Appendix, IsUMap: Manifold Learning and Data Visualization leveraging Vietoris-Rips Filtrations

Parvaneh Joharinad[1,2] **Hannaneh Fahimi**[1,2] **Lukas Silvester Barth**[2] **Janis Keck**[2, 3] **Jürgen Jost**[2]

[1] Center for Scalable Data Analytics and Artificial Intelligence (ScaDS,AI) Dresden/Leipzig, Germany
[2] Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany [3] Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
parvaneh.joharinad@mis.mpg.de, fatemeh.fahimi@mis.mpg.de, lukas.barth@mis.mpg.de, janis.keck@maxplanckschools.de, jjost@mis.mpg.de

A Metric spaces and VR-complexes

We consider a discrete data set $X = \{x_1, x_2, \ldots, x_N\}$, where distinct points can be quantitatively compared via some proximity measure. This proximity is best formulated as a metric (or distance function).

Definition 1 *A metric d on a set X is a map $d : X \times X \to \mathbb{R}_{\geq 0}$ which assigns to each pair (x, y) of points in X a non-negative real number, satisfying the following properties:*

1. $d(x, y) \geq 0$, *and* $d(x, y) = 0$ *iff* $x = y$;
2. $d(x, y) = d(y, x)$; *and*
3. $d(x, z) \leq d(x, y) + d(y, z)$

for all $x, y, z \in X$.

In practice, the function d that quantifies the similarity between points may sometimes fail to satisfy one of the properties of a metric. However, theoretically, there usually exists a semi-metric obtained by relaxing one of these properties. When $d(x, y) = 0$ may also happen for some points $x \neq y$, we get a pseudo-metric. We get an *uber*-metric when distances may be infinite, and an extended metric when the triangle inequality need not hold when $d(x, z) = \infty$. Actually, we shall simply write "metric space" below when more precisely, we typically mean uber-metric spaces (UM for short).
The metric d defines a *topology* on X by considering open subsets obtained from finite intersections of metric open balls

$$U(x, r) := \{y \in X : d(x, y) < r\}.$$

A *topological space* is a set X along with a family $\mathcal{O}$ of subsets, the *open sets*, satisfying the following properties:

1. $\emptyset, X \in \mathcal{O}$.
2. If $U_1, U_2 \in \mathcal{O}$, then also $U_1 \cap U_2 \in \mathcal{O}$;
3. for any index set I, if $(U_\iota)_{\iota \in I} \subset \mathcal{O}$, then also $\bigcup_{\iota \in I} U_\iota \in \mathcal{O}$:

When we have topologies, we can define continuity of maps. Thus, a map $F : X \to Y$, where X and Y are both topological spaces, is *continuous* if and only if the inverse image $F^{-1}(\mathcal{O})$ of every open $\mathcal{O} \subset Y$ is open in X.
There is also a method for constructing scale dependent graphs or simplicial complexes from a dataset X that is equipped with a metric d. For every scale $r \geq 0$, two data points are connected by an edge whenever their distance is $\leq r$. We can then construct the clique complex of the resulting graph, called the *Vietoris-Rips complex*, by filling in triangles, tetrahedrons, and so forth whenever all the required edges exist in the graph.

In combinatorics, an abstract *simplicial complex* with a vertex set X is a family of finite subsets in $\mathcal{P}(X)$, the power set of X, which is closed under taking subsets. This combinatorial structure models relations with a hereditary property, where the relation extends to all subsets of a subset σ that satisfy the relation. Here is the formal definition of the Vietoris-Rips complex:

Definition 2 *For a metric space (X, d) and $r \geq 0$, the Vietoris-Rips complex $VR(X, r)$ is the simplicial complex with vertex set X, where every finite subset $\{x_0, x_1, \ldots, x_n\}$ of X spans a simplex if its diameter is $\leq r$, i.e the distance of any pair of points in the set is not larger than r.*

As we vary r, a family of simplicial complexes emerges, evolving as the scale increases, i.e.

$$VR(X, r) \subseteq VR(X, r'), \text{ for } r \leq r'.$$

To visualize an abstract simplicial complex K, we turn it into a topological space $|K|$, its geometric realization. While there is not a unique geometric realization for every simplicial complex, the standard geometric realization provides a canonical way to construct it, especially when the vertex set X is finite.

