Non-isotropic Persistent Homology: Leveraging the Metric Dependency of PH

Vincent P. Grande
RWTH Aachen University, Germany
grande@cs.rwth-aachen.de

Michael T. Schaub.
RWTH Aachen University, Germany
schaub@cs.rwth-aachen.de

Abstract
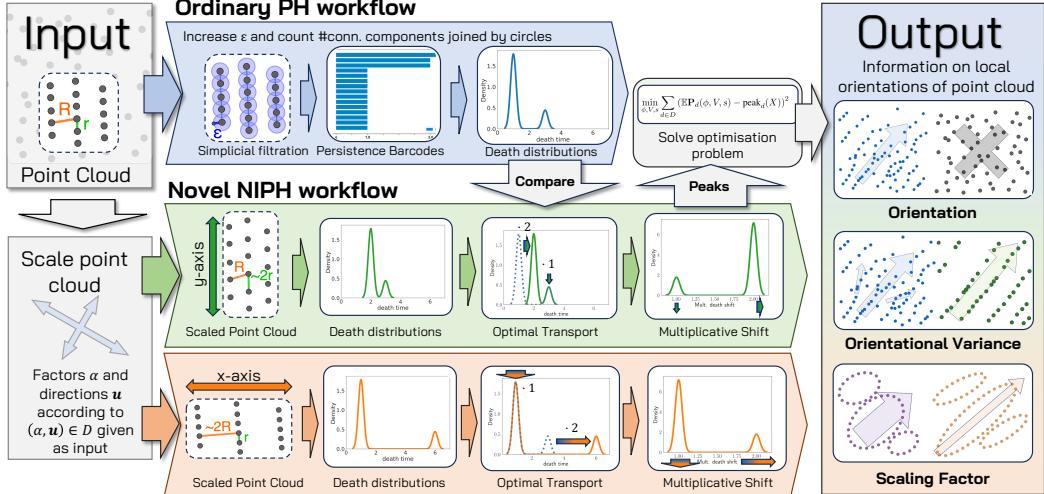
Persistent Homology is a widely used topological data analysis tool that creates a concise description of the topological properties of a point cloud based on a specified filtration. Most filtrations used for persistent homology depend (implicitly) on a chosen metric, which is typically agnostically chosen as the standard Euclidean metric on $\mathbb{R}^n$. Recent work has tried to uncover the "true" metric on the point cloud using distance-to-measure functions, in order to obtain more meaningful persistent homology results. Here we propose an alternative look at this problem: we posit that information on the point cloud is lost when restricting persistent homology to a single (correct) distance function. Instead, we show how by varying the distance function on the underlying space and analysing the corresponding shifts in the persistence diagrams, we can extract additional topological and geometrical information. Finally, we numerically show that non-isotropic persistent homology can extract information on orientation, orientational variance, and scaling of randomly generated point clouds with good accuracy and conduct some experiments on real-world data.

1 Introduction

Over the last decades, topological data analysis (TDA) has proven to provide a valuable toolkit for extracting information out of complex data sets. Most notably, persistent homology (PH) provides a straight-forward way to extract topological information across different scales from a point cloud. The resulting persistence diagrams and persistence barcodes form a metric space and have been used for many interesting applications, see for example Edelsbrunner and Harer [2008]. An important motivation for our work is to combine the robust topological descriptors PH provides of (point-cloud) data with more refined geometric notions of orientation and preferred dimensions.

In order to form the simplicial filtration used for computing PH, we need to specify a distance function on the point cloud, which is typically chosen as a metric on $\mathbb{R}^n$; most often the standard ℓ_2 metric on $\mathbb{R}^n$ is used. As in topology virtually all reasonable metrics on $\mathbb{R}^n$ are equivalent and induce the same topology, it would seem that the choice of metric for PH is irrelevant. However, this is not true: because the metric controls the birth-time of simplices, changing the metric alters the birth and death times of the topological features of the constructed simplicial complexes. A change in metric may even eradicate certain topological features or introduce new topological features. The role of the underlying metric for PH is often not further investigated, even though it has been acknowledged in recent work such as Chazal et al. [2011], Anai et al. [2020], although from a different view. However, this influence is treated by the authors as a problem that they attempt to fix by introducing an improved Euclidean distance that makes PH more robust to outliers. Stated differently, the aim is to construct a single metric that leads to the "best possible" simplicial filtration that provides an accurate topological description of the underlying space (point cloud).

In this work, we take a different perspective on the influence of the choice of a metric on PH. Rather than trying to find the best possible metric, we ask: how does PH change as we vary the metric, and
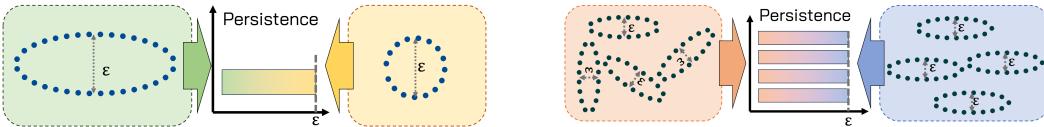
Figure 1: Schematic of Non-Isotropic Persistent Homology (NIPH). Step 1. Produce different versions of the input point cloud by applying directional scaling in direction $\mathbf{u}$ and scaling-factor s according to elements in input set D. **Step 2.** Compute PH of desired degree (the diagram displays 0th persistent homology) for input point cloud and each of the scaled point clouds. **Step 3.** Compute optimal transport between weighted death time distributions of input point cloud and each of the scaled point clouds. **Step 4.** Compute multiplicative factor of shift for each death time. Extract maxima and use optimisation problem to compute preferred orientations, scaling-factor, and orientational variance of point cloud, as seen in the illustrations on the right.

can we exploit the induced changes in PH to extract additional information from the data? We name this approach non-isotropic persistent homology (NIPH) — see Figure 1. Note that NIPH not only provides us with a means to assess how robust (sensitive) the results of PH are to changes in the metric. In addition, non-isometric persistent homology (NIPH) harvests the rich information of how the persistence diagram of a point cloud changes when we change the underlying metric, to extract new information that is not apparent from any single PH analysis (See Figure 2). This may be seen as analogous to cases in physics, where taking the derivative of one important physical quantity often yields another physical quantity of interest.

As a concrete example, the notion of a preferred orientation is perhaps the most intuitive quantity that is sensitive to, and therefore measurable by a change in the metric — certain coordinate rescalings will show virtually no effect, whereas others will strongly alter the PH diagram. Indeed, orientations and preferred direction are omnipresent in data: Animal biologists study the intricate patterns of coordinated swarm behaviour of insects, birds, and marine life [Schneider, 1995]. Soft and condensed matter physicist study the patterns and emerging orientations in molecules shifting between solid and fluid state [Marder, 2010]. Microbiologists are interested in the patterns and directions viral RNA infiltrates cells and macrophages come to the rescue [Steenblock et al., 2021]. Judging based on this prevalence of notions of orientations and direction in applications, making persistent homology sensitive to orientations presents itself to be a fruitful endeavour.

From an abstract point of view, a set of points X only carries a geometric meaning in conjunction with a metric d. The metric encodes information on the distance between individual pairs of points. Taken together, this then determines the overall geometry of the space. Importantly, the metric determines

Figure 2: Data features not captured by standard PH. *Left:* Persistent homology will not distinguish circles and ellipses by their death time ε. *Right:* Persistent homology cannot distinguish orientations and orientational variances of the data set by death times ε.

the topology of the underlying space as well. From a different perspective, this means that we can *change* the geometry of a space by *changing* the metric of the space. Making use of this simple, but powerful insight is the key idea behind NIPH.

Organisation of the paper. In Section 2, we will give an overview of some of the core concepts of NIPH, including Simplicial Complexes, the Vietoris–Rips filtration, Persistent Homology, and Optimal Transport. In Section 3, we give an overview over the NIPH algorithm. We provide an introductory example in Section 3.1, and the general method in Section 3.2, highlight further approaches using the same framework in Section 3.3, and finally give theoretical guarantees in Section 3.4. Finally, we validate the performance of NIPH in experiments on synthetic and real-world data in Section 4. The code to replicate all experiments in this paper can be found here.

Related Work. Chazal et al. [2011] introduced distance-to-measure filtrations to adapt the distance function (or the simplicial filtration) of the point cloud to reduce the effect of outliers (cf. Anai et al. [2020]). There has also been work on extracting geometric information using persistent homology, such as the persistent homology transform Turner et al. [2014]. However, this work focussed on extracting shapes of $2d$ surfaces and $3d$ objects in $3d$ space, whereas we focus on local geometric information encoded in point clouds. In Hofer et al. [2019] and Carriere et al. [2021], the authors introduce the notion of differentiating persistent homology diagrams. However, their goal is to utilize this differentiability to make PH accessible to machine learning tasks, rather than to extract geometrical information on the point cloud. In Stucki et al. [2023], the authors consider the problem of matching persistence diagrams as well, but for persistence diagrams from the same metric space.

2 Theoretical Background

Our method of Non-Isotropic Persistent Homology draws from many ideas of Algebraic Topology, Topological Data Analysis, Computational Geometry, and Transport Theory. We give a brief overview over the main concepts needed for NIPH, but direct the interested reader to the more comprehensive introductions to Algebraic Topology [Bredon et al., 1993, Hatcher, 2002], Topological Data Analysis [Chazal and Michel, 2021], Computational Geometry [Preparata and Shamos, 2012], and Optimal Transport [Peyré and Cuturi, 2019].

Simplicial Complexes. Simplicial Complexes are higher-order generalisations of graphs, originally constructed in algebraic topology to capture the shapes (homotopy types) of topological spaces.

