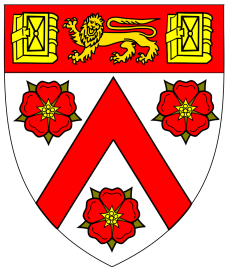
Topological and geometric inference of data



Adam P. Goucher

Supervisors: Prof. J. A. D. Aston
Prof. J. A. Rasmussen

Department of Pure Mathematics and Mathematical Statistics
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Trinity College August 2020

Declaration

This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation is original and is not substantially the same, in whole or in part, as any work previously or concurrently submitted for consideration for any other degree or diploma at the University of Cambridge or any other university or similar institution.

Adam P. Goucher
August 2020

Abstract

The overarching problem under consideration is to determine the structure of the subspace on which a distribution is supported, given only a finite noisy sample thereof. The special case in which the subspace is an embedded manifold is given particular attention owing to its conceptual elegance, and asymptotic bounds are obtained on the admissible level of noise such that the manifold can be recovered up to homotopy equivalence.

Attention is turned on how to accomplish this in practice. Following ideas from topological data analysis, simplicial complexes are used as discrete analogues of spaces suitable for computation. By utilising the prior assumption that the data lie on a manifold, topologically inspired techniques are proposed for refining the simplicial complex to better approximate this manifold. This is applied to the problem of nonlinear dimensionality reduction and found to improve accuracy of reconstructing several synthetic and real-world datasets.

The second chapter focuses on extending this work to the case where the ambient space is non-Euclidean. The interfaces between topological data analysis, functional data analysis, and shape analysis are thoroughly explored. Lipschitz bounds are proved which relate several metrics on the space of positive semidefinite matrices; they are then interpreted in the context of topological data analysis. This is applied to diffusion tensor imaging and phonology.

The final chapter explores the case where the points are non-uniformly distributed over the embedded subspace. In particular, a method is proposed to overcome the shortcomings of witness complex construction when there are large deviations in the density. The theory of multidimensional persistence is leveraged to provide a succinct setting in which the structure of the data can be interpreted as a generalised stratified space.

'Let no-one unversed in geometry enter here'

Plato's Academy

Table of contents

Chapter 1

Introduction

The statistical field of *manifold learning* is concerned with the problem where we have some dataset X lying in a high-dimensional ambient space A, typically $\mathbb{R}^D$. The data are assumed to be independent, identically distributed random variables of the form $X_i := Y_i + \varepsilon_i$, where Y_i is sampled from an unknown lower-dimensional smoothly embedded manifold $M \subsetneq A$ and ε_i is a small error term. The problem is to determine M.

The method of *principal component analysis*, first described in [11], works very effectively when M is an affine subspace of A. The more difficult and interesting case is when M is nonlinearly embedded in the ambient space. In [12], Genovese, Perone-Pacifico, Verdinelli and Wasserman consider this problem where ε_i is Gaussian, with the objective being to recover M as closely as possible by Hausdorff distance.

Nonlinear dimensionality reduction considers a similar problem of trying to find a map $\theta : X \to \mathbb{R}^d$ with $d < D$, such that for $x, y \in X$, the distance between x and y is small if and only if the distance between $\theta(x)$ and $\theta(y)$ is small. Popular techniques include *Isomap* [14], *t-SNE*, and, more recently, *UMAP*.

If we knew the manifold M, then we could construct θ by composing the 'orthogonal projection' $\pi : X \to M$ (mapping each point to the closest point on the manifold) with an embedding $\phi : M \to \mathbb{R}^d$. If M is a smooth n-manifold, then such an embedding is guaranteed to exist if $d \geq 2n$ by the Whitney embedding theorem (and this is tight, as shown by taking $M = \mathbb{RP}^{2^k}$) [19]. In the other direction, an n-manifold can only be embedded into $\mathbb{R}^d$ provided $d \geq n$ (and this is also tight, as shown by taking $M = \mathbb{R}^n$). This poses an obstacle for the pursuit of nonlinear dimensionality reduction: estimating the intrinsic dimensionality n of the manifold

M is not sufficient to determine the optimal d for which we can hope to map X into $\mathbb{R}^d$; instead, it is dependent on the topology of M.

Consequently, we argue that determining M and its topology should be seen as a fundamental precursor to performing dimensionality reduction, as it is necessary even to find the optimal embedding dimension. Finding the topology of a space given a finite sample is one of the questions of the relatively new field of *topological data analysis*.

1.1 Overview

We begin with a literature review of topological data analysis and nonlinear dimensionality reduction. This is followed by three substantive chapters (Chapters 3, 4, and 5), the main results of which are summarised here.

1.1.1 Overview of Chapter 3

Chapter 3 introduces the problem of detecting the embedded submanifold $M \subseteq A$, concentrating on the case where $A = \mathbb{R}^D$ is Euclidean and each sampled point is observed together with spherical Gaussian noise of variance σ^2. Given a large finite sample X of points obtained in this manner, we propose to eliminate the effect of the noise by discarding 'isolated points'; that is to say, we choose a density estimator f and threshold δ, and restrict to the points $\{x \in X : f(x) \geq \delta\}$. We establish bounds for the maximum value of σ^2 in terms of the dimension and normal injectivity radius of M such that some thickening of the above set admits a deformation retract onto M. In particular, we obtain tight asymptotic bounds for the case where M is the unit sphere:

Theorem. *Let $X := Y + \varepsilon$ be the sum of a random variable Y uniformly distributed on the surface of a sphere $S^n \subseteq \mathbb{R}^D$ and a Gaussian random variable ε with zero mean and covariance matrix $\sigma^2 I$ (where I is the $D \times D$ identity matrix).*

Then there exists a superlevel set of the probability density function of X admitting a deformation retraction onto the support of Y if and only if the following inequality holds:

$$e^{-\tau}\, {}_0F_1\left(; \frac{1+n}{2}; \tau^2\right) > 1$$

where $\tau = \frac{1}{2\sigma^2}$ and ${}_0F_1$ is the 'confluent hypergeometric limit function'.

We also determine an asymptotic version (as $n \to \infty$) of the above inequality, dispensing with the use of the non-elementary function:

Theorem. *Let n be the dimension, and consider the critical value of σ such that the centre and boundary of the sphere are equiprobable. Then the following asymptotic holds for $n \to \infty$:*

$$\frac{1}{2\sigma^2} = \frac{n-1}{4} z_0 + \frac{\log(1+z_0^2)(1+\sqrt{1+z_0^2})}{8z_0 - 4(1+\sqrt{1+z_0^2})} + o(1)$$

where z_0 is the positive real root of the equation:

$$\sqrt{1+z^2} + \log(2) - 1 - \log(1+\sqrt{1+z^2}) = \frac{1}{2} z$$

This result is then used to establish a more general result on arbitrary smoothly embedded manifolds, and to provide conditions for which we can recover the manifold from a finite noisy sample. In particular, the following is proved:

Theorem. *Suppose we have a compact smoothly embedded n-manifold $M \subset \mathbb{R}^D$ which has a normal injectivity radius of h. Further, let P be a finite subset of the r_2-thickening of M. Then, provided that we have $r_2 < \frac{1}{2}(2-\sqrt{2})h$, and the r_1-thickening of P covers M (where $r_1 = \sqrt{2}(h-r_2) - h$), then the R-thickening of P and the associated Čech complex are both homotopy equivalent to M where $R = \frac{1}{\sqrt{2}}(h - r_2)$.*

To utilise this theorem, we find conditions under which P can be obtained by sampling enough random points from X and restricting to a superlevel set of f:

Theorem. *Let $X := Y + \varepsilon$ be the sum of a random variable Y supported on an n-manifold $M \subseteq \mathbb{R}^D$ and a Gaussian random variable ε with zero mean and covariance matrix $\sigma^2 I$ (where I is the $D \times D$ identity matrix).*

Then $f(x) \geq \beta f(y)$ for all $x \in M$ and $y \notin M_{r_2}$ provided the following inequality holds:

$$\frac{h^2}{2\sigma^2} \geq \frac{n-1}{4} z_0 + \frac{\frac{1}{4}\log(1+z_0^2) + \log(K\beta) - \log(\alpha)}{\left(1+\sqrt{1+z_0^2}\right)^{-1} z_0 - \left(1 - \frac{r_2^2}{2h^2}\right)} + o(1)$$

where $K, z_0 > 0$ are universal constants, M_{r_2} is the r_2-thickening of M, h is the normal injectivity radius of M, and α is the (appropriately normalised) minimum density of the random variable Y on its support M.

The remainder of the chapter examines the practicalities of performing manifold learning in this way, especially for large datasets where a Čech complex is infeasible to construct and an approximation such as a witness complex is required. Topologically-inspired approaches for removing noise from the witness complex are proposed; they are shown both to improve the efficacy of manifold learning and to remove the majority of short-lived 'noise' bars from the persistent homology barcode. These ideas are reëxamined in Chapter 5.

1.1.2 Overview of Chapter 4

Chapter 4 explores the situation where the ambient space A is non-Euclidean; we demonstrate how to generalise the ideas from the previous chapter. Particular attention is given to spaces of covariance operators, where new Lipschitz constants are proved relating various important metrics on these spaces. In particular, we prove the following:

Theorem. *Let H_1, H_2 be $d \times d$ covariance matrices. Then we have:*

$$d_S(H_1, H_2) \leq d_H(H_1, H_2) \leq \sqrt{2} d_S(H_1, H_2)$$

where d_S denotes Procrustes distance and d_H denotes square-root distance.

These have immediate corollaries in terms of the persistent homology of X with respect to these metrics: changing from one metric to the other can only perturb the coordinates of the points in the persistence diagram by a (multiplicative) factor of at most $\sqrt{2}$. As an application, these ideas are applied to the analysis of diffusion tensor imaging and (by extending the result from finite-dimensional matrices to infinite-dimensional trace-class positive-semidefinite operators) to speech datasets.

We also define a continuum of metrics on the space of positive-definite matrices, which interpolate between the log-Euclidean metric and the affine-invariant Riemannian metric:

$$d_p(A, B) := \left\| \frac{1}{p} \log(A^{-p/2} B^p A^{-p/2}) \right\|_2$$

The following is proved:

Theorem. *The log-p distances interpolate between the log-Euclidean distance and the affine-invariant Riemannian distance, in the sense that:*

- $d_0(A, B) = d_L(A, B)$;
- $d_1(A, B) = d_R(A, B)$.

Moreover, when A and B commute, all of the log-p distances coincide.

We proceed to show that $d_p(A, B)$ is a monotone-increasing function of p, and therefore that the affine-invariant Riemannian distance is never smaller than the log-Euclidean distance. Moreover, provided the condition numbers of the two matrices are bounded, a Lipschitz constant in the other direction is established.

1.1.3 Overview of Chapter 5

Chapter 5 addresses the issue of the free parameter δ. Instead of choosing it arbitrarily, we propose taking a filtration with respect to this density parameter. This fits naturally in the framework of multidimensional persistence, and so we generalise our ideas to this setting. Experiments are performed firstly on a synthetic dataset, where we demonstrate the efficacy of the approach even in the presence of heavy-tailed Cauchy noise; we then turn our attention to larger real-world datasets.

The chapter substantially deals with practical considerations involving building witness complexes when the ambient density is highly variable. A 'conformal' variant of the sequential maxmin procedure for choosing landmark points is proposed. This is formulated as taking a 'sequential maxmin sequence' (or SMMS) with respect to a semi-metric (metric without the triangle inequality) $d^\star$, and proved to asymptotically fill the space with the same density as the original distribution:

Theorem. *Let M be a compact n-manifold endowed with a Riemannian metric g which induces the path metric $d : M \times M \to \mathbb{R}_{\geq 0}$. Let $f : M \to \mathbb{R}$ be a density function with uniformly continuous logarithm, and furthermore suppose f is normalised such that:*

$$\int_M f(x)\, d\mu_g = 1$$

where μ_g is the standard Riemannian measure. This allows us to define a probability measure on the Borel subsets of M:

$$\lambda(A) := \int_A f(x)\, d\mu_g$$

Setting $\ell(x) := f(x)^{-1/n}$, we define the semimetric as before:

$$d^\star(x,y) := \frac{d(x,y)}{\sqrt{\ell(x)\ell(y)}}$$

Then there exist constants $0 < c, C < \infty$ *such that, for any Jordan measurable subset* $A \subseteq M$ *and any SMMS* $x_1, x_2, \ldots$ *chosen with respect to* $(X, d^\star)$*, there exists* $m_0 \in \mathbb{N}$ *such that for all* $m \geq m_0$ *we have:*

$$c\lambda(A) \leq \frac{1}{m}\sum_{i=1}^{m}[x_i \in A] \leq C\lambda(A)$$

where the Iverson bracket $[x_i \in A]$ *is defined to be* 1 *if* $x_i \in A$ *and* 0 *otherwise.*

The Jordan measurability criterion is shown to be largely necessary (in particular, weaker conditions such as 'Borel measurable' or even 'open' does not suffice).

Also, the constants c, C cannot be eliminated, even when the density function f is constant: a one-dimensional space is exhibited where C/c cannot be taken to be less than 2, and similarly a four-dimensional space is exhibited where the ratio cannot be taken to be less than 4. It remains open if there is a dimension-independent choice of C, c which work universally for all manifolds.

The chapter concludes by experimentation on various synthetic and real-world datasets.

Chapter 2

Background

2.1 Topological data analysis

2.1.1 Simplicial complexes and homology

Homology was introduced by Henri Poincaré in [5]; a more modern exposition is given in Chapter 2 of Hatcher's [6]. Most relevant to this thesis is *simplicial homology*, which associates a sequence of modules to an object known as a *simplicial complex*.

Definition 2.1.1. *A* simplicial complex *is a set $\mathcal{K} \subseteq \mathcal{P}(V)$ of* simplices *identified with non-empty subsets of some finite* vertex set *V, satisfying the following properties:*

- *If $\emptyset \neq \tau \subseteq \sigma \in \mathcal{K}$, then $\tau \in \mathcal{K}$.*
- *The singleton set $\{v\}$ is an element of $\mathcal{K}$ for all $v \in V$.*

Beyond this combinatorial structure, a simplicial complex $\mathcal{K}$ gives rise to a (compact, Hausdorff, and metrisable) topological space $|\mathcal{K}|$.

Definition 2.1.2. *The* polyhedron *of a simplicial complex $\mathcal{K}$ is the subspace $|\mathcal{K}| \subseteq [0,1]^V$ containing all points $x \in [0,1]^V$ satisfying both of the following conditions:*

- *$\sum_{v \in V} x_v = 1$;*
- *The support $\{v \in V : x_v > 0\}$ is an element of the set $\mathcal{K}$ of simplices.*

$|\mathcal{K}|$ inherits the subspace topology from the (compact, Hausdorff, and metrisable) space $[0,1]^V$.

Whenever a statement is made about the topology of one or more simplicial complexes, such as describing a simplicial complex as 'connected' or asserting that two simplicial complexes are 'homeomorphic', this is to be regarded as a statement about the topology of the corresponding polyhedra.

Definition 2.1.3. *Fix a commutative ring R, assumed to be $\mathbb{Z}$ unless otherwise specified, and let $\mathcal{K}$ be a simplicial complex. For a nonnegative integer n, an* n-chain *is a formal R-linear combination of n-simplices in $\mathcal{K}$. The R-module of all n-chains is denoted C_n.*

As V is a finite set, we can identify it with an initial segment of the positive integers, endowed with the usual total order. For each n, the *boundary map* $\partial_n : C_n \to C_{n-1}$ is defined as the linear map which acts on the basis elements as such:

$$\{v_0, v_1, \ldots, v_n\} \mapsto \sum_{i=0}^{n} (-1)^i \left(\{v_0, v_1, \ldots, v_n\} \setminus \{v_i\}\right)$$

where $v_0 < v_1 < \cdots < v_n$ by assumption.

Lemma 2.1.4. *The boundary maps satisfy $\partial_n \circ \partial_{n+1} = 0$ for all $n \geq 1$.*

Proof. It suffices to show that every simplex in C_{n+1} is in the kernel of $\partial_n \circ \partial_{n+1}$; the result would follow by linearity. Composing the definitions of the boundary maps, we get:

$$\{v_0, v_1, \ldots, v_{n+1}\} \mapsto \sum_{i=0}^{n+1} (-1)^i \sum_{j=0}^{n+1} [i \neq j](-1)^{j-[j>i]} \left(\{v_0, v_1, \ldots, v_n\} \setminus \{v_i, v_j\}\right)$$

where $[\phi]$ is the Iverson bracket (defined to be 1 if ϕ is true, and 0 if ϕ is false). Rearranging, we get:

$$\sum_{i \neq j} (-1)^{i+j-[j>i]} \left(\{v_0, v_1, \ldots, v_n\} \setminus \{v_i, v_j\}\right)$$

The terms with $i < j$ cancel out exactly with the corresponding terms with $j < i$, so the entire sum evaluates to zero. □

Definition 2.1.5. *The* space of n-cycles *is defined to be the kernel of ∂_n, and the* space of n-boundaries *is defined to be the image of ∂_{n+1}. It follows from the previous assertion that the latter is contained within the former, enabling one to define the* nth simplicial homology group *to be the quotient R-module.*

Homology is a topological invariant, in that any two homeomorphic simplicial complexes have an identical (up to isomorphism) sequence of homology groups. More generally, it is a homotopy invariant: any two homotopy-equivalent simplicial complexes have isomorphic homology. This is proved by showing that simplicial homology is isomorphic to another definition of homology, *singular homology*, for which homotopy invariance follows immediately from the definition.

2.1.2 Simplicial complexes associated with point cloud data

A point cloud dataset is a discrete metric space X, so it is not entirely obvious how to impose a meaningfully interesting topology on such a dataset. In [10], Vietoris introduced the *Vietoris-Rips complex*: a simplicial complex $\mathcal{R}_r$ which contains a vertex for each point $x \in X$ and a simplex for each finite subset $\sigma \subseteq X$ with diameter at most r. Unlike the original point cloud, which just has a zeroth Betti number β_0 equal to the number of connected components and no nontrivial homology beyond that, the Vietoris-Rips complex can have an arbitrary finitely-supported sequence of finite Betti numbers as demonstrated in [68].

The following table details salient properties of the different types of simplicial complexes. The column labelled 'arbitrary metric?' specifies whether the complex can be constructed from just knowing the distances between points and making no assumptions on the metric space (such as it being $\mathbb{R}^D$, or a normed vector space, or a Riemannian manifold). The remaining two columns pertain specifically to the Euclidean case when $X \subseteq \mathbb{R}^D$: 'k-simplices' is an upper bound on the number of k-dimensional simplices in the complex, and 'faithful?' specifies whether the complex is necessarily homotopy-equivalent to some subset of the ambient space.

Complex	k-simplices	arbitrary metric?	faithful?
Vietoris-Rips	$\binom{n}{k+1}$	Yes	No
Čech	$\binom{n}{k+1}$	No	Yes
Alpha	$O(n^{\lfloor D/2 \rfloor})$	No	Yes
Witness, W_∞	N	Yes	Yes
Witness, W_r	$O(N^{\max(1, \frac{k+1}{r+1})})$	Yes	Asymptotically

Vietoris-Rips complexes

An equivalent formulation of the Vietoris-Rips complex is as the clique complex of its 1-skeleton. In particular, we can construct a graph G_r by connecting points

$x, y \in X$ with an edge whenever $d(x, y) \leq r$. This has the effect that the connected components in G_r correspond to clusters in X.

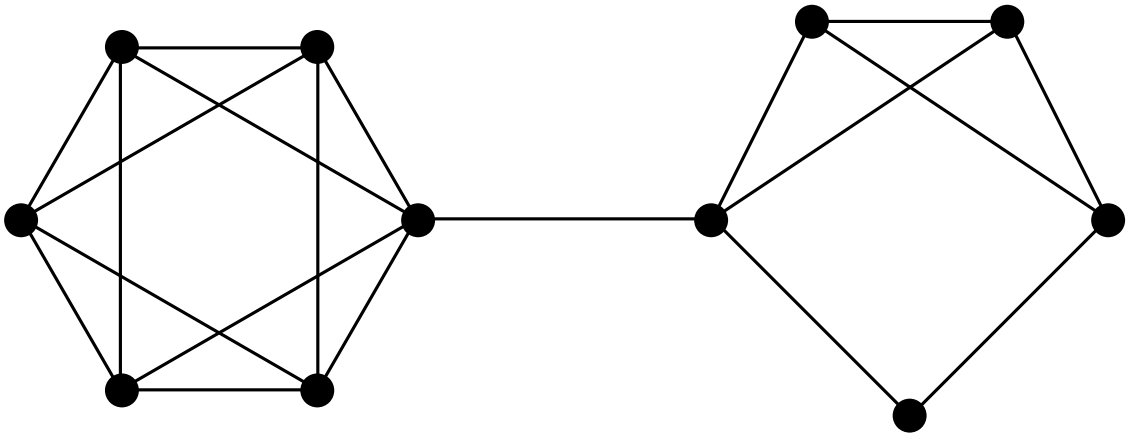


Fig. 2.1 A graph obtained by connecting every pair of points in X within a particular threshold distance r.

Unfortunately, this only captures information about the connected components (or equivalently H_0). To ascertain the higher homology groups, we need to somehow upgrade G_r into a simplicial complex. The most straightforward way is to form the *clique complex* of the graph:

Definition 2.1.6. *Let G be a graph with vertex-set V and edge-set E. The* clique complex *C is the abstract simplicial complex with vertex-set V, where a simplex $\sigma \subseteq V$ is included in C if and only if the subgraph of G induced by σ is a complete graph.*

That is to say, we begin with our graph G, and attach an n-simplex $[v_0, v_1, \ldots, v_n]$ whenever every pair $(v_i, v_j) \in E$. Note that, in particular, the vertices and edges of C are precisely those of G, so we can recover G by taking the 1-skeleton of C. We can equivalently define the clique complex as the unique largest simplicial complex with 1-skeleton G.

In the case where the graph is obtained by connecting pairs of vertices within a threshold distance of r, as in Figure 2.2, the clique complex is called a *Vietoris-Rips complex*. This contains a simplex $\sigma \subseteq V$ if and only if the diameter $d(\sigma) \leq r$.

If we look closely at the hexagonal arrangement of vertices in Figure 2.2, there is an edge for every non-antipodal pair of vertices. The clique complex corresponding to this subgraph is an octahedron, which is homeomorphic to the sphere S^2. Note that this has a spuriously non-trivial second homology – given that our set X was

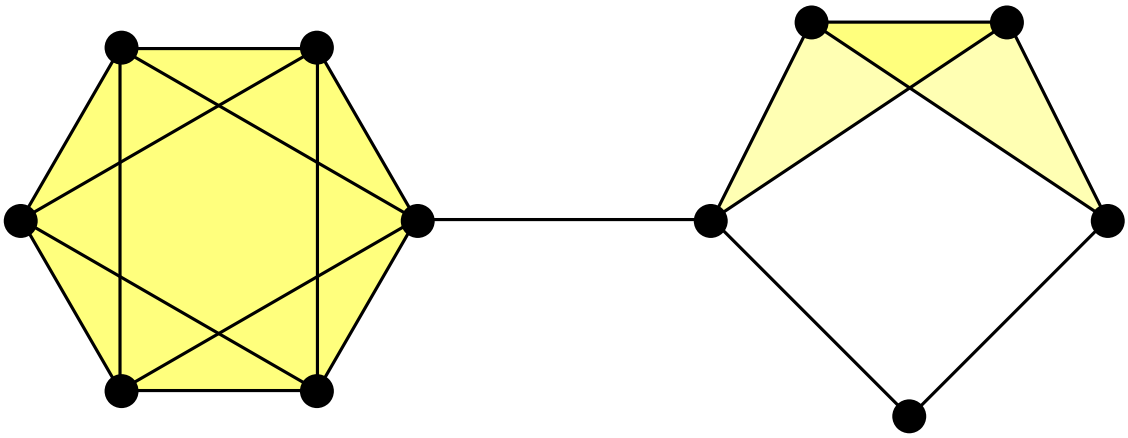


Fig. 2.2 The clique complex of the graph in Figure 2.1

sampled from some distribution on the plane, we would want the second homology to reflect that.

The phenomenon generalises: if we take the vertices of a regular $2n$-gon of radius 1, then for all $2\cos(\frac{\pi}{2n}) \leq r < 2$, the Vietoris-Rips complex $\mathcal{R}_r$ is homeomorphic to S^{n-1} and therefore has non-vanishing $(n-1)$th homology.

Another unfortunate property of Vietoris-Rips complexes is that the number of simplices can grow very large. In particular, when r exceeds the diameter of X, the number of simplices grows to $2^{|X|} - 1$. Unless X is very small, computing the homology of the Vietoris-Rips filtration is infeasible.

Čech complexes

One problem with Vietoris-Rips complexes is that they do not accurately convey the topology of the underlying space: we remarked that a regular $2n$-gon in the plane has a Vietoris-Rips complex with non-trivial $(n-1)$th homology, whereas no subset of the plane has any homology groups beyond H_1. We shall instead consider a sublevel filtration, which we know has meaningful topology, and associate a homotopy-equivalent filtration of simplicial complexes.

These spaces are unions of closed balls, rather than simplicial complexes. However, it is possible to associate a natural simplicial complex to each of these spaces, which we will see is homotopy-equivalent to the original space:

Definition 2.1.7. *The* Čech complex $\mathcal{C}_r(X)$ *contains a simplex* $\sigma \subseteq X$ *whenever the closed* $\frac{r}{2}$*-balls centred on the vertices mutually overlap, i.e.* $\bigcap_{x\in\sigma} B(x, \frac{r}{2}) \neq \emptyset$.

Note that the 1-skeleton (underlying graph) of $\mathcal{C}_r$ agrees with that of the Vietoris-Rips complex $\mathcal{R}_r$. However, the Čech complex has only a subset of the simplices of $\mathcal{R}_r$. Compare Figures 2.2 and 2.3, and note in particular that the octahedron is replaced with an annulus.

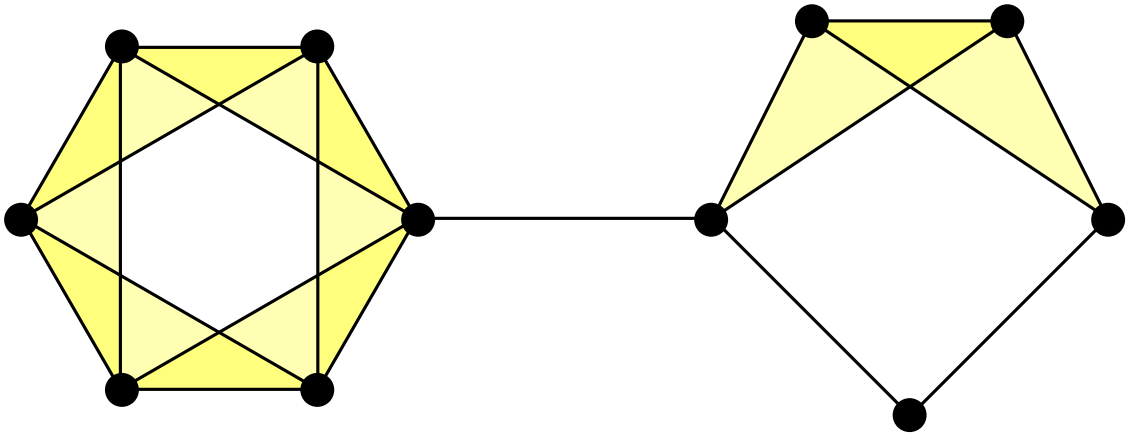


Fig. 2.3 The Čech complex $\mathcal{C}_r$ for the same value of r as the Vietoris-Rips complex in Figure 2.2

To show that $X_{r/2}$ is homotopy-equivalent to $\mathcal{C}_r(X)$ relies on a result called the *nerve theorem*.

Definition 2.1.8. *Let K be a paracompact space. A* good cover *of K is a collection of (wlog non-empty) open sets $\{U_\alpha : \alpha \in I\}$ such that:*

- $\bigcup_{\alpha \in I} U_\alpha = K$;
- $\bigcap_{\alpha \in J} U_\alpha$ *is either empty or contractible for any finite $J \subseteq I$.*

The nerve *of a good cover is the abstract simplicial complex with vertex-set I and a simplex for each finite set $J \subseteq I$ such that $\bigcap_{\alpha \in J} U_\alpha$ is non-empty (ergo contractible).*

Subject to these conditions, the *nerve theorem* proved by Borsuk in [34] states that the nerve of a good cover of K is homotopy-equivalent to K. We are usually interested in the case where K is a metric space and the open sets are balls in this metric:

Corollary 2.1.9. *Let $X \subset V$ be a finite subset of a finite-dimensional normed space V. Then the $\frac{r}{2}$-thickening $X_{r/2}$ is homotopy-equivalent to the Čech complex $\mathcal{C}_r(X)$.*

Proof. Note that the claim involves closed balls, whereas the statement of the nerve theorem involves a covering by open sets. To account for this, it suffices to choose r' slightly larger than r such that the associated nerves are equal. This can be accomplished as long as r' is less than $\min(\Delta \cap (r, \infty))$ where Δ is the (finite) set of diameters of minimal bounding spheres of finite subsets of X.

Now everything becomes straightforward: balls in V are necessarily convex (therefore contractible), and intersections of convex sets are also convex. So taking K to be $X_{r/2}$ and our good cover to be open balls of radius $r'/2$ centred on the vertices of X, the result follows. □

The case where $X_{r/2}$ covers all of M and deformation retracts onto M is of interest, because the resulting Čech complex is therefore homotopy equivalent to M. That is to say, it is possible to recover M (up to homotopy equivalence, at least) from a sufficiently well-chosen finite set X sampled from the ambient space in which M resides. This dissertation is largely an exploration of this idea, examining under what conditions and how we can guarantee the existence of a suitable set X.

Whilst the Čech complex addresses the topological problems with the Vietoris-Rips complex, it is still large and unwieldy. Indeed, we see in the next section that $\mathcal{R}_r \subseteq \mathcal{C}_{r\sqrt{2}}$, so the Čech complex has the same asymptotic complexity as the Vietoris-Rips complex. As such, both Čech and Vietoris-Rips complexes are interesting only for their theoretical value; practical applications favour the smaller *alpha complex* and *witness complex*.

Alpha complexes

Earlier, we mentioned that all $2^{|X|} - 1$ simplices will appear in the Čech filtration. When $|X|$ is even moderately large, the complex is intractably large to store or manipulate. *Alpha complexes* are an alternative which are homotopic to Čech complexes whilst being much smaller.

For each ball $B_{r/2}(v)$ in the $r/2$-thickening of X, consider its intersection $B_{r/2}(v) \cap \Omega_v$ with the closed Voronoi cell Ω_v (set of points at least as close to v as to any other point in X). The nerve of this covering was introduced by Edelsbrunner and Mücke in [33].

Definition 2.1.10. *The* alpha complex $\mathcal{A}_\alpha$ *of a set X contains a simplex with vertices $\sigma \subseteq X$ whenever there is a ball B_r with radius $r \leq \alpha/2$ such that σ lies on the boundary of B and the interior of B contains no points in X.*

Note that this condition is equivalent to there being at least one point (the centre of the ball B_r) which lies in the intersection of the sets $B_{\alpha/2}(v) \cap \Omega_v$, so the alpha complex is exactly the nerve of the 'Voronoi covering' described above.

