

DISCUSSION PAPER

2017/19

Data depth and rank-based tests for covariance and
spectral density matrices

Chau, J., Ombao, H. and R. von Sachs

Data depth and rank-based tests for covariance and spectral density matrices

Joris Chau*, Hernando Ombao[†] and Rainer von Sachs[‡]

Abstract

In multivariate time series analysis, objects of primary interest to study cross-dependences in the time series are the autocovariance or spectral density matrices. Non-degenerate covariance and spectral density matrices are necessarily Hermitian and positive definite, and our primary goal is to develop new methods to analyze samples of such matrices. The main contribution of this paper is the generalization of the concept of statistical data depth for collections of covariance or spectral density matrices by exploiting the geometric properties of the space of Hermitian positive definite matrices as a Riemannian manifold. This allows one to naturally characterize most central or outlying matrices, but also provides a practical framework for rank-based hypothesis testing in the context of samples of covariance or spectral density matrices. First, the desired properties of a data depth function acting on the space of Hermitian positive definite matrices are presented. Second, we propose two computationally efficient pointwise and integrated data depth functions that satisfy each of these requirements. Several applications of the developed methodology are illustrated by the analysis of collections of spectral matrices in multivariate brain signal time series datasets.

Keywords: Data depth, Hermitian positive definite matrices, Riemannian manifold, Spectral density matrix, Rank-based hypothesis tests, Multivariate time series.

1 Introduction

In multivariate time series analysis, we study covariance and correlation structures in a time series by means of its autocovariance matrices in the time domain, or its spectral density matrix in the frequency domain. In this work, our goal is to develop data exploration and inference tools for random collections or samples of such matrices. The objects of interest, i.e. non-degenerate autocovariance or spectral density matrices, are necessarily elements of the space of Hermitian positive definite (PD) matrices. The space of Hermitian PD matrices, although very well-structured, is inherently non-Euclidean and it is generally difficult to apply standard Euclidean-based statistical procedures

*Corresponding author, j.chau@uclouvain.be, Institute of Statistics, Biostatistics, and Actuarial Sciences, Université catholique de Louvain, Voie du Roman Pays 20, B-1348, Louvain-la-Neuve, Belgium.

[†]Department of Statistics, University of California at Irvine, Bren Hall 2206, Irvine, CA, 92697, United States. Department of Applied Mathematics and Computational Science, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia.

[‡]Institute of Statistics, Biostatistics, and Actuarial Sciences, Université catholique de Louvain, Voie du Roman Pays 20, B-1348, Louvain-la-Neuve, Belgium.

(e.g. regression, clustering or inference procedures), as common matrix operations, such as matrix subtraction or extrapolation, may be unstable or could break down. Recent works that generalize statistical procedures for data in the space of symmetric or Hermitian PD matrices, taking into account the non-Euclidean geometry of the space, include Smith (2000), Pennec et al. (2006), Fletcher et al. (2009), Fletcher et al. (2011) and Said et al. (2015) among others. In the context of multivariate spectral analysis, Holbrook et al. (2016) recently considered Bayesian spectral matrix estimation and inference based on geodesic Lagrangian Monte Carlo methods, and Chau and von Sachs (2017) considered wavelet-based spectral estimation and clustering of spectral matrices on the Riemannian manifold of Hermitian PD matrices. The main contribution of this work is the generalization of a notion of statistical data depth for samples of Hermitian PD matrices, as a means of data exploration and inference for samples of covariance or spectral density matrices. In particular, in the context of spectral analysis in biomedical or neuroscientific applications –where we typically record (large amounts of) data for multiple subjects or trials in an experiment– this new data depth concept can be used as an exploratory data tool to identify most central or outlying spectral matrices (as in Liu et al. (1999) or Sun and Genton (2012) in a Euclidean context); or as a means of inference, by way of rank-based hypothesis testing, replacing the ordinary ranks by the depth-induced ranks (as in Liu and Singh (1993), Chenouri and Small (2012), or (Mosler, 2002, Chapter 5) for samples of Euclidean vectors). The latter provides, for instance, a practical framework to test for significant differences in spectral matrix behavior between different subject groups or treatment conditions in an experiment.

Although many different data depths have been proposed and studied in the literature over the years, most data depth functions are constructed in the first place for vector observations in the Euclidean space $\mathbb{R}^d$. One exception is Liu and Singh (1992), where the authors consider depth functions for directional data observations taking values on a circle or a sphere. For an overview of different Euclidean data depth functions and their properties, we refer to e.g. Liu et al. (1999), Zuo and Serfling (2000), or Mosler (2002). The space of $(d \times d)$ -dimensional Hermitian (not necessarily PD) matrices $(\mathbb{H}_{d \times d}, +, \cdot_S)$ together with matrix addition and matrix scalar multiplication is a real vector space, and each Hermitian matrix bijectively maps to a vector in $\mathbb{R}^{d^2}$ by taking its real-valued component vector with respect to a basis of the space of Hermitian matrices. Data depth values in a sample of Hermitian matrices can then be calculated using any ordinary Euclidean data depth function applied to the real basis component vectors of the Hermitian matrices, as long as the computed depth values do not depend on the chosen basis. In contrast, the space of Hermitian PD matrices $(\mathbb{P}_{d \times d}, +, \cdot_S)$ is not a vector space, due to the nonlinear positive definite constraints. Moreover, embedding the space of Hermitian PD matrices as a cone in a Euclidean space endowed with the Euclidean distance results in an incomplete metric space. As a consequence, computation of data depth values by means of standard Euclidean depth functions is not so straightforward. To illustrate, according to Zuo and Serfling (2000), a proper depth function should be monotone

non-increasing moving outwards from a well-defined center. However, moving away from a central point along a straight line is not always well-defined in the cone of Hermitian PD matrices, as the boundary of the space lies at a finite distance. Instead of embedding the space of Hermitian PD matrices in a Euclidean space, a more natural approach is to consider the space as a Riemannian manifold (see Pennec et al. (2006), Fletcher et al. (2011) and Chau and von Sachs (2017)). In particular, the space of Hermitian PD matrices endowed with a specific invariant Riemannian distance function becomes a *complete* metric space and each Hermitian PD matrix lies at an infinite distance to the boundary of the space consisting of all singular matrices.

In the preliminary Section 2, we introduce the necessary geometric tools and notions to develop proper *manifold* data depths acting directly on the space of Hermitian PD matrices as a geodesically complete manifold. In Section 3, we present the desired properties a manifold depth function should satisfy, and we propose two computationally efficient manifold depth functions that satisfy each of these properties. In addition, we consider integrated depth functions that act on curves of Hermitian PD matrices, such as spectral density matrices. In Section 4, we examine several applications of the introduced depth functions in the context of rank-based hypothesis testing, which is applied in Section 5 to analyze collections of spectral matrices in two brain signal time series datasets of multivariate local field potential (LFP) recordings. The technical proofs and simulated power analyses of the rank-based tests, with comparisons to competing tests in the literature, can be found in the supplementary material. The accompanying R-code, containing all the necessary tools to compute the manifold data depths and to perform rank-based hypothesis testing on the geodesic manifold, is publicly available in the R-package `pdSpecEst` on CRAN, (Chau (2017)).

2 Preliminaries

2.1 Geometry of HPD matrices

In order to construct a notion of data depth for observations in the space of HPD (Hermitian positive definite) matrices, we consider this space as a Riemannian manifold, see e.g. Pennec et al. (2006), (Bhatia, 2009, Chapter 6), or Smith (2000) for more details. Let us denote $\mathcal{M} := \mathbb{P}_{d \times d}$ for the space of $(d \times d)$ -dimensional HPD matrices, a well-studied d^2 -dimensional differentiable Riemannian manifold. The tangent space $\mathcal{T}_p(\mathcal{M})$ at a point, i.e. a matrix, $p \in \mathcal{M}$ can be identified by the real vector space $\mathcal{H} := \mathbb{H}_{d \times d}$ of $(d \times d)$ -dimensional Hermitian matrices, and the Frobenius inner product on $\mathcal{H}$ induces the natural (also invariant, Fisher information or Fisher-Rao) Riemannian metric on the manifold $\mathcal{M}$ given by the smooth family of inner products:

$$\langle h_1, h_2 \rangle_p = \text{Tr}((p^{-1/2} * h_1)(p^{-1/2} * h_2)), \quad \forall p \in \mathcal{M} \quad (2.1)$$

with $h_1, h_2 \in \mathcal{T}_p(\mathcal{M})$. Here and throughout this paper, $y^{1/2}$ always denotes the Hermitian square root matrix of $y \in \mathcal{M}$, and we write $y * x := y^* x y$ for matrix congruence transformation, where

$*$ denotes the conjugate transpose of a matrix. The natural Riemannian distance on $\mathcal{M}$ derived from the Riemannian metric above is given by:

$$\delta(p_1, p_2) = \|\text{Log}(p_1^{-1/2} * p_2)\|_F \quad (2.2)$$

where $\|\cdot\|_F$ denotes the matrix Frobenius norm and $\text{Log}(\cdot)$ is the (ordinary) matrix logarithm. By (Bhatia, 2009, Theorem 6.1.6 and Prop. 6.2.2), the shortest curve with respect to the Riemannian distance function joining any two points $p_1, p_2 \in \mathcal{M}$, i.e. the *geodesic* segment, is unique and $(\mathcal{M}, \delta)$ is a complete separable metric space, which by the Hopf-Rinow Theorem implies that every geodesic curve can be extended indefinitely. The mapping $x \mapsto a * x$ is an isometry for each invertible matrix $a \in \text{GL}(d, \mathbb{C})$, i.e. it is distance-preserving, in the sense that:

$$\delta(p_1, p_2) = \delta(a * p_1, a * p_2), \quad \forall a \in \text{GL}(d, \mathbb{C})$$

In order to characterize convex sets and measures of centrality in the metric space $(\mathcal{M}, \delta)$, which play an important role in the construction of the manifold data depths, we also introduce the notions of exponential and logarithmic maps, relating the manifold $\mathcal{M}$ to its tangent spaces $\mathcal{T}_p(\mathcal{M})$.

First, by (Pennec et al. (2006)) the exponential maps $\text{Exp}_p : \mathcal{T}_p(\mathcal{M}) \rightarrow \mathcal{M}$ are (locally) diffeomorphic maps from the tangent space at a point $p \in \mathcal{M}$ to the manifold given by,

$$\text{Exp}_p(h) = p^{1/2} * \text{Exp}\left(p^{-1/2} * h\right), \quad \forall h \in \mathcal{T}_p(\mathcal{M})$$

where $\text{Exp}(\cdot)$ denotes the ordinary matrix exponential. In particular, $\gamma(t) = \text{Exp}_p(th)$ for $t \geq 0$ parametrizes a geodesic curve emanating from p with unit tangent vector h , such that $\gamma(0) = p$. Since $\mathcal{M}$ is a geodesically complete manifold and minimizing geodesics are always unique, it follows by (do Carmo, 1992, Chapter 13) that for each $p \in \mathcal{M}$ the image of the exponential map Exp_p is the entire manifold $\mathcal{M}$, and the exponential maps are in fact *global* diffeomorphisms. In the other direction, the logarithmic maps $\text{Log}_p : \mathcal{M} \rightarrow \mathcal{T}_p(\mathcal{M})$ are global diffeomorphisms from the manifold to the tangent space at the point $p \in \mathcal{M}$, given by the inverse exponential maps:

$$\text{Log}_p(\tilde{p}) = p^{1/2} * \text{Log}\left(p^{-1/2} * \tilde{p}\right)$$

The Riemannian distance function can now also be expressed in terms of the logarithmic map as:

$$\delta(p_1, p_2) = \|\text{Log}_{p_1}(p_2)\|_{p_1} = \|\text{Log}_{p_2}(p_1)\|_{p_2}, \quad \forall p_1, p_2 \in \mathcal{M} \quad (2.3)$$

where throughout this work $\|h\|_p := \langle h, h \rangle_p$ is the norm of $h \in \mathcal{T}_p(\mathcal{M})$ induced by the Riemannian metric.

As there exist unique geodesic curves connecting any two points $p_1, p_2 \in \mathcal{M}$, geodesically convex sets are well-defined. We say that a subset $\mathcal{K} \subseteq \mathcal{M}$ is convex or geodesically convex if for each pair

of points $p_1, p_2 \in \mathcal{K}$, the geodesic segment $[p_1, p_2]$ lies entirely in $\mathcal{K}$. If $\mathcal{S} \subseteq \mathcal{M}$, then the convex hull of $\mathcal{S}$, denoted by $\text{conv}(\mathcal{S})$, is the smallest convex set containing $\mathcal{S}$. This set is conveniently expressed as,

$$\text{conv}(\mathcal{S}) := \left\{ p \in \mathcal{S} : p = \text{Exp}_p \left(\int_{\mathcal{S}} \text{Log}_p(x) g(x) \lambda(dx) \right), g : \mathcal{S} \rightarrow [0, 1], \int_{\mathcal{S}} g(x) \lambda(dx) = 1 \right\}$$

where λ is the Lebesgue measure on the finite-dimensional metric space $(\mathcal{M}, \delta)$ and g is a measurable function. For more details on the construction of (approximate) convex hulls on the manifold $\mathcal{M}$, we refer to Fletcher et al. (2011).

2.2 Probability distributions and random variables

Let a random variable on the manifold of HPD matrices $X : \Omega \rightarrow \mathcal{M}$ be a measurable function from some probability space $(\Omega, \mathcal{A}, \nu)$ to the measurable space $(\mathcal{M}, \mathcal{B}(\mathcal{M}))$, where $\mathcal{B}(\mathcal{M})$ is the Borel algebra, i.e. the smallest σ -algebra containing all open sets in the complete separable metric space $(\mathcal{M}, \delta)$. In the following, we always work directly with the induced probability on $\mathcal{M}$, $\nu(B) = \nu(\{\omega \in \Omega : X(\omega) \in B\})$. By $P(\mathcal{M})$, we denote the set of all probability measures on $(\mathcal{M}, \mathcal{B}(\mathcal{M}))$ and $P_p(\mathcal{M})$ denotes the subset of probability measures in $P(\mathcal{M})$ that have finite moments of order p with respect to the Riemannian distance, i.e. the L^p -Wasserstein space, (see e.g. (Villani, 2009, Definition 6.4)). That is,

$$P_p(\mathcal{M}) := \left\{ \nu \in P(\mathcal{M}) : \exists y_0 \in \mathcal{M}, \text{ s.t. } \int_{\mathcal{M}} \delta(y_0, x)^p \nu(dx) < \infty \right\}$$

Note that if $\int_{\mathcal{M}} \delta(y_0, x)^p \nu(dx) < \infty$ for some $y_0 \in \mathcal{M}$ and $1 \leq p < \infty$, this is true for any $y \in \mathcal{M}$. This follows by the triangle inequality and the fact that $\delta(p_1, p_2) < \infty$ for any $p_1, p_2 \in \mathcal{M}$, as $\int_{\mathcal{M}} \delta(y, x)^p \nu(dx) \leq 2^p (\delta(y, y_0)^p + \int_{\mathcal{M}} \delta(y_0, x)^p \nu(dx)) < \infty$. For a sequence of probability measures $(\nu_n)_{n \in \mathbb{N}}$ in $P(\mathcal{M})$, $\nu_n \xrightarrow{w} \nu$ denotes weak convergence to the probability measure ν in the usual sense, i.e. $\int_{\mathcal{M}} \phi(x) \nu_n(dx) \rightarrow \int_{\mathcal{M}} \phi(x) \nu(dx)$ for every continuous and bounded function $\phi : \mathcal{M} \rightarrow \mathbb{R}$, and a sequence $(\nu_n)_{n \in \mathbb{N}}$ is said to be uniformly integrable if $\lim_{K \rightarrow \infty} \sup_{n \in \mathbb{N}} \int_{\mathcal{M}} \delta(y_0, x) \mathbf{1}_{\{\delta(y_0, x) > K\}} \nu_n(dx) = 0$ for some $y_0 \in \mathcal{M}$. Note that if $(\nu_n)_{n \in \mathbb{N}}$ is uniformly integrable for some $y_0 \in \mathcal{M}$, then the sequence is uniformly integrable for any $y \in \mathcal{M}$. Finally, we use the notation $\text{conv}(\nu) := \text{conv}(\text{supp}(\nu))$ for the convex hull of the support of the measure ν on $\mathcal{M}$, and $\text{rint}(\text{conv}(\nu))$ and $\text{rd}(\text{conv}(\nu))$ for its relative interior and relative boundary.

2.3 Geometric mean

To characterize the center of a random variable X with probability measure ν , one important measure of centrality is the *geometric* (also Karcher or Fréchet) mean, see e.g. Pennec (2006), which turns out to be the point of maximum depth in the manifold zonoid depth introduced in

Section 3.2. The set of geometric means is given by the points that minimize the second moment with respect to the Riemannian distance,

$$\mu = \mathbb{E}_\nu[X] := \arg \min_{y \in \text{supp}(\nu)} \int_{\mathcal{M}} \delta(y, x)^2 \nu(dx)$$

If $\nu \in P_2(\mathcal{M})$, then at least one geometric mean will exist as the above expectation is finite for $y \in \mathcal{M}$. Moreover, since the manifold $\mathcal{M}$ is a space of non-positive curvature with no cut-locus (see Pennec et al. (2006) or Skovgaard (1984)), by (Le, 1995, Proposition 1) the geometric mean μ is unique for any distribution $\nu \in P_2(\mathcal{M})$. Recall that the cut-locus at a point $p \in \mathcal{M}$ is the complement of the image of the exponential map Exp_p , which is the empty set for each $p \in \mathcal{M}$ as the image of Exp_p is the entire manifold $\mathcal{M}$. By (Pennec, 2006, Corollary 1), the geometric mean can also be represented by the point $\mu \in \mathcal{M}$ that satisfies,

$$\mathbf{E}_\nu[\text{Log}_\mu(X)] = \mathbf{0}_{d \times d} \quad (2.4)$$

where $\mathbf{0}_{d \times d}$ is the zero matrix. The sample geometric mean of a set of manifold-valued observations minimizes a sum of squared Riemannian distances and can be computed efficiently through a gradient descent algorithm as in Pennec (2006).

2.4 Geometric median

A second measure of centrality of primary interest is the geometric median as in e.g. Fletcher et al. (2009), which is the point of maximum depth in the geodesic distance depth defined in Section 3.4. In contrast to the geometric mean, the geometric median minimizes an expected value of non-squared distances, and the set of geometric medians of a random variable X with probability measure ν is given by:

$$m = \text{GM}_\nu(X) := \arg \min_{y \in \text{supp}(\nu)} \int_{\mathcal{M}} \delta(y, x) d\nu(x)$$

It can be shown that on $\mathcal{M}$, a space with non-positive curvature and no cut-locus, the geometric median exists and is also unique for any distribution $\nu \in P_1(\mathcal{M})$. The proof runs along the same lines as that of (Fletcher et al., 2009, Theorem 1), combined with an application of Leibniz's integral rule. Furthermore, similar to the characterization of the geometric mean in eq.(2.4), the geometric median is uniquely characterized by the point $m \in \mathcal{M}$ that satisfies,

$$\mathbf{E}_\nu \left[\frac{\text{Log}_m(X)}{\delta(m, X)} \right] = \mathbf{0}_{d \times d} \quad (2.5)$$

Remark If the distribution ν of a random variable X is *centrally symmetric* around $\mu \in \mathcal{M}$ in the sense that $\text{Log}_\mu(X) \stackrel{d}{=} -\text{Log}_\mu(X)$, then the geometric mean and median coincide and are equal to μ . Here, equality in distribution ($\stackrel{d}{=}$) is read as equality in terms of the joint distribution of all matrix components. The claim for the geometric mean follows by observing that $\mathbf{E}_\nu[\text{Log}_\mu(X)] = \mathbf{0}_{d \times d}$,

which implies that μ is the unique geometric mean of the random variable X . For the geometric median, if X is centrally symmetric around μ , then X is also *angularly symmetric* around μ in the sense that $\text{Log}_\mu(X)/\|\text{Log}_\mu(X)\|_\mu \stackrel{d}{=} -\text{Log}_\mu(X)/\|\text{Log}_\mu(X)\|_\mu$, writing $\|\text{Log}_\mu(X)\|_\mu = \langle \text{Log}_\mu(X), \text{Log}_\mu(X) \rangle_\mu$. Since $\|\text{Log}_\mu(X)\|_\mu = \delta(\mu, X)$ by eq.(2.3), $\mathbf{E}_\nu[\text{Log}_\mu(X)/\delta(\mu, X)] = \mathbf{0}_{d \times d}$, implying that μ is also the unique geometric median of the random variable X .

3 Data depth for HPD matrices

Before introducing the data depth measures, we first present the desired properties a proper *manifold* data depth function acting directly on the space of $(d \times d)$ -dimensional HPD matrices should satisfy. These properties can be viewed as the natural manifold generalizations of the properties in Zuo and Serfling (2000) for depth functions acting on vectors in the Euclidean space $\mathbb{R}^d$. With in mind the application to spectral matrices, we also consider integrated analogs of the depth properties for depth functions acting on curves of HPD matrices $y(t) \in \mathcal{M}$ with $t \in \mathcal{I} \subset \mathbb{R}$.

3.1 Depth properties

Below, we denote $D(\nu, y)$ for the depth of a point $y \in \mathcal{M}$ with respect to a distribution $\nu \in P(\mathcal{M})$; or $\text{iD}(\nu, y)$ for the integrated depth of a curve $y := (y(t))_{t \in \mathcal{I}}$ with respect to a curve of marginal measures $\nu := (\nu(t))_{t \in \mathcal{I}}$, such that $\nu(t) \in P(\mathcal{M})$ for each $t \in \mathcal{I}$. If a nonnegative bounded function $D(\cdot, \cdot)$ or $\text{iD}(\cdot, \cdot)$ satisfies the pointwise (resp. integrated) properties **P.1** to **P.4**, we say that it is a proper data depth function on the manifold $\mathcal{M}$.

P.1 (*Congruence invariance*) The depth function should be invariant under matrix congruence transformation of the form $x \mapsto a * x$, with $a \in \text{GL}(d, \mathbb{C})$. That is, for each $a \in \text{GL}(d, \mathbb{C})$,

$$D(\nu, y) = D(\nu_a, a * y), \quad \forall y \in \mathcal{M} \quad (3.1)$$

where ν_a is the distribution of the transformed random variable $a * X$, such that X is distributed according to ν . Generalizing this property for an integrated depth function $\text{iD}(\nu, y)$, we require that the same property holds pointwise for each $t \in \mathcal{I}$. In this case, $a := (a(t))_{t \in \mathcal{I}}$ is a curve of invertible matrices, with $a(t) \in \text{GL}(d, \mathbb{C})$ for each $t \in \mathcal{I}$.