Definition 2.1 (Simplicial Complex)**.** A simplicial complex X consists of a set of vertices V and a set of finite non-empty subsets of V, called the simplices S, such that **(i)** S is closed under taking non-empty subsets and **(ii)** for every $v \in V$ the singleton set $\{v\}$ is contained in S. We call the simplices with $k + 1$ elements the k-simplices.

A graph can be considered as a simplicial complex with the nodes being the 0-simplices and the edges being the 1-simplices, connecting two nodes. Intuitively, 2-simplices are represented by triangles between three nodes, 3-simplices are tetrahedra between four nodes, and so on.

Vietoris–Rips filtration. The Vietoris–Rips (VR) filtration is a tool from computational topology that turns a point cloud into a filtration of simplicial complexes.

Definition 2.2 (Vietoris–Rips Complex)**.** Given a point cloud X with distance function d and a parameter $r \geq 0$, the associated simplicial complex $\mathrm{VR}_r(X) = (X, S)$ is given by the vertex set X and the set of simplices
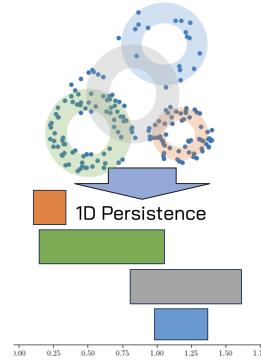
$$S = \left\{ \sigma \subset X : \max_{x,y \in \sigma} d(x, y) \leq r \right\}.$$

For $r \leq r'$, we then have the canonical inclusion $\mathrm{VR}_r(X) \subset \mathrm{VR}_{r'}(X)$. Intuitively, we can construct $\mathrm{VR}_r(X)$ by connecting all points with a distance of at most r with edges, and finally adding all available higher-order simplices (i.e. the ones where all edges between vertices are already included).

Persistent Homology. Homology is a tool from algebraic topology that describes the "shape" of a topological space. From a perspective of computational topology, homology groups have two major advantages over the homotopy groups, another popular concept of algebraic topology. Firstly,

the linear algebraic definition of homology groups makes them easily accessible for computational methods [Bauer, 2021]. Secondly, the theory of homology groups is well understood, whereas the computation of higher homotopy groups of even the most basic non-trivial topological spaces, spheres, is an active research topic today (See for example Ravenel [2023]).

Intuitively, n-dimensional homology encodes n-dimensional holes in the topological space. 0-dimensional holes are connected components, 1-dimensional holes are loops, 2-dimensional holes are cavities in $3d$-space and so on. Given a point cloud X, persistent homology keeps track of these homological features across all stages of the associated VR filtrations. Persistent homology summarises these findings in a persistence diagram that, for each connected component/loop/cavity/etc./ F, stores r and r' such that the generalised loop F first occurred in the step $\text{VR}_r(X)$ of the VR filtration and last occurred in $\text{VR}_{r'}(X)$. (Note: for 0-dimensional features, the birth time r is always 0.) For loops for instance, the birth time marks the time the loop is closed, whereas the death time r' marks the filtration step where the loop is filled by higher-order simplices. We use the implementation by The GUDHI Project [2015].

Optimal Transport. At its core, optimal transport is an optimisation problem. Intuitively, the most basic ("earth movers") formulation asks for the "cheapest" way to move one arrangement of mass to another, desired arrangement. More formally we have the following:

Definition 2.3 (Optimal transport). Given two measurable spaces A and B, a distribution μ on A, a distribution ν on B, and a joint cost function $c\colon A \times B \to \mathbb{R}_{\geq 0}$, optimal transport looks for a joint distribution π on $A \times B$ with marginal distributions μ and ν that minimises

$$C(\pi) = \int_{A \times B} c(a,b)d\pi + r(\pi)$$

over the space of all such distributions where r is some regularisation function (often an entropic regularisation). π is then called the *transport plan* and $C(\pi)$ the *transport distance*.

Optimal transport has seen widespread application in the wider ML community: Rubner et al. [2000] in Computer Vision, Kolouri et al. [2020] in graph learning, Kandasamy et al. [2018] for neural architectures, Gramfort et al. [2015] in neuroimaging, etc. In TDA, optimal transport has long been used to compare homological information across spaces. This has been done using the ℓ_∞ bottleneck distance Kerber et al. [2017] on the level of persistence diagrams, or more directly on the level of persistence landscapes Bubenik et al. [2015].

3 Methods

3.1 Introductory Example
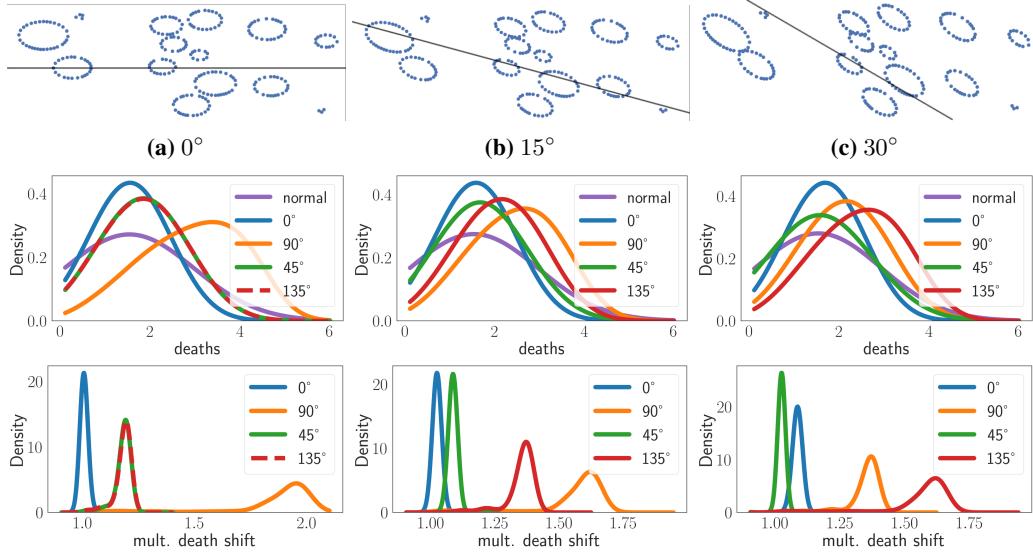
As an introductory example, we can consider the following class of metrics on $\mathbb{R}^2$:

Definition 3.1. For real $\alpha, \beta > 0$, we define the associated metric $d_{\alpha,\beta}$ on $\mathbb{R}^2$ as follows:

$$d_{\alpha,\beta}\colon \left((x_1,y_1)^\top, (x_2,y_2)^\top\right) \mapsto \sqrt{\alpha\left(x_1 - x_2\right)^2 + \beta\left(y_1 - y_2\right)^2},$$

which is the standard Euclidean distance for $\alpha = \beta = 1$. Picking α and β amounts to (de-)prioritising the x- and y-axis for our metric. When choosing $\alpha \gg \beta$, the distance between two points is almost entirely determined by the distance of their x-values. When choosing $\beta \gg \alpha$, the distance between two points is almost entirely determined by the distance of their y-values, and differences in the x-coordinate have almost no influence.

To build some intuition, assume that we compute ordinary 1-dimensional PH with $\alpha = \beta = 1$ first and compare it with PH associated to $d_{\alpha,\beta}$ for $\alpha = 0.5$ and $\beta = 1$. On a data set consisting of points sampled from circles with radius $\varepsilon/2$, the death times would remain almost constant: although points are now farther apart in the x-axis, simplices with edge length of ε are enough to cover all circles in the y-direction. On the other hand, if the data set consisted of ellipses stretched by a factor of 2 along the y-axis the persistent homology would change with the change in metric. Because the change of the metric is orthogonal to the orientation of the ellipses, their death times will double (Figure 2).

Figure 3: Illustrative Experiments. Point cloud with ellipses of various radii in orientation $\varphi = 0°$, $15°$ and $30°$, scaling factor $s = 2$ and orientational variance $V = 0$. *Top:* Part of point cloud. *Middle:* Death density diagram of 1-dimensional persistent homology. *Bottom:* Multiplicative shift diagram used to extract information on orientation, orientational variance, and scaling. In the left diagram, the blue curve represents scaling parallel to the directional scaling of the ellipses. Thus there is no change in the death times and the peak of the curve is at ~ 1.0. The orange curve represents scaling in a direction almost orthogonal to the scaling of the data points, hence the death times are multiplied by the factor of the scaling. This is represented by the peak almost reaching 2.0. The red and green curve represent scaling roughly at a direction of $45°$ to the original scaling in the point cloud. When we change the orientation of the point cloud the orange peak will move to the left and the the red peak will move to the right. The mult. death shift diagrams are easy to interpret and concise.

This simple example motivates why looking at the change of PH under different metrics can be a powerful tool to enrich the standard tools of persistent homology with notions of orientation and preferred directions. We give a more rigorous account of our mathematical models below.
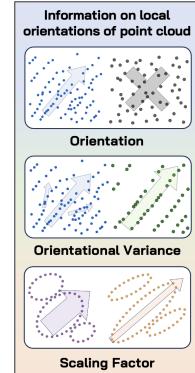
3.2 General Method

NIPH takes as input a point cloud $X \in \mathbb{R}^n$ and a set of scaling directions (see Algorithm 1), and computes a mult. shift diagram, from which a variety of interesting properties can be computed. For instance, NIPH can extract from this the presence or absence of a **preferred orientation**, **orientational variance**, and **scaling factor** (as shown on the *right*):

Orientation. Two point clouds with similar 0-dimensional PH. *Left:* The point cloud has additional structure: a preferred orientation. *Right:* There is no preferred orientation in the point cloud. NIPH can detect the differences in the amount of orientation in both point clouds across arbitrary scales.

Orientational variance. Two point clouds with similarly structured 0-dimensional PH. Both have additional structure and a preferred dimension. *Left:* There is a comparably large variance in the local preferred direction. *Right:* The alignment of the point data with the preferred direction is very strong. NIPH can detect the differences in the amount of variance of orientation in both point clouds.