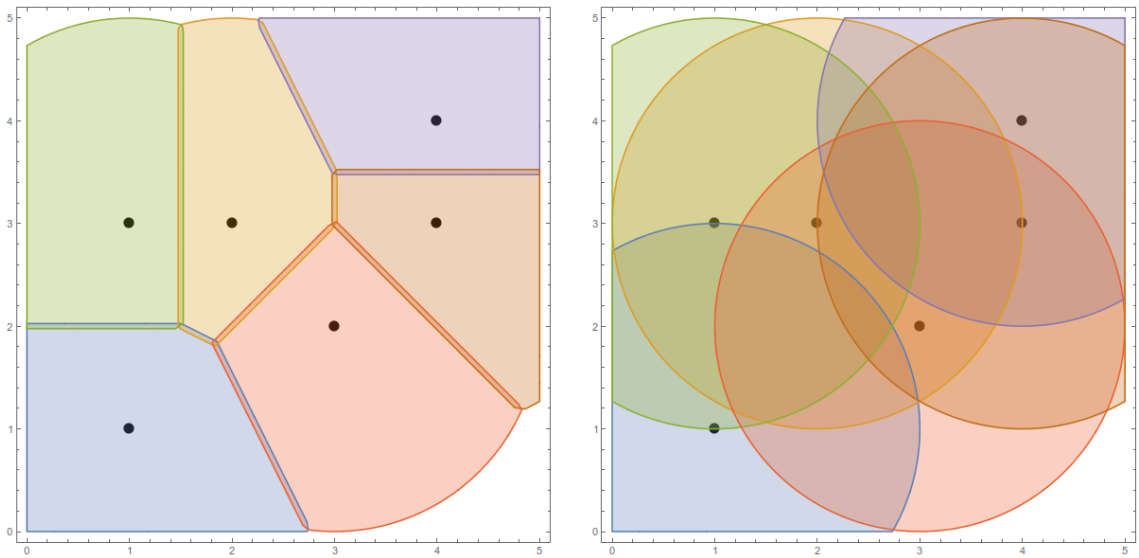


Fig. 2.4 The alpha complex (left) and Čech complex (right) of the same finite set X with the same radius parameter. The apparent overlaps between the Voronoi regions are for visualisation purposes only; in reality, the regions overlap only at their (measure-zero!) boundary.

Unlike in the Čech complex, where every non-empty finite subset $\sigma \subseteq X$ determines a simplex in $\mathcal{C}_r$ for sufficiently large r, the sets of points which can form simplices in the alpha complex are very restricted. For example, when X is a set of points in general position in $\mathbb{R}^D$, the dimensions of the simplices are no greater than D. It immediately follows that the number of simplices in the *Delaunay triangulation* (the limiting alpha complex when $\alpha \to \infty$, which is geometrically the dual of the Voronoi partition) has a polynomial upper bound $\binom{|X|}{D+1} + \binom{|X|}{D} + \cdots + \binom{|X|}{1}$.

It is possible to improve upon this bound. To do so, we 'lift' our set $X \subseteq \mathbb{R}^D$ to $\mathbb{R}^{D+1}$ by prepending an extra coordinate given by the sum of squares of the existing coordinates:

$$(x_1, x_2, \ldots, x_D) \mapsto (x_1^2 + x_2^2 + \cdots + x_D^2, x_1, \ldots, x_D)$$

Then the Delaunay triangulation is obtained by projecting the lower convex hull of these points back into $\mathbb{R}^D$ by deleting the additional coordinate x_0. Consequently, the

number of k-simplices in the Delaunay triangulation is bounded above by the number of k-simplices in a convex polytope in $\mathbb{R}^{D+1}$. This was established by McMullen in [38], the celebrated *Upper Bound Theorem*.

Theorem 2.1.11 (Upper Bound Theorem)**.** *A convex polytope with n vertices in $\mathbb{R}^{D+1}$ has no more k-dimensional faces than does the* cyclic polytope $\Delta(D+1,n)$*, namely:*

$$f_k(\Delta(D+1,n)) = \sum_{r=0}^{D+1} \binom{r}{D-k} \binom{n-1-\max(r,D+1-r)}{\min(r,D+1-r)}$$

When $k < \lceil \frac{D}{2} \rceil$, every set of $k+1$ vertices induces a k-simplex; hence, $f_k(\Delta(D+1,n)) = \binom{n}{k+1}$. When k is larger than this, we can derive an asymptotic estimate by fixing D, allowing n to grow large, and discarding any non-dominant terms in the summation:

$$f_k(\Delta(D+1,n)) = \frac{a_{k,D}+o(1)}{\lceil \frac{D}{2} \rceil !} n^{\lceil \frac{D}{2} \rceil}$$

where $a_{k,D}$ is a constant given by:

$$a_{k,D} := \begin{cases} \binom{\frac{D+1}{2}}{D-k} & \text{if } D \text{ odd;} \\ \binom{\frac{D}{2}}{D-k} + \binom{\frac{D}{2}+1}{D-k} & \text{otherwise} \end{cases}$$

In particular, the total number of simplices is asymptotically:

$$\sum_k f_k(\Delta(D+1,n)) = (2+(-1)^D + o(1)) \frac{(2n)^{\lceil \frac{D}{2} \rceil}}{\lceil \frac{D}{2} \rceil !}$$

There are two immediate drawbacks to the alpha complex. Firstly, the number of simplices in the Delaunay complex grows exponentially with the dimension. Secondly, the existence and uniqueness of Delaunay triangulations depends on the ambient space; whilst they are guaranteed for all finite sets of points in general position in Euclidean space, $A = \mathbb{R}^D$, there are non-Euclidean Riemannian 3-manifolds in which this is not the case [39].

2.1.3 Witness complexes

Witness complexes were introduced by Carlsson and de Silva in [31]. Instead of taking the vertex-set to be the entirety of X, we carefully choose a much smaller

subset $L \subseteq X$, termed *landmark points*, and use the remaining points as 'witnesses' to determine which subsets of the landmark points induce simplices.

Definition 2.1.12. *A k-simplex with vertices $\sigma \subseteq L$ is said to be* (weakly) α-witnessed *if there exists some point $w \in X$ (the* witness*) and radius $r \leq \alpha$ such that $B_r(w) \cap L = \sigma$. We say that the simplex is* strongly α-witnessed *if σ lies exactly on the boundary of $B_r(w)$.*

A simplex $\sigma \subseteq L$ is included in the witness complex $\mathcal{W}_\alpha^\infty$ *if and only if every non-empty face $\tau \subseteq \sigma$ is weakly α-witnessed.*

Moreover, the following theorem gives a relation between weakly and strongly witnessed points whenever $X = \mathbb{R}^d$ is the whole of Euclidean space:

Theorem 2.1.13. *A simplex σ is strongly α-witnessed by some point in $\mathbb{R}^d$ if and only if every non-empty face $\tau \subseteq \sigma$ is weakly α-witnessed by some point in $\mathbb{R}^d$.*

Consequently, when $X = \mathbb{R}^d$, the witness complex $\mathcal{W}_\alpha^\infty$ is precisely the alpha complex $\mathcal{A}_\alpha$ of the landmark set L. Whereas computing the alpha complex relies on knowledge of the ambient space M, the witness complex can be computed from the matrix of distances between L and X (recall that $L \subseteq X \subseteq M$, where X is typically finite and M is typically infinite).

Note that for each k, a particular witness point $w \in X$ can witness at most one k-simplex. This gives an upper bound of $|X|$ simplices of each dimension k.

Choosing the landmark points

This leaves the question of how to choose an appropriate $L \subseteq X$. The two prevailing approaches are *random sampling* (where any of the $\binom{|X|}{|L|}$ subsets is chosen with equal probability) or *sequential maxmin* (where points are chosen greedily to be maximally distant from the closest existing landmark point).

When X is randomly sampled from a non-uniform distribution on its support, sequential maxmin has a shortcoming: the landmark points L are uniformly spaced in a way that we shall make precise in the final chapter, and therefore not representative of the distribution of X. We proceed to introduce a modification of sequential maxmin that overcomes this problem.

Completed witness complexes

If X is too small compared with L, there are often k-simplices in the alpha complex of L which fail to be included in the witness complex. This problem grows with the

dimension k: not only are k-simplices less likely to be witnessed, but also we require all of their proper faces to similarly be witnessed.

In [31], this is circumvented by taking the r-skeleton of the witness complex and 'completing' it by including a face of dimension $r+1$ or greater whenever all of its proper faces are present. The resulting construction is deemed $\mathcal{W}^r$. We shall now proceed to bound the number of k-simplices by $O(N^{\max(1,\frac{k+1}{r+1})})$.

Recall that the r-skeleton of $\mathcal{W}^r$ agrees with $\mathcal{W}$, so for each $k \leq r$, there are at most $|X|$ k-simplices. We can use this to bound the number of simplices in higher dimensions by appealing to the *Kruskal-Katona theorem* first proved independently in [66] and [67]. To formalise this, we need to introduce a few definitions:

Definition 2.1.14. *Let $X^{(k)}$ denote the collection of size-k subsets of the finite set $X := \{1, 2, \ldots, n\}$. The* colexicographical order *(or* colex*) on $X^{(k)}$ is obtained by assigning, to each $S \in X^k$, a score of:*

$$score(S) = \sum_{i \in S} \omega^i$$

where ω is an arbitrary integer larger than n. This induces a total order on $X^{(k)}$ where lower-scoring subsets appear before higher-scoring ones.

Definition 2.1.15. *Let $A \subseteq X^{(k)}$ be a collection of k-subsets of X. Define the* lower shadow *$\partial A \subseteq X^{(k-1)}$ to be the collection of $(k-1)$-subsets of X which can be obtained by deleting an element from a member of A.*

Theorem 2.1.16 (Kruskal-Katona)**.** *Suppose $|A| = |C|$ and C is an initial segment of colexicographical order. Then $|\partial A| \geq |\partial C|$.*

Proof. We induct on n, the size of our ground set, proving the result uniformly for all k. The idea is to massage A with a sequence of 'compressions', each of which cannot increase the size of its lower shadow, until we have either C (in which case we're done) or something so close to C that we can compare them directly.

Let $i \in X$ be fixed. We partition $X^{(k)}$ into two families, namely:

- $+_i^k = \{S \in X^{(k)} : i \in S\}$
- $-_i^k = \{S \in X^{(k)} : i \notin S\} = (X \setminus \{i\})^{(k)}$

As noted, $-_i^k$ is precisely $(X \setminus \{i\})^{(k)}$. Similarly, $+_i^k$ is isomorphic to $(X \setminus \{i\})^{(k-1)}$ as a lattice; we can remove i from every element of $+_i^k$ to obtain the equivalence.

Now, given an arbitrary $A \subseteq X^{(k)}$, we define its *i-compression* to be the collection $B \subseteq X^{(k)}$ satisfying:

- $|A \cap +_i^k| = |B \cap +_i^k|$ and the latter is an initial segment of colex restricted to $+_i^k$.
- $|A \cap -_i^k| = |B \cap -_i^k|$ and the latter is an initial segment of colex restricted to $-_i^k$.

Note that we have the following, where d_i denotes the isomorphism from $+_i^k$ to $(X \setminus \{i\})^{(k-1)}$ obtained by deleting i:

- $(\partial A) \cap +_i^{k-1} = (\partial(A \cap +_i^k)) \cap +_i^{k-1}$
- $(\partial A) \cap -_i^{k-1} = d_i(A \cap +_i^k) \cup \partial(A \cap -_i^k)$

By the inductive hypothesis, we have $|(\partial A) \cap +_i^{k-1}| \geq |(\partial B) \cap +_i^{k-1}|$ and $|\partial(A \cap -_i^k)| \geq |\partial(B \cap -_i^k)|$. Also, $d_i(A \cap +_i^k)$ and $d_i(B \cap +_i^k)$ are the same size. Finally, we appeal to the fact that initial segments of colex are nested (a smaller initial segment is a subset of a larger initial segment) to conclude that $|(\partial A) \cap -_i^{k-1}| \leq |(\partial B) \cap -_i^{k-1}|$.

Consequently, B has a smaller lower shadow than A. Also, B is 'closer to colex' in that the sum of the positions in colex of the elements of B is at least as small as that of A (with equality if and only if $A = B$). So we can repeatedly i-compress our set A for different $i \in X$ until it is i-compressed for every $i \in X$.

Hence, assume without loss of generality that A is i-compressed for every i.

If A is an initial segment of colex then we are done, so assume that there exist $x \in A$ and $y \in X^{(k)} \setminus A$ such that y precedes x. If either x or y share an element i, then A is not i-compressed and we obtain a contradiction. Similarly, if either $X \setminus x$ or $Y \setminus y$ share an element i, then again A is not i-compressed. Consequently, x and y are complements (their disjoint union is X).

We now show that such a pair x, y must be unique. If there is another $x' \in A$ colexicographically preceded by some $y' \in X^{(k)} \setminus A$, then we can assume without loss of generality that y' precedes y. Thence, we deduce that x, y' is a third such complementary pair, implying that $x = x'$ and $y = y'$ as desired.

Hence, $A = (X \setminus \{n\})^{(k)} \setminus \{\{1, \ldots, n-1\}\} \sqcup \{\{n\}\}$. It is apparent that its lower shadow is not smaller than that of $(X \setminus \{n\})^{(k)}$ (the corresponding initial segment of colex).

This completes the induction step. The base case of $n = 1$ is trivial (every subset of X is an initial segment of colexicographical order), so the result follows. □

The importance of Kruskal-Katona arises from the fact that in a simplicial complex, the collection of $(k-1)$-simplices contains the lower shadow of the collection of k-simplices. Consequently, we get a tight bound on the number of k-simplices as a function of the number of $(k-1)$-simplices; equality holds when the simplices form an initial segment of colexicographical order.

Let n_{k-1} be the number of $(k-1)$-simplices. By considering the general form of an initial segment of colexicographical order, we see that the number of k-simplices is therefore upper-bounded by:

$$n_k \leq \sum_{r \in R} \binom{r}{k+1}$$

where the set R gives the unique way to write n_{k-1} as a sum of binomial coefficients $\binom{r}{k}$:

$$n_{k-1} =: \sum_{r \in R} \binom{r}{k}$$

In particular, $n_k^k \leq n_{k-1}^{k+1}$. For an r-completed witness complex, we can show by induction that $n_k \leq N^{\frac{k+1}{r+1}}$, where N is the number of witness points (an upper bound for n_r).

2.1.4 Persistent homology

In 2002, Edelsbrunner, Letscher, and Zomorodian [25] introduced the notion of *persistence*: instead of computing the Betti numbers of a single Čech complex, $\mathcal{C}_r$, one can instead compute how many homology generators present in $\mathcal{C}_r$ 'persist' in some later Čech complex $\mathcal{C}_s$ (where $s > r$). More precisely, the authors defined the kth *persistent Betti number* $\beta_k^{r,s}$ to be the rank of the linear map on the kth homology induced by the inclusion map $\mathcal{C}_r \hookrightarrow \mathcal{C}_s$. (Technically, the authors used a different notation, where the second superscript was equal to $s-r$ instead of s, but we anachronistically adopt the more modern notation to ensure consistency.)

The first systematic treatment of *persistent homology*, one of the cornerstones of topological data analysis, began in 2004 with Zomorodian and Carlsson's seminal paper, [21].

As homology associates a sequence of groups to a topological space, persistent homology gives information about a *filtration* of nested spaces. The spaces are nested

in the sense that we have inclusion maps $\iota_{r,s} : K_r \to K_s$ for each pair of reals $r \leq s$, where they compose in the obvious way: if $r \leq s \leq t$, we have $\iota_{r,t} = \iota_{s,t} \circ \iota_{r,s}$.

One natural way in which a filtration can arise is as the sublevel sets of a function $f : A \to \mathbb{R}$ defined on the ambient space. In particular, let $K_r := \{x \in M : f(x) \leq r\}$. A familiar example is when $X \subseteq A$ is a finite set and $f : A \to \mathbb{R}$ is half of the distance to the closest point in X; the sublevel sets are precisely the thickenings $X_{r/2}$, which we know are homotopy-equivalent to the corresponding Čech complexes $\mathcal{C}_r$.

Definition 2.1.17. *Let $s \in \mathbb{R}$. If we can find $\epsilon > 0$ such that for all $r, t \in (s-\epsilon, s+\epsilon)$, the inclusion map $\iota_{r,t}$ induces an isomorphism between the homology of K_r and K_t, we say that s is a* homological regular value*.*

Otherwise, s is a homological singular value*.*

Less formally, the homological singular values can be thought of as times in the filtration where the homology abruptly changes. In [40], this importance of this notion is demonstrated by the *critical value lemma*.

Lemma 2.1.18. *Suppose the interval $[s, t]$ contains no homological singular values. Then the inclusion map $\iota_{s,t}$ induces an isomorphism of homology.*

Proof. Suppose otherwise. Then, we can perform repeated interval bisection to find a descending chain of nested closed intervals $[s_i, t_i]$ such that ι_{s_i,t_i} does not induce an isomorphism. By compactness, these intersect at a common point $p \in [s, t]$. For any open interval $(p - \epsilon, p + \epsilon)$, we can find some i such that the interval $[s_i, t_i]$ has length less than ϵ, and therefore lies entirely within $(p-\epsilon, p+\epsilon)$. However, ι_{s_i,t_i} does not induce an isomorphism, so p must be a homological singular value. □

If a function $f : A \to \mathbb{R}$ has only finitely many homological critical values, and the homology of each K_s is finite-dimensional, f is described as a *tame* function. In that case, the homology only changes finitely many times, so we can henceforth suppose that the spaces in our filtration are indexed by the naturals:

$$K_0 \subseteq K_1 \subseteq K_2 \subseteq K_3 \subseteq \cdots$$

Given a field $\mathbb{F}$, we can consider the chain complex of K_i. Following the notation in Weibel [1], we denote this by $C^i_\bullet$. The inclusion maps from each space to the next induce chain maps between the corresponding chain complexes:

$$C_\bullet^0 \xrightarrow{x_0} C_\bullet^1 \xrightarrow{x_1} C_\bullet^2 \xrightarrow{x_2} C_\bullet^3 \xrightarrow{x_3} \cdots$$

The *persistence complex* is the formal product $\mathbf{C}_\bullet = \prod_{i=0}^{\infty} C_\bullet^i$. We combine all of the x_i into a *shift map* $x : \mathbf{C}_\bullet \to \mathbf{C}_\bullet$.

By letting the monomial x^n denote the n-fold composition of x, the persistence complex is naturally a $\mathbb{F}[x]$-module. It is also a chain complex, since we can define a boundary map ∂ on $\mathbf{C}_\bullet$ which acts elementwise on each $C_\bullet^i$.

We can compute the kth homology $H_k(\mathbf{C})$ of the persistence complex, giving a $\mathbb{F}[x]$-module. With $\mathbb{F}$ being a field, it follows that the ring $R = \mathbb{F}[x]$ is a principal ideal domain. The Structure Theorem states that every finitely generated module M over a principal ideal domain R can be written as a direct sum:

$$M = \bigoplus_i R/(p_i)$$

where (p_i) is a primary ideal (possibly the zero ideal). These direct summands in the case of the polynomial ring come in one of two forms:

- Free elements $x^s\mathbb{F}[x]$, where a homology generator first appears in K_s and is present in all later complexes;
- Torsion elements $x^s(\mathbb{F}[x]/\langle x^{t-s}\rangle)$, where a homology generator first appears in K_s and subsequently first disappears in K_t.

We represent these elements as intervals $[s, \infty)$ or $[s, t)$, whose endpoints are the homological critical values at which the homology generator is created and subsequently destroyed. For each k, the collection of these intervals is a *barcode*.

The number of bars in the H_k barcode which contain the point s gives the Betti number β_k of the space K_s. It is convenient to generalise this to the *persistent Betti number*, β_s^t, which counts the number of bars which fully contain the interval $[s, t]$. We can think of this as the number of homology generators which exist at 'time' s (meaning present in the space K_s in the filtration) and continue to exist at time t. More formally, it is the dimension of the image of $H_k(K_s)$ under the homology map induced by $\iota_{s,t}$.

Longer bars correspond to more persistent topological features, which are more relevant to describing the global geometric structure of X. Shorter bars can be either due to sampling noise (if, say, X is a random sample), or indicative of small-scale

topological features. A large block of Emmental cheese would have no long bars, since the large-scale structure is that of a cube, but would possess many short H_2 bars corresponding to the spherical voids within the cheese.

2.1.5 Computing persistent homology

There are algorithms for computing the persistent homology over a field $\mathbb{F}$ of a filtration $\mathcal{K}$ of simplicial complexes. In particular, let a *chronological order* on the simplices be a total order $\preccurlyeq$ satisfying the following properties:

- If $\sigma \in \mathcal{K}_s$ and $\tau \notin \mathcal{K}_s$ for some s, then $\sigma \preccurlyeq \tau$;
- If $\sigma \subseteq \tau$, then $\sigma \preccurlyeq \tau$.

Then, beginning with the empty simplicial complex and adding the simplices in this order, the introduction of a k-simplex increases the Euler characteristic by $(-1)^k$. Since the Euler characteristic is also the alternating sum of Betti numbers, it follows that the introduction of this k-simplex either decrements β_{k-1} (if the boundary of the simplex was in a nontrivial homology class prior to the introduction of the simplex) or increments β_k (otherwise). That is to say, each simplex corresponds uniquely to an endpoint of a bar in the homology barcode.

Let D be the upper-triangular matrix specifying the boundary map ∂ in the basis given by the simplices in chronological order. By applying column operations to D (of the form 'add a constant multiple of one column to a later column'), we can *reduce* D to a matrix R with the property that the bottommost nonzero entries of the nonzero columns all occupy distinct rows. Each of these bottommost nonzero entries corresponds to a finite bar in the homology barcode: if it occupies row i and column j, then it is generated by the ith simplex and killed by the jth. By the previous argument, any unpaired simplices are generators of unbounded bars in the homology barcode.

Optimisations

Note that if we partition the simplices according to their dimension, D has the structure of a block matrix, and the reduction procedure can be applied to each of the nonzero blocks independently.

The application of the Structure Theorem was the only place where we needed $\mathbb{F}$ to be a field, rather than a ring. As mentioned earlier, we often choose $\mathbb{F}_2$ for various computational reasons:

- $\mathbb{F}_2$ is a field of characteristic 2, so the orientations of simplices can be safely ignored. More precisely, permuting the vertices of a simplex does not change the corresponding element of the chain complex.
- Gaussian elimination can be performed without involving division (since the unique invertible element of $\mathbb{F}_2$ is 1).
- Each matrix entry can be stored in a single bit of memory.
- Provided the matrix is column-major (and the columns each occupy a contiguous region of memory), we can subtract one column from another very efficiently using $\lceil n/w \rceil$ applications of the `XOR` machine instruction, where n is the number of elements in each column and w is the number of bits processed by each instruction. [1]
- Moreover, an adaptation of the *Method of Four Russians* algorithm for Gaussian elimination can be applied for the finite field $\mathbb{F}_q$, giving a further speedup proportional to $\log_q(n)$. This speedup is greatest when the finite field is as small as possible; the unique smallest field is $\mathbb{F}_2$.

Even after utilising all of the advantages of $\mathbb{F}_2$, reducing the matrix for the boundary map (restricted to dimension-d simplices) takes memory $O(mn)$ and time $\tilde{O}(nm^2)$, where m is the number of dimension-d simplices and n is the number of dimension-$(d-1)$ simplices. The paper [37] mentions that the same barcode can be computed by using row operations (taking time $\tilde{O}(mn^2)$, which is smaller for typical datasets as we tend to have many more dimension-d simplices than dimension-$(d-1)$ simplices). The authors proceed to suggest computing persistent *cohomology*, which yields the same barcode as persistent homology, as a more efficient alternative in practice.

Another approach is to simplify the simplicial complex to have fewer simplices whilst having the same homology. This can be done most effectively by working with the more general collection of CW-complexes: whilst the minimal simplicial complex homeomorphic to the torus has 7 vertices, 21 edges, and 14 triangles, the minimal

[1] A modern CPU with *Advanced Vector Extensions* has $w = 256$.

CW-complex homeomorphic to the torus has merely 1 vertex, 2 edges (loops attached at that vertex), and 1 square face. Given an initial simplicial complex, finding a simpler homotopy-equivalent CW-complex can be accomplished using *discrete Morse theory* [52]. This technique has been extended to the efficient computation of persistent homology by Mischaikow and Nanda in [53] and implemented in the software *Perseus* by the same authors.

Application to dataset of natural images

In [54], Mumford, Lee, and Pedersen explored the space of high-contrast 3×3 patches taken from greyscale natural images. They started with van Hateren's dataset of 4167 greyscale natural images, where the value of each pixel is given by the logarithm of the light intensity. For each image, 5000 random 3×3 squares of pixels were selected.

Each 3×3 square of pixels is regarded as a vector in a 9-dimensional space. The *contrast*, or *D-norm*, of the vector v is given by the square-root of the sum, over all pairs i, j of adjacent pixels, of the squared difference $|x_i - x_j|^2$ between their values. Subtraction of the mean of the pixel values corresponds to projection onto an 8-dimensional subspace on which the (squared) D-norm is a positive-definite quadratic form. Consequently, one can proceed to recover an inner product on this space:

$$\langle u, v \rangle := \frac{1}{4} \left(\|u + v\|_D^2 - \|u - v\|_D^2 \right).$$

The authors took a particular orthonormal basis with respect to this inner product, thereby identifying the space with $\mathbb{R}^8$. Of the 5000 random patches sampled from each image, the 1000 of highest contrast were retained, and the remainder discarded. Repeating across all the images, a dataset of 4167000 vectors in $\mathbb{R}^8$ were obtained; these were projected onto the unit sphere S^7 by normalising (dividing by the norm, *i.e.* contrast).

It was observed in [54] that this dataset $\mathcal{M}$ is far from being a uniform distribution over S^7. In particular, the authors partitioned the sphere into 17520 Voronoi cells, defined by the 17520 vectors of norm $\sqrt{8}$ in the E_8 lattice (described in [69]); they found that the distribution of points amongst these cells substantially differed from what one would expect if the points were sampled from the uniform measure on S^7.

The non-uniformity was explained in [41], in which Carlsson, Ishkhanov, de Silva, and Zomorodian explored the persistent homology of spaces obtained from $\mathcal{M}$.

2.1.6 Persistence diagrams and stability

One of the more important aspects of persistent homology, versus other methods for pattern recognition, is the idea of *stability*: small changes in the input result in small changes in the output. This is made rigorous by defining metrics on the input and output space, and showing that the map is Lipschitz (indeed, with Lipschitz constant equal to 1):

- The input space is taken to be $L_\infty(M)$, the space of bounded functions $M \to \mathbb{R}$, endowed with the uniform norm $d(f,g) = \sup_{x \in M} |f(x) - g(x)|$.
- The output space is the space of *persistence diagrams* endowed with a suitable metric (typically *bottleneck distance*, although we shall discuss other choices).

Definition 2.1.19. *A persistence diagram $D \in \mathcal{P}$ is considered to be a multiset of points in $\mathbb{R}^2$, where each point on the diagonal $\Delta := \{(x,x) : x \in \mathbb{R}\}$ appears with infinite multiplicity and each bar $[s,t]$ in the barcode gives rise to the off-diagonal point (s,t).*

Now we are ready to describe these metrics. The *bottleneck distance* is given by minimising, over all bijections $\eta : D \leftrightarrow D'$, the maximum distance between a point in D and its respective partner in D':

$$d_B(D,D') = \inf_{\eta : D \leftrightarrow D'} \sup_{x \in D} \|x - \eta(x)\|_\infty$$

The p-Wasserstein distance is defined by:

$$d_{W(p)}(D,D') = \inf_{\eta : D \leftrightarrow D'} \left(\sum_{x \in D} \|x - \eta(x)\|_\infty^p \right)^{1/p}$$

Much as the uniform norm on Euclidean space is given by the limit of the p-norm as $p \to \infty$, so the bottleneck distance is the limit of p-Wasserstein distances. The stability of bottleneck distance is proved by Edelsbrunner, Harer, and Cohen-Steiner in [40]:

Theorem 2.1.20 (Bottleneck stability)**.** *Let $f, g : M \to \mathbb{R}$ be continuous tame functions on a triangulable space M. Denote the persistence diagrams of the ith*

homology by $\Phi_i(f)$ and $\Phi_i(g)$, respectively. Then the bottleneck distance satisfies $d_B(\Phi_i(f), \Phi_i(g)) \leq \|f - g\|_\infty$.

Key in the proof is the following lemma which allows one to count the total multiplicity of points which lie in a rectangular region of the persistence diagram.

Lemma 2.1.21 (Generalised k-triangle lemma)**.** *Let $w < x < y < z$ be homological regular values. Then the total multiplicity of points in the persistence diagram within the box $[w, x] \times [y, z]$ is given by the alternating sum of persistent Betti numbers at the vertices of this box:*

$$|D \cap ([w, x] \times [y, z])| = \beta_w^z - \beta_x^z + \beta_x^y - \beta_w^y$$

Proof. Consider a bar $[a, b]$ in the barcode. If its left endpoint a does not lie between w and x, then its contributions to β_w^z and β_x^z will cancel out, as will its contributions to β_w^y and β_x^y. By symmetry, if b does not lie between y and z, the contributions of this bar to β_x^y and β_x^z will cancel out, as will the contributions to β_w^y and β_w^z.

Consequently, for $[a, b]$ to contribute to the alternating sum of persistent Betti numbers, we must have $w < a < x < y < b < z$. In that case, it is easy to see that the bar is included in the count for β_x^y but not for the other three Betti numbers, and therefore contributes $+1$ to the alternating sum.

Summing over all bars, we obtain the above identity.

□

In [40], it is deduced from the special case where $w = -\infty$ and $z = \infty$ (the *k-triangle lemma*), but here we have instead proved it directly from the formulation of a persistent Betti number as counting the number of bars which contain a given interval.

This allows statements in terms of multiplicities of points in persistence diagrams to be translated into statements about persistent Betti numbers, which ultimately boils down to dimension-counting. In particular, instrumental in the proof of bottleneck stability is the *box lemma*:

Lemma 2.1.22 (Box lemma, [40])**.** *Let f, g be tame functions with persistence diagrams $\Phi_d(f), \Phi_d(g)$, and suppose $\|f - g\|_\infty \leq \epsilon$. Suppose $a < b < c < d$ are homological regular values of g, and $a + \epsilon < b - \epsilon < c + \epsilon < d - \epsilon$ are homological regular values of f. Let $R = [a, b] \times [c, d]$ and $R' = [a + \epsilon, b - \epsilon] \times [c + \epsilon, d - \epsilon]$. Then the following bound holds:*

$$|\Phi_d(f) \cap R'| \leq |\Phi_d(g) \cap R|$$

This establishes bottleneck stability in the 'easy case' where the shortest distance between two distinct points in either $\Phi_d(f)$ or $\Phi_d(g)$ is greater than 2ϵ, since each point has exactly one candidate within a distance of ϵ giving an obvious bijection with distance ϵ. The authors extend this to a full proof of bottleneck stability by linearly interpolating between f and g to obtain a sequence of intermediate functions $f = h_0, h_1, \ldots, h_{k-1}, h_k = g$ where adjacent functions differ by a distance of $\frac{\epsilon}{k}$ in the uniform norm. Provided the distance between two distinct points in each $\Phi_d(h_i)$ is at most $2\frac{\epsilon}{k}$, and appealing to the triangle inequality for bottleneck distance, the result follows.

In [30], a stability result is proved for p-Wasserstein distance. However, unlike bottleneck distance, the Lipschitz constant is not 1; instead, it depends on properties of the ambient space M.

2.1.7 Multidimensional persistence

Recall that a filtration on a simplicial complex $\mathcal{K}$ is a family of subcomplexes $\mathcal{K}_t$ indexed by a parameter $t \in I$. The index set I is a totally-ordered set, typically $\mathbb{R}$, and $\mathcal{K}_s \subseteq \mathcal{K}_t$ whenever $s \leq t$.

A multifiltration generalises this by replacing I with a product space $I := I_1 \times \cdots \times I_k$ of k totally-ordered sets. I has the partial order where $s \leq t$ if each coordinate $s_i \leq t_i$. Hence, for a multifiltration, we have the condition that $\mathcal{K}_{s_1,\ldots,s_k} \subseteq \mathcal{K}_{t_1,\ldots,t_k}$ whenever $s_i \leq t_i$ for all $1 \leq i \leq k$. The special case of $k = 1$ corresponds to a filtration, and the case of $k = 2$ is usually called a *bifiltration*.