To heuristically motivate this property, for a depth function acting on the Euclidean space $\mathbb{R}^d$, it is desirable that it is affine invariant $D(\nu, y) = D(\nu_{a,b}, ay + b)$ for each $y \in \mathbb{R}^d$, where $\nu_{a,b}$ is the distribution of the random vector $aX + b$, with $a \in \text{GL}(d, \mathbb{R})$, $b \in \mathbb{R}^d$ and X distributed according to ν . In the context of covariance and spectral density matrices, we are concerned with second-order behavior of random variables. For a random vector X with covariance matrix Σ , the covariance matrix of the affine transformation $aX + b$ is $a * \Sigma = a \Sigma a^T$. The natural equivalent of the affine invariance for depth functions acting on symmetric or HPD matrices, such as covariance or spectral

matrices, then becomes the congruence invariance in eq.(3.1). Another way to put this is that a depth function acting on the covariance or spectral matrix of a data vector X should be invariant under a change in the coordinate system of the data space of X . In particular, in spectral analysis of multivariate time series, this property ensures that a reordering or rescaling of the components of the multivariate time series data does not change the depth.

P.2 (*Maximality at center*) The depth function should attain its maximum value, i.e. deepest point or point with maximum depth, at a well-defined unique *center* of the distribution, such as the geometric mean or median, which are characterized as points of central and angular symmetry respectively. Let $\mu \in \mathcal{M}$ be a unique central point of the distribution ν , then,

$$D(\nu, \mu) = \sup_{y \in \mathcal{M}} D(\nu, y)$$

Similarly, for an integrated depth function, the maximum value should be attained at a well-defined unique central curve $\mu(t)$ with $t \in \mathcal{I}$, such as the curve of geometric means or medians.

P.3 (*Monotonicity relative to center*) As $y \in \mathcal{M}$ moves away from the deepest point μ along a geodesic curve emanating from μ , the depth of the point y with respect to the distribution ν should be monotone non-increasing. Let $\text{Exp}_\mu(th)$, $t \geq 0$, be the geodesic emanating from μ with unit tangent vector h . Then,

$$D(\nu, \text{Exp}_\mu(t_1 h)) \geq D(\nu, \text{Exp}_\mu(t_2 h)), \quad \forall 0 \leq t_1 \leq t_2$$

For an integrated depth function, let $s_1(t), s_2(t)$ be real-valued curves over $\mathcal{I}$, such that $0 \leq s_1(t) \leq s_2(t)$ for each $t \in \mathcal{I}$. Denote $y_1 := (\text{Exp}_{\mu(t)}(s_1(t)h(t)))_{t \in \mathcal{I}}$ and $y_2 := (\text{Exp}_{\mu(t)}(s_2(t)h(t)))_{t \in \mathcal{I}}$, where $h(t) \in \mathcal{T}_{\mu(t)}(\mathcal{M})$ is a curve of unit tangent vectors. Then, for each such curves $s_1(t), s_2(t)$,

$$\text{iD}(\nu, y_1) \geq \text{iD}(\nu, y_2)$$

P.4 (*Vanishing at infinity*) The depth of a point $y \in \mathcal{M}$ should approach zero as the point y converges to a singular matrix, i.e. a matrix with zero or infinite eigenvalues,

$$\lim_{M \rightarrow \infty} \sup_{\|\text{Log}(y)\|_F \geq M} D(\nu, y) = 0$$

Similarly, for an integrated depth function, if the curve $y(t)$ converges to a curve of singular matrices for each $t \in \mathcal{I}$, then the integrated depth should approach zero.

Below, we give two additional continuity properties, which although not strictly required are nonetheless useful to derive asymptotic results in subsequent applications, such as rank-based hypothesis testing as in Section 4 or the construction of asymptotic depth-based confidence sets in the same spirit as Yeh and Singh (1997).

(P.5) (Continuity in y) Let $(y_n)_{n \in \mathbb{N}}$ be a convergent sequence with $y_n \in \mathcal{M}$ for each $n \in \mathbb{N}$, such that $y_n \rightarrow y$ in the sense that $\delta(y_n, y) \rightarrow 0$. Then the depth function is continuous in y as,

$$\lim_{n \rightarrow \infty} D(\nu, y_n) = D(\nu, y)$$

(P.6) (Uniform continuity in ν) The depth function is uniformly continuous in terms of the probability measure ν in the sense that if $(\nu_n)_{n \in \mathbb{N}}$ is a uniformly integrable sequence of probability measures, such that $\nu_n \xrightarrow{w} \nu$. Then,

$$\sup_{y \in \mathcal{M}} |D(\nu_n, y) - D(\nu, y)| \rightarrow 0, \quad \text{as } n \rightarrow \infty$$

3.2 Manifold zonoid depth

As geodesic convex hulls are well-defined on the geodesically complete manifold $\mathcal{M}$, there exist natural manifold generalizations of the notions of simplicial depth or convex hull peeling depth (Liu et al. (1999)) for multivariate observations in a Euclidean space. However, the simplicial depth requires the computation of possibly many (approximate) convex hulls, which quickly becomes computationally infeasible, especially for higher-dimensional matrices. Instead, we propose a straightforward manifold generalization of another depth measure based on trimmed convex depth regions, the zonoid depth (e.g. Mosler (2002)). The manifold zonoid depth can be computed using the same tools as the ordinary zonoid depth for Euclidean vectors and its computation remains efficient, also for higher-dimensional HPD matrices.

In a Euclidean context, let ζ be a probability measure on $(\mathbb{R}^d, \mathcal{B}^d)$ with finite first moment, then the *Euclidean zonoid α -trimmed region*, with $0 < \alpha \leq 1$ is defined as the set,

$$D_\alpha(\zeta) := \left\{ \int_{\mathbb{R}^d} xg(x) d\zeta(x) : g : \mathbb{R}^d \rightarrow [0, 1/\alpha] \text{ measurable, s.t. } \int_{\mathbb{R}^d} g(x) d\zeta(x) = 1 \right\}$$

If $\alpha = 0$, we set $D_0(\zeta) = \mathbb{R}^d$. By (Mosler, 2002, Chapter 3), $D_\alpha(\zeta)$ is convex and monotone decreasing in α , creating a nested sequence of convex sets for decreasing values $\alpha_1 \geq \dots \geq \alpha_n$. If $\alpha = 1$, $D_\alpha(\zeta)$ consists of the single point $\mathbf{E}_\zeta[X]$, the Euclidean mean of the distribution ζ . The Euclidean zonoid depth of a point $y \in \mathbb{R}^d$ with respect to a distribution ζ is characterized by the smallest α -trimmed region still containing y ,

$$\text{ZD}_{\mathbb{R}^d}(\zeta, y) := \sup \{ \alpha : y \in D_\alpha(\zeta) \}$$

The zonoid data depth is extended to the manifold as follows.

Definition 3.1. (*Manifold zonoid depth*) Let $\nu \in P_2(\mathcal{M})$ and let ζ_y be the probability measure on $(\mathbb{R}^{d^2}, \mathcal{B}(\mathbb{R}^{d^2}))$ of the random variable $\text{Log}_y(X) \in \mathcal{T}_y(\mathcal{M}) \cong \mathbb{R}^{d^2}$ as a d^2 -dimensional random real

basis component vector, where X has probability measure ν . The manifold zonoid depth of a point $y \in \mathcal{M}$ with respect to the distribution ν is defined as:

$$\text{ZD}_{\mathcal{M}}(\nu, y) := \sup \left\{ \alpha : \vec{0} \in D_{\alpha}(\zeta_y) \right\} \quad (3.2)$$

where $\vec{0}$ is a d^2 -dimensional zero vector, and $D_{\alpha}(\zeta_y)$ is the Euclidean zonoid α -trimmed region of the distribution ζ_y on $(\mathbb{R}^{d^2}, \mathcal{B}(\mathbb{R}^{d^2}))$. Equivalently, the manifold zonoid depth can be written as,

$$\text{ZD}_{\mathcal{M}}(\nu, y) = \sup \left\{ \alpha : y \in D_{\alpha}^{\mathcal{M}}(\nu) \right\}$$

where $D_{\alpha}^{\mathcal{M}}(\nu)$ is the *manifold zonoid α -trimmed region* defined as,

$$D_{\alpha}^{\mathcal{M}}(\nu) = \left\{ y \in \mathcal{M} : y = \text{Exp}_y \left(\int_{\mathcal{M}} \text{Log}_y(x) g(x) \nu(dx) \right), g : \mathcal{M} \rightarrow [0, 1/\alpha], \int_{\mathcal{M}} g(x) \nu(dx) = 1 \right\}$$

with g a measurable function.

Remark Computation of the manifold zonoid depth is straightforward via the definition, since $\text{ZD}_{\mathcal{M}}(\nu, y) = \text{ZD}_{\mathbb{R}^{d^2}}(\zeta_y, 0)$, which can be calculated directly by the Euclidean zonoid depth as in (Mosler, 2002, Chapter 4). Note that if $(e_1, \dots, e_{d^2})$ is an orthonormal basis of the Euclidean space $(\mathcal{H}, \langle \cdot, \cdot \rangle_F)$, then an orthonormal basis of $(\mathcal{T}_y(\mathcal{M}), \langle \cdot, \cdot \rangle_y)$ is simply $(y^{1/2} * e_1, \dots, y^{1/2} * e_{d^2})$. In fact, the basis components of $\text{Log}_y(x) \in \mathcal{T}_y(\mathcal{M})$ can be computed directly using only an orthonormal basis of $(\mathcal{H}, \langle \cdot, \cdot \rangle_F)$, since $\langle \text{Log}_y(x), y^{1/2} * e_i \rangle_y = \langle \text{Log}(y^{-1/2} * x), e_i \rangle_F$

Theorem 3.1. *The manifold zonoid depth is a proper data depth function in the sense of Section 3.1, satisfying properties P.1–P.4 for distributions in $P_2(\mathcal{M})$. The uniquely existing point of maximum depth is characterized by the geometric mean of the distribution.*

In order to show that the continuity properties P.5 and P.6 also hold for the manifold zonoid depth, we need the following lemma.

Lemma 3.2. *Let $\nu \in P_2(\mathcal{M})$. Then, $\bigcup_{0 < \alpha \leq 1} D_{\alpha}^{\mathcal{M}}(\nu) = \text{conv}(\nu)$. In particular, for each $y \in \text{conv}(\nu)$, $\text{ZD}_{\mathcal{M}}(\nu, y) > 0$ by definition of the manifold zonoid depth.*

Theorem 3.3. *The manifold zonoid depth is continuous in y as in P.5 for $y \in \text{conv}(\nu)$ and $\nu \in P_2(\mathcal{M})$, i.e. if $\delta(y_n, y) \rightarrow 0$ with $y_n \in \mathcal{M}$ for all $n \in \mathbb{N}$, then,*

$$\lim_{n \rightarrow \infty} \text{ZD}_{\mathcal{M}}(\nu, y_n) = \text{ZD}_{\mathcal{M}}(\nu, y)$$

The manifold zonoid depth is uniformly continuous in ν as in P.6 for $y \in \text{rint}(\text{conv}(\nu))$ and $(\nu_n)_{n \in \mathbb{N}}$ in $P_2(\mathcal{M})$ uniformly integrable. If $\nu_n \xrightarrow{w} \nu$, then,

$$\sup_{y \in \text{rint}(\text{conv}(\nu))} |\text{ZD}_{\mathcal{M}}(\nu_n, y) - \text{ZD}_{\mathcal{M}}(\nu, y)| \rightarrow 0, \quad \text{as } n \rightarrow \infty$$

3.3 Integrated manifold zonoid depth

A straightforward generalization of the pointwise manifold zonoid depth in Definition 3.1 to compute the depth of a curve $y(t) \in \mathcal{M}$ with respect to a collection of marginal measures $\nu(t)$ for $t \in \mathcal{I} \subset \mathbb{R}$ is to consider the *integrated* manifold zonoid depth. The integrated depth is given by,

$$\text{iZD}_{\mathcal{M}}(\nu, y) := \int_{\mathcal{I}} \text{ZD}_{\mathcal{M}}(\nu(t), y(t)) dt = \int_{\mathcal{I}} \sup \{ \alpha : \mathbf{0}_{d \times d} \in D_{\alpha}(\zeta_y(t)) \} dt$$

where $\zeta_y(t)$ is the probability measure of the random variable $\text{Log}_{y(t)}(X(t)) \in \mathcal{T}_{y(t)}(\mathcal{M}) \cong \mathbb{R}^{d^2}$, such that $X(t)$ has probability measure $\nu(t)$. Note that such an integrated depth function is essentially analogous to e.g. the construction of the modified band depth (MBD), where we integrate pointwise Euclidean simplicial depths $y(t)$ over a functional domain $t \in \mathcal{I}$. For more details on functional data depth in a Euclidean context, we refer to López-Pintado and Romo (2009) or Sun and Genton (2012). The integrated versions of the properties **P.1** to **P.6** continue to hold for the integrated manifold zonoid depth and are straightforward generalizations of their pointwise analogs.

Theorem 3.4. *The integrated manifold zonoid depth is a proper integrated depth function in the sense of Section 3.1, satisfying the integrated versions of properties **P.1–P.4** for collections of marginal distributions $\nu(t) \in P_2(\mathcal{M})$ for $t \in \mathcal{I}$. The uniquely existing curve of maximum depth is characterized by the curve of pointwise geometric means of the marginal distributions.*

Proposition 3.5. *Let $y(t) \in \text{conv}(\nu(t))$, $\nu(t) \in P_2(\mathcal{M})$ and $y_n(t) \in \mathcal{M}$ for each $t \in \mathcal{I}$, such that $y_n(t) \rightarrow y(t)$ uniformly in t , i.e. $\sup_{t \in \mathcal{I}} \delta(y_n(t), y(t)) \rightarrow 0$. Then the integrated manifold zonoid depth is continuous in y as in **P.5** in the sense that,*

$$\lim_{n \rightarrow \infty} \text{iZD}_{\mathcal{M}}(\nu, y_n) = \text{iZD}_{\mathcal{M}}(\nu, y)$$

If $y(t) \in \text{rint}(\text{conv}(\nu))$, $(\nu_n(t))_{n \in \mathbb{N}}$ in $P_2(\mathcal{M})$ is a uniformly integrable sequence of measures uniform in t , and $\nu_n(t) \xrightarrow{w} \nu(t)$ uniformly in t . Then,

$$\sup_{y \in \text{rint}(\text{conv}(\nu))} |\text{iZD}_{\mathcal{M}}(\nu_n, y) - \text{iZD}_{\mathcal{M}}(\nu, y)| \rightarrow 0, \quad \text{as } n \rightarrow \infty$$

Here $y \in \text{rint}(\text{conv}(\nu))$ is read as $y(t) \in \text{rint}(\text{conv}(\nu(t)))$ for each $t \in \mathcal{I}$, and the uniform weak convergence $\nu_n(t) \xrightarrow{w} \nu(t)$ says that $\sup_{t \in \mathcal{I}} |\mathbb{E}_{\nu_n(t)}[\phi(X)] - \mathbb{E}_{\nu(t)}[\phi(X)]| \rightarrow 0$ for every continuous and bounded function $\phi : \mathcal{M} \rightarrow \mathbb{R}$

3.4 Geodesic distance depth

As a second notion of data depth on the geodesically complete manifold $\mathcal{M}$, we consider the *geodesic distance depth*, which can be viewed as the natural analog on the metric space $(\mathcal{M}, \delta)$ of the arc distance depth Liu and Singh (1992), defined for data observations on circles and spheres. The geodesic distance depth is straightforward to calculate, as it only requires the computation of

Riemannian distances between HPD matrices, an operation that remains computationally efficient also for higher-dimensional matrices.

Definition 3.2. (*Geodesic distance depth*) Let $\nu \in P_1(\mathcal{M})$, then the geodesic distance depth of a point $y \in \mathcal{M}$ with respect to the distribution ν is defined as:

$$\text{GDD}(\nu, y) = \exp \left(- \int_{\mathcal{M}} \delta(y, x) \nu(dx) \right) \quad (3.3)$$

Theorem 3.6. *The geodesic distance depth is a proper data depth function in the sense of Section 3.1, satisfying P.1–P.4 for distributions in $P_1(\mathcal{M})$. The uniquely existing point of maximum depth is characterized by the geometric median of the distribution.*

Theorem 3.7. *The geodesic distance depth is continuous in y as in P.5 for $y \in \text{cl}(\mathcal{M})$, the closure of $\mathcal{M}$, and $\nu \in P_1(\mathcal{M})$. That is, if $\delta(y_n, y) \rightarrow 0$ with $y_n \in \mathcal{M}$ for all $n \in \mathbb{N}$, then,*

$$\lim_{n \rightarrow \infty} \text{GDD}(\nu, y_n) = \text{GDD}(\nu, y)$$

The geodesic distance depth is uniformly continuous in ν as in P.6 for $y \in \mathcal{M}$ and $(\nu_n)_{n \in \mathbb{N}}$ uniformly integrable. If $\nu_n \xrightarrow{w} \nu$, then,

$$\sup_{y \in \mathcal{M}} |\text{GDD}(\nu_n, y) - \text{GDD}(\nu, y)| \rightarrow 0, \quad \text{as } n \rightarrow \infty$$

Remark In practice, we typically wish to compute the empirical depth of each observation in a sample $y \in \{x_1, \dots, x_n\}$ with respect to the empirical distribution ν_n of the sample $\{x_1, \dots, x_n\}$ itself. In this case, it suffices to compute the $(n \times n)$ -dimensional distance matrix with (i, j) -th entry $\delta(x_i, x_j)$. This matrix is fully determined by $n(n-1)/2$ components, as the diagonal entries are zero and $\delta(x_i, x_j) = \delta(x_j, x_i)$. In particular, in online applications where the depths need to be updated each time a new observation enters the sample, we simply add one extra column and row to the distance matrix for each newly recorded observation and update the depth values.

Remark A third notion of data depth on the geodesic manifold $\mathcal{M}$, closely related to the geodesic distance depth, is the *manifold spatial depth*. This can be viewed as the natural manifold generalization of the spatial depth in e.g. Vardi and Zhang (2000) or Serfling (2002). For a distribution $\nu \in P_1(\mathcal{M})$ and a point $y \in \mathcal{M}$, the manifold spatial depth is given by:

$$\text{SD}(\nu, y) = 1 - \left\| \int_{\mathcal{M}} \frac{\text{Log}_y(x)}{\delta(y, x)} \nu(dx) \right\|_y = 1 - \left\| \int_{\mathcal{M}} \frac{\text{Log}(y^{-1/2} * x)}{\delta(y, x)} \nu(dx) \right\|_F$$

The manifold spatial depth attains its maximum value $\text{SD}(\nu, m) = 1$ at the geometric median, since $\mathbf{E}_\nu \left[\frac{\text{Log}_m(x)}{\delta(m, x)} \right] = \mathbf{0}_{d \times d}$ by eq.(2.5), and the depth is lower bounded by zero, which is a direct consequence of the triangle inequality combined with the fact that $\|\text{Log}_y(x)\|_y = \delta(y, x)$. The manifold spatial depth is closely associated to the geodesic distance depth in the sense that it is based on the gradient of the distance function, i.e. the gradient of $f_x(y) = \delta(y, x)$ for fixed x is given by $\text{grad} f_x(y) = \frac{\text{Log}_y(x)}{\delta(y, x)}$, see Fletcher et al. (2009). Further theoretical properties of the manifold spatial depth are a topic of future research.

3.5 Integrated geodesic distance depth

In order to generalize the pointwise geodesic distance depth to the depth of a curve $y(t) \in \mathcal{M}$, with respect to a collection of marginal measures $\nu_t = \nu(t)$ for $t \in \mathcal{I} \subset \mathbb{R}$, we replace the pointwise expected distance in eq.(3.3) by an integrated expected distance as:

$$\text{iGDD}(\nu, y) = \exp \left(- \int_{\mathcal{I}} \int_{\mathcal{M}} \delta(y(t), x) \nu_t(dx) dt \right)$$

The integrated versions of the properties **P.1** to **P.6** continue to hold for the integrated geodesic distance depth and are straightforward generalizations of their pointwise analogs as in the case of the integrated manifold zonoid depth.

Theorem 3.8. *The integrated geodesic distance depth is a proper function depth function in the sense of Section 3.1, satisfying the integrated versions of properties **P.1–P.4** for collections of marginal distribution $\nu(t) \in P_1(\mathcal{M})$ for $t \in \mathcal{I}$. The uniquely existing curve of maximum depth is characterized by the curve of pointwise geometric medians of the marginal distributions.*

Proposition 3.9. *Let $y(t) \in \text{cl}(\mathcal{M})$ and $\nu(t) \in P_1(\mathcal{M})$ for each $t \in \mathcal{I}$, such that $y_n(t) \rightarrow y(t)$ uniformly in t , i.e. $\sup_{t \in \mathcal{I}} \delta(y_n(t), y(t)) \rightarrow 0$. Then the integrated geodesic distance depth is continuous in y as in **P.5** in the sense that,*

$$\lim_{n \rightarrow \infty} \text{iGDD}(\nu, y_n) = \text{iGDD}(\nu, y)$$

If $y(t) \in \mathcal{M}$, $(\nu_n(t))_{n \in \mathbb{N}}$ in $P_1(\mathcal{M})$ is a uniformly integrable sequence of measures uniform in t , and $\nu_n(t) \xrightarrow{w} \nu(t)$ uniformly in t . Then,

$$\sup_{y \in \mathcal{M}} |\text{iGDD}(\nu_n, y) - \text{iGDD}(\nu, y)| \rightarrow 0, \quad \text{as } n \rightarrow \infty$$

where $y \in \mathcal{M}$ is read as $y(t) \in \mathcal{M}$ for each $t \in \mathcal{I}$.

4 Rank-based tests for samples of HPD matrices

In this section, we further develop the application of the manifold data depths in the context of rank-based hypothesis testing for samples of covariance or spectral density matrices. In particular, for random samples of HPD matrices, we construct manifold analogs of several common univariate rank-based tests. This is in the same spirit as (Mosler, 2002, Chapter 5), Liu and Singh (1993) and Chenouri and Small (2012), where the authors extend the univariate Wilcoxon rank-sum and Kruskal-Wallis test for samples of Euclidean vectors by replacing the ordinary ranks by the ranks induced by the Euclidean data depth values. In addition, we propose a generalization of the Wilcoxon signed-rank test for samples of paired observations in the space of HPD matrices. This hypothesis test is not based on data depth, but on a specific manifold difference score between HPD

matrix-valued observations. In Appendix II in the supplementary material we investigate the size and power of the proposed rank-based tests based on various simulated scenarios, and –whenever possible– compare the performance with available benchmark procedures in the literature. All of the proposed tests are available in the R-package `pdSpecEst`, Chau (2017).