Scaling factor. Two point clouds with similarly structured 1-dimensional PH. Both have additional structure and are scaled along the same direction. The ellipses of the *left* point cloud are scaled differently than on the *right*.

The distance function. For every direction of an n-dimensional space given by a unit vector $\mathbf{u}$ we can pick an orthonormal basis with first basis vector $\mathbf{u}$. Let U be the associated

base transformation matrix. For a scaling-factor $\alpha > 0$ we can define the scaling function $S_{\mathbf{u},\alpha} \colon x \mapsto U^{-1} \operatorname{diag}((\alpha, 1, \ldots, 1)) U x$.

Definition 3.2. We compute the distance function $d_{\mathbf{u},\alpha}$ associated to the pair $(\mathbf{u}, \alpha)$ as follows:

$$d_{\mathbf{u},\alpha} \colon (x_1, x_2) \mapsto \| S_{\mathbf{u},\alpha}(x_1) - S_{\mathbf{u},\alpha}(x_2) \|_2.$$

We note that this definition generalises definition 3.1 for $\beta = 1$ given in the introduction and does not depend on the concrete choice of U. Intuitively, $d_{\mathbf{u},\alpha}$ scales the distances in the direction of $\mathbf{u}$ by a factor of α while leaving the orthogonal directions untouched. Setting $\alpha = 0$ would amount to the distance obtained by projecting on the subspace orthogonal to $\mathbf{u}$, which would possibly violate the $d(x,y) = 0 \Leftrightarrow x = y$ condition on a metric. This is why we have excluded $\alpha = 0$.

Definition 3.3. Let X be a point cloud in n-dimensional space $X \subset \mathbb{R}^n$, $\mathbf{u} \in \mathbb{R}^n$ a unit vector and $\alpha > 0$ a real scaling-factor. We denote by $X_{\mathbf{u},\alpha}$ the scaled data-set $X_{\mathbf{u},\alpha} := \{ S_{\mathbf{u},\alpha}(x) \mid x \in X \}$.

It is easy to see that $(X_{\mathbf{u},\alpha}, d)$ is isometric to $(X, d_{\mathbf{u},\alpha})$ where d is the Euclidean distance on $\mathbb{R}^n$. With this observation in mind, we refer to the metric spaces $(X, d_{\mathbf{u},\alpha})$ as the scaled point clouds.

Computing persistence diagrams. We compute the persistence diagrams $\mathbf{P}$ and $\mathbf{P}_{\mathbf{u},\alpha}$ on the point clouds X and $X_{\mathbf{u},\alpha} := (X, d_{\mathbf{u},\alpha})$ for all $(\mathbf{u}, \alpha) \in D$ using a Vietoris-Rips filtration on X and $X_{\mathbf{u},\alpha}$.

Death distributions. In many cases, the death of homology classes carries more structural information than their birth time. This is because birth time is mainly controlled by point densities, whereas death times measure distances between clusters or the sizes of loops. Hence we transform the persistence diagrams to density plots over the death time. In the case of 1-dimensional homology, we can weigh the points according to the difference or quotient of death and birth time, i.e., longer lived loops carry more weight. We denote by $\mathbf{D}$ and $\mathbf{D}^{(\mathbf{u},\alpha)}$ the (weighted) vectors of death times. E.g., for a grid with grid length a the 0-dim death distribution $\mathbf{D}$ would just contain entries of a.

Computing the multiplicative shift. We are interested in how persistent homology and the death distributions change when changing the metric of the underlying data set. Hence we propose to compute an optimal transport based matching between the death distribution diagrams of X and of the $X_{\mathbf{u},\alpha}$. We can then compute for every point in the original persistence diagram the factor by which the death time was multiplied in the optimal transport matching. We can then present these multiplicative death shift factors as another density plot (See Figure 3):

Definition 3.4 (Multiplicative Death Shift). Let $T^{(\mathbf{u},\alpha)}$ denote the solution transport matrix of optimal transport between $\mathbf{D}$ and $\mathbf{D}^{(\mathbf{u},\alpha)}$. We denote by the i-th multiplicative death shift $\mathbf{ms}_i^{\mathbf{u},\alpha}$ associated to scaling $\mathbf{u}, \alpha$ for $1 \leq i \leq |\mathbf{P}|$ the following expression

$$\mathbf{ms}_i^{\mathbf{u},\alpha}(X) := \exp \left(\sum_j T_{i,j}^{\mathbf{u},\alpha}(X) \ln \mathbf{D}_j^{\mathbf{u},\alpha}(X) / \mathbf{D}_i(X) \right).$$

In the most simple case of an $1:1$ transport plan T, where $t^{\mathbf{u},\alpha}(i)$ denotes the destination of i, this reduces to the much simpler form of

$$\mathbf{ms}_i^{\mathbf{u},\alpha}(X) = \mathbf{D}_{t^{\mathbf{u},\alpha}(i)}^{\mathbf{u},\alpha} / \mathbf{D}_i(X).$$

The multiplicative shift is a scale-free measure of how much the persistence of a homology class is affected by the scaling of the point cloud. Now let w denote a vector of weights for the homological features of X in PH. Then we denote by $\mathbf{MS}_{\mathbf{u},\alpha}$ the multiplicative death shift diagram, given by the density diagram

$$\mathbf{MS}_{\mathbf{u},\alpha}(X) := \operatorname{density} \left\{ (\mathbf{ms}_i^{\mathbf{u},\alpha}(X), w_i) \mid 1 \leq i \leq |\mathbf{P}| \right\}.$$

$\mathbf{MS}_{\mathbf{u},\alpha}(X)$ is basically the combination of all the individual $\mathbf{ms}_i^{\mathbf{u},\alpha}(X)$. We furthermore denote the x-value of the maximum of $\mathbf{MS}_{\mathbf{u},\alpha}(X)$ by $\operatorname{peak}_{\mathbf{u},\alpha}(X)$, the most prominent mult. shift value.

Extracting the orientations. The multiplicative death shift density diagrams obtained in the previous step are an interesting object in their own right. To illustrate the utility of these diagrams, we give an example of extracting information on orientations (φ), orientational variance V, and scaling factor s from these shift diagrams. The idea behind this is that we can compute how the multiplicative shift diagrams should behave under given parameters (φ, V, α). We do this using the model of rectangles (closely matching the behaviour of ellipses, but having a concise analytical formulation) for 1-dimensional homology and grids for 0-dimensional homology.

Algorithm 1 Non-isotropic persistent homology (NIPH)

Input: Point cloud X, list of directions and scale-factors D
Compute persistent homology diagram $\mathbf{P}$ of X
for $d \in D$ **do**
    Compute scaled data set X_d according to $d = (\mathbf{u}, \alpha)$.
    Compute persistent homology $\mathbf{P}_d$ of X_d.
    Solve optimal transport from $\mathbf{P}$ to $\mathbf{P}_d$.
    Compute mult. shift diagram $\mathbf{MS}_d$ associated to d
end for
Post-Processing, e.g. solve optimisation problem for best orientation φ, scaling-factor s, and orientational variance V matching the maxima of $\mathbf{MS}_d$.
    Output: φ, s, V

Definition 3.5 (Expected Peak). Let $\mathbb{EP}_{\mathbf{u},\alpha}(\varphi, V, s)$ be a function such that for a point cloud Y sampled from rectangles (a grid) with scaling factor s, orientation φ and orientational variance V we have that $\mathbf{MS}_{\mathbf{u},\alpha}(Y)$ in 1-dimensional (0-dimensional) persistent homology takes it maximum value at $\mathbb{EP}_{\mathbf{u},\alpha}(\varphi, V, s)$.

$\mathbb{EP}$ has an analytic description depending on the considered homology dimension, but can be determined via sampling as well (See Appendix D.1). For D being the set of sampling directions and scalings, NIPH solves the optimisation problem

$$\min_{\varphi, V, s} \sum_{d \in D} \left(\mathbb{EP}_d(\varphi, V, s) - \mathrm{peak}_d(X) \right)^2$$

to obtain an estimated orientation φ, orientational variance V, and scaling factor α of the point cloud X explaining the witnessed multiplicative shift diagrams best. We provide an in-depth discussion of choices of sampling directions and sampling scaling factors $\mathbf{u}$ in practice in Appendix E.

3.3 Additional approaches

The underlying idea of NIPH — to leverage the metric dependency of persistent homology — allows for great flexibility and extensions beyond orientations, orientational variance and scaling. We highlight two such approaches to extract information from point clouds using the NIPH framework.

Exotic metrics and outlier detection. The family of metrics we considered so far was based on re-weighing the influence of different directions of our ambient space $\mathbb{R}^d$. However, we can construct more intricate families of metrics for specific tasks.

Definition 3.6 (Outlier metric, cf. Anai et al. [2020]). Let $X \subset \mathbb{R}^d$ be a finite point cloud and $\mu \colon X \to \mathbb{R}_{>0}$ a function. We then define a metric d_μ on X by $d_\mu(x, y) = 2d(x, y)/(\mu(x) + \mu(y))$.

In the previous definition, for a $\delta > 0$, we can define f to be $f_\delta(x) = \sum_{y \neq x, y \in X} \exp(-d(x, y)^2/\delta)$ and then set $\mu_\delta(x) = f_\delta(x) \cdot |X|/ \sum_{y \in X} f_\delta(y)$. For $\delta \to 0$, this returns the original metric, whereas increasing δ will increasingly make outliers further apart and decrease the distance in dense parts of the point cloud. Adopting this metric and tracking the persistence and multiplicative shift diagrams gives insight into the outlier structure of the base point cloud X.