A *persistence module* over I is a functor M from I to the category of vector spaces over $\mathbb{F}$; that is to say, it associates a vector space M_s to each point in $s \in I$ and provides maps $\iota_{s,t} : M_s \to M_t$ whenever $s \leq t$ which compose ($\iota_{r,t} = \iota_{r,s} \circ \iota_{s,t}$ whenever $r \leq s \leq t$). When each $I_i \subseteq \mathbb{N}$, the persistence module has the structure of a graded module over $\mathbb{F}[x_1, \ldots, x_k]$.

Note that when $k \geq 2$ this is not a principal ideal domain, with the ideal $\langle x_1, x_2 \rangle$ not generated by any one element. Consequently, the Structure Theorem does not apply, and there is no convenient analogue of a barcode in the multi-parameter (or *multidimensional*) case. Carlsson and Zomorodian further showed in [44] that there is no complete discrete invariant for multidimensional persistence; that is to say, if the

base field $\mathbb{F}$ is uncountable, then so is the set of isomorphism classes of persistence modules. Contrast this with unidimensional persistence, in which there are only countably many barcodes irrespective of the cardinality of the field $\mathbb{F}$.

The authors also proposed a discrete invariant, the *rank invariant*, which is equivalent to the barcode for unidimensional persistence; however, it is necessarily an incomplete invariant for $k \geq 2$ by the previous theorem. For a multifiltration $\mathcal{K}$ and non-negative integer i, the *rank invariant* $\rho_{\mathcal{K},i}(s,t)$ is defined to be the rank of the map between the ith homology groups, $H_i(\mathcal{K}_s)$ and $H_i(\mathcal{K}_t)$, induced by the inclusion map $\iota_{s,t}$.

Note that when $k = 1$, this is identical to the ith persistent Betti number, so (by the aforementioned generalised k-triangle lemma) two filtrations have the same rank invariants if and only if they have the same persistence diagrams or equivalently homology barcodes.

Interleaving distance

In [45], Michael Lesnick takes an alternative approach: instead of attempting to find a topological invariant (such as a barcode) for multidimensional persistence, the author directly defines a pseudometric (the *interleaving distance*) between multidimensional persistence modules.

The poset I is taken to be $\mathbb{R}^k$. Given a persistence module M and $\epsilon \geq 0$, the *shift* $M(\epsilon)$ has vector spaces $M(\epsilon)_s := M_{s+\epsilon}$ (where addition $s + \epsilon$ is interpreted elementwise) and the map from $M(\epsilon)_s$ to $M(\epsilon)_t$ is defined to be the map from $M_{s+\epsilon}$ to $M_{t+\epsilon}$. The *ϵ-transition morphism* is defined as the map from M to $M(\epsilon)$ given by $\iota_{s,s+\epsilon}$ for each M_s.

Modules M and N are described as *ϵ-interleaved* if there exist morphisms $f : M \to N(\epsilon)$ and $g : N \to M(\epsilon)$ where the compositions $g(\epsilon) \circ f$ and $f(\epsilon) \circ g$ are equal to the 2ϵ-transition morphisms $M \to M(2\epsilon)$ and $N \to N(2\epsilon)$, respectively. Clearly, if $\delta \geq \epsilon$ then ϵ-interleaved modules are also δ-interleaved; this leads to the notion of the *interleaving distance* $d_I(M,N)$ as the infimum ϵ such that the two modules are ϵ-interleaved.

The author proves that when the base field $\mathbb{F}$ is either a prime finite field $\mathbb{F}_p$ or the rationals $\mathbb{Q}$, then d_I has the same stability property as bottleneck distance (which coincides with d_I in the case $k = 1$). Moreover, the interleaving distance has the universality property that if d is another stable pseudometric, we have $d(M,N) \leq d_I(M,N)$ for all pairs M, N of persistence modules.

2.2 Shape spaces

So far, we have primarily concentrated on points in Euclidean space, $X \subseteq \mathbb{R}^d$, whilst alluding to the fact that certain data more naturally inhabit non-Euclidean spaces. An example of this occurs when we study configurations of points modulo isometries.

Suppose we have k labelled points $X \subseteq \mathbb{R}^d$, represented as a $k \times d$ matrix K. The configuration space is dk-dimensional, but subtracting the mean corresponds to projecting down to a $d(k-1)$-dimensional subspace. This can be isometrically identified with $\mathbb{R}^{d(k-1)}$. We make this explicit as follows.

Definition 2.2.1. *Let $k \geq 2$ be a positive integer. Then the* demeaning matrix [2] *H_k is the orthogonal symmetric self-inverse $k \times k$ matrix:*

$$H_k := \begin{pmatrix} c & c & c & c & \cdots & c & c \\ c & a & b & b & \cdots & b & b \\ c & b & a & b & \cdots & b & b \\ c & b & b & a & & b & b \\ \vdots & \vdots & \vdots & & \ddots & & \vdots \\ c & b & b & b & & a & b \\ c & b & b & b & \cdots & b & a \end{pmatrix}$$

where $c = \frac{1}{\sqrt{k}}$, $b = \frac{1}{\sqrt{k}-k}$, and $a = 1 + \frac{1}{1-\sqrt{k}} + \frac{1}{\sqrt{k}}$. The reduced demeaning matrix *H_k' is the $(k-1) \times k$ matrix obtained by deleting the first row from H_k.*

Then, premultiplying the configuration by the reduced demeaning matrix yields the *pre-size-and-shape* $L = H_k'K$, which is a $(k-1) \times d$ matrix. Two labelled point sets X_1 and X_2 share the same pre-size-and-shape matrix L if and only if X_1 is a translate of X_2.

Often we want to quotient further, by either rotations, reflections, scaling, or some combination thereof. The *pre-shape* is obtained by normalising the pre-size-and-shape to have unit Hilbert-Schmidt norm. The prefix 'pre-' indicates that we have yet to quotient by $SO(d)$, the resulting spaces being the *shape space* and *size-and-shape space*. If we quotient by the full orthogonal group $O(d)$ as opposed to merely $SO(d)$, we obtain the *reflection shape space* and *reflection size-and-shape space*, respectively.

The relationships are summarised in Figure 2.5.

[2] Based on the negative answer to http://mathoverflow.net/questions/262091/, it appears that this matrix does not have a standard name. We introduce this terminology since the matrix isolates the mean vector from the shape information.

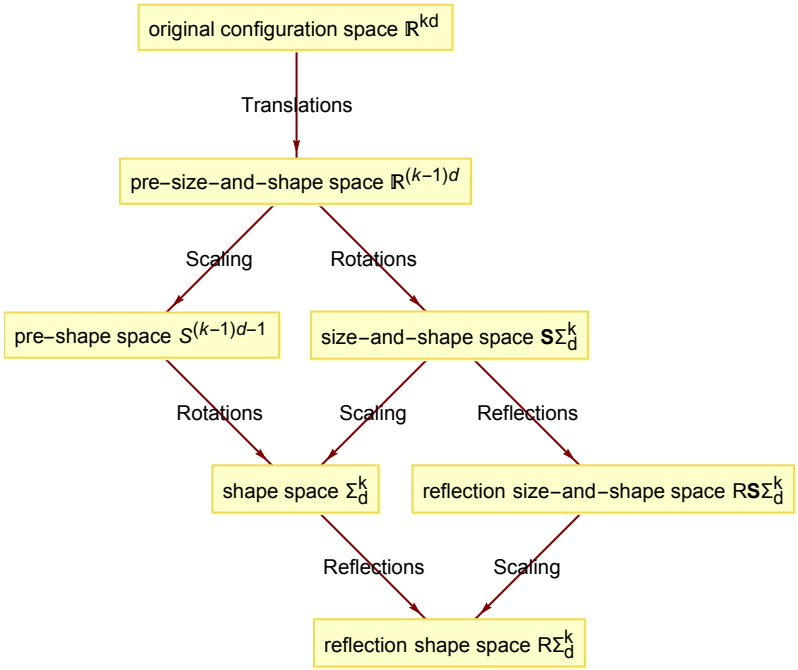


Fig. 2.5 Lattice of spaces corresponding to quotienting by different groups of isometries and similarity transforms. A similar diagram occurs in [64].

The pre-size-and-shape space $\mathbb{R}^{(k-1)d}$ and pre-shape space $S^{(k-1)d-1}$ each come equipped with natural metrics: the Euclidean distance in the case of the pre-size-and-shape space, and the geodesic distance in the case of the pre-shape space. The other four spaces each arise as the quotient of one of these spaces by the action of an isometry group G (either $O(d)$ or $SO(d)$).

Lemma 2.2.2. *Let X be either the pre-size-and-shape space $\mathbb{R}^{(k-1)d}$ or the pre-shape space $S^{(k-1)d-1}$ and let G be either $O(d)$ or $SO(d)$. Denote the aforementioned natural metric on the space X by ℓ, and let $\pi : X \to X/G$ be the quotient map.*

Consider any pair of points $x, y \in X/G$ in the quotient space, and let $\tilde{x}_1, \tilde{x}_2$ be arbitrary points in the fibre $\pi^{-1}(x)$. Define $\tilde{y}_1, \tilde{y}_2$ to be the points in the fibre $\pi^{-1}(y)$ closest to $\tilde{x}_1$ and $\tilde{x}_2$, respectively.

Then $\ell(\tilde{x}_1, \tilde{y}_1) = \ell(\tilde{x}_2, \tilde{y}_2)$, and therefore $\ell(\tilde{x}_1, \tilde{y}_1)$ is dependent only on the points $x, y \in X/G$ and not in the particular choice of preimage $\tilde{x}_1$.

Proof. This follows from the fact that the group G of isometries acts transitively on each fibre, so we can find an isometry g mapping $\tilde{x}_1$ to $\tilde{x}_2$. The image $g(\tilde{y}_1)$ remains on the fibre $\pi^{-1}(y)$, so it follows that:

$$\ell(\tilde{x}_2, g(\tilde{y}_1)) \geq \ell(\tilde{x}_2, \tilde{y}_2)$$

by the defining property of $\tilde{y}_2$ being the closest point on $\pi^{-1}(y)$ to $\tilde{x}_2$. The left-hand-side of this inequality is equal to $\ell(\tilde{x}_1, \tilde{y}_1)$ as g is an isometry.

$$\ell(\tilde{x}_1, \tilde{y}_1) \geq \ell(\tilde{x}_2, \tilde{y}_2)$$

By symmetry, the inequality also holds in the other direction, and therefore the two distances must be equal. □

Consequently, the spaces Σ_d^k, $S\Sigma_d^k$, $R\Sigma_d^k$, and $RS\Sigma_d^k$ are all endowed with the structure of a metric space. Away from singularities (images of points where the derivative of π is not full-rank), this is a Riemannian metric as demonstrated in [61]. The geodesics in the quotient space are the images (under π) of geodesics in the top space.

2.2.1 Examples of shape spaces

When $d = 1$, $SO(d)$ is the trivial group and therefore the shape space Σ_1^k coincides with the pre-shape space S^{k-2}. For the reflection shape space, we further quotient by reflection, ergo $R\Sigma_1^k = \mathbb{RP}^{k-2}$ is real projective space.

When $d = 2$, the pre-size-and-shape space can be viewed as $(k-1)$-tuples of complex numbers. Two such tuples represent the same shape if they can be related by a scaling and rotation, or equivalently if they are equal up to multiplication by a complex scalar. Consequently, $\Sigma_2^k = \mathbb{CP}^{k-2}$, and the metric is exactly the Fubini-Study metric. The reflection shape space is obtained by identifying tuples which are complex conjugates of each other: so $R\Sigma_2^k = \mathbb{CP}^{k-2}/\sim$. This is smooth except at the submanifold corresponding to $\mathbb{RP}^{k-2}$.

For $d \geq 3$, the shape spaces have singularities and are no longer homogeneous; this is discussed in [61].

As for the size-and-shape counterparts, they are cones on the shape spaces. The distance from the apex of the cone is given by the Hilbert-Schmidt norm of the pre-size-and-shape matrix L.

2.2.2 Geodesics and Procrustes distances

Earlier, we mentioned that the geodesics in the shape spaces Σ_d^k, $S\Sigma_d^k$, $R\Sigma_d^k$, and $RS\Sigma_d^k$ are precisely the projections of the geodesics in the original space under π. This means that given a pair of representatives (or *icons*) x, y of a pair of shapes, the quotient distance in the appropriate shape space is given by:

$$\ell'(\pi(x), \pi(y)) = \min\{\ell(x, \sigma(y)) : \sigma \in G\}$$

where G is the relevant isometry group (either $SO(d)$ or $O(d)$), ℓ is the geodesic distance in the top space, and ℓ' is the corresponding distance in the quotient space as described in Lemma 2.2.2. This particular choice of metric ℓ' is called the *Procrustes distance*.

For the aforementioned spaces, this can be computed effectively. In terms of matrices L_1, L_2 in the pre-shape (or pre-size-and-shape) space, we want to find the (optionally special) orthogonal matrix R which minimises the Hilbert-Schmidt norm of the matrix $L_1 - L_2R$. In that case, we can compute the Procrustes distance as either the straight-line or great-circle distance depending on whether the top space is $\mathbb{R}^{(k-1)d}$ or the pre-shape sphere $S^{(k-1)d-1}$:

$$\ell(x, y) = \begin{cases} \|L_1 - L_2R\| & \text{for size-and-shape;} \\ 2\arcsin(\frac{1}{2}\|L_1 - L_2R\|) & \text{for shape.} \end{cases}$$

So, how do we find the optimal matrix $R \in G$ (where G is either $SO(d)$ or $O(d)$)? This is called the *orthogonal Procrustes problem*. Somewhat remarkably, Schönemann discovered an exact solution using singular value decomposition in [62].

Theorem 2.2.3 (Schönemann, 1966)**.** *Let UDV^T be the singular value decomposition of $L_2^T L_1$, where D is diagonal matrix of singular values in descending order. Then, the optimal $R \in O(d)$ is given by $R = UV^T$, and the optimal $R \in SO(d)$ is given by UJV^T, where $J = diag(1, 1, \ldots, 1, \pm 1)$ and the sign on the last term is chosen to ensure R has determinant 1.*

2.2.3 Generalised Procrustes analysis

In statistics, when we have a collection $x_1, x_2, \ldots, x_n$ of points in Euclidean space, we are often interested in the mean and covariance matrix. It is not inconceivable that the same properties are interesting when we are operating in a shape space, but

actually calculating the 'mean shape' is less easy: we are not in some vector space where we can just calculate $\frac{1}{n}(x_1 + x_2 + \cdots + x_n)$. Instead, we need to use a different characterisation of mean which generalises to other Riemannian manifolds.

Definition 2.2.4. *Given a set $x_1, x_2, \ldots, x_n$ of points lying in a metric space (M, d), a* Fréchet mean *is a point $y \in M$ which minimises the sum of squared distances, $\sum_{i=1}^{n} d(x_i, y)^2$.*

This coincides with our notion of mean in Euclidean space. The *generalised Procrustes algorithm* attempts to locate this iteratively. At a point $y \in M$, we have a tangent space T_y and the *exponential map* $\exp_y : T_y \to M$ defined by mapping the tangent vector v to the endpoint of the geodesic which starts at y, runs parallel to v, and has a length of $|v|$.

With this in place, we define $\Phi(y)$ to be the image (under the exponential map) of the mean of the preimages of the points. That is to say:

$$\Phi(y) := \exp_y \left(\frac{1}{n} \sum_{i=1}^{n} \exp_y^{-1}(x_i) \right)$$

In [63], it is noted that the right-hand side is proportional to the gradient of the function $F(y) = \sum_{i=1}^{n} d(x_i, y)^2$. Hence, as a Fréchet mean is a minimiser of F, we have $\nabla F = 0$ and thus the Fréchet mean is a fixed point of Φ.

The generalised Procrustes algorithm iteratively applies Φ, starting from x_1 without loss of generality, with the hope that the iteration converges to a Fréchet mean. It is shown in [63] that if the points are bounded by a sufficiently small ball, the map Φ is a contraction mapping and therefore the iteration converges to a (unique) Fréchet mean. Certain positively-curved spaces, such as the sphere S^2, do not have a unique Fréchet mean in all cases; for example, the two-point set consisting of two antipodal points has an infinite family of Fréchet means forming a great circle. Note that $S^2 = \mathbb{CP}^1 = \Sigma_2^3$ is the size-and-shape space for configurations of three points in $\mathbb{R}^2$.

It is informative to contextualise this by seeing how the generalised Procrustes algorithm would apply to the shape spaces. If the points are represented by pre-shape icons $L_1, L_2, \ldots, L_n$, the Procrustes iteration is given by:

$$\Phi(Y) := \frac{1}{n} \sum_{i=1}^{n} L_i R_i$$

where $R_i \in G$ is the solution to the orthogonal Procrustes problem of most closely matching L_i onto Y. As discussed, this can be solved using the singular value decomposition, and does not involve explicitly applying the exponential map or its inverse.

Tangent-space analysis and principal components

Once we have the Fréchet mean y, we can analyse the distribution of shapes $x_1, x_2, \ldots, x_n$ in terms of their preimages under $\exp_y$. In particular, the tangent space has an inner product (the Riemannian metric) so we can define the sample covariance operator:

$$\Sigma(u, v) = \frac{1}{n-1} \sum_{i=1}^{n} \langle \exp_y^{-1}(x_i), u \rangle \langle \exp_y^{-1}(x_i), v \rangle$$

In [64], this covariance operator is decomposed in terms of its principal components, which are vectors in the tangent space at y. They visualise this by applying the exponential map to these vectors to obtain shapes characterising each principal component.

The iterative algorithm for computing Fréchet means and the subsequent principal component analysis both generalise to other Riemannian manifolds. The only place where we utilised the special structure of shape spaces was in the closed-form iteration involving singular value decomposition. For other manifolds we would need to explicitly switch between the tangent spaces and the manifold itself by means of the exponential map and its inverse.

2.3 Metrics on covariance matrices

In the previous section, we studied spaces of configurations of points up to some group of isometries. We had a natural choice of metric on these spaces, namely the Procrustes distance obtained by taking the quotient metric. We are not always so fortunate in the existence of such a natural metric, and in many cases there are multiple equally convincing candidates. It is therefore desirable to determine to what extent the methods of topological data analysis are affected by the choice of underlying metric.

The space of positive-semidefinite $d \times d$ matrices is particularly rich in terms of candidate metrics. Covariance matrices and Riemannian metrics are examples

of positive-semidefinite matrices (or equivalently quadratic forms), as are diffusion tensors [65].

2.3.1 Flat metrics

We can firstly view this space as a convex cone lying in the space of symmetric matrices. The apex of the cone is the zero matrix, and the axis is the ray of positive scalar multiples of the identity matrix. The slices of the cone orthogonal to this axis are spaces comprising matrices of constant trace.

The ambient space of symmetric matrices is endowed with the Hilbert-Schmidt norm induced by the inner product $\langle A, B\rangle = \mathrm{tr}(A^T B)$. Following [65], we shall refer to the metric induced by this norm as the *Euclidean* metric on the space of positive-semidefinite matrices.

$$d_E(A, B) = \|A - B\|_2$$

Whilst very natural, it is not geodesically complete; geodesics in this space stop abruptly at the boundary (the singular matrices with determinant 0). In an attempt to remedy this, it is tempting to take either the matrix logarithm (which sends the boundary to infinity) or square-root (which causes the geodesics to smoothly deflect away from the boundary like parabolic arcs). This defines the *logarithmic* and *square-root* metrics:

$$d_L(A, B) = \|\log(A) - \log(B)\|_2$$
$$d_H(A, B) = \|\sqrt{A} - \sqrt{B}\|_2$$

Computationally, these metrics can be obtained by writing each matrix in the form UDU^T (where U is orthogonal and D is diagonal), and applying the respective operation (logarithm or square-root) elementwise to D. In [65], Dryden *et al* also consider the *power-Euclidean* metric where the eigenvalues are raised to some power $\alpha \in (0, 1]$. The Euclidean and square-root metrics are recovered by taking α to be 1 and $\frac{1}{2}$, respectively.

Yet another flat metric, which is computationally faster to compute, is the *Cholesky* metric where each matrix A is mapped to its lower Cholesky factor L satisfying $LL^T = A$. An interesting caveat is that the Cholesky distance between

two quadratic forms is dependent on our choice of orthonormal basis, so it would not be appropriate for inherently isotropic data such as diffusion tensors.

2.3.2 Procrustes metric

In both the Cholesky and square-root metrics, we map each matrix A to some particular L (either square-root or lower Cholesky factor) satisfying $LL^T = A$. The Procrustes distance $d_S(A_1, A_2)$ is defined to be the infimum distance $\|L_1 - L_2\|$ over all choices L_1, L_2 such that $L_i L_i^T = A_i$. By definition, the Procrustes distance is no greater than either the square-root or Cholesky distance, and in general is smaller.

Note that $LL^T = L'L'^T$ if and only if $L' = LR$ for some orthogonal matrix R. Consequently, $d_S(A_1, A_2)$ is the distance in $RS\Sigma_d^{d+1}$ between the shapes represented by pre-size-and-shape matrices L_1, L_2. Our choice of representatives L_1, L_2 is arbitrary; we choose the Cholesky factor since it is computationally easier to compute than the square-root.

Unlike the Cholesky distance, the Procrustes distance is again isotropic (independent of the choice of orthogonal basis). As it is isometric to the size-and-shape space $RS\Sigma_d^{d+1}$, rather than Euclidean space, all of the sectional curvatures are positive. It is more difficult to compute Čech and alpha complexes (compared with the 'flat' metrics); even finding the minimum metric ball bounding a set of points is no longer a convex optimisation problem.

2.3.3 Affine-invariant Riemannian metric

In [58], a negatively-curved metric is proposed on the space of positive-definite matrices (the singular matrices being sent to infinity). It is the metric obtained from giving the space an affine-invariant Riemannian metric; the space becomes a geodesically-complete manifold without boundaries. In particular, given positive-definite matrices A, B, the distance is defined as follows:

$$d_R(A, B) = \|\log(A^{-\frac{1}{2}} B A^{-\frac{1}{2}})\|_2$$

As with the positively-curved Procrustes metric, we can use the generalised Procrustes algorithm to compute Fréchet means, and perform tangent space analysis to compute more sophisticated statistics. Due to the negative sectional curvature of the space, the Fréchet mean is guaranteed to be unique.

2.4 Nonlinear dimensionality reduction

2.4.1 Multidimensional scaling

Multidimensional scaling, or MDS, is a 'global' method of nonlinear dimensionality reduction, which attempts to find an embedding $\phi : (X, d) \to \mathbb{R}^n$ that preserves distances as closely as possible. More precisely, *metric MDS* seeks to minimise the following loss function (known as *stress*):

$$\sum_{x,y \in X} \left(d(x, y) - \|\phi(x) - \phi(y)\|_2\right)^2$$

In [13], de Leeuw proposed an iterative algorithm for optimising this loss function. Note that this is not necessarily guaranteed to converge to a global optimum.

In the special case where X is a finite subset of a Euclidean space and d is the Euclidean metric, MDS is actually equivalent to PCA as mentioned in [14]: the global optimum of the stress function is attained by projecting onto the first n principal components. As such, applying MDS directly to the matrix of Euclidean distances cannot find a nonlinear embedding.

2.4.2 Isomap

Instead of using the subspace metric, Tenenbaum and de Silva proposed in [14] the idea of applying metric MDS to an estimate of the intrinsic geodesic distances along the submanifold of interest. The *Isomap* algorithm accomplishes this as follows:

- Connect each point to its k nearest neighbours to obtain a 'neighbourhood graph'. The edges are labelled with the corresponding distances (in the ambient, usually Euclidean, metric space). It is assumed that k is sufficiently large as to ensure the graph is connected.
- Perform Dijkstra's algorithm to compute the pairwise path distances through this graph.
- Apply metric MDS to the resulting matrix of pairwise path distances.

Isomap can be considered a generalisation of multidimensional scaling, degenerating into the latter when $k = |X| - 1$ and the neighbourhood graph is complete. Unlike multidimensional scaling, however, Isomap is capable of finding nonlinear embeddings.

2.4.3 Locally linear embedding

Locally linear embedding, or LLE, was introduced in [15]. Suppose the ambient space is Euclidean, so $X \subset \mathbb{R}^D$. Each vertex $x_i \in X$ is expressed as an affine combination of its k nearest neighbours; that is to say, we find weights w_{ij} (which are zero unless x_j is in the k nearest neighbours of x_i) minimising the loss function:

$$\left\| x_i - \sum_{x_j \in X} w_{ij} x_j \right\|_2^2$$

subject to the condition that $\sum_{x_j \in X} w_{ij} = 1$. This is a least squares problem solvable with linear algebra. If necessary, a ridge-like penalty term can be incorporated into this quadratic form to aid regularisation.

Once these weights have been established, embedded vectors $\phi(x_i)$ are constructed by minimising the loss function:

$$\sum_{x_i \in X} \left\| \phi(x_i) - \sum_{x_j \in X} w_{ij} \phi(x_j) \right\|_2^2$$

To prevent the obvious degeneracy of ϕ being identically zero, the embedded vectors are constrained to be normalised with identity covariance matrix. Moreover, the translation invariance of the problem allows one to additionally impose the simplifying constraint that the embedded vectors have zero mean without loss of generality.

This can be generalised to non-Euclidean spaces. In particular, Appendix C of [17] describes how the weights w_{ij} about each point x_i can be determined from the matrix of pairwise distances in the set $\{x_i\} \cup \Gamma(x_i)$ of the point and its neighbourhood.

2.4.4 Other approaches

The cornucopia of methods for nonlinear dimensionality reduction is rather too vast to adequately cover in detail; however, it is worth noting some of the more novel approaches.

In [7], Hinton and Salakhutdinov applied a deep neural network to the problem of nonlinear dimensionality reduction. In particular, the authors took an *autoencoder*, i.e. a neural network comprising an 'encoder' $E : A \to \mathbb{R}^n$ from the ambient space to a *latent space* composed with a 'decoder' $D : \mathbb{R}^n \to A$. Using stochastic gradient descent trained to reproduce the identity function on $X \subseteq A$, the network attempts

to find a continuous map $X \to \mathbb{R}^n$ which is approximately invertible. Equivalently, it can be viewed as a (lossy) compression algorithm mapping vectors in the ambient space to vectors in the much lower-dimensional latent space.

Another popular method is t-SNE, proposed by Hinton and van der Maaten in [8]. It is generally applied for embedding multidimensional data into two dimensions for the purposes of visualisation. It does not perform as well when the ambient space is of extremely high dimensionality, so it is recommended in [9] to use an autoencoder to reduce to $\mathbb{R}^{32}$ before applying t-SNE on the latent embedding in order to visualise the result.

2.5 Functional data

In classical statistics, our data are usually a collection of points in Euclidean space. We discussed how for certain data, such as shapes and covariance matrices, it can be more natural to think of these points as lying in a non-Euclidean manifold (possibly with boundary). An independent generalisation is to consider data in infinite-dimensional spaces; for instance, the ambient space $\mathbb{R}^n$ can be replaced with an infinite-dimensional Hilbert space $\mathcal{H}$.

One particular source of infinite-dimensional data is where each data point is a function f, typically with assumptions about continuity or smoothness. The output from a thermometer will be a (continuous) function of time giving the temperature $T(t)$ at time t. As with covariance matrices, there are often many distances used to compare two functions: uniform norm and L_2 norm being among the most common metrics.

2.5.1 Core concepts

In [18], Horváth and Kokoszka review many of the concepts behind *functional data analysis*. A few of these ideas are relevant here, so we shall discuss them briefly.

Function space

Up until now, our data have usually been a finite subset $X \subseteq M$ of a finite-dimensional metric space, such as Euclidean space, or a Riemannian manifold, or subset thereof. In functional data analysis, we typically begin with the space $\mathcal{L}_p[K]$

of p-power-integrable functions on a compact measurable space K. Most commonly, $p \in \{1, 2, \infty\}$ and K is an interval $[a, b] \subseteq \mathbb{R}$.

Technically, $\mathcal{L}_p[K]$ is actually the metric space obtained by quotienting out the pseudo-metric space $L_p[K]$ of p-power-integrable functions, with distance function $d(f, g) = \|f - g\|_p$.

The vast majority of functions in $\mathcal{L}_p[K]$ are insensibly wild. For example, returning to the thermometer example, our function is assumed to be continuous. Making assumptions on continuity or smoothness restricts us down to a more manageable space $M \subsetneq \mathcal{L}_p[K]$.

The case $p = 2$ is particularly important, since $\mathcal{L}_2[K]$ has an inner product $\langle f, g\rangle := \int f(x)g(x)$: rather than being merely a Banach space, it is a Hilbert space. This property turns out to be necessary for many statistical applications; consequently, we henceforth operate in a space $M \subseteq \mathcal{L}_2[K]$ unless otherwise specified.

In particular, it is mentioned in [18] that for a Hilbert space H, a symmetric positive-definite Hilbert-Schmidt operator Ψ admits a decomposition as follows:

$$\Psi(f) = \sum_{j=1}^{\infty} \lambda_j \langle f, v_j\rangle v_j$$

where the λ_j are the eigenvalues and v_j are orthonormal eigenfunctions: $\langle v_i, v_j\rangle = \delta_{ij}$.

Dimension reduction

Even though our space is infinite-dimensional, we can only manipulate finite amounts of data. Usually, we represent a function f by its values sampled at some finite subset $S \subseteq K$. If K is an interval, we typically choose the points in S to be equally spaced; similarly, if K is a product of intervals, it is convenient to take S to be a regular grid.

It is worth noting that $|S|$, the dimension of the space of interest, is often too large to be manipulable. If f is (say) the temperature map of a human organ, and we sample 100 points in each of the three dimensions, then $|S| = 10^6$. Whilst storing a vector of $d = 10^6$ elements is straightforward on a modern machine, it is impractical to perform linear algebra: matrices would have $d^2 = 10^{12}$ entries, and even multiplying matrices takes $d^3 = 10^{18}$ operations. More sophisticated algorithms can reduce this to $O(d^{2.807})$ (Strassen) or even $O(d^{2.375})$ (Coppersmith-Winograd), but this is still too impractical for statistical applications.

Computational efficiency is not the only reason why we may want to reduce the dimensionality: representing f by its values sampled at points of S ignores assumptions such as continuity. Stated another way, we are taking our basis functions to be 'peaks' supported at individual points of S. In the limit as $|S| \to \infty$, the Lipschitz constants associated with the basis functions similarly increase without bound. These are less sensible choices of function than (say) low-degree polynomials, splines, and wavelets.

Consider the thermometer example. If we take temperature readings every 5 seconds as opposed to every 10 seconds, we do not expect to have twice as much information: the gradual nature of temperature changes mean that we get diminishing returns from increasing the sampling frequency. It stands to reason that we can reduce the dimensionality without discarding much data.

Assuming continuity, one approach is to express our function (sampled at d points) in the Fourier basis and retain only the first $k \ll d$ terms. In the thermometer example, taking the first 80 Fourier coefficients of a day's worth of 10-second samples will reduce the dimension from 8640 to 80.

Often, it is desirable to reduce the dimension even further, by (for example) projecting the dataset onto its first $p \ll k$ *functional principal components*. The second projection is subtly different, because our basis depends on the data itself rather than being a fixed (e.g. Fourier) basis.

Mean functions and covariance operators

Given a finite sample of functions, $f_1, \ldots, f_n$, we can compute the sample mean function $\mu = \frac{1}{n} \sum_{i=1}^{n} f_i$. Similarly, as in the finite-dimensional case, one can compute the sample covariance operator $C(g) = \frac{1}{n} \sum_{i=1}^{n} \langle f_i - \mu, g \rangle (f_i(x) - \mu(x))$. The *(empirical) functional principal components*, which were alluded to above but not defined, are the eigenfunctions of this operator.

When we project to the space spanned by the first p EFPCs, we need some method of choosing p. Two common choices are below:

- Choose p to be a fixed number, such as 5;
- Choose p to be the minimum number to capture some proportion of the variance (in [18], the authors recommend 85 percent). That is to say, we choose the

minimal p satisfying $\sum_{i=1}^{p} \lambda_i \geq 0.85 \sum_{i=1}^{\infty} \lambda_i$, where λ_i is the ith largest eigenvalue of the covariance operator.