Let assign a total order on the vertex set X of K which is assumed to be finite, that is we assume that $X = \{x_1, \ldots, x_N\}$. We then simply employ the standard $(N-1)$-simplex in $\mathbb{R}^N$

$$\Delta_{N-1} := \left\{ (t_0, \ldots, t_{N-1}) \in \mathbb{R}^N \,\middle|\, \sum_{i=0}^{N-1} t_i = 1 : t_i \geq 0 \right\} \tag{1}$$

to construct the standard geometric realization of K, i.e. $|K|$, as we explain below.

Δ_{N-1} is a simplex with vertex set consisting of elements in the standard basis $\{e_1, e_2, \ldots, e_N\}$ of $\mathbb{R}^N$. Every face spanned by $\{e_{i_0}, e_{i_1}, \ldots, e_{i_k}\}$ is the convex hull of the corresponding vertices. Δ_{N-1} is equipped with a canonical metric, i.e. the restriction of Euclidean distance to Δ_{N-1}. $|K|$ is a subcomplex of Δ_{N-1} defined as the collection of faces spanned by $\{e_{i_0}, e_{i_1}, \ldots, e_{i_k}\}$ for which $\{x_{i_0}, x_{i_1}, \ldots, x_{i_k}\}$ form a k-simplex in K. $|K|$ inherits the canonical metric of Δ_{N-1}, making it a metric space.

Although this process yields a metric realization for each $VR(X,r)$, this metric assigns a distance between vertices of $|VR(X,r)|$ which differs from that given by d. Specifically, the Euclidean distance between every pair (e_i, e_j) is equal to $\sqrt{2}$ while the distance between corresponding vertices in X (that is $d(x_i, x_j)$) is not necessarily constant. While the geometric realization (specifically, the standard geometric realization) of Vietoris-Rips complexes does not accurately reflect the geometry of X, it can recover its topology. This is supported by the following *Hausmann theorem*, c.f (Hausmann 1995) (see also (Latschev 2001)):

Theorem 1 *Let (M, g) be a Riemannian manifold that is compact (or more generally, satisfies some technical condition that essentially amounts to a positive lower bound on the injectivity radius), and $VR(M, r)$ be the Vietoris-Rips complex at scale r, corresponding to the metric d defined by the Riemannian metric tensor g. Then there exists $r_0 > 0$ such that for every $r \leq r_0$, the geometric realization of $VR(M, r)$ is homotopy equivalent to M.*

Homotopy equivalence is important because in algebraic topology many concepts are homotopy invariant, that is, their values are invariant under homotopy equivalence. More generally, the widely-used method of persistent homology in topological data analysis records the homology of the scale-dependent combinatorial model $(VR(X,r))_{r\geq 0}$ to recover topological features of a dataset (X, d). Here, however, one is interested both in invariances and in changes of homology, the appearance and disappearance of homology generators, to identify scales of particular interest.
Indeed, for every scale r, the object $VR(X,r)$ forms a simplicial complex, enabling the computation of corresponding simplicial homology groups $H_*(VR(X,r))$ and their generators, i.e., non-trivial cycles which represent topological features.

The inclusion map $VR(X,r) \subseteq VR(X,r')$ for $r \leq r'$ induces morphisms $H_*(VR(X,r)) \to H_*(VR(X,r'))$ between homology groups, facilitating the tracking of topological features across the scaling parameter. Persistent features, enduring over substantial intervals, indicate intrinsic topological features of X, whereas those that vanish quickly are indicative of noise. This method is justified and supported by the equivalence of Čech homology and singular homology of a topological space under certain conditions, c.f. (Wallace 2007), that is more general than the Hausmann theorem.