Tracking the orientation of death simplices. Persistent Homology equips us with yet another tool: For every persistence class, we also get the death simplex, i.e. the simplex that connects the connected components, closes the loop, etc. We can then look at the orientation φ_σ of the longest edge of the death simplex σ. By considering the entire distribution of orientations, we can again infer information on the geometry of the point cloud. However, limiting this analysis to the snapshot of the base point cloud with standard metric is prone to noise: Even a circle in one-dimensional homology, despite having no preferred orientation, will have a death edge with an orientation. Hence we propose to apply the NIPH framework to the analysis of death simplices: By tracking the distributions of death edge orientations over different metrics on the point cloud, we can distinguish noise from real orientations, and obtain the underlying scaling factor of the orientations for free.

3.4 Theoretical Guarantees

Theorem 3.7 (Theoretical guarantees)**.** *Given points on an orthogonal $n_1 \times n_2$ grid with $n_1, n_2 > 1$ in $\mathbb{R}^2$ rotated by φ with distances d_1 and d_2 with $d_1 < d_2$. Applying 0-dimensional unweighted NIPH with scaling factor $s \leq d_2/d_1$ in direction ψ will then yield a multiplicative shift diagram with a peak at $s_1 := \sqrt{(s^2-1)\cos^2(\psi - \varphi) + 1}$ with weight $(n_1 - 1)n_2$ and a second peak at $s_2 := \sqrt{(s^2-1)\sin^2(\psi - \varphi) + 1}$ with weight $n_2 - 1$. In particular, when scaling in direction φ the maximum will appear at a value of s, while scaling in an orthogonal direction will produce a maximum at value 1.*

Remark 3.8. This theorem explains how we can use the NIPH framework to extract orientations of point clouds: By probing different scaling directions and the associated mult. shift diagrams, we can look for the diagrams with the largest x-value of the peak. This will then approximate the orientation of the point cloud. Guarantees in the setting with noise and the proofs can be found in Appendix A.

4 Experiments

Orientation, orientational variance, and scaling factor. We ran experiments on synthetic data to verify that NIPH can indeed infer information on orientation, scaling, and variance inside a data set. Given an angle φ, an orientational variance V, and a scaling factor s, we created a point cloud resembling ellipses or rectangles with sizes between 0.2 and 2. We scaled the different axes of the ellipses/rectangles with factor s, oriented them in the general direction of φ and added independent Gaussian noise with variance V and mean 0 to each orientation of the individual ellipses/rectangles.

We then used the NIPH pipeline to infer the parameters of the direction of orientation φ, the orientational variance V, and the scaling factor s from the point cloud data. We report the prediction accuracy of NIPH in Figure 4 and Figure 5. We note that even with high orientational variance in the synthetic data, NIPH is able to predict orientation high accuracy, and obtains scaling and orientational variance with medium accuracy. PCA cannot infer any of these parameter, see Appendix B.
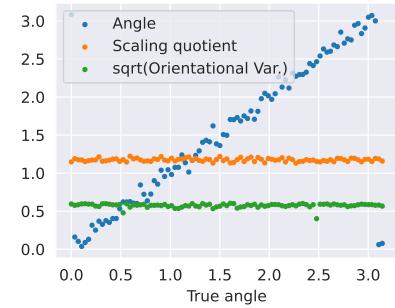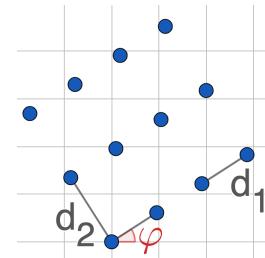
Road networks. We used 0-dimensional NIPH on road network data from Manhattan, New York City, U.S.A., and Tübingen, Germany. We first extracted the road networks from openstreetmap data (OpenStreetMap contributors [2023]). We then created a point cloud by choosing random locations of 130000 cars (NYC) and 43000 cars (Tübingen) on the respective road networks. We then ran the NIPH-pipeline for 0-dimensional homology on the point clouds restricted to specific road types. This allows us to extract information on the predominant orientations in the point cloud, see Figure 6. The point clouds generated from the Manhattan road network have a very strong orientation, whereas the point clouds resembling the Tübingen road network have virtually no preferred orientation.

Additionally, we used NIPH on 15 U.S. and European cities to extract the orientational strength of point clouds sampled on the road network. On average, U.S. cities have a higher strength of orientation than their European counterparts (Figure 7 *Right*). This experiment demonstrates that

$\sqrt{V} = \text{std}(\varphi)$	$\sqrt{\text{MSE}}\ \varphi$	$\sqrt{\text{MSE}}$ scaling	$\sqrt{\text{MSE}}\ \sqrt{V}$
$0°$	$0.3°$	0.09	0.26
$5°$	$0.8°$	0.11	0.18
$10°$	$1.6°$	0.15	0.12
$15°$	$2.3°$	0.18	0.08
$20°$	$3.0°$	0.19	0.07
$25°$	$4.0°$	0.19	0.07
$30°$	$5.1°$	0.19	0.08
$35°$	$6.2°$	0.20	0.08
$40°$	$8.0°$	0.20	0.08
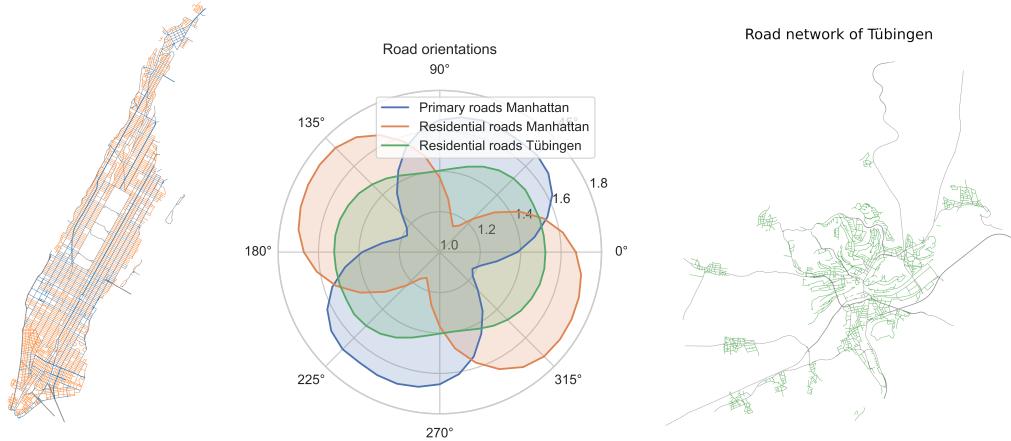
Figure 4: Quantitative performance of NIPH. We have run NIPH on a point cloud sampled from 200 oriented rectangles with different orientational variances and $s = 2$. We show the root of the mean squared error of the predictionsof NIPH depending on the orientational variance of X.
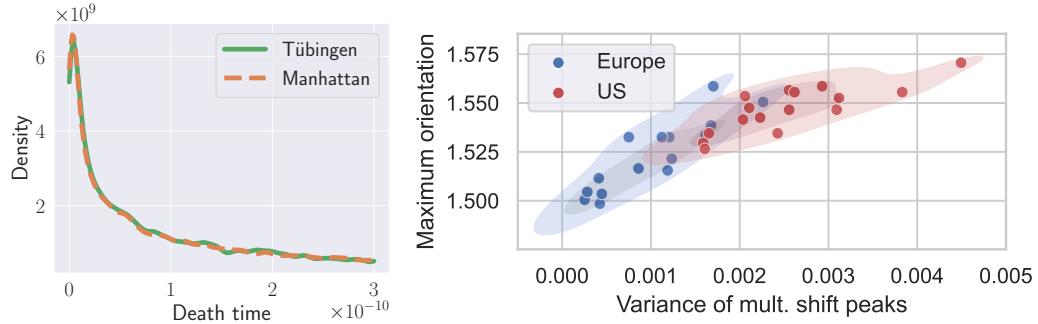
Figure 5: Predictions of NIPH in high orientational variance setting. x-axis: True φ. *Blue:* φ. *Green:* $\sqrt{V}$. (True value: 0.5). *Orange:* Scaling s. (True value: 1.5).

Figure 6: *Left and right:* Road networks of Manhattan and Tübingen. Different colours denote different type of roads. *Middle:* Strength of orientation along different directions of different road types in Manhattan and Tübingen. Residential roads ("streets") and primary roads ("avenues") have a very strong orientation in orthogonal directions. Residential roads in Tübingen have a very faint bias of an east–west orientation. The strength of orientation is measured by the x-value of the peaks in the corresponding mult. shift diagrams.

Figure 7: *Left:* 0-dimensional PH density plot corr. to Figure 6. Virtually no difference exists between the data sets exists, underscoring the relevance of NIPH. *Right:* 2 measures of orientation strength of points clouds sampled on residential roads in 15 small European and U.S. cities. While outliers exist, the grid-like structure of U.S. cities results in a higher strength of orientation in the point clouds.

NIPH is a powerful tool to extract notions of orientation and other information inaccessible to ordinary PH. The 0th PH density diagrams are virtually identical for Tübingen and Manhattan (Figure 7 *Left*).

5   Discussion

Limitations and future work.   Because persistent homology in higher dimensions is computationally expensive, for very large point clouds, only 0-dimensional NIPH is feasible. However, landmark sampling or down-sampling can often mitigate this problem. In certain cases, especially for inhomogeneous data sets, the optimal transport will produce an incorrect matching between the persistent homology classes, falsifying the results of the algorithm (See Appendix C for an in-depth analysis). Future work will consider other methods to match homology classes, more deeply rooted in the algebraic machinery behind PH, including tools like vineyards [Cohen-Steiner et al., 2006]. Furthermore, using the orientations of the death simplices (Cf. Section 3.3) does not even require the computation of a matching. The flexibility of the NIPH framework allows for a whole avenue of research using different hand-crafted families of metrics for exciting new applications.

