paindaveine, and Thomas Verdebout. On the detection of hierarchical signals on hyperspheres. 2025. submitted.

10 Blanche Buet, Gian Paolo Leonardi, and Simon Masnou. Weak and approximate curvatures of a measure: a varifold perspective. *Nonlinear Anal.*, 222:Paper No. 112983, 34, 2022. doi:10.1016/j.na.2022.112983.

11 Frédéric Chazal, David Cohen-Steiner, Leonidas J Guibas, Facundo Mémoli, and Steve Y Oudot. Gromov-hausdorff stable signatures for shapes using persistence. In *Computer Graphics Forum*, volume 28, pages 1393–1403. Wiley Online Library, 2009.

12 Frédéric Chazal, Vin de Silva, and Steve Oudot. Persistence stability for geometric complexes. *Geom. Dedicata*, 173:193–214, 2014. doi:10.1007/s10711-013-9937-z.

13 Frédéric Chazal, Leonidas J. Guibas, Steve Y. Oudot, and Primoz Skraba. Persistence-based clustering in riemannian manifolds. *J. ACM*, 60(6):41:1–41:38, November 2013. URL: http://doi.acm.org/10.1145/2535927, doi:10.1145/2535927.

14 Théo Dumont, Théo Lacombe, and François-Xavier Vialard. On the existence of monge maps for the gromov–wasserstein problem. *Foundations of Computational Mathematics*, 25(2):463–510, 2025.

15 Bruno Ebner, Norbert Henze, and Simos Meintanis. A unified approach to goodness-of-fit testing for spherical and hyperspherical data. *Statist. Papers*, 65(6):3447–3475, 2024. doi:10.1007/s00362-024-01529-1.

16 Nicolas Fournier and Arnaud Guillin. On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Related Fields*, 162(3-4):707–738, 2015. doi:10.1007/s00440-014-0583-7.

17 Eduardo García-Portugués and Thomas Verdebout. An overview of uniformity tests on the hypersphere. 04 2018.

18 Andreas Greven, Peter Pfaffelhuber, and Anita Winter. Convergence in distribution of random metric measure spaces (λ-coalescent measure trees). *Probability Theory and Related Fields*, 145(1):285–322, 2009.

19 Misha Gromov. *Metric structures for Riemannian and non-Riemannian spaces.* Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, english edition, 2007. Based on the 1981 French original, With appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael Bates.

20 Wassily Hoeffding. A class of statistics with asymptotically normal distribution. *Ann. Math. Statistics*, 19:293–325, 1948. doi:10.1214/aoms/1177730196.

21 Marie Hušková and Simos G. Meintanis. Tests for the multivariate k-sample problem based on the empirical characteristic function. *J. Nonparametr. Stat.*, 20(3):263–277, 2008. doi:10.1080/10485250801948294.
22 Facundo Mémoli. Gromov–Wasserstein distances and the metric approach to object matching. *Foundations of Computational Mathematics*, 11(4):417–487, 2011.
23 Robert Osada, Thomas Funkhouser, Bernard Chazelle, and David Dobkin. Shape distributions. *ACM Transactions on Graphics (TOG)*, 21(4):807–832, 2002.
24 Gabriel Peyré, Marco Cuturi, and Justin Solomon. Gromov-wasserstein averaging of kernel and distance matrices. In *International conference on machine learning*, pages 2664–2672. PMLR, 2016.
25 The GUDHI Project. *GUDHI User and Reference Manual*. GUDHI Editorial Board, 3.11.0 edition, 2025. URL: https://gudhi.inria.fr/doc/3.11.0/.
26 Lord Rayleigh. Xxxi. on the problem of random vibrations, and of random flights in one, two, or three dimensions. *Philosophical Magazine Series 1*, 37:321–347, 1919. URL: https://api.semanticscholar.org/CorpusID:123909479.
27 Robert J. Serfling. *Approximation theorems of mathematical statistics*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1980.
28 Karl-Theodor Sturm. On the geometry of metric measure spaces. 2006.
29 Thomas Verdebout. On some validity-robust tests for the homogeneity of concentrations on spheres. *J. Nonparametr. Stat.*, 27(3):372–383, 2015. doi:10.1080/10485252.2015.1041945.
30 Thomas Verdebout. On the efficiency of some rank-based test for the homogeneity of concentrations. *J. Statist. Plann. Inference*, 191:101–109, 2017. doi:10.1016/j.jspi.2017.05.009.
31 Cédric Villani et al. *Optimal transport: old and new*, volume 338. Springer, 2008.
32 Ulrike von Luxburg. A tutorial on spectral clustering. *Stat. Comput.*, 17(4):395–416, 2007. doi:10.1007/s11222-007-9033-z.
33 Jonathan Weed and Francis Bach. Sharp asymptotic and finite-sample rates of convergence of empirical measures in Wasserstein distance. *Bernoulli*, 25(4A):2620–2648, 2019. doi:10.3150/18-BEJ1065.

A Proofs

A.1 Proof of Gromov's mm spaces reconstruction theorem

In this section, we provide lemmas together with their proofs that we use to prove the Gromov's mm spaces reconstruction theorem, Theorem 1.

A.1.1 Lemma 8 and its proof

▶ **Lemma 8** (Isometries extension theorem). *Let $(\mathscr{X}, \mathrm{d}_{\mathscr{X}})$ and $(\mathscr{Y}, \mathrm{d}_{\mathscr{Y}})$ be two metric spaces, with $(\mathscr{Y}, \mathrm{d}_{\mathscr{Y}})$ complete. Let A be a dense subset of $\mathscr{X}$ and $f : A \to \mathscr{Y}$ an isometry. Then, there exists a unique isometry $g : \mathscr{X} \to \mathscr{Y}$ that coincides with f on A.*

Proof. For every sequence $(x_n)_{n\in\mathbb{N}}$ in A converging to $x \in \mathscr{X}$, $(f(x_n))_{n\in\mathbb{N}}$ is Cauchy in the complete space $\mathscr{Y}$ since f is an isometry, thus $(f(x_n))_{n\in\mathbb{N}}$ converges to some limit $g(x) \in \mathscr{Y}$. Note that this limit does not depend on the sequence $(x_n)_{n\in\mathbb{N}}$. The map g built in this way is an isometry. ◀

A.1.2 Lemma 9 and its proof

▶ **Lemma 9.** *If $\mathrm{P}_{\mathscr{X},r} = \mathrm{P}_{\mathscr{Y},r}$ for every $r \in \mathbb{N}$, then, for every $n \in \mathbb{N}^*$:*

$$M_{\mathscr{X},n_\sharp}(\mathrm{P}_{\mathscr{X}}^{\otimes n}) = M_{\mathscr{Y},n_\sharp}(\mathrm{P}_{\mathscr{Y}}^{\otimes n}),$$

with $M_{\mathscr{X},n}$ *and* $M_{\mathscr{Y},n}$ *defined in* (1).