2.5.2 Speech spectrograms and frequency covariance matrices

In [59], the authors examine a set of speech recordings of the words 'one' through to 'ten' spoken in five different Romance languages: American Spanish, French, Portuguese, Iberian Spanish, and Italian. There are 50 language/word ordered pairs, each of which has at least one recording. The total number of recordings is 219, since some ordered pairs are represented by more than one recording.

By applying a local Fourier transform, the recordings were converted into spectrograms (smooth functions of time and frequency, which lie in the Hilbert space $L_2([0, F] \times [0, T])$). The spectrogram was modified by applying the following in succession:

- **Smoothing**: the original sound samples are noisy, whereas a spectrogram is ideally a smooth surface. The authors approach this by using smoothing splines.
- **Alignment**: different speakers enunciate words at different speeds, so the data are not directly comparable. To address this, each spectrogram is warped in the time dimension by applying an order-preserving bijection to the time axis (chosen to align the sounds optimally, subject to a penalty according to how this warping departs from the identity function).

Whilst being continuous surfaces in theory, the spectrograms are represented as 8100-dimensional vectors (by sampling on a grid with 81 frequency divisions and 100 points in time).

For each of the 50 language/word ordered pairs, the authors computed the mean and covariance matrix of the time-slices of the spectrograms (if there are n recordings, this gives $100n$ time-slices, each of which is an 81-dimensional vector). This yields a dataset X comprising 50 different 81×81 covariance matrices – one for each language/word ordered pair.

Chapter 3

Topological approaches to manifold learning

3.1 Detectability

Recall that we have some unknown subspace $M \subseteq A$ of the ambient space, and each Y_i is an independent identically distributed random variable distributed according to a measure μ on M. Moreover, the observations $X_i := Y_i + \varepsilon_i$ are afflicted by noise.

One problem when ε_i is unbounded is that, in the limit where we sample infinitely many data points, X will be dense in the ambient space A. Following [41], we need to remove these outliers before constructing a simplicial complex approximating M, by taking super-level sets with respect to some estimate of the probability density function f of the random variables X_i.

In 2006, Niyogi, Smale, and Weinberger approached a similar problem in [2], but with several key differences to the work in this chapter:

- In [2], the noise model ε_i has bounded support with radius no greater than the normal injectivity radius of the manifold. On the other hand, this chapter allows for noise with unbounded support, with particular interest in spherical Gaussian noise.

- The paper [2] takes, as input, a finite set of points X_i. By contrast, this chapter also requires a density estimate of the distribution at each point X_i. As such, to use this methodology in practice, one would need to first use a density estimating algorithm (such as the kernel density estimator by Parzen [42] and

Rosenblatt [70]) to obtain a density estimate at each point X_i from the set of points.

- In [2], the noise model is such that the manifold is guaranteed to be *detectable* (to use the terminology introduced in this chapter), and the paper obtains a result on the number n of points that must be sampled to correctly deduce the homology of the original manifold with probability $1 - \delta$. Instead, this chapter somewhat complements this by instead examining under what conditions it is possible to correctly deduce the homology at all, in the asymptotic regime where the number of sampled points $n \to \infty$ and the density estimate converges to the actual probability density function of X_i.

- In the simplest nontrivial case where M is the unit sphere and ε_i is spherial Gaussian with variance σ^2, the maximum value of σ^2 such that the manifold remains detectable is determined exactly in a closed form involving the *confluent hypergeometric limit function*.

In 2011, Niyogi, Smale, and Weinberger published a sequel [3] which does include a Gaussian noise model. However, it is not equivalent to the noise model studied in this chapter. In particular, the noise vector ε_i in [3] is restricted to the normal space at the point Y_i on the manifold, whereas the noise in this chapter is isotropic and (almost surely) has a tangential component as well as a normal component.

This makes a huge difference. For the unit sphere, the normal lines all 'focus' at the origin. This causes a singularity under the 'normal space noise' model in [3], where the probability density at the origin diverges to infinity. The authors circumvent these isolated *hotspots* by 'choosing with care the size of the neighborhoods for cleaning the data'. By contrast, isotropic noise does not result in hotspots where the density diverges, and indeed this is key to the particular formulation of manifold detectability in this chapter.

In [12], Genovese, Perone-Pacifico, Verdinelli, and Wasserman consider the same additive noise model as this chapter, but their objective is different: they are interested in recovering the manifold to within a specified Hausdorff distance rather than determining its topology up to homotopy equivalence. This is still relevant to this chapter because we show that a particular subset P of our sampled points X_i is contained within the r_2-thickening M_{r_2} of the manifold, and similarly the manifold M is contained within the r_1-thickening of P; the Hausdorff distance is upper-bounded by the maximum of r_1 and r_2 by definition. We use this 'asymmetric Hausdorff

bound' as a stepping-stone to finding a thickening of P which deformation-retracts onto our manifold, and it is convenient to let r_1 and r_2 be different (as opposed to a Hausdorff bound, where $r_1 = r_2$).

If we could estimate the probability density function f exactly, and then build a Čech complex supported on the high-density points of X (those with $f(X_i) \geq \delta$ for some threshold δ), we could recover M up to homotopy equivalence provided the following holds:

Definition 3.1.1. *We say (M, μ) is* detectable *in the presence of noise ε if there exists some δ and radius $R > 0$ such that $S(\delta)_R$ (the R-thickening of the super-level set $S(\delta)$) both contains M as a subset and admits a deformation retraction to M. This implies, in particular, that it is homotopy equivalent to M – and, by the nerve theorem, so is the corresponding Čech complex.*

It is worth justifying why we choose the stronger condition of $S(\delta)_R$ admitting a deformation retraction to M, rather than merely being homotopy equivalent to M. Specifically, it avoids the situation in Figure 3.1 where the large annulus is only 'accidentally' homotopy equivalent to the small annulus.

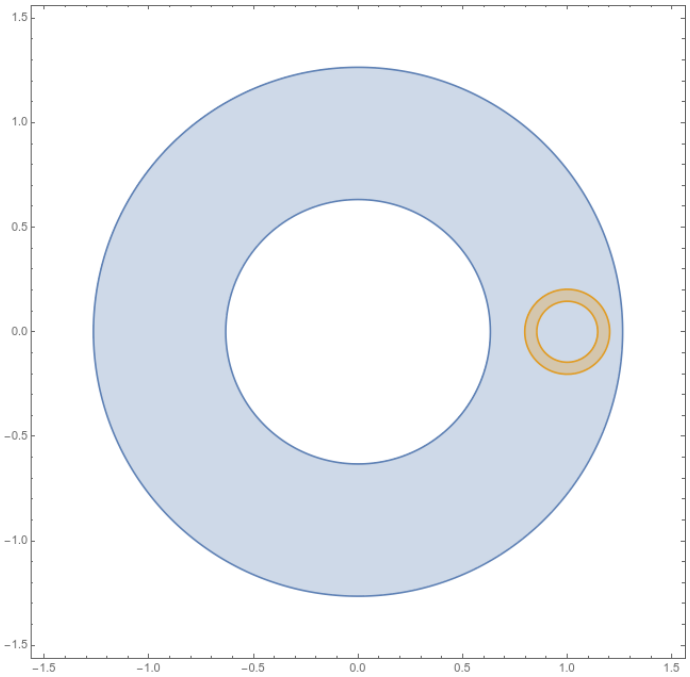


Fig. 3.1 The larger annulus is a superset of the smaller annulus and is homotopy equivalent to it; however, the large annulus does not admit a deformation retraction to the smaller annulus.

In particular, we examine under what constraints the unit n-sphere is detectable, before moving on to the general case of arbitrary n-manifolds. It turns out that, as n increases, we need the inverse variance $\tau := \frac{1}{2\sigma^2}$ to increase linearly with n; moreover, the gradient and offset of this linear function are determined exactly using asymptotic analysis. The bounds for arbitrary n-manifolds are only a constant factor poorer than those for the n-dimensional sphere.

3.1.1 When is the sphere detectable?

Suppose M is the unit sphere, S^n, embedded in $\mathbb{R}^D$ as the intersection of S^{D-1} with an $(n+1)$-dimensional linear subspace. Let μ be the uniform measure on M. The probability density function of the resulting distribution is given by the following expression:

$$f(x_0) = \int_{S^n} e^{-\frac{|x-x_0|^2}{2\sigma^2}} \mu(x) dx$$

where we have scaled μ to have an integral of 1.

In this section, we show that the super-level sets of f are homotopy equivalent to either S^n or a point, and determine necessary and sufficient conditions in terms of σ and n for this to be the case. Finally, we examine how the critical value of σ changes asymptotically as we let the dimension n approach infinity.

Before we do that, however, it is necessary to establish various identities and a lemma about the confluent hypergeometric limit function ${}_0F_1$, which shows up when we compute the integral necessary to evaluate the density function f.

Definition 3.1.2. *The* confluent hypergeometric limit function *is a meromorphic function ${}_0F_1 : \mathbb{C}^2 \to \mathbb{C}$ of two complex variables defined by the series:*

$${}_0F_1(;b;z) := \sum_{k=0}^{\infty} \frac{z^k}{(b)_k k!}$$

where $(b)_k = b(b+1)(b+2)\dots(b+k-1)$ is the Pochhammer symbol.

It has a *regularised* version ${}_0\tilde{F}_1$, where it is scaled by the reciprocal of $\Gamma(b)$, which is in turn related to the better-known *modified Bessel function of the first kind*. In particular:

$${}_0F_1(;\nu+1;\frac{z^2}{4}) = \Gamma(\nu+1){}_0\tilde{F}_1(;\nu+1;\frac{z^2}{4}) = \Gamma(\nu+1)\left(\frac{2}{z}\right)^{\nu} I_\nu(z)$$

We now establish that ${}_0F_1$ is logarithmically concave in the argument z. This is important in constraining the shape of the super-level sets of the probability density function f.

Lemma 3.1.3. *${}_0F_1(;\nu+1;z)$ is logarithmically concave as a function of z, where we restrict to $\nu \geq -\frac{1}{2}$ and $z > 0$.*

Proof. It suffices to evaluate the second derivative $\frac{\partial^2}{\partial z^2}\log({}_0F_1(;\nu+1;z))$ and show that it is nowhere positive.

It is straightforward to show from the series definition that the first derivative is given by the quotient:

$$\frac{{}_0\tilde{F}_1(;\nu+2;z)}{{}_0\tilde{F}_1(;\nu+1;z)}$$

and the second derivative is:

$$\frac{{}_0\tilde{F}_1(;\nu+1;z){}_0\tilde{F}_1(;\nu+3;z) - {}_0\tilde{F}_1(;\nu+2;z)^2}{{}_0\tilde{F}_1(;\nu+1;z)^2}$$

As the denominator is non-negative, it suffices to show that the numerator is non-positive. Equivalently, we want to show that:

$${}_0\tilde{F}_1(;\nu+1;z){}_0\tilde{F}_1(;\nu+3;z) \leq {}_0\tilde{F}_1(;\nu+2;z)^2$$

Since z is positive, we can express this in terms of modified Bessel functions:

$$I_\nu(2\sqrt{z})I_{\nu+2}(2\sqrt{z}) \leq I_{\nu+1}(2\sqrt{z})^2$$

This is known, and in particular is proved for all $\nu \geq -2$ in Theorem 7 of [48]. It follows that ${}_0F_1$ is logarithmically concave in the area of interest. □

We are now ready to establish necessary and sufficient conditions for the sphere to be detectable:

Theorem 3.1.4. *The sphere is detectable if and only if $f(0)$ is strictly less than $f(1)$, where 0 is the origin and 1 is an arbitrary point in M.*

Proof. Let Π be the $(n+1)$-dimensional subspace containing M, and π be orthogonal projection onto Π. The homotopy given by:

$$H_t(x) = t\pi(x) + (1-t)x$$

restricts to any super-level set of f, so every super-level set admits a deformation retract onto its intersection with Π. Hence, we can assume without loss of generality that $D = n + 1$. In that case, we have:

$$\frac{f(x_0)}{f(0)} = \frac{\int_{-1}^{1}(1-x^2)^{n/2-1}\exp(-\tau(1+r^2-2xr))\,\mathrm{d}x}{\int_{-1}^{1}(1-x^2)^{n/2-1}\exp(-\tau)\,\mathrm{d}x} = e^{-\tau r^2}{}_0F_1\left(;\frac{1+n}{2};\tau^2 r^2\right)$$

where $r = |x_0|$, $\tau = \frac{1}{2\sigma^2}$, and ${}_0F_1$ is the confluent hypergeometric limit function.

Viewing this as a function in r^2, we observe that it is a product of the logarithmically concave functions $x \mapsto e^{-\tau x}$ and $x \mapsto {}_0F_1(;\frac{1+n}{2};\tau^2 x)$, and therefore itself logarithmically concave. Consequently, every super-level set of f must intersect every ray from the origin in an interval, and therefore be either a ball or annulus.

We are particularly interested in the smallest super-level set which contains the unit sphere M. This set is an annulus (and therefore deformation retracts onto M) if and only if it does not contain the origin; otherwise, it is a ball and therefore has the incorrect homotopy type. The result follows. □

The proof actually gives us more information, namely an inequality expressing exactly when the sphere is detectable:

$$e^{-\tau}{}_0F_1\left(;\frac{1+n}{2};\tau^2\right) > 1$$

In the next section we will explore how the critical value of τ changes as $n \to \infty$.

Asymptotics for large n

Theorem 3.1.5. *Let n be the dimension, and consider the critical value of σ such that the centre and boundary of the sphere are equiprobable. Then the following asymptotic holds for $n \to \infty$:*

$$\frac{1}{2\sigma^2} = \frac{n-1}{4}z_0 + \frac{\log(1+z_0^2)(1+\sqrt{1+z_0^2})}{8z_0 - 4(1+\sqrt{1+z_0^2})} + o(1)$$

where z_0 is the positive real root of the equation:

$$\sqrt{1+z^2} + \log(2) - 1 - \log(1+\sqrt{1+z^2}) = \frac{1}{2}z$$

Proof. The following asymptotic expansion for the modified Bessel function of the first kind is given in [49]:

$$I_\nu(\nu z) \sim \frac{e^{\nu\eta}}{\sqrt{2\pi\nu\sqrt{1+z^2}}} \sum_{k=0}^{\infty} \frac{U_k(p)}{\nu^k}$$

where $\eta := \sqrt{1+z^2} + \log(z) - \log(1+\sqrt{1+z^2})$ and $p := \frac{1}{\sqrt{1+z^2}}$, and each U_k is a degree-$3k$ polynomial defined by a recurrence relation. Crucially, $U_0(p) = 1$ so the summation tends to 1 as $\nu \to \infty$.

Using Stirling's formula to asymptotically approximate the gamma function, we can obtain an analogous asymptotic expansion for ${}_0F_1$:

$${}_0F_1\left(;\nu+1;\frac{(\nu z)^2}{4}\right) \sim (1+z^2)^{-\frac{1}{4}} \left(\frac{2}{z}e^{\eta-1}\right)^\nu \sum_{k=0}^{\infty} \frac{U_k(p)}{\nu^k}$$

Taking logarithms, evaluating η and simplifying, we get:

$$\log\left({}_0F_1\left(;\nu+1;\frac{(\nu z)^2}{4}\right)\right) = -\frac{1}{4}\log(1+z^2) + \nu\left(\sqrt{1+z^2} + \log(2) - 1 - \log(1+\sqrt{1+z^2})\right) + o(1)$$

Our inequality for the detectability of the sphere then becomes:

$$\sqrt{1+z^2} + \log(2) - 1 - \log(1+\sqrt{1+z^2}) - \frac{z}{2} \geq \frac{1}{\nu}\left(\frac{1}{4}\log(1+z^2) + o(1)\right)$$

where we have set $\nu = \frac{n-1}{2}$ and $z = \frac{2\tau}{\nu}$. The right-hand-side of this equation tends to zero, so we can solve for z for large values of ν and obtain the solution $z_0 \cong 3.30479977$ to the equation:

$$\sqrt{1+z^2} + \log(2) - 1 - \log(1+\sqrt{1+z^2}) = \frac{1}{2}z$$

However, we can do slightly better than just the constant of proportionality. In particular, set $z = z_0 + \frac{k}{\nu}$ and expand the original equation to order ν^{-1}:

$$k\left(\frac{z_0}{1+\sqrt{1+z_0^2}} - \frac{1}{2}\right) = \frac{1}{4}\log(1+z_0^2) + o(1)$$

Solving for k establishes a stronger condition for detectability, where the difference between the actual and asymptotic values for τ tends to zero:

$$\tau = \frac{n-1}{4} z_0 + \frac{\log(1+z_0^2)(1+\sqrt{1+z_0^2})}{8z_0 - 4(1+\sqrt{1+z_0^2})} + o(1)$$

This is precisely the statement of the theorem. □

As this involved a non-trivial amount of algebraic manipulation, partially manual and partially using computer algebra software, it is worth verifying that the approximation holds for large n:

Dimension n	**Actual** τ	**Asymptotic** τ	**Discrepancy**
1	1.357170	1.279153	7.802×10^{-2}
10	8.725988	8.714952	1.104×10^{-2}
100	83.074071	83.072947	1.124×10^{-3}
1000	826.653009	826.652896	1.126×10^{-4}
10000	8262.452395	8262.452384	1.126×10^{-5}

This asymptotic approximation is surprisingly accurate even in the case $n = 1$. The discrepancy between the actual result and the approximation appears experimentally to be improvable from $o(1)$ to $O(n^{-1})$.

3.1.2 Smoothly embedded compact manifolds

Suppose $M \subset \mathbb{R}^D$ is now a smoothly embedded compact n-manifold. This is strictly more general than the previous case of a unit n-sphere. Underpinning this is the *tubular neighbourhood theorem*. The version we quote here, for smoothly embedded compact manifolds, is proved in [50].

Definition 3.1.6. *Let $M \subseteq \mathbb{R}^D$ be a C^2-embedded compact n-manifold. The* normal exponential map *is the function $\exp^\perp : NM \to \mathbb{R}^D$ from the normal bundle to the ambient space given by:*

$$\exp^\perp(x, v) := x + v$$

Theorem 3.1.7 (Tubular neighbourhood theorem)**.** *Let $M \subseteq \mathbb{R}^D$ be a C^2-embedded compact n-manifold. Then there exists $h > 0$ such that $\exp^\perp$, when restricted to the open tube $\{(x, v) : |v| < h\}$, is a diffeomorphism onto its image. h shall be deemed the* normal injectivity radius *of M.*

Observe that the principal curvatures must be bounded above by h^{-1} in absolute value. The tubular neighbourhood theorem is stronger, however, since it imposes global constraints on the manifold: h also measures how far the manifold is from self-intersecting.

Analogous to how there is the notion of 'differentiability at point', one can similarly define what it means to have a normal injectivity radius of h at a point:

Definition 3.1.8. *Let $M \in \mathbb{R}^D$ be a C^2-embedded compact n-manifold. Let $P \in M$ be a point, T_P be its n-dimensional tangent space, and U_P be its $(D-n)$-dimensional orthogonal complement. Then we say that M* has normal injectivity radius h at P *if, for every point $x \in U_P$ with $d(x, P) = h$, the open radius-h ball $B_h(x)$ centred on x is disjoint from the manifold M.*

Lemma 3.1.9. *If M is a C^2-embedded compact manifold with a normal injectivity radius of h, then it has a normal injectivity radius of h at every point $P \in M$.*

For every $\varepsilon > 0$, the $(h-\varepsilon)$-thickening of M admits a deformation retract onto M. We are more interested in the case where we have a finite set P which approximates M, and want to show that some R-thickening of P also necessarily admits a deformation retract onto M.

Lemma 3.1.10. *Suppose P is a finite set such that $M \subseteq P_{r_1}$ and $P \subseteq M_{r_2}$. Then P_R admits a deformation retract onto M provided the following inequality holds:*

$$h - r_2 \geq \sqrt{R^2 + (h - R + r_1)^2}$$

Proof. Firstly, note that this instantly implies that $h - R \geq r_2$, and therefore that the right-hand-side strictly exceeds R; consequently, $h - R > r_2$, which means that P_R is contained in the $(h-\varepsilon)$-thickening of M. By the tubular neighbourhood theorem, this means that each point $x \in P_R$ has a closest point $\pi(x) \in M$, and that π is continuous.

We want to say that $H_t(x) := t\pi(x) + (1-t)x$ is a deformation retract onto M. It is necessary, however, to show that the fibres of π are convex (or, at the very least, are star-neighbourhoods around the 0-section). To accomplish this, we will use the fact that the $(R - r_1)$-thickening of M is contained within P_R.

Recall that P_R is a union of radius-R balls centred on the points $x \in P$. It suffices to show, therefore, that H_t restricted to each $B_R(x) \cup M_{R-r_1}$ is a deformation retract onto M. The fibres of π are the unions of two coplanar metric $(D-n)$-balls

(one from intersecting $B_R(x)$ with the normal space; the other from intersecting M_{R-r_1} with the normal space). These are star-neighbourhoods around the 0-section provided that the boundaries of these balls do not intersect at an acute angle. This is demonstrated more clearly in Figure 3.2.

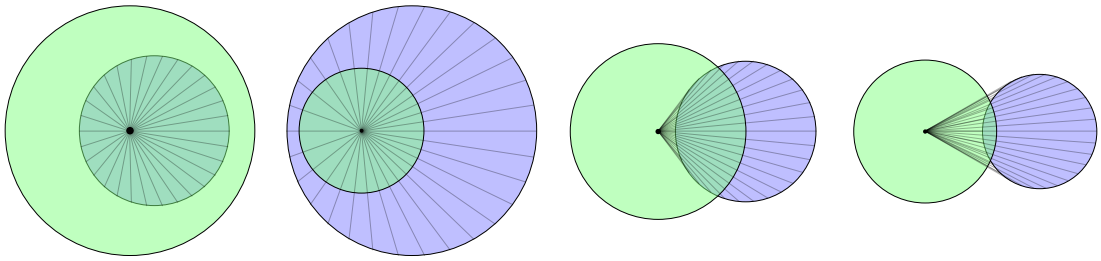


Fig. 3.2 In each of the first two diagrams, one disc completely contains the other so the diagram can be linearly retracted to the centre of the green disc. In the third diagram, they intersect transversely with an obtuse angle of incidence; as such, they can still be linearly retracted to the origin. In the fourth diagram, the angle of incidence is acute, and the grey lines escape the union of the two discs.

Equivalently, in terms of the total space rather than each fibre in isolation, it suffices to show that the boundaries of M_{R-r_1} and $B_R(x)$ do not intersect at an acute angle.

Suppose otherwise. Then there is some point y where the boundary of M_{R-r_1} intersects the boundary of $B_R(x)$ at an acute angle. The curvature constraint on M implies that $\pi(x)$ and $\pi(y)$ lie on some $(D-1)$-sphere of centre O and radius $h' \geq h$ tangent to M at $\pi(x)$. The acute angle is extremised when y is coplanar with O and x, thus we henceforth assume this without loss of generality.

We have that $\pi(x)$ and $\pi(y)$ are at a distance of h' from O, whence it follows that $|O-x| \geq h' - r_2$ and $|O-y| = h' - R + r_1$. But then the acuteness of the intersection of the surfaces gives the inequality:

$$|O-x|^2 < |O-y|^2 + |y-x|^2$$

which immediately contradicts the inequality we assumed in our premise. □

Analytic optimisation shows us that, provided $r_2 < \frac{1}{2}(2-\sqrt{2})h$, we can take $r_1 = \sqrt{2}(h-r_2) - h$ and $R = \frac{1}{\sqrt{2}}(h-r_2)$. This gives the following corollary which is applicable to topological data analysis:

Corollary 3.1.11. *Suppose we have a compact smoothly embedded n-manifold $M \subset \mathbb{R}^D$ which has a normal injectivity radius of h. Further, let P be a finite subset of the r_2-thickening of M. Then, provided that we have $r_2 < \frac{1}{2}(2-\sqrt{2})h$, and the r_1-thickening of P covers M (where $r_1 = \sqrt{2}(h-r_2) - h$), then the R-thickening of P and the associated Čech complex are both homotopy equivalent to M where $R = \frac{1}{\sqrt{2}}(h - r_2)$.*

In practice, r_2 is the radius which depends on the Gaussian noise parameter σ, and we can make r_1 as small as we like by sampling sufficiently many points and choosing the subset of high enough density to ensure all points lie within M_{r_2} with some sufficiently high probability p. The rest of this section is concerned with finding explicit bounds for how many points we need to sample as a function of M and p.

Explicit bounds for smoothly embedded manifolds

We have shown that, provided $r_2 < \frac{1}{2}(2-\sqrt{2})h$ and $f(x) > f(y)$ for every $x \in M$ and $y \notin M_{r_2}$, then after sampling arbitrarily many points, removing outliers appropriately, and constructing a Čech complex, it will be homotopy equivalent to M.

It is natural to ask whether there are any natural sufficient conditions which ensure that $f(x) > f(y)$ for all $x \in M$ and $y \notin M_{r_2}$. We do this by bounding both sides of the inequality:

- We know that $f(y) = \mathbb{E}_x e^{-\tau d(x,y)^2} \leq e^{-\tau r_2^2}$, where we have omitted the normalisation factor for convenience. Unless most of the measure μ on the manifold M is concentrated near the closest point to y, this bound is rather weak. In particular, it can be improved considerably when μ is the uniform distribution over M.

- We shall show that, provided $n \geq 2$, $f(x) \geq K^{-1}\alpha e^{-2\tau h^2}{}_0F_1(\frac{1+n}{2}, \tau^2 h^4)$, where α is the ratio of the minimum density of μ on M divided by the density of the uniform measure of the n-sphere of radius h. $K > 0$ is a universal constant which may be taken to be 12.75.

To prove the latter, it suffices to show that the n-sphere is the 'worst case scenario' up to a constant factor (in a way we make precise below) amongst all manifolds of uniform density and normal injectivity radius h. That is to say, the n-sphere minimises the volume of intersection between M and a radius-r ball centred on a point of the manifold, up to an absolute dimension-independent constant factor.

Lemma 3.1.12. *Let $n \geq 1$ and consider a C^1 embedded n-manifold $M \subset \mathbb{R}^D$ (where $D > n$), which passes through the origin with a normal injectivity radius of h at the origin. Let $r > 0$ be a positive real. Then there exists a manifold M'' which satisfies the same conditions as M, occupies an $(n+1)$-dimensional linear subspace of $\mathbb{R}^D$, is axisymmetric, and has a smaller area when restricted to the ball of radius r centred on the origin:*

$$|B_r(0) \cap M''| \leq |B_r(0) \cap M|$$

where $B_r(0)$ is the radius-r ball centred on the origin, and $|\ |$ denotes the n-dimensional Lebesgue measure.

Proof. Let the tangent space of M at the origin be denoted T, and let U be its $(D-n)$-dimensional orthogonal complement. We endow the space $S^{n-1} \subset T$ of unit tangent vectors with a measure given by:

$$\mu(X) := |B_r(0) \cap M \cap \{u + \rho t : u \in U, t \in X, \rho \in \mathbb{R}_{\geq 0}\}|$$

Then $|B_r(0) \cap M|$ is an integral over the space S^{n-1}, and we can let $\hat{t}$ be a minimiser of the integrand. It follows that we can replace M with the axisymmetric M' without increasing its area:

$$M' := \{u + \rho t : t \in S^{n-1}, \rho \in \mathbb{R}_{\geq 0}, u \in U, u + \rho\hat{t} \in M\}$$

Note that M' need not be smooth (as singularities may appear on U), but still has a normal injectivity radius of h at the origin. If M' has multiple connected components, discard the components other than the one that contains the origin. Then, after parametrising the geodesics from the origin in terms of arc-length, s, we get:

$$M' = \{u(s) + \rho(s)t : t \in S^{n-1}, s \in [0, \ell]\}$$

where ℓ is the arc-length of the geodesic measured from the origin to either the boundary of the radius-r ball or to the subspace U (whichever comes sooner), and $u : [0, \ell] \to U$ and $\rho : [0, \ell] \to \mathbb{R}^{\geq 0}$ are continuously differentiable functions.

The other simplification that we can apply without increasing the area is to make $u(s)$ monotonic. Specifically, let $v'(s) = |u'(s)|$ be the Euclidean norm of the

derivative of $u(s)$, and let $v : [0, \ell] \to \mathbb{R}_{\geq 0}$ be the antiderivative of $v'(s)$. Let $\hat{u} \in U$ be an arbitrary unit vector, and define:

$$M'' = \{v(s)\hat{u} + \rho(s)t : t \in S^{n-1}, s \in [0, \ell]\}$$

We have $|B_r(0) \cap M''| \leq |B_r(0) \cap M'| \leq |B_r(0) \cap M|$. Also, $v(s)$ is a monotone-increasing function, and $M'' \subseteq V \oplus \langle \hat{u} \rangle$ now lies in a $(n+1)$-dimensional subspace of $\mathbb{R}^D$. Consequently, we can henceforth assume that $D = n + 1$ without loss of generality.

The condition of having normal injectivity radius of h implies that M'' does not intersect the open ball of radius h tangent to M'' at the origin. Recall that the manifold is either closed (and may be non-C^1 at the second intersection with the axis of symmetry U) or has a boundary on the surface of $B_r(0)$. In the former case, it follows that M'' encloses the ball of radius h, so has area at least as equal to the radius-h sphere, and we are done. Otherwise, M'' is already everywhere C^1, and we are also done. □

Theorem 3.1.13. *There exists a universal constant K such that for all $D > n \geq 1$ and $r, h > 0$, for every compact C^1-embedded n-manifold $M \subset \mathbb{R}^D$ which passes through the origin with a normal injectivity radius of h at the origin, we have:*

$$|B_r(0) \cap S^n(h)| \leq K|B_r(0) \cap M|$$

where $S^n(h)$ denotes an n-sphere of radius h tangent to M at the origin.

Proof. By the previous lemma, we can assume without loss of generality that $D = n+1$ and M is axisymmetric, parametrised as:

$$M = \{v(s)\hat{u} + \rho(s)t : t \in S^{n-1}, s \in [0, \ell]\}$$

where $v(s)$ is a monotone-increasing function of arc-length s along the geodesic. The following conditions hold:

- $v(s), \rho(s) \geq 0$ with equality if and only if $s = 0$;
- $v(s)^2 + \rho(s)^2 \leq r^2$ with equality if and only if $s = \ell$;
- $v'(s) \geq 0$ (monotonicity);
- $v'(s)^2 + \rho'(s)^2 = 1$ (unit speed);

- $(v(s) - h)^2 + \rho(s)^2 \geq h^2$ (normal injectivity radius).

The area $|B_r(0) \cap M|$ can now be expressed as an integral, where the scaling factor in front of the integral is the area of the unit sphere S^{n-1}:

$$|B_r(0) \cap M| = \frac{2\pi^{n/2}}{\Gamma(n/2)} \int_0^\ell \rho(s)^{n-1}\, ds$$

To complete the proof, we use a tripartite case-bash, considering the cases of $r \leq h$, $h \leq r \leq \sqrt{2}h$, and $\sqrt{2}h \leq r$, which together cover the positive real axis. The corresponding bounds on K obtained from these cases are 2, $1 + \pi/(2 - \sqrt{3})$, and $2 + 2\pi/(2 - \sqrt{2})$, respectively.

Let $S = S^n(h)$ be the radius-h sphere tangent to M at the origin, such that M and S lie on the same side of the common tangent space; this is well defined from the monotonicity property. The condition of having a normal injectivity radius of h at the origin means that M is disjoint from the open ball whose boundary is S. For ease of exposition, we shall henceforth use *vertical* to mean 'parallel to the common tangent space T of M, S' and *horizontal* to mean 'perpendicular to the common tangent space T'. The *horizontal coordinate* of a point P in the ambient space is the signed perpendicular distance from P to T, where we adopt the convention that all points on M and S have nonnegative horizontal coordinate. We say Q is *left* (resp. *right*) of P if its horizontal coordinate is less than (resp. greater than) that of P.