Conclusion.   We have introduced NIPH, a novel method building on top of persistent homology. NIPH extracts additional topological and geometrical information by varying the distance function on the underlying space and analysing the corresponding shifts in the persistence diagrams. We have verified the performance of NIPH on a synthetic data set and on real-world road network data.

Acknowledgments

References

Herbert Edelsbrunner and John Harer. Persistent homology—a survey. *Contemporary mathematics*, 453(26):257–282, 2008. 1

Frédéric Chazal, David Cohen-Steiner, and Quentin Mérigot. Geometric inference for probability measures. *Foundations of Computational Mathematics*, 11:733–751, 2011. 1, 3

Hirokazu Anai, Frédéric Chazal, Marc Glisse, Yuichi Ike, Hiroya Inakoshi, Raphaël Tinarrage, and Yuhei Umeda. Dtm-based filtrations. In *Topological Data Analysis: The Abel Symposium 2018*, pages 33–66. Springer, 2020. 1, 3, 7

SS Schneider. Swarm movement patterns inferred from waggle dance activity of the neotropical african honey bee in costa rica. *Apidologie*, 26(5):395–406, 1995. 2

Michael P Marder. *Condensed matter physics*. John Wiley & Sons, 2010. 2

Charlotte Steenblock, Stefanie Richter, Ilona Berger, Marko Barovic, Janine Schmid, Undine Schubert, Natalia Jarzebska, Anne von Mässenhausen, Andreas Linkermann, Annette Schürmann, et al. Viral infiltration of pancreatic islets in patients with covid-19. *Nature communications*, 12(1):3534, 2021. 2

Katharine Turner, Sayan Mukherjee, and Doug M Boyer. Persistent homology transform for modeling shapes and surfaces. *Information and Inference: A Journal of the IMA*, 3(4):310–344, 2014. 3

Christoph Hofer, Roland Kwitt, Marc Niethammer, and Mandar Dixit. Connectivity-optimized representation learning via persistent homology. In *International Conference on Machine Learning*, pages 2751–2760. PMLR, 2019. 3

Mathieu Carriere, Frédéric Chazal, Marc Glisse, Yuichi Ike, Hariprasad Kannan, and Yuhei Umeda. Optimizing persistent homology based functions. In *International conference on machine learning*, pages 1294–1303. PMLR, 2021. 3

Nico Stucki, Johannes C Paetzold, Suprosanna Shit, Bjoern Menze, and Ulrich Bauer. Topologically faithful image segmentation via induced matching of persistence barcodes. In *International Conference on Machine Learning*, pages 32698–32727. PMLR, 2023. 3

G.E. Bredon, J.H. Ewing, F.W. Gehring, and P.R. Halmos. *Topology and Geometry*. Graduate Texts in Mathematics. Springer, New York, 1993. 3

Allen Hatcher. *Algebraic Topology*. Cambridge University Press, Cambridge, 2002. 3

Frédéric Chazal and Bertrand Michel. An introduction to topological data analysis: fundamental and practical aspects for data scientists. *Frontiers in artificial intelligence*, 4:108, 2021. 3

Franco Preparata and Michael Shamos. *Computational geometry: an introduction*. Springer Science & Business Media, 2012. 3

Gabriel Peyré and Marco Cuturi. Computational optimal transport: With applications to data science. *Foundations and Trends® in Machine Learning*, 11(5-6):355–607, 2019. 3

Ulrich Bauer. Ripser: efficient computation of vietoris–rips persistence barcodes. *Journal of Applied and Computational Topology*, 5(3):391–423, 2021. 4

Douglas C Ravenel. *Complex cobordism and stable homotopy groups of spheres*, volume 347. American Mathematical Society, 2023. 4

The GUDHI Project. *GUDHI User and Reference Manual*. GUDHI Editorial Board, 2015. 4

Yossi Rubner, Carlo Tomasi, and Leonidas J Guibas. The earth mover's distance as a metric for image retrieval. *International journal of computer vision*, 40:99–121, 2000. 4

Soheil Kolouri, Navid Naderializadeh, Gustavo K Rohde, and Heiko Hoffmann. Wasserstein embedding for graph learning. *arXiv preprint arXiv:2006.09430*, 2020. 4

Kirthevasan Kandasamy, Willie Neiswanger, Jeff Schneider, Barnabas Poczos, and Eric P Xing. Neural architecture search with bayesian optimisation and optimal transport. *Advances in neural information processing systems*, 31, 2018. 4

Alexandre Gramfort, Gabriel Peyré, and Marco Cuturi. Fast optimal transport averaging of neuroimaging data. In *Information Processing in Medical Imaging: 24th International Conference, IPMI 2015, Sabhal Mor Ostaig, Isle of Skye, UK, June 28-July 3, 2015, Proceedings 24*, pages 261–272. Springer, 2015. 4

Michael Kerber, Dmitriy Morozov, and Arnur Nigmetov. Geometry helps to compare persistence diagrams, 2017. 4

Peter Bubenik et al. Statistical topological data analysis using persistence landscapes. *J. Mach. Learn. Res.*, 16(1):77–102, 2015. 4

OpenStreetMap contributors. Planet dump retrieved from https://planet.osm.org. https://www.openstreetmap.org, 2023. 8

David Cohen-Steiner, Herbert Edelsbrunner, and Dmitriy Morozov. Vines and vineyards by updating persistence in linear time. In *Proceedings of the twenty-second annual symposium on Computational geometry*, pages 119–126, 2006. 9

A   Proofs and Theoretical Guarantees for Noisy Grids

We will now give a proof of Theorem 3.7 and Theorem A.1 proving guarantees for noisy grids:

Theorem 3.7 (Theoretical guarantees). *Given points on an orthogonal $n_1 \times n_2$ grid with $n_1, n_2 > 1$ in $\mathbb{R}^2$ rotated by φ with distances d_1 and d_2 with $d_1 < d_2$. Applying 0-dimensional unweighted NIPH with scaling factor $s \leq d_2/d_1$ in direction ψ will then yield a multiplicative shift diagram with a peak at $s_1 := \sqrt{(s^2-1)\cos^2(\psi-\varphi)+1}$ with weight $(n_1-1)n_2$ and a second peak at $s_2 := \sqrt{(s^2-1)\sin^2(\psi-\varphi)+1}$ with weight $n_2 - 1$. In particular, when scaling in direction φ the maximum will appear at a value of s, while scaling in an orthogonal direction will produce a maximum at value 1.*

Proof. The NIPH pipeline consists of multiple steps: First we will compute the persistent 0-homology of the base space. Because of the grid structure, we will obtain $(n_1-1)n_2$ homology classes with a death time of d_1, $n_2 - 1$ classes with a death time of d_2 and 1 class with death time ∞. Now we consider the scaled point cloud.: After rotating by $-\psi$, some geometric calculations reveal a scaling factor of

$$s_1 := \sqrt{s^2\cos^2(\psi-\varphi)+\sin^2(\psi-\varphi)} = \sqrt{(s^2-1)\cos^2(\psi-\varphi)+1} \leq s \leq d_2/d_1$$

in the d_1 direction and of $s_2 := \sqrt{(s^2-1)\sin^2(\psi-\varphi)+1} \geq 1$ in the d_2 direction. We notice that because of our choice of s, we still have that $s_1 d_1 \leq s_2 d_2$. Thus, the persistence diagram of the scaled point cloud consists of $(n_1-1)n_2$ classes with a death time of $s_1 d_1$, $n_2 - 1$ classes with a death time of $s_2 d_2$, and one class with death time ∞. Now, unweighted optimal transport in the $1d$ setting has a straight-forward solution inducing a $1:1$ mapping of the $(n_1-1)n_2$ points from d_1 to $s_1 d_1$ with a multiplicative shift of s_1 and a $1:1$ mapping of the $n_2 - 1$ points from d_2 to $s_2 d_2$ with a mult. shift of s_2. This then induces the multiplicative shift diagram as described in the theorem. $\square$

Theorem A.1 (Guarantees for noisy grids). *Assume we are in the situation of Theorem 3.7 but with an additive random noise ε with $|\varepsilon| < \delta$ added to every point independently, with $d_1 + \delta < d_2 - \delta$ and $s \leq (d_2 - 2\delta)/(d_1 + 2\delta)$. The mult. shift diagram will then have a mass of $(n_1-1)n_2$ in the interval $[(s_1 d_1 - 2s\delta)/(d_1 + 2\delta), (s_1 d_1 + 2s\delta)/(d_1 - 2\delta)]$.*
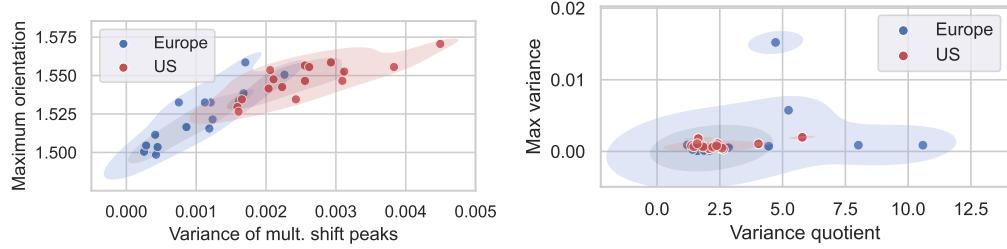
Proof. The proof of the theorem in the noisy case works essentially the same as in the noise-free case. The important observation is that before the shift, both points vary at most by a distance of δ from their original grid position, whereas after the shift, both points vary at most by a distance of $s\delta$. Dividing these bounds yields a new approximation of the multiplicative shift. $\square$
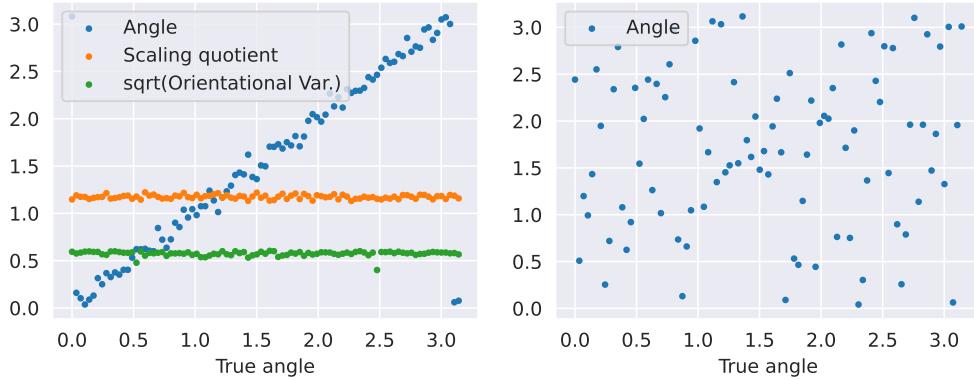
Figure 8: *Top:* Road networks of Manhattan and Tübingen. Different colours denote different type of roads. Red rectangle denotes borders of smaller dataset *Middle:* Strength of orientation along different directions of different road types in Manhattan and Tübingen. Residential roads ("streets") and primary roads ("avenues") have a very strong orientation in orthogonal directions. *Left:* Entire dataset, *Right:* Dataset restricted to rectangle. Note that NIPHcan extract the same orientations on both datasets. *Bottom:* Angles of principal directions obtained by PCA. Although the local structure is unchanged when restricting to the rectangle, the resulting PCA direction is entirely different. Note furthermore PCA estimates primary and residential roads in Manhattan to have the same orientation which is entirely false.