4.1 Homogeneity of distributions for independent samples

4.1.1 Two or more independent samples

Suppose that we observe $K \geq 2$ independent samples of random HPD matrices, and we wish to test whether the distributions of the matrices are homogeneous between samples. In particular, for $k = 1, \dots, K$, assume that $X_{k1}, \dots, X_{kn_k} \stackrel{\text{iid}}{\sim} \nu_k$ with $\nu_k \in P_2(\mathcal{M})$, then we wish to test the null hypothesis $H_0 : \nu_1 = \dots = \nu_K$. To build the test statistic, we compute the depth $d_N(x_{ki}) := D(\nu_{\cup_{ki} x_{ki}}, x_{ki})$ of each observation x_{ki} with respect to the empirical distribution $\nu_{\cup_{ki} x_{ki}}$ of the pooled sample $\bigcup_{k=1}^K \{x_{k1}, \dots, x_{kn_k}\}$, where the total number of observations is denoted by $N := \sum_{k=1}^K n_k$. Here, the depth $D(\cdot)$ can be substituted by either the (integrated) manifold zonoid depth or the (integrated) geodesic distance depth. The rank $r_N(x_{ki})$ of an observation is now defined as the usual univariate rank of the depth $d_N(x_{ki})$ in the complete set of depth values $\{d_N(x_{11}), \dots, d_N(x_{1n_1}), \dots, d_N(x_{K1}), \dots, d_N(x_{Kn_K})\}$. In the remainder of this section, we assume that for sufficiently large samples there are no tied ranks. If this is not the case, we may choose to break ties at random or to assign an averaged rank simultaneously to the tied depth values.

Proposition 4.1. (Manifold Wilcoxon rank-sum test) *In the case of $K = 2$ independent samples, consider the test statistic given by the sum of ranks of the first sample,*

$$T_{1,N} = \sum_{i=1}^{n_1} r_N(x_{1i})$$

Under the null hypothesis $H_0 : \nu_1 = \nu_2$, the exact distribution of $T_{1,N}$ is the same as in the univariate Wilcoxon rank-sum test, i.e. the distribution of the sum of n numbers randomly drawn without replacement from the integers $\{1, \dots, N\}$. Asymptotically, as $N \rightarrow \infty$ and $n_1/N \rightarrow p \in (0, 1)$, under the null hypothesis,

$$\left(\frac{n_1 n_2 (N+1)}{12} \right)^{-1/2} \left(T_{1,N} - \frac{n_1 (N+1)}{2} \right) \xrightarrow{d} N(0, 1)$$

*Furthermore, $T_{1,N}$ is invariant under congruent transformation of the observations $\{a * x_1, \dots, a * x_n, a * y_1, \dots, a * y_m\}$, for any $a \in \text{GL}(d, \mathbb{C})$.*

Proposition 4.2. (Manifold Kruskal-Wallis test) *For $K > 2$ independent samples, consider,*

$$T_{2,N} = \frac{12}{N(N+1)} \sum_{k=1}^K n_k \left(\bar{R}_{N,k} - \frac{N+1}{2} \right)^2$$

where $\bar{R}_{N,k} = \frac{1}{n_k} \sum_{i=1}^{n_k} r_N(x_{ki})$. Under the null hypothesis $H_0 : \nu_1 = \dots = \nu_K$, the asymptotic distribution of $T_{2,N}$ is the same as in the univariate Kruskal-Wallis test, i.e. if $n_k/N \rightarrow p_{k0} \in (0, 1)$ for each $k = 1, \dots, K$, then $T_{2,N} \xrightarrow{d} \chi_{K-1}^2$. Furthermore, $T_{2,N}$ is invariant under congruent transformation of the observations $(a * x_{ki})$ for all $k = 1, \dots, K$, $i = 1, \dots, n_k$, with $a \in \text{GL}(d, \mathbb{C})$.

Remark In Appendix II, the simulated power of the two manifold rank-based tests above is compared to the simulated power of two benchmark permutation tests based on the Riemannian distance between geometric sample means as in Pigoli et al. (2014) and Cabassi et al. (2017). Under several simulated scenarios, the manifold rank-based tests are observed to outperform the benchmark permutation tests in terms of power. In particular, the benchmark permutation tests are not able to detect a *change in scale* of the generating distributions on the manifold, as they do not take into account the spread of the data on the manifold. In contrast, the manifold rank-based tests are able to detect such a scale change, see the supplementary material for more details.

4.1.2 A single independent sample

In the data example in Section 5.1, our goal is to test if a sequence of independent HPD matrices is stationary in the sense that the marginal distributions of the matrices do not change across the sample. This amounts to testing the null hypothesis of randomness or exchangeability $H_0 : \nu_1 = \dots = \nu_n$ for a sequence of independent random variables $X_i \sim \nu_i$ on the manifold, with $\nu_i \in P_2(\mathcal{M})$ for $i = 1, \dots, n$. To construct such a test, we consider a depth-based version of the Bartels-von Neumann test (Bartels (1982)) to test against the alternative hypothesis of first-order (lag-1) autocorrelation, such as a smooth trend. If instead we wish to test for randomness against the alternative of serial autocorrelation at (several) higher-order lags $\ell = 1, \dots, L$, the Bartels-von Neumann rank statistic may be replaced by for instance a depth-based version of the Portmanteau rank statistic as in Dufour and Roy (1986). In the proposition below, $r_n(x_i)$ denotes the usual rank of $d_n(x_i)$ in the set of depth values $\{d_n(x_1), \dots, d_n(x_n)\}$, and $d_n(x_i) := D(\nu_{\cup_i x_i}, x_i)$ is the depth of x_i with respect to the empirical probability measure $\nu_{\cup_i x_i}$ of the set of observations $\{x_1, \dots, x_n\}$.

Proposition 4.3. (Manifold Bartels-von Neumann test) *Given independent observations $\{x_1, \dots, x_n\}$ on the manifold, consider the test statistic,*

$$T_{3,n} = \frac{\sum_{i=1}^{n-1} (r_n(x_{i+1}) - r_n(x_i))^2}{\sum_{i=1}^n (r_n(x_i) - (n+1)/2)^2}$$

*Under the null hypothesis $H_0 : \nu_1 = \dots = \nu_n$, the asymptotic distribution of $T_{3,n}$ is the same as in the univariate Bartels-von Neumann test, i.e. $(T_{3,n} - 2)/\sigma_n \xrightarrow{d} N(0, 1)$ as $n \rightarrow \infty$, where $\sigma_n^2 = \frac{4(n-2)(5n^2-2n-9)}{5n(n+1)(n-1)^2}$. Furthermore, $T_{3,n}$ is invariant under congruent transformation of the observations $\{a * x_1, \dots, a * x_n\}$ for any $a \in \text{GL}(d, \mathbb{C})$*

4.2 Homogeneity of distributions for paired samples

In practice, samples of HPD matrices are not always independent and the rank-based tests for independent samples in the previous section may break down. In the data example in Section 5.2, the data consists of spectral matrices corresponding to repeated measurements of the *same* subjects under two treatment conditions, and we still wish to test the null hypothesis of homogeneous distributions between samples. In this section, we therefore construct a test for homogeneity of distributions in the context of matched or paired samples of HPD matrices. Suppose that $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \nu_x$ and $Y_1, \dots, Y_n \stackrel{\text{iid}}{\sim} \nu_y$, are two equal-sized random samples on the manifold $\mathcal{M}$, with $\nu_x, \nu_y \in P_2(\mathcal{M})$, where (X_i, Y_i) , for $i = 1, \dots, n$, are paired random variables, e.g. repeated measurements of the same subject. In order to test the null hypothesis $H_0 : \nu_x = \nu_y$, we propose a straightforward manifold generalization of the univariate Wilcoxon signed-rank test. Instead of computing paired differences between observations by ordinary matrix subtraction, we consider specific manifold difference scores. Let $x, y \in \mathcal{M}$, then a natural measure of the difference between x and y is given by $\text{Log}(y^{-1/2} * x)$. This is heuristically explained as follows; $\text{Log}(y^{-1/2} * x)$ is an element of the tangent space $\mathcal{T}_{\text{Id}}(\mathcal{M})$, as $\text{Log}(y^{-1/2} * x) = \text{Log}_{\text{Id}}(y^{-1/2} * x)$. The norm in $\mathcal{T}_{\text{Id}}(\mathcal{M})$ induced by the Riemannian metric is the ordinary Frobenius norm $\|\cdot\|_F$, but $\|\text{Log}(y^{-1/2} * x)\|_F$ equals the Riemannian distance $\delta(x, y)$ by the definition in eq.(2.2). Since the norm of $\text{Log}(y^{-1/2} * x)$ equals the distance between x and y , $\text{Log}(y^{-1/2} * x)$ can naturally be interpreted as a difference between x and y on the Riemannian manifold. In particular, we consider the trace of the manifold difference, since the trace of the manifold difference is a real-valued scalar that follows a symmetric distribution around zero under the null hypothesis.

Proposition 4.4. (Manifold Wilcoxon signed-rank test) *Consider the test statistic given by,*

$$T_{4,n} = \sum_{i=1}^n \text{sgn}(\text{diff}(x_i, y_i)) \cdot r(|\text{diff}(x_i, y_i)|)$$

where $\text{diff}(x_i, y_i) = \text{Tr}(\text{Log}(y_i^{-1/2} * x_i))$ and $r(|\text{diff}(x_i, y_i)|)$ is the usual rank of $|\text{diff}(x_i, y_i)| \in \mathbb{R}_+$ in the set $\cup_{i=1}^n |\text{diff}(x_i, y_i)|$. Here it is implicitly assumed that $\text{diff}(x_i, y_i) \neq 0$ for each $i = 1, \dots, n$, otherwise exclude the zero difference scores and consider the reduced sample size n_r . Under the null hypothesis $H_0 : \nu_x = \nu_y$, the distribution of $T_{4,n}$ is the same as in the univariate Wilcoxon signed-rank test, with expectation zero and variance $n(n+1)(2n+1)/6$. Furthermore, $T_{4,n}$ is invariant under congruence transformation of the observations $\{a * x_1, \dots, a * x_n\}$ and $\{a * y_1, \dots, a * y_n\}$ for any $a \in \text{GL}(d, \mathbb{C})$.

4.2.1 Equal spectral matrices of two multivariate time series

In the context of multivariate spectral analysis, the manifold signed-rank test immediately provides an asymptotically valid test for equality of spectral matrices of two stationary multivariate time

series. Consider two independent strictly stationary (cf. (Brillinger, 1981, Assumption 2.6.1)) time series $(X_t)_{t=1,\dots,2T}$ and $(Y_t)_{t=1,\dots,2T}$, and let $\{\hat{f}_x(\omega_1), \dots, \hat{f}_x(\omega_T)\}$ and $\{\hat{f}_y(\omega_1), \dots, \hat{f}_y(\omega_T)\}$ be *positive definite* periodograms of the two time series at the Fourier frequencies $\omega_k = \pi k/T \in (0, \pi]$. In particular, one approach to obtain positive definite periodograms is through multitaper spectral estimation with a fixed number of $B \geq d$ tapering functions, where d denotes the dimension of the time series (see e.g. Dai and Guo (2004)). Let $\{k_T\}$ be a sequence of integers, such that $\omega_{k:T} = \pi k_T/T \rightarrow \omega_k$ as $T \rightarrow \infty$ for $k = 1, \dots, K$. By (Dai and Guo, 2004, Lemma 1), for strictly stationary time series $(X_t)_{t=1,\dots,2T}$ and $(Y_t)_{t=1,\dots,2T}$, the positive definite multitaper periodograms $\hat{f}(\omega_{k:T})$ are asymptotically independent at the Fourier frequencies $\{\omega_k\}_{k=1,\dots,K}$, where the dot may be replaced by either the x - or the y -sample. Furthermore, if $\omega_k \not\equiv 0 \pmod{\pi}$, $\hat{f}(\omega_{k:T})$ is asymptotically distributed as $W_d^c(B, B^{-1}f(\omega_k))$, a complex Wishart distribution of dimension d with B degrees of freedom. Therefore, for each $k = 1, \dots, K$ with $\omega_k \not\equiv 0 \pmod{\pi}$, it follows that,

$$f_{\cdot,k}^{-1/2} * \hat{f}_{\cdot,k} \xrightarrow{d} W_c^d(B, B^{-1}I_d), \quad \text{as } T \rightarrow \infty \quad (4.1)$$

writing $f_{\cdot,k} := f(\omega_k)$ and $\hat{f}_{\cdot,k} := \hat{f}(\omega_{k:T})$. Under the null hypothesis of equal spectral matrices $H_0 : f_{x,k} = f_{y,k}$, for each $k = 1, \dots, K$ with $\omega_k \not\equiv 0 \pmod{\pi}$, it follows that,

$$\begin{aligned} \text{Tr}(\text{Log}(\hat{f}_{y,k}^{-1/2} * \hat{f}_{x,k})) &= \text{Tr}(\text{Log}((f_{y,k}^{-1/2} * \hat{f}_{y,k})^{-1/2} * (f_{x,k}^{-1/2} * \hat{f}_{x,k}))) \\ &\xrightarrow{d} \text{Tr}(\text{Log}(W_y^{-1/2} * W_x)), \quad \text{as } T \rightarrow \infty \end{aligned} \quad (4.2)$$

where $W_y, W_x \stackrel{\text{iid}}{\sim} W_c^d(B, B^{-1}I_d)$ are complex Wishart matrices *independent of the underlying spectra*. In the first equality, we use the invariance of the difference scores under matrix congruence transformation combined with the fact that $f_{y,k} = f_{x,k}$. The second step then follows by an application of the continuous mapping theorem. By the above arguments, under the null hypothesis, the manifold difference scores $\text{Tr}(\text{Log}(\hat{f}_{y,k}^{-1/2} * \hat{f}_{x,k}))$ for $k = 1, \dots, K$ are asymptotically i.i.d. and their asymptotic distribution is symmetric around zero by the same argument as in the proof of Proposition 4.3. The ordinary Wilcoxon-signed rank test based on the manifold difference scores now provides an asymptotically valid test for equality of the underlying spectra $f_x(\omega)$ and $f_y(\omega)$.

Remark Under the same stationarity assumptions, Jentsch and Pauly (2015) and Dette and Paparoditis (2009) consider permutation tests for the equality of spectral matrices of multivariate time series based on the integrated L_2 -distance between smoothed periodograms. As pointed out in Jentsch and Pauly (2015), these tests are usually not consistent for non-smoothed periodograms, due to the use of non-consistent spectral estimators. As a consequence, the permutation tests in Jentsch and Pauly (2015) and Dette and Paparoditis (2009) require consistent smoothing of the periodograms, introducing an additional tuning parameter in the test procedure. In contrast, other than the minimum amount of smoothing required to ensure positive definiteness of the periodograms, the manifold signed-rank test does not require consistent estimators of the underlying

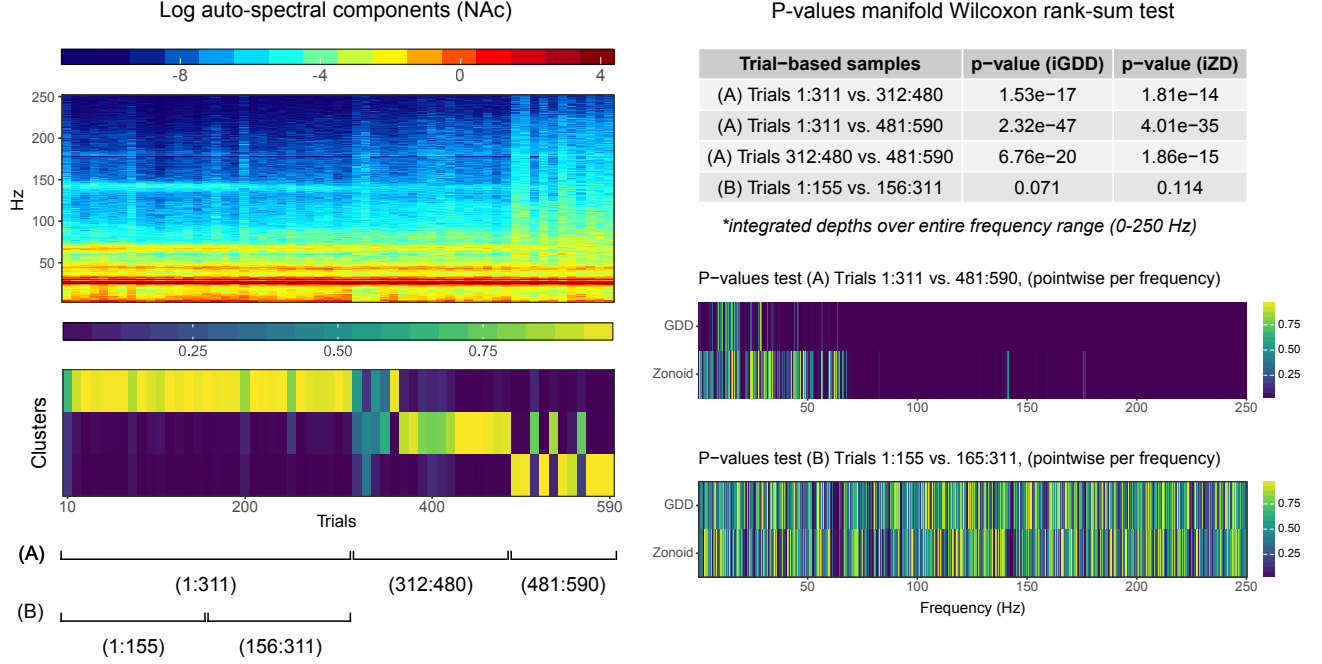


Figure 1: Top-left: NAc auto-spectra of (the matrix logarithm) of block-averaged periodograms. Bottom-left: probabilistic cluster assignments for $K = 3$ clusters of block-averaged periodograms. Top- and bottom-right: p-values of manifold Wilcoxon rank-sum test for two samples of periodograms, either integrated over frequency (top) or pointwise per frequency (bottom).

spectra. Moreover, the asymptotic test remains computationally efficient also for a large number of Fourier frequencies or higher-dimensional spectral matrices, whereas the exact tests in Jentsch and Pauly (2015) and Dette and Paparoditis (2009) become computationally more burdensome, as we need to calculate many computationally expensive randomized test statistics. At the same time, the asymptotic test is able to outperform the permutation test in Jentsch and Pauly (2015) in terms of power under several different simulated scenarios, in particular in the case of a *multiplicative* shift in the baseline spectrum as shown in Appendix II in the supplementary material.

5 Analysis of local field potential data

5.1 Associative learning experiment with a macaque

In a first data example, we apply the proposed rank-based tests as a means of inference to analyze evolving spectral behavior in a brain signal dataset consisting of multivariate local field potential (LFP) time series recorded for many trials over the course of an associative learning experiment with a macaque, see Gorrostrieta et al. (2012), Fiecas and Ombao (2016), or Chau and von Sachs (2016) for more details. After preprocessing of the LFP time series data, there remain a total of 590

approximately stationary 2-dimensional time series trials of length 2048 sampled at 1000 Hz, thus roughly corresponding to 2 seconds of data per trial. The two components of the time series trials correspond to LFP measurements in the hippocampus (Hc) and nucleus accumbens (NAc) regions of the macaque’s brain, see e.g. (Fiecas and Ombao, 2016, Figure 1) for the recorded LFP time series at several trials throughout the experiment. In Chau and von Sachs (2017), we performed wavelet-based probabilistic clustering of trial-specific spectral matrices on the Riemannian manifold of HPD matrices and in this section we wish to further validate the obtained cluster results. Figure 1 shows the probabilistic cluster assignments for $K = 3$ clusters based on 59 block-averaged periodogram matrices (averaged over 10 adjacent time series trials) with reference to the NAc auto-spectral components of the block-averaged periodogram matrices, which give an indication as to why we obtain the given cluster assignments. The clusters suggest gradually evolving spectral behavior over the course of the learning experiment, and in particular the trial-specific spectral behavior in the LFP time series towards the end of the associative learning experiment appears to differ significantly from the spectral behavior at the start of the experiment.

Significant difference between clusters of trials In order to test for significant differences between clusters, we apply the manifold Wilcoxon rank-sum test for homogeneity of distributions between two independent samples of periodogram matrices, either simultaneously over the entire frequency range or pointwise per frequency. Note that it is implicitly assumed that the LFP time series are independent between trials. Using the R-package `pdSpecEst` (Chau (2017)), the raw trial-specific periodogram matrices are pre-smoothed by a multitaper spectral estimator with $B = 2$ Slepian tapering functions to ensure positive definiteness at each individual frequency. Next, we separate the noisy trial-specific periodograms into two independent samples based on several different scenarios specified in the right-hand side of Figure 1. In particular, we test for pairwise differences in spectral behavior between three observed phases in the experiment corresponding to trials 1 to 311 (start of the experiment), trials 312 to 480 (middle of the experiment), and trials 481 to 590 (end of the experiment) respectively. The table in Figure 1 shows that the integrated tests over the entire frequency range (0–250 Hz) –for both integrated depth measures– confidently reject the null hypothesis of homogeneity of distributions of the periodograms at any reasonable significance level. To compare, we run the same test for two samples containing the trials 1 to 155 and trials 156 to 311 respectively. This corresponds to trials contained *within* the first phase of the experiment. At the significance level $\alpha = 0.05$, we cannot reject the null hypothesis of homogeneous distributions in the two trial-based samples, suggesting that the spectral behavior does not significantly evolve within the first phase of the experiment. The bottom right-hand image in Figure 1 displays p-values of the tests pointwise per frequency, comparing the distributional behavior of periodogram matrices between trials 1 to 311 (start) and trials 481 to 590 (end); and between trials 1 to 155 (start) and trials 156 to 311 (start) respectively. In the first scenario,

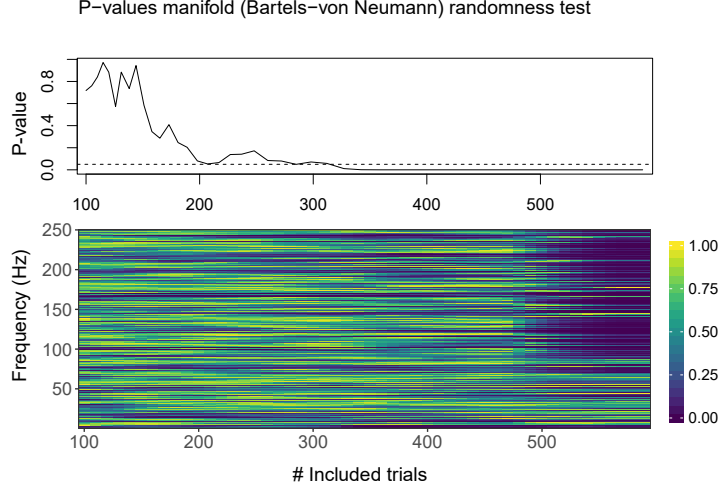


Figure 2: Top: p-values of manifold Bartels-von Neumann test for one sample of periodograms integrated over frequency based on the iGDD. The horizontal (dashed) line indicates the level $\alpha = 0.05$. Bottom: p-values pointwise per frequency based on the (pointwise) GDD.

the pointwise p-values are approximately zero for most frequencies above 70 Hz using the zonoid depth and for most frequencies above 30 Hz using the geodesic distance depth. In the second scenario, there is no clear indication of non-homogeneous distributions between the two samples, except around the 60 Hz frequency band for both the geodesic distance and zonoid depth or the 145 Hz frequency band for the zonoid depth. It may be interesting to further investigate whether this behavior is explained by actual electronic activity in the macaque’s brain or if this is caused by recording artifacts in the data. Overall, the obtained results largely confirm our a priori expectations based on the evolving spectral characteristics in the top-left image of Figure 1.