A.1 Weighted Simplicial Complexes and One-Parameter-Filtrations

As mentioned in the main text, the Vietoris-Rips-filtration is a special case of a family of one-parameter filtrations. These filtrations consist of a collection of complexes $(S_r)_{r\geq 0}$, where $\sigma \in S_r \implies \sigma \in S_{r'}$, that is, they are monotone with respect to the parameter. Note that such filtrations, where r is not necessarily a distance radius, often arise in practice, such as from a complex network via parameter-dependent relationships like the clique complex of a social network. Persistent homology has also been used for such a filtrations if the parameter-dependent correlation is defined in a way that the generated simplicial complexes grow as the parameter increases, c.f. (Horak, Maletić, and Rajković 2009; Myers, Munch, and Khasawneh 2019; Battiston et al. 2020).

Now, given a filtration (S_r), one may canonically obtain a weighted simplicial complex from it by assigning to each simplex σ the least scale at which this simplex first appears in the filtration, that is,

$$W(\sigma) = \inf\{r : \sigma \in S_r\},$$

where one may additionally include the weight ∞ for simplices σ which never appear. One immediately observes that the VR-filtration is a special case of this construction. Also, note that this assignment ensures that the weight of each simplex is equal to the maximum of the weights of its $1-$dimensional faces, i.e. edges, which fulfills the monotonicity requirement of the weights we have imposed for weighted simplicial complexes. Conversely, one may construct a filtration from a weighted complex by letting

$$S_r = \{\sigma \in S | W(\sigma) \leq r\}.$$

This implies that one may transfer between these two descriptions without losing information and hence, one may perform operations in the setting that is more natural for the desired outcome.

The question then arises whether there exists a distance relation on the set of nodes of a complex network that leads to such a filtration. More precisely, we seek a metric space (Y, d) that encompasses the set of nodes X such that $VR(Y,r)$ reflects the geometry of our 1-parameter filtration. This is the question we aim to address utilizing the theory of metric realization introduced in (Spivak 2009).

For this purpose, we require a combinatorial model, a weighted simplicial complex, to represent the correlation between data points.This model will be utilized to construct the metric realization of the corresponding complex. In fact, as our considerations on recovering the topology imply, we want to construct a complex that only retains information up to a certain scale - the radius that ensures homotopy through the Hausmann theorem 1, ideally. When we restrict the construction in this way, we obtain the following weighted simplicial complex:

Definition 3 *Let $(VR(X,r))_{0\leq r\leq R}$ be a portion of the Vietoris-Rips filtration of the metric space (X, d). Then we define the corresponding weighted simplicial complex as a full simplex whose vertices are the elements of X, with the weights assigned*

by

$$W(\sigma) := \begin{cases} \min\{0 \leq r \leq R; \sigma \in VR(X,r)\}, & \textit{if } \sigma \in VR(X,R) \\ \infty & \textit{otherwise} \end{cases} \tag{2}$$

A.2 Weighted and Fuzzy Simplicial Complexes

As noted before and in the main text, we sometimes want to transfer to the category of fuzzy simplicial complexes, where a family of merging operations is readily defined. We now want to explain how to transfer between weighted and fuzzy simplicial complexes, and how this transfer modifies the properties of those complexes which correspond to VR-filtrations. Recall that a fuzzy simplicial complex is a classical fuzzy set, with the particular property that the strength function does not decrease as one traverses from a simplex to its faces. That is, for the fuzzy weight $\mu(\sigma) \in [0,1]$ we have $\sigma \subset \sigma' \implies \mu(\sigma) \geq \mu(\sigma')$.