B   Comparison to Principal Component Analysis

Principal component analysis (PCA) is a standard tool in machine learning and data science to determine the principal directions in high-dimensional point clouds. However, PCA only considers the global shape of the data, whereas NIPH is able to extract the orientation based on local structures, which is important in many real-world applications. When using PCA on the data sets introduced in Section 4, we noticed that the principle direction inferred by PCA depends on the shape of the

Figure 9: Orientational strength of road grids in U.S. and European cities extracted by *Left:* NIPH and *Right:* PCA.
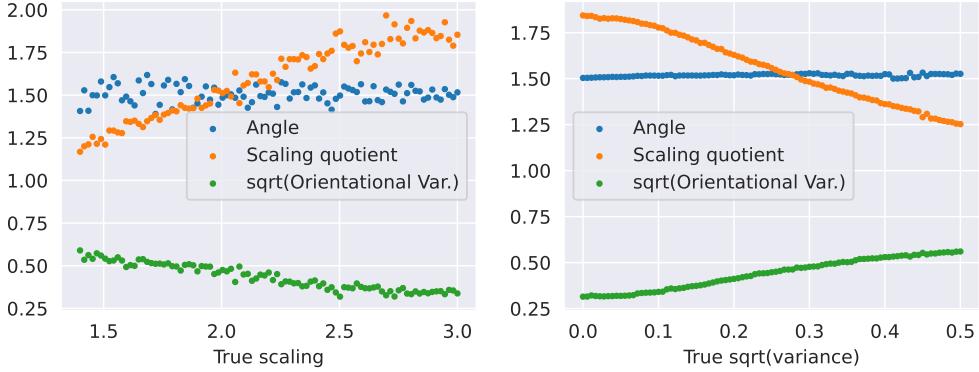
Figure 10: Trying to recover the orientation of synthetic data sets. *x-axis:* True orientation. *Blue on the y-axis:* Inferred orientation. *Left:* Results obtained with NIPH, *Right:* Results obtained with PCA. It is clear that NIPH outperforms PCA by a wide margin.

considered subset/cutout of the data and not on the local structures in the data. In Figure 8 we show that PCA does not infer correct orientations in the Manhatten/Tübingen road networks. In Figure 9 we have tried to replicate the results of Figure 7 using PCA. We try to determine the orientational variance by the maximum variance of the data set in any direction (y-axis) and the quotient of the variance of the data in the first and second principal direction. As shown in the diagram, PCA fails to capture meaningful information. In Figure 10 we have compared the effectiveness of NIPH with the effectiveness of PCA in uncovering hidden 1-dimensional orientations on synthetic datasets. NIPH clearly outperforms PCA.
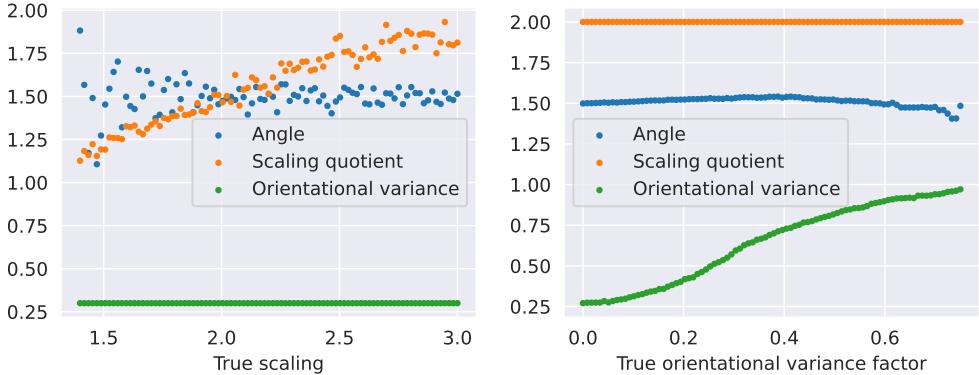
C   Limitations

Across all our experiments, NIPH was able to extract the underlying orientation of the point cloud with high accuracy. However, in certain parameter regimes, NIPH systematically underestimated the scaling factor and overestimated the orientational variance, see Figure 11. However, because of the systematic nature of the error, this may be mitigated by training an ML model on the NIPH outputs to infer the true parameters. We claim that there are two separate effects that likely contribute to these results:

i) Firstly, both increasing the orientational variance and decreasing the scaling factor results in decreasing the x-value of the peaks of the multiplicative death shift density diagrams sampled in the direction of the primary orientation of the data set. For the scaling factor, this is simply the case because a small scaling factor limits the maximum possible multiplicative shift. For the orientational variance, this is the case because a higher orientational variance means that fewer of the ellipses perfectly align with the sampling direction, thus decreasing the maximal possible multiplicative shift. Although low scaling factor and high orientational variance induce different behaviour in the other peaks, the effects on the peaks with the largest x-values are the most pronounced and thus are the main contributors to the parameter estimation of NIPH.

Figure 11: *Left:* Estimations of NIPH for fixed orientation angle $\phi = 1.5$, orientational variance $\sqrt{V} = 0.2$ and varying scaling factor. *Right:* Estimation of NIPH for fixed orientation angle $\phi = 1.5$, scaling factor $s = 2$ and varying orientational variance. NIPH systematically overestimates the orientational variance and underestimates the scaling factor. This effect is more pronounced for high orientational variances and low scaling factors.

Figure 12: Performance of NIPH when being told one correct parameter. *Left:* Estimations of NIPH for fixed orientation angle $\phi = 1.5$, fixed *known* orientational variance $\sqrt{V} = 0.3$ and varying scaling factor. *Right:* Estimation of NIPH for fixed orientation angle $\phi = 1.5$, *known* scaling factor $s = 2$ and varying orientational variance.

ii) The second effect leading to this over- and underestimation is the noise introduced to the optimal transport maps due to the varying size of the ellipses, with the smaller principal axis ranging from a length of 0.2 to 2. Because the cost of the optimal transport assignment increases with higher multiplicative shifts, this likely decreases the x-value of the peaks of the multiplicative death shift density diagrams. Because we have no mathematical model to factor in this effect, the optimisation algorithm tries to approximate this effect by increasing the orientational variance. (In our experiments by roughly 0.25, see Figure 12.)

To further study these two effects, we have analysed the performance of NIPH if the optimisation algorithm gets told one correct parameter. Because NIPH infers the correct orientation with high accuracy even in the base case, we have implemented this analysis only for the scaling factor and the orientational variance, see Figure 12. It appears that providing NIPH with the correct orientational variance does not improve the performance, but even worsens it (*Left* diagram). This supports the second claim: Because NIPH now cannot model some of the noise as noise stemming from orientation variance, its predictive performance is decreased. The *right* diagram shows that giving NIPH the correct scaling factor does not decrease performance. However, now the orientational variance is roughly constantly overestimated by a value of 0.25. This supports the first claim: In the previous experiments (Figure 11 *Right*), overestimation of the orientational variance decreased, as the

underestimation of the scaling quotient increased. When preventing underestimation of the scaling quotient, the orientational variance now is consequently overestimated.

Future work could either work on training an ML approach to compensate for these systematic errors, or to give a more complex mathematical model that better incorporates all sources of noise. However, the main goal of the current paper is to introduce the novel idea of changing the underlying metric of point clouds and then quantifying changes in the PH to infer geometric information on the point cloud and to provide a first proof-of-concept. We have demonstrated that NIPH can harvest rich information, can accurately infer the orientation of a point cloud, and can give a sensible estimation of the scaling factor and orientational variance. We believe that this already serves as a proof-of-concept of the powerful ideas of NIPH, leaving some fine-tuning on implementational details to future work.