Proof. The image of the map $M_{\mathscr{X},n}$ is contained in $\mathbb{R}_+^d$ with $d = \frac{5n^2-n}{2}$. According to the monotone class lemma, proving equality of the probability measures $M_{\mathscr{X},n_\sharp}(\mathrm{P}_{\mathscr{X}}^{\otimes n})$ and $M_{\mathscr{Y},n_\sharp}(\mathrm{P}_{\mathscr{Y}}^{\otimes n})$ boils down to prove that they coincide on the π-system $\{]a_1,b_1[\times\ldots]a_d,b_d[\mid (a_i,b_i)\in\mathbb{Q}^2, a_i<b_i\}$. Let $d'=2n^2$. For every $i\in\{1,\ldots,d'\}$, we approximate $\mathbb{1}_{]a_i,b_i[}$ with a bounded continuous function $\alpha_{\ell,i}$ equal to 1 on $]a_i-\frac{1}{\ell},b_i+\frac{1}{\ell}[$ and 0 on $]a_i,b_i[^c$ for ℓ large enough. So, $\lim_{\ell\to\infty}\int_{[0,1]^{d'}\times\mathbb{R}_+^{d-d'}}\prod_{i=1}^{d'}\alpha_{\ell,i}(t_i)\prod_{i=d'+1}^{d}\mathbb{1}_{]a_i,b_i[}(t_i)\mathrm{d}M_{\mathscr{X},n_\sharp}(\mathrm{P}_{\mathscr{X}}^{\otimes n})(t_1,\ldots,t_d)$ equals $M_{\mathscr{X},n_\sharp}(\mathrm{P}_{\mathscr{X}}^{\otimes n})(]a_1,b_1[\times\ldots\times]a_d,b_d[)$. It remains to prove equality of the integrals associated to both measures, for every $\alpha_{\ell,i}$s. The Stone-Weierstrass theorem entails that the integrals $\int_{[0,1]^{d'}\times\mathbb{R}_+^{d-d'}}\alpha(t_1,\ldots,t_{d'})\prod_{i=d'+1}^{d}\mathbb{1}_{]a_i,b_i[}(t_i)\mathrm{d}M_{\mathscr{X},n_\sharp}(\mathrm{P}_{\mathscr{X}}^{\otimes n})(t_1,\ldots,t_d)$ are equal for the two measures $M_{\mathscr{X},n_\sharp}(\mathrm{P}_{\mathscr{X}}^{\otimes n})$ and $M_{\mathscr{Y},n_\sharp}(\mathrm{P}_{\mathscr{Y}}^{\otimes n})$, for every bounded continuous function α defined on $[0,1]^{d'}$ since they coincide for every polynomial function α, as noted in [19, Theorem 3 1/2.5.]. This last results follows from the fact that $\mathrm{P}_{\mathscr{X},r}$ and $\mathrm{P}_{\mathscr{Y},r}$ coincide for every $r\in\mathbb{N}^*$ and from Fubini-Tonelli theorem. For instance, for $r=2$, we get:

$$\int_{\mathscr{X}}\mathrm{P}_{\mathscr{X}}(\mathrm{B}(x,\rho))\mathrm{d}\mathrm{P}_{\mathscr{X}}(x)=\int_{\mathscr{X}\times\mathscr{X}}\mathbb{1}_{\mathrm{d}_{\mathscr{X}}(x,x')\leq q}\mathrm{d}\mathrm{P}_{\mathscr{X}}^{\otimes 2}(x,x')$$
$$=\int_{\mathscr{Y}\times\mathscr{Y}}\mathbb{1}_{\mathrm{d}_{\mathscr{Y}}(y,y')\leq q}\mathrm{d}\mathrm{P}_{\mathscr{Y}}^{\otimes 2}(y,y')=\int_{\mathscr{Y}}\mathrm{P}_{\mathscr{Y}}(\mathrm{B}(y,\rho))\mathrm{d}\mathrm{P}_{\mathscr{Y}}(y).$$

◀

A.1.3 Lemma 10 and its proof

▶ **Lemma 10.** *Let* $\mathbf{x}^*$ *be as defined in Section 2.1.2. Then, we can build a sequence* $(\mathbf{y}^n)_{n\in\mathbb{N}^*}$ *of sequences* $\mathbf{y}^n=(y_1^n,y_2^n,\ldots,y_n^n,a,a,\ldots)$ *in* $\mathscr{Y}^{\mathbb{N}}$ *with a fixed point* $a\in\mathscr{Y}$*, such that for every* $n\in\mathbb{N}^*$, $\|M_{\mathscr{X},n}(\mathbf{x}_n^*)-M_{\mathscr{Y},n}(\mathbf{y}_n^n)\|_\infty\leq\frac{1}{n}$*. That is,*

- $\forall\, i,j\in\{1,\ldots,n\}$, $|\mathrm{P}_{\mathscr{X}}(\mathrm{B}(x_i^*,\phi_j(i)))-\mathrm{P}_{\mathscr{Y}}(\mathrm{B}(y_i^n,\phi_j(i)))|\leq\frac{1}{n}$,
- $\forall\, j,\ell\in\{1,\ldots,n\}$, $|\mathrm{P}_{\mathscr{X}}(\bigcap_{i=1}^{\ell}\mathrm{B}(x_i^*,\phi_j(i)))-\mathrm{P}_{\mathscr{Y}}(\bigcap_{i=1}^{\ell}\mathrm{B}(y_i^n,\phi_j(i)))|\leq\frac{1}{n}$,
- $\forall\, i,j\in\{1,\ldots,n\}$, $|\mathrm{d}_{\mathscr{X}}(x_i^*,x_j^*)-\mathrm{d}_{\mathscr{Y}}(y_i^n,y_j^n)|\leq\frac{1}{n}$.