Let now $\Phi : [0,\infty] \to [0,1]$ be a monotonically decreasing, continuous function such that $\Phi(0) = 1$ and $\Phi(\infty) = 0$. Then for any simplicial complex (S, W), we obtain a fuzzy simplicial complex $(S, \Phi \circ W)$. Conversely, we obtain back a weighted complex by applying the inverse $\Phi^{-1}(w)$.
We now want to identify the conditions to ensure that the inverse of $\Phi : [0,\infty] \to [0,1]$ (strictly decreasing with $\Phi(0) = 1$ and $\Phi(\infty) = 0$) converts a fuzzy simplicial complex with strength function w into a weighted simplicial complex with weight function W, as described in Equation 2, corresponding to some extended metric d on the vertex set X.

1- The value of w on vertices must be 1 and < 1 for other simplices. This ensures that the first property of a metric (being reflexive) is satisfied.

2- The symmetry of the metric is guaranteed as there is no direction on 1-simplices (edges).

3- The triangle inequality for the extended metric is satisfied if and only if for any pair of vertices (x, z), either $w([x, z]) = 0$ or

$$w([x,z]) \geq \Phi\left(\Phi^{-1}(w[x,y]) + \Phi^{-1}(w[y,z])\right), \forall y \in X$$

4- The weight $w(\sigma)$ of each simplex should be equal to the minimum of the weights on its 1-dimensional faces.

We here set the strength of simplices, that are not present, to 0.
The last condition could be relaxed to only imply that $w(\sigma)$ is at most equal to the minimum of the weights on its 1-dimensional faces. This, for instance, is the case for the Čech complexes of open balls centered at points in X and with varying radii.
Therefore, if assuming that (S, w) is a fuzzy simplicial complex on the vertex set X with the strength function w satisfying conditions 1,2,4 above, then one can search for a proper function $\Psi : [0,1] \to [0,\infty]$ (strictly decreasing with $\Psi(1) = 0$ and $\Psi(0) = \infty$) such that the triangle inequality is satisfied. That is, for (x, z) with $w([x, z]) \neq 0$, one has $\Psi(w([x, z])) \leq \Psi(w([x, y])) + \Psi(w([y, z]))$.

Example. Let the fuzzy simplicial complex (S, w) with vertex set X satisfy conditions 1,2,4 above. If for every (x, z) with $w([x, z]) \neq 0$

$$w([x,z]) \geq w([x,y])w([y,z]),\ \forall y \in X,$$

Then the map $\Psi = -\log$ converts S to a Vietoris-Rips filtration.

This fuzzy simplicial complex model proves especially advantageous when combining two or more such complexes as we describe in Section C.

B Metric realizations

As outlined in Section A, for each scale r, there exists a geometric realization of the simplicial complex $VR(X, r)$, denoted by $|VR(X, r)|$, where each k-simplex is represented by the standard geometric simplex Δ^k as defined in 1.
Inspired by this geometric approach, Spivak (Spivak 2009) constructed the metric realization of a fuzzy simplicial set. We use this to obtain a realization for the weighted simplicial complex of Definition 2. Since we are interested in finite data sets equipped with discrete metrics, we will not discuss non-finite metric spaces.
Here, we present a version of metric realization for a weighted simplicial complex slightly modified from that of (Spivak 2009). We specifically tailor this method to accommodate the weighted Vietoris-Rips complex obtained from $(VR(X, r))_{0\leq r\leq R}$, which represents a subset of the entire spectrum $(VR(X, r))_{r\geq 0}$.
In the sequel, we assume that (S, W) represents an admissible weighted simplicial complex on the vertex set X:

Definition 4 *A weight function W on the simplicial complex S on the vertex set X, assigning to each simplex $\sigma \in S$ a non-negative finite weight $W_\sigma := W(\sigma)$, is called* admissible *if it satisfies the following conditions:*

1. $W_x = 0$ for every vertex $x \in X$, and for any other simplex $\sigma \in S$, W_σ is strictly positive.