D   Implementational details

We implemented NIPH using python. The code to replicate all experiments in this paper can be found under https://git.rwth-aachen.de/netsci/publication-2023-non-isotropic-persistent-homology. We generated the density plots using gaussian kernel densities estimates and set the bandwidth parameter using Scott's rule. For solving the optimisation problem, we used dual annealing from the SciPy optimisation library. We determined the peaks of the multiplicative death shift density curves using a grid search. For Figure 7 *Right*, we compared point clouds sampled from the residential road networks of ~ 15 U.S. and European cities. The cities were chosen based on a size of roughly 50000–100000 inhabitants and city boundaries correctly represented by openstreetmaps. We used the European Cities Tübingen, Göttingen, Cambridge UK, Meran, Södertälje, Belfort, Namur, Delft, Quedlinburg, Passau, Jena, Legnica, Flensburg, Görlitz, and King's Lynn and West Norfolk. For U.S. cities, we used Albany NY, South Bend IN, Boulder CO, Davenport IA, Cupertino CA, Nampa ID, Cheyenne WY Wheaton IL, Utica NY, Santa Fe NM, Bismarck ND, Coon Rapids MN, Sioux City IA, Great Falls MT, Rapid City SD, and Sunrise FL. For NIPH, we sampled 15 different directions and 9 different scaling factors between 1.2 and 2.5. We have sampled the locations of the cars uniformly and independently at random across the entire road network. Hence the expected number of cars on a certain road segment across a city only depended on the length of the segment.

For the synthetic experiments as shown in Figure 5, for each dataset we have generated 200 ellipses with random centre point in a 3000×3000 square with each 100 points. We controlled the principal direction of the ellipses and added normal distributed noise to this principal direction based on the given orientational variance. We uniformly set the length of the smaller principal axis to a value between 0.2 and 2 and scaled the larger principal axis accordingly. An excerpt of the datasets can be found in the top of Figure 3.

D.1   Analytic expressions for expected multiplicative shift

For 1-dimensional homology, the expected multiplicative shift for a true orientation ϕ, sampling orientation ψ, true scaling factor s, sampling scaling factor S and orientational variance $V = 0$ is the minimum of the shift in the scaling direction and the shift orthogonal to the scaling direction

$$\mathbb{E}\mathbf{P}_{\psi,S}(\varphi, 0, s) := \min\left(\frac{S}{\sqrt{1 + (S^2 - 1) \cdot \cos^2(\varphi - \psi)}}, \frac{sS}{\sqrt{1 + (S^2 - 1) \cdot \sin^2(\varphi - \psi)}} \right).$$

For non-zero orientational variance V, we have to fold this function:
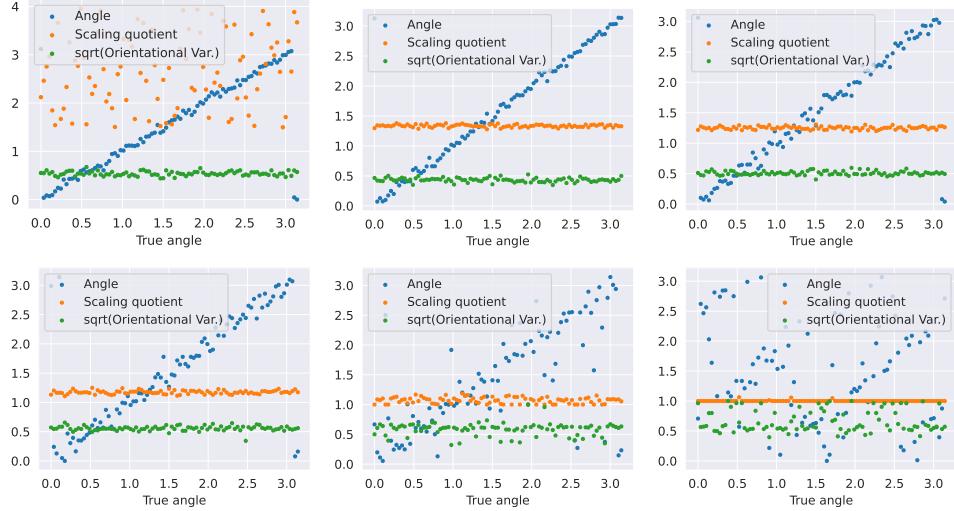
$$\mathbb{E}\mathbf{P}_{\psi,S}(\varphi, V, s) = \text{argmax}_x \frac{1}{\sqrt{V2\pi}} \int_{-\infty}^{\infty} e^{-t^2/(2V)} \delta_{x=\mathbb{E}\mathbf{P}_{\psi,S}(\varphi+t,0,s)} dt$$

We can approximate this using the weighted integral of the peaks:
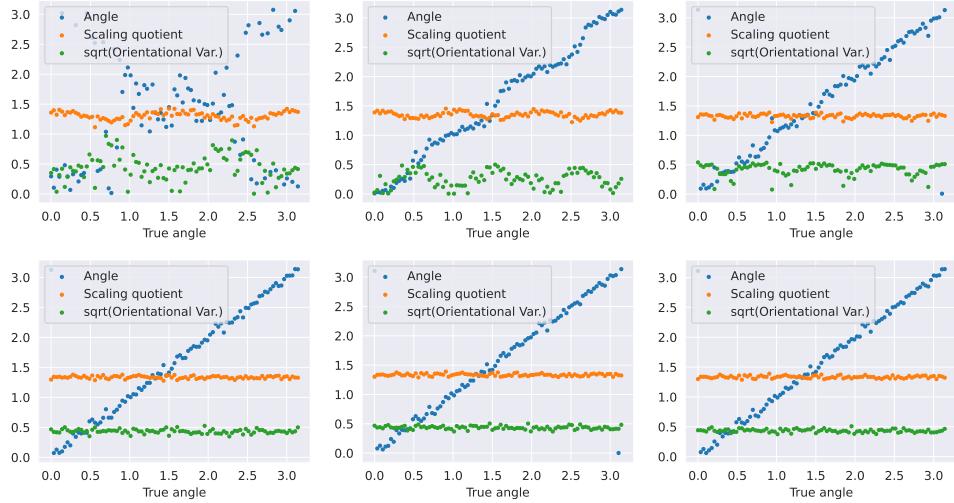
$$\mathbb{E}\mathbf{P}_{\psi,S}(\varphi, V, s) = \frac{1}{\sqrt{V2\pi}} \int_{-\infty}^{\infty} e^{-t^2/(2V)} \mathbb{E}\mathbf{P}_{\psi,S}(\varphi + t, 0, s) dt.$$

In the 0-dimensional case, we do the analogue for the following expected shift function:

$$\mathbb{E}\mathbf{P}_{\psi,S}(\varphi, 0, s) := \min\left(\sqrt{S^2 \cos^2(\varphi - \psi)^2 + \sin^2(\varphi - \psi)}, s \right).$$

Figure 13: Performance of NIPH for different probing scaling factors and 8 **probing directions.** We have a ground truth orientational standard deviation of 0.5, scaling factor of 1.5 and varying angle. *Top left:* Probing scaling factor 1.5, *Top centre:* Probing scaling factor 2.0, *Top right:* Probing scaling factor 2.5, *Bottom left:* Probing scaling factor 3.0, *Bottom centre:* Probing scaling factor 4.0, *Bottom right:* Probing scaling factor of 8.0.
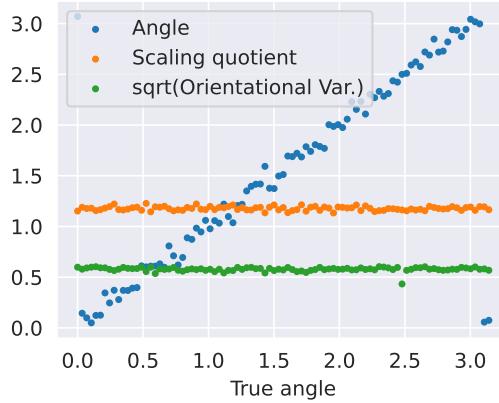
Figure 14: Performance of NIPH for different probing scaling factors and 8 **probing directions.** We have a ground truth orientational standard deviation of 0.5, scaling factor of 1.5 and varying angle. Numbers of probing angles from left to right *First row:* 2–3–4, *Second row:* 8–16–32.

E   Selection of Probing Directions and Scaling Factors

In this section, we will analyse how different parameter choices affect the performance of NIPH. The presented form of NIPH does not have many parameters, and mainly the choice of sampling directions and probing scaling factor has an impact on the result. In Figure 13, we have repeated the experiment of Figure 5 with 8 probing angles and varying probing scaling factors. We make the following observations:

i) Picking a probing scaling factor that is not above the true scaling factor of the data set results in NIPH failing to extract meaningful scaling information. However, the extraction of the orientational variance and the underlying orientation still performs well.