Proof. Since $\mathrm{Supp}(\mathrm{P}_{\mathscr{X}})=\mathscr{X}$, $\mathbf{x}_n^*=(x_1^*,x_2^*,\ldots,x_n^*)\in\mathrm{Supp}\left(\mathrm{P}_{\mathscr{X}}^{\otimes n}\right)$. It yields that for every $\eta>0$, $\mathrm{P}_{\mathscr{X}}^{\otimes n}(\mathrm{B}(\mathbf{x}_n^*,\eta))>0$. Moreover, according to Lemma 11, $M_{\mathscr{X},n}$ is continuous on $\mathscr{X}^n$. Thus, for every $\epsilon>0$, there exists $\eta>0$ such that $\mathrm{B}(\mathbf{x}_n^*,\eta)\subset\{\mathbf{z}_n=(z_1,z_2,\ldots,z_n)\in\mathscr{X}^n\mid\|M_{\mathscr{X},n}(\mathbf{x}_n^*)-M_{\mathscr{X},n}(\mathbf{z}_n)\|_\infty\leq\epsilon\}$. It holds that $M_{\mathscr{X},n}(\mathbf{x}_n^*)\in\mathrm{Supp}\left(M_{\mathscr{X},n_\sharp}(\mathrm{P}_{\mathscr{X}}^{\otimes n})\right)$. According to Lemma 9, the distributions $M_{\mathscr{X},n_\sharp}(\mathrm{P}_{\mathscr{X}}^{\otimes n})$ and $M_{\mathscr{Y},n_\sharp}(\mathrm{P}_{\mathscr{Y}}^{\otimes n})$ coincide. Thus, $M_{\mathscr{X},n}(\mathbf{x}_n^*)\in\mathrm{Supp}\left(M_{\mathscr{Y},n_\sharp}(\mathrm{P}_{\mathscr{Y}}^{\otimes n})\right)$, and $\forall\epsilon>0$, $M_{\mathscr{Y},n_\sharp}(\mathrm{P}_{\mathscr{Y}}^{\otimes n})(\mathrm{B}(M_{\mathscr{X},n}(\mathbf{x}_n^*),\epsilon))>0$. It yields that for every $n\in\mathbb{N}^*$, there exists $\mathbf{y}_n^n=(y_1^n,y_2^n,\ldots,y_n^n)\in\mathscr{Y}^n=\mathrm{Supp}\left(\mathrm{P}_{\mathscr{Y}}^{\otimes n}\right)$ that satisfies $\|M_{\mathscr{Y},n}(\mathbf{y}_n^n)-M_{\mathscr{X},n}(\mathbf{x}_n^*)\|_\infty\leq\frac{1}{n}$.

◀

A.1.4 Lemma 11 and its proof

▶ **Lemma 11.** *For every* $n\in\mathbb{N}^*$*, the map* $M_{\mathscr{X},n}:\mathscr{X}^n\to\mathbb{R}_+^{\frac{5n^2-n}{2}}$ *defined in* (1) *is continuous.*

Proof. We equip the space $\mathscr{X}^n$ with the metric $\mathrm{d}_{\mathscr{X}}^n$ defined for every $\mathbf{x}_n=(x_1,x_2,\ldots,x_n)$ and $\mathbf{z}_n=(z_1,z_2,\ldots,z_n)$ by $\mathrm{d}_{\mathscr{X}}^n(\mathbf{x}_n,\mathbf{z}_n)=\max_{1\leq i\leq n}(\mathrm{d}_{\mathscr{X}}(x_i,z_i))$. Consider a converging sequence $(\mathbf{z}_n^k)_{k\in\mathbb{N}}=((z_1^k,z_2^k,\ldots,z_n^k))_{k\in\mathbb{N}}$ in $\mathscr{X}^n$, with limit $\mathbf{x}_n=(x_1,x_2,\ldots,x_n)$. Then

$|\mathrm{d}_{\mathscr{X}}(z_i^k, z_j^k) - \mathrm{d}_{\mathscr{X}}(x_i, x_j)| \leq 2\mathrm{d}^n_{\mathscr{X}}(\mathbf{x}_n, \mathbf{z}^k_n)$ converges to zero when k goes to infinity. Moreover, $\mathrm{P}_{\mathscr{X}}(\mathrm{B}(z_i^k, \phi_j(i))) \leq \mathrm{P}_{\mathscr{X}}(\mathrm{B}(x_i, \phi_j(i) + \mathrm{d}^n_{\mathscr{X}}(\mathbf{x}_n, \mathbf{z}^k_n)))$. Thus, $\limsup_{k\to\infty} \mathrm{P}_{\mathscr{X}}(\mathrm{B}(z_i^k, \phi_j(i))) \leq \mathrm{P}_{\mathscr{X}}(\overline{\mathrm{B}}(x_i, \phi_j(i)))$. As well, $\liminf_{k\to\infty} \mathrm{P}_{\mathscr{X}}(\mathrm{B}(z_i^k, \phi_j(i))) \geq \mathrm{P}_{\mathscr{X}}(\mathrm{B}(x_i, \phi_j(i)))$. Consequently, by definition of ϕ_j, $\mathrm{P}_{\mathscr{X}}(\mathrm{B}(x_i, \phi_j(i))) = \mathrm{P}_{\mathscr{X}}(\overline{\mathrm{B}}(x_i, \phi_j(i))) = \lim_{k\to\infty} \mathrm{P}_{\mathscr{X}}(\mathrm{B}(z_i^k, \phi_j(i)))$. The same result holds for the intersections of balls. Thus, the map $M_{\mathscr{X},n}$ is continuous. ◀

A.1.5 Lemma 12 and its proof

▶ **Lemma 12.** *Let $(\mathbf{y}^n)_{n\in\mathbb{N}^*}$ be a $\mathscr{Y}^{\mathbb{N}}$-valued sequence defined as in Lemma 10. Then, we can build a subsequence of $(\mathbf{y}^n)_{n\in\mathbb{N}^*}$ such that for every $i \in \mathbb{N}^*$, the ith coordinate of $\mathbf{y}^n$, y_i^n, converges to a point y_i^* in $\mathscr{Y}$.*

Proof. According to Lemma 10, $(\mathbf{y}^n)_{n\in\mathbb{N}^*}$ is built such that for every $i, j \in \{1, \dots, n\}$, $|\mathrm{P}_{\mathscr{X}}(\mathrm{B}(y_i^n, \phi_j(i))) - \mathrm{P}_{\mathscr{X}}(\mathrm{B}(x_i^*, \phi_j(i)))| \leq \frac{1}{n}$. For every $\epsilon > 0$, there exists $N \in \mathbb{N}^*$ such that $0 < \phi_N(i) < \epsilon$. Thus, for every $n \geq N$, $\mathrm{P}_{\mathscr{X}}(\mathrm{B}(y_i^n, \epsilon)) \geq \mathrm{P}_{\mathscr{X}}(\mathrm{B}(x_i^*, \phi_N(i))) - \frac{1}{n}$. And for n large enough, $\mathrm{P}_{\mathscr{X}}(\mathrm{B}(y_i^n, \epsilon)) \geq \frac{1}{2}\mathrm{P}_{\mathscr{X}}(\mathrm{B}(x_i^*, \phi_N(i)))$ which is positive since $x_i^* \in \mathrm{Supp}(\mathrm{P}_{\mathscr{X}})$.

Thus, from any subsequence of $(y_i^n)_{n\in\mathbb{N}^*}$ we extract a converging subsequence, according to Lemma 13. In order to make the sequences $(y_i^n)_{n\in\mathbb{N}^*}$ for $i \in \mathbb{N}^*$ converge simultaneously, we apply a diagonal process.