- $\mathbf{r \leq h}$: The axis-parallel projection from M to S can increase areas by at most a factor of 2. This can be seen by observing that the integrand $\rho(s)^{n-1}$ is unaffected, but the Jacobian of the projection map picks up a factor of at most $\sec(\theta) \leq 2$, where $\theta \leq \frac{\pi}{3}$ is the angle between the normal of S and the horizontal.

- $\mathbf{h \leq r \leq \sqrt{2}h}$: We split $B_r(0) \cap S$ into two portions: the spherical cap to the left of the coordinate $v(\ell)$, and the annular region to the right. The area of M is an upper bound for that of the spherical cap (because normal projection onto S cannot increase area). Moreover, let ρ_0 be the maximum distance of any point on $B_r(0) \cap S$ from the horizontal axis through the origin, and let $a = \rho(\ell) - \rho_0$ be the amount by which M exceeds that. On the interval $s \in [\ell - a, \ell]$, we have $\ell(s) \geq \rho_0$. This defines an annular region of M which is at least $(2 - \sqrt{3})/\pi$ times the area of the aforementioned annular region of $B_r(0) \cap S$. This gives an overall value of $K = 1 + \pi/(2 - \sqrt{3})$ for this case.

- $\sqrt{\mathbf{2}}\mathbf{h} \leq \mathbf{r}$: This case proceeds identically to the previous case, except we only consider the left-hand hemisphere of S instead of the whole of $B_r(0) \cap S$. By the same logic, we can show that this hemisphere has area at most $1 + \pi/(2 - \sqrt{2})$ times the area of $B_r(0) \cap M$; doubling this to cover the whole sphere yields the bound of $K = 2 + 2\pi/(2 - \sqrt{2})$.

The result follows. For concreteness, note that $K = 12.75$ is sufficient for all three cases, and only slightly larger than both $2 + 2\pi/(2 - \sqrt{2})$ and $1 + \pi/(2 - \sqrt{3})$. □

We obtain the following corollary:

Corollary 3.1.14. *The ratio $f(x)/f(y)$, where $x \in M$ and $y \notin M_{r_2}$, is bounded below by:*

$$K^{-1}\alpha e^{\tau(r_2^2 - 2h^2)}{}_0F_1\left(\frac{1+n}{2}, \tau^2 h^4\right)$$

where α is the ratio of the minimum density of μ on M, divided by the density of the uniform measure of the n-sphere of radius h (in other words, scaled so that $\alpha = 1$ in the case of $S^n(h)$. K is the absolute constant from the previous theorem.

At this point, we are ready to prove the detectability result for arbitrary smoothly embedded compact manifolds (generalising, at the expense of worse bounds, the result proved for the unit sphere).

Theorem 3.1.15. *Let $X := Y + \varepsilon$ be the sum of a random variable Y supported on an n-manifold $M \subseteq \mathbb{R}^D$ and a Gaussian random variable ε with zero mean and covariance matrix $\sigma^2 I$ (where I is the $D \times D$ identity matrix).*

Then $f(x) \geq \beta f(y)$ for all $x \in M$ and $y \notin M_{r_2}$ provided the following inequality holds:

$$\frac{h^2}{2\sigma^2} \geq \frac{n-1}{4} z_0 + \frac{\frac{1}{4}\log(1+z_0^2) + \log(K\beta) - \log(\alpha)}{\left(1 + \sqrt{1+z_0^2}\right)^{-1} z_0 - \left(1 - \frac{r_2^2}{2h^2}\right)} + o(1)$$

where z_0 is a universal constant, h is the normal injectivity radius of M, and α is the (appropriately normalised) minimum density of the random variable Y on its support M.

Proof. We proceed as before: in order for the expression in Corollary 3.1.14 to exceed $\beta > 1$, we want the logarithm to be positive. Set $\nu = \frac{n-1}{2}$ and $z = \frac{2\tau}{\nu} h^2$, and apply the asymptotic formula for ${}_0F_1$, recovering the following inequality:

$$\sqrt{1+z^2} + \log(2) - 1 - \log(1+\sqrt{1+z^2}) - z\left(1 - \frac{r_2^2}{2h^2}\right) \geq \frac{1}{\nu}\left(\frac{1}{4}\log(1+z^2) + \log(K\beta) - \log(\alpha) + o(1)\right)$$

The bugbear is that we require $r_2 < \frac{1}{2}(2-\sqrt{2})h$, so the solution z_0 is correspondingly larger than in the case of the sphere. Indeed, for this value of r_2 we get a threshold of $z_0 \approxeq 118.591$ which translates to a bound on σ which is six times worse than the case of the sphere.

It is, thankfully, only a constant factor worse, so we still get a concrete result that our sampling methodology will necessarily work assuming the following conditions (which holds asymptotically for large n, as well as in practice for small n):

$$\frac{h^2}{2\sigma^2} \geq \frac{n-1}{4} z_0 + \frac{\frac{1}{4}\log(1+z_0^2) + \log(K\beta) - \log(\alpha)}{\left(1+\sqrt{1+z_0^2}\right)^{-1} z_0 - \left(1 - \frac{r_2^2}{2h^2}\right)} + o(1)$$

□

For a concrete example of α for an n-manifold that is not the unit sphere, take the torus T^n embedded in $\mathbb{R}^{2n}$ as the Cartesian product of n unit circles. The torus has a measure of $(2\pi)^n$ whereas the n-sphere has a measure of $2\frac{\sqrt{\pi}^{n+1}}{\Gamma((n+1)/2)}$, so a uniform distribution over the torus would have an associated α of:

$$\alpha = \left(\Gamma((n+1)/2)\sqrt{4\pi}^{n-1}\right)^{-1}$$

In this case, $-\log(\alpha)$ grows superlinearly as a function of n owing to the presence of the gamma function. Specifically, the second term in the expression for the lower bound on $\frac{1}{2\sigma^2}$ grows at a rate of $\Theta(n \log n)$ and therefore eventually dominates the linear term.

3.2 Witness complexes and modifications thereof

The previous analysis uses the Čech complex. For large datasets, it is infeasible to explicitly compute a full Čech complex. As mentioned in the previous chapter, when $2r$ is greater than the diameter of the set X, the Čech complex includes a simplex

for every non-empty subset of X; even for a relatively small dataset with $|X| = 30$, there are already more than one billion simplices.

As such, it is preferable to use a more computationally tractable alternative, such as the witness complexes described in [31]. Unfortunately, whilst they are convenient for computation, they are often not homotopy equivalent to the r-thickening of X (or equivalently to the Čech complex $\mathcal{C}_r$) because omission of simplices from the original Čech complex often creates 'holes' which introduce additional features in the homology. To address this, we propose a pair of methods to 'fill in' these holes.

One sources of holes is as follows: if four vertices of L form a planar cyclic quadrilateral, the set of points that could witness either of the diagonals is of measure zero; consequently, quadruples of landmark points which are close to being concyclic tend to yield quadrilateral 'holes' in the witness complex where neither diagonal is witnessed. These holes, visible in Figure 3.3, are problematic because they yield spurious infinite bars in the H_1 homology.

3.2.1 Squared witness complex

We depart from [31] by considering the squared graph G^2 of the 1-skeleton G of the witness complex. That is to say, two landmark vertices $x, y \in L$ are adjacent in G^2 if there is either a point witnessing the edge xy or a 'common neighbour' $z \in L$ such that each of the edges xz and yz are witnessed. This addresses the problem of the near-cyclic quadrilaterals at the minor expense of increasing the number of edges.

Proceeding as before, one can take the clique complex $\mathcal{K}$ of G^2, including a simplex for every clique (complete subgraph) in G^2. As in a Vietoris-Rips complex, a filtration $\{\mathcal{K}_r : r \geq 0\}$ can be naturally defined on $\mathcal{K}$ by including only the simplices whose underlying vertex-sets have diameter not exceeding r.

The motivation for squared witness complexes is similar to the motivation for *bounded Kan filling* introduced in the next section. The salient difference is that squared witness complexes remove infinite bars in the homology barcode, whereas bounded Kan filling removes finite bars. These two approaches are complementary, as is seen in the analysis of the final dataset (natural images).

3.2.2 Bounded Kan filling

As the points in X are typically sampled randomly from the manifold we are trying to detect, the number of points in a region of the manifold follows a Poisson distribution.

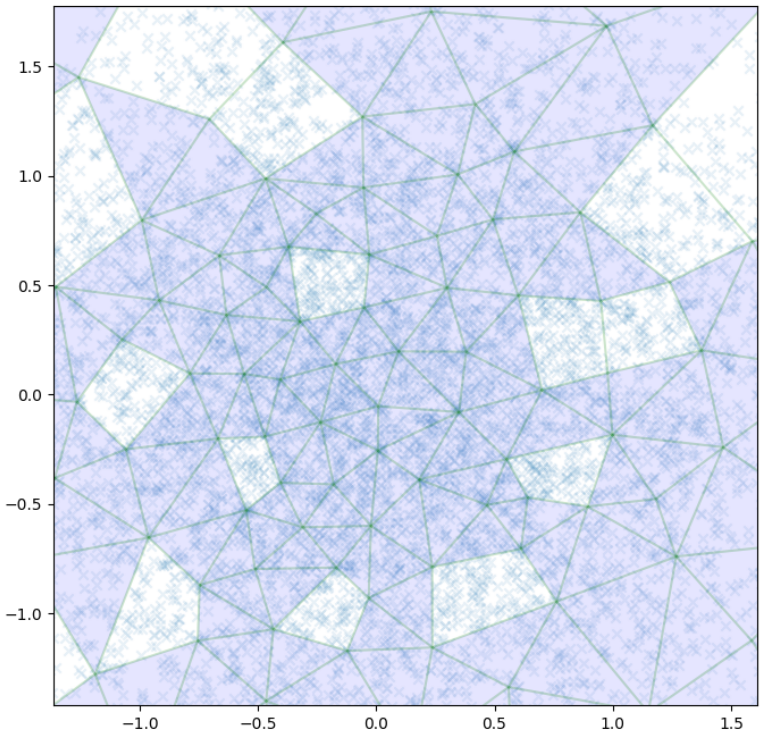


Fig. 3.3 Near-cyclic quadrilaterals fail to be witnessed, causing spurious features in the first homology.

This means that often there are relatively large regions of the manifold devoid of points, causing holes to appear as in the left-hand complex in Figure 3.4.

Not only are they damaging for topological data analysis, but also they are slightly detrimental to Isomap: the estimate of geodesic distances across the manifold is distorted since the shortest path must traverse around the hole instead of passing through it. This theoretical argument is confirmed by the experimental results later in this chapter.

Increasing the length scale further would eventually lead to additional edges bridging the inner and outer layers of the Swiss roll, which would compromise Isomap much more severely by making distant points on the manifold appear adjacent. In topological data analysis, this is remedied by observing features which persist over a range of length scales; there is no obvious way to apply this idea to Isomap.

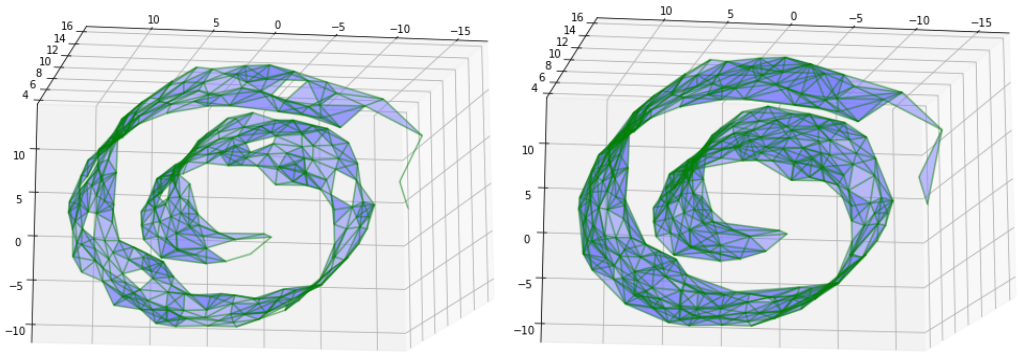


Fig. 3.4 Witness complexes of the Swiss roll dataset, constructed with and without bounded Kan filling.

Instead, using another technique similar to the squared witness complex, *bounded Kan filling* connects points which are close in both the intrinsic geometry of the graph and in the ambient geometry, resulting in the hole-free right-hand image. It operates on the filtration as follows:

- If a simplex σ appears at time t – that is to say, t is the infimum $t := \inf\{s : \sigma \in \mathcal{K}_s\}$ – and its codimension-1 faces appear at times $t_0, \ldots, t_{n-1}, t_n$ (in chronological order), then we 'bring forward' both the simplex and its most recent codimension-1 face to appear at time t_{n-1}. This is performed iteratively.
- We impose that, for each simplex σ, the new time t' is constrained to be at least ct, where t is the original time and $0 < c < 1$ is a universal constant. In practice, $c = \frac{1}{2}(\sqrt{5} - 1)$ works very well; in the following section, we demonstrate that it is necessary and sufficient to ensure, for Vietoris-Rips, Čech, alpha, and squared witness complexes in $\mathbb{R}^D$, that all cycles of length 4 and 5 in the original graph are triangulated.

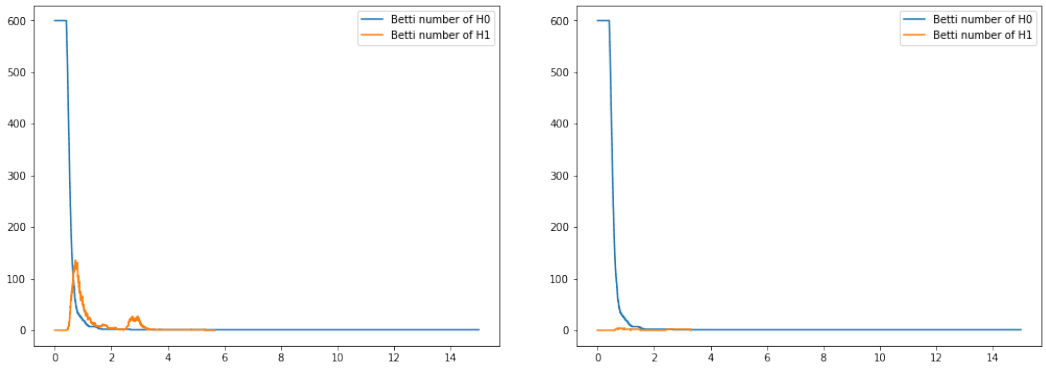
By judiciously choosing the order of simplices to be updated, and using appropriate data structures, this can be implemented as an algorithm with running time $O(nd^2 + n \log n)$ (in the random-access model), where n is the number of vertices and d is the maximum dimension of any simplex. Details of this algorithm are described later. This is dominated by the time complexity of the linear algebra involved in Isomap, so bounded Kan filling has negligible overhead for large datasets.

The adjective 'bounded' refers to the presence of the constant c. If $\mathcal{K}$ is the original filtration and $\mathcal{K}'$ is the filtration after bounded Kan filling, one obtains a

Lipschitz result entirely analogous to the relationship between Vietoris-Rips and Čech complexes:

$$\mathcal{K}_{ch}(X) \subseteq \mathcal{K}'_h(X) \subseteq \mathcal{K}_h(X)$$

Consequently, bounded Kan filling can only perturb the (logarithmic) persistence diagram by a bottleneck distance of at most $\log(c^{-1})$; this implies that persistent features in the homology barcode will remain after bounded Kan filling. Conversely, experimental results confirm that, in practice, it causes the majority of short 'noise' bars to disappear, thereby 'cleaning' the barcode such that the topology can be inferred from the homology of a single simplicial complex, rather than by needing to compute persistent homology. This is reflected in the plots of the Betti numbers of the Swiss roll before and after bounded Kan filling:



At the end of this chapter, we apply this idea to manifold learning, where we need a single neighbourhood graph. Another feasible application is multidimensional persistence, where the simplicial complexes $\mathcal{K}_{r,p}$ are described by a 2-parameter filtration (for instance, witness complexes with length-scale r, restricted to a proportion p of the 'highest-density' points as in [41]); one could apply bounded Kan filling with respect to r followed by persistent homology with respect to p. This is explored in the final chapter of the dissertation.

Choice of constant c

We mentioned that $c = \frac{1}{2}(\sqrt{5} - 1)$ works well in practice. Here we shall prove that, if there are any 5-cycles in a Čech or Vietoris-Rips complex, then they are null-homotopic in the corresponding Kan-filled complex. It is easy to see that the claim is equivalent to the following:

Theorem 3.2.1. *Let P be a pentagon in $\mathbb{R}^n$, and suppose that every edge has length at most 1. Then we can triangulate P by adding two diagonals of length at most $\phi := \frac{1}{2}(\sqrt{5}+1)$*

Proof. It is shown in [28] that the sum of external angles of any closed (not necessarily planar) polygon is at least 2π, by induction on the number of vertices. This implies that one of those angles is at least $\frac{2\pi}{5}$, and the corresponding internal angle $\widehat{V_{i-1}V_iV_{i+1}}$ is at most $\frac{3\pi}{5}$. By the cosine rule, it follows that the diagonal $V_{i-1}V_{i+1}$ has length bounded above by ϕ.

What remains to show is that one of the diagonals of the quadrilateral $V_{i-2}V_{i-1}V_{i+1}V_{i+2}$ has length at most ϕ. This is a consequence of Ptolemy's inequality, which states that:

$$|V_{i-2}V_{i+1}||V_{i-1}V_{i+2}| \leq |V_{i-1}V_{i+1}||V_{i-2}V_{i+2}| + |V_{i-2}V_{i-1}||V_{i+1}V_{i+2}|$$

with equality if and only if the quadrilateral is a planar cyclic quadrilateral. The right-hand-side of the inequality is bounded above by $\phi + 1$, which implies that at least one of the diagonals is no greater than $\sqrt{\phi+1} = \phi$. □

3.2.3 Computational complexity

Here we show that it is possible to perform the bounded Kan filling algorithm in time $O(nd^2)$, where n is the number of simplices and d is the maximum dimension of any simplex, assuming the simplices are already sorted in order of appearance. (Otherwise, the preprocessing step will take time $O(n \log n)$, which may dominate if $d^2 = o(\log n)$.)

Preprocessing

The preprocessing step prepares the data structure on which the algorithm will operate. It takes time $O(nd)$ if the simplices are initially sorted, or $O(n \log n)$ otherwise.

1. If necessary, sort the list of simplices into ascending order according to the point t at which each simplex appears in the filtration. This takes time $O(n \log n)$ in the worst case using mergesort.

2. Walk through the list of simplices and produce an ordered list T of unique arrival times. This takes time $O(n)$.

3. Define $T' := T \cup cT$, where $cT := \{ct : t \in T\}$. The ordered set T' can be produced in time $O(n)$ by interleaving T and cT with the procedure used in a single iteration of mergesort.

4. Let the elements of T' be labelled $\{t_1, t_2, t_3, \ldots, t_m\}$, where $m \leq 2n$. We construct an array of initially empty doubly-linked lists, $\{L_1, L_2, \ldots, L_m\}$.

5. Populate the doubly-linked lists with the simplices such that if a simplex appears at time t_i, it is appended to list L_i. This can be accomplished in time $O(n)$ by simultaneously walking through the list of simplices and array of doubly-linked lists; we never backtrack.

6. Annotate each simplex σ with its list of *children* (codimension-1 faces) and *parents* (simplices which contain σ as a codimension-1 face).

7. Also annotate each simplex with two indices: a **mutable** index i such that L_i is the list which contains σ, and an **immutable** index j such that $t_j = ct_i$ *ab initio*. The purpose of j is to provide a lower bound for i. (To make it clear as to the simplex to which we are referring in the description of the algorithm, i and j will henceforth be subscripted with the simplex name σ.)

Iterative procedure

We define two operations on simplices:

- **Update:** we take a simplex σ and let ρ, π be the children with largest and second-largest index i, respectively. Now we update $i_\sigma \leftarrow \max(j_\sigma, i_\pi)$, and update $i_\rho \leftarrow \min(i_\rho, i_\sigma)$. Move these simplices into the appropriate linked lists.

- **Signal:** we take a simplex σ and update each of its parents. Then move σ out of its linked list and place it in the 'finished pile' L_0.

Clearly, an update step takes time $O(d)$, and a signal step takes **amortized** time $O(d^2)$. To conduct the algorithm, we walk along the array of linked lists, repeatedly applying the signal step to the topmost element until the list is fully consumed before moving on to the next linked list in the array. The total time is $O(nd^2)$ as claimed.

The algorithm works because, whenever we place a simplex σ on the finished pile, we are assured that no simplex τ will subsequently ever receive an index $i_\tau < i_\sigma$. This can be proved by strong induction on the index i. After completing the procedure, we obtain a new filtration where the simplex σ appears at time t_{i_σ}.

3.3 Application to manifold learning

The efficacy of bounded Kan filling is demonstrated by improving the reconstruction accuracy of two methods of nonlinear dimensionality reduction: Isomap and LLE. We measure reconstruction accuracy in terms of the Procrustes distance between the ground truth and the low-dimensional embedding.

3.3.1 Tenenbaum's faces

Tenenbaum prepared a dataset of 698 monochrome 64×64 synthetic images of faces. Each image has three independent uniformly-distributed parameters: the vertical pose, horizontal pose, and lighting direction. This dataset is embedded naturally in $\mathbb{R}^{4096}$, and by projecting it onto the first 160 principal components this is reduced to a linear subspace $\mathbb{R}^{160}$.

For each $k \in \mathbb{N}$, we created the *k-nearest neighbour graph* G_k by including an edge between $x, y \in X$ if and only if either x is in the k nearest neighbours of y, or y is in the k nearest neighbours of x, or both. This is the identical to the graph used in Isomap and its conformal counterpart, but differs from the digraph used in LLE in that it is symmetrised.

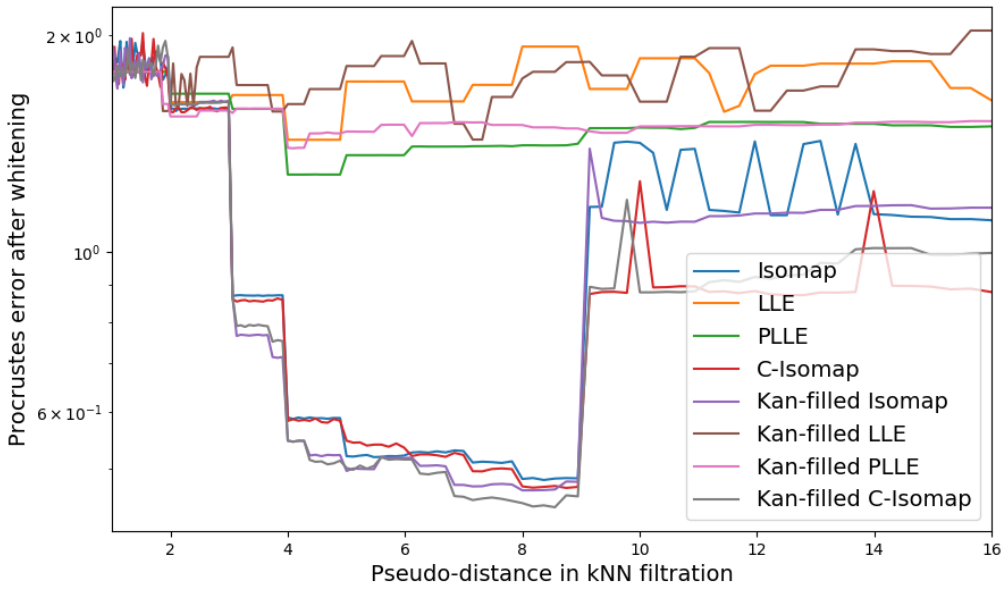
G_k can be viewed as the 1-skeleton of its clique complex $\mathcal{R}_k$. Analogously, we let G'_k be the 1-skeleton of the simplicial complex obtained from applying bounded Kan filling to the filtration of $\mathcal{R}_k$.

We experimented with the Isomap [14], C-Isomap [16], and LLE [15] on both G_k and G'_k for various positive real values of k. We also include a modification of LLE in which the weights are constrained to be positive, as briefly alluded to in [15], which shall be abbreviated as PLLE. This is slower to compute as it requires convex optimisation (typically an iterative method) as opposed to direct linear algebra, but we shall see that the reconstruction error is consistently lower than unconstrained LLE.

Note that the graph G_k only depends on the integer part of k, whereas G'_k depends on the integer parts of both k and $c^{-1}k$, where $c = \frac{1}{2}(\sqrt{5} - 1)$ is the parameter used for bounded Kan filling. We sample 125 geometrically uniformly spaced values of $k \in [1, 16]$.

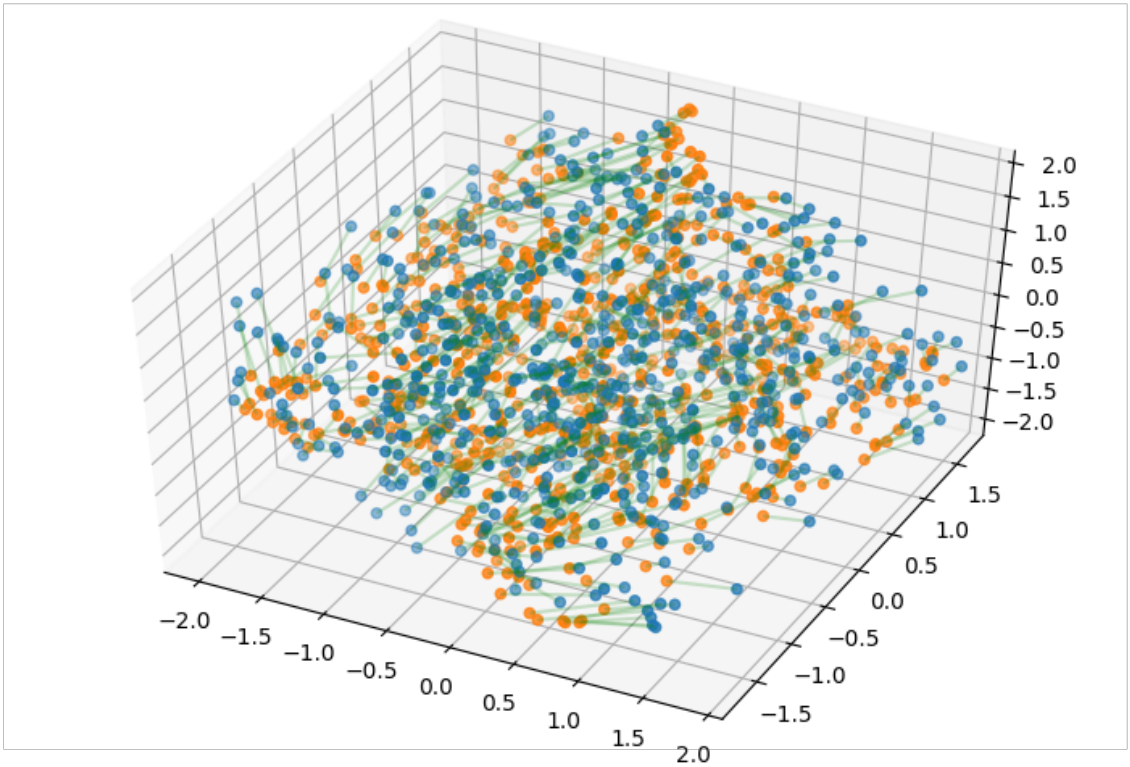
For each combination of graph and method of nonlinear dimensionality reduction, we obtain an embedding from $\mathbb{R}^{160}$ to $\mathbb{R}^3$. This is then whitened (affinely transformed

to have zero mean vector and identity covariance matrix) and compared with the (also whitened) three-dimensional 'ground truth' using Procrustes distance:



This is consistent with the results in [14], which found that LLE gave a very distorted embedding of the dataset of faces. Our results further show that this is only marginally improved by constraining the weights to be positive, and bounded Kan filling offers no further improvement. On the other hand, bounded Kan filling marginally improves the performance of both the original and conformal variants of Isomap in the trough in which reconstruction is most accurate.

The best reconstruction is with conformal Isomap after bounded Kan filling; the original and reconstructed datasets are visualised below and superimposed as to minimise the mean squared error:



3.3.2 Swiss roll

The noiseless Swiss roll described in [20] is generated by sampling a set S of 1600 points (x, y) from a Gaussian mixture model of four equiprobable unit-variance normal distributions centred on the points $\{\frac{15}{2}, \frac{25}{2}\} \times \{\frac{15}{2}, \frac{25}{2}\}$. This is embedded in $\mathbb{R}^3$ using a parametrisation of an Archimedean spiral:

- **3D embedding:** $Z := \{(x \cos x, y, x \sin x) : (x, y) \in S\}$

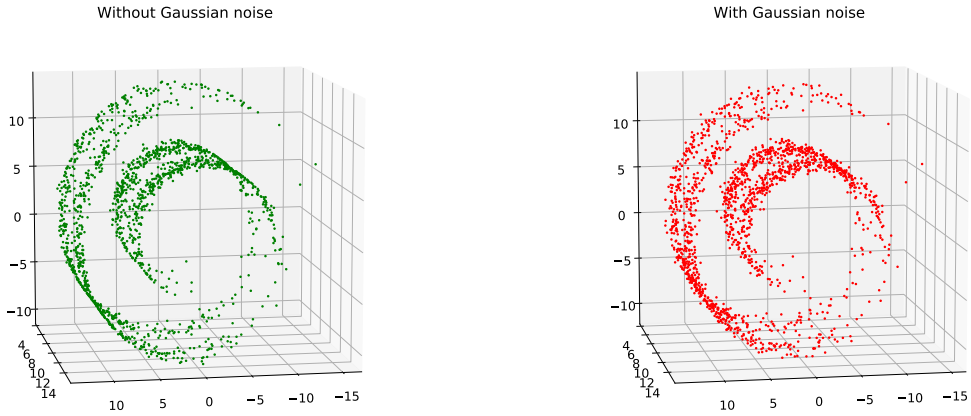
If this were to be isometrically 'unravelled' into $\mathbb{R}^2$, one would obtain the following embedding:

- **2D embedding:** $Y := \{(\frac{1}{2}\left(x\sqrt{1+x^2} + \log(x + \sqrt{1+x^2})\right), y) : (x, y) \in S\}$

This two-dimensional embedding is the target 'ground truth', and the accuracy of nonlinear dimensionality reduction can be assessed by comparing the output of Isomap with this target embedding; the error is the Procrustes distance between them.

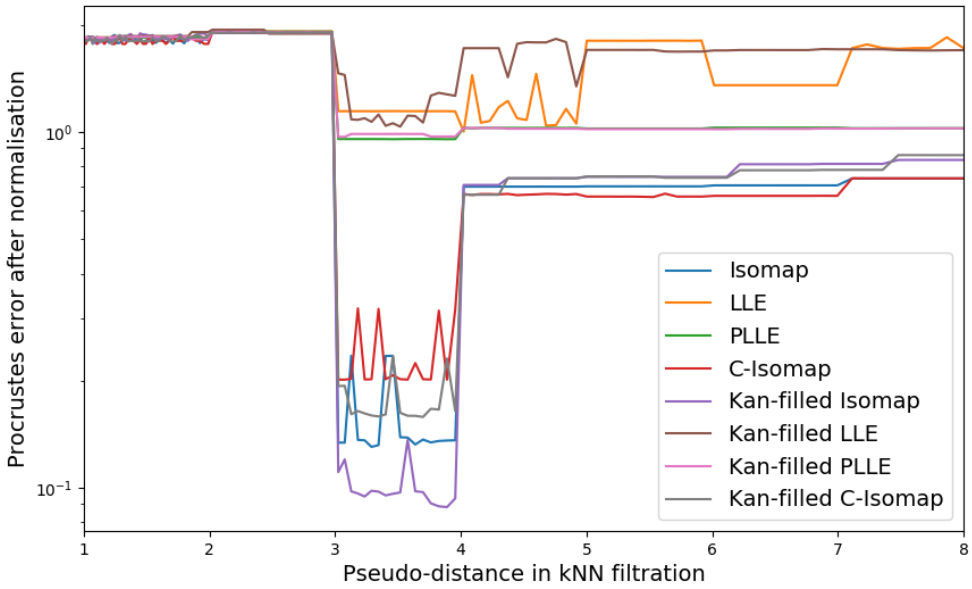
Our objective, given a set of points in the embedding in $\mathbb{R}^3$ (optionally with added noise), is to find the isometric map into the lower-dimensional space $\mathbb{R}^2$. Below

is a random 3D embedding Z, together with a noisy variant X obtained by adding an independent spherical Gaussian $N(0, \sigma^2 I)$ with $\sigma := 0.3$ to every point in the dataset Z.