16
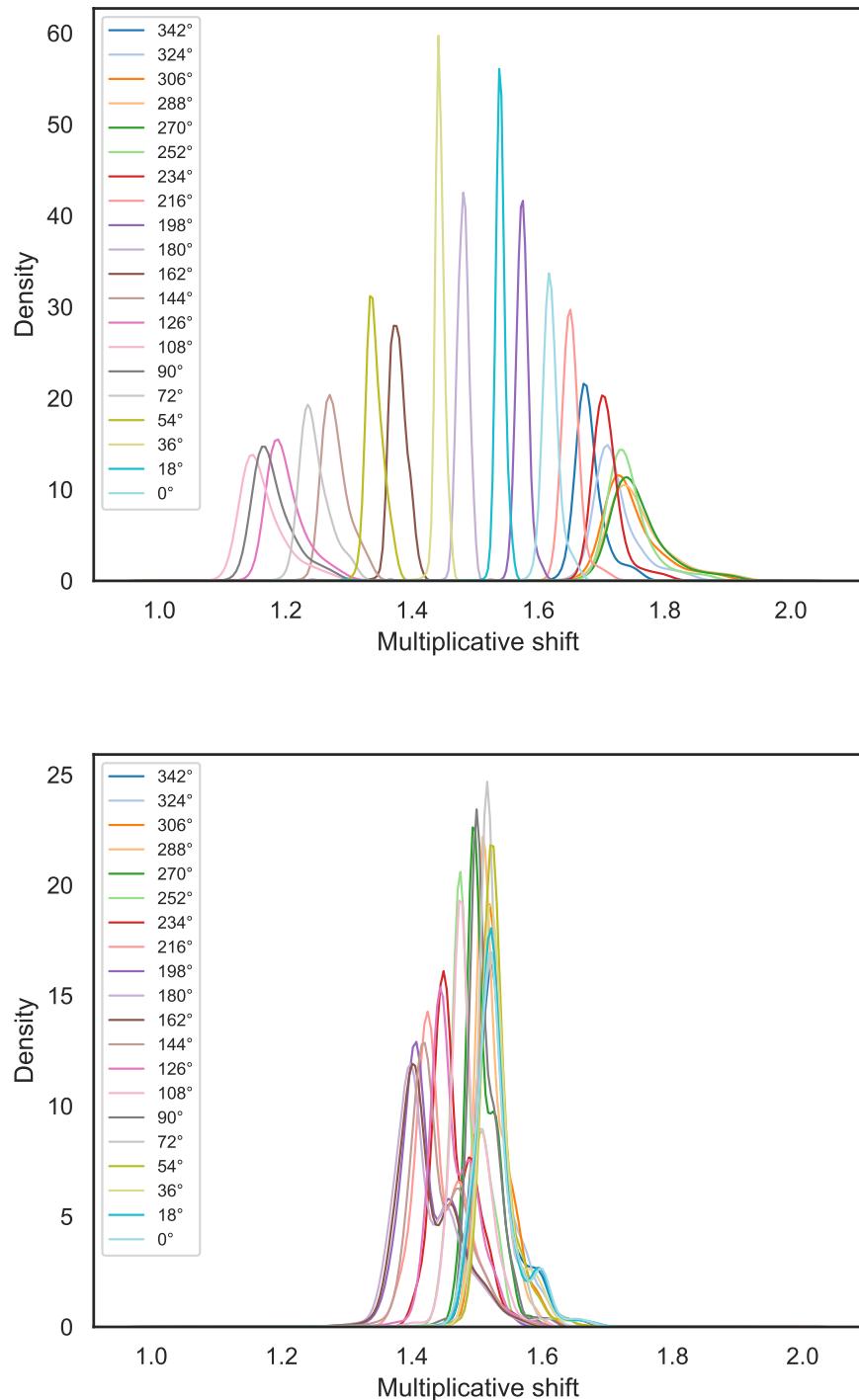
Dataset	$\sqrt{\text{MSE}}\ \varphi$	$\sqrt{\text{MSE}}$ scaling	$\sqrt{\text{MSE}}\ \sqrt{V}$
#2 probing directions	35°	0.21	13°
#3 probing directions	6.3°	**0.16**	16°
#4 probing directions	4.7°	**0.17**	7.7°
#8 probing directions	2.9°	**0.17**	4.3°
#16 probing directions	**2.6°**	**0.17**	4.1°
#32 probing directions	**2.6°**	**0.17**	4.0°
1.5 probing scaling factor	3.0°	1.4	3.5°
2 probing scaling factor	2.9°	**0.17**	4.3°
2.5 probing scaling factor	5.0°	0.25	**1.9°**
3 probing scaling factor	6.8°	0.33	4.0°
4 probing scaling factor	22°	0.43	7.9°
8 probing scaling factor	47°	0.50	14°
Combined scaling factors	3.8°	0.32	4.6°

Figure 15: Quantitative performance of NIPH using different sampling parameters

Figure 16: Performance of NIPH for a combination of the probing scaling factors and 8 **probing directions.** We have a ground truth orientational standard deviation of 0.5, scaling factor of 1.5 and varying angle.
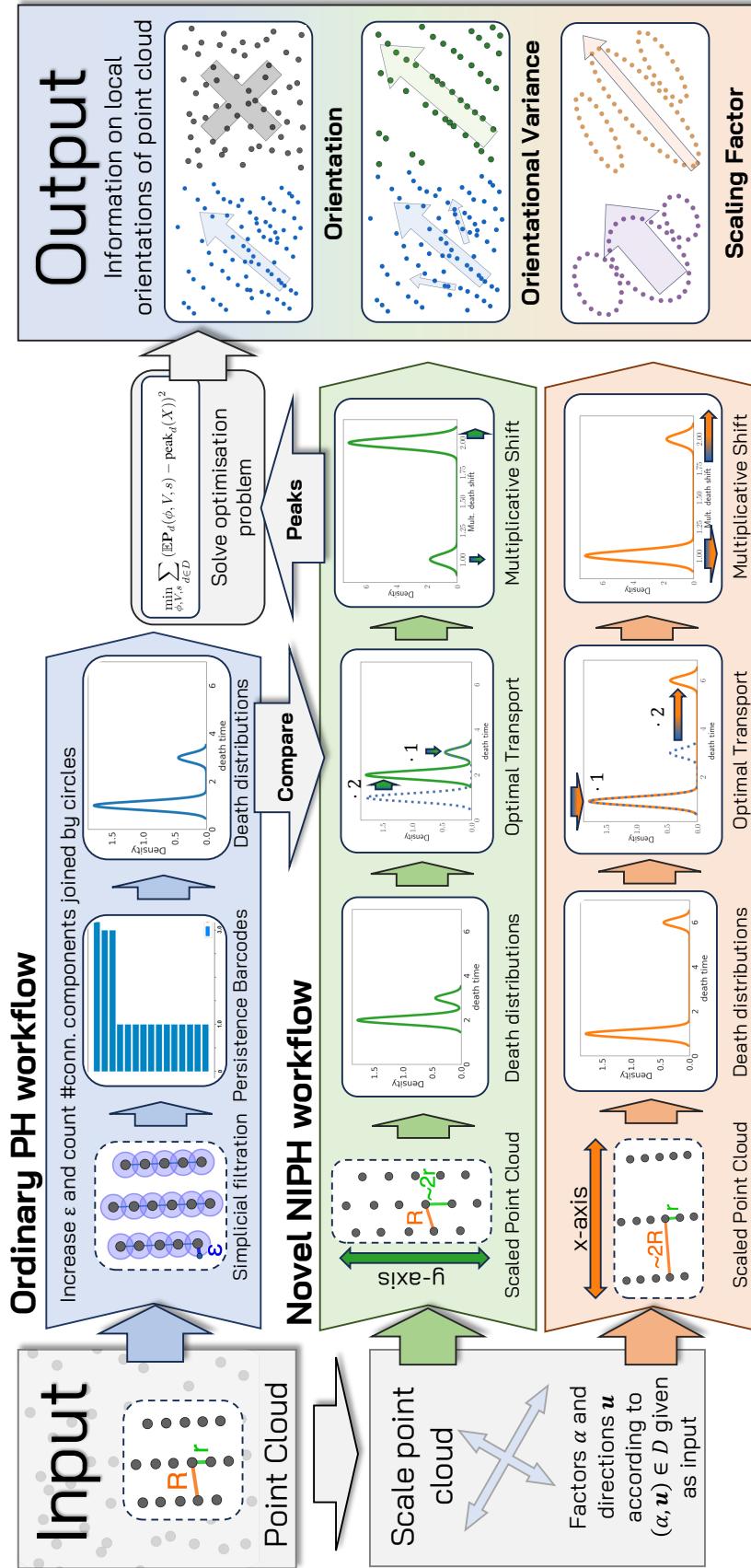
  ii) Picking probing scaling factors that are significantly larger than the underlying scaling factor results in bad overall performance.

  iii) There is a wide range of possible probing scaling directions for which NIPH produces good results.

  iv) Sampling with different probing scaling factors during one iteration of NIPH produces good results comparable with the best individual parameter choice, see Figure 16. This solves the problem of choosing a fixed probing scaling factor in practice.

In Figure 14, we have repeated the experiment of Figure 5 with a varying number of probing angles and a fixed probing scaling direction of 2. The quality of the approximation increases with the number of probing directions used. In our case, the the quality increase above 8 scaling directions is only minuscule. In the case of a low number of sampling directions, the errors are periodic. We compare the root mean squared errors of the different parameter selections in Figure 15.

F Figures

Figure 17: Mult. Shift Diagrams used to construct Figure 6 *Top:* Residential Roads in Manhattan. *Bottom:* Residential Roads in Tübingen.

Figure 18: Schematic of Non-Isotropic Persistent Homology (NIPH). Step 1. Produce different versions of the input point cloud by applying directional scaling in direction **u** and scaling-factor s according to elements in input set D. **Step 2.** Compute PH of desired degree (the diagram displays 0th persistent homology) for input point cloud and each of the scaled point clouds. **Step 3.** Compute optimal transport between weighted death time distributions of input point cloud and each of the scaled point clouds. **Step 4.** Compute multiplicative factor of shift for each death time. Extract maxima and use optimisation problem to compute preferred orientations, scaling-factor, and orientational variance of point cloud, as seen in the illustrations on the right. **Orientation.** Two point clouds with similar 0-dimensional PH. *Left:* The point cloud has additional structure: a preferred orientation. *Right:* There is no preferred orientation in the point cloud. NIPH can detect the differences in the amount of orientation in both point clouds across arbitrary scales. **Orientational variance.** Two point clouds with similarly structured 0-dimensional PH. Both have additional structure and a preferred dimension. *Left:* There is a comparably large variance in the local preferred direction. *Right:* The alignment of the point data with the preferred direction is very strong. NIPH can detect the differences in the amount of variance of orientation in both point clouds. **Scaling factor.** Two point clouds with similarly structured 1-dimensional PH. Both have additional structure and are scaled along the same preferred dimension. The ellipses of the *left* point cloud are scaled differently than the ellipses on the *right*.