◀

A.1.6 Lemma 13 and its proof

▶ **Lemma 13.** *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in a complete metric space $(\mathscr{X}, \mathrm{d}_{\mathscr{X}})$ equipped with a Borel probability measure $\mathrm{P}_{\mathscr{X}}$, such that for every $\epsilon > 0$ there exists $c > 0$ and $N > 0$ such that for every $n \geq N$, $\mathrm{P}_{\mathscr{X}}(\mathrm{B}(x_n, \epsilon)) > c$. Then, a converging subsequence can be extracted from $(x_n)_{n\in\mathbb{N}}$.*

Proof. For $\epsilon > 0$, let c and N be defined such that for every $n \geq N$, $\mathrm{P}_{\mathscr{X}}(\mathrm{B}(x_n, \epsilon)) > c$. Let $L(\epsilon) = \sup\{m \in \mathbb{N}^* \mid \exists n_1, n_2, \dots n_m \geq N, \forall\, i \neq j \in \{1, \dots, m\}, \mathrm{B}(x_{n_i}, \epsilon) \cap \mathrm{B}(x_{n_j}, \epsilon) = \emptyset\}$. Since $L(\epsilon) \leq \frac{1}{c} < \infty$, let $(x_{n^*_\ell})_{1\leq \ell \leq L(\epsilon)}$ be $L(\epsilon)$ points satisying the previous condition. Then, by construction, $\mathscr{X} = \bigcup_{\ell=1}^{L(\epsilon)} \mathrm{B}(x_{n^*_\ell}, 2\epsilon)$. Thus, there is a ball of radius 2ϵ containing an infinite subsequence of $(x_n)_{n\in\mathbb{N}^*}$. We denote any such infinite subsequence by $\mathcal{S}\left((x_n)_{n\in\mathbb{N}}, \epsilon\right)$. To build a converging subsequence $(z_n)_{n\in\mathbb{N}^*}$, we proceed as follows. First set $z_1 = x_1$ and keep $S_1 = \mathcal{S}\left((x_n)_{n\in\mathbb{N}}, 1\right) \setminus \{z_1\}$. For $n \in \mathbb{N}^*$, select for z_{n+1} the first element of $S_{n+1} = \mathcal{S}\left(S_n, \frac{1}{n+1}\right) \setminus \{z_n\}$. Then $(z_n)_{n\in\mathbb{N}^*}$ is a subsequence of $(x_n)_{n\in\mathbb{N}^*}$ which is Cauchy in a complete metric space, thus converging. ◀

A.1.7 Lemma 14 and its proof

▶ **Lemma 14.** *Let $\mathbf{y}^* = (y_i^*)_{i\in\mathbb{N}^*}$ be a sequence defined as in Lemma 12. Then,*

1. $\forall\, j, \ell \in \mathbb{N}^*$, $\mathrm{P}_{\mathscr{Y}}\left(\bigcap_{i=1}^{\ell} \mathrm{B}(y_i^*, \phi_j(i))\right) = \mathrm{P}_{\mathscr{X}}\left(\bigcap_{i=1}^{\ell} \mathrm{B}(x_i^*, \phi_j(i))\right)$
2. $\forall\, i, j \in \mathbb{N}^*$, $\mathrm{d}_{\mathscr{X}}(x_i^*, x_j^*) = \mathrm{d}_{\mathscr{Y}}(y_i^*, y_j^*)$.

Proof. Point 2 is a direct consequence of the triangular inequality, Lemma 10 and the definition of $\mathbf{y}^*$ given by Lemma 12. We now focus on the proof of point 1. According to Lemma 10, with the notation therein, we get that for every $n \in \mathbb{N}^*$,

$$\forall\, j, \ell \in \{1, \dots, n\}, \left|\mathrm{P}_{\mathscr{X}}\left(\bigcap_{i=1}^{\ell} \mathrm{B}\left(x_i^*, \phi_j(i)\right)\right) - \mathrm{P}_{\mathscr{Y}}\left(\bigcap_{i=1}^{\ell} \mathrm{B}\left(y_i^n, \phi_j(i)\right)\right)\right| \leq \frac{1}{n}.$$

For $j,i\in\mathbb{N}^*$, we can define $\boldsymbol{\phi}^{j,i}=\left(\phi_{n_k^{j,i}}(i)\right)_{k\in\mathbb{N}^*}$, an increasing sequence converging to $\phi_j(i)$, which is a subsequence of $(\phi_n(i))_{n\in\mathbb{N}^*}$ defined in Section 2.1.2. Now, let $j,\ell\in\mathbb{N}^*$. Then, for every $k\in\mathbb{N}^*$, there exists some $m\in\mathbb{N}$ satisfying $\mathrm{d}_{\mathscr{Y}}(y_i^*,y_i^m)\leq\phi_j(i)-\phi_{n_k^{j,i}}(i)$ and $m\geq n_k^{j,i}$ for every $1\leq i\leq\ell$. Then, we get that $\bigcap_{i=1}^\ell\mathrm{B}\left(y_i^*,\phi_j(i)\right)\supset\bigcap_{i=1}^\ell\mathrm{B}\left(y_i^m,\phi_{n_k^{j,i}}(i)\right)$, and $\mathrm{P}_{\mathscr{Y}}\left(\bigcap_{i=1}^\ell\mathrm{B}\left(y_i^*,\phi_j(i)\right)\right)\geq\mathrm{P}_{\mathscr{Y}}\left(\bigcap_{i=1}^\ell\mathrm{B}\left(y_i^m,\phi_{n_k^{j,i}}(i)\right)\right)\geq\mathrm{P}_{\mathscr{X}}\left(\bigcap_{i=1}^\ell\mathrm{B}\left(x_i^*,\phi_{n_k^{j,i}}(i)\right)\right)-\frac{1}{m}$. By making m and then k go to infinity, we get that $\mathrm{P}_{\mathscr{X}}\left(\bigcap_{i=1}^\ell\mathrm{B}\left(x_i^*,\phi_j(i)\right)\right)\leq\mathrm{P}_{\mathscr{Y}}\left(\bigcap_{i=1}^\ell\mathrm{B}\left(y_i^*,\phi_j(i)\right)\right)$. Similarly, $\mathrm{P}_{\mathscr{X}}\left(\bigcap_{i=1}^\ell\overline{\mathrm{B}}\left(x_i^*,\phi_j(i)\right)\right)\geq\mathrm{P}_{\mathscr{Y}}\left(\bigcap_{i=1}^\ell\mathrm{B}\left(y_i^*,\phi_j(i)\right)\right)$. Since the functions ϕ_j are R-valued, $\mathrm{P}_{\mathscr{X}}\left(\bigcap_{i=1}^\ell\overline{\mathrm{B}}\left(x_i^*,\phi_j(i)\right)\right)=\mathrm{P}_{\mathscr{X}}\left(\bigcap_{i=1}^\ell\mathrm{B}\left(x_i^*,\phi_j(i)\right)\right)$ and the result follows. ◀