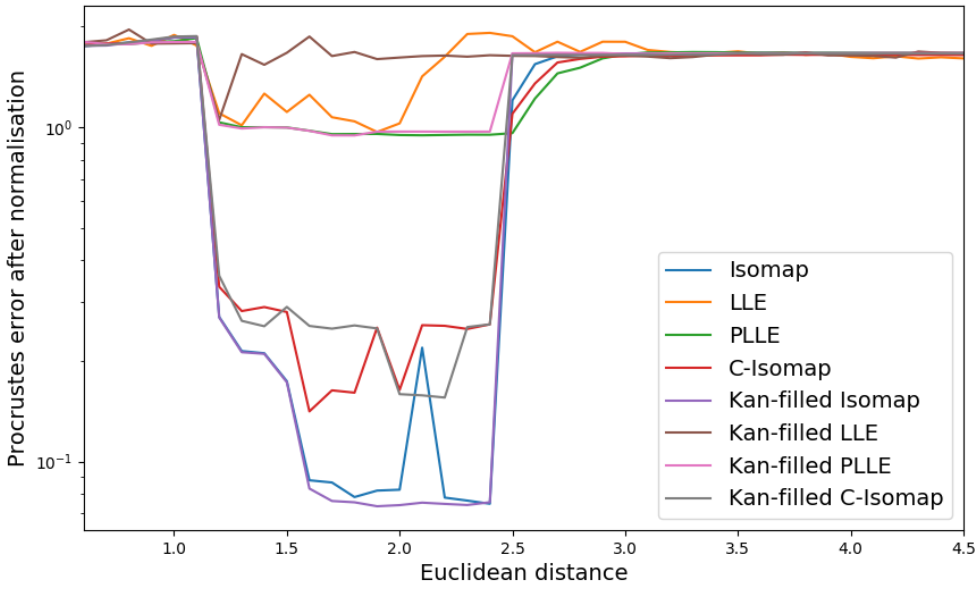


The results that follow are performed on the noisy embedding X, as it is a more realistic problem. We use sequential maxmin to sample a set of 600 landmark points, compute an alpha complex on those landmark points, and then embed it. The remaining 1000 (non-landmark) points are embedded by using PLLE to interpolate between the $k = 10$ nearest landmark points.

We observe results that are qualitatively similar to the other dataset: bounded Kan filling substantially improves the performance of both the original and conformal variants of Isomap in the $k \in (3, 4)$ interval in which reconstruction is most accurate:



The above experiments have used the k-nearest neighbours graph, for consistency with [14], [16], and, to the greatest extent possible whilst ensuring that the graph is symmetric, [15]. An alternative is to use the Euclidean metric for defining our filtration; we find that the results are similar, with Isomap performing very well in the optimal interval, and the Kan-filled variant outperforming it slightly further:



In all of these plots, there are occasional upward 'spikes' in the reconstruction error, especially for Isomap and its variants. These errors are caused by a topological defect, where poor connectivity near the middle of the manifold can cause one end of the embedding to be flipped relative to the other end. Figure 3.5 shows the bijection between the original points and the reconstructed points, where the upper-right end is correctly oriented and the lower-left end is flipped.

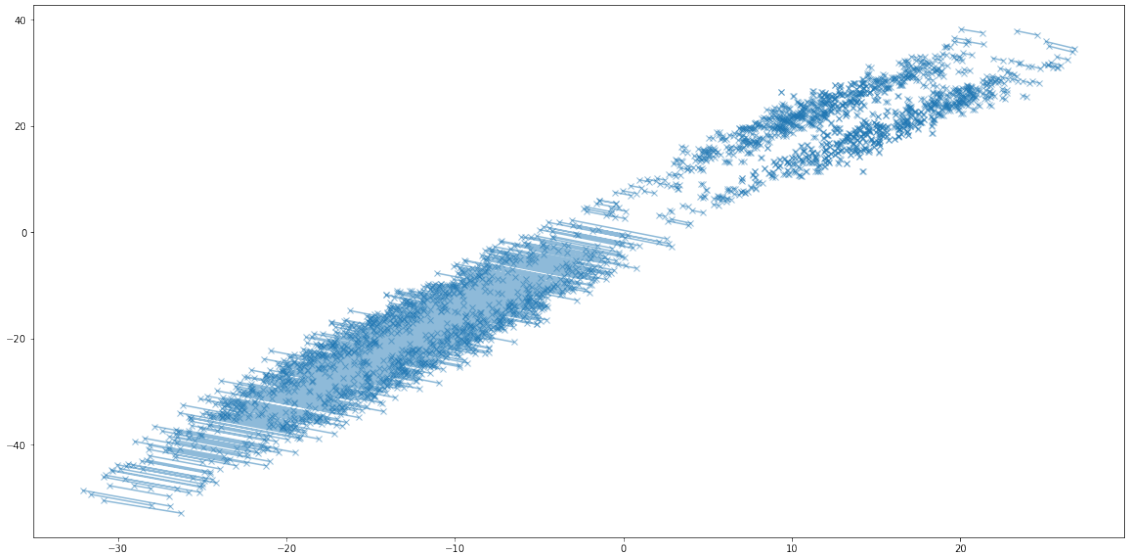


Fig. 3.5 Optimal matching between the ground truth and a defective 'twisted' reconstruction. The lower-left end is flipped relative to the upper-right end.

A sufficient condition for this problem to occur can be formulated in terms of the connectivity, or rather lack thereof, of the neighbourhood graph used by Isomap. In particular, let $\phi : X \to \mathbb{R}^d$ be the ground truth embedding and suppose that the neighbourhood graph (supported on X) fails to be d-vertex-connected. As such, X can be written as the union $X_1 \cup X_2$, where the intersection $X_1 \cap X_2$ contains at most $d-1$ points and there are no edges between $X_1 \setminus X_2$ and $X_2 \setminus X_1$. Let $\Pi \in \mathbb{R}^d$ be an arbitrary hyperplane containing the intersection $X_1 \cap X_2$, the involution $\rho_\Pi : \mathbb{R}^d \to \mathbb{R}^d$ be reflection in Π, and define a twisted embedding $\phi' : X \to \mathbb{R}^d$ as follows:

$$\phi'(x) = \begin{cases} \phi(x) & \text{if } x \in X_1; \\ \rho_\Pi(\phi(x)) & \text{otherwise.} \end{cases}$$

Observe that the shortest-path metrics induced by ϕ and ϕ' are identical, so Isomap is incapable of distinguishing between them. The situation in Figure 3.5 is an example of this phenomenon in the case where $d = 2$.

Chapter 4

Non-Euclidean and functional data

4.1 Statistics on non-Euclidean manifolds

The previous chapter considered a model in which points are sampled from an unknown subspace $M \subseteq A$ of the ambient space, with each Y_i being an independent identically distributed random variable distributed according to a measure μ on M. We introduced noise by taking $X_i := Y_i + \varepsilon_i$, where each ε_i is independently identically distributed according to some noise distribution (typically Gaussian, although we consider other distributions in the next chapter).

We proceed to generalise this idea to when A is a non-Euclidean space. For concreteness, let us suppose A is a smooth Riemannian manifold. One obstacle to immediately generalising this to an arbitrary ambient space A is that it need not carry the structure of a vector space, so we cannot directly add the random variables Y_i and ε_i or equivalently convolve their distributions. The closest analogue is to take $\varepsilon_i \in T_{Y_i}$ to be in the tangent space of A at the point Y_i, and then define:

$$X_i := \exp_{Y_i}(\varepsilon_i)$$

where $\exp_{Y_i} : T_{Y_i} \to A$ is the *exponential map* introduced in the background material.

As tempting as this definition is, it no longer makes sense to speak of the ε_i as being 'identically distributed', as the domain of the distribution is the tangent space T_{Y_i} which depends on Y_i. Instead, therefore, we shall let each ε_i be independently and identically distributed according to a noise distribution on a vector space $V = \mathbb{R}^{\dim A}$,

and define (for each $x \in A$) an isometric isomorphism $f_x : V \to T_x$. Then our previous definition becomes:

$$X_i := \exp_{Y_i}(f_{Y_i}(\varepsilon_i))$$

How do we define the isometric isomorphisms f_x? If the distribution of ε_i is spherically symmetric, then the distribution of $f_x(\varepsilon_i)$ is independent of the choice of f_x, so we may choose the isomorphism arbitrarily.

More generally, the Levi-Civita connection gives us the notion of *parallel transport*: given a choice of path $\gamma \in A$ with endpoints x and y, we obtain an identification between the tangent spaces T_x and T_y. As such, we can choose f_{x_0} arbitrarily for some basepoint $x_0 \in A$ and obtain f_x by composing with the parallel transport map along an arbitrary curve γ_x with endpoints x_0 and x. Provided the distribution of ε_i is invariant under a group of isometries known as the *holonomy group* of A, then the distribution of $f_x(\varepsilon_i)$ is independent of the choice of the path γ_x. In the Euclidean case, the holonomy group is trivial so there are no constraints on the distribution of ε_i and thus this is indeed a generalisation of the scenario in the previous chapter.

4.2 Metrics on spaces of positive-definite matrices

In this section, we shall establish inequalities relating the various metrics (described in the background material and in [65]) on the space of positive-definite matrices. To recap, given positive-definite matrices A_1 and A_2, the following distances are defined:

- **Euclidean**: $d_E(A_1, A_2) := \|A_1 - A_2\|_2$;
- **Square-root**: $d_H(A_1, A_2) := \|\sqrt{A_1} - \sqrt{A_2}\|_2$;
- **Log-Euclidean**: $d_L(A_1, A_2) := \|\log(A_1) - \log(A_2)\|_2$;
- **Procrustes**: $d_S(A_1, A_2) := \inf\|L_1 - L_2\|_2$ where the infimum is over all choices L_1, L_2 such that $L_i L_i^T = A_i$;
- **Affine-invariant Riemannian**: $d_R(A_1, A_2) := \|\log(A_1^{-1/2} A_2 A_1^{-1/2})\|_2$.

In particular, we establish absolute Lipschitz bounds relating the square-root and Procrustes distances, as well as conditional Lipschitz bounds (depending on condition number) relating the log-Euclidean and affine-invariant Riemannian metrics. We

also introduce a continuously-parametrised family of metrics intermediate between the log-Euclidean and affine-invariant Riemannian metrics.

4.2.1 Square-root and Procrustes metrics

In both the Cholesky and square-root metrics, we map each matrix A to some particular L (either square-root or lower Cholesky factor) satisfying $LL^T = A$. The Procrustes distance $d_S(A_1, A_2)$, as mentioned in [65], is defined to be the infimum distance $\|L_1 - L_2\|$ over all choices L_1, L_2 such that $L_i L_i^T = A_i$. By definition, the Procrustes distance is no greater than either the square-root or Cholesky distance, and in general is smaller.

Note that $LL^T = L'L'^T$ if and only if $L' = LR$ for some orthogonal matrix R. Consequently, $d_S(A_1, A_2)$ is the distance in $RS\Sigma_d^{d+1}$ between the shapes represented by pre-size-and-shape matrices L_1, L_2. Our choice of representatives L_1, L_2 is arbitrary; we choose the Cholesky factor since it is computationally easier to compute than the square-root.

Unlike the Cholesky distance, the Procrustes distance is again isotropic (independent of the choice of orthogonal basis). As it is isometric to the size-and-shape space $RS\Sigma_d^{d+1}$, rather than Euclidean space, all of the sectional curvatures are positive. For this reason, it is more difficult to compute Čech and alpha complexes.

Nonetheless, we can find Lipschitz constants which bound the ratio between the square-root and Procrustes distance.

Theorem 4.2.1. *Let H_1, H_2 be $d \times d$ covariance matrices. Then we have:*

$$d_S(H_1, H_2) \leq d_H(H_1, H_2) \leq \sqrt{2} d_S(H_1, H_2)$$

where d_S denotes Procrustes distance and d_H denotes square-root distance.

Proof. The first of these inequalities follows immediately from the definitions.

For the second inequality, it suffices to show that for all $\epsilon > 0$, we have $d_H(H_1, H_2) \leq \sqrt{\frac{2}{1-\epsilon}} d_S(H_1, H_2)$. Let $\delta > 0$ be small, possibly dependent on ϵ. If $d_S(H_1, H_2)$ is larger than δ, we can draw the Procrustes geodesic between H_1 and H_2 and consider the $k-1$ intermediate points spaced at regular intervals:

$$H_1 =: J_0, J_1, J_2, \ldots, J_k := H_2$$

Provided k is sufficiently large, we have $d_S(J_i, J_{i+1}) < \delta$. Now, if we can prove that $d_H(J_i, J_{i+1}) \leq \sqrt{\frac{2}{1-\epsilon}} d_S(J_i, J_{i+1})$ for each i, the triangle inequality for the square-root metric will imply the desired result.

Consequently, we may assume without loss of generality that $d_S(H_1, H_2) \leq \delta$. Since δ can be chosen to be arbitrarily small, much smaller than ϵ, we will be able to do a tangent-space approximation and linearise.

Let X_1 and X_2 be the positive-semidefinite square-roots of H_1 and H_2, respectively. Suppose that R is the orthogonal matrix which minimises the distance between X_1R and X_2. Let $Y = \frac{1}{2}(X_1R + R^T X_1)$ be the orthogonal projection of X_1R onto the space of symmetric matrices, and let Z be the orthogonal projection of X_1R onto the line ℓ through X_1 and X_2.

Since ℓ is an affine subspace of the space of symmetric matrices, we necessarily have the following:

$$\|X_1R - Y\| \leq \|X_1R - Z\|$$

by the property that the orthogonal projection Q of P onto a space is the closest point on that space to the original point P.

If we let θ be the angle at X_1 between the line ℓ and the line segment between X_1 and X_1R, then the following result holds:

$$\frac{\|X_2 - X_1R\|}{\|X_2 - X_1\|} \geq \cos(\theta) = \frac{\|X_1R - Z\|}{\|X_1R - X_1\|}$$

Now, if we could prove that $\|X_1R - Y\|^2 \geq \frac{1-\epsilon}{2}\|X_1R - X_1\|^2$, then we could chain this with the previous inequalities to show that $\|X_2 - X_1R\|^2 \geq \frac{1-\epsilon}{2}\|X_2 - X_1\|^2$, which is exactly the result we want to prove.

By Pythagoras' theorem, $\|X_1R - X_2\|^2 = \|X_1R - Y\|^2 + \|Y - X_2\|^2$, so equivalently we need to show that $\|Y - X_1\|^2 \leq \frac{1+\epsilon}{2}\|X_1R - X_1\|^2$. We will show that this holds for any sufficiently small rotation R (we know that R must be small since H_1, H_2 are close by assumption).

Since X_1 is symmetric, this is just the assertion that $X_1(R - I)$ is shrinked by a factor of at least $\sqrt{\frac{2}{1+\epsilon}}$ when projected onto the space of symmetric matrices. Firstly, we will take advantage of the ϵ of slack to absorb all but the leading-order terms, and linearise. Hence, it suffices to show that for any positive-definite matrix X_1 and antisymmetric matrix A (approximating $R - I$), we have:

$$\|X_1 A\| \geq \sqrt{2}\|\pi(X_1 A)\|$$

where π is the projection onto the space of symmetric matrices. By an orthogonal change of basis, we may assume without loss of generality that X_1 is a diagonal matrix with non-negative entries. It is straightforward to see that the diagonal of $X_1 A$ is identically zero, and for each pair of entries $(X_1 A)_{ij}, (X_1 A)_{ji}$, precisely one is positive (or zero) and the other is negative (or zero).

Let $\{a, -b\}$ be one such pair of entries. When we project onto the space of symmetric matrices, $X_1 A$ maps to $\frac{1}{2}(X_1 A - A X_1)$, which has entries of $\{\frac{1}{2}(a - b), \frac{1}{2}(a-b)\}$ in the corresponding positions. Since $a^2 + b^2 \geq (a-b)^2$, we can conclude that:

$$a^2 + b^2 \geq 2\left[\left(\frac{a-b}{2}\right)^2 + \left(\frac{a-b}{2}\right)^2\right]$$

Summing over all $\binom{n}{2}$ such pairs, one for each pair of off-diagonal entries, we get $\|\frac{1}{2}(X_1 A - A X_1)\|^2 \leq \frac{1}{2}\|X_1 A\|^2$ or equivalently $\|\pi(X_1 A)\| \leq \frac{1}{\sqrt{2}}\|X_1 A\|$. This is precisely what we needed to prove; the result therefore follows. □

There are a couple of repercussions worth mentioning. Firstly, bottleneck stability means that the logarithmic persistence diagrams of a set X of positive-semidefinite matrices, with respect to the square-root and Procrustes metrics, differ by a bottleneck distance of at most $\frac{1}{2}\log 2$. In the same way that the Vietoris-Rips complex can be used as a lazier, more easily-computable approximation to a Čech complex, we can use the square-root metric as an approximation to the Procrustes metric. As well as being able to compute pairwise distances much more quickly, we can also compute bounding balls (and therefore Delaunay complexes) exactly.

Secondly, since the bound is independent of the dimension d, it naturally extends to the infinite-dimensional case: the square-root distance between two trace-class covariance operators is at most $\sqrt{2}$ times larger than the corresponding Procrustes distance. We shall later discuss covariance operators in greater detail.

It is also tempting to ask whether the bound is tight, or whether it can be improved. It transpires that we can get pairs of matrices with distance ratio arbitrarily close to $\sqrt{2}$, namely the following pair of idempotent positive-semidefinite matrices:

$$M_0 := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, M_\varepsilon := \begin{pmatrix} 1-\varepsilon & \sqrt{\varepsilon-\varepsilon^2} \\ \sqrt{\varepsilon-\varepsilon^2} & \epsilon \end{pmatrix}$$

4.2.2 Log-Euclidean, Riemannian, and intermediate metrics

When the matrices A and B commute (that is to say, they are simultaneously diagonalisable), we have $\log(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}) = \log(B) - \log(A)$ and therefore the log-Euclidean distance (the Hilbert-Schmidt norm of the right-hand side) coincides exactly with the distance in the affine-invariant Riemannian metric (the Hilbert-Schmidt norm of the left-hand side). This suggests that it may be possible to obtain Lipschitz constants bounding the ratio $d_L(A,B)/d_R(A,B)$, analogous to the bounds we obtained for $d_S(A,B)/d_H(A,B)$.

These distances are the Hilbert-Schmidt norms of, respectively, the matrices $\log(B) - \log(A)$ and $\log(A^{-1/2}BA^{-1/2})$. If we let $A = \exp(X)$ and $B = \exp(Y)$, where X and Y are symmetric matrices, then these are the norms of the matrix logarithms of $\exp(Y - X)$ and $\exp(-X/2)\exp(Y)\exp(-X/2)$, respectively. This is reminiscent of the celebrated *Golden-Thompson inequality* [56], which states that the trace of $\exp(Y - X)$ is no greater than the trace of $\exp(-X/2)\exp(Y)\exp(-X/2)$. Unfortunately, it does not immediately prove anything about the norms of the matrix logarithms, so a different approach is required.

Firstly, we propose a common generalisation of both metrics, providing a method of smoothly interpolating between them. In particular, we define the log-p distance (where $p \in (0,1]$) to be:

$$d_p(A,B) := \left\| \frac{1}{p} \log(A^{-p/2}B^pA^{-p/2}) \right\|_2$$

This can be seen to be a metric, as $d_p(A,B) = p^{-1}d_R(A^p,B^p)$ and we already know that d_R is a metric.

The definition can be extended to the entirety of the closed interval by defining $d_0(A,B)$ to be the limit $\lim_{p\to 0} d_p(A,B)$, whose existence is demonstrated by showing that it is equal to the log-Euclidean distance $d_L(A,B)$:

Theorem 4.2.2. *The log-p distances interpolate between the log-Euclidean distance and the affine-invariant Riemannian distance, in the sense that:*

- $d_0(A,B) = d_L(A,B)$*;*

- $d_1(A, B) = d_R(A, B)$.

Moreover, when A and B commute, all of the log-p distances coincide.

Proof. As A and B are positive-definite symmetric matrices, we can express them in the form $A = \exp(X)$ and $B = \exp(Y)$. Then the formula for the log-p distance becomes:

$$d_p(A, B) := \left\| \frac{1}{p} \log \left(\exp\left(-\frac{p}{2}X\right) \exp(pY) \exp\left(-\frac{p}{2}X\right) \right) \right\|_2$$

When A and B commute, so do X and Y (as they are all simultaneously diagonalisable), so the distance simplifies to the following expression which does not depend on p:

$$d_p(A, B) := \|Y - X\|_2$$

When A and B do not commute, the situation is more complicated, but the Lie product formula in [4] implies that this still holds in the limit as $p \to 0$, so $d_0(A, B) = d_L(A, B)$ as claimed. The other fact, that $d_1(A, B) = d_R(A, B)$, holds by definition. □

We shall also prove that $d_p(A, B)$ is a monotone-increasing function of p, and therefore that the affine-invariant Riemannian distance is never smaller than the log-Euclidean distance. This depends on the Araki-Lieb-Thirring inequality proved in [57]:

Theorem 4.2.3 (Araki-Lieb-Thirring inequality)**.** *Let A, B be positive semi-definite $n \times n$ matrices, and suppose that $k \geq 0$ and $r \geq 1$ are reals. Then the following inequality holds:*

$$\mathrm{tr}[(A^{1/2} B A^{1/2})^{rk}] \leq \mathrm{tr}[(A^{r/2} B^r A^{r/2})^k]$$

If A and B are moreover invertible (and therefore positive definite), then the Araki-Lieb-Thirring inequality can be strengthened. Firstly, by a change of variables $A \mapsto A^{-1}$, $r = \frac{t}{s}$, $q = rk$, we can restate the inequality in a more symmetrical form:

Corollary 4.2.4. *Let A, B be positive definite $n \times n$ matrices, and suppose that $q \geq 0$ and $0 < s \leq t$ are reals. Then the following inequality holds:*

$$\mathrm{tr}[(A^{-s/2} B^s A^{-s/2})^{q/s}] \leq \mathrm{tr}[(A^{-t/2} B^t A^{-t/2})^{q/t}]$$

Secondly, the condition $q \geq 0$ can actually be dispensed. In particular, if we let $X = A^{-s/2}B^{s/2}$, then the left-hand-side is just the sum of the (q/s)th eigenvalues of XX^T. Both XX^T and X^TX share the same eigenvalues (these being the squares of the singular values of X), so we can rewrite the inequality as follows:

$$\mathrm{tr}[(B^{s/2}A^{-s}B^{s/2})^{q/s}] \leq \mathrm{tr}[(B^{t/2}A^{-t}B^{t/2})^{q/t}]$$

Using the fact that $(CD)^{-1} = D^{-1}C^{-1}$, this can be rewritten as:

$$\mathrm{tr}[(B^{-s/2}A^{s}B^{-s/2})^{-q/s}] \leq \mathrm{tr}[(B^{-t/2}A^{t}B^{-t/2})^{-q/t}]$$

Interchanging the rôles of the matrices A and B establishes Corollary 4.2.4 with q replaced with $-q$; consequently, the result holds for all $q \in \mathbb{R}$. For convenience, let us restate this strengthened version:

Lemma 4.2.5 (Bidirectional Araki-Lieb-Thirring inequality)**.** *Let A, B be positive definite $n \times n$ matrices, and suppose that $q \in \mathbb{R}$ and $0 < s \leq t$ are reals. Then the following inequality holds:*

$$\mathrm{tr}[(A^{-s/2}B^{s}A^{-s/2})^{q/s}] \leq \mathrm{tr}[(A^{-t/2}B^{t}A^{-t/2})^{q/t}]$$

This is almost, but not quite, what we need to prove that the log-p metrics are monotone-increasing as a function of p. Specifically, this inequality relates the traces of the qth powers of the matrices $(A^{-s/2}B^{s}A^{-s/2})^{1/s}$ and $(A^{-t/2}B^{t}A^{-t/2})^{1/t}$, whereas we would like the same relation between the Hilbert-Schmidt norms of the logarithms of the matrices. Equivalently, we would like to prove the following:

Theorem 4.2.6. *Let A, B be positive definite $n \times n$ matrices, and suppose that $s \leq t$ are positive reals. Then $d_s(A, B) \leq d_t(A, B)$, which is to say that:*

$$\mathrm{tr}[\log((A^{-s/2}B^{s}A^{-s/2})^{1/s})^2] \leq \mathrm{tr}[\log((A^{-t/2}B^{t}A^{-t/2})^{1/t})^2]$$

Proof. We utilise the fact that $\log(\lambda)^2 = \lim_{q\to 0} q^{-2}(\lambda^q + \lambda^{-q} - 2)$ for all reals $\lambda > 0$. The same expression also holds for Λ being a positive definite matrix, by choosing a diagonal basis and applying this equation elementwise. As such, it suffices to prove the following and obtain the desired result by allowing $q \to 0$:

$$\operatorname{tr}[q^{-2}((A^{-s/2}B^sA^{-s/2})^{q/s} + (A^{-s/2}B^sA^{-s/2})^{-q/s} - 2I)] \\ \leq \operatorname{tr}[q^{-2}((A^{-t/2}B^tA^{-t/2})^{q/t} + (A^{-t/2}B^tA^{-t/2})^{-q/t} - 2I)]$$

By linearity of trace, this follows from invoking Lemma 4.2.5 twice, with exponents of q and $-q$. □

Fixing t and taking the limit as $s \to 0$ extends this result to the case where s, t are nonnegative (rather than strictly positive) reals. As a special case, the log-Euclidean distance is upper-bounded by the distance in the affine-invariant Riemannian metric:

$$\operatorname{tr}[(\log(B) - \log(A))^2] \leq \operatorname{tr}[\log(A^{-1/2}BA^{-1/2})^2]$$

What about the other direction? Are the metrics Lipschitz-equivalent?

Obstruction to Lipschitz equivalence

In the pursuit of a potential counterexample, observe that the metrics are identical when A and B commute, so we would like to find examples of matrices which are 'as far away as possible' from commuting. Given that the centre of $GL(n, \mathbb{R})$ consists exactly of the scalar multiples of the identity, or equivalently the matrices whose condition number h (ratio between largest and smallest eigenvalues) is 1, a potentially fruitful fountain of counterexamples is amongst matrices with very large condition numbers.

Moreover, if we can find a pair of matrices A, B with a large ratio $\frac{d_R(A,B)}{d_L(A,B)} \geq K$, then we can find matrices A', B' within any prescribed distance of each other which similarly satisfy $\frac{d_R(A',B')}{d_L(A',B')} \geq K$; this follows from taking a geodesic (with respect to d_L) between A and B, dissecting it into $N \gg 1$ intervals of equal length (again, with respect to d_L), and letting A', B' be the endpoints of the interval with the largest separation (with respect to d_R).

Given that, it makes sense to consider the 2×2 matrices A and $B = UAU^T$, where A is the diagonal matrix with entries h and 1 and U is a rotation matrix extremely close to the identity:

$$U := \frac{1}{1+\varepsilon^2}\begin{pmatrix} 1-\varepsilon^2 & 2\varepsilon \\ -2\varepsilon & 1-\varepsilon^2 \end{pmatrix}$$

In this regime with $\varepsilon \to 0$, we can approximate the log-p distance as:

$$d_p(A, B) = \left\| \frac{1}{p}(A^{-p/2}B^pA^{-p/2} - I) \right\|_2 (1 + O(\varepsilon))$$

The lack of logarithm makes this expression more convenient for manipulation. Using computer algebra software, the squared Hilbert-Schmidt norm inside the above expression can be shown to be $8\varepsilon^2(h^p + h^{-p} - 2)(1 + O(\varepsilon))$; consequently, we have:

$$d_p(A, B) = \frac{1}{p}\sqrt{8}\varepsilon(h^{p/2} - h^{-p/2})(1 + O(\varepsilon))$$

The ratio between $d_t(A, B)$ and $d_s(A, B)$ is therefore given by:

$$\frac{d_t(A, B)}{d_s(A, B)} = \frac{t^{-1}(h^{t/2} - h^{-t/2})}{s^{-1}(h^{s/2} - h^{-s/2})}(1 + O(\varepsilon))$$

which can be made arbitrarily large (whenever $t > s$) by taking h to be large. That is to say, no pair of these metrics are Lipschitz equivalent.

In the absence of a universal Lipschitz constant, perhaps it is still possible to bound the ratio $\frac{d_t(A,B)}{d_s(A,B)}$ by some function of the condition numbers of the matrices involved.

Dependence on condition number

It transpires that the matrices A, B in the previous section, in the limit as $\varepsilon \to 0$, maximise the ratio $\dfrac{d_t(A, B)}{d_s(A, B)}$ for a given condition number h. More precisely, the following theorem holds:

Theorem 4.2.7. *If A, B are* any *pair of positive definite $n \times n$ matrices, each with condition number upper-bounded by h, then we have:*

$$1 \leq \frac{d_t(A, B)}{d_s(A, B)} \leq \frac{t^{-1}(h^{t/2} - h^{-t/2})}{s^{-1}(h^{s/2} - h^{-s/2})}$$

Recall that the lower bound has already been established. To establish the upper bound, it suffices to prove it pointwise; that is to say, at every point Σ in the space of positive-definite matrices and every tangent vector $W \in T_\Sigma$, the ratio of the norms of W in the inner products on T_Σ induced by the Riemannian metrics has this upper bound.

Lemma 4.2.8. *Suppose V is a vector space endowed with two inner products, and let $\{x_1, \ldots, x_D\}$ be a basis of orthogonal vectors with respect to both inner products. Then the ratio:*

$$\frac{\langle v, v\rangle_1}{\langle v, v\rangle_2}$$

attains its maximum at one of the basis vectors x_i.

Proof. Suppose $v = a_1 x_1 + \cdots + a_D x_D$. Then, by orthogonality of the basis vectors, we have:

$$\frac{\langle v, v\rangle_1}{\langle v, v\rangle_2} = \frac{a_1^2\langle x_1, x_1\rangle_1 + a_2^2\langle x_2, x_2\rangle_1 + \cdots + a_D^2\langle x_D, x_D\rangle_1}{a_1^2\langle x_1, x_1\rangle_2 + a_2^2\langle x_2, x_2\rangle_2 + \cdots + a_D^2\langle x_D, x_D\rangle_2}$$

Now let x_i be the basis vector which maximises the ratio $K^2 := \langle x_i, x_i\rangle_1 / \langle x_i, x_i\rangle_2$. Each term in the numerator can be verified to be at most K^2 times the corresponding term in the denominator, and therefore the ratio is bounded above by K^2. Since $v = x_i$ itself attains that bound, the result follows. □

The utility of this lemma arises from taking $V = T_\Sigma$ to be the tangent space, $\langle u, v\rangle_1$ to be the inner product on V induced by the Riemannian metric d_t, and $\langle u, v\rangle_2$ to be the inner product induced by d_s. By exhibiting a computationally convenient set of $D = \frac{1}{2}n(n+1)$ basis vectors which are simultaneously orthogonal with respect to every log-p metric, it suffices to show that:

$$\sqrt{\frac{\langle x_i, x_i\rangle_1}{\langle x_i, x_i\rangle_2}} \leq \frac{t^{-1}(h^{t/2} - h^{-t/2})}{s^{-1}(h^{s/2} - h^{-s/2})}$$

for each of the $\frac{1}{2}n(n+1)$ basis vectors x_i.