2. *The weight is non-increasing when moving from a simplex to its facets. More restrictively, we can assume that the weight of each simplex is the maximum of the weights on its* 1*-skeleton.*

Remark. The weight function (2) corresponding to the Vietoris-Rips filtration has the following additional property

3. If $[x, z] \in S$, then for every $y \in X$ we have

$$W_{[x,z]} \leq W_{[x,y]} + W_{[y,z]}, \tag{3}$$

where we assume an infinite weight for $[x, y]$ or $[y, z]$ when not appearing in S.

However, this assumption is not necessary for metric realization. The process of metric realization resolves any violations of the triangle inequality by the weight function. In fact, even if starting with a weighted simplicial complex (X, W) on the vertex set X that fail to satisfy (3), the metric realization will ultimately define a metric on X whose Vietoris-Rips filtration is identical with S while the weight function W is being improved to satisfy the triangle inequality (3).

The process of metric realization consists of two main steps. First, we construct the metric realization of each simplex (σ, W_σ). Second, we glue different simplices together to obtain the realization of the entire simplicial complex (S, W). For each $k-$simplex $\sigma \in S$ with weight W_σ, we map σ to $\Delta^k_{W_\sigma}$, the simplex whose vertices lie on the coordinates axes at a distance of W_σ from the origin, that is,

$$\Delta^k_{W_\sigma} = \{(t_0, \ldots, t_k) \in \mathbb{R}^{k+1} | \sum_{i=0}^{n} t_i = W_\sigma : t_i \geq 0\}. \tag{4}$$

It is equipped with the induced Euclidean metric.

Thus, $\Delta^k_{W_\sigma}$ is isometric to Δ^k equipped with the Euclidean metric scaled by W_σ and we denote this scaled metric by d_{W_σ}. For the second step, i.e. gluing different simplices, we use the general procedure for gluing different metric spaces. In metric geometry, when gluing two or more metric spaces, the first step involves fixing the gluing parts, c.f. (Burago, Burago, and Ivanov 2001). For simplicity, let's consider two metric spaces (X, d_X) and (Y, d_Y), although this procedure can be generalized to more than two spaces.

Definition 5 *Let (X, d_X) and (Y, d_Y) be two metric spaces. If there is a bijection $\phi : X' \subset X \to Y' \subset Y$, the* gluing *of X and Y along X' and Y' is obtained as follows.*
First, one takes the disjoint union $Z := X \bigsqcup Y$ of X and Y as sets and defines an extended metric d_Z on Z as follows.

$$d_Z(z, z') = \begin{cases} d_X(z, z'), & z, z' \in X \\ d_Y(z, z'), & z, z' \in Y \\ \infty & \textit{else.} \end{cases} \tag{5}$$

Next, gluing is performed by identifying points $x \in X$ and $y \in Y$ whenever (x, y) is on the graph of ϕ. In other words, the bijection ϕ defines an equivalence relation, $x \sim y$ if $y = \phi(x)$. The gluing metric space is $(Z/d_\sim, d_\sim)$ with the quotient metric $d_\sim$, which is defined by the following general procedure.
Let (Z, d_Z) be a metric space with an equivalence relation $\sim$. We then let

$$d_\sim(p, q) := \inf \sum_{i=1}^{\ell} d_Z(p_{i-1}, q_i) \tag{6}$$

where the infimum is taken over all sequences of the form $(p_0, q_1, \cdots, p_{l-1}, q_l)$ with $p_0 = p, p_\ell = q, p_i \sim q_i$ for $i = 1, \ldots \ell - 1$. $d_\sim$ defines a pseudo metric on Z and the quotient metric space $(Z/d_\sim, d_\sim)$ is obtained by identifying points with zero $d_\sim$ distance.