A.1.8 Lemma 15 and its proof

▶ **Lemma 15** (Equality of $\tilde{\psi}(\mathscr{X})$ and $\mathscr{Y}$)**.** *Let $\tilde{\psi}$ be the isometry defined on $\mathscr{X}$ that is the extension of the isometry that sends $\mathbf{x}^*$ to $\mathbf{y}^*$, as defined in Lemma 10 and Section 2.1.2. Then, $\tilde{\psi}(\mathscr{X})=\mathscr{Y}$.*

Proof. First note that $\tilde{\psi}(\mathscr{X})$ is a closed subset of the complete space $\mathscr{Y}$. For the sake of contradiction, assume that $\tilde{\psi}(\mathscr{X})\neq\mathscr{Y}$ and choose $y\in\mathscr{Y}\backslash\tilde{\psi}(\mathscr{X})$. Then, for some $\epsilon>0$, $\mathrm{B}(y,\epsilon)\subset\mathscr{Y}\backslash\tilde{\psi}(\mathscr{X})$. Since $y\in\mathrm{Supp}(\mathrm{P}_{\mathscr{Y}})=\mathscr{Y}$, $\mathrm{P}_{\mathscr{Y}}(\mathrm{B}(y,\frac{\epsilon}{2}))>0$. Let $\rho\in R$ (with R defined in Section 2.1.2) be such that $\rho<\frac{\epsilon}{2}$, then $\mathrm{P}_{\mathscr{Y}}(\mathscr{Y}\backslash\mathrm{B}(y,\frac{\epsilon}{2}))\geq\lim_{n\to\infty}\mathrm{P}_{\mathscr{Y}}(\bigcup_{i\leq n}\mathrm{B}(y_i^*,\rho))$. The Poincaré's formula and Lemma 14 (with j chosen such that for every $i\in\{1,\ldots,n\}$, $\phi_j(i)=\rho$) yield $\mathrm{P}_{\mathscr{X}}(\bigcup_{i\leq n}\mathrm{B}(x_i^*,\rho))=\mathrm{P}_{\mathscr{Y}}(\bigcup_{i\leq n}\mathrm{B}(y_i^*,\rho))$. By making n go to infinity, it holds that $1=\mathrm{P}_{\mathscr{X}}(\mathscr{X})=\mathrm{P}_{\mathscr{Y}}(\mathscr{Y}\backslash\mathrm{B}(y,\frac{\epsilon}{2}))$, which is a contradiction. ◀

A.2 Proof of results from Section 2.2

A.2.1 Proof of Proposition 3

First recall that the function $x\in\mathbb{R}\mapsto\exp(ix)\in\mathbb{C}$ is 1-Lipschitz: $\forall x,y\in\mathbb{R}$, $|\exp(ix)-\exp(iy)|\leq|x-y|$. As a consequence, for every $\rho\geq0$ and $\theta\in\mathscr{S}^{\frac{r(r-1)}{2}-1}$ for $r\geq2$, the difference $\Delta_r(\rho\theta)$ between the characteristic functions at $\rho\theta$ is bounded by:

$$\begin{aligned}\Delta_r(\rho\theta)&:=|\phi_{\mathscr{X},r}(\rho\theta)-\phi_{\mathscr{Y},r}(\rho\theta)|\\&=\left|\int_{\mathscr{X}^r}\exp(i\langle\rho\theta,D_{\mathscr{X},r}(\mathbf{x}_r)\rangle)\mathrm{dP}_{\mathscr{X}}^{\otimes r}(\mathbf{x}_r)-\int_{\mathscr{Y}^r}\exp(i\langle\rho\theta,D_{\mathscr{Y},r}(\mathbf{y}_r)\rangle)\mathrm{dP}_{\mathscr{Y}}^{\otimes r}(\mathbf{y}_r)\right|\\&\leq\rho\int_{\mathscr{X}^r\times\mathscr{Y}^r}\|D_{\mathscr{X},r}(\mathbf{x}_r)-D_{\mathscr{Y},r}(\mathbf{y}_r)\|\,\mathrm{d}\pi^{\otimes r}(\mathbf{x}_r,\mathbf{y}_r),\end{aligned}$$

according to the Cauchy-Schwarz inequality and since $\|\theta\|=1$, for any transport map $\pi\in\Pi(\mathrm{P}_{\mathscr{X}},\mathrm{P}_{\mathscr{Y}})$ (as defined in Section 2.2.2). Moreover, using the fact that for non negative real numbers $x_1,\ldots,x_D$, $\sqrt{\sum_{i=1}^D x_i^2}\leq\sqrt{D}\sum_{i=1}^D x_i$ with $D=\frac{r(r-1)}{2}$, we get that:

$$\Delta_r(\rho\theta)\leq\rho\left(\frac{r(r-1)}{2}\right)^{\frac{3}{2}}\int_{\mathscr{X}\times\mathscr{Y}}\int_{\mathscr{X}\times\mathscr{Y}}|\mathrm{d}_{\mathscr{X}}(x,x')-\mathrm{d}_{\mathscr{Y}}(y,y')|\,\mathrm{d}\pi(x,y)\mathrm{d}\pi(x',y')\tag{11}$$

$$\leq2\rho\left(\frac{r(r-1)}{2}\right)^{\frac{3}{2}}\left(\int_{\mathscr{X}}\mathrm{d}_{\mathscr{X}}(x,x_0)\mathrm{dP}_{\mathscr{X}}(x)+\int_{\mathscr{Y}}\mathrm{d}_{\mathscr{X}}(y,y_0)\mathrm{dP}_{\mathscr{Y}}(y)\right),\tag{12}$$

for some $x_0 \in \mathscr{X}$ and some $y_0 \in \mathscr{Y}$. Alternatively, we also get that:

$$\Delta_r(\rho\theta) \leq \rho\sqrt{\int_{\mathscr{X}^r\times\mathscr{Y}^r}\sum_{1\leq i<j\leq r}|\mathrm{d}_{\mathscr{X}}(x_i,x_j)-\mathrm{d}_{\mathscr{Y}}(y_i,y_j)|^2\,\mathrm{d}\pi^{\otimes r}(\mathbf{x}_r,\mathbf{y}_r)} \quad (13)$$

$$= \rho\sqrt{\sum_{1\leq i<j\leq r}\int_{\mathscr{X}\times\mathscr{Y}}\int_{\mathscr{X}\times\mathscr{Y}}|\mathrm{d}_{\mathscr{X}}(x_i,x_j)-\mathrm{d}_{\mathscr{Y}}(y_i,y_j)|^2\,\mathrm{d}\pi(x_i,y_i)\mathrm{d}\pi(x_j,y_j)} \quad (14)$$

$$= \rho\sqrt{\frac{r(r-1)}{2}}\sqrt{\int_{\mathscr{X}\times\mathscr{Y}}\int_{\mathscr{X}\times\mathscr{Y}}|\mathrm{d}_{\mathscr{X}}(x_1,x_2)-\mathrm{d}_{\mathscr{Y}}(y_1,y_2)|^2\,\mathrm{d}\pi(x_1,y_1)\mathrm{d}\pi(x_2,y_2)} \quad (15)$$

$$\leq 2\rho\sqrt{\frac{r(r-1)}{2}}\sqrt{\int_{\mathscr{X}}\mathrm{d}^2_{\mathscr{X}}(x,x_0)\mathrm{dP}_{\mathscr{X}}(x)+\int_{\mathscr{Y}}\mathrm{d}^2_{\mathscr{Y}}(y,y_0)\mathrm{dP}_{\mathscr{Y}}(y)}. \quad (16)$$

This proves point 1 under the two sets of hypotheses. The symmetry (point 2) and the triangular inequality (point 3) are direct. The additional points concerning the isomorphisms are a direct consequence of the Gromov's Theorem, recalled in Theorem 1, and the fact that two probability distributions on a Euclidean space are equal if and only if their characteristic functions coincide on this space.