Evolving spectral behavior over trials Instead of comparing the distributional behavior of periodograms in two or more samples of trials, we can also directly test for an underlying trend in the distributions of the trial-specific periodograms by applying the manifold Bartels-von Neumann test. In this case, we consider the trial-specific periodograms over the course of the learning experiment as a single independent sample, and we test the null hypothesis of randomness of the periodograms against the alternative of lag-1 serial correlation, such as a smooth trend in the underlying spectral matrices. Figure 2 displays obtained p-values of the manifold Bartels-von Neumann test, both simultaneous over the entire frequency range (0–250 Hz) and pointwise per frequency, for an accumulative sample of LFP time series trials. The values on the x-axes correspond to the number of trials from the start of the experiment included in the test, with a maximum of 590, in which case all time series trials are included. The p-values in Figure 2 are based on the integrated (resp. pointwise) geodesic distance depth. For the integrated (resp. pointwise) zonoid depth the results are similar although somewhat less pronounced especially towards the end

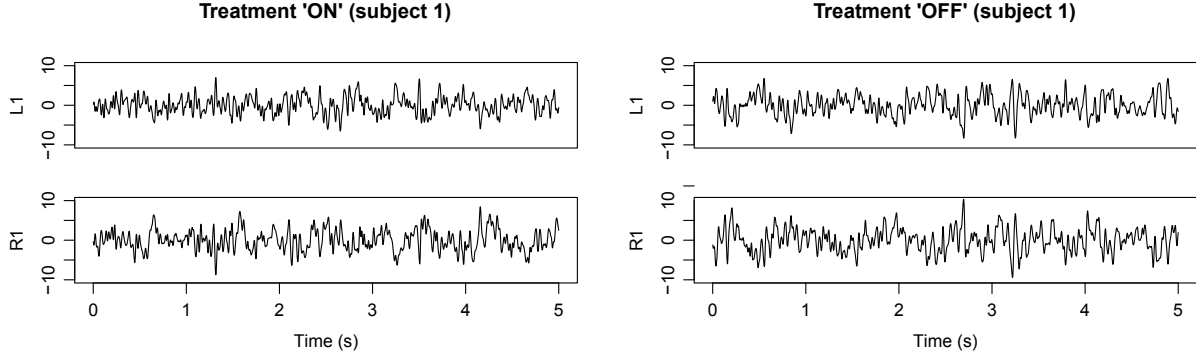


Figure 3: Local field potential (LFP) time series recordings (5 seconds, at rest) of a patient with Parkinson’s disease at the L1- and R1-channel under no treatment and dopaminergic treatment, downsampled to 50 Hz (Butterworth low-pass filter).

of the experiment for the pointwise tests. Based on these test results, at the level $\alpha = 0.05$, the integrated test consistently rejects the null hypothesis of randomness for more than approximately 300 trials included from the start of the experiment. Note that this more or less corresponds to the gradual change from the first to the second cluster of trials in Figure 1. For most frequencies above 90 Hz, the pointwise tests consistently reject the null hypothesis of randomness for more than approximately 480 trials included from the start of the experiment. This corresponds to the jump from the second to the third cluster of trials in Figure 1.

5.2 Parkinson’s disease dopaminergic treatment

In this second application, our task is to analyze treatment differences in a dataset of multivariate LFP time series trials recorded for patients with advanced Parkinson’s disease in resting state. The LFP time series are recorded under no particular treatment (treatment ‘off’), and under antiparkinsonian dopamine replacement therapy (treatment ‘on’), see Zénon et al. (2016) for more details. The goal of the analysis is to investigate whether the spectral characteristics in the LFP time series traces significantly differ between treatment conditions, and if so, at which particular frequencies or frequency bands. After preprocessing of the data, there remain a total of 22 multivariate time series traces of dimension $d = 8$ sampled at 1000 Hz, cut off at the shared minimum length of 115424 time series observations. The time series are recorded for 11 patients with Parkinson’s disease under both the treatment ‘on’ and treatment ‘off’ conditions, and the eight recorded LFP time series channels L1 – L4 and R1 – R4 correspond to four electrodes on the left side and four electrodes on the right side of the brain. Figure 3 shows a short LFP time series segment of a particular patient recorded at the L1- and R1-channel under both treatment conditions. Assuming second-order stationary of the time series recordings, we compute positive definite periodograms via multitaper spectral estimators with $B = d = 8$ Slepian tapering functions to guarantee positive

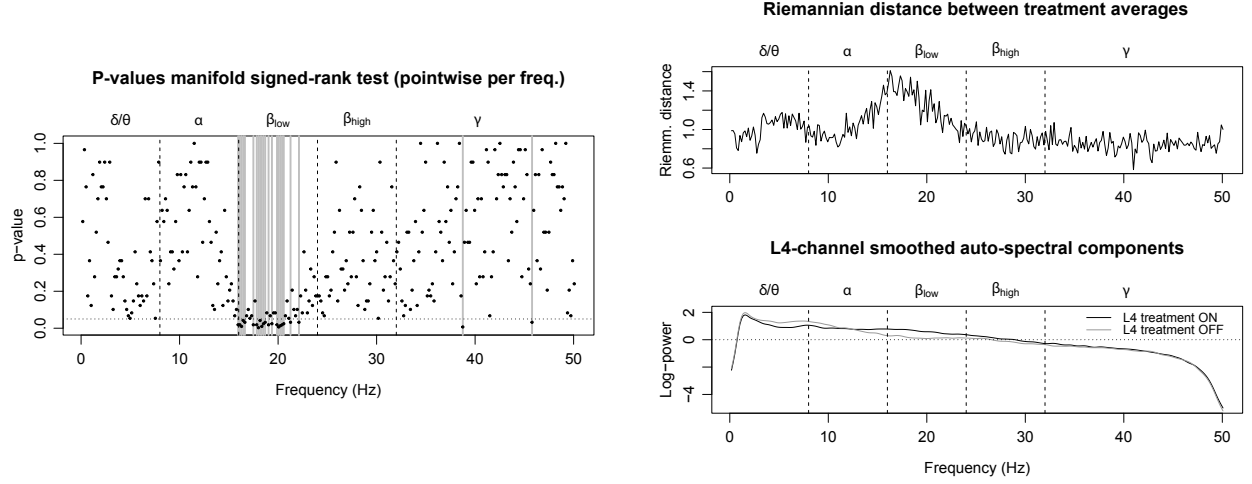


Figure 4: Left: p-values of manifold Wilcoxon signed-rank test for paired samples pointwise per frequency, with a horizontal dashed line at the level $\alpha = 0.05$. Top-right: Riemannian distance pointwise per frequency between the two treatment-specific group-mean periodograms. Bottom-right: L4-channel auto-spectral components of smoothed treatment-specific group-mean periodograms.

definiteness at each individual frequency.

Treatment differences in spectral power As the LFP time series are recorded under the two treatment conditions for the same subjects, assuming independence between subjects, we test for homogeneity of the distributions of the periodogram matrices using the manifold Wilcoxon signed-rank test for matched or paired samples of HPD matrices. In the left-hand image in Figure 4, we display p-values of the Wilcoxon signed-rank test applied pointwise per frequency in the frequency range 0–50 Hz, as the spectral power largely fades out for frequencies above 50 Hz. As two reference images, we also plot the Riemannian distance between the two geometric mean periodogram matrices across subjects within the two treatment groups (treatment ‘on’ and treatment ‘off’), and we display the L4-channel auto-spectral components of the geometric mean periodogram matrices across subjects in the two treatment groups, as the difference in spectral behavior between the treatment conditions is most pronounced in this channel. Here, the geometric mean periodogram matrices have been smoothed by a multitaper spectral estimator. The vertical dashed lines indicate several frequency bands of interest to the neuroscientist: the δ – and θ -band (0.2–8 Hz), the α -band (8–16 Hz), the lower β -band (16–24 Hz), the upper β -band (24–32 Hz) and the γ -band (> 32 Hz). At the significance level $\alpha = 0.05$, the pointwise manifold signed-rank tests reject homogeneity of the distributions of the periodograms primarily within the 16–21 Hz frequency range (i.e. lower β -band), where the spectral (log-)power of the LFP time series of subjects under dopaminergic treatment exceeds that of subjects without treatment. These conclusions also agree with the Rie-

mannian distances between the within-group averages in the reference image on the top, which are largest in the lower β -band. In future work, it may be of interest to further investigate the neurological processes that cause the observed treatment differences in spectral behavior in the lower β -band, but this is outside of the scope of this illustrating example.

6 Conclusion and outlook

In this work we developed new data depth concepts for matrix-valued observations in the space of symmetric or HPD matrices. The sample data depth values are straightforward to compute and remain computationally efficient also for relatively high-dimensional matrices, with implementations directly available in the R-package `pdSpecEst`, Chau (2017). As such, the data depths serve as an easy-to-use data exploration tool, but also provide a powerful framework for nonparametric rank-based hypothesis testing in the context of random samples of HPD matrices. None of the introduced rank-based hypothesis tests require resampling or bootstrapping of test statistics and the proposed tests remain computationally efficient for large samples of matrix-valued observations or relatively high-dimensional matrices. At the same time, they are able to compete with –or outperform– available benchmark permutation tests in terms of power, as illustrated in the supplementary material under various simulated scenarios.

At the time of writing of this paper, we learned of the recent work of Paindaveine and Van Bever (2017), in which the authors construct halfspace depths for symmetric PD matrices in the context of deepest scatter, covariance and shape matrices. Although developed from a different perspective, the authors arrive at similar desired properties for depth functions acting on the space of symmetric PD matrices. Unfortunately, the manifold halfspace depth does not necessarily satisfy all of the desired properties, (in particular **P.3** in Section 3.1 may fail to hold). Another practical issue is the fact that the manifold halfspace depth is computationally quite expensive, and as the authors point out more efficient algorithms are already needed for (3×3) -dimensional matrix-valued observations. Finally, in Paindaveine and Van Bever (2017) the authors note that it is of primary interest to further investigate depth-based hypothesis testing for samples of covariance matrices, which has been developed in Section 4 of this work.

The potential applications of data depth in the context of covariance or spectral density matrices are far more diverse than only rank-based hypothesis testing considered in this paper, see e.g. Serfling (2006). In terms of inference, an interesting application is the construction of confidence (or credible) regions for estimated covariance or spectral matrix based on depth trimmed regions in the same spirit as Yeh and Singh (1997). This could serve as a more natural alternative to matrix-componentwise confidence regions as considered in e.g. Dai and Guo (2004) or Fiecas and Ombao (2016). Another application of data depth for large collections of covariance or spectral matrices,

as in e.g. diffusion tensor imaging, is the construction of depth-based (functional) boxplots similar to e.g. Sun and Genton (2012) in a Euclidean context. This allows for straightforward characterization of central regions in the data and detection of outlying covariance or spectral matrices as *single objects* in the space of HPD matrices. To conclude, in addition to the presented manifold depth functions, our intention is to further examine the theoretical properties and performance of the manifold spatial depth briefly mentioned in Section 3.4. Implementations of the manifold spatial depth are already available in the R-package `pdSpecEst` alongside with the manifold zonoid and geodesic distance depth. Preliminary simulations show that the manifold spatial depth and the geodesic distance depth are comparable in terms of computation time and performance.

References

- Bagdonavicius, V., J. Kruopis, and M. Nikulin (2011). *Nonparametric Tests for Complete Data*. Hoboken: John Wiley & Sons.
- Bartels, R. (1982). The rank version of von Neumann’s ratio test for randomness. *Journal of the American Statistical Association* 77(377), 40–46.
- Bhatia, R. (2009). *Positive Definite Matrices*. Princeton university press.
- Brillinger, D. (1981). *Time Series: Data Analysis and Theory*. San Francisco: Holden-Day.
- Cabassi, A., D. Pigoli, P. Secchi, and P. Carter (2017). Permutation tests for the equality of covariance operators of functional data with applications to evolutionary biology. *ArXiv preprint ArXiv:1701.05870*.
- Chau, J. (2017). *pdSpecEst: An Analysis Toolbox for Hermitian Positive Definite Matrices*. R package version 1.1.0.
- Chau, J. and R. von Sachs (2016). Functional mixed effects wavelet estimation for spectra of replicated time series. *Electronic Journal of Statistics* 10(2), 2461–2510.
- Chau, J. and R. von Sachs (2017). Positive-definite multivariate spectral estimation: a geometric wavelet approach. *ArXiv preprint 1701.03314*.
- Chenouri, S. and C. Small (2012). A nonparametric multivariate multisample test based on data depth. *Electronic Journal of Statistics* 6, 760–782.
- Dai, M. and W. Guo (2004). Multivariate spectral analysis using Cholesky decomposition. *Biometrika* 91(3), 629–643.
- Dette, H. and E. Paparoditis (2009). Bootstrapping frequency domain tests in multivariate time series with an application to comparing spectral densities. *Journal of the Royal Statistical Society: Series B* 71(4), 831–857.
- do Carmo, M. (1992). *Riemannian Geometry*. Boston: Birkhäuser.
- Dufour, J. and R. Roy (1986). Generalized portmanteau statistics and tests of randomness. *Communications in Statistics - Theory and Methods* 15(10), 2953–2972.
- Dyckerhoff, R. (2016). Convergence of depths and depth-trimmed regions. *ArXiv preprint 1611.08721*.
- Fiecas, M. and H. Ombao (2016). Modeling the evolution of dynamic brain processes during an associative learning experiment. *Journal of the American Statistical Association* 111(516), 1440–1453.
- Fletcher, P., J. Moeller, J. Phillips, and S. Venkatasubramanian (2011). Horoball hulls and extents in positive definite space. In *Workshop on Algorithms and Data Structures*, pp. 386–398.
- Fletcher, P., S. Venkatasubramanian, and S. Joshi (2009). The geometric median on Riemannian manifolds with application to robust atlas estimation. *NeuroImage* 45(1), S143–S152.
- Gorrostieta, C., H. Ombao, R. Prado, S. Patel, and E. Eskandar (2012). Exploring dependence between brain signals in a monkey during learning. *Journal of Time Series Analysis* 33(5), 771–778.
- Holbrook, A., S. Lan, A. Vandenberg-Rodes, and B. Shahbaba (2016). Geodesic Lagrangian Monte Carlo over the space of positive definite matrices: with application to Bayesian spectral density estimation.

- ArXiv preprint 1612.08224*.
- Jentsch, C. and M. Pauly (2015). Testing equality of spectral densities using randomization techniques. *Bernoulli* 21(2), 697–739.
- Le, H. (1995). Mean size-and-shapes and mean shapes: a geometric point of view. *Advances in Applied Probability* 27(1), 44–55.
- Liu, R., J. Parelius, K. Singh, et al. (1999). Multivariate analysis by data depth: descriptive statistics, graphics and inference. *The Annals of Statistics* 27(3), 783–858.
- Liu, R. and K. Singh (1992). Ordering directional data: concepts of data depth on circles and spheres. *The Annals of Statistics*, 1468–1484.
- Liu, R. and K. Singh (1993). A quality index based on data depth and multivariate rank tests. *Journal of the American Statistical Association* 88(421), 252–260.
- López-Pintado, S. and J. Romo (2009). On the concept of depth for functional data. *Journal of the American Statistical Association* 104(486), 718–734.
- Mosler, K. (2002). *Multivariate Dispersion, Central Regions, and Depth: The Lift Zonoid Approach*. New York: Springer Science & Business Media.
- Paindaveine, D. and G. Van Bever (2017). Halfspace depths for scatter, concentration and shape matrices. *ArXiv preprint 1704.06160*.
- Pennec, X. (2006). Intrinsic statistics on Riemannian manifolds: basic tools for geometric measurements. *Journal of Mathematical Imaging and Vision* 25(1), 127–154.
- Pennec, X., P. Fillard, and N. Ayache (2006). A Riemannian framework for tensor computing. *International Journal of Computer Vision* 66(1), 41–66.
- Pigoli, D., J. Aston, I. Dryden, and P. Secchi (2014). Distances and inference for covariance operators. *Biometrika* 101(2), 409–422.
- Said, S., L. Bombrun, Y. Berthoumieu, and J. Man-ton (2015). Riemannian Gaussian distributions on the space of symmetric positive definite matrices. *ArXiv preprint 1507.01760*.
- Serfling, R. (2002). A depth function and a scale curve based on spatial quantiles. In *Statistical Data Analysis Based on the L1-Norm and Related Methods*, pp. 25–38. Basel: Birkhäuser Verlag.
- Serfling, R. (2006). Depth functions in nonparametric multivariate inference. *DIMACS Series in Discrete Mathematics and Theoretical Computer Science* 72.
- Skovgaard, L. (1984). A Riemannian geometry of the multivariate normal model. *Scandinavian Journal of Statistics* 11(4), 211–223.
- Smith, S. (2000). Intrinsic Cramér-Rao bounds and subspace estimation accuracy. In *Proceedings of the IEEE Sensor Array and Multichannel Signal Processing Workshop*, pp. 489–493. IEEE.
- Sun, Y. and M. Genton (2012). Functional box-plots. *Journal of Computational and Graphical Statistics* 20(2), 316–334.
- Vardi, Y. and C. Zhang (2000). The multivariate L1-median and associated data depth. *Proceedings of the National Academy of Sciences* 97(4), 1423–1426.
- Villani, C. (2009). *Optimal Transport: Old and New*. Berlin: Springer-Verlag.
- Yeh, A. and K. Singh (1997). Balanced confidence regions based on Tukeys depth and the bootstrap. *Journal of the Royal Statistical Society: Series B* 59(3), 639–652.
- Zénon, A., Y. Duclos, R. Carron, T. Witjas, C. Baunez, J. Régis, J. Azulay, P. Brown, and A. Eusebio (2016). The human subthalamic nucleus encodes the subjective value of reward and the cost of effort during decision-making. *Brain* 139(6), 1830–1843.
- Zuo, Y. and R. Serfling (2000). General notions of statistical depth function. *Annals of statistics*, 461–482.

7 Appendix I: Proofs

7.1 Proof of Theorem 3.1

Proof. **P.1** This is a direct consequence of the claim that the following two events are equivalent:

$$\{\mathbf{0}_{d \times d} \in D_\alpha(\zeta_y)\} \Leftrightarrow \{\mathbf{0}_{d \times d} \in D_\alpha(\zeta_{a,y})\}, \quad 0 \leq \alpha \leq 1 \quad (7.1)$$

with ζ_y the probability measure of $\text{Log}_y(X)$ and $\zeta_{a,y}$ the probability measure of $\text{Log}_{a*y}(a*X)$, where X has probability measure ν . Here, the Euclidean zonoid trimmed region $D_\alpha(\zeta_y)$ is represented as a set of $(d \times d)$ -dimensional Hermitian matrices, instead of an equivalent set of d^2 -dimensional real basis component vectors, as in Section 3.2, and $\mathbf{0}_{d \times d}$ is the zero matrix. For $\alpha = 0$, the equivalence in eq.(7.1) is true by definition, since $D_0(\zeta_y) = D_0(\zeta_{a,y}) = \mathbb{R}^{d \times d}$.

Suppose that $\mathbf{0}_{d \times d} \in D_\alpha(\zeta_y)$ for some $0 < \alpha \leq 1$. Noting that $\mathcal{T}_y(\mathcal{M})$ can be identified by the real vector space of Hermitian matrices $\mathcal{H}$ for each $y \in \mathcal{M}$, by definition of the zonoid α -trimmed region, there exists a measurable function $\tilde{g} : \mathcal{H} \rightarrow [0, \frac{1}{\alpha}]$, such that,

$$\int_{\mathcal{H}} \tilde{g}(z) \zeta_y(dz) = 1, \quad \int_{\mathcal{H}} z \tilde{g}(z) \zeta_y(dz) = \mathbf{0}_{d \times d}$$

It is straightforward to verify that for each $a \in \text{GL}(d, \mathbb{C})$ and $x, y \in \mathcal{M}$, $\text{Log}_{a*y}(a*X) = a*\text{Log}_y(x)$. Define $g(z) = \tilde{g}(a^{-1} * z)$, then $g : \mathcal{H} \rightarrow [0, \frac{1}{\alpha}]$ is a measurable function such that,

$$\int_{\mathcal{H}} g(z) \zeta_{a,y}(dz) = \int_{\mathcal{H}} g(a * z) \zeta_y(dz) = \int_{\mathcal{H}} \tilde{g}(z) \zeta_y(dz) = 1$$

and,

$$\begin{aligned} \int_{\mathcal{H}} z g(z) \zeta_{a,y}(dz) &= \int_{\mathcal{H}} (a * z) g(a * z) \zeta_y(dz) = \int_{\mathcal{H}} (a * z) \tilde{g}(z) \zeta_y(dz) \\ &= a * \left(\int_{\mathcal{H}} z \tilde{g}(z) \zeta_y(dz) \right) \\ &= a * \mathbf{0}_{d \times d} = \mathbf{0}_{d \times d} \end{aligned}$$

Therefore $\mathbf{0}_{d \times d} \in D_\alpha(\zeta_{a,y})$. The other direction follows by a similar argument, using that $a \neq \mathbf{0}_{d \times d}$.

P.2 The zonoid trimmed region $D_1(\zeta_y)$ contains the single point $\mathbf{E}_\nu[\text{Log}_y(X)]$ by construction. The deepest point $y \in \mathcal{M}$ is therefore characterized by the point that satisfies $\mathbf{E}_\nu[\text{Log}_y(X)] = \mathbf{0}_{d \times d}$. By eq.(2.4) in the main document, on the Riemannian manifold $\mathcal{M}$ with $\nu \in P_2(\mathcal{M})$, this point is the uniquely existing geometric expectation of the distribution ν .

P.3 Using the equivalent definition $\text{ZD}_{\mathcal{M}}(\nu, y) = \sup\{\alpha : y \in D_\alpha^{\mathcal{M}}(\nu)\}$, by construction $D_\alpha^{\mathcal{M}}(\nu)$ is a geodesically convex set that contains the geometric mean $\mu := \mathbb{E}_\nu[X]$ for each $\alpha \in [0, 1]$. Also, $D_{\alpha_1}^{\mathcal{M}}(\nu) \subseteq D_{\alpha_2}^{\mathcal{M}}(\nu)$ for each $1 \geq \alpha_1 \geq \alpha_2 \geq 0$. Combining the above arguments, it follows that a geodesic curve $\text{Exp}_\mu(th)$, with $t \geq 0$ increasing, has monotone non-increasing depth as it moves

further away from the center μ .