In the simplicial complex (S, W), each simplex σ carries a weight $W(\sigma)$, which may differ from the weights assigned to its corresponding boundary faces $\partial\sigma$. For example, if $\sigma = x_{i_0}, \ldots, x_{i_k}$ is a k-simplex, then for each $0 \leq j \leq k$, the simplex $\partial_j\sigma := x_{i_0}, \ldots, \widehat{x_{i_j}}, \ldots, x_{i_k}$ represents one of its $(k-1)$-dimensional facets. These facets are mapped to their respective metric simplices $\Delta^{k-1}_{W_{\partial_j\sigma}}$ (or equivalently $(\Delta^{k-1}, d_{W_{\partial_j\sigma}})$) during the realization process. The following lemma then specifies the metric on the entire simplex obtained by attaching all the corresponding faces.

Lemma 1 *The mapping $\phi_j : \Delta^{k-1}_{W_{\partial_j\sigma}} \to \Delta^k_{W_\sigma}$, defined as*

$$(t_0, \ldots, t_{k-1}) \mapsto \frac{W_\sigma}{W_{\partial_j\sigma}}(t_0, t_1, \ldots, t_{j-1}, 0, t_{j+1}, \ldots, t_{k-1}), \tag{7}$$

is a bijection onto a face of $\Delta^k_{W_\sigma}$. The resulting metric realization is the gluing metric of Definition 5 induced by the gluing maps ϕ_j (which maps facets to their corresponding co-faces) and the identity map on shared faces (which attaches two simplices that share a face).

Proof. Multiplying by $\frac{1}{W_{\partial_j\sigma}}$ maps $\Delta^{k-1}_{W_{\partial_j\sigma}}$ to the standard simplex Δ^{k-1}, which is then mapped to the $j-$th face of Δ^k (obtained by excluding the vertex e_{j+1}) that finally is being mapped to the $j-$th face of $\Delta^k_{W_\sigma}$ by multiplying by W_σ. This process effectively shrinks the j-th face of $\Delta^k_{W_\sigma}$ to become isometric to $\Delta^{k-1}_{W_{\partial_j\sigma}}$. □

Remark. The gluing of the corresponding face induced by ϕ_j results in the shrinking of that face proportional to its weight. This process is iteratively extended to all the faces of a simplex. Therefore, one can start from the 1-skeleton and develop the gluing of (S, W) by iteratively attaching simplices. At the end, the distance between each pair of points is obtained by (6).

Corollary. *Since the weight decreases as traversing from simplex to each of its facets, the distance between points in the interior of each simplex is obtained by possibly traversing through its faces. Thus, the shortest path connecting two interior points of a simplex never travels through its higher dimensional co-faces.*
For pairs of vertices, as the exceptional case, the distance is realized by a shortest path consisting of edges (i.e. the graph distance on the $1-$skeleton). Therefore, in application, one could only restrict the realization to the realization of the $1-$skeleton of (S, W) if, at the end, the aim is to only obtain the restriction of this realization to the vertex set.

Example. Starting from a fuzzy simplicial complex (S, w) that satisfies the conditions outlined in the example presented at the end of Section A, applying the function $\Psi = -log$ converts the strength function w into a weight function W, thereby yielding an admissible simplicial complex (S, W). The metric realization of (S, W) is the same as that of (S, w) as outlined in (Spivak 2009).

C Combining metric spaces

In this section we will use the Vietoris-Rips filtration and the resulting admissible simplicial complex model to combine two (or more) different metrics on the same set X.

Suppose we have two metric spaces $X_1 := (X, d_1)$ and $X_2 := (X, d_2)$, both defined over the finite point set X. Additionally, let $(VR(X_i, r))_{0\leq r\leq R_i}$ represent a portion of the Vietoris-Rips filtration for each metric space X_i, $i = 1, 2$. Correspondingly, we denote their resulting admissible simplicial complexes by (S^i, W^i) for $i = 1, 2$. To combine these two metric spaces, we need to define a process that combines their respective Vietoris-Rips filtrations and admissible simplicial complexes. The goal is to create a unified representation that incorporates the information from both metrics. While we have a canonical way to directly combine (S^1, W^1) and (S^2, W^2), we develop a probabilistic merging procedure for the corresponding fuzzy simplicial complexes.