Firstly, we shall construct a set of $\frac{1}{2}n(n+1)$ mutually orthogonal geodesics emanating from the point Σ. In particular, assume Σ is the diagonal matrix $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, and consider the following $\frac{1}{2}n(n+1)$ matrices:

- For each $1 \leq i \leq n$, the matrix A whose entries are all zero with the exception of the diagonal entry $A_{ii} = 1$;
- For each $1 \leq i < j \leq n$, the matrix A whose entries are all zero with the exception of $A_{ij} = 1$ and $A_{ji} = -1$.

Then, as mentioned in [58], each curve of the form $\Gamma_A(t) = \exp(tA)\Sigma\exp(tA)^T$ is a geodesic with respect to the affine-invariant Riemannian metric, with corresponding

tangent vector $W = A\Sigma + \Sigma A^T$. The inner product $\langle W_1, W_2 \rangle$ on the tangent space is defined to be equal to the matrix inner product $\langle \Sigma^{-1/2} W_1 \Sigma^{-1/2}, \Sigma^{-1/2} W_2 \Sigma_{-1/2} \rangle$. However, since Σ is a diagonal matrix, the support (set of nonzero entries) of each of $\Sigma^{-1/2} W \Sigma^{-1/2}$ coincides exactly with the support of A; the orthogonality of these geodesics therefore follows from the fact that each of the matrices described above has disjoint support.

The result follows for the other log-p distances, as the map $\Sigma \to \Sigma^p$ induces an isometry between the space of positive-definite matrices endowed with d_p to the same space endowed with the metric $p^{-1} d_R$, and this isometry preserves (up to reparametrisation) the geodesics described above. As isometries are conformal, the geodesics remain orthogonal under the change of metric. The tangent vectors parallel to these geodesics are therefore an orthogonal basis with respect to every log-p distance.

Along one of the 'radial geodesics' $\Gamma_A(t)$, where A is a diagonal matrix with one nonzero entry, every point on the geodesic commutes with every other point, so all log-p distances coincide. Hence, we can restrict attention to the 'circular geodesics' $\Gamma_A(t)$, where A is an antisymmetric matrix with two nonzero entries: $A_{ij} = 1$ and $A_{ji} = -1$. The ratio of the lengths of the tangent vector with respect to each of the metrics is given by the limiting ratio:

$$\lim_{t \to 0} \frac{d_t(\Sigma, \exp(tA)\Sigma \exp(tA)^T)}{d_s(\Sigma, \exp(tA)\Sigma \exp(tA)^T)}$$

Conveniently, this is what we already computed in order to demonstrate the nonexistence of a universal Lipschitz constant. In particular, the matrix $\exp(tA)$ is a unitary matrix that differs from the identity only on a 2-dimensional subspace, where it induces a small rotation. As such, the previous computation still holds:

$$\frac{d_t(\Sigma, \exp(tA)\Sigma \exp(tA)^T)}{d_s(\Sigma, \exp(tA)\Sigma \exp(tA)^T)} = \frac{t^{-1}(k^{t/2} - k^{-t/2})}{s^{-1}(k^{s/2} - k^{-s/2})}(1 + O(\varepsilon))$$

Here, $k \geq 1$ is the ratio between the eigenvalues λ_i and λ_j corresponding to the two coordinates where $\exp(tA)$ differs from the identity. Across all choices of i, j, the value k (and therefore the expression) is maximised when λ_i and λ_j are the largest and smallest eigenvalues, respectively; in this case, $k = h$ is the condition number of Σ. The result follows.

Note that this proof has shown that:

$$1 \leq \frac{d_t(A,B)}{d_s(A,B)} \leq \frac{t^{-1}(h^{t/2} - h^{-t/2})}{s^{-1}(h^{s/2} - h^{-s/2})}$$

where h is the *maximum* of the condition numbers of the two matrices. Given the fact that the ratio is identically 1 whenever either A or B is the identity (i.e. the *minimum* of the condition numbers is 1), it suggests that it may be possible to strengthen the theorem to provide a bound that depends on the minimum of the two condition numbers.

4.3 Infinite-dimensional trace-class operators

In [60], the authors generalise the Euclidean, square-root, and Procrustes distances to the trace-class operators on any Hilbert space. By contrast, it is mentioned that the log-Euclidean and affine-invariant Riemannian metrics do not generalise to trace-class operators on infinite-dimensional Hilbert spaces: the eigenvalues of such an operator tends to zero, so their logarithms must approach $-\infty$ without bound. We note that the same argument extends to the log-p distances for all $p \in [0, 1]$.

The result $d_S(H_1, H_2) \leq d_H(H_1, H_2) \leq \sqrt{2} d_S(H_1, H_2)$ proved for finite-dimensional positive-semidefinite matrices can be seen to generalise to infinite-dimensional trace-class operators.

Theorem 4.3.1. *Suppose that H_1 and H_2 are infinite-dimensional trace-class operators. Then $d_S(H_1, H_2) \leq d_H(H_1, H_2) \leq \sqrt{2} d_S(H_1, H_2)$, where d_S and d_H are the Procrustes and square-root distances, respectively.*

Proof. The first side of the inequality follows immediately from the definitions of d_S and d_H, as with the finite-dimensional case. We therefore concentrate on the second side of the inequality.

Let $e_1, e_2, \ldots$ be an eigenbasis for H_1 with decreasing eigenvalues, and $e'_1, e'_2, \ldots$ be an analogous eigenbasis for H_2. Let V_n be the finite-dimensional vector space spanned by the first n vectors in the interleaved sequence $e_1, e'_1, e_2, e'_2, \ldots$.

For each $n \in \mathbb{N}$, we define the operator $H_1^{(n)}(u, v)$ to be $H_1(\pi_n(u), \pi_n(v))$ where π_n is the orthogonal projection onto the subspace V_n. The operators $H_2^{(n)}$ are defined analogously. Note that $H_1^{(2n)}(u, v)$ is upper-bounded by $H_1(u, v)$ and is lower-bounded by the restriction of $H_1(u, v)$ to the first n terms in its spectral decomposition. Consequently, the sequence converges in the square-root metric; for any $\varepsilon > 0$ we can

find n such that $d_H(H_1^{(n)}, H_1) \leq \varepsilon$ and analogously $d_H(H_2^{(n)}, H_2) \leq \varepsilon$. Since we have already established that Procrustes distance d_S is upper-bounded by square-root distance, the same follows for d_S. By the triangle inequality, it follows that:

$$d_H(H_1, H_2) \leq d_H(H_1^{(n)}, H_2^{(n)}) + 2\varepsilon$$

and similarly:

$$d_S(H_1^{(n)}, H_2^{(n)}) \leq d_S(H_1, H_2) + 2\varepsilon$$

As V_n is a finite-dimensional vector space, we have:

$$d_H(H_1^{(n)}, H_2^{(n)}) \leq \sqrt{2} d_S(H_1^{(n)}, H_2^{(n)})$$

which immediately implies:

$$d_H(H_1, H_2) - 2\varepsilon \leq \sqrt{2}(d_S(H_1, H_2) + 2\varepsilon)$$

As this is true for all $\varepsilon > 0$, it must also hold in the limit when $\varepsilon = 0$, thereby establishing the desired result. □

In the final section of this chapter, this is applied to the dataset of speech spectrogram frequency covariance operators from [59].

4.3.1 Topological data analysis of speech data

We analyse the set of frequency covariance matrices from [59]. Since X is so small ($|X| = 50$), we build a Vietoris-Rips complex instead of a witness complex. More specifically, we compute the 3-skeleton of the Vietoris-Rips complex, and ignore any higher-dimensional simplices. This way, our simplicial complex only contains $\binom{50}{1} + \binom{50}{2} + \binom{50}{3} + \binom{50}{4} = 251175$ simplices, as opposed to the 2^{50} simplices present in the full Vietoris-Rips complex. This is adequate for accurately determining the persistent homology groups H_0, H_1, and H_2, but not any higher.

After computing the skeleton of the Vietoris-Rips complex, we determine its persistent homology over $\mathbb{F}_2$ by column reduction of the boundary map matrices. Also, we keep track of the simplices responsible for creating and destroying each homology generator, rather than just the times at which the generator is created and destroyed. Keeping track of this information is helpful for visualising clusters.

In particular, we can plot the results as a *decorated dendrogram*. This is similar to a barcode, but where the H_0 bars are replaced with a dendrogram and the higher homology bars are drawn along the branches of the dendrogram corresponding to the connected components in which they are formed. Figure 4.1 shows the decorated dendrograms for both the square-root and Procrustes distance.

Fig. 4.1 Decorated dendrograms showing the persistent homology of the Čech filtration associated with a dataset of 50 covariance matrices. The red and green bars represent H_1 and H_2 homology generators, respectively. The label D_n corresponds to the word $n \mod 10$ spoken in the language $\lfloor n/10 \rfloor$.

Observe that (up to irrelevant reordering of the leaves) the dendrograms are very similar. There are many transient H_1 bars in addition to a couple of H_2 bars which

form after all of the connected components are merged. We attribute this structural similarity to the Lipschitz result we established earlier: the square-root distance is at least as long as, and no more than $\sqrt{2}$ times longer than, the Procrustes distance.

Also to note is that there is some apparent affinity for words from the same language to have covariance matrices which cluster together (in the figure, words are coloured according to the language). On the other hand, there does not appear to be an analogous clustering of the same word across different languages. These observations agree with the analysis in [59].

Chapter 5

Datasets of variable density

The approach thus far has been to choose a density threshold δ and restrict to the subset of X where some density estimator exceeds δ. This, however, suffers from much the same drawbacks as choosing an arbitrary length scale r when building a Čech complex; we obtain far more information by considering a filtration of simplicial complexes as a function of r.

This naturally gives rise to a bifiltration, indexed by the distance parameter r and density parameter δ. The density Čech complex $\mathcal{K}_{r,\delta}$ contains a simplex σ if and only if the diameter of σ is not larger than r and every vertex $x \in \sigma$ has density estimator $\geq \delta$. This bifiltration has a property called *separability* which we shall proceed to introduce.

5.1 Multifiltrations

Given k filtrations $(\mathcal{K}^1, \mathcal{K}^2, \ldots, \mathcal{K}^k)$ of the same simplicial complex $\mathcal{K}$, respectively on index sets $I_1, \ldots, I_k$, we can define a multifiltration by:

$$\mathcal{K}_{s_1,\ldots,s_k} := \bigcap_{i=1}^{k} \mathcal{K}^i_{s_i}$$

We shall describe multifiltrations obtained in this way as *separable*. Separable multifiltrations are convenient from a computational perspective, as each simplex can be described by a k-tuple of arrival times. This is not possible in a general multifiltration where the indices for which the simplex is present may form an arbitrary upward-closed set in the poset $I_1 \times \cdots \times I_k$.

Equivalently, a separable multifiltration of a simplicial complex $\mathcal{K}$ can be characterised by a function $\gamma : \mathcal{K} \to I_1 \times \cdots \times I_k$ by taking the *sublevel sets*:

$$\mathcal{K}_s := \{\sigma \in \mathcal{K} : \gamma(\sigma) \leq s\}$$

where $I_1 \times \cdots \times I_k$ is endowed with the partial order such that $(s_1, \ldots, s_k) \leq (t_1, \ldots, t_k)$ if and only if $s_i \leq t_i$ for each coordinate i. The required condition on γ is that $\gamma(\tau) \leq \gamma(\sigma)$ whenever the simplex τ is a face of the simplex σ. Equivalently, if we impose a partial order on the simplicial complex $\mathcal{K}$ by ordering the simplices by inclusion, then this is precisely the statement that γ is a functor from the poset $\mathcal{K}$ to the poset $P := I_1 \times \cdots \times I_k$.

Bounded Kan filling can be interpreted in this setting. Firstly, we introduce an operator on a nested pair $Y \subseteq X$ of simplicial complexes which yields an intermediate simplicial complex Z; secondly, this is used to generalise the bounded Kan filling construction from the first chapter to apply to arbitrary posets (and therefore multifiltrations).

Definition 5.1.1. *Let X be a simplicial complex, and $Y \subseteq X$ be a subcomplex. Then the* bounded Kan filling*, $Y \triangleright X$, is the smallest simplicial complex Z satisfying:*

- *$Y \subseteq Z$;*
- *If $\sigma \in X$ is an n-simplex with $n \geq 2$ and at most one of its proper faces $\tau < \sigma$ is not in Z, then $\sigma \in Z$.*

Equivalently, $Y \triangleright X$ can be seen as the eventual fixed point of an iterative process. We begin with $Z_0 := Y$ and define the successor $Z_{\alpha+1}$ to be $Z_\alpha \cup \sigma$ where σ is an arbitrary simplex in $X \setminus Z_\alpha$ such that at most one proper face $\tau < \sigma$ is not in Z_α. (If there is no such σ, then $Z_{\alpha+1} = Z_\alpha$ is the fixed point.) If α is a limit ordinal, we define Z_α to be the union of Z_β for all $\beta < \alpha$.

In the case where X is a finite simplicial complex, this process terminates after finitely many steps: $Y = Z_0 < Z_1 < \cdots < Z_n = Y \triangleright X$. It should be stressed that the reverse of this sequence, $Z_n, Z_{n-1}, \ldots, Z_1, Z_0$, is not *quite* the same as a *simplicial collapse*, owing to the use of 'at most one' instead of 'exactly one' [proper face] in the above definition. This is an important distinction, as the following theorem holds for bounded Kan filling but not for (the reverse of) simplicial collapse; a counterexample is obtained by taking the simplicial 3-ball $B_{18,124}$ mentioned in [55].

Theorem 5.1.2. *Let X be a finite simplicial complex homeomorphic to a n-dimensional ball, where $n \leq 2$. Let $Y \subseteq X$ be a subcomplex, and let $Z \subseteq X$ be the result of bounded Kan filling Y within X. Then $Z = X$ if and only if Y is both connected and contains all vertices of X.*

Proof. One direction is immediate: bounded Kan filling cannot change the number of vertices or connected components, so Z is connected and contains all vertices only if Y does.

The other direction is more involved. Separate arguments are necessary for $n = 1$ and $n = 2$. By induction on the number of simplices in $X \setminus Y$ (which is a set of simplices, but not necessarily a simplicial complex), we can assume that Y has already been Kan-filled (and is therefore equal to Z by definition); it remains to show that $X \setminus Y$ must be empty.

For $n = 1$, X must be a subdivision of an interval, so the only connected subset which contains all vertices is X itself; as such, the result holds vacuously.

For $n = 2$, let H be the planar dual of the 1-skeleton of X. The graph H contains a vertex r corresponding to the 'face at infinity', together with a vertex for every triangle in X. Let $G \subseteq H$ be the subgraph corresponding to $X \setminus Y$: each vertex (other than r) corresponds to a triangle not in Y, and two vertices are connected by an edge if and only if the corresponding faces share an edge in X but not in Y.

We claim that G is acyclic. In particular, if G contains a minimal cycle C, then by the Jordan curve theorem it would partition the vertex-set of X into two components (the inside and outside), violating the assumption that Y is connected and contains every vertex.

As such, G is a forest. If it contains no vertices other than r, then Y contains every triangle in X and therefore they are equal. If $v \neq r$ is an isolated point, then the corresponding triangle in X has three edges in Y and therefore Y is not already Kan-filled. Hence, we can assume that at least one of the connected components of G is a tree with at least two vertices. Such a tree must have at least two leaves, one of which is therefore not r. This leaf l corresponds to a triangle in X with two edges in Y, so Y is again not Kan-filled.

The result follows. □

One may hope that this holds for all n, but unfortunately there is a counterexample for $n = 3$: a simplicial ball X with 120 vertices, 720 edges, 1200 triangles, and 599 tetrahedra, where Y is a Hamiltonian path (linear spanning tree) and $Y = Z$. It is

necessary and sufficient that no two edges in Y are incident with the same triangle in X.

In particular, let D be the regular four-dimensional convex polytope with 120 vertices of the following form, where $\phi = \frac{1}{2}(\sqrt{5}+1)$:

- 8 vertices of the shape $(1, 0, 0, 0)$;
- 16 vertices of the shape $\frac{1}{2}(1, 0, 0, 0)$;
- 96 vertices of the shape $\frac{1}{2}(\phi, \phi^{-1}, 0, 1)$;

where we take even permutations and arbitrary sign changes of the coordinates. The boundary ∂D is homeomorphic to the 3-sphere S^3, and removing any one of the 600 tetrahedra from ∂D results in a simplicial 3-ball; let X be a simplicial 3-ball obtained in this manner.

The 2-skeleton of X is the same as the 2-skeleton of D. We shall find an enumeration $v_0, v_1, \ldots, v_{119}$ of the vertices of D such that:

- v_i and v_{i+1} share a common edge (for all $0 \leq i \leq 118$);
- v_i and v_{i+2} do not share a common edge (for all $0 \leq i \leq 117$).

If Y is the 1-dimensional simplicial complex with the vertices v_i and edges $\{v_i, v_{i+1}\}$, then the first condition implies that it is contained in D (and therefore contained in X, because their 2-skeleta agree) and the second condition implies that no two edges in Y are contained within the same triangle in D (and similarly in X). As such, Kan filling has no effect on Y, despite Y being connected and containing every vertex in X.

It remains to exhibit such a Hamiltonian path $v_0, v_1, \ldots, v_{119}$. We identify each vertex (w, x, y, z) with the quaternion $w + xi + yj + zk$, which endows the vertex-set of D with the structure of a group. The group is the order-120 *binary icosahedral group* with an order-2 centre ± 1, modulo which it is isomorphic to the order-60 simple group A_5. The vertices p and q are connected by an edge if and only if $p^{-1}q$ has real part $\frac{1}{2}\phi$.

Consider the following 119-symbol word:

$$a^9ba^9ba^9ba^9ca^9da^9da^9da^9ea^9ca^9ea^9ea^9$$

Let the vertex v_i ($0 \leq i \leq 119$) be the quaternion obtained by evaluating the length-i initial segment of this word, where the symbols correspond to the following quaternions:

- a $= \frac{1}{2}(\phi, \phi^{-1}, 0, 1)$;
- b $= \frac{1}{2}(\phi, -\phi^{-1}, 0, 1)$;
- c $= \frac{1}{2}(\phi, 0, 1, \phi^{-1})$;
- d $= \frac{1}{2}(\phi, 0, -1, \phi^{-1})$;
- e $= \frac{1}{2}(\phi, 1, \phi^{-1}, 0)$.

Then it can be verified that the v_i are all distinct elements of the binary icosahedral group (and therefore every element is visited exactly once). Moreover, v_i is adjacent to v_{i+1} because $v_i^{-1}v_{i+1}$ is either a, b, c, d, or e, all of which have real part $\frac{1}{2}\phi$. Finally, v_i is never adjacent to v_{i+2} because $v_i^{-1}v_{i+2}$ belongs to the set:

$$\{a^2, ab, ba, ac, ca, ad, da, ae, ea\}$$

and none of these quaternions have real part $\frac{1}{2}\phi$.

Counterexamples similarly exist for all higher dimensions n. In particular, given the counterexample X, Y in dimension $n = 3$, we obtain an analogous counterexample X', Y' in dimension $n = 4$ by adjoining a new vertex w and letting X' be the cone with apex w and base X, and similarly letting Y' be the cone with apex w and base Y. Iterating this, we obtain a counterexample for every $n \geq 3$.

Definition 5.1.3. *Suppose $\mathcal{K}$ is a simplicial complex endowed with a functor $\gamma : \mathcal{K} \to P$ into a poset P; this induces a filtration $\{\mathcal{K}_s : s \in P\}$ of sublevel sets. Let $\theta : P \to P$ be a function which is both (weakly) increasing and satisfies $\theta(x) \geq x$ ($\forall x \in P$).*

Then the bounded Kan filling *of the filtration is given by $\mathcal{K}_s^\theta := \mathcal{K}_s \triangleright \mathcal{K}_{\theta(s)}$.*

The familiar case of bounded Kan filling from the previous chapter corresponds to when $P = (\mathbb{R}^+, \leq)$ and $\theta(x) = cx$ for some $c > 1$.

It is immediate from the definition that $\mathcal{K}_s \subseteq \mathcal{K}_s^\theta \subseteq \mathcal{K}_{\theta(s)}$. In particular, this means that the filtrations $\{\mathcal{K}_s : s \in P\}$ and $\{\mathcal{K}_s^\theta : s \in P\}$ are *θ-interleaved*.

5.1.1 Bifiltered witness complexes

Suppose we have a set X with a distinguished set $L \subseteq X$ of landmark points. Then, given some function $f : X \to \mathbb{R}$ (such as the distance to the kth nearest neighbour of the point), we can induce a second filtration on the simplices as $f(\sigma) := \inf\{f(x) : x \text{ witnesses of } \sigma\}$. A separable bifiltration is then obtained as the intersection of this with the usual distance filtration.

If X is of non-uniform density – an example being Carlsson's Klein bottle [41] with the high-density 'three-circle space' embedded inside it – then landmark sampling by sequential maxmin samples disproportionately many points from regions of comparatively low density. At this point, it is worth exploring exactly why this happens and what can be done to circumvent it, so that we can proceed with our endeavour to analyse the persistent homology of a bifiltered witness complex.

5.2 Landmark sampling, revisited

When sampling the subset $L \subseteq X$ to form the landmark points of a witness complex, there are two commonly-used alternatives: random sampling and sequential maxmin.

The drawback of random sampling is that it often produces triangles in the Delaunay triangulation which are difficult to witness, consequently leading to unwanted gaps in the complex. Sequential maxmin, on the other hand, necessarily produces Delaunay triangles with angles bounded away from 0 and π, and we shall establish this soon in a generalised setting.

One instance where sequential maxmin is less than ideal is when the set X is distributed with non-uniform density. In these cases, the procedure selects disproportionately many points from large regions of low density, and very few from small regions of high density. Indeed, when a set of $|X| = 833400$ witness points was subsampled by sequential maxmin to obtain $|L| = 4167$ landmark points, only 91 of those were in the upper half of the original points in terms of density (calculated by distance to kth nearest neighbour); the remaining 4076 landmark points were sampled from regions of below-median density.

This situation is particularly undesirable in the case of a bifiltered witness complex, where we have a persistence parameter corresponding to density: the dearth of landmark points in high-density regions severely limits the accuracy with

which the witness complex can approximate the geometry of those high-density regions.

5.2.1 Conformal sequential maxmin

One way to overcome the shortcomings of sequential maxmin in cases where density is variable is to perform sequential maxmin with respect to a different metric – one with respect to which X has uniform density. If we had a Riemannian metric on the ambient n-manifold M, we could define a new Riemannian metric (conformally equivalent to the old one) by scaling the inner product at each point:

$$\langle u, v \rangle_x^\star := \ell(x)^{-2} \langle u, v \rangle_x$$

where $f(x) = \ell(x)^{-n}$ is the probability density function from which X has been sampled.

Whilst conceptually elegant, this is awkward in practice for several reasons:

- The ambient space may not be a Riemannian manifold;
- We only have access to a finite sample X, not the probability distribution, so this would involve density estimation anyway;
- Reconstructing a distance metric from a Riemannian metric involves finding the minimum integral over all paths between two points, which may be intractable to compute exactly.

Instead, we simplify this by omitting the involvement of Riemannian metrics and density estimation, and instead replace it with a direct definition of pseudo-distance between two points:

$$d^\star(x, y) := \frac{d(x, y)}{\sqrt{\ell(x)\ell(y)}}$$

where $d(x, y)$ is the original (typically Euclidean) distance between x and y, and $\ell : X \to (0, \infty)$ is the 'typical length scale'. Note that this is the same definition as is used in the conformal variant of Isomap introduced in [16].

In the case where $X = M$ is an n-manifold with a known density function $f : M \to \mathbb{R}$, we can take $\ell(x) = f(x)^{-1/n}$. In the case where X is a finite set, we can take $\ell(x)$ to be the distance between x and its kth nearest neighbour in X. The

sequential maxmin procedure can then be repeated with $d^\star$ instead of d, which yields a more density-uniform selection of landmark points as elucidated in Figure 5.1.

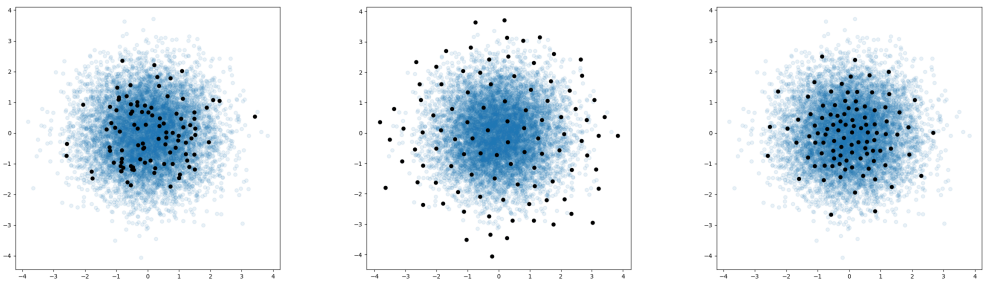


Fig. 5.1 The left panel shows random sampling, which reflects the density of the original distribution but often causes subsampled points to occur very close together (as a result of the birthday paradox) as well as large unoccupied voids. The middle panel is sequential maxmin, which gives a more uniform distribution of sampled points but completely ignores the density of the original distribution. The right panel is the conformal variant we propose, which has both desirable properties.

It is worth noting that $d^\star$ does not necessarily satisfy the triangle inequality. As such, it is necessary to formulate the theory of sequential maxmin in this more general setting:

Definition 5.2.1. *A* semimetric space *is a pair (X, d) comprising a set X endowed with a function $d : X \times X \to \mathbb{R}_{\geq 0}$ satisfying:*

- $d(x, y) = d(y, x) > 0$ *($\forall x, y \in X$ with $x \neq y$)*
- $d(x, x) = 0$ *($\forall x \in X$)*

Definition 5.2.2. *Let (X, d) be a semimetric space. Then a* sequential maxmin sequence*, henceforth abbreviated to* SMMS*, is a (finite or infinite) sequence of points in X:*

$$x_1, x_2, x_3, \ldots$$

such that each point x_j is 'as far as possible from the previous points' in the sense that:

$$\min_{i<j} d(x_i, x_j) = \sup_{y \in X} \min_{i<j} d(x_i, y)$$

Observe that if X is compact with respect to some topology, and d is continuous with respect to the same topology, then the supremum on the right-hand side of the above definition is always attained by at least one point; as such, an infinite SMMS can be constructed iteratively. [1]

Lemma 5.2.3. *Let $f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ be a strictly increasing function satisfying $f(0) = 0$. Then $(X, f \circ d)$ is also a semimetric space, and $x_1, x_2, \ldots$ is an SMMS for (X, d) if and only if it is an SMMS for $(X, f \circ d)$.*

Proof. By the property that $f(0) = 0$ combined with the strictly increasing nature of the function, we have $f(d(x, y)) > 0$ if and only if $d(x, y) > 0$; as (X, d) is a semimetric space, this is equivalent to $x \neq y$ as required. Moreover, $d(x, y) = d(y, x)$ implies that $f(d(x, y)) = f(d(y, x))$, so $(X, f \circ d)$ satisfies both properties needed to be a semimetric space.

Observe that f is an isomorphism between the image of d and the image of $f \circ d$, considered as totally-ordered sets. The definition of SMMS depends only on the order relationship between the distances, thereby establishing the lemma. □

The utility of this lemma is that when proving results about sequential maxmin sequences on the unit sphere S^n, it is immaterial whether we use the subspace metric (where antipodal points are a distance of 2 apart) or the path metric (where they are a distance of π apart).

5.2.2 Uniform density of conformal sequential maxmin

Let M be a compact n-manifold endowed with a Riemannian metric g which induces the path metric $d : M \times M \to \mathbb{R}_{\geq 0}$. Let $f : M \to \mathbb{R}$ be a density function with uniformly continuous logarithm, and furthermore suppose f is normalised such that:

$$\int_M f(x)\, d\mu_g = 1$$

where μ_g is the standard Riemannian measure. This allows us to define a probability measure on the Borel subsets of M:

$$\lambda(A) := \int_A f(x)\, d\mu_g$$

[1] As the choice of x_j may be non-unique, in that the supremum could be attained by more than one point, this construction technically relies on the axiom of dependent choice.

Setting $\ell(x) := f(x)^{-1/n}$, we define the semimetric as before:

$$d^\star(x,y) := \frac{d(x,y)}{\sqrt{\ell(x)\ell(y)}}$$

We are now in a position to prove the following statement, which justifies the choice of semimetric $d^\star$ for the purposes of conformal sequential maxmin.

Theorem 5.2.4. *There exist constants $0 < c, C < \infty$ such that, for any Jordan measurable subset $A \subseteq M$ and any SMMS $x_1, x_2, \ldots$ chosen with respect to $(M, d^\star)$, there exists $m_0 \in \mathbb{N}$ such that for all $m \geq m_0$ we have:*

$$c\lambda(A) \leq \frac{1}{m}\sum_{i=1}^{m}[x_i \in A] \leq C\lambda(A)$$

where the Iverson bracket $[x_i \in A]$ is defined to be 1 if $x_i \in A$ and 0 otherwise.

Proof. Fix some arbitrary $\kappa > 1$.

Define an *h-cube* to be the image of $[0,h]^n$ under a continuous map $\phi : [0,h]^n \to M$ which is a diffeomorphism onto its image (viewed as a Riemannian manifold) and obeys the following Lipschitz condition with respect to the semimetric $d^\star$:

$$\kappa^{-1}d^\star(\phi(x),\phi(y)) < \|x-y\| < \kappa d^\star(\phi(x),\phi(y))$$

The Jordan-measurability of A means that, for all $\epsilon > 0$, we can find some $h > 0$ and a finite collection of disjoint h-cubes, each a subset of A, with total measure at least $\lambda(A) - \epsilon$. Likewise, we can find a finite collection of h-cubes which cover A and have total measure at most $\lambda(A) + \epsilon$.

Consequently, to establish the result in general, it suffices to prove it for the special case where A is a single h-cube. Suppose we have a finite SMMS, $x_1, x_2, \ldots, x_{m-1}, x_m$, where m has been chosen sufficiently large such that the $d^\star$ distance between x_{m-1} and x_m is $r \leq \frac{h}{4\kappa}$.

It follows that, for every pair x_i, x_j of points in the SMMS, we have:

$$\kappa^{-1}r < \|\phi^{-1}(x_i) - \phi^{-1}(x_j)\|$$

Similarly, for every point $y \in A$, we can guarantee that there exists some x_i such that:

$$\|\phi^{-1}(x_i) - \phi^{-1}(y)\| \leq \kappa r$$

The first inequality implies that every subcube of $[0,h]^n$ of sidelength $\frac{r}{\kappa\sqrt{n}}$ contains at most one preimage of a point in the SMMS (because the diameter of the subcube is $\kappa^{-1}r$). The second inequality implies that every subcube of sidelength $2\kappa r$ contains at least one preimage of a point in the SMMS (because the subcube contains a ball of radius κr).

Consequently, the h-cube must contain between $\lfloor \frac{h}{2\kappa r} \rfloor^n$ and $\lceil \frac{h\sqrt{d}}{\kappa^{-1}r} \rceil^n$ points from the SMMS. As the expressions inside the floor and ceiling functions are each at least 2, the application of the floor and ceiling functions can only perturb them by a factor of at most 2. As such, the ratio between the upper and lower bounds on the number of points in the h-cube is bounded above by $(8\kappa^2\sqrt{d})^n$, which is independent of the value of r.

It remains to calculate $\lambda(\phi([0,h]^n))$. Let ℓ^+ and ℓ^- be upper and lower bounds, respectively, for the function ℓ restricted to this h-cube. Each eigenvalue of the Jacobian of the diffeomorphism ϕ is bounded above by $\kappa\frac{\ell^+}{\ell^-}$ and below by its reciprocal, $\kappa^{-1}\frac{\ell^-}{\ell^+}$. It follows that the determinant is bounded above and below by the nth power of these, and therefore we have:

$$\left(h\kappa^{-1}\frac{\ell^-}{\ell^+}\right)^n < \lambda(\phi([0,h]^n)) < \left(h\kappa\frac{\ell^+}{\ell^-}\right)^n$$

The uniform continuity of $\log \ell$ implies that these bounds differ by a constant factor $\left(\kappa\frac{\ell^+}{\ell^-}\right)^2$ which can be made arbitrarily close to κ^2 (by reducing h and concomitantly reducing r) and therefore arbitrarily close to 1 (by initially choosing κ close to 1). Combining this with the bounds on the number of SMMS points inside the h-cube, the result follows.