P.4 With the same notation as above, for $\alpha \in (0, 1]$ we claim that the sets $D_\alpha^\mathcal{M}(\nu)$ are closed and bounded, and therefore also compact by the Hopf-Rinow theorem. The fact that the sets are closed follows directly from the definition of $D_\alpha^\mathcal{M}(\nu)$. The fact that they are bounded is seen as follows; for $\alpha > 0$, by construction $D_\alpha^\mathcal{M}(\nu) \subset \mathcal{M}$. Therefore, if $y \in D_\alpha^\mathcal{M}(\nu)$, necessarily $\delta(\text{Id}, y) < \infty$, which follows by the fact that both Id and y are elements of $\mathcal{M}$, combined with (Bhatia, 2009, Theorem 6.1.6). Let $(y_n)_{n \in \mathbb{N}}$ be an unbounded sequence, such that $\|\text{Log}(y_n)\|_F \rightarrow \infty$ as $n \rightarrow \infty$. The divergence $\|\text{Log}(y_n)\|_F \rightarrow \infty$ implies in particular also that $\delta(\text{Id}, y_n) \rightarrow \infty$, which violates the boundedness (or compactness) of $D_\alpha^\mathcal{M}(\nu)$ for $a \in (0, 1]$, and therefore we must have $\lim_{n \rightarrow \infty} \text{ZD}_\mathcal{M}(\nu, y_n) = \lim_{n \rightarrow \infty} \sup\{\alpha : y_n \in D_\alpha^\mathcal{M}(\nu)\} = 0$. $\square$

7.2 Proof of Lemma 3.2

Proof. By definition of the manifold zonoid trimmed regions $D_\alpha^\mathcal{M}(\nu) = \{y \in \mathcal{M} : \mathbf{0}_{d \times d} \in D_\alpha(\zeta_y)\}$ with $D_\alpha(\zeta_y)$ as in eq.(7.1). The distribution ζ_y has finite first moment with respect to the Riemannian metric in $\mathcal{T}_y(\mathcal{M})$, since

$$\begin{aligned} \int_{\mathcal{T}_y(\mathcal{M})} \|z\|_y \zeta_y(dz) &= \int_{\mathcal{M}} \|\text{Log}_y(x)\|_y \nu(dx) \\ &= \int_{\mathcal{M}} \delta(y, x) \nu(dx) < \infty \end{aligned}$$

using eq.(2.3) in the main document and the fact that $\nu \in P_2(\mathcal{M}) \subset P_1(\mathcal{M})$. By (Mosler, 2002, Theorem 3.13) for a probability measure ζ_y defined on $\mathcal{T}_y(\mathcal{M}) \cong \mathbb{R}^{d^2}$ with finite first moments,

$$\bigcup_{\alpha > 0} D_\alpha(\zeta_y) = \text{conv}_{\mathcal{T}_y(\mathcal{M})}(\zeta_y)$$

where $\text{conv}_{\mathcal{T}_y(\mathcal{M})}(\zeta_y)$ denotes the convex hull of the support of ζ_y in $\mathcal{T}_y(\mathcal{M}) \cong \mathbb{R}^{d^2}$, based on the Riemannian metric on $\mathcal{T}_y(\mathcal{M})$, i.e. a rescaled Euclidean metric. Using the above result, we write out,

$$\begin{aligned} \bigcup_{\alpha > 0} D_\alpha^\mathcal{M}(\nu) &= \bigcup_{\alpha > 0} \{y \in \mathcal{M} : \mathbf{0}_{d \times d} \in D_\alpha(\zeta_y)\} \\ &= \{y \in \mathcal{M} : \mathbf{0}_{d \times d} \in \bigcup_{\alpha > 0} D_\alpha(\zeta_y)\} \\ &= \{y \in \mathcal{M} : \mathbf{0}_{d \times d} \in \text{conv}_{\mathcal{T}_y(\mathcal{M})}(\zeta_y)\} \\ &= \left\{ y \in \mathcal{M} : \exists g : \text{supp}(\nu) \rightarrow [0, 1] \text{ measurable, s.t.} \right. \\ &\quad \left. \int_{\text{supp}(\nu)} \text{Log}_y(x) g(x) \lambda(dx) = \mathbf{0}_{d \times d} \text{ and } \int_{\text{supp}(\nu)} g(x) \lambda(dx) = 1 \right\} \\ &= \text{conv}(\nu) \end{aligned}$$

where the last step follows by definition $\text{conv}(\nu)$ as the geodesic convex hull of the support of ν on the manifold. $\square$

7.3 Proof of Theorem 3.3

7.3.1 Continuity in y (P.5)

Proof. We argue that the map $y \mapsto \text{ZD}_{\mathcal{M}}(\nu, y)$ is both upper- and lower-semicontinuous for $y \in \text{conv}(\nu)$.

Upper-semicontinuity: the map is upper-semicontinuous if and only if for each $\alpha \in [0, 1]$ the sets $\{y \in \text{conv}(\nu) : \text{ZD}_{\mathcal{M}}(\nu, y) < \alpha\}$ are open in $\text{conv}(\nu)$ or equivalently the sets $\{y \in \text{conv}(\nu) : \text{ZD}_{\mathcal{M}}(\nu, y) \geq \alpha\}$ are closed in $\text{conv}(\nu)$. If $\alpha = 0$, $\{y \in \text{conv}(\nu) : \text{ZD}_{\mathcal{M}}(\nu, y) \geq \alpha\} = \text{conv}(\nu)$, and $\text{conv}(\nu)$ is closed in itself. If $\alpha > 0$, note that we can rewrite $\{y \in \text{conv}(\nu) : \text{ZD}_{\mathcal{M}}(\nu, y) \geq \alpha\} = \{y \in \text{conv}(\nu) : y \in D_{\alpha}^{\mathcal{M}}(\nu)\}$, since on the one hand, if $y \in D_{\alpha}^{\mathcal{M}}(\nu)$, then $\text{ZD}_{\mathcal{M}}(\nu, y) = \sup\{\beta : y \in D_{\beta}^{\mathcal{M}}(\nu)\} \geq \alpha$, and on the other hand, if $\text{ZD}_{\mathcal{M}}(\nu, y) = \beta \geq \alpha$, then $y \in D_{\beta}^{\mathcal{M}}(\nu) \subseteq D_{\alpha}^{\mathcal{M}}(\nu)$ by nestedness of the manifold zonoid trimmed regions. For each $\alpha > 0$, by construction $D_{\alpha}^{\mathcal{M}}(\nu)$ is closed, therefore $\{y \in \text{conv}(\nu) : \text{ZD}_{\mathcal{M}}(\nu, y) \geq \alpha\}$ is also closed.

Lower-semicontinuity: the map is lower-semicontinuous if and only if for each $\alpha \in [0, 1]$ the sets $\{y \in \text{conv}(\nu) : \text{ZD}_{\mathcal{M}}(\nu, y) \leq \alpha\}$ are closed in $\text{conv}(\nu)$ or equivalently the sets $\{y \in \text{conv}(\nu) : \text{ZD}_{\mathcal{M}}(\nu, y) > \alpha\}$ are open in $\text{conv}(\nu)$. If $\alpha = 1$, $\{y \in \text{conv}(\nu) : \text{ZD}_{\mathcal{M}}(\nu, y) > \alpha\} = \emptyset$, and the empty set is open in $\text{conv}(\nu)$. If $\alpha = 0$, $\{y \in \text{conv}(\nu) : \text{ZD}_{\mathcal{M}}(\nu, y) > \alpha\} = \text{conv}(\nu)$ by Lemma 3.2, and $\text{conv}(\nu)$ is open in itself. If $0 < \alpha < 1$, note that we can rewrite $\{y \in \text{conv}(\nu) : \text{ZD}_{\mathcal{M}}(\nu, y) > \alpha\} = \{y \in \text{conv}(\nu) : y \in D_{\alpha+}^{\mathcal{M}}(\nu)\}$, where,

$$D_{\alpha+}^{\mathcal{M}}(\nu) := \left\{ y \in \mathcal{M} : y = \text{Exp}_y \left(\int_{\mathcal{M}} \text{Log}_y(x) g(x) \nu(dx) \right), g : \mathcal{M} \rightarrow [0, 1/\alpha), \int_{\mathcal{M}} g(x) \nu(dx) = 1 \right\}$$

with g measurable. To see that the set-equivalence is true: on the one hand, if $y \in D_{\alpha+}^{\mathcal{M}}(\nu)$, then $\text{ZD}_{\mathcal{M}}(\nu, y) = \sup\{\beta : y \in D_{\beta}^{\mathcal{M}}(\nu)\} > \alpha$, since $[0, 1/\alpha) \subset [0, 1/\alpha]$. On the other hand, if $\text{ZD}_{\mathcal{M}}(\nu, y) = \beta > \alpha$, take $\epsilon > 0$ sufficiently small such that $\beta > \beta - \epsilon > \alpha$, then $[0, \frac{1}{\beta}] \subset [0, \frac{1}{\beta - \epsilon}] \subset [0, \frac{1}{\alpha}]$. As a consequence, $y \in D_{\beta}^{\mathcal{M}}(\nu) \subseteq D_{\alpha+}^{\mathcal{M}}(\nu)$ by nestedness of the manifold zonoid trimmed regions. For $0 < \alpha < 1$, distinguish between two cases: (i) $D_{\alpha+}^{\mathcal{M}}(\nu) = \text{conv}(\nu)$, then the set is open as $\text{conv}(\nu)$ is open in itself, (ii) $D_{\alpha+}^{\mathcal{M}}(\nu) \subset \text{conv}(\nu)$. In this case, writing $r\partial D_{\alpha+}^{\mathcal{M}}(\nu)$ for the relative boundary of the geodesic convex set $D_{\alpha+}^{\mathcal{M}}(\nu)$ in $\text{conv}(\nu)$, we note that $r\partial D_{\alpha+}^{\mathcal{M}}(\nu) = r\partial D_{\alpha}^{\mathcal{M}}(\nu)$. Here, the relative boundary of $D_{\alpha}^{\mathcal{M}}(\nu)$ is characterized by those points in $D_{\alpha}^{\mathcal{M}}(\nu)$ for which the weighting function g attains the maximum value $\frac{1}{\alpha}$. Since $D_{\alpha+}^{\mathcal{M}}(\nu) \cap r\partial D_{\alpha+}^{\mathcal{M}}(\nu) = D_{\alpha+}^{\mathcal{M}}(\nu) \cap r\partial D_{\alpha}^{\mathcal{M}}(\nu) = \emptyset$, it follows that $D_{\alpha+}^{\mathcal{M}}(\nu)$ is open. By combining the above arguments, we conclude that $\{y \in \text{conv}(\nu) : \text{ZD}_{\mathcal{M}}(\nu, y) > \alpha\}$ is open for each $\alpha \in [0, 1]$.

Since the map $y \mapsto \text{ZD}_{\mathcal{M}}(\nu, y)$ is both upper- and lower-semicontinuous on $\text{conv}(\nu)$ it is also continuous on $\text{conv}(\nu)$. $\square$

7.3.2 Uniform continuity in ν (P.6)

Proof. Pointwise convergence of depths: first, we show pointwise convergence of $\text{ZD}_{\mathcal{M}}(\nu_n, y)$ to $\text{ZD}_{\mathcal{M}}(\nu, y)$ for each $y \in \text{rint}(\text{conv}(\nu))$, where $\text{rint}(\text{conv}(\nu))$ denotes the relative interior of the geodesic convex set $\text{conv}(\nu)$. We note that $y \in \text{rint}(\text{conv}(\nu))$ if and only if $\mathbf{0}_{d \times d} \in \text{rint}(\text{conv}_{\mathcal{T}_y(\mathcal{M})}(\zeta_y))$, where $\text{conv}_{\mathcal{T}_y(\mathcal{M})}(\zeta_y)$ is the convex hull of the support of ζ_y in $\mathcal{T}_y(\mathcal{M})$ as in the proof of Lemma 3.2. This is seen as follows: by Lemma 3.2, $y \in \text{conv}(\nu)$ if and only if $\exists \alpha > 0$, such that $y \in D_{\alpha}^{\mathcal{M}}(\nu)$, but this is equivalent to $\mathbf{0}_{d \times d} \in D_{\alpha}(\zeta_y)$ which holds if and only if $\mathbf{0}_{d \times d} \in \text{conv}_{\mathcal{T}_y(\mathcal{M})}(\zeta_y)$ by (Mosler, 2002, Theorem 3.13). Since the sets $\{y : y \in \text{conv}(\nu)\}$ and $\{y : \mathbf{0}_{d \times d} \in \text{conv}_{\mathcal{T}_y(\mathcal{M})}(\zeta_y)\}$ are equivalent their relative interiors are equivalent as well. By Definition 3.1, $\text{ZD}_{\mathcal{M}}(\nu_n, y) = \text{ZD}_{\mathbb{R}^{d^2}}(\zeta_y^n, \vec{0})$, where ζ_y^n is the distribution of $\text{Log}_y(X)$ as a d^2 -dimensional real basis component vector, with $X \sim \nu_n$, such that $\zeta_y^n \xrightarrow{w} \zeta_y$. Similarly, $\text{ZD}_{\mathcal{M}}(\nu, y) = \text{ZD}_{\mathbb{R}^{d^2}}(\zeta_y, \vec{0})$. By the same argument as in the proof of Lemma 3.2, we know that $\zeta_y^n, \zeta_y \in P_1(\mathcal{T}_y(\mathcal{M}))$ for each $n \in \mathbb{N}$, where $P_1(\mathcal{T}_y(\mathcal{M}))$ denotes the set of probability measures on $\mathcal{T}_y(\mathcal{M})$ with finite first moment, i.e. if $\zeta \in P_1(\mathcal{T}_y(\mathcal{M}))$ then $\int_{\mathcal{T}_y(\mathcal{M})} \|z\|_y d\zeta(z) < \infty$. Furthermore, the sequence of measures $(\zeta_y^n)_{n \in \mathbb{N}}$ is uniformly integrable with respect to the Riemannian metric in $\mathcal{T}_y(\mathcal{M})$, since for any $y \in \mathcal{M}$,

$$\lim_{K \rightarrow \infty} \sup_{n \in \mathbb{N}} \int_{\mathcal{T}_y(\mathcal{M})} \|z\|_y \mathbf{1}_{\{\|z\|_y > K\}} \zeta_y^n(dz) = \lim_{K \rightarrow \infty} \sup_{n \in \mathbb{N}} \int_{\mathcal{M}} \delta(y, x) \mathbf{1}_{\{\delta(y, x) > K\}} \nu_n(dx) = 0$$

By (Mosler, 2002, Theorem 4.6), under these conditions, for $y \in \text{rint}(\text{conv}(\nu))$ or equivalently $\mathbf{0}_{d \times d} \in \text{rint}(\text{conv}_{\mathcal{T}_y(\mathcal{M})}(\zeta_y))$, it follows that,

$$\text{ZD}_{\mathcal{M}}(\nu_n, y) = \text{ZD}_{\mathbb{R}^{d^2}}(\zeta_y^n, \vec{0}) \rightarrow \text{ZD}_{\mathbb{R}^{d^2}}(\zeta_y, \vec{0}) = \text{ZD}_{\mathcal{M}}(\nu, y), \quad \text{as } n \rightarrow \infty \quad (7.2)$$

Uniform convergence of depths: uniform depth convergence now follows from the pointwise depth convergence above by a generalized version of the proof of (Dyckerhoff, 2016, Theorem 4.8) for the complete metric space $(\mathcal{M}, \delta)$, using Lemma 3.2 and the fact that $\text{ZD}_{\mathcal{M}}(\nu, y)$ is a *normed* geodesically convex depth, continuous in y by the first part of Theorem 3.3. Since the proof is completely analogous to the proof of (Dyckerhoff, 2016, Theorem 4.8), we omit the details here. Note that the only required modification is to replace the Euclidean metric space by the complete metric space $(\mathcal{M}, \delta)$. In particular, Euclidean open balls, convex sets and convergence are replaced by geodesic open balls, geodesic convex sets and convergence in the Riemannian distance function respectively.

By the generalized proof of (Dyckerhoff, 2016, Theorem 4.8), the depths $(\text{ZD}_{\mathcal{M}}(\nu_n, y_0))_{n \in \mathbb{N}}$ are continuously convergent for $y_0 \in \text{rint}(\text{conv}(\nu))$. That is, for $y_n \rightarrow y_0$ in the sense that $\delta(y_n, y_0) \rightarrow 0$, also $\lim_{n \rightarrow \infty} \text{ZD}_{\mathcal{M}}(\nu_n, y_n) = \text{ZD}(\nu, y_0)$. By (Dyckerhoff, 2016, Proposition A.1), since $\mathcal{M}$ is a metric

space, continuous convergence of the depths implies compact convergence, i.e. for every compact set $M \subseteq \text{rint}(\text{conv}(\nu))$,

$$\lim_{n \rightarrow \infty} \sup_{y \in M} |\text{ZD}_{\mathcal{M}}(\nu_n, y) - \text{ZD}_{\mathcal{M}}(\nu, y)| = 0$$

Consequently, by (Dyckerhoff, 2016, Theorem 4.4), compact convergence implies uniform convergence, since the arguments in the proof of (Dyckerhoff, 2016, Theorem 4.4) continue to hold for the manifold zonoid depth defined on the complete metric space $\mathcal{M}$, where closed and bounded subsets are compact. $\square$

7.4 Proof of Theorem 3.4 and Proposition 3.5

Proof. Properties **P.1**–**P.4** follow directly by Theorem 3.1, using the definition of the depth as the integrated pointwise zonoid depth (integrated over $t \in \mathcal{I}$).

For the first part (**P.5**) of Proposition 3.5: using that $\sup_{t \in \mathcal{I}} \delta(y_n(t), y(t)) \rightarrow 0$, by the first part of Theorem 3.3, $\text{ZD}_{\mathcal{M}}(\nu(t), y_n(t)) \rightarrow \text{ZD}_{\mathcal{M}}(\nu(t), y(t))$ uniformly over $t \in \mathcal{I}$. By definition of the integrated manifold zonoid depth also,

$$|\text{iZD}_{\mathcal{M}}(\nu, y_n) - \text{iZD}_{\mathcal{M}}(\nu, y)| \leq \int_{\mathcal{I}} |\text{ZD}_{\mathcal{M}}(\nu(t), y_n(t)) - \text{ZD}_{\mathcal{M}}(\nu(t), y(t))| dt \rightarrow 0$$

by the pointwise convergence and the fact that the depth function $\text{ZD}_{\mathcal{M}}(\cdot, \cdot) \in [0, 1]$ is bounded.

For the second part (**P.6**) in Proposition 3.5: under the given assumptions, by the second part of Theorem 3.3,

$$\sup_{y(t) \in \text{rint}(\text{conv}(\nu(t)))} |\text{ZD}_{\mathcal{M}}(\nu_n(t), y(t)) - \text{ZD}_{\mathcal{M}}(\nu(t), y(t))| \rightarrow 0, \quad \text{uniformly for } t \in \mathcal{I}$$

and similarly as above,

$$\begin{aligned} \sup_{y \in \text{rint}(\text{conv}(\nu))} |\text{iZD}_{\mathcal{M}}(\nu_n, y) - \text{iZD}_{\mathcal{M}}(\nu, y)| &\leq \\ \sup_{y \in \text{rint}(\text{conv}(\nu))} \int_{\mathcal{I}} |\text{ZD}_{\mathcal{M}}(\nu_n(t), y(t)) - \text{ZD}_{\mathcal{M}}(\nu(t), y(t))| dt &\rightarrow 0 \end{aligned}$$

using the pointwise convergence and the fact that the depth function $\text{ZD}_{\mathcal{M}}(\cdot, \cdot) \in [0, 1]$ is bounded. $\square$

7.5 Proof of Theorem 3.6

Proof. **P.1** This follows directly from the definition of the depth by the fact that the map $x \mapsto a * x$ with $a \in \text{GL}(d, \mathbb{C})$ is distance preserving, i.e. $\delta(a * x, a * y) = \delta(x, y)$ for each $x, y \in \mathcal{M}$.

P.2 Since $\int_{\mathcal{M}} \delta(y, x) \nu(dx) \geq 0$ and $\exp(-z)$ is strictly decreasing in $z \geq 0$, the point of maximum

depth is attained at $y = \arg \min_{z \in \text{supp}(\nu)} \int_{\mathcal{M}} \delta(z, x) \nu(dx)$. By eq.(2.5) in the main document, on the Riemannian manifold $\mathcal{M}$ with $\nu \in P_1(\mathcal{M})$, this point is the uniquely existing geometric median of the distribution ν .

P.3 By the proof of (Fletcher et al., 2009, Theorem 1) and an application of Leibniz's integral rule, $y \mapsto \mathbf{E}_\nu[\delta(y, X)]$ is a (strictly) convex function, and by **P.2** it attains its unique minimum at $m := \text{GM}\nu(X)$. This implies that $\mathbf{E}_\nu[\delta(\text{Exp}_m(th), X)]$ is a nondecreasing function for $t \geq 0$, where $\text{Exp}_m(th)$ is a geodesic curve emanating from m with unit tangent vector h . As a consequence $\text{GDD}(\nu, \text{Exp}_m(th)) = \exp(-\mathbf{E}_\nu[\delta(\text{Exp}_m(th), X)])$ is monotone non-increasing for $t \geq 0$.