A.2.2 Proof of Proposition 4

For $p = 2$, according to (15) and the definition of the Gromov-Wasserstein distance, we get that:

$$\Delta_r(\rho\theta) \leq \rho\sqrt{\frac{r(r-1)}{2}}\mathscr{GW}_2(\Xi_{\mathscr{X}},\Xi_{\mathscr{Y}}).$$

Moreover, when the spaces $(\mathscr{X},\mathrm{d}_{\mathscr{X}})$ and $(\mathscr{Y},\mathrm{d}_{\mathscr{Y}})$ are equal, since $|\mathrm{d}_{\mathscr{X}}(x_1,x_2)-\mathrm{d}_{\mathscr{Y}}(y_1,y_2)|^2 \leq |\mathrm{d}_{\mathscr{X}}(x_1,y_1)+\mathrm{d}_{\mathscr{X}}(x_2,y_2)|^2 \leq 2\mathrm{d}_{\mathscr{X}}(x_1,y_1)^2+2\mathrm{d}_{\mathscr{X}}(x_2,y_2)^2$, it follows that for every transport map $\pi \in \Pi(\mathrm{P}_{\mathscr{X}},\mathrm{P}_{\mathscr{Y}})$, $\Delta_r(\rho\theta) \leq 2\rho\sqrt{\frac{r(r-1)}{2}}\sqrt{\int_{\mathscr{X}\times\mathscr{Y}}\mathrm{d}_{\mathscr{X}}(x,y)^2\mathrm{d}\pi(x,y)}$. Therefore,

$$\Delta_r(\rho\theta) \leq 2\rho\sqrt{\frac{r(r-1)}{2}}\mathscr{W}_2(\mathrm{P}_{\mathscr{X}},\mathrm{P}_{\mathscr{Y}}).$$

The inequalities follow using Fubini-Tonelli theorem.

For $p = 1$, the results follows from (11).

A.2.3 Proof of Proposition 6

First, we use the bounded differences inequality [6, Theorem 6.2]. If $g(X_1,\ldots,X_n) = \mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}},\Xi_{\mathscr{X}_n})$, we get that $|g(X_1,\ldots,X_i,\ldots,X_n)-g(X_1,\ldots,X_i',\ldots,X_n)| \leq c_i$ for every $X_i, X_i' \in \mathscr{X}$, with $c_i = \sum_{r=2}^{+\infty}\alpha_r \times 2\frac{n^r-(n-1)^r}{n^r} \leq \frac{2}{n}\sum_{r=2}^{+\infty} r\alpha_r$, since $n^r-(n-1)^r$ is the number of configurations of r-uplets in $\mathscr{X}_n$ in which X_i does not appear. Therefore, for every $\ell > 0$, with probability at least $1-n^{-\ell}$:

$$\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}},\Xi_{\mathscr{X}_n}) \leq \mathbb{E}[\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}},\Xi_{\mathscr{X}_n})] + \sqrt{\frac{2\ell\log(n)}{n}}\sum_{r=2}^{+\infty} r\alpha_r.$$

Now, we focus on the expectation term, that becomes, using Fubini-Tonelli Theorem:

$$\mathbb{E}[\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}},\Xi_{\mathscr{X}_n})] = \sum_{r=2}^{+\infty}\alpha_r\int_{\rho=0}^{+\infty}\int_{\theta\in\mathscr{S}^{\frac{r(r-1)}{2}-1}}\mathbb{E}\,|\phi_{\mathscr{X},r}(\rho\theta)-\phi_{\mathscr{X}_n,r}(\rho\theta)|\,\mathrm{d}\mu_r(\theta)\mathrm{d}\omega(\rho).$$

If $r \geq n$, we have that:

$$\mathbb{E}\left|\phi_{\mathscr{X},r}(\rho\theta) - \phi_{\mathscr{X}_n,r}(\rho\theta)\right| \leq 2.$$

Assume that $r < n$, then, we have that:

$$\begin{aligned}\mathbb{E}\left|\phi_{\mathscr{X},r}(\rho\theta) - \phi_{\mathscr{X}_n,r}(\rho\theta)\right| &= \mathbb{E}\left|\phi_{\mathscr{X},r}(\rho\theta) - \frac{1}{n^r}\sum_{1\leq i_1,\ldots,i_r\leq n}\exp(i\langle D_{i_1,\ldots,i_r},\rho\theta\rangle)\right|\\ &\leq \frac{A_n^r}{n^r}\mathbb{E}\left|\frac{1}{C_n^r}\sum_{1\leq i_1<\ldots<i_r\leq n}(\phi_{\mathscr{X},r}(\rho\theta) - Y_{i_1,\ldots,i_r})\right| + 2\frac{n^r - A_n^r}{n^r}.\end{aligned}$$

where $D_{i_1,\ldots,i_r} = (\mathrm{d}_{\mathscr{X}}(X_{i_i}, X_{i_j}))_{1\leq i<j\leq r}$ and $Y_{i_1,\ldots,i_r} = \frac{1}{r!}\sum_{\sigma\in\mathfrak{S}_r}\exp(i\langle D_{\sigma(i_1),\ldots,\sigma(i_r)},\rho\theta\rangle)$ is symmetric, $\mathfrak{S}_r$ the set of all permutations of r elements, $A_n^r = \frac{n!}{(n-r)!}$ and $C_n^r = \frac{n!}{r!(n-r)!}$, since $n^r - A_n^r$ is the number of r-uplets with non distinct indices. Notice that $\mathbb{E}[Y_{i_1,\ldots,i_r}] = \phi_{\mathscr{X},r}(\rho\theta)$. Consequently, the expectation of the U-statistic is bounded by:

$$\begin{aligned}\mathbb{E}\left|\frac{1}{C_n^r}\sum_{1\leq i_1<\ldots<i_r\leq n}(\phi_{\mathscr{X},r}(\rho\theta) - Y_{i_1,\ldots,i_r})\right| &\leq \sqrt{\mathbb{V}\left(\frac{1}{C_n^r}\sum_{1\leq i_1<\ldots<i_r\leq n}Y_{i_1,\ldots,i_r}\right)}\\ &\leq \sqrt{\frac{r}{n}},\end{aligned}$$

according to [27, Lemma A, Section 5.2.1], since $\mathbb{V}(Y_{1,\ldots,r}) \leq 1$.

We also get that:

$$2\sum_{r=2}^{+\infty}\alpha_r\mathbb{1}_{r\geq n} \leq \frac{2}{\sqrt{n}}\sum_{r=2}^{+\infty}\sqrt{r}\alpha_r.$$

Moreover, for every $n \in \mathbb{N}^*$, the sequence $(u_{r,n})_{r\geq 2}$ defined by $u_{r,n} = 2\frac{n^r - A_n^r}{n^r}$ satisfies: $u_{r,n} - u_{r-1,n} = \frac{2(r-1)}{n}\frac{A_n^{r-1}}{n^{r-1}} \leq 2\frac{r-1}{n}$, with $u_{2,n} = \frac{2}{n}$, so that for every $r \geq 2$, $u_{r,n} \leq \frac{2}{n}\sum_{k=1}^{r-1}k = \frac{r(r-1)}{n} \leq \frac{r^2}{n}$.

A.2.4 Proof of Example 7

Let $\Xi_{\mathscr{X}} = (\mathscr{X}, \mathrm{d}_{\mathscr{X}}, \mathrm{P}_{\mathscr{X}})$ denote the uniform graph with $|\mathscr{X}| = V > 1$ vertices.

We get that:

$$
\begin{aligned}
&\mathbb{E}\left[\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}},\Xi_{\mathscr{X}_n})\right]\\
&\geq \alpha_2\int_0^{+\infty}\left|\mathbb{E}_{(X,X')\sim \mathrm{P}_{\mathscr{X}}^{\otimes 2}}\left[\exp(i\mathrm{d}_{\mathscr{X}}(X,X')\rho)\right]-\frac{1}{n^2}\sum_{1\leq \ell,j\leq n}\exp\left(i\mathrm{d}_{\mathscr{X}}(X_\ell,X_j)\rho\right)\right|\mathrm{d}\omega(\rho)\\
&=\alpha_2\left(\int_0^{+\infty}|1-\exp(i\rho)|\,\mathrm{d}\omega(\rho)\right)\mathbb{E}\left|\frac{1}{V}-\frac{1}{n^2}\sum_{1\leq \ell,j\leq n}\mathbb{1}_{X_\ell=X_j}\right|\\
&=\alpha_2\left(\int_0^{+\infty}|1-\exp(i\rho)|\,\mathrm{d}\omega(\rho)\right)\frac{\sigma_V}{\sqrt{n}}\mathbb{E}\left|\frac{\sqrt{n}}{\sigma_V}\left(\frac{1}{n^2}\sum_{1\leq \ell,j\leq n}\mathbb{1}_{X_\ell=X_j}-\frac{1}{V}\right)\right|\\
&\geq \alpha_2\left(\int_0^{+\infty}|1-\exp(i\rho)|\,\mathrm{d}\omega(\rho)\right)\frac{\sigma_V}{\sqrt{n}}\frac{n-1}{n}\Bigg(\mathbb{E}\left|\frac{\sqrt{n}}{\sigma_V}\left(\frac{1}{C_n^2}\sum_{1\leq \ell<j\leq n}\mathbb{1}_{X_\ell=X_j}-\frac{1}{V}\right)\right|\\
&-\left|\frac{\sqrt{n}}{\sigma_V}\frac{1}{n(n-1)}\sum_{\ell=1}^n\left(1-\frac{1}{V}\right)\right|\Bigg)
\end{aligned}
$$

According to the central limit theorem for U-statistics [20], see also [27, Theorem A, Section 5.5], for $\sigma_V=2\sqrt{\frac{1}{V}\left(1-\frac{1}{V}\right)}$, it follows that for some absolute constant $C>0$:

$$
\lim_{n\to+\infty}\inf\ \mathbb{E}\left|\frac{\sqrt{n}}{\sigma_V}\left(\frac{1}{C_n^2}\sum_{1\leq \ell<j\leq n}\mathbb{1}_{X_\ell=X_j}-\frac{1}{V}\right)\right|\geq C,
$$

and $\left|\frac{\sqrt{n}}{\sigma_V}\frac{1}{n^2}\sum_{\ell=1}^n\left(1-\frac{1}{V}\right)\right|=O\left(\frac{\sqrt{V}}{\sqrt{n}}\right)$, so that,

$$
\lim_{n\to+\infty}\inf\ \frac{\sqrt{n}}{\sigma_V}\mathbb{E}\left[\mathrm{d}_{\boldsymbol{\alpha},\boldsymbol{\mu},\omega}(\Xi_{\mathscr{X}},\Xi_{\mathscr{X}_n})\right]\geq C\alpha_2\int_0^{+\infty}|1-\exp(i\rho)|\,\mathrm{d}\omega(\rho).
$$

B Additional numerical experiments

This section is complementary to Section 3.1.

We focus on the permutation procedure, that consists, given a permutation $\sigma\in\mathfrak{S}_{n+m}$ of the pooled sample $(Z_1,\ldots,Z_{n+m})=(X_1,\ldots,X_n,Y_1,\ldots,Y_m)$, to consider two new samples $\mathscr{X}_n^b=(Z_{\sigma(1)},\ldots,Z_{\sigma(n)})$ and $\mathscr{Y}_m^b=(Z_{\sigma(n+1)},\ldots,Z_{\sigma(n+m)})$. More precisely, we proceed as follows:

1. Calculate $T_{n,m,obs}=\mathrm{d}(\mathscr{X}_n,\mathscr{Y}_m)$ from the original samples $\mathscr{X}_n$ and $\mathscr{Y}_m$.
2. Generate B permutation samples $\mathscr{X}_n^b$ and $\mathscr{Y}_m^b$ from $\mathscr{X}_n\cup\mathscr{Y}_m$.
3. Calculate $T_{n,m}^b=\mathrm{d}(\mathscr{X}_n^b,\mathscr{Y}_m^b)$ for each permutation samples $\mathscr{X}_n^b$ and $\mathscr{Y}_m^b$, $b\in\{1,\ldots,B\}$.
4. Approximate the p-value of the test by $\hat{p}=\frac{\mathrm{Card}\{b\in\{1,\ldots,B\},\,T_{n,m}^b\geq T_{n,m,obs}\}}{B}$.

The power of the test with level $\alpha\in\{0.05,0.1\}$ is estimated by $\hat{\mathrm{P}}_\alpha=\frac{\mathrm{Card}\{m\in\{1,\ldots,M\},\,\hat{p}_m\leq\alpha\}}{M}$, after $M=100$ independent replications of the experiment.

$p2$	sample size	$d^{boot,2}$	$d^{boot,5}$	$\tilde{d}^{boot,2}$	$\tilde{d}^{boot,3}$	d^{shp}	$d^{dtm,0.05}$	$d^{dtm,0.5}$	$d^{dtm,1}$
1.4	10	0.11	0.05	0.05	0.18	0.12	0.00	0.10	0.10
	20	0.13	0.12	0.07	0.13	0.08	0.00	0.09	0.08
	50	0.39	0.44	0.04	0.11	0.23	0.10	0.07	0.16
1.45	10	0.11	0.08	0.09	0.08	0.07	0.00	0.11	0.06
	20	0.07	0.07	0.07	0.02	0.04	0.00	0.05	0.06
	50	0.11	0.18	0.04	0.06	0.09	0.07	0.10	0.08
1.5	10	0.07	0.08	0.04	0.06	0.05	0.00	0.06	0.03
	20	0.07	0.02	0.06	0.07	0.05	0.00	0.06	0.03
	50	0.05	0.05	0.07	0.04	0.03	0.05	0.05	0.05
1.55	10	0.05	0.05	0.09	0.04	0.05	0.00	0.06	0.02
	20	0.09	0.09	0.05	0.07	0.04	0.00	0.06	0.05
	50	0.08	0.13	0.04	0.09	0.05	0.05	0.02	0.04
1.6	10	0.08	0.06	0.04	0.05	0.04	0.00	0.05	0.05
	20	0.11	0.14	0.05	0.03	0.06	0.00	0.04	0.05
	50	0.25	0.39	0.07	0.07	0.25	0.10	0.06	0.16

Table 7 l_p-balls, $\alpha = 0.05$

$p2$	sample size	$d^{boot,2}$	$d^{boot,5}$	$\tilde{d}^{boot,2}$	$\tilde{d}^{boot,3}$	d^{shp}	$d^{dtm,0.05}$	$d^{dtm,0.5}$	$d^{dtm,1}$
1.4	10	0.18	0.27	0.07	0.1	0.20	0.00	0.17	0.16
	20	0.23	0.20	0.12	0.21	0.13	0.00	0.11	0.12
	50	0.47	0.63	0.09	0.15	0.33	0.13	0.16	0.22
1.45	10	0.12	0.14	0.11	0.12	0.13	0.00	0.15	0.12
	20	0.14	0.13	0.11	0.06	0.20	0.00	0.14	0.13
	50	0.17	0.28	0.12	0.13	0.20	0.12	0.13	0.15
1.5	10	0.11	0.10	0.14	0.14	0.08	0.00	0.12	0.11
	20	0.10	0.06	0.1	0.13	0.08	0.00	0.11	0.04
	50	0.14	0.13	0.1	0.11	0.06	0.09	0.08	0.09
1.55	10	0.08	0.11	0.09	0.13	0.15	0.00	0.11	0.09
	20	0.15	0.19	0.12	0.14	0.11	0.00	0.09	0.14
	50	0.12	0.18	0.09	0.1	0.12	0.09	0.04	0.07
1.6	10	0.15	0.13	0.08	0.11	0.08	0.00	0.09	0.12
	20	0.15	0.25	0.09	0.1	0.09	0.00	0.12	0.07
	50	0.39	0.53	0.14	0.13	0.37	0.18	0.15	0.25

Table 8 l_p-balls, $\alpha = 0.1$

κ_2	sample size	$d^{boot,2}$	$d^{boot,5}$	$\tilde{d}^{boot,2}$	$\tilde{d}^{boot,3}$	d^{shp}	$d^{dtm,0.05}$	$d^{dtm,0.5}$	$d^{dtm,1}$
1.0	10	0.19	0.22	0.16	0.19	0.21	0.00	0.17	0.21
	20	0.32	0.28	0.31	0.32	0.32	0.00	0.30	0.34
	50	0.63	0.66	0.41	0.56	0.68	0.3	0.64	0.67
1.5	10	0.09	0.10	0.05	0.06	0.12	0.00	0.10	0.11
	20	0.09	0.06	0.09	0.09	0.09	0.00	0.06	0.09
	50	0.20	0.16	0.16	0.12	0.21	0.09	0.20	0.21
2.0	10	0.05	0.08	0.07	0.07	0.06	0.00	0.05	0.06
	20	0.06	0.04	0.05	0.05	0.08	0.00	0.08	0.08
	50	0.17	0.14	0.1	0.08	0.15	0.08	0.14	0.14
2.5	10	0.07	0.07	0.05	0.07	0.08	0.00	0.08	0.09
	20	0.11	0.13	0.12	0.11	0.13	0.00	0.14	0.13
	50	0.14	0.13	0.09	0.14	0.15	0.04	0.14	0.15
3.0	10	0.13	0.11	0.14	0.12	0.13	0.00	0.15	0.13
	20	0.20	0.20	0.15	0.17	0.23	0.00	0.18	0.24
	50	0.35	0.35	0.24	0.29	0.39	0.09	0.34	0.37

Table 9 von Mises distributions, $\alpha = 0.05$

κ_2	sample size	$d^{boot,2}$	$d^{boot,5}$	$\tilde{d}^{boot,2}$	$\tilde{d}^{boot,3}$	d^{shp}	$d^{dtm,0.05}$	$d^{dtm,0.5}$	$d^{dtm,1}$
1.0	10	0.25	0.26	0.27	0.27	0.26	0.00	0.25	0.26
	20	0.43	0.44	0.38	0.44	0.40	0.00	0.42	0.42
	50	0.74	0.80	0.54	0.62	0.81	0.45	0.78	0.83
1.5	10	0.16	0.15	0.12	0.13	0.17	0.00	0.17	0.16
	20	0.13	0.11	0.15	0.13	0.12	0.00	0.13	0.12
	50	0.29	0.26	0.21	0.19	0.32	0.15	0.28	0.30
2.0	10	0.11	0.13	0.11	0.11	0.11	0.00	0.07	0.11
	20	0.12	0.13	0.13	0.08	0.11	0.00	0.12	0.10
	50	0.22	0.27	0.12	0.14	0.22	0.22	0.2	0.23
2.5	10	0.12	0.13	0.12	0.1	0.13	0.00	0.17	0.14
	20	0.15	0.17	0.18	0.22	0.17	0.00	0.17	0.17
	50	0.21	0.23	0.14	0.21	0.24	0.12	0.27	0.24
3.0	10	0.19	0.19	0.22	0.23	0.20	0.00	0.2	0.2
	20	0.31	0.31	0.27	0.22	0.31	0.00	0.25	0.34
	50	0.48	0.43	0.33	0.38	0.54	0.19	0.46	0.52

Table 10 von Mises distributions, $\alpha = 0.1$