□

Note that it was not particularly efficient to partition $[0,h]^n$ into cubes when the dimension n is large: the ratio of the diameter of the cube to its inradius is $2\sqrt{n}$. We can instead partition $[0,h]^n$ by taking the Voronoi partition with respect to an arbitrary lattice; the quantity $2\sqrt{n}$ can then be replaced by twice the *packing-covering ratio* of the lattice. Rogers proved in [51] that, irrespective of the dimension, this can be made as low as 3; this results in absolute dimension-independent bounds on the quantities $c^{1/n}$ and $C^{1/n}$ in the above theorem.

5.2.3 Necessity of conditions

One may then ask whether these constant factors are necessary at all. In this section, we provide an explicit construction of a smooth 1-manifold whereby the ratio between C and c cannot be reduced below $\sqrt{2}$. This example has a uniform density function f, so the distinction between ordinary sequential maxmin and its conformal counterpart is irrelevant.

We also show in passing that the 'Jordan measurable' criterion cannot be relaxed to 'Lebesgue measurable' or even 'open'.

Definition 5.2.5. *A* segmentation *of a semimetric space (X, d) is a function $\sigma : X \to I$ (where I is an arbitrary, typically finite, index set) such that:*

$$((\sigma(x) = \sigma(y)) \wedge (\sigma(x) \neq \sigma(z))) \implies (d(x, y) < d(x, z))$$

Lemma 5.2.6. *Let (X, d) be a semimetric space and $x_1, x_2, x_3, \ldots$ be a SMMS. Then, for any segmentation $\sigma : X \to I$ and $\alpha \in I$, the restriction of $x_1, x_2, x_3, \ldots$ to the fibre $F_\alpha := \sigma^{-1}(\{\alpha\})$ is an SMMS in (F_α, d).*

Proof. Suppose otherwise. Then there exists $n \in \mathbb{N}$ such that:

$$\min_{i<n, x_i \in F_\alpha} d(x_i, x_n) \neq \sup_{y \in F_\alpha} \min_{i<n, x_i \in F_\alpha} d(x_i, y)$$

so we can find $y \in F_\alpha$ such that:

$$\min_{i<n, x_i \in F_\alpha} d(x_i, x_n) < \min_{i<n, x_i \in F_\alpha} d(x_i, y)$$

However, as σ is a segmentation of X, it follows that both sides of the inequality remain unchanged when we drop the $x_i \in F_\alpha$ condition (as every point in F_α is closer to x_n than every point outside F_α; the same is true of y). That is to say:

$$\min_{i<n} d(x_i, x_n) < \min_{i<n} d(x_i, y)$$

ergo $x_1, x_2, x_3, \ldots$ cannot be an SMMS in the original space (X, d). Contradiction. □

Lemma 5.2.7. *Let $x_1, x_2, \ldots$ be a SMMS of the unit circle $S^1 \subseteq \mathbb{C}$ endowed with either the path or subspace metric. Then the set $\{x_1 x_1^{-1}, x_2 x_1^{-1}, \ldots, x_{2^k} x_1^{-1}\}$ is the set of 2^kth roots of unity.*

Proof. The proof proceeds by induction on k. The base case $k = 0$ holds by definition; $x_1 x_1^{-1} = 1$ is indeed the unique 1st root of unity. As the isometric automorphisms of S^1 form a transitive group, we can assume without loss of generality that $x_1 = 1$. By the inductive hypothesis, the points $x_1, \ldots, x_{2^{k-1}}$ coincide with the 2^{k-1}th roots of unity. Then the remaining 2^kth roots of unity are precisely the points maximally distant from the first 2^{k-1} points. As such, an inner induction on i shows that the points $x_{2^{k-1}}, x_{2^{k-1}+1}, \ldots, x_{2^{k-1}+i}$ form a subset of the primitive 2^kth roots of unity. The result follows. □

Corollary 5.2.8. *The condition 'Jordan measurable' cannot be relaxed to 'Lebesgue measurable' in the previous theorem.*

Proof. Take $M \subseteq \mathbb{R}^3$ to be the disjoint union of the circles $\{z = 0\} \cap \{(x-3)^2 + y^2 = 1\}$ and $\{y = 0\} \cap \{(x+3)^2 + z^2 = 1\}$, endowed with the Euclidean metric inherited from $\mathbb{R}^3$. This admits a segmentation into its two connected components, so every SMMS on M must be some interleaving of an SMMS on each of the two connected components.

Note that the point $(-4, 0, 0)$ is maximally distant from every point on the first circle, and the point $(4, 0, 0)$ is maximally distant from every point on the second circle. Consequently, it follows that the second point in any SMMS on M must be $(\pm 4, 0, 0)$, and all subsequent points in that connected component will have algebraic coordinates.

Consequently, the measure-0 set $A := M \cap \overline{\mathbb{Q}}^3$ of algebraic points will capture at least half of the points sampled by sequential maxmin, contradicting the property of its measure being equal to the asymptotic proportion of sampled points in A. □

We cannot even relax 'Jordan measurable' to 'open', as it is easy to find open sets of arbitrarily small measure which contain A as a subset: fix an enumeration of A, and place an open ball of measure $\epsilon 2^{-n}$ around the nth point; a union bound shows that the total measure of the union of these balls cannot exceed ϵ.

Theorem 5.2.9. *The set* $\{(x, y) \in \mathbb{R}^2 : (x^2 + y^2 - 1)((x-6)^2 + y^2 - 2) = 0\}$ *endowed with the subspace metric is a* 1*-manifold such that no SMMS has an asymptotic uniform density.*

Proof. This 1-manifold admits a segmentation into its two connected components, which are disjoint circles (of radii 1 and $\sqrt{2}$). As with the previous two-circle space,

any SMMS must be an interleaving of sequential maxmin sequences on the two connected components.

It remains to determine how the sequences are interleaved. We claim that for sufficiently large $n \in \mathbb{N}$, the first 2^{n+1} points are partitioned equally between the two circles; the result would follow immediately from this claim, since one of the circles has a measure $\sqrt{2}$ times greater than the other.

To verify the claim, note that the pigeonhole principle implies that there is at least one circle with at least 2^n of those points. The introduction of a (2^n+1)th point would be a distance of at most $\sqrt{2}\pi 2^{-n}$ from the nearest existing point, whereas the spacing between the first 2^n points on the other circle is at least $(2-o(1))\pi 2^{-n}$. Consequently, both circles 'fill up' with 2^n points before a (2^n+1)th point is added to either circle. □

Observe that the proof used circles of radii 1 and $\sqrt{2}$, but could equally work with circles of radius 1 and $2-\epsilon$ for arbitrarily small fixed positive $\epsilon > 0$. (The value of 'sufficiently large' $n \in \mathbb{N}$ in the details of the proof increases concomitantly.) As such, the ratio C/c between the constants cannot be reduced below 2. This is tight for 1-manifolds.

The circle S^1 can be viewed as the quotient $\mathbb{R}/\mathbb{Z}$ of the real line by the integer lattice. An analogous construction for n-manifolds is to take the torus obtained by quotienting $\mathbb{R}^n$ by a lattice Λ. This raises the question: what properties does Λ need to have to ensure an analogue of Lemma 5.2.7 holds? Sufficient conditions are:

- $\Lambda \cup D$ is geometrically similar to Λ, where D is the set of *deep holes* maximally distant from the nearest point in Λ;
- The minimum distance between two deep holes is at least as large as the distance between a deep hole and the nearest lattice point.

In addition to $\mathbb{Z}$, which gives the lower bound $C/c \geq 2$ for 1-manifolds, there are higher-dimensional lattices with these properties. The square lattice $\mathbb{Z}^2$ provides a bound of $C/c \geq 2$; the hexagonal lattice A_2 provides a stronger lower bound of $C/c \geq 3$. For 4-manifolds, we can take the quotient of $\mathbb{R}^4$ with the D_4 *lattice* (described in [69]), $\mathbb{Z}^4 \cup (\mathbb{Z}+\frac{1}{2})^4$, which establishes a lower bound of $C/c \geq 4$. We conjecture that there is no dimension-independent upper bound on C/c.

5.3 Experimental results

We now apply the methods explored herein to suitably non-uniform datasets. The space of natural images is a good choice, as it was shown in [41] to exhibit different topologies at different density scales. To begin with, a more tractable synthetic dataset is explored to assess how the presence of noise affects the accuracy of the persistent homology barcode.

5.3.1 Synthetic data

A good test-case for these methods, which has the property of having a non-uniform density, is the set $\{Y_i : i \in \{1, 2, \ldots, N\}\}$ where each point Y_i is independently identically distributed according to the following distribution:

$$Y_i = \frac{M_i Z_i}{\|M_i Z_i\|_2}$$

where $Z_i \sim N(0, I)$ is a standard trivariate Gaussian, and the matrix M_i is distributed from a discrete set of idempotent diagonal matrices:

- The identity matrix, $diag(1, 1, 1)$, with probability $\frac{5}{8}$;
- The projection matrix $diag(1, 1, 0)$ with probability $\frac{1}{8}$;
- The projection matrix $diag(1, 0, 1)$ with probability $\frac{1}{8}$;
- The projection matrix $diag(0, 1, 1)$ with probability $\frac{1}{8}$;

In other words, each point Y_i is chosen from a mixture distribution: it has a probability of $\frac{5}{8}$ of being sampled from the uniform measure on the unit sphere S^2, and a probability of $\frac{1}{8}$ of being sampled from the uniform measure on each of the three 'equators' obtained by intersecting a coordinate plane with the unit sphere.

It is instructive to add a small amount of noise: we shall apply our methods to the dataset $\{X_i := Y_i + \varepsilon_i : i \in \{1, \ldots, N\}\}$ where we have added an additional 'noise term' ε_i to each point in the dataset. Figure 5.2 is a plot where ε_i is taken from a spherically-symmetric Cauchy distribution.

We use the distance to the kth nearest neighbour with $k = 50$ as the length scale for applying conformal sequential maxmin, construct a witness complex, and then take a filtration with respect to this same density estimator. There are seven long

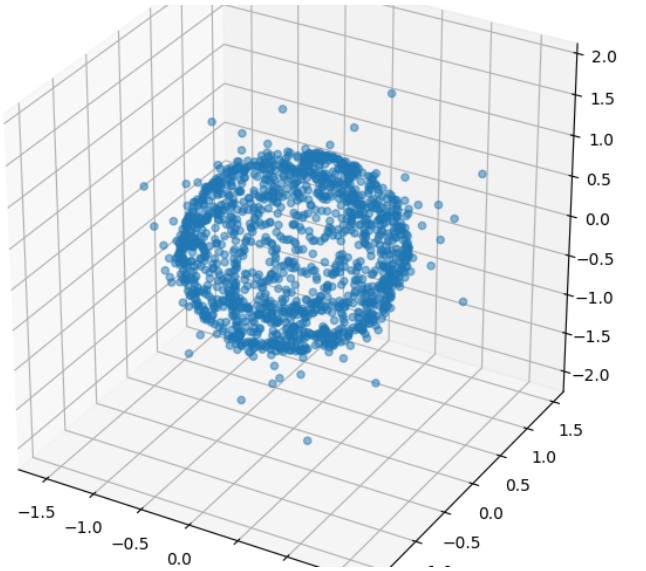


Fig. 5.2 The synthetic dataset, sampled with Cauchy noise. The points outside a small bounding box are not shown – the Cauchy distribution is heavy-tailed and many of the points are so distant from the origin that attempting to plot all of them would result in the unit sphere occupying an imperceptibly tiny region of the plot.

persistent H_1 bars in the barcode Figure 5.3 (which is precisely what we expect since the union of the three circles has $\beta_1 = 7$) which promptly disappear and are replaced with a single H_2 generator (the surface of the sphere). The only problem is a second persistent H_0 bar; this is caused by an isolated point in the heavy tails of the noise distribution, and is an artefact of the conformal sequential maxmin procedure. Such an artefact is easily removable by deleting any isolated points from the simplicial complex.

5.3.2 Natural images

The space of natural images explored in [54] and [41] can be given the structure of a trifiltered simplicial complex. Specifically, let $\mathcal{C}_{r,k^{-1},s}$ be the Čech complex obtained by the union of radius-$\frac{1}{2}r$ balls centred on the 'high-density' points $X_{k,s} \subseteq X$ defined as:

$$X_{k,s} := \{x \in X \text{ such that } |B_s(x) \cap X| > k\}$$

Evidently, this is a trifiltration, as $\mathcal{C}_{r,k^{-1},s}$ is increasing as a function of each of the three parameters r, k^{-1}, and s. Note that this is not, in general, separable:

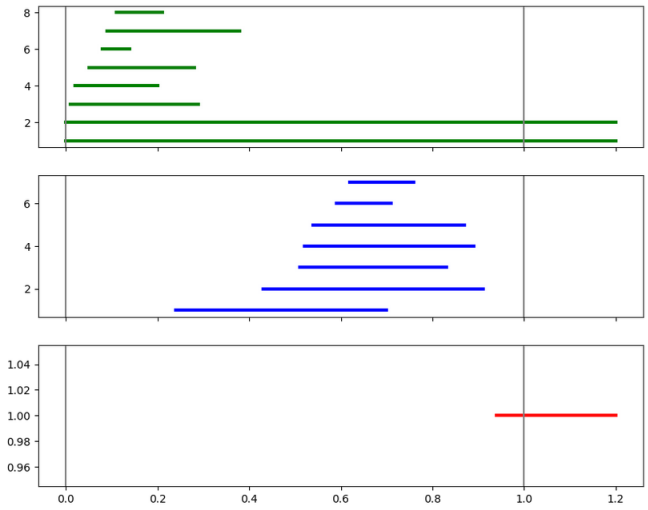


Fig. 5.3 Persistent homology barcode of the density filtration of the synthetic dataset.

it is possible for a vertex to be present in both X_{k_1,s_1} and X_{k_2,s_2}, but absent from $X_{\max(k_1,k_2),\min(s_1,s_2)}$. For the experiments that follow, we take the 'slice' at $k = 100$ and focus on the resulting (separable) bifiltration. We also truncate the bifiltration and restrict attention to the densest 10% of points (so $|X| = 416700$).

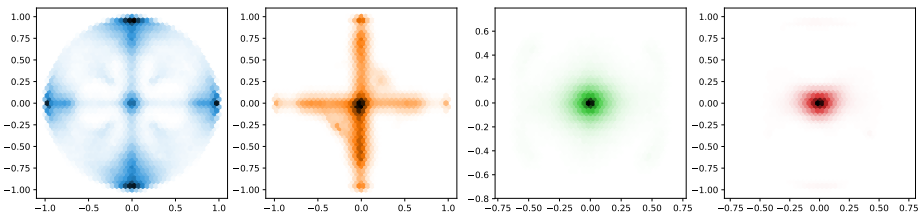


Fig. 5.4 Projections of the 8-dimensional Mumford dataset onto four pairs of coordinate axes with respect to the Discrete Cosine Transform basis.

Plotted in Figure 5.4 are the joint distributions of each of the four pairs (x_0, x_1), (x_2, x_3), (x_4, x_5), and (x_6, x_7) of coordinates (in the Discrete Cosine Transform basis from [54]) of the points of X, revealing that the majority of the variability is supported on the first four coordinates.

As discussed in [41], the points in $\mathcal{M}$ of highest density cluster around three circles. The circles S_v and S_h are visible in each of the first two plots as a pair of vibrant bands (one vertical and one horizontal); the third circle is visible around the circumference of the first plot. These features should be detectable using persistent

homology, and consequently this can be used to test the relative efficacy of different choices of simplicial complex.

Owing to the large number of points, we consider both conventional and squared witness complexes. We take $|L| = 400$ landmark points sampled from the full space of $|X| = 416700$ witness points, so approximately 0.1% of the points are landmarks.

Figure 5.5 depicts the H_1 barcodes for a conventional witness complex (left), squared witness complex (middle), and its Kan-filled counterpart (right). The five long bars present in the latter two cases are the homology generators of the 'three-circle space' [41] consisting of two disjoint circles S_v and S_h joined by a third circle S_{lin} which intersects each of the other circles in exactly two points.

The conventional witness complex spuriously has a sixth long H_1 generator persisting to infinity, which conjecturally is caused by a nearly-cyclic quadrilateral. In the squared witness complex, it has finite length, and in the Kan-filled counterpart is considerably shorter (and would disappear entirely if a smaller value of c were used).

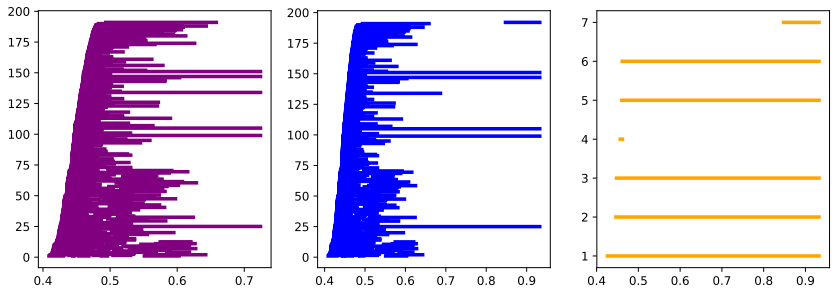


Fig. 5.5 First homology barcodes for a conventional witness complex, squared witness complex, and Kan-filled squared witness complex.

Virtually all of the short-lived noise bars are completely eradicated by bounded Kan filling, leaving the homology generators corresponding to genuine topological features. Figure 5.6 shows both β_1 (left) and β_2 (right) for the two filtrations of the squared witness complex, with the latter being empty for the Kan-filled complex.

This is highly desirable. If we take an arbitrary point in the Kan-filled filtration, corresponding to a vertical line in Figure 5.5, it is likely to intersect none of the short β_1 bars and thereby exactly capture the homology of the three-circle space without requiring the use of persistence. The same was true of the synthetic dataset we explored beforehand.

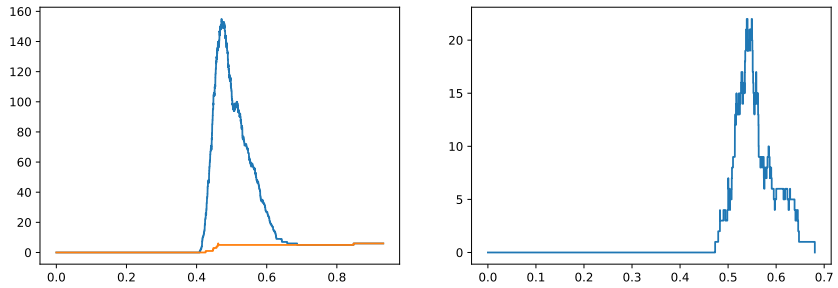


Fig. 5.6 Betti numbers β_1 and β_2 for the squared witness complex with (orange) and without (blue) bounded Kan filling.

Being able to compute ordinary homology instead of persistent homology does not seem immediately useful, since both are equally easy to compute by using elementary row/column operations to convert the boundary map matrix into a suitably reduced form. However, it is particularly useful for a multifiltration: we can apply this approach to fix one of the parameters (and allow the other parameter to vary as usual) so that one-dimensional persistence can be used instead of multidimensional persistence.

Chapter 6

Potential for future research

Aside from general research directions, such as enhancing the manifold detection methods introduced herein, there are several specific questions left open in this dissertation:

6.1 Specific open questions

In Chapter 4, we introduced the log-p distances on the space of positive definite matrices and proved that, for $0 \leq s \leq t \leq 1$, we have:

$$1 \leq \frac{d_t(A,B)}{d_s(A,B)} \leq f_{s,t}(\max(h_A, h_B))$$

where $f_{s,t}$ is some function depending on the *maximum* of the condition numbers of the matrices A and B. We mentioned that it is plausible that the Lipschitz constant can be strengthened to be a function of the *minimum* of the two condition numbers:

$$1 \leq \frac{d_t(A,B)}{d_s(A,B)} \leq g_{s,t}(\min(h_A, h_B))$$

The proof does not immediately apply to the strengthened statement, however, so this remains a conjecture.

In Chapter 5, the following was proved:

Theorem. *Let M be a compact n-manifold endowed with a Riemannian metric g which induces the path metric $d : M \times M \to \mathbb{R}_{\geq 0}$. Let $f : M \to \mathbb{R}$ be a density function*

with uniformly continuous logarithm, and furthermore suppose f is normalised such that:

$$\int_M f(x)\, d\mu_g = 1$$

where μ_g is the standard Riemannian measure.

Then there exist constants $0 < c, C < \infty$ such that, for any Jordan measurable subset $A \subseteq M$ and any SMMS $x_1, x_2, \ldots$ chosen with respect to $(M, d^\star)$, there exists $m_0 \in \mathbb{N}$ such that for all $m \geq m_0$ we have:

$$c\lambda(A) \leq \frac{1}{m} \sum_{i=1}^{m} [x_i \in A] \leq C\lambda(A)$$

where the Iverson bracket $[x_i \in A]$ is defined to be 1 if $x_i \in A$ and 0 otherwise.

We remarked that a result of Rogers in [51] implies that there exists an absolute dimension-independent bound on each of $c^{1/n}$ and $C^{1/n}$. We conjectured that there is not a dimension-independent bound on c and C, but this remains to be proved one way or the other.

6.2 Further theory to be developed

Chapter 3 mentions that the manifold detection framework differs from [2] in several aspects, including that it takes as input a density estimate evaluated at the sample points. Given a particular choice of density estimator, such as a kernel density estimator, it should be possible to determine asymptotically how many points need to be sampled in order for the manifold to be determined up to homotopy equivalence. This would make the research in Chapter 3 more directly comparable to [2] and its sequel [3], extending the results by supporting an isotropic additive noise model by contrast with the 'normal-space noise' in the latter paper.

References

[1] C. Weibel, *An Introduction to Homological Algebra*, Cambridge University Press, 1994.

[2] P. Niyogi, S. Smale, and S. Weinberger, *Finding the Homology of Submanifolds with High Confidence from Random Samples*, Discrete Computational Geometry, 2008.

[3] P. Niyogi, S. Smale, and S. Weinberger, *A topological view of unsupervised learning from noisy data*, SIAM Journal on Computing, 2011.

[4] S. Lie and F. Engel, *Theorie der Transformationsgruppen I*, Leipzig, 1888.

[5] H. Poincaré, *Analysis Situs*, Journal de l'École Polytechnique, 1895.

[6] A. Hatcher, *Algebraic Topology*, http://pi.math.cornell.edu/~hatcher/AT/ATchapters.html, 2001.

[7] G. E. Hinton and R. R. Salakhutdinov, *Reducing the Dimensionality of Data with Neural Networks*, Science (Vol 313, pp. 504–507), 2006.

[8] L. J. P. van der Maaten and G. E. Hinton, *Visualizing Data Using t-SNE* Journal of Machine Learning Research (Vol 9, pp. 2579–2605), 2008.

[9] https://blog.keras.io/building-autoencoders-in-keras.html

[10] L. Vietoris, *Über den höheren Zusammenhang kompakter Räume und eine Klasse von zusammenhangstreuen Abbildungen*, Mathematische Annalen, 1927.

[11] H. Hotelling, *Analysis of a complex of statistical variables into principal components*, Journal of Educational Psychology, 1933.

[12] C. R. Genovese, M. Perone-Pacifico, I. Verdinelli, and L. Wasserman, *Manifold estimation and singular deconvolution under Hausdorff loss*, The Annals of Statistics, 2012.

[13] J. de Leeuw, *Applications of convex analysis to multidimensional scaling*, Recent Developments in Statistics, 1977.

[14] J. B. Tenenbaum, V. de Silva, and J. C. Langford, *A Global Geometric Framework for Nonlinear Dimensionality Reduction*, Science (Vol 290, pp. 2319), 2000.

[15] S. T. Roweis and L. K. Saul, *Nonlinear Dimensionality Reduction by Locally Linear Embedding*, Science (Vol 290, pp. 2323), 2000.

[16] V. de Silva and J. B. Tenenbaum, *Global versus local methods in nonlinear dimensionality reduction*, Neural Information Processing Systems, 2002.

[17] L. K. Saul and S. T. Roweis, *An Introduction to Locally Linear Embedding*, Journal of Machine Learning Research, 2001.

[18] L. Horváth and P. Kokoszka, *Inference for Functional Data with Applications*. Springer Series in Statistics, 2012.

[19] H. Whitney, *The self-intersections of a smooth n-manifold in 2n-space*, Annals of Mathematics, 1944.

[20] D. Surendran, http://people.cs.uchicago.edu/~dinoj/manifold/swissroll.html, 2004.

[21] A. Zomorodian and G. Carlsson, *Computing Persistent Homology*, 20th ACM Symposium on Computational Geometry, 2004.

[22] A. Talwalkar, S. Kumar, and H. Rowley, *Large-Scale Manifold Learning*, Computer Vision and Pattern Recognition, 2008.

[23] K. Florek, J. Łukaszewicz, J. Perkal, Hugo Steinhaus, and S. Zubrzycki, *Sur la liaison et la division des points d'un ensemble fini*. Colloquium Mathematicae 2.3-4, 1951.

[24] R. Sibson, *SLINK: an optimally efficient algorithm for the single-link cluster method*. The Computer Journal, 1973.

[25] H. Edelsbrunner, D. Letscher, and A. Zomorodian, *Topological persistence and simplification*. Discrete Computational Geometry, 2002.

[26] S. Bhattacharya and R. Ghrist, *Path Homotopy Invariants and their Application to Optimal Trajectory Planning*. Proceedings of IMA Conference on Mathematics of Robotics, 2015.

[27] M. Gidea and Y. Katz, *Topological Data Analysis of Financial Time Series: Landscapes of Crashes*. 2017.

[28] J. W. Milnor, *On the total curvature of knots*, Annals of Mathematics, 52(2), pp. 248-257, 1950.

[29] S. Bhattacharya, R. Ghrist, and V. Kumar, *Persistent Homology for Path Planning in Uncertain Environments*. IEEE Transactions on Robotics, 2015.

[30] D. Cohen-Steiner, H. Edelsbrunner, J. Harer, and Y. Mileyko, *Lipschitz Functions have L_p-stable Persistence*. Foundations of Computational Mathematics, 2010.

[31] G. Carlsson and V. de Silva, *Topological approximation by small simplicial complexes*. Computational Geometry, 2003.

[32] Vitaliy Kurlin, *A fast persistence-based segmentation of noisy 2D clouds with provable guarantees*. Pattern Recognition Letters, 2015.

[33] H. Edelsbrunner and E. P. Mücke, *Three-dimensional alpha shapes*. ACM Transactions on Graphics, 13:1, 43-72, 1994.

[34] K. Borsuk, *On the imbedding of systems of compacta in simplicial complexes*. Fund. Math. 35, 217-234, 1948.

[35] R. Ghrist, *Barcodes: the persistent topology of data*. AMS Bulletin, 2008.

[36] V. de Silva and R. Ghrist, *Coverage in sensor networks via persistent homology*. Alg. & Geom. Topology, 2006.

[37] V. de Silva, D. Morozov, and M. Vejdemo-Johansson, *Dualities in persistent (co)homology*. Inverse Problems, Volume 7, Number 12, 2011.

[38] P. McMullen, *The maximum numbers of faces of a convex polytope*. Mathematika 17, 179-184, 1970.

[39] J. D. Boissonnat, R. Dyer, A. Ghosh, N. Martynchuk, *An obstruction to Delaunay triangulations in Riemannian manifolds*, https://arxiv.org/abs/1612.02905, 2016.

[40] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer, *Stability of Persistence Diagrams*. Discrete and Computational Geometry, 2007.

[41] G. Carlsson, T. Ishkhanov, V. de Silva, A. Zomorodian, *On the local behaviour of spaces of natural images*. International Journal of Computer Vision, 2008.

[42] E. Parzen, *On the estimation of a probability density function and the mode*. Annals of Mathematical Statistics, 33:1065-1076, 1962.

[43] Y. P. Mack and M. Rosenblatt, *Multivariate k-nearest neighbor density estimates*. Journal of Multivariate Analysis, 1979.

[44] G. Carlsson and A. Zomorodian, *The Theory of Multidimensional Persistence*. Discrete and Computational Geometry, 2009.

[45] M. Lesnick, *The Theory of the Interleaving Distance on Multidimensional Persistence Modules*, Foundations of Computational Mathematics, 2015.

[46] A. Criminisi, J. Shotton, and E. Konukoglu, *Decision Forests for Classification, Regression, Density Estimation, Manifold Learning and Semi-Supervised Learning*, Microsoft Research technical report, 2011.

[47] A. N. Gorban and A. Zinovyev, *Principal Manifolds And Graphs In Practice: From Molecular Biology To Dynamical Systems*, International Journal of Neural Systems, 2010.

[48] J. Segura, *Bounds for ratios of modified Bessel functions and associated Turán-type inequalities*, Journal of Mathematical Analysis and Applications, 2010.

[49] F. W. J. Olver and L. C. Maximon, *Digital Library of Mathematical Functions*, Section 10.41(ii).

[50] John M. Lee, *Introduction to smooth manifolds*, Graduate Texts in Mathematics (218. New York, Springer. xvii, 628 p), 2002.

[51] C. A. Rogers, *A note on coverings and packings*. Journal of the London Mathematical Society, 1950.

[52] R. Forman, *A user's guide to discrete Morse theory*. Séminaire Lotharingien de Combinatoire, 2002.

[53] V. Nanda and K. Mischaikow, *Morse Theory for Filtrations and Efficient Computation of Persistent Homology*. Discrete and Computational Geometry, 2013.

[54] D. Mumford, A. Lee, K. Pedersen, *The nonlinear statistics of high-contrast patches in natural images*. International Journal of Computer Vision, 2003.

[55] B. Benedetti, F. H. Lutz, *Knots in collapsible and non-collapsible balls*. Electronic Journal of Combinatorics, 2013.

[56] C. J. Thompson, *Inequalities and partial orders on matrix spaces*. Indiana University Mathematics Journal, 1971.

[57] H. Araki, *On an inequality of Lieb and Thirring*. Letters in Mathematical Physics, 1990.

[58] X. Pennec, P. Fillard, and N. Ayache, *A Riemannian Framework for Tensor Computing*. International Journal of Computer Vision, 2006.

[59] Davide Pigoli, P. Z. Hadjipantelis, J. S. Coleman, and J. A. D. Aston, *The analysis of Acoustic Phonetic Data: exploring differences in the spoken Romance languages*. 2015.

[60] Davide Pigoli, J. A. D. Aston, I. L. Dryden, and P. Secchi, *Distances and inference for covariance operators*. Biometrika, 2014.

[61] H. Le and D. Kendall, *The Riemannian structure of Euclidean shape spaces: a novel environment for statistics*. The Annals of Statistics, 1993.

[62] P. H. Schönemann, *A generalized solution of the orthogonal Procrustes problem*. Psychometrika, 1966.

[63] Huiling Le, *Locating Fréchet Means with Application to Shape Spaces*. Advances in Applied Probability, 2001.

[64] I. L. Dryden and K. V. Mardia, *Statistical Shape Analysis*. Wiley, 1998.

[65] I. L. Dryden, A. Koloydenko, and D. Zhou, *Non-Euclidean statistics for covariance matrices, with applications to diffusion tensor imaging*. The Annals of Applied Statistics, 2009.

[66] J. B. Kruskal, *The number of simplices in a complex*. Mathematical Optimization Techniques, 1963.

[67] G. O. H. Katona, *A theorem of finite sets*. Theory of Graphs, 1968.

[68] Adam P. Goucher, https://mathoverflow.net/a/303065/39521. MathOverflow, 2018.

[69] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*. Grundlehren der mathematischen Wissenschaften 290, 1988.

[70] M. Rosenblatt, *Remarks on Some Nonparametric Estimates of a Density Function*. The Annals of Mathematical Statistics 27(3), 1956.