P.4 Let $(y_n)_{n \in \mathbb{N}}$ be an unbounded sequence such that $\|\text{Log}(y_n)\|_F \rightarrow \infty$ as $n \rightarrow \infty$, then also $\delta(y_n, x) = \|\text{Log}(x^{-1/2} * y_n)\|_F \rightarrow \infty$ for each $x \in \mathcal{M}$, and as a consequence $\text{GDD}(\nu, y_n) = \exp(-\mathbf{E}_\nu[\delta(y_n, X)]) \rightarrow 0$. $\square$

7.6 Proof of Theorem 3.7

7.6.1 Continuity in y (P.5)

Proof. First, suppose that $(y_n)_{n \in \mathbb{N}}$ is an unbounded sequence $\|\text{Log}(y_n)\|_F \rightarrow \infty$ as $n \rightarrow \infty$, i.e. $y_n \rightarrow y$, where y is a singular matrix. Since $\text{GDD}(\nu, y) = 0$, by **P.4** in Theorem 3.6, $\lim_{n \rightarrow \infty} \text{GDD}(\nu, y_n) = \text{GDD}(\nu, y)$. Second, suppose that $(y_n)_{n \in \mathbb{N}}$ is a bounded sequence, i.e. $\sup_{n \in \mathbb{N}} \|\text{Log}(y_n)\|_F = \sup_{n \in \mathbb{N}} \delta(y_n, \text{Id}) < \infty$. Since $\nu \in P_1(\mathcal{M})$, there exists an $y_0 \in \mathcal{M}$ such that $\int_{\mathcal{M}} \delta(y_0, x) \nu(dx) < \infty$. By the triangle inequality,

$$\int_{\mathcal{M}} \sup_{n \in \mathbb{N}} \delta(y_n, x) \nu(dx) \leq \sup_{n \in \mathbb{N}} \delta(y_n, \text{Id}) + \delta(\text{Id}, y_0) + \int_{\mathcal{M}} \delta(y_0, x) \nu(dx) < \infty \quad (7.3)$$

using that $\delta(y_0, \text{Id}) < \infty$ as both Id and y_0 are elements of $\mathcal{M}$, (see (Bhatia, 2009, Theorem 6.1.6)). We show continuity directly from the definition of the geodesic distance depth. The function $z \mapsto \exp(-z)$ is continuous in z , also the function $z \mapsto \delta(z, x)$ is continuous in z , since $\delta(z, x) = \|\text{Log}(x^{-1/2} * z)\|_F$ is a composition of continuous functions. Furthermore, by the dominated convergence theorem, $\lim_{n \rightarrow \infty} \int_{\mathcal{M}} \delta(y_n, x) \nu(dx) = \int_{\mathcal{M}} \lim_{n \rightarrow \infty} \delta(y_n, x) \nu(dx)$, since $\int_{\mathcal{M}} \sup_{n \in \mathbb{N}} \delta(y_n, x) \nu(dx) < \infty$. Combining these arguments, it follows that $\lim_{n \rightarrow \infty} \text{GDD}(\nu, y_n) = \text{GDD}(\nu, y)$. $\square$

7.6.2 Uniform continuity in ν (P.6)

Proof. We start by noting that the uniform integrability condition implies in particular that $\nu_n \in P_1(\mathcal{M})$ for each $n \in \mathbb{N}$. Also, since $z \mapsto \delta(y, z)$ is continuous in z , by the continuous mapping theorem $\delta(y, X_n) \xrightarrow{d} \delta(y, X)$, with $X_n \sim \nu_n$ and $X \sim \nu$, and by Vitali's convergence theorem $\int_{\mathcal{M}} \delta(y, x) \nu_n(dx) \rightarrow \int_{\mathcal{M}} \delta(y, x) \nu(dx)$ for any $y \in \mathcal{M}$. Note that the convergence implies in

particular also that $\nu \in P_1(\mathcal{M})$. For two measures $\mu, \nu \in P_1(\mathcal{M})$ define their L^1 -Wasserstein distance as:

$$W_1(\mu, \nu) = \inf_{\gamma \in \Gamma(\mu, \nu)} \int_{\mathcal{M} \times \mathcal{M}} \delta(y, x) \gamma(dy, dx)$$

where $\Gamma(\mu, \nu)$ denotes the collection of all probability measures on $\mathcal{M} \times \mathcal{M}$ with marginal measures μ and ν . Substituting $\mu = \delta_y$, the point measure in y , it follows that $W_1(\delta_y, \nu) = \int_{\mathcal{M}} \delta(y, x) \nu(dx)$. Therefore, a sufficient condition to ensure uniform convergence in $y \in \mathcal{M}$ of $\int_{\mathcal{M}} \delta(y, x) \nu_n(dx)$ to $\int_{\mathcal{M}} \delta(y, x) \nu(dx)$, is $W(\nu_n, \nu) \rightarrow 0$, since

$$\sup_{y \in \mathcal{M}} \left| \int_{\mathcal{M}} \delta(y, x) \nu_n(dx) - \int_{\mathcal{M}} \delta(y, x) \nu(dx) \right| = \sup_{y \in \mathcal{M}} |W_1(\delta_y, \nu_n) - W_1(\delta_y, \nu)| \leq W_1(\nu_n, \nu) \quad (7.4)$$

where the last step follows by the reverse triangle inequality for the L^1 -Wasserstein distance. The manifold $\mathcal{M}$ is a complete separable metric space, and therefore by (Villani, 2009, Theorem 6.9) a necessary and sufficient condition for $W_1(\nu_n, \nu) \rightarrow 0$ is that the sequence of probability measures ν_n converges weakly in $P_1(\mathcal{M})$ to ν , i.e. (i) $\nu_n \xrightarrow{w} \nu$ and (ii) $\int_{\mathcal{M}} \delta(y, x) \nu_n(dx) \rightarrow \int_{\mathcal{M}} \delta(y, x) \nu(dx)$ for any $y \in \mathcal{M}$. Condition (i) holds by assumption, and condition (ii) has already been shown above.

The function $z \rightarrow \exp(-z)$ is uniformly continuous for $z \geq 0$, therefore the uniform convergence of the geodesic distance depth follows as well since,

$$\sup_{y \in \mathcal{M}} |\text{GDD}(\nu_n, y) - \text{GDD}(\nu, y)| = \sup_{y \in \mathcal{M}} |\exp(-\mathbf{E}_{\nu_n}[\delta(y, X)]) - \exp(-\mathbf{E}_{\nu}[\delta(y, X)])| \xrightarrow{n \rightarrow \infty} 0$$

□

7.7 Proof of Theorem 3.8 and Proposition 3.9

Proof. Properties **P.1–P.4** follow directly by the pointwise depth properties in Theorem 3.6, using the definition of the depth in terms of the integrated Riemannian distance (integrated over $t \in \mathcal{I}$).

For the first part (**P.5**) of Proposition 3.9: using that $\sup_{t \in \mathcal{I}} (\delta(y_n(t), y(t)) \rightarrow 0$, by the first part of the proof in Theorem 3.7 also,

$$\sup_{t \in \mathcal{I}} |\mathbf{E}_{\nu(t)}[\delta(y_n(t), X)] - \mathbf{E}_{\nu(t)}[\delta(y(t), X)]| \xrightarrow{n \rightarrow \infty} 0$$

and as a direct consequence $\lim_{n \rightarrow \infty} \int_{\mathcal{I}} \mathbf{E}_{\nu(t)}[\delta(y_n(t), X)] dt = \int_{\mathcal{I}} \mathbf{E}_{\nu(t)}[\delta(y(t), X)] dt$. Using again that $z \mapsto \exp(-z)$ is continuous in z , the composition converges as well and we conclude that $\lim_{n \rightarrow \infty} \text{iGDD}(\nu, y_n) = \text{iGDD}(\nu, y)$.

For the second part (**P.6**) of Proposition 3.9. Denote by $\xi_{n,y}(t)$ and $\xi_y(t)$ respectively the distributions of $\delta(y(t), X_n(t))$ and $\delta(y(t), X(t))$, such that $X_n(t) \sim \nu_n(t)$ and $X(t) \sim \nu(t)$. Let $\phi : \mathbb{R} \rightarrow \mathbb{R}$

be a continuous and bounded function and write $y \in \mathcal{M}$ for a curve with $y(t) \in \mathcal{M}$ for each $t \in \mathcal{I}$. Then for any curve $y \in \mathcal{M}$,

$$\sup_{t \in \mathcal{I}} |\mathbf{E}_{\xi_{n,y}(t)}[\phi(X)] - \mathbf{E}_{\xi_y(t)}[\phi(X)]| = \sup_{t \in \mathcal{I}} |\mathbf{E}_{\nu_n(t)}[\phi(\delta(y(t), X))] - \mathbf{E}_{\nu(t)}[\phi(\delta(y(t), X))]| \xrightarrow{n \rightarrow \infty} 0$$

where the last step follows by the fact that for each $t \in \mathcal{I}$ the composition $x \mapsto \phi(\delta(y(t), x))$ is again a continuous and bounded function, and the fact that $\nu_n(t) \xrightarrow{w} \nu(t)$ uniformly in t . Thus, for any curve $y \in \mathcal{M}$, the weak convergence $\xi_{n,y}(t) \xrightarrow{w} \xi_y(t)$ holds as well uniformly in t . By the uniform integrability of $(\nu_n(t))_{n \in \mathbb{N}}$ uniformly in t , combined with Vitali's convergence theorem, it follows that for each curve $y \in \mathcal{M}$,

$$\sup_{t \in \mathcal{I}} |\mathbf{E}_{\nu_n(t)}[\delta(y(t), X)] - \mathbf{E}_{\nu(t)}[\delta(y(t), X)]| \rightarrow 0 \quad \text{as } n \rightarrow \infty \quad (7.5)$$

By the same argument as in the second part of the proof of Theorem 3.7, a sufficient condition for uniform convergence in $y \in \mathcal{M}$ of $\int_{\mathcal{I}} \mathbf{E}_{\nu_n(t)}[\delta(y(t), X)] dt$ to $\int_{\mathcal{I}} \mathbf{E}_{\nu(t)}[\delta(y(t), X)] dt$ is the condition $\sup_{t \in \mathcal{I}} W_1(\nu_n(t), \nu(t)) \rightarrow 0$. Again by (Villani, 2009, Theorem 6.9), the convergence $\sup_{t \in \mathcal{I}} W_1(\nu_n(t), \nu(t)) \rightarrow 0$ is implied by the conditions (i) $\nu_n(t) \xrightarrow{w} \nu(t)$ uniformly in t , which holds by assumption and (ii) the convergence in eq.(7.5) pointwise in $y \in \mathcal{M}$.

The function $z \rightarrow \exp(-z)$ is uniformly continuous for $z \geq 0$, therefore the uniform convergence of the integrated geodesic distance depth follows as well,

$$\begin{aligned} \sup_{y \in \mathcal{M}} |\text{iGDD}(\nu_n, y) - \text{iGDD}(\nu, y)| &= \\ \sup_{y \in \mathcal{M}} \left| \exp \left(- \int_{\mathcal{I}} \mathbf{E}_{\nu_n(t)}[\delta(y(t), X)] dt \right) - \exp \left(- \int_{\mathcal{I}} \mathbf{E}_{\nu(t)}[\delta(y(t), X)] dt \right) \right| &\xrightarrow{n \rightarrow \infty} 0 \end{aligned}$$

□

7.8 Proof of Proposition 4.1

Proof. The derivation of the exact null distribution of $T_{1,N}$ is the same as for the classical univariate Wilcoxon rank-sum test, see e.g. (Bagdonavicius et al., 2011, Section 4.5) or (Mosler, 2002, Section 5.2). To derive the asymptotic null distribution, assume that the null hypothesis holds. Let $\nu := \nu_1 = \nu_2$ and,

$$T_1 = \sum_{i=1}^{n_1} r(x_{1i})$$

where $r(x_{1i})$ denotes the rank of the population depth $d(x_{1i}) := D(\nu, x_{1i})$ in the pooled sample of population depth values $\{d(x_{11}), \dots, d(x_{1n_1}), d(x_{21}), \dots, d(x_{2n_2})\}$. We can rewrite the Wilcoxon rank-sum test statistic T_1 in terms of the Mann-Whitney test statistic $U = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \mathbf{1}_{\{d(x_{1i}) > d(x_{2j})\}}$ as in the univariate case (see e.g. (Bagdonavicius et al., 2011, Section 4.5.1)),

$$T_1 = U + \frac{n_1(n_1 + 1)}{2}$$

Recalling that $N = n_1 + n_2$, by the proof of (Bagdonavicius et al., 2011, Theorem 4.7), as $N \rightarrow \infty$ and $n_1/N \rightarrow p \in (0, 1)$, under the null hypothesis,

$$\left(\frac{n_1 n_2 (N + 1)}{12} \right)^{-1/2} \cdot \left(U - \frac{n_1 n_2}{2} \right) \xrightarrow{d} N(0, 1) \quad (7.6)$$

Let us write $U_N := \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \mathbf{1}_{\{d_N(x_{1i}) > d_N(x_{2j})\}}$, with $d_N(z)$ the empirical depth as in Section 4.1.1, such that $T_{1,N} = U_N + \frac{n_1(n_1+1)}{2}$. Since $\nu = \nu_1 = \nu_2 \in P_2(\mathcal{M})$, it follows that $\nu_N \xrightarrow{w} \nu$ as $N \rightarrow \infty$, where ν_N denotes the empirical probability measure of the pooled sample $\{x_{11}, \dots, x_{1n_1}, x_{21}, \dots, x_{2n_2}\}$. Also, the sequence of empirical measures $(\nu_N)_{N \in \mathbb{N}}$ in $P_2(\mathcal{M})$ is uniformly integrable. By the uniform convergence in Theorem 3.3 for the manifold zonoid depth or by Theorem 3.7 for the geodesic distance depth,

$$\sup_{z \in \text{rint}(\text{conv}(\nu))} |d_N(z) - d(z)| \rightarrow 0, \quad \text{as } N \rightarrow \infty \quad (7.7)$$

Assuming (w.l.o.g.) that $P_\nu(Z \in \text{r}\partial\text{conv}(\nu)) = 0$, it follows that $(U_N - U) \xrightarrow{a.s.} 0$ by the uniform convergence above, since,

$$\begin{aligned} |U_N - U| &\leq \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} |\mathbf{1}_{\{d_N(X_{1i}) > d_N(X_{2j})\}} - \mathbf{1}_{\{d(X_{1i}) > d(X_{2j})\}}| \\ &\leq n_1 n_2 \cdot \sup_{i,j} |\mathbf{1}_{\{d_N(X_{1i}) > d_N(X_{2j})\}} - \mathbf{1}_{\{d(X_{1i}) > d(X_{2j})\}}| \end{aligned}$$

and the supremum on the right-hand side equals zero almost surely for N sufficiently large by eq.(7.7), using that $P_\nu(X_{1i} \in \text{rint}(\text{conv}(\nu))) = P_\nu(X_{2j} \in \text{rint}(\text{conv}(\nu))) = 1$ for all i, j . To derive the asymptotic distribution of $T_{1,N}$, we rewrite,

$$\begin{aligned} \left(\frac{n_1 n_2 (N + 1)}{12} \right)^{-1/2} \left(T_{1,N} - \frac{n_1 (N + 1)}{2} \right) &= \\ &= \left(\frac{n_1 n_2 (N + 1)}{12} \right)^{-1/2} \left(U - \frac{n_1 n_2}{2} \right) + \left(\frac{n_1 n_2 (N + 1)}{12} \right)^{-1/2} (U_N - U) \end{aligned}$$

The first term on the right-hand side converges to a standard normal distribution by eq.(7.6), the second term converges to zero almost surely, thus the asymptotic normality of the left-hand side follows as well by Slutsky's Lemma.

For the congruence invariance of the test statistic; by **P.1** in Theorem 3.1 and Theorem 3.6, the depth values for either the manifold zonoid depth or the geodesic distance depth are invariant under matrix congruence transformation for each $a \in \text{GL}(d, \mathbb{C})$. As a consequence, $r_N(a * z) = r_N(z)$ for each $z \in \{x_{11}, \dots, x_{1n_1}, x_{21}, \dots, x_{2n_2}\}$, and $T_{1,N}$ is invariant under matrix congruent transformation of the observations as well. $\square$

7.9 Proof of Proposition 4.2

Proof. Assume that the null hypothesis holds. Let $\nu := \nu_1 = \dots = \nu_K$ and,

$$T_2 = \frac{12}{N(N+1)} \sum_{k=1}^K n_k \left(\bar{R}_{k\cdot} - \frac{N+1}{2} \right)^2$$

where $\bar{R}_{k\cdot} = \frac{1}{n_k} \sum_{i=1}^{n_k} r(x_{ki})$, such that $r(x_{ij})$ is the rank of the population depth $d(x_{ij}) := D(\nu, x_{ij})$ in the pooled sample of population depths $\{d(x_{11}), \dots, d(x_{1n_1}), \dots, d(x_{K1}), \dots, d(x_{Kn_K})\}$. The derivation of the asymptotic null distribution of T_2 is the same as in the univariate case (see e.g. (Bagdonavicius et al., 2011, Section 4.8)), and it follows that if $n_k/N \rightarrow p_{k0} \in (0, 1)$ for each k ,

$$T_2 \xrightarrow{d} \chi_{K-1}^2 \quad (7.8)$$

By the same argument as in the proof of Proposition 4.1, using the uniform convergence in Theorem 3.3 for the manifold zonoid depth or Theorem 3.7 for the geodesic distance depth, $(T_{2,N} - T_2) \xrightarrow{a.s.} 0$. Combining this with eq.(7.8), under the null hypothesis also,

$$T_{2,N} \xrightarrow{d} \chi_{K-1}^2$$

We note that the same asymptotic null distribution of a Euclidean depth-based Kruskal-Wallis test also appears in Chenouri and Small (2012), where the depth function is assumed to satisfy the Euclidean equivalent of **P.6**.

The congruence invariance follows by the same argument as in the proof of Proposition 4.1. $\square$

7.10 Proof of Proposition 4.3

Proof. Assume that the the null hypothesis holds. Let $\nu := \nu_1 = \dots, \nu_n$ and,

$$T_3 = \frac{\sum_{i=1}^{n-1} (r(x_{i+1}) - r(x_i))^2}{\sum_{i=1}^n (r(x_i) - (n+1)/2)^2}$$

where $r(x_i)$ is the rank of the population depth $d(x_i) := D(\nu, x_i)$ with respect to the sample of population depth values $\{d(x_1), \dots, d(x_n)\}$, which are i.i.d. random variables under the null hypothesis. The derivation of the asymptotic null distribution of T_3 is the same as in the univariate case, see Bartels (1982), i.e.

$$(T_3 - 2)/\sigma_n \xrightarrow{d} N(0, 1), \quad \text{as } n \rightarrow \infty$$

with σ_n^2 as given in the proposition. For each $n \in \mathbb{N}$, let $\nu_{x,n}$ be the empirical probability measure of the observed sample $\{x_1, \dots, x_n\}$, such that $\nu_{x,n} \xrightarrow{w} \nu$ as $n \rightarrow \infty$. As in the proof of Proposition 4.1, the sequence of empirical measures $(\nu_{x,n})_{n \in \mathbb{N}}$ in $P_2(\mathcal{M})$ is uniformly integrable, and

$$\sup_{z \in \text{rint}(\text{conv}(\nu))} |d_n(z) - d(z)| \rightarrow 0, \quad \text{as } n \rightarrow \infty \quad (7.9)$$

Assume (w.l.o.g.) that $P_\nu(Z \in \text{r}\delta\text{conv}(\nu)) = 0$ and that for n sufficiently large there are no tied ranks, implying that $\sum_{i=1}^n (r(x_i) - (n+1)/2)^2 = n(n^2 - 1)/12$, see e.g. Bartels (1982). Substituting $r_n(x_i) = \sum_{j=1}^n \mathbf{1}\{d_n(x_i) \geq d_n(x_j)\}$, we write out,

$$\begin{aligned} |T_{3,n} - T_3| &= \frac{12}{n(n^2 - 1)} \sum_{i=1}^{n-1} \left[\left(\sum_{j=1}^n [\mathbf{1}\{d_n(X_{i+1}) \geq d_n(X_j)\} - \mathbf{1}\{d_n(X_i) \geq d_n(X_j)\}] \right)^2 \right. \\ &\quad \left. - \left(\sum_{j=1}^n [\mathbf{1}\{d(X_{i+1}) \geq d(X_j)\} - \mathbf{1}\{d(X_i) \geq d(X_j)\}] \right)^2 \right] \end{aligned}$$

For n sufficiently large, uniformly for $1 \leq i \leq n-1$, we have that $\sum_{j=1}^n [\mathbf{1}\{d_n(X_{i+1}) \geq d_n(X_j)\} - \mathbf{1}\{d_n(X_i) \geq d_n(X_j)\}] = \sum_{j=1}^n [\mathbf{1}\{d(X_{i+1}) \geq d(X_j)\} - \mathbf{1}\{d(X_i) \geq d(X_j)\}]$ almost surely by the uniform convergence in eq.(7.9), and we conclude that $(T_{3,n} - T_3) \xrightarrow{\text{a.s.}} 0$ as $n \rightarrow \infty$. Write $F_{T_{3,n}}(\cdot)$, $F_{T_3}(\cdot)$ and $F_{N(2, \sigma_n^2)}(\cdot)$ for the cumulative distribution functions of $T_{3,n}$, T_3 and a random variable with distribution $N(2, \sigma_n^2)$ respectively. Combining the previous arguments, for each $x \in \mathbb{R}$ at which the distribution functions are continuous,

$$\lim_{n \rightarrow \infty} |F_{T_{3,n}}(x) - F_{N(2, \sigma_n^2)}(x)| \leq \lim_{n \rightarrow \infty} |F_{T_{3,n}}(x) - F_{T_3}(x)| + \lim_{n \rightarrow \infty} |F_{T_3}(x) - F_{N(2, \sigma_n^2)}(x)| = 0$$

and we conclude that $(T_{3,n} - 2)/\sigma_n \xrightarrow{d} N(0, 1)$.

The congruence invariance follows by the same argument as in the proof of Proposition 4.1. $\square$

7.11 Proof of Proposition 4.4

Proof. In this proof, we make multiple use of the identity $\text{Tr}(\text{Log}(y^{-1/2} * x)) = \text{Tr}(\text{Log}(x)) - \text{Tr}(\text{Log}(y))$, which follows immediately from $\text{Tr}(\text{Log}(ab)) = \text{Tr}(\text{Log}(a)) + \text{Tr}(\text{Log}(b))$ and $\text{Tr}(\text{Log}(a^t)) = t\text{Tr}(\text{Log}(a))$ for any $a, b \in \mathcal{M}$ and $t \in \mathbb{R}$. First, for random variables $X_i \sim \nu_x$ and $Y_i \sim \nu_y$ with $i = 1, \dots, n$, under the null hypothesis $H_0 : \nu_x = \nu_y$, the distributions of $\text{diff}(X_i, Y_i) = \text{Tr}(\text{Log}(Y_i^{-1/2} * X_i))$ are symmetric around zero since,

$$\begin{aligned} \text{Tr}(\text{Log}(Y_i^{-1/2} * X_i)) &\stackrel{d}{=} \text{Tr}(\text{Log}(X_i^{-1/2} * Y_i)) \\ &\stackrel{d}{=} \text{Tr}(\text{Log}(Y_i)) - \text{Tr}(\text{Log}(X_i)) \\ &\stackrel{d}{=} -\text{Tr}(\text{Log}(Y_i^{-1/2} * X_i)) \end{aligned}$$

which also indirectly implies that $\mathbf{E}[\text{Tr}(\text{Log}(Y_i^{-1/2} * X_i))] = 0$ under the null hypothesis. In fact, the latter property remains true under the weaker assumption of equality of the geometric means of the distributions ν_x and ν_y , i.e. $\mathbb{E}_{\nu_x}[X_i] = \mathbb{E}_{\nu_y}[Y_i]$.

The null distribution of $T_{4,n}$ is now derived in the same way as for the classical univariate Wilcoxon signed-rank test statistic, using that the difference scores $\text{diff}(X_i, Y_i)$ are real-valued scalars following a symmetric distribution around zero, see e.g. (Bagdonavicius et al., 2011, Section 4.6).