Such merging employs *triangular conorms* (briefly *t-conorms*) which are dual to *triangular norms* (briefly *t-norms*), under the order-reversing operation that assigns $1 - a$ to a in $[0, 1]$. t-norms are binary operations introduced by Menger (Menger and Menger 1942) for generalizing transitivity criteria in the context of probabilistic metric spaces. Later, Zadeh (Zadeh 1965) used t-norms to generalize logical operations in fuzzy logic, a particular class of multi-valued logics. The definitions are as follows

Definition 6 *A t-norm is a function $T : [0, 1] \times [0, 1] \to [0, 1]$, satisfying the following properties :*

- *Commutativity: $T(a, b) = T(b, a)$,*
- *Monotonicity: $T(a, b)$ is non-decreasing in either variable. i.e., $T(a, b) \leq T(c, d)$ if $a \leq c$ and $b \leq d$,*
- *Associativity: $T(a, T(b, c)) = T(T(a, b), c)$,*
- *Identity element 1: $T(a, 1) = a$*

The name triangular norm refers to the fact that, in the context of probabilistic metric spaces, t-norms generalize the triangle inequality of ordinary metric spaces, c.f. (Menger and Menger 1942).

Definition 7 *A t-conorm is a binary operation $T^{co} : [0, 1] \times [0, 1] \to [0, 1]$ that has the same properties as a t-norm, but its identity element is 0 (i.e., $T^{co}(a, 0) = a$).*

We can apply the t-(co)-norms iteratively and obtain something akin to a product

Definition 8 *For T either a t-norm or a t-conorm, and $\alpha_1, ..., \alpha_n \in [0, 1]$ we define their T-product as follows*

$$\begin{aligned} \prod^T : [0,1]^n &\to [0,1], \\ (\alpha_1, \ldots, \alpha_n) &\mapsto T(\alpha_1(T(\alpha_2, (\cdots T(\alpha_{n-1}, \alpha_n))))). \end{aligned} \tag{8}$$

Note that by associativity, a T-product does not depend on the order. We also note that the monotonicity carries over to this T-product:

Lemma 2 *If $\alpha^i \leq \beta^i$ for $1 \leq i \leq n$, then $\prod^T \alpha_i \leq \prod^T \beta_i$.*

When merging fuzzy simplicial complexes, this property ensures that the resulting object remains a fuzzy simplicial complex. In other words, the combination of probabilistic weights on each simplex does not exceed the combination of probabilistic weights on any of its faces. One can define isomorphic counterparts of t-norms and t-conorms for combining metric-related weights. For instance, The t-conorm $T(a, b) := \max\{a, b\}$, called the canonical t-conorm, has an isomorphic counterpart $\hat{T} : [0, \infty] \times [0, \infty] \to [0, \infty]$, with properties of commutativity, associativity, monotonicity and identity element 0, defined by $\hat{T}(A, B) := \min\{A, B\}$. This t-conorm is the one that we will use in the IsUMap pipeline, as its metric counterpart facilitates the gluing of every pair of metric graphs (G, W^1) and (G, W^2) as outlined in Definition 5.

With a function $\Phi : [0, \infty] \to [0, 1]$ that is strictly decreasing with $\Phi(0) = 1$ and $\Phi(\infty) = 0$, we convert the metric weights into probabilistic weights. We assign a probability weight of zero to simplices that do not appear in (S, W). This means that simplices which are not present in S are assumed to only appear at infinity. This way, the fuzzy simplicial complex corresponding to every admissible simplicial complex (S, W) is the complete simplex on X with the strength function

$$w(\sigma) = \begin{cases} \Phi(W(