Computational Statistics and Data Analysis 224 (2026) 108447

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Metric Oja depth, new statistical tool for estimating the most central objects

Vida Zamanifarizhandi *, Joni Virta

Department of Mathematics and Statistics, University of Turku, Finland

ARTICLE INFO

Keywords:
Object data
Metric Oja depth
Statistical depth
Optimization
Metric statistics

ABSTRACT

The Oja depth (simplicial volume depth) is one of the classical statistical techniques for measuring the central tendency of data in multivariate space. Despite the widespread emergence of object data like images, texts, matrices or graphs, a well-developed and suitable version of Oja depth for object data is lacking. To address this shortcoming, a novel measure of statistical depth, the metric Oja depth applicable to any object data, is proposed. Two competing strategies are used for optimizing metric depth functions, i.e., for finding the deepest objects among the sample. The performance of the metric Oja depth is compared with three other depth functions (half-space, lens, and spatial) in diverse data scenarios.

1. Introduction

1.1. Background

Nowadays, data are being generated in high volumes, at high speeds, and with great diversity. A key aspect of modern data analysis is that we no longer deal exclusively with data existing in Euclidean spaces. Instead, data are currently taking on more complex formats, such as images, graphs, matrices, etc., all of which reside in non-Euclidean spaces and are collectively known as *object data*. Methodology for object data, known as object data analysis or metric statistics (Dubey et al., 2024), are currently abundant in statistical literature; see, e.g., Dubey and Müller (2022), Virta et al. (2022), Zhang et al. (2024), Zhu and Müller (2025). Object data are typically modelled as residing in an arbitrary metric space $(\mathcal{X}, d)$ and we also adopt this approach.

As with any data, the analysis of object data begins with exploratory data analysis on the raw data. This can take various forms: calculating descriptive statistics, dimension reduction, detecting outliers etc. However, despite being straightforward for Euclidean data, these tasks can be surprisingly difficult in the context of object data, due to the lack of structure in $(\mathcal{X}, d)$. In this work, we focus on one of the most fundamental exploratory statistical tasks, location/mean/average estimation, using *depth functions* as our tool of choice.

The foundation for depth functions was introduced in 1975 by Tukey who proposed his seminal half-space depth for multivariate data to rank observations and reveal features of the underlying data distribution (Tukey, 1975). Apart from this, numerous depth functions for data in Euclidean space have been proposed. E.g., the convex hull peeling depth (onion depth) (Barnett, 1976; Eddy, 1981), simplicial volume depth (Oja depth) (Oja, 1983) and several others; see Table 1 and the comprehensive addressing in Mosler and Mozharovskyi (2022).

* The corresponding author.
E-mail address: vizama@utu.fi (V. Zamanifarizhandi).

https://doi.org/10.1016/j.csda.2026.108447
Received 15 November 2024; Received in revised form 21 June 2026; Accepted 22 June 2026
Available online 23 June 2026

Table 1
List of depth functions. In the bottom row, the metric β-skeleton depth has been considered only for $\beta = 2$, where the concept reduces to lens depth. The listed complexities are for computing depths on the full training sample of size n; For metric half-space depth, the complexity assumes the use of (Dai and Lopez-Pintado, 2023, Algorithm 1) with the sample itself as anchors. The complexities consist of two parts, distance matrix formation and depth computation, where in the former, h denotes the cost of computing the distance between a single pair of objects in $(\mathcal{X}, d)$; see Section 3.2 for further details.

Author	Year	Depth function	Has a metric version	Complexity
Mahalanobis (Mahalanobis, 1936)	1936	Mahalanobis depth (distance)	No	–
Tukey (Tukey, 1975)	1975	Half-space/location/Tukey depth	Yes (Dai and Lopez-Pintado, 2023)	$\mathcal{O}(n^2\ h) + \mathcal{O}(n^3)$
Barnett, Eddy (Barnett, 1976; Eddy, 1981)	1976, 1981	Convex hull peeling/onion depth	No	–
Liu; Zuo & Serfling (Liu, 1992; Zuo and Serfling, 2000)	1992, 2000	Projection depth	No	–
Liu (Liu, 1990)	1990	Simplicial depth	No	–
Oja, Zuo & Serfling (Oja, 1983)	1983, 2000	Simplicial volume depth/Oja depth	Yes, in the current paper	$\mathcal{O}(n^2\ h) + \mathcal{O}(n^4)$
Koshevoy & Mosler (Koshevoy and Mosler, 1997)	1997	Zonoid depth	No	–
Vardi & Zhang, Serfling (Vardi and Zhang, 2000; Serfling, 2002)	2000, 2002	Spatial depth	Yes (Virta, 2026)	$\mathcal{O}(n^2\ h) + \mathcal{O}(n^3)$
Liu & Modarres (Liu and Modarres, 2011)	2011	Lens depth	Yes (Kleindessner and Von Luxburg, 2017; Cholaquidis et al., 2023; Geenens et al., 2023)	$\mathcal{O}(n^2\ h) + \mathcal{O}(n^3)$
Yang & Modarres (Yang and Modarres, 2018)	2018	β-skeleton depths	Not for general β, see the caption	–

In recent years, object/metric versions of various depth functions D have been proposed; see the second-to-last column of Table 1. That is, given an object $X \in \mathcal{X}$ and a distribution P taking values in $\mathcal{X}$, the depth $D(X; P)$ describes how central the object X is w.r.t. P, much in the same way as for the Euclidean depths earlier. Two key properties of these extensions are that (a) they depend on the data only through the metric d, making them applicable to *any* form of object data, regardless in which specific metric space they live, and (b) if $(\mathcal{X}, d)$ is a Euclidean space, then the original Euclidean version of the depth is recovered, showing that these metric depth functions are "true" generalizations of their classical counterparts. Metric lens depth was developed in Kleindessner and Von Luxburg (2017), Cholaquidis et al. (2023), Geenens et al. (2023), metric half-space depth was proposed in Dai and Lopez-Pintado (2023) and metric spatial depth in Virta (2026).

Most of the depth functions in Table 1, in particular all three that have object versions, are *robust* (Maronna et al., 2019), meaning that their performance is not skewed by the presence of possible outliers. In Section 2 we list the robustness properties of the existing metric depth functions, and more details on the robustness of Euclidean depths can be found in (Mozharovskyi and Valla, 2025, Table 1) and Zuo and Serfling (2000), Mozharovskyi (2022). Note that not all location estimation procedures are robust. For example, the Fréchet mean (Fréchet, 1948), which is a generalization of the concept of average and a classical estimator of location for object data, is not robust, but is instead affected greatly by data outliers. Despite the popularity of Fréchet mean and similar tools, in this work we concentrate solely on robust methods. This is because object data are very diverse, which makes recognizing any possible outliers very difficult. Robust analysis methods protect against the influence of outliers and leads to reliable results, regardless of the exact type of outliers and whether they are detected in the data or not.

1.2. Our contributions

This work focuses on the robust location estimation of object data using depth functions, and its main contributions are as follows.

(i) We propose a new depth function for object data, the metric Oja depth, and extensively study its theoretical properties. In particular, we characterize its range of values and show the sample consistency of its maximizers under specific assumptions.

(ii) We study the estimation of the deepest objects (data location). That is, given a sample $X_1, \dots, X_n$ of object data, we aim to find the object $X \in \mathcal{X}$ which has the maximal depth value w.r.t. the sample. Despite the fundamental nature of this problem, in the earlier works on metric depths it has been considered only by Dai and Lopez-Pintado (2023), who proposed finding the optimum among the sample points $X_1, \dots, X_n$, or using more elaborate manifold optimization to find the deepest out-of-sample object, as implemented in the R-package `MHD` (Dai, 2024). As a competing approach, we propose using non-linear Euclidean optimizers along with a PCA-reduced coordinate representation of the object data to estimate the deepest out-of-sample object.

(iii) We conduct extensive simulation and data comparisons between our proposed depth and its competitors in estimating the deepest object in several data scenarios. In particular, we investigate how much the in-sample estimation can be improved by using non-linear optimization to achieve out-of-sample estimation. The results show that, at the cost of greater computational complexity, metric Oja depth surpasses its competitors in several scenarios, especially when combined with our proposed optimization strategies.

1.3. Contents

We review the three existing metric depth functions in Section 2. In Section 3, we introduce the proposed new metric depth function and develop its theoretical properties. In Section 4, we evaluate and compare it with three other depth functions, using two simulation scenarios. In Section 5, we further test our metric depth function on a real dataset, using permutation and rank tests to obtain inferential results. Finally, future recommendations are compiled in Section 6.

2. Review of existing metric depth functions

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and take $(\mathcal{X}, d)$ to be a complete and separable metric space where our data reside. Recall that a metric $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ quantifies distances between objects residing in the set $\mathcal{X}$. Formally, metrics are required to satisfy the following axioms: Positivity, $d(x, y) \geq 0$ with equality if and only if $x = y$; Symmetry, $d(x, y) = d(y, x)$; and Triangle inequality, $d(x, y) \leq d(x, z) + d(z, y)$ (Burago et al., 2001; Gromov, 2007). Further, we let P be a probability measure defined on the Borel sets of $\mathcal{X}$. Given a fixed, non-random object $x \in \mathcal{X}$, the metric depth functions answer the question "how central is the object x with respect to the distribution P?".

To tackle this question, Kleindessner and Von Luxburg (2017), Cholaquidis et al. (2023), Geenens et al. (2023) proposed metric lens depth, defined as

$$D_L(x) := \mathbb{P}\big(d(X_1, X_2) > \max\big\{d(X_1, x), d(X_2, x)\big\}\big),$$

where X_1, X_2 are independent random variables with the distribution P. Drawing an analogy from Euclidean statistics, $D_L(x)$ essentially gives the probability that the side corresponding to X_1, X_2 is the longest in a "triangle" drawn using the objects x, X_1, X_2. Intuitively, the probability is small (large) if x is located far away from (close to) the bulk of the distribution P, making D_L behave as expected from a depth function.

Dai and Lopez-Pintado (2023) gave a similar treatment to the classical half-space depth and defined the metric half-space depth as

$$D_H(x) = \inf_{x_1, x_2 \in \mathcal{X},\ d(x_1, x) \leq d(x_2, x)} \mathbb{P}(d(X, x_1) \leq d(X, x_2)).$$

Every pair of objects x_1, x_2 in the infimum divides the space $\mathcal{H}$ in two subsets (objects closer to x_1 than x_2 and vice versa). The depth $D_H(x)$ is defined to be the smallest possible probability mass of a subset produced in this way and containing the object x. As with the metric lens depth, it is easy to see that, for an outlying object x, it is easy to find a subset which contains x but has very small P-probability mass, making $D_H(x)$ small. An approximative algorithm, that we also use in this paper to compute the metric half-space depth, is given in (Dai and Lopez-Pintado, 2023, Algorithm 1).

The metric spatial depth was proposed in Virta (2026) as

$$D_S(x) = 1 - \frac{1}{2}\mathbb{E}\left[\mathbb{I}(d(X_1, x) \neq 0, d(X_2, x) \neq 0)\left\{\frac{d^2(X_1, x) + d^2(X_2, x) - d^2(X_1, X_2)}{d(X_1, x)d(X_2, x)}\right\}\right],$$

where $X_1, X_2 \sim P$ are independent. The quantity $D_S(x)$ takes values in $[0, 2]$ and is best interpreted via the endpoints of this interval, which are characterized by equality cases of the triangle inequality between specific permutations of x and two random objects X_1, X_2. Namely, the value 0 is reached if and only if the three points are almost surely on a "line" with x never in the middle (that is, the probability of $\{d(X_1, x) = d(X_1, X_2) + d(X_2, x)\} \cup \{d(X_2, x) = d(X_2, X_1) + d(X_1, x)\}$ is 1). The value 2 is attained if and only if the points are almost surely on a "line" with x in the middle (i.e., the event $d(X_1, X_2) = d(X_1, x) + d(x, X_2)$ holds with probability 1). More details on interpreting the metric spatial depth are given in Virta (2026).

Throughout this work, we refer to metric lens depth, metric half-space depth and metric spatial depth as MLD, MHD and MSD, respectively. All of these have the property of *vanishing at infinity* meaning that, for objects far enough away from the bulk of the distribution P, the depth values approach zero. All three also satisfy specific forms of continuity (small changes in x cause small changes in the depth) and both MLD and MHD are invariant to data transformations that preserve the ordering of distances. Regarding robustness properties, MHD and MSD are robust, the former in the sense of having a high breakdown point and the latter in the sense of having a bounded influence function. Whereas, MLD is also likely to be highly robust, being based only on comparisons between distances and not their absolute sizes, but this aspect of MLD has not yet been studied in the literature. For detailed statements of these, and some other properties of the three depth functions, including sample consistency results, we refer the reader to the papers (Cholaquidis et al., 2023; Geenens et al., 2023; Dai and Lopez-Pintado, 2023; Virta, 2026).

3. New depth for object data

3.1. Population-level properties

For any three objects $x_1, x_2, x_3 \in \mathcal{X}$, we use the notation $L(x_1, x_2, x_3)$ to denote the event that $d(x_1, x_3) = d(x_1, x_2) + d(x_2, x_3)$. On an intuitive level, $L(x_1, x_2, x_3)$ means that x_2 resides "between" x_1 and x_3, since under this event the distance from x_1 to x_3 is the same as when taking a detour through x_2. Hence, in the sequel we use phrases such as "x_2 lies in between x_1 and x_3" to indicate that $L(x_1, x_2, x_3)$ holds. For arbitrary objects $x_0, x_1, x_2, x_3 \in \mathcal{X}$, we denote by $B_3(x_0, x_1, x_2, x_3)$ the 3×3 matrix whose (k, ℓ)-element equals

$$\frac{1}{2}\big\{d^2(x_0, x_k) + d^2(x_0, x_\ell) - d^2(x_k, x_\ell)\big\}. \tag{1}$$

Letting $X_1, X_2, X_3 \sim P$ be i.i.d. random objects, we propose to measure the depth of a fixed object $x \in \mathcal{X}$ with respect to the distribution P as

$$D_{O3}(x) := \frac{1}{1 + \mathbb{E}\left[\left\{|B_3(x, X_1, X_2, X_3)| + 4d^2(x, X_1)d^2(x, X_2)d^2(x, X_3)\right\}^{1/2}\right]}. \tag{2}$$

The subscript O refers to "Oja" due to the close connection between $D_{O3}(x)$ and the classical Oja depth explained in Appendix A. By Theorem B.1 in Appendix B, taking the square root in the denominator of (2) is well-defined. The intuitive meaning of $D_{O3}(x)$ is not clear from its definition (2) alone, so we next present a collection of results that allow better understanding it, starting with a moment condition that guarantees its existence.

Theorem 1. *Assume that, for some $a \in \mathcal{X}$, we have $\mathbb{E}\{d(X, a)\} < \infty$. Then $D_{O3}(x)$ exists as well-defined.*

By "well-defined" in Theorem 1 we mean that the expected value used to compute $D_{O3}(x)$ exists as finite. The condition $\mathbb{E}\{d(X, a)\} < \infty$ is analogous to requiring a univariate random variable to have a finite mean and is a rather mild condition as even the Fréchet mean already requires the existence of second moments. The choice of a in Theorem 1 is completely arbitrary as, if the condition $\mathbb{E}\{d(X, a)\} < \infty$ holds for some a, then, by the triangle inequality, it holds for all $a \in \mathcal{X}$.

Theorems 2 and 3 next establish the range of $D_{O3}(x)$ to be $[0, 1]$ and give interpretations to these endpoints.

Theorem 2. *$D_{O3}(x)$ takes values in $[0, 1]$. Moreover, if $D_{O3}(x) = 1$, then*

$$P\{L(X_1, x, X_2) \cup L(X_2, x, X_3) \cup L(X_3, x, X_1)\} = 1,$$

where $X_1, X_2, X_3 \sim P$ are i.i.d.

Theorem 2 implies that large values of $D_{O3}(x)$ are indicative of the centrality of the object x with respect to the distribution P. That is, under the extremal case $D_{O3}(x) = 1$ the object x must almost surely lie in between of at least two of three randomly sampled objects X_1, X_2, X_3 from P.

We say that a sequence of objects $x_n \in \mathcal{X}$ is divergent if there exists $a \in \mathcal{X}$ such that $d(x_n, a) \to \infty$ when $n \to \infty$. Intuitively, a divergent sequence is such that it moves "towards infinity" eventually getting arbitrarily far from any fixed object. Clearly, such sequences do not exist in every metric space.

Theorem 3. *Assume that, for some $a \in \mathcal{X}$, we have $\mathbb{E}\{d^2(X, a)\} < \infty$. Let x_n be a divergent sequence of objects in $\mathcal{X}$. Then, $D_{O3}(x_n) \to 0$ as $n \to \infty$.*

The second moment condition in Theorem 3 is used to avoid certain pathological behavior; see the proof of the result in Appendix B. Theorem 3 states that small values of $D_{O3}(x)$ are reached by outlying objects, i.e., points that are located far away from the bulk of the distribution P.

It is typical for depth functions to satisfy various invariance properties (Dai and Lopez-Pintado, 2023; Geenens et al., 2023). Namely, one usually expects that the depth of a point x w.r.t. to distribution P does not change if x and P are both subjected to a specific class of transformations. Being based on pair-wise distances, D_{O3} inherits such properties directly from the metric d. That is, if $d(gx_1, gx_2) = d(x_1, x_2)$ for all $g \in \mathcal{G}$ where $\mathcal{G}$ is a group of transformations on $\mathcal{X}$, then also D_{O3} is invariant to $\mathcal{G}$. In particular, if $x_0 \in \mathcal{X}$ has maximal depth w.r.t. P, then gx_0 has maximal depth w.r.t. to P_g, the distribution of gX where $X \sim P$.

3.2. Sample-level properties

Let $X_1, \dots, X_n$ be a random sample of objects from P. To estimate depths in practice, a natural sample counterpart of D_{O3} is the third-order U-statistic,

$$D_{O3,n}(x) := \frac{1}{1 + \frac{1}{\binom{n}{3}} \sum_{(i,j,k)} h(x, X_i, X_j, X_k)}, \tag{3}$$

where the summation runs over all unordered triples (i, j, k) of indices from $\{1, \dots, n\}$ with distinct elements and

$$h(x, X_1, X_2, X_3) := \{|B_3(x, X_1, X_2, X_3)| + 4d^2(x, X_1)d^2(x, X_2)d^2(x, X_3)\}^{1/2}.$$

Standard results on U-statistics (Lee, 2019) combined with the continuous mapping theorem and Theorem 1 give the following point-wise convergence result which states that the depth of any single object x can be accurately estimated using $D_{O3,n}$.

Theorem 4. *Assume that, for some $a \in \mathcal{X}$, we have $\mathbb{E}\{d(X, a)\} < \infty$. Then, for any fixed $x \in \mathcal{X}$, we have $D_{O3,n}(x) \to_p D_{O3}(x)$ as $n \to \infty$.*

A stronger result is obtained if the metric space is totally bounded, in which case the convergence is uniform in $\mathcal{X}$.

Theorem 5. *Assume that the metric space $(\mathcal{X}, d)$ is totally bounded. Then,*

$$\sup_{x \in \mathcal{X}} |D_{O3,n}(x) - D_{O3}(x)| \to_p 0,$$

as $n \to \infty$.

The classical M-estimator argument (Van der Vaart, 2000, Theorem 5.7) then shows that the point with maximal depth, assuming that it is unique, can be consistently estimated with the sample depth function $D_{O3,n}$.

Corollary 6. *Assume that the metric space $(\mathcal{X}, d)$ is totally bounded and that $x \mapsto D_{O3}(x)$ has unique maximizer x_0. Then any sequence $x_n \in \mathcal{X}$ of maximizers of $x \mapsto D_{O3,n}(x)$ satisfies $x_n \to_p x_0$ as $n \to \infty$.*

The sample maximizers in Corollary 6 exist since $D_{O3,n}$ is a continuous function on a totally bounded and complete, and hence compact, set. The assumption of a unique population maximizer is discussed further in Section 6. Also MLD and MHD satisfy analogous consistency properties; see Cholaquidis et al. (2023), Geenens et al. (2023), Dai and Lopez-Pintado (2023).

We conclude the subsection by discussing computational aspects of the proposed $D_{O3,n}$. Obtaining the depths of all sample points comprises of (A) computing the distance matrix $\{d(X_i, X_j)\}_{i,j=1}^n$ and, (B) using this to compute the depths of the sample points with (3). The cost of step (A), distance computation, depends on the metric space in question and is essentially $O(n^2\, h)$ where h is the cost of computing the distance between a single pair of objects in $(\mathcal{X}, d)$. For example, both in $\mathbb{R}^p$ with the Euclidean metric and on the unit sphere in $\mathbb{R}^p$ with the arc length metric, forming the distance matrix is an $O(n^2 p)$-operation, whereas on the manifold of $p \times p$ positive definite matrices with either the log-Euclidean or Riemannian metric, this operation is $O(n^2 p^3)$. Importantly, step (A) is identical for all metric depth functions, implying that their complexities can be compared based on step (B) only, detailed next.

Step (B), computing the actual depths, is done based on the $n \times n$ distance matrix and, as such, the complexity of this step depends only on n and not on other aspects of the sample. In particular, the dimensionality of the data plays no role in this step. The computational complexity of our proposed depth is from (3) seen to be $\mathcal{O}(n^4)$, whereas the complexity of MLD, MSD and MHD is $\mathcal{O}(n^3)$, for computing the depths of the full sample (assuming (Dai and Lopez-Pintado, 2023, Algorithm 1) with the sample itself as anchors is used for MHD). The time efficiency of our proposed depth could be improved, at the cost of accuracy, by using partial U-statistics to randomly choose which elements to include in the triple sum in (3). However, we have not resorted to this compromise here. The previous information on the complexities has been summarized also in the final column of Table 1. The complexities of the Euclidean counterparts of these depths and other Euclidean depths can be found in (Mozharovskyi, 2022, Table 1.9).

The computation of MLD and MHD does not rely on the actual distances $d(X_i, X_j)$ but only on indicators of the form $\mathbb{I}\{d(X_i, X_j) \leq d(X_i, X_k)\}$; see Section 2. This means that MLD and MHD (a) are invariant to transformations that preserve these indicators, and (b) can be computed even if one knows only the binary values of these triple comparisons. The latter viewpoint for MLD was taken in Kleindessner and Von Luxburg (2017) to find deepest objects based on binary dissimilarity data. Our proposed depth, on the other hand, uses the actual values $d(X_i, X_j)$ in its computation, meaning that it cannot be computed based on indicators alone. This limits its usefulness in such dissimilarity scenarios but, at the same time, allows our depth to use additional information from the data, which later manifests as improved performance in several experiments compared with the other methods. This behavior is perfectly in line with the original Oja depth, which also depends on the numerical values of the distances $\|X_i - X_j\|$.

3.3. Auxiliary depth-like quantity

Our proposed depth D_{O3} is based on the 3-dimensional Oja depth; see Appendix A for details, and a natural question is whether also the p-dimensional Oja depth can be used for similar purposes with $p \neq 3$. In this section we briefly investigate the version with $p = 2$. Letting $B_2(x_0, x_1, x_2)$ denote the 2×2 top left sub-matrix of $B_3(x_0, x_1, x_2, x_3)$, we define

$$D_{O2}(x) := \frac{1}{1 + \mathbb{E}\left\{|B_2(x, X_1, X_2)|^{1/2}\right\}}, \tag{4}$$

which, as shown in Appendix A, is a metric generalization of the classical 2-dimensional Oja depth. As with D_{O3}, also D_{O2} is well-defined as soon as the first moment exists, i.e., $\mathbb{E}\{d(X, a)\} < \infty$ for some $a \in \mathcal{X}$ (proof omitted). However, the next result, which follows directly from Theorem B.1(i) in Appendix B shows that $D_{O2}(x)$ does not actually measure the centrality of x w.r.t. P and is thus not a proper depth measure.

Theorem 7. *$D_{O2}(x)$ takes values in $[0, 1]$. Moreover, $D_{O2}(x) = 1$ if and only if*

$$P\{L(x, X_1, X_2) \cup L(X_2, x, X_1) \cup L(X_1, X_2, x)\} = 1,$$

where $X_1, X_2 \sim P$ are i.i.d.

Intuitively, $D_{O2}(x)$ takes a large value if, for random X_1, X_2 and the fixed object x, one of the three objects is typically located between the other two. This, however, does not guarantee that x is central with respect to P, as it is possible that X_1 (or X_2) is the in-between point. Thus, large values of $D_{O2}(x)$ do not characterize centrality. As an extreme example, if $(\mathcal{X}, d)$ is the one-dimensional Euclidean space, then $D_{O2}(x) = 1$ for all $x \in \mathbb{R}$. Consequently, we do not pursue the theoretical properties of D_{O2} further. However, we still include $D_{O2}(x)$ in our data examples, where it performs surprisingly well in locating the centers of distributions, see Section 6 for more discussion on this.

4. Simulation examples

4.1. In-sample optimization

Throughout the paper, we use the following abbreviations for the depth functions: MOD3 (metric Oja depth), MOD2 (the depth-like counterpart of MOD3 defined in Section 3.3), MHD (metric half-space depth), MLD (metric lens depth), and MSD (metric spa-

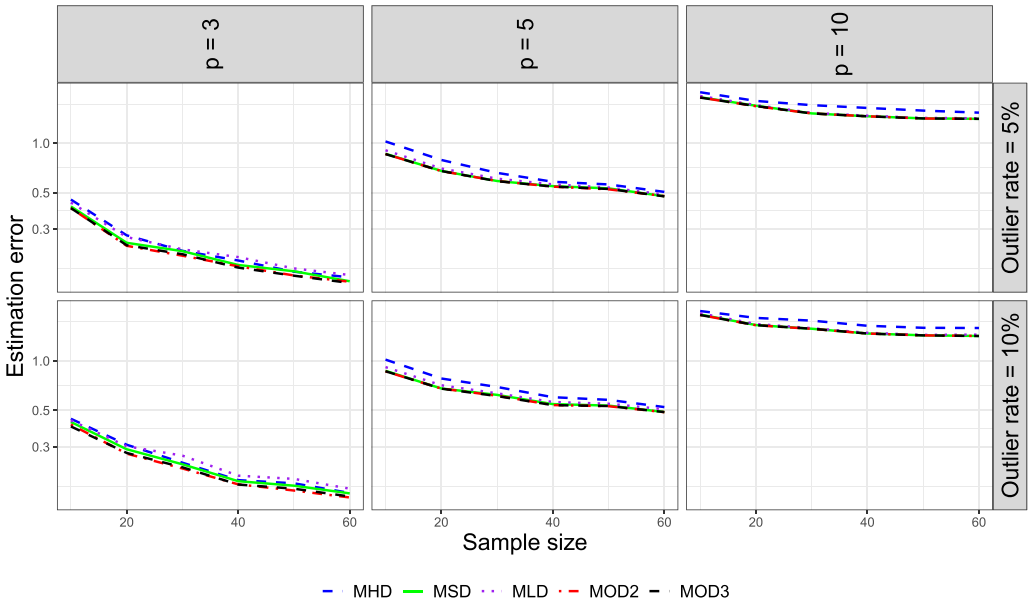

Fig. 1. The average estimation errors for each of the five methods in the correlation matrix simulation. The scale of the y-axis is logarithmic.

tial depth). The depths were implemented in `Rcpp` (Eddelbuettel and Balamuta, 2018) and are available at https://github.com/vidazamani/Depth-functions-for-Object-Data.

In our first experiment, we compare the depths with simulations in the context of location estimation. Given a sample $X_1, \dots, X_n \in \mathcal{X}$ from a distribution P, each sample depth function D_n is used to determine an in-sample estimate of the deepest object, i.e., $\hat{\mu} := X_{i_0}$ where $i_0 = \operatorname{argmin}_{i=1,\dots,n} D_n(X_i)$. The depths are compared based on the averages, over 200 replications, of the estimation errors $d(\hat{\mu}, \mu)$ where $\mu \in \mathcal{X}$ is the true location of the distribution P. In addition to the error, we also evaluate the computation times of the depths. When measuring the running times, we disregard the computation of the interobject distances, as this step is common to all competing methods.

4.1.1. *Correlation matrix dataset*

In our first simulation, we consider the Riemannian manifold $(\mathcal{X}, d)$ of $p \times p$ correlation matrices as the data space, where $d(X_1, X_2) = \|\operatorname{Log}(X_1^{-1/2} X_2 X_1^{-1/2})\|_F$ is the affine invariant Riemannian metric; see Bhatia (2009). Such data are commonly encountered, e.g., in neuroscience in the form of brain connectivity matrices (Simeon et al., 2022).

The random $p \times p$ correlation matrices $X_1, \dots, X_n$ were generated as $X_i := \operatorname{diag}(S_i)^{-1/2} S_i \operatorname{diag}(S_i)^{-1/2}$, where $S_i := U_i D_i U_i'$, U_i is a random orthogonal matrix and D_i is a diagonal matrix with all positive diagonal elements, independent of U_i. The diagonal matrices D_i were generated as follows

$$D_i = \operatorname{diag}\left(\exp\big(\mathcal{N}(\nu, 1)\big), \underbrace{\exp\big(\mathcal{N}(-\nu, 1)\big), \dots, \exp\big(\mathcal{N}(-\nu, 1)\big)}_{p-1 \text{ times}} \right).$$

The parameter ν was determined as follows: each X_i had the probability $\varepsilon > 0$ of being an outlier, in which case $\nu = 3$, whereas the remaining $1 - \varepsilon$ proportion of the data (the bulk) used $\nu = 0$. The matrices U_i are uniformly random; see Stewart (1980) for details. This choice ensures that the eigenvectors of S_i do not favor any direction. Consequently, the distribution of S_i is rotationally symmetric around the half-line $\{CI_p \mid C \geq 0\}$ and, intuitively, the most central matrix w.r.t. this distribution has to reside on this half-line. As the correlation matrices X_i are obtained from the S_i by scaling the diagonal elements to unity, any reasonable depth measure should estimate the central object of the distribution of X_i to be $\mu = I_p$. Note that both the bulk and the outliers share the same true most central object (identity matrix) but the high variance of the outliers makes the estimation more difficult for larger ε.

The simulation, performed over 200 iterations, involves three parameters, the contamination proportion $\varepsilon = 0.05, 0.10$, dimension $p = 3, 5, 10$ and sample size $n = 10, 20, 30, 40, 50, 60$, leading to a total of 36 cases. Note that n was deliberately chosen as low, since object data scenarios typically have much smaller sample sizes than in Euclidean data analysis. The average estimation errors are shown in Fig. 1, grouped by ε and p. In all cases, as n increases, the estimation error gradually decreases. Moreover, the complexity of the data space grows as the dimension p of the matrices increases, leading to a decline in accuracy. Increasing the percentage of outliers from 5% to 10% did not lead to significant changes in the results, indicating the stability and robustness of the estimators. In all cases, MOD3 and MOD2 demonstrated the best performance, while MHD demonstrated the weakest performance across all cases.

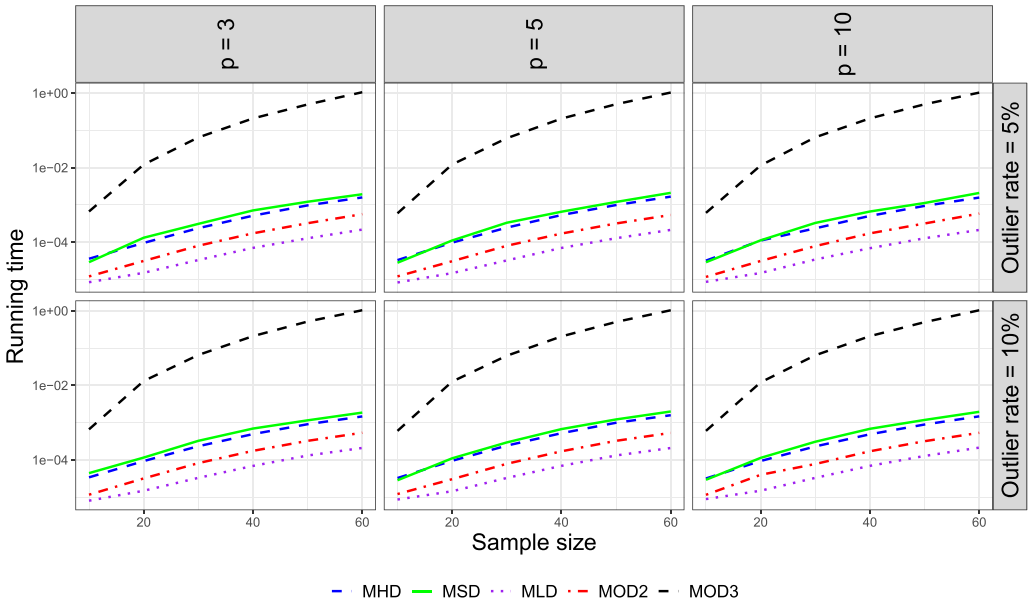

Fig. 2. The running times (in seconds) of each method in the correlation matrix simulation. The scale of the y-axis is logarithmic.

However, when $p = 3$, MHD outperforms MLD as the sample size increases, whereas for small sample sizes, MLD performs slightly better than MHD. This observation is more clearly illustrated in an alternative version of Fig. 1, provided in Appendix C, Figure C.2, where the results are presented relative to MSD. MSD was the third most efficient method and close to MOD3 and MOD2 in performance. From the timing results given in Fig. 2, we observe that neither p nor ε affect the running time. This is because neither has an effect on the size of the distance matrix; see the discussion in Section 3.2. Additionally, MOD3 has a relatively high time cost; see Fig. 2, meaning that it essentially offers improved performance at the cost of speed.

4.1.2. *Hypersphere dataset*

In this experiment, we generated samples of points on a p-dimensional unit hypersphere $\mathcal{X}$. As a metric d, we used the usual arc length distance. Each point X_i was simulated by first generating the p-dimensional random vector $Z_i \sim \mathcal{N}_p(\lambda 1_p, I_p)$, where 1_p is a vector of ones and $\lambda \neq 0$, and then taking $X_i := Z_i/\|Z_i\|$. By symmetry, this strategy leads to the most central object always being equal to $\mu = \text{sign}(\lambda)(1/\sqrt{p})1_p$ and the absolute value of the parameter λ controls the spread of the points: the larger the value of $|\lambda|$, the more closely the points are concentrated around μ. For the outliers we used $\lambda = -1$ and for the bulk $\lambda = 5$, meaning that the two distributions have their most central points on the opposite sides of the hypersphere; see Figure C.1 in Appendix C for an illustration.

The simulation thus has three parameters, ϵ, p and n, which are set to the same values as in the earlier simulation. As shown in Fig. 3 and its relative version presented in Appendix C, the performance of each metric depth function on the hypersphere dataset is very similar to its performance on the correlation data. We have omitted the timing results for the sphere simulation, as they were visually almost identical to those in Fig. 2.

We also conducted an additional simulation in the same setting with $\epsilon = 0.10$, $n = 100, 200$, and $p = 50, 100, 200, 400, 800, 1600$ to investigate the effect of increasing the data dimensionality p on the estimate. The average estimation errors over 100 replicates are plotted in Fig. 4. The results indicate that: (i) MSD, MLD, MOD2, MOD3 yield almost identical results and, as expected, their performance improves as n increases. (ii) MHD is worse than the other four and changing n does not appear to affect it. The standard half-space depth is known to degenerate in higher dimensions (Dutta et al., 2011) and it is therefore possible that MHD suffers from the same issue. A non-trivial solution could be to regularize its computation in a suitable way. (iii) For all methods, the estimation error increases with p. From the viewpoint of efficiency, for example $n = 100$ and $p = 500$ yields roughly the same accuracy as $n = 200$ and $p = 650$. (iv) As a baseline, we also considered an estimator which picks a random object from the sample and uses it as the estimate of the location. Regardless of p, this leads to estimation errors of 0.413 and 0.4078 for $n = 100$ and $n = 200$, respectively, showing that all estimators improve considerably upon this simple benchmark.

4.2. *Out-of-sample optimization*

The in-sample optimization studied in the previous section might come across as a naive approach, but in many metric spaces it is the best one can do, since the lack of Euclidean structure prevents the use of standard optimization algorithms. However, despite being non-Euclidean, some metric spaces admit coordinate representations/encodings in Euclidean spaces. Such spaces include, e.g., unit spheres (stereographic projection), the positive definite manifold (Cholesky decomposition) and the space of L_2-functions (Karhunen-

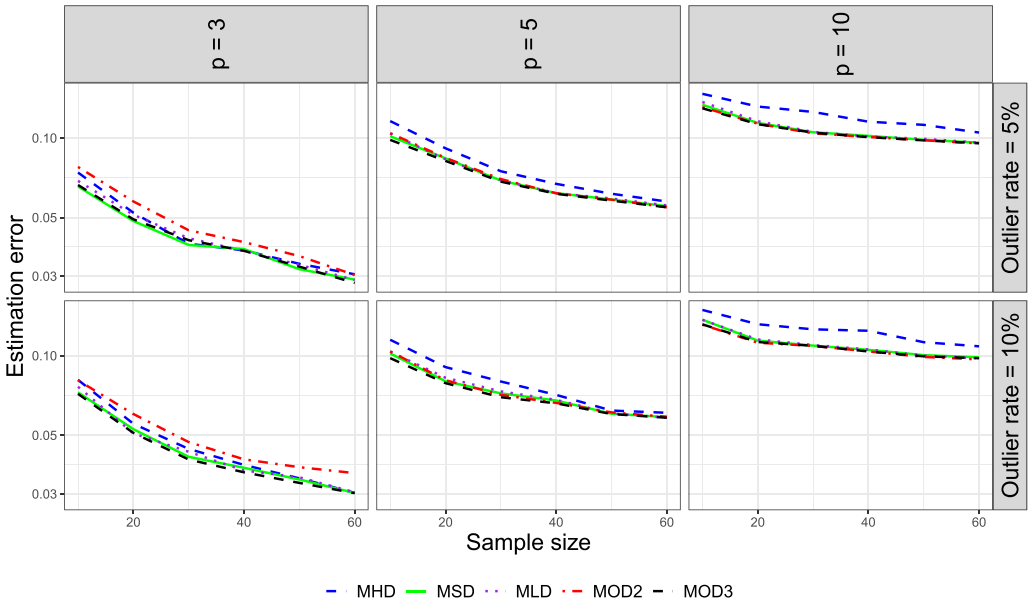

Fig. 3. The average estimation errors for each of the five methods in the hypersphere simulation. The scale of the y-axis is logarithmic.

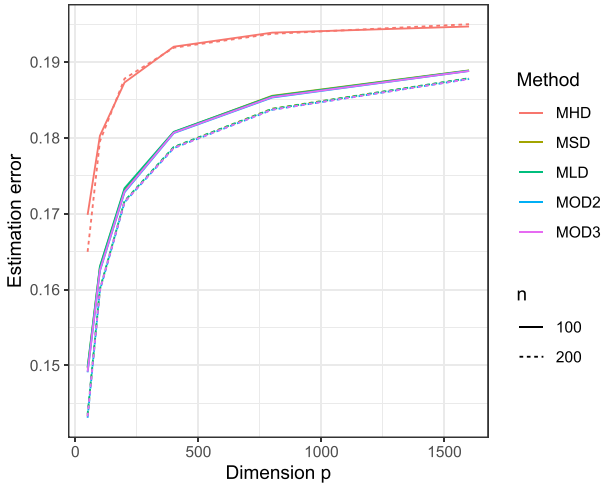

Fig. 4. The effect of increasing the dimension of the unit sphere to the in-sample estimation error, estimated using 100 replications. The lines of MSD, MLD, MOD2 and MOD3 are almost perfectly overlapping.

Loève expansion, assuming we truncate it). In these cases, Euclidean non-linear optimization can be used to find the deepest point. To formalize the earlier, assume that there exists a bijective map $f : \mathcal{X} \to S$ where $S \subseteq \mathbb{R}^q$ for some q. Then, the problem of finding the deepest object of a sample $X_1, \dots, X_n \in \mathcal{X}$ can be formulated as

$$\max_{v \in S} D_n(f^{-1}(v)),$$

where D_n is some sample metric depth function and v is a Euclidean vector. After finding the optimizer v_0, it can be mapped to $\mathcal{X}$ simply as $f(v_0)$. Since the map $v \mapsto D_n(f^{-1}(v))$ is highly non-linear, we propose using bounded *Nelder-Mead* (NMKB) and *Limited-memory BFGS* (L-BFGS-B) for the optimization. NMKB is a robust and derivative-free optimizer, which makes it generally well-suited for non-smooth functions like metric depths. However, it can be slow in high-dimensional problems. On the other hand, L-BFGS-B uses approximated gradients to find the optimal solutions and stores only the most recent steps during the optimization process, making it faster and more memory-efficient. See Byrd et al. (1995) for more details.

We employed here the implementation provided by the R-package `dfoptim` for NMKB and R's base `optim()` function for L-BFGS-B, configuring the algorithm with the following tuning parameters. The initial values were set using the top five in-sample deepest points

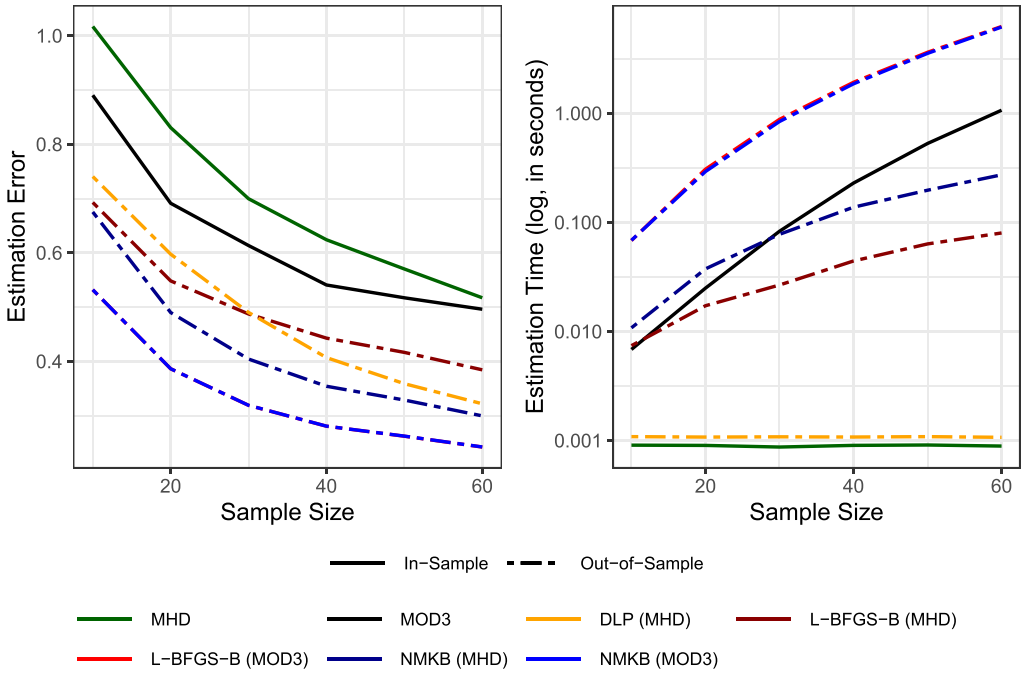

Fig. 5. Comparing the performance of different optimization methods on MOD3 and MHD. DLP refers to the out-of sample manifold optimization in the R-package `MHD`. Left: estimation error for 7 different cases. L-FBSG-B yielded exactly the same result on MOD3 as NMKB (light red line). Right: Estimation Time. The scale of the y-axis is logarithmic. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and the search space was defined with a width of 0.10 for each dimension. To further alleviate computational burden, we apply a dimension reduction via PCA to the representations in the space S to map them to an r-dimensional space with $r \ll q$. As training data for the dimension reduction, we use the images $f(X_1), \dots, f(X_n)$ of the sample. As data, we generate a sample of correlation matrices similarly as described in Section 4.1.1, meaning that $(\mathcal{X}, d)$ is now the manifold of positive definite matrices having unit diagonal. We define the map f such that $f(X_i)$ is the vector of length $p(p+1)/2$ containing the non-zero elements in the Cholesky decomposition of X_i. Since the used optimization algorithms can produce solutions lying outside of the set S (that is, solutions which are covariance matrices, but not correlation matrices), we manually scale the final optimizer $f(\upsilon_0)$ into a proper correlation matrix.

In this simulation, the matrix dimension p was set to 5, and the contamination proportion ε was fixed to 10%. The sample size, as in previous simulations in Sections 4.1.1 and 4.1.2, was taken to be $n = 10, 20, 30, 40, 50, 60$. A new parameter introduced in this simulation was the number of principal components, determined by a threshold $\texttt{TSH} \in (0, 1]$, representing the minimum proportion of total variance to be explained by the chosen principal components. Thus, the number of principal components is $r = \max(2, \{k : \sum_{i=1}^{k} \text{ExplainedVariance}_i \geq \text{TSH}\})$. In this simulation `TSH` was set to be 0.9. The whole simulation was repeated 500 times using both MHD and MOD3. As a further competitor we took the MHD-based out-of-sample estimator in the R-package `MHD` (Dai, 2024), see also Dai and Lopez-Pintado (2023), which is based on Riemannian optimization of MHD and denoted by DLP hereafter. Finally, also the in-sample versions of MHD (computed as a byproduct of DLP in the package `MHD`) and MOD3 were included.

As observed in the results shown in Fig. 5 (left), the use of the two introduced optimization algorithms (NMKB and L-BFGS-B), combined with dimension reduction, significantly reduces the estimation error compared to the in-sample estimators represented by the solid lines. We further observe that NMKB has the best performance, surpassing the other two optimization strategies, DLP and L-BFGS-B. Moreover, MOD3 exhibits lower estimation error than MHD in both out-of-sample and in-sample cases. However, as shown in the Fig. 5 (right), MOD3 involves a higher computational cost compared to MHD. Additionally, despite the substantial computational burden imposed by using the NMKB algorithm, we also observe that the MOD3 in-sample estimator tends to follow a similar trend to its optimization counterparts.

5. Real data example

In this section, we evaluate the performances of the depth functions using real data. The secondary aim of this experiment is to demonstrate how the depth functions, whose analytically complex form prevents theoretical inferential results, can still be used for statistical inference via relying on permutation and rank tests. Similar depth-based tests have also been considered earlier in the context of complex and high-dimensional data. For instance, López-Pintado and Wrobel (2017) proposed several robust two-sample permutation tests for imaging data based on multivariate volume depth. Moreover, a depth-based global two-sample envelope test

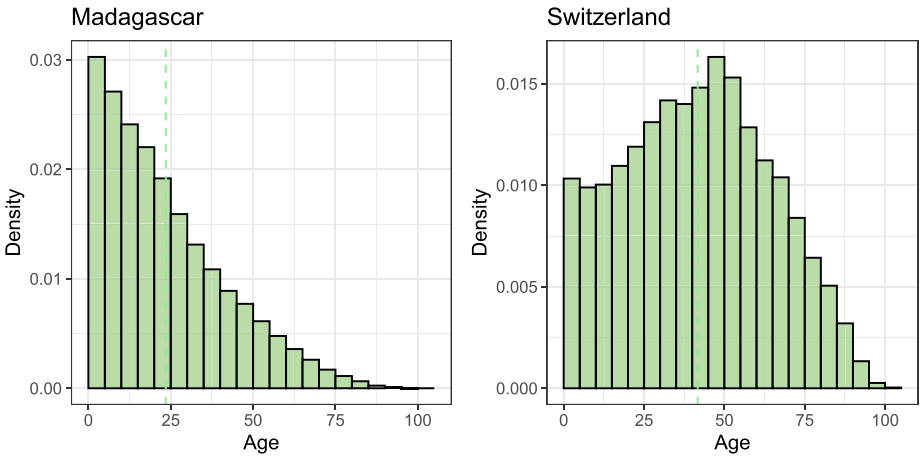

Fig. 6. The deepest age histograms among African (left) and European (right) countries, with respect to L_2-Wasserstein distance and MOD3. The dashed green lines depict the means of the two distributions.

was developed in López-Pintado and Qian (2021) for functional and image data, and Dai and Lopez-Pintado (2023) used depth-based rank testing proposed in Chenouri and Small (2012) to compare samples of functional connectivity matrices.

For this experiment, the `Age_pyramids_2014` data from R-package `HistDAWass` were chosen. This dataset contains histogram-valued data representing the age distribution of the population in each country worldwide, reported for both sexes combined as well as separately for males and females. We model the histograms as objects in the Wasserstein space $(\mathcal{X}, d)$ where d is the L_2-Wasserstein distance, implemented in the `HistDAWass`-package.

The current analysis was conducted on a subset of data that includes both sexes across all African and European countries, accounting for a total of 95 countries. Based on the geographic locations of the two continents, we hypothesize that their age distributions differ. This is visually confirmed in Fig. 6, where the deepest age distribution, w.r.t. MOD3, among African countries (left) shows a different structure from the corresponding country in Europe (right). Thus, our research question is: Is there a significant difference in the age distribution between these two groups? To study this, we next conduct a permutation test to assess whether the difference between the deepest histograms in the two continents is statistically significant. Letting $f_1, f_2 \in \mathcal{X}$ denote the deepest in-sample histograms of the two groups, we use the L_2-Wasserstein distance between them, $t = d(f_1, f_2)$, as the test statistic. To simulate the distribution of t under the null hypothesis (i.e., assuming no difference between the two groups), we randomly permute the group labels of the $n = 95$ histograms and compute the test statistic value. This reshuffling was done a total of 500 times, yielding $t_1^*, \dots, t_{500}^*$, using which the p-value of the test is computed as $\#\{t \le t_i^*\}/500$. Because the experiment produces very small p-values, we rescaled the y-axis in Fig. 7 to display $-\log_{10}(p\text{-value})$ instead. Under this transformation, taller bars correspond to smaller p-values and therefore provide stronger evidence against the null hypothesis. All depth functions indicate a difference between the groups at the 0.05 significance level; see the black bars in Fig. 7 (left).

We next take the previous result of a difference between the continents as a ground truth and continue the experiment by (a) contaminating the original data set, and (b) performing the same test for the contaminated data set and observing which of the depths still let us obtain the correct conclusion (group difference). This experiment therefore mimics the practical scenario where the data have been partially wrongly recorded. To perform the contamination, we randomly select k countries and swap their labels (a European country is labeled an African country and vice versa), using two different choices, $k = 12, 16$. To reduce randomness in the choice of the swap, we repeat the simulation 10 times, using 500 iterations in the test each time and the resulting average p-values are shown as black bars in Fig. 7 (middle and right). Under medium contamination (12 out of 95 countries mislabeled), all depths continue to yield the correct decision. Whereas, under high contamination (16 out of 95 countries mislabeled) both MHD and MOD2 identify the groups as indistinct. This is likely due to the fact that the algorithm used for MHD is approximate (Dai and Lopez-Pintado, 2023). In addition, MOD2 is not a "true" depth function and therefore does not necessarily measure the centrality of its input object; see the discussion after Theorem 7. The most consistent performance across swaps is given by MOD3 and MSD, which likely relates to the fact that they both use numerical distance information in their computation, unlike MHD and MLD which only use indicators; see Section 3.2. The results also align with the simulation studies in Section 4.1 where MSD was almost always as good as MOD3.

Moreover, following Chenouri and Small (2012), Dai and Lopez-Pintado (2023), we implemented depth-based rank tests using MLD, MHD, MSD, MOD2, and MOD3 for the same null hypothesis of no difference between the continents. The idea behind the tests is to inspect the ranks of the observation depths, which should show no group pattern under the null hypothesis. Two versions of the test are proposed in Chenouri and Small (2012): the K-statistic and the H-statistic. Of these, the H-statistic computes the depth rankings separately relative to each group and then averages these, making it more powerful than the K-statistic which instead pools the data for the rankings. For this reason, the H-statistic was selected for this study. Being depth-based, the H-statistic inherits all invariance properties of the underlying depth function, and we employed the asymptotic null distribution described in Chenouri and Small (2012).

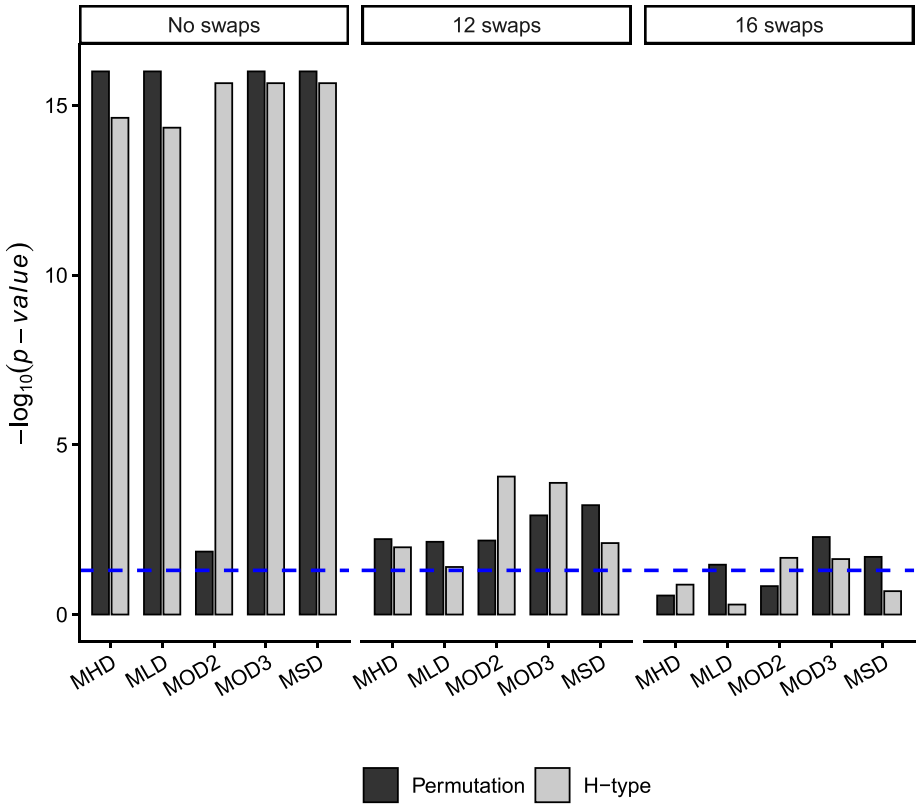

Fig. 7. p-values of depth-based permutation and rank test across all metric depth functions in the case of no contamination (left panel), 12 swaps (middle panel) and 16 swaps (right panel). Significance level of $-\log_{10}(0.05)$ is indicated by blue dashed line. All p-values smaller than 10^{-16} were recorded as 10^{-16}.

The p-values of the rank test (H-statistic) are shown in light gray in Fig. 7. As seen in the left panel, the depth-based rank test rejects the null hypothesis for the original data in all models, giving the same result as the permutation test. The corresponding numerical values are reported in Table C.1 in Appendix C. The results of the rank test (H-statistic) under the swapping experiment in Fig. 7 (right) show that only MOD2 and MOD3 manage to reject the null hypothesis under the strongest contamination (16 swaps). Overall, among the depth functions considered, MOD3 appears to be the most stable, showing strong performance under both permutation and rank-based testing framework.

6. Conclusions

Possible future research topics include: (a) Selecting and setting appropriate values for the hyper-parameters (take initial values as an example) used in the optimization algorithms in Section 4.2. (b) Properly investigating why D_{O2}, despite not being a measure of depth, manages to empirically estimate central objects on the manifold of positive definite matrices and on the unit sphere. Theorem 7 states that D_{O2} assigns large depths not only for central points but also for points x which have a high probability of $L(X_1, X_2, x)$ occurring. Intuitively, the second case arises most easily if x is a gross outlier. However, this intuition proves wrong: approximating the situation as $d^2(x, X_2) - d^2(x, X_1) =: \varepsilon$ and $d^2(X_1, X_2) =: \delta$ for some $\varepsilon, \delta > 0$ negligible compared to $d^2(x, X_1)$, one can verify that $|B_2(x, X_1, X_2)| \approx d^2(x, X_1)d^2(X_1, X_2)$, implying that outlyingness actually increases the determinant and brings the depth down. Hence, one possible reason for why MOD2 was successful in finding the true deepest points, is that "spurious" deepest points actually appear to be difficult to construct and, as such may be rare also in practice. (c) Depth functions are typically expected to be maximized at a symmetric center of a distribution. However, unlike, e.g., the metric half-space depth, D_{O3} does not have a natural companion concept of symmetry using which such a center could be defined. One possibility would be to use the quantities $A(x, X_1, X_2) := \{d^2(x, X_1) + d^2(x, X_2) - d^2(X_1, X_2)\}/\{2d(x, X_1)d(x, X_2)\}$, which satisfy $A(x, X_1, X_2) \geq -1$ with equality if and only if x lies in between of X_1 and X_2 (Virta, 2025, Theorem 1). This suggests that if there exists an object x_0 such that

$$A(x_0, X_1, X_2) <_S A(x, X_1, X_2), \quad \text{for all } x \in \mathcal{X} \setminus \{x_0\}, \tag{5}$$

where $<_S$ denotes a suitable form of strict stochastic dominance, then x_0 is a natural candidate for the most central point. Informally, (5) means that x_0 is more likely to satisfy $d(X_1, X_2) \approx d(X_1, x_0) + d(x_0, X_2)$ than any other point in $\mathcal{X}$. Accordingly, we conjecture that

if x_0 satisfies (5), then it also uniquely maximizes D_{O3}. (d) The metric half-space depth has a so-called center outward monotonicity property (Dai and Lopez-Pintado, 2023, Theorem 1(d)), which states that in certain regular enough geodesic spaces MHD is non-increasing when going away from the deepest point along geodesics. Hence, a natural continuation to point (c) above would be to study whether the analogous holds for metric Oja depth in geodesic spaces under, e.g, the condition (5).

Acknowledgments

The work of VZ was supported by the Finnish Doctoral Program Network in Artificial Intelligence, AI-DOC (decision number VN/3137/2024-OKM-6). The work of VZ and JV was supported by the Research Council of Finland (Grants 347501, 353769, 368494).

Supplementary materials

The supplementary appendices contain proofs of the technical results, additional simulation plots, and other material. The methods and simulation code (in R) are accessible at https://github.com/vidazamani/Depth-functions-for-Object-Data.

References

Barnett, V., 1976. The ordering of multivariate data. J. R. Stat. Soc. A 139 (3), 318–344. With discussion.
Bhatia, R., 2009. Positive Definite Matrices. Princeton University Press.
Burago, D., Burago, Y., Ivanov, S., 2001. A Course in Metric Geometry. Vol. 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI.
Byrd, R.H., Lu, P., Nocedal, J., Zhu, C., 1995. A limited memory algorithm for bound constrained optimization. SIAM J. Sci. Comput. 16 (5), 1190–1208.
Chenouri, S., Small, C.G., 2012. A nonparametric multivariate multisample test based on data depth. Electron. J. Stat. 6, 760–782.
Cholaquidis, A., Fraiman, R., Gamboa, F., Moreno, L., 2023. Weighted lens depth: some applications to supervised classification. Can. J. Stat. 51 (2), 652–673.
Dai, X., 2024. MHD: Metric Halfspace Depth. R package version 0.1.2. https://CRAN.R-project.org/package=MHD.
Dai, X., Lopez-Pintado, S., 2023. Tukey's depth for object data. J. Am. Stat. Assoc. 118 (543), 1760–1772.
Dubey, P., Chen, Y., Müller, H.-G., 2024. Metric statistics: exploration and inference for random objects with distance profiles. Ann. Stat. 52 (2), 757–792.
Dubey, P., Müller, H.-G., 2022. Modeling time-varying random objects and dynamic networks. J. Am. Stat. Assoc. 117 (540), 2252–2267.
Dutta, S., Ghosh, A.K., Chaudhuri, P., 2011. Some intriguing properties of Tukey's half-space depth. Bernoulli 17 (4), 1420–1434.
Eddelbuettel, D., Balamuta, J.J., 2018. Extending R with C++: a brief introduction to Rcpp. Am. Stat. 72 (1), 28–36.
Eddy, W.F., 1981. Graphics for the multivariate two-sample problem: comment. J. Am. Stat. Assoc. 76 (374), 287–289.
Fréchet, M., 1948. Les éléments aléatoires de nature quelconque dans un espace distancié. Annales de l'Institut Henri Poincaré 10, 215–310.
Geenens, G., Nieto-Reyes, A., Francisci, G., 2023. Statistical depth in abstract metric spaces. Stat. Comput. 33 (2), 46.
Gromov, M., 2007. Metric Structures for Riemannian and Non-Riemannian Spaces. Birkhäuser, Boston.
Kleindessner, M., Von Luxburg, U., 2017. Lens depth function and k-relative neighborhood graph: versatile tools for ordinal data analysis. J. Mach. Learn. Res. 18 (58), 1–52.
Koshevoy, G., Mosler, K., 1997. Zonoid trimming for multivariate distributions. Ann. Stat. 25 (5), 1998–2017.
Lee, A.J., 2019. U-Statistics: Theory and Practice. Routledge.
Liu, R.Y., 1990. On a notion of data depth based on random simplices. Ann. Stat. 18 (1), 405–414.
Liu, R.Y., 1992. Data depth and multivariate rank tests. In: Dodge, Y. (Ed.), L1-Statistics Analysis and Related Methods. North-Holland, Amsterdam, pp. 279–294.
Liu, Z., Modarres, R., 2011. Lens data depth and median. J. Nonparametric Stat. 23 (4), 1063–1077.
López-Pintado, S., Qian, K., 2021. A depth-based global envelope test for comparing two groups of functions with applications to biomedical data. Stat. Med. 40 (7), 1639–1652.
López-Pintado, S., Wrobel, J., 2017. Robust nonparametric tests for imaging data based on data depth. Stat 6 (1), 405–419.
Mahalanobis, P.C., 1936. On the generalized distance in statistics. Proc. Natl. Inst. Sci. 2 (1), 49–55.
Maronna, R.A., Martin, R.D., Yohai, V.J., Salibián-Barrera, M., 2019. Robust Statistics: Theory and Methods (with R). John Wiley & Sons.
Mosler, K., Mozharovskyi, P., 2022. Choosing among notions of multivariate depth statistics. Stat. Sci. 37 (3), 348–368.
Mozharovskyi, P., 2022. Data Depth: Computation, Applications, and Beyond. Habilitation thesis. Institut Polytechnique de Paris.
Mozharovskyi, P., Valla, R., 2025. Anomaly detection using data depth: multivariate case. Int. J. Data Sci. Anal. 20, 5171–5196. https://doi.org/10.1007/s41060-025-00784-1
Oja, H., 1983. Descriptive statistics for multivariate distributions. Stat. Probab. Lett. 1 (6), 327–332.
Serfling, R., 2002. A depth function and a scale curve based on spatial quantiles. In: Dodge, Y. (Ed.), Statistical Data Analysis Based on the L1-Norm and Related Methods. Birkhäuser, Basel. Statistics for Industry and Technology, pp. 25–38.
Simeon, G., Piella, G., Camara, O., Pareto, D., 2022. Riemannian geometry of functional connectivity matrices for multi-site attention-deficit/hyperactivity disorder data harmonization. Front. Neuroinf. 16, 769274.
Stewart, G.W., 1980. The efficient generation of random orthogonal matrices with an application to condition estimators. SIAM J. Numer. Anal. 17, 403–409.
Tukey, J.W., 1975. Mathematics and the picturing of data. In: James, R. (Ed.), Proceedings of the International Congress of Mathematicians. Canadian Mathematical Congress, pp. 523–531.
Van der Vaart, A.W., 2000. Asymptotic Statistics. Vol. 3. Cambridge University Press.
Vardi, Y., Zhang, C.-H., 2000. The multivariate l_1-median and associated data depth. Proc. Natl. Acad. Sci. 97 (4), 1423–1426.
Virta, J., 2025. Measure of shape for object data. J. Nonparametric Stat. 38 (2), 675–695.
Virta, J., 2026. Spatial depth for data in metric spaces. Scand. J. Stat. 53 (2), 684–711. https://doi.org/10.1111/sjos.70054
Virta, J., Lee, K.-Y., Li, L., 2022. Sliced inverse regression in metric spaces. Stat. Sin. 32, 2315–2337.
Yang, M., Modarres, R., 2018. β-Skeleton depth functions and medians. Commun. Stat. - Theory Methods 47 (20), 5127–5143.
Zhang, Q., Xue, L., Li, B., 2024. Dimension reduction for Fréchet regression. J. Am. Stat. Assoc. 119, 1–15.
Zhu, C., Müller, H.-G., 2025. Geodesic optimal transport regression. Biometrika. 113 (1), p.asaf086.
Zuo, Y., Serfling, R., 2000. General notions of statistical depth function. Ann. Stat. 28 (2), 461–482.