To show that the test statistic is invariant under congruence transformation of the observations by $a \in \text{GL}(d, \mathbb{C})$ decompose,

$$\begin{aligned} \text{Tr}(\text{Log}((a * y_i)^{-1/2} * (a * x_i))) &= \text{Tr}(\text{Log}(a * x_i)) - \text{Tr}(\text{Log}(a * y_i)) \\ &= \text{Tr}(\text{Log}(x_i)) - \text{Tr}(\text{Log}(y_i)) + \text{Tr}(\text{Log}(a^* a)) - \text{Tr}(\text{Log}(a^* a)) \\ &= \text{Tr}(\text{Log}(y_i^{-1/2} * x_i)) \end{aligned}$$

Since $T_{4,n}$ depends only on the difference scores $\text{Tr}(\text{Log}((a * y_i)^{-1/2} * (a * x_i)))$ for $i = 1, \dots, n$, the invariance of the test statistic follows immediately.

Change in scale symmetric error distribution In the simulation experiment in Figure 10, we note that under a change in scale for data generated from a *symmetric* distribution, the manifold Wilcoxon signed-rank test has no power, as the distributions of the difference scores $\text{diff}(X_i, Y_i) = \text{Tr}(\text{Log}(Y_i^{-1/2} * X_i))$ remain symmetric around zero. With the same notation as in the simulation setup in Figure 10, assume that the paired samples are generated as $X_i = \text{Exp}_{\mu_i} \left(\sum_{k=1}^{d^2} z_{ki}^x e_{ki} \right)$ and $Y_i = \text{Exp}_{\mu_i} \left(\sum_{k=1}^{d^2} t z_{ki}^y e_{ki} \right)$ for $i = 1, \dots, n$, with (z_{ki}^x) and (z_{ki}^y) symmetric mean-zero i.i.d. error terms, then we verify that $\text{diff}(X_i, Y_i) \stackrel{d}{=} -\text{diff}(X_i, Y_i)$. For ease of notation, denote $Z_i^x := \sum_{k=1}^{d^2} z_{ki}^x e_{ki}$ and $Z_i^y := \sum_{k=1}^{d^2} z_{ki}^y e_{ki}$, then for each $i = 1, \dots, n$,

$$\begin{aligned} \text{Tr}(\text{Log}(Y_i^{-1/2} * X_i)) &= \log(\det(\mu_i^{1/2} * \text{Exp}(\mu_i^{-1/2} * Z_i^x))) - \log(\det(\mu_i^{1/2} * \text{Exp}(\mu_i^{-1/2} * t Z_i^y))) \\ &= \log(\det(\text{Exp}(\mu_i^{-1/2} * Z_i^x))) - \log(\det(\text{Exp}(\mu_i^{-1/2} * t Z_i^y))) \\ &\stackrel{d}{=} \log(\det(\text{Exp}(\mu_i^{-1/2} * (-Z_i^x)))) - \log(\det(\text{Exp}(\mu_i^{-1/2} * (-t Z_i^y)))) \\ &= -\log(\det(\text{Exp}(\mu_i^{-1/2} * Z_i^x))) + \log(\det(\text{Exp}(\mu_i^{-1/2} * t Z_i^y))) \\ &= -\text{Tr}(\text{Log}(Y_i^{-1/2} * X_i)) \end{aligned}$$

where we use the identity $\log(\det(\text{Exp}(-A))) = -\log(\det(\text{Exp}(A)))$ for $A \in \mathcal{T}_{\text{Id}}(\mathcal{M})$ and the fact that $Z_i^x \stackrel{d}{=} -Z_i^x$ and $Z_i^y \stackrel{d}{=} -Z_i^y$ for each $i = 1, \dots, n$. $\square$

8 Appendix II: Power analysis of rank-based tests

In the simulation experiments below, we assess the power of the proposed tests in Section 4.1–4.2 in the main document under several different scenarios. First, in Figures 5- to 10-(a) and Figures 5- to 10-(c) if existing, we consider the power at the significance level $\alpha = 0.05$ based on different variants of a *change in location*. Second, in Figures 5- to 10-(b) and Figures 5- to 10-(d) if existing, we consider the power at level $\alpha = 0.05$ based on different variants of a *change in scale*. In Figure 11, when distinguishing between the spectra of two multivariate time series, we examine two types of location shifts of a baseline spectral matrix. In Figures 5 to 10, for each replication $b = 1, \dots, B$ of the individual simulation experiments we generate a random

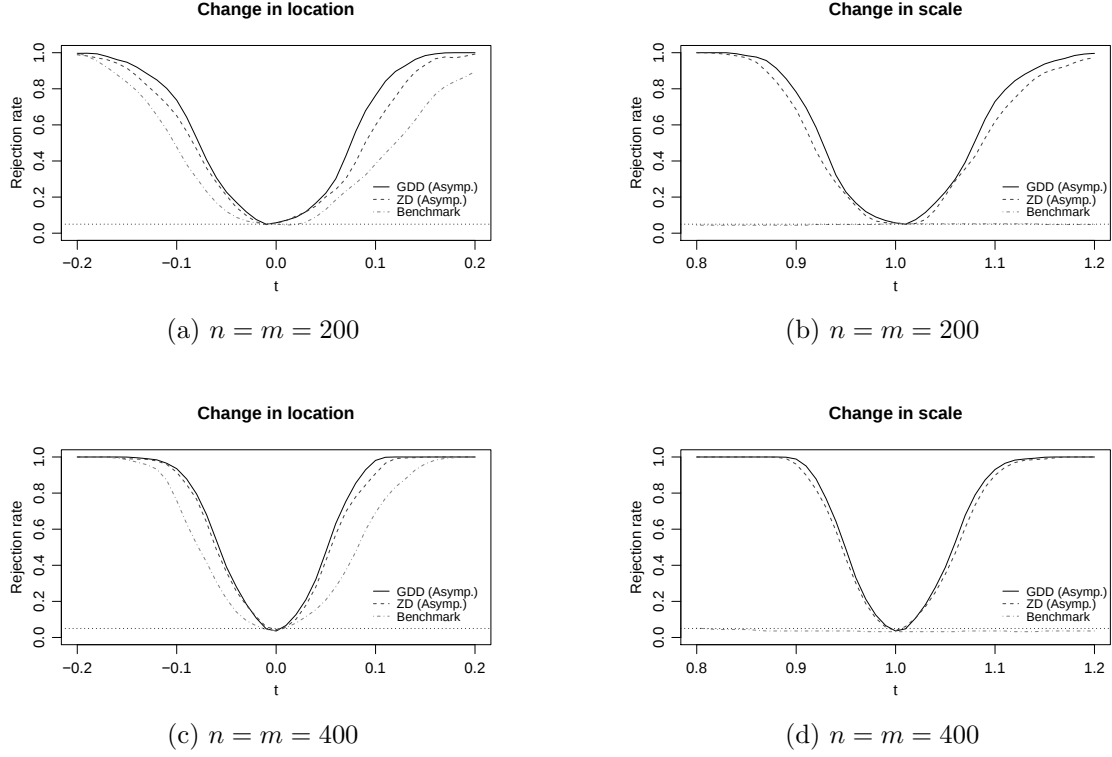


Figure 5: Rejection rates of the manifold Wilcoxon rank-sum test and the benchmark permutation test for two independent samples of (2×2) -dimensional random HPD matrices based on 500 instances for each value t and 500 permutations for the benchmark test. The horizontal (dotted) line indicates the level $\alpha = 0.05$.

location parameter $\mu_b = \text{Exp}_{\text{Id}}\left(\sum_{k=1}^{d^2} z_{kb} e_k\right)$, with $(z_{kb})_{k=1, \dots, d^2}$ i.i.d. an standard normal vector and $(e_k)_{k=1, \dots, d^2}$ an orthonormal basis of $(\mathcal{T}_{\text{Id}}(\mathcal{M}, \langle \cdot, \cdot \rangle_F))$, the space of Hermitian matrices endowed with the matrix Frobenius inner product. For notational simplicity we write $\mu := \mu_b$ and describe the data generating processes conditional on a particular value of $\mu \in \mathcal{M}$.

Manifold Wilcoxon rank-sum test (Figure 5) We consider the performance of the manifold Wilcoxon rank-sum test based on two independent random samples of HPD matrices, where the first independent sample $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \nu_x$ is distributed as $X_i = \text{Exp}_{\mu}\left(\sum_{k=1}^{d^2} z_{ki} e_{k0}\right)$ for $i = 1, \dots, n$, with $(e_{k0})_{k=1, \dots, d^2}$ an orthonormal basis of $\mathcal{T}_{\mu}(\mathcal{M})$. In the case of a change in location, given $t \in [-0.5, 0.5]$, the second independent sample $Y_1, \dots, Y_m \stackrel{\text{iid}}{\sim} \nu_y(t)$ is distributed as the first sample, but with a *location shift* of the geometric mean along a geodesic on the manifold $\mu(t) = \text{Exp}_{\mu}(t \sum_{k=1}^{d^2} e_{k0})$. That is, $Y_j = \text{Exp}_{\mu(t)}\left(\sum_{k=1}^{d^2} z_{kj} e_{k0}\right)$ for $j = 1, \dots, m$ with $(z_{kj})_{k=1, \dots, d^2}$ i.i.d. standard normal. At $t = 0$, the null hypothesis $H_0 : \nu_x = \nu_y(t)$ is true, and the rejection rate should equal $\alpha = 0.05$. In the case of a change in scale, given $t \in [0.8, 1.2]$, the second independent

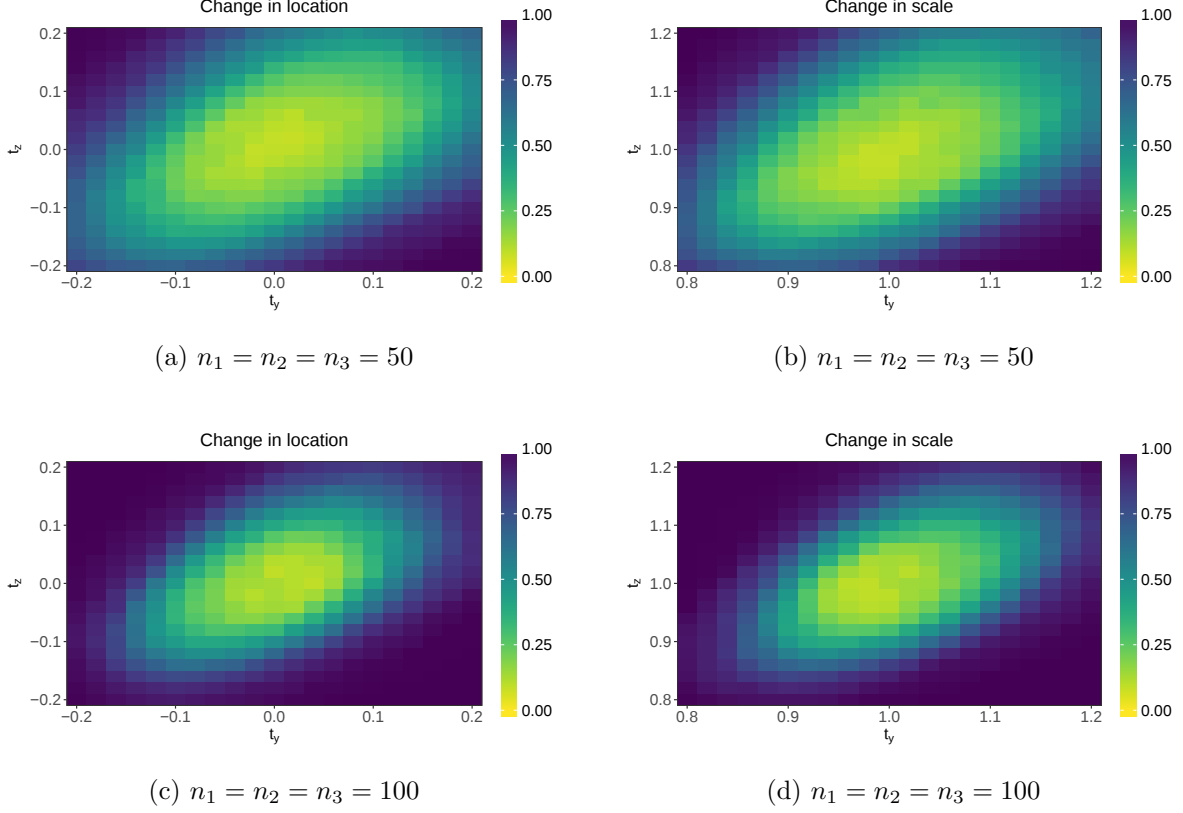


Figure 6: Rejection rates of the manifold Kruskal-Wallis test based on geodesic distance depth (GDD) for three independent samples of (2×2) -dimensional random HPD matrices based on 250 instances for each combination (t_y, t_z) .

sample $Y_1, \dots, Y_m \stackrel{\text{iid}}{\sim} \nu_y(t)$ is distributed as $Y_j = \text{Exp}_\mu\left(t \sum_{k=1}^{d^2} z_{kj} e_{k0}\right)$ for $j = 1, \dots, m$. Thus, instead of varying the underlying geometric mean, we vary the *scale* of the distribution. At $t = 1$, the null hypothesis $H_0 : \nu_x = \nu_y(t)$ is true.

As a benchmark to the manifold rank-sum test, we consider the permutation test of Pigoli et al. (2014) based on the Riemannian distance between geometric sample means, described in more detail in Section 9.1 below. In Figure 5, we observe that under a change in location, the manifold rank-sum test (for both depth functions) outperforms the benchmark procedure. Moreover, since the benchmark procedure takes into account only the geometric means of the data and not the spread of the data, the benchmark test has no power under a change in scale, whereas the manifold depth-based test *does* have power under a change in scale.

Manifold Kruskal-Wallis test (Figures 6 and 7) We consider the performance of the manifold Kruskal-Wallis test in the context of three independent samples $X_1, \dots, X_{n_1} \stackrel{\text{iid}}{\sim} \nu_x$, $Y_1, \dots, Y_{n_2} \stackrel{\text{iid}}{\sim} \nu_y(t_y)$ and $Z_1, \dots, Z_{n_3} \stackrel{\text{iid}}{\sim} \nu_z(t_z)$. The first sample is distributed as in the previous paragraph. In the

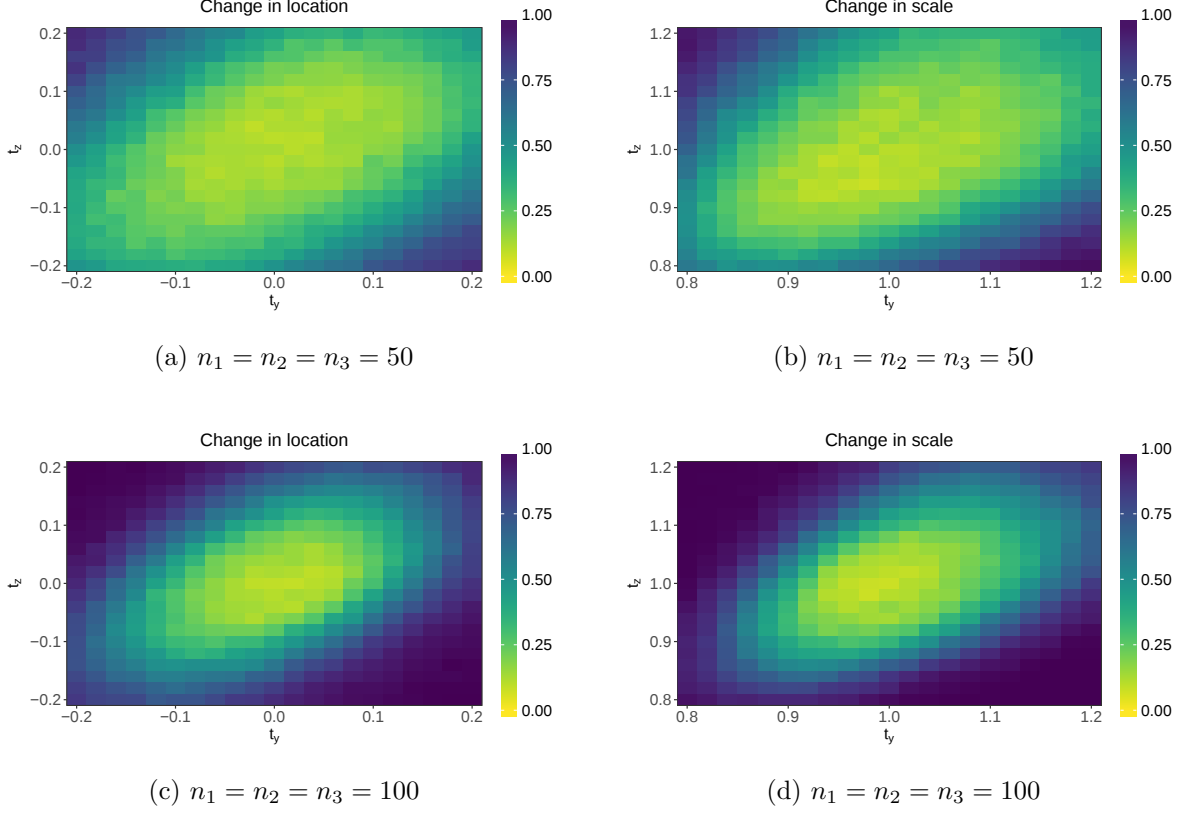


Figure 7: Rejection rates of the manifold Kruskal-Wallis test based on zonoid depth (ZD) for three independent samples of (2×2) -dimensional random HPD matrices based on 250 instances for each combination (t_y, t_z) .

case of a change in location, given $t_y, t_z \in [-0.5, 0.5]$, the second and third samples are generated according to a location shift in the underlying geometric mean along a geodesic as in the previous paragraph, with $\mu(t_y) = \text{Exp}_\mu(t_y \sum_{k=1}^{d^2} e_{k0})$ and $\mu(t_z) = \text{Exp}_\mu(t_z \sum_{k=1}^{d^2} e_{k0})$. For $j = 1, \dots, n_2$ and $\ell = 1, \dots, n_3$, we generate $Y_j = \text{Exp}_{\mu(t_y)}\left(\sum_{k=1}^{d^2} z_{kj} e_{k0}\right)$ and $Z_\ell = \text{Exp}_{\mu(t_z)}\left(\sum_{k=1}^{d^2} z_{k\ell} e_{k0}\right)$, with $(z_{kj})_k$ and $(z_{k\ell})_k$ i.i.d. standard normal. Again, at $t_y = t_x = 0$, the null hypothesis $H_0 : \nu_x = \nu_y(t_y) = \nu_z(t_z)$ is true, and the rejection rate should equal $\alpha = 0.05$. In the case of a change in scale, given $t_y, t_z \in [0.8, 1.2]$, we extend the simulation setup for the manifold rank-sum test based on a change in scale with one additional independent sample in the same way as under a change in location.

In Figures 6 and 7, we display the obtained rejection rates based on the geodesic distance depth and the manifold zonoid depth respectively. As a benchmark to the proposed manifold Kruskal-Wallis test, we consider the permutation test of Cabassi et al. (2017), which is a generalization of the two-sample test in Pigoli et al. (2014) to more than two independent samples. The test in Cabassi et al. (2017) is described in more detail in Section 9.1. The rejection rates

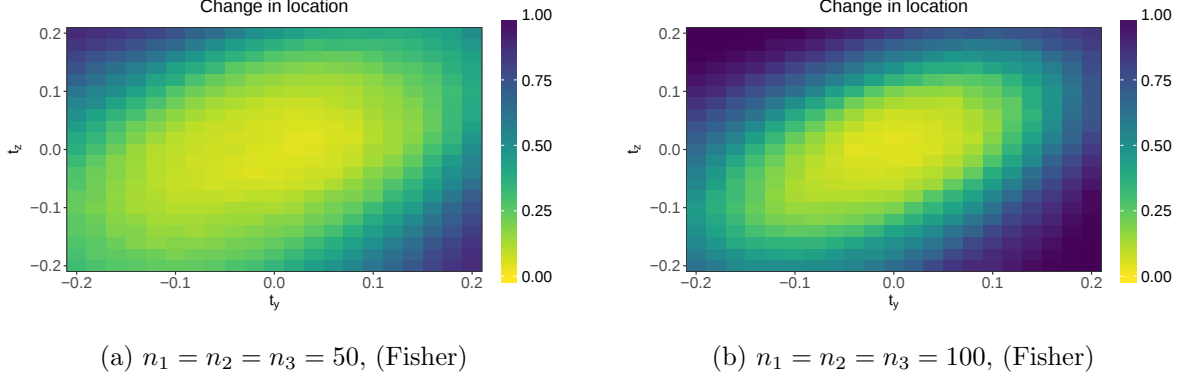


Figure 8: Rejection rates of the benchmark permutation test using Fisher's combining function for three independent samples of (2×2) -dimensional random HPD matrices based on 250 instances with 250 permutations per instance for each combination (t_y, t_z) .

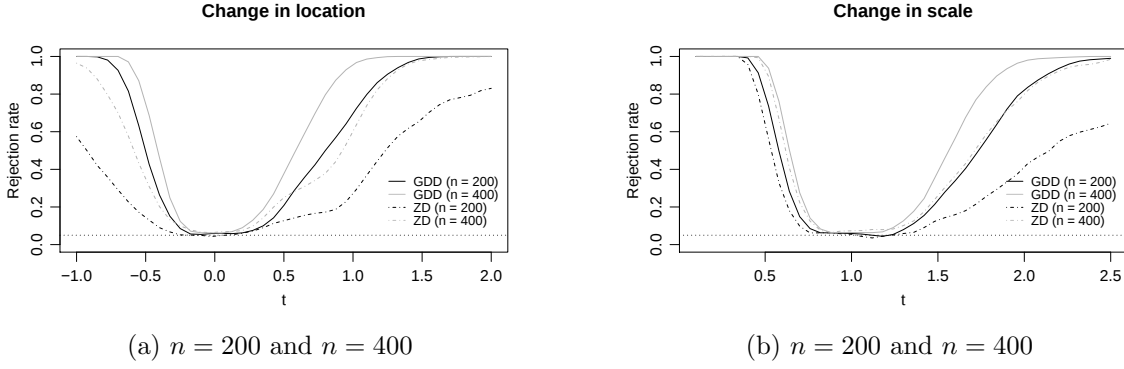


Figure 9: Rejection rates of the manifold Bartels-von Neumann test for a single independent sample of (2×2) -dimensional random HPD matrices based on 250 instances for each value t . The horizontal (dotted) line indicates the level $\alpha = 0.05$.

of the benchmark procedure under a change in location based on Fisher's combining function $\Psi(p_1, \dots, p_N) = -2 \sum_{k=1}^N \log(p_k)$ are shown in Figure 8. We also computed the rejection rates of the benchmark procedure based on a different combining function (Tippet's combining function $\Psi(p_1, \dots, p_N) = 1 - \min(\log(p_1), \dots, \log(p_N))$), but since the obtained rejection rates are practically equivalent to Figure 8 they have not been included here. Similar to the previous paragraph, the manifold Kruskal-Wallis test (for both depth functions) is able to compete with –or outperform– the benchmark procedure in terms of power. The rejection rates of the benchmark test under a change in scale are omitted, since the permutation test has no power under a change in scale for the same reason as in the previous paragraph.

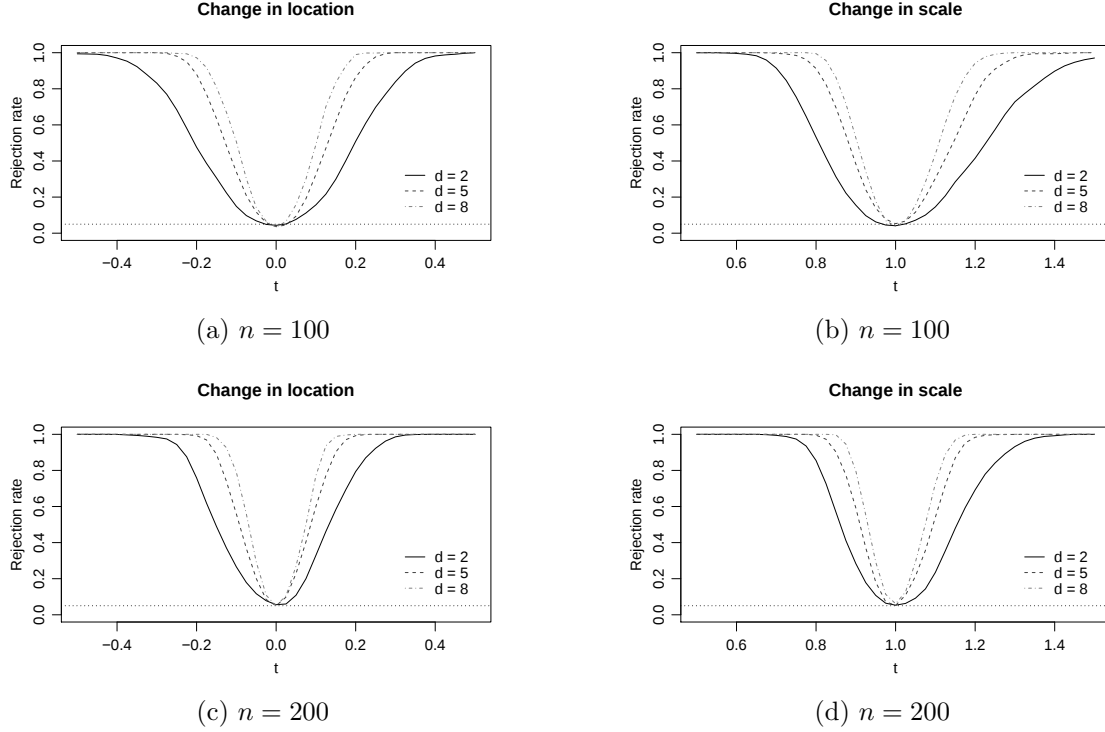


Figure 10: Rejection rates of the manifold Wilcoxon signed-rank test for paired samples of $(d \times d)$ -dimensional random HPD matrices based on 500 instances for each value t . The horizontal (dotted) line indicates the level $\alpha = 0.05$.

Manifold Bartels-von Neumann test (Figure 9) We consider the performance of the manifold Bartels-von Neumann test for a single sample $X_1, \dots, X_n$ of independent observations, with $X_i \sim \nu_i(t)$ for $i = 1, \dots, n$, where the distributions $\nu_i(t)$ change according to a trend in the location or a trend in the scale. In the case of a trend in location, given $t \in [-1, 2]$, we generate $X_i = \text{Exp}_{\mu(t_i)}\left(\sum_{k=1}^{d^2} z_{ki} e_{ki}\right)$, with $\mu(t_i) = \text{Exp}_{\mu}\left(\frac{i}{n} t \sum_{k=1}^{d^2} e_{k0}\right)$, $(z_{ki})_i$ i.i.d. standard normal and $(e_{ki})_i$ an orthonormal basis of $\mathcal{T}_{\mu(t_i)}(\mathcal{M})$. At $t = 0$, there is no trend in the location and the null hypothesis is true. For increasing values of $|t|$ the trend is more pronounced. In the case of a trend in the scale, given $t \in [0.1, 2.5]$, we generate $X_i = \text{Exp}_{\mu}\left(\left(\frac{n-i}{n} + \frac{i}{n} t\right) \sum_{k=1}^{d^2} z_{ki} e_{k0}\right)$ for $i = 1, \dots, n$, such that for $t = 1$ the null hypothesis is true, and for increasing values of $|\log(t)|$ the trend in the scale is more pronounced. We did not find potential benchmark procedures in the literature that allow for comparison of the performance of the proposed manifold test for randomness.

Manifold Wilcoxon signed-rank test (Figure 10) To study the performance of the manifold Wilcoxon signed-rank test, we consider a similar setup as for the manifold Wilcoxon rank-sum test, but with the main difference that the two equal-sized samples $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \nu_x$ and $Y_1, \dots, Y_n \stackrel{\text{iid}}{\sim} \nu_y(t)$ are no more independently generated. To introduce between-sample depen-

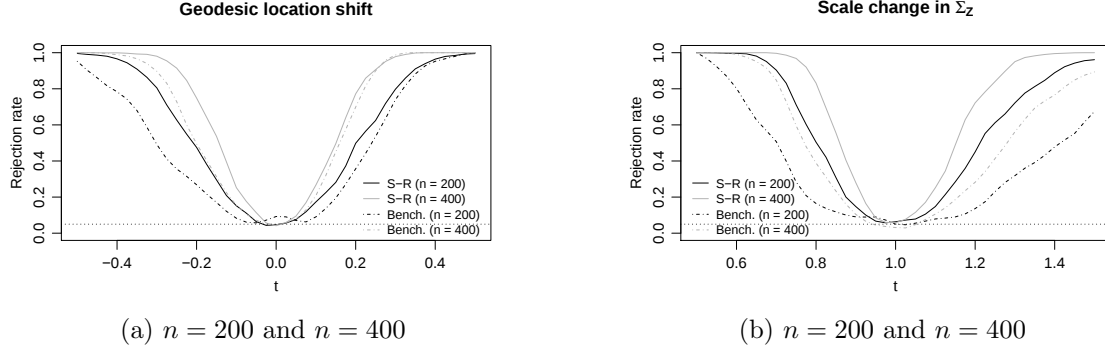


Figure 11: Rejection rates of the manifold signed-rank test (S-R) for equal spectra of two multivariate time series and the benchmark permutation test (Bench.) for a baseline (2×2) -dimensional VAR(1) process. For each value of t , based on 500 instances of the signed-rank test and 125 instances of the benchmark test with 250 permutations per instance. The horizontal (dotted) line indicates the level $\alpha = 0.05$.

dence, we incorporate a *trial-specific* random effect in the underlying geometric means. In the case of a change in location, given $t \in [-0.5, 0.5]$, the paired samples are generated according to $X_i = \text{Exp}_{\mu_i} \left(\sum_{k=1}^{d^2} z_{ki}^x e_{ki} \right)$ and $Y_i = \text{Exp}_{\mu_i(t)} \left(\sum_{k=1}^{d^2} z_{ki}^y e_{ki} \right)$ for $i = 1, \dots, n$, with $(z_{ki}^x)_k$ and $(z_{ki}^y)_k$ i.i.d. standard normal and $(e_{ki})_k$ an orthonormal basis of $\mathcal{T}_{\mu_i}(\mathcal{M})$. Here, the random trial-specific geometric means are generated as $\mu_i = \text{Exp}_{\mu} \left(\sum_{k=1}^{d^2} u_{ki} e_{k0} \right)$ for $i = 1, \dots, n$, with $(u_{ki})_k$ i.i.d. standard normal. The location shift for the second sample along a geodesic on the manifold is given by $\mu_i(t) = \text{Exp}_{\mu_i} \left(t \sum_{k=1}^{d^2} e_{ki} \right)$, such that $\mu_i(t)$ moves away from $\mu_i = \mu_i(0)$ along a geodesic for increasing values of $|t|$. In the case of a change in scale, given $t \in [0.5, 1.5]$, for $i = 1, \dots, n$ the observations in the paired samples are generated according to $X_i = \text{Exp}_{\mu_i} \left(\sum_{k=1}^{d^2} \zeta_{ki}^x e_{ki} \right)$ and $Y_i = \text{Exp}_{\mu_i} \left(\sum_{k=1}^{d^2} t \zeta_{ki}^y e_{ki} \right)$ with μ_i as before. It is important to point out that we do not consider the error terms to be normally distributed. Instead, we consider $(\zeta_{ki}^x)_k$ and $(\zeta_{ki}^y)_k$ to be i.i.d. exponentially distributed with rate $\lambda = 1$. The reason is that if the error terms are sampled from a symmetric distribution, the test has no power under the alternative of a change in scale, since the distribution of the difference scores remain symmetric around zero. This is shown at the end of the proof of Proposition 4.4 above. If the error terms are sampled from a non-symmetric (skewed) distribution, we *do* have power under the alternative of a change in scale as seen in Figure 10. Note that we did not find potential benchmark procedures in the literature that allow for performance comparison of the proposed signed-rank test for homogeneity in distribution of paired samples of HPD matrices.

Testing for equality of spectral matrices (Figure 11) To conclude, we study the performance of the manifold signed-rank test to distinguish between the spectra of two stationary multi-

variate time series. The first time series trace $\vec{X}_1, \dots, \vec{X}_{2n}$ is generated from a causal 2-dimensional vector autoregressive VAR(1) process $\vec{X}_j = \Phi_1 \vec{X}_{j-1} + \vec{Z}_j$. Here, the eigenvalues of the coefficient matrix $\Phi_1 \in \mathbb{R}^{2 \times 2}$ are such that $|\lambda_1|, |\lambda_2| < 1$, and $\vec{Z}_j$ is an independent Gaussian white noise process with innovation covariance matrix Σ_Z . In particular, we consider observations generated from a baseline VAR(1) process with parameters,

$$\Phi_1 = \begin{pmatrix} 0.5 & 0.1 \\ -0.1 & 0.5 \end{pmatrix} \quad \text{and} \quad \Sigma_Z = \begin{pmatrix} 0.5 & 0.2 \\ 0.2 & 1 \end{pmatrix}$$

The HPD spectral density matrix of a VAR(1) process at the Fourier frequencies $\omega_\ell = \pi\ell/n \in (-\pi, \pi]$ takes the form:

$$f(\omega_\ell) = \frac{1}{2\pi} \{(\text{Id} - \Phi_1 \exp(-i\omega_\ell))\}^{-1} \Sigma_Z \{(\text{Id} - \Phi_1 \exp(i\omega_\ell))^*\}^{-1} \quad (8.1)$$

In Figure 11-(a), given $t \in [-0.5, 0.5]$, we consider a geodesic location shift of the baseline spectrum of the first time series trace according to $f_t(\omega_\ell) = \text{Exp}_{f(\omega_\ell)}\left(t \sum_{k=1}^4 e_{k\ell}\right)$ for $\ell = -(n-1), \dots, n$, with $(e_{k\ell})_{k=1, \dots, 4}$ an orthonormal basis of $\mathcal{T}_{f(\omega_\ell)}(\mathcal{M})$. The shifted spectral matrices are illustrated in Figure 12 below for several different values of t . We generate the second time series trace $\vec{Y}_1, \dots, \vec{Y}_{2n}$ with underlying spectrum $f_t(\omega_\ell)$ based on the Cramér representation (Brillinger, 1981, Section 4.6), through the inverse Fourier transform:

$$\vec{Y}_j = \frac{1}{\sqrt{2n}} \sum_{\ell=-(n-1)}^n f_t^{1/2}(\omega_\ell) \exp(i\omega_\ell j) \vec{\xi}_\ell, \quad \text{for } j = 1, \dots, 2n$$

where $\vec{\xi}_\ell$ is a 2-dimensional complex standard normal vector, such that $\vec{\xi}_\ell = \vec{\xi}_{-\ell}^*$. For $\ell = \{0, n\}$, $\vec{\xi}_\ell$ is a real standard normal vector. At $t = 0$, the null hypothesis $H_0 : \{f(\omega_\ell) = f_t(\omega_\ell), \forall \ell = 1, \dots, n\}$ is true and the rejection rate should equal $\alpha = 0.05$. In Figure 11-(b), given $t \in [0.5, 1.5]$, we generate the second sample $\vec{Y}_1, \dots, \vec{Y}_{2n}$ based on a change in the scale of the innovation covariance matrix Σ_Z according to $\Sigma_Z(t) = t\Sigma_Z$, such that the null hypothesis is true for $t = 1$. In terms of the underlying spectral matrix, rescaling Σ_Z results in a (multiplicative) location shift of the baseline spectrum as seen from eq.(8.1).

As a benchmark to the proposed manifold signed-rank test, we consider the permutation test in Jentsch and Pauly (2015) based on the integrated L_2 -distance between smoothed periodograms, which is described in more detail in Section 9.2 below. The test procedure in Jentsch and Pauly (2015) requires a consistent spectral estimate, therefore we smooth the raw periodogram matrices using a Bartlett-Priestley kernel as in (Jentsch and Pauly, 2015, Section 4). For the bandwidth of the kernel, we compute an oracle bandwidth minimizing the L_2 -risk of the kernel spectral density estimator with respect to the true known spectral matrix. For the manifold signed-rank test, we only smooth the raw periodogram matrices by a minimum amount to guarantee positive definiteness of the periodogram matrices at each frequency. In particular, we compute the Bartlett spectral

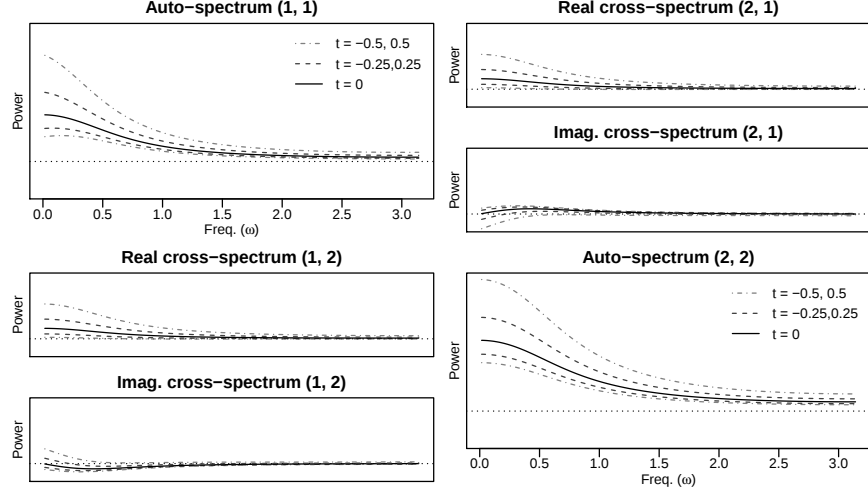


Figure 12: (2×2) -dimensional spectral matrix of second generated time series trace according to the simulation setup in Figure 11 based on a geodesic location shift of the baseline spectral matrix of a VAR(1) process for several different values of t .

estimator by averaging the (2×2) -dimensional raw periodogram matrices of $B = 2$ non-overlapping time series segments, which is a specific trivial instance of a multitaper spectral estimator with non-overlapping uniform tapering windows. From the rejection rates in Figure 11, we observe that the manifold Wilcoxon signed-rank test outperforms the benchmark permutation test in terms of power under each of the considered scenarios. This is perhaps not surprising, as both the geodesic location shift (Figure 12) and the change in scale of the innovation matrix result in multiplicative changes of the baseline spectral matrix. As such, the distance between periodograms on the Riemannian manifold is relatively large over the entire frequency range, whereas the L_2 -distance between (smoothed) periodograms is only large at frequencies with high spectral power as apparent from Figure 12.

9 Appendix III: Additional material

9.1 Benchmark tests Pigoli et al. (2014) and Cabassi et al. (2017)

As a benchmark for the performance of the manifold Wilcoxon rank-sum test, we consider the pointwise version of the two-sample permutation test of Pigoli et al. (2014) based on the Riemannian distance function and within-sample empirical geometric means. Although this permutation test is constructed in a more general functional analysis context, it remains valid in the non-functional (pointwise) setting considered here. Let $X_{11}, \dots, X_{1n_1} \stackrel{\text{iid}}{\sim} \nu_1$ and $X_{21}, \dots, X_{2n_2} \stackrel{\text{iid}}{\sim} \nu_2$ be two independent samples on the manifold $\mathcal{M}$, then we compute a p-value of the permutation test (B permutations) as follows:

1. Compute $\Sigma_1^{(0)}, \Sigma_2^{(0)}$ the within-sample geometric means of $\{x_{11}, \dots, x_{1n_1}\}$ and $\{x_{21}, \dots, x_{2n_2}\}$ and the Riemannian distance $\delta(\Sigma_1^{(0)}, \Sigma_2^{(0)})$.
2. For $b = 1, \dots, B$, randomly sample n_1 and n_2 observations without replacement from the pooled sample $\{x_{11}, \dots, x_{1n_1}, x_{21}, \dots, x_{2n_2}\}$. Compute $\Sigma_1^{(b)}, \Sigma_2^{(b)}$ the within-sample geometric means of the permuted samples and the Riemannian distance $\delta(\Sigma_1^{(b)}, \Sigma_2^{(b)})$.
3. Compute the p-value of the test as $p = \frac{1}{B} \sum_{b=1}^B \mathbf{1}\{\delta(\Sigma_1^{(b)}, \Sigma_2^{(b)}) \geq \delta(\Sigma_1^{(0)}, \Sigma_2^{(0)})\}$.

As a benchmark for the performance of the manifold Kruskal-Wallis test, we consider the pointwise version of the multi-sample permutation test of Cabassi et al. (2017) again based on the Riemannian distance function and within-sample empirical geometric means. This test essentially generalizes the two-sample test above by combining the partial p-values into a global p-value according to some combining function. For $k = 1, \dots, K$, let $X_{k1}, \dots, X_{kn_k} \stackrel{\text{iid}}{\sim} \nu_k$ be independent random samples on the manifold $\mathcal{M}$, then we compute a p-value of the permutation test (B permutations) as follows:

1. For each pair $1 \leq i < j \leq K$, compute $\delta(\Sigma_i^{(0)}, \Sigma_j^{(0)})$ with $\Sigma_i^{(0)}, \Sigma_j^{(0)}$ the within-sample geometric means of $\{x_{i1}, \dots, x_{in_i}\}$ and $\{x_{j1}, \dots, x_{jn_j}\}$.
2. For $b = 1, \dots, B$, randomly sample grouped observations of size $n_1, \dots, n_K$ without replacement from the pooled sample $\{x_{11}, \dots, x_{1n_1}, \dots, x_{K1}, \dots, x_{Kn_K}\}$. Compute $\delta(\Sigma_i^{(b)}, \Sigma_j^{(b)})$ for each $1 \leq i < j \leq K$, with $\Sigma_i^{(b)}, \Sigma_j^{(b)}$ the within-sample geometric means of the resampled groups $\{x_{i1}^{(b)}, \dots, x_{in_i}^{(b)}\}$ and $\{x_{j1}^{(b)}, \dots, x_{jn_j}^{(b)}\}$.
3. Let $\lambda_{ij}(d) = \frac{1}{B} \sum_{b=1}^B \mathbf{1}\{\delta(\Sigma_i^{(b)}, \Sigma_j^{(b)}) \geq d\}$. Compute the partial p-values of the test as $p_{ij}^{(0)} = \lambda_{ij}(\delta(\Sigma_i^{(0)}, \Sigma_j^{(0)}))$ and for each $b = 1, \dots, B$ also compute $p_{ij}^{(b)} = \lambda_{ij}(\delta(\Sigma_i^{(b)}, \Sigma_j^{(b)}))$ for $1 \leq i < j \leq K$.
4. Combine the $p_{ij}^{(0)}$ via the combining function Ψ to obtain the observed global test statistic $T_\Psi^{(0)} = \Psi(p_{1,2}^{(0)}, p_{1,3}^{(0)}, \dots, p_{K-1,K}^{(0)})$. For each $b = 1, \dots, B$, also compute the b -th combined test statistic as $T_\Psi^{(b)} = \Psi(p_{1,2}^{(b)}, p_{1,3}^{(b)}, \dots, p_{K-1,K}^{(b)})$.
5. Compute the p-value of the combined test as $p = \frac{1}{B} \sum_{b=1}^B \mathbf{1}\{T_\Psi^{(b)} \geq T_\Psi^{(0)}\}$.

For more details about these benchmark randomization tests, we refer to Pigoli et al. (2014) and Cabassi et al. (2017).

9.2 Benchmark test Jentsch and Pauly (2015)

As a benchmark procedure for the manifold Wilcoxon signed-rank test for equality of the spectral matrices of two multivariate stationary time series, we consider the permutation test of Jentsch and Pauly (2015) based on the integrated L_2 -distance between kernel smoothed periodogram matrices.

Let $\vec{X}_{11}, \dots, \vec{X}_{1n}$ and $\vec{X}_{21}, \dots, \vec{X}_{2n}$ be (strictly) stationary multivariate time series, and denote by $I_1(\omega_\ell)$ and $I_2(\omega_\ell)$ respectively the raw periodogram matrices of the two time series at the Fourier frequencies $\omega_\ell = 2\pi\ell/n \in (-\pi, \pi]$. The permutation test (with B permutations) for equality of the two spectral matrices proceeds as follows:

1. Compute the test statistic,

$$T_n^{(0)} = nh^{1/2} \int_{-\pi}^{\pi} \sum_{i=1}^2 \left\| \frac{1}{n} \sum_{\ell=-\lfloor (n-1)/2 \rfloor}^{\lfloor n/2 \rfloor} K_h(\omega - \omega_\ell) (I_i(\omega_\ell) - \tilde{I}(\omega_\ell)) \right\|^2 d\omega$$

where K_h is a kernel function with bandwidth h and $\tilde{I}(\omega_\ell) = \frac{1}{2}(I_1(\omega_\ell) + I_2(\omega_\ell))$ is the pooled periodogram matrix.

2. For each $b = 1, \dots, B$, let $\pi_\ell = (\pi_\ell(1), \pi_\ell(2))'$ be a random permutation of the data labels $\{1, 2\}$ for each $\omega_\ell \in (-\pi, \pi]$ and compute the test statistic of the permuted sample,

$$T_n^{(b)} = nh^{1/2} \int_{-\pi}^{\pi} \sum_{i=1}^2 \left\| \frac{1}{n} \sum_{\ell=-\lfloor (n-1)/2 \rfloor}^{\lfloor n/2 \rfloor} K_h(\omega - \omega_\ell) (I_{\pi_\ell(i)}(\omega_\ell) - \tilde{I}(\omega_\ell)) \right\|^2 d\omega$$

3. Compute the p-value of the test as $p = \frac{1}{B} \sum_{b=1}^B \mathbf{1}\{T_n^{(b)} \geq T_n^{(0)}\}$.

For more details about this randomization test, we refer to Jentsch and Pauly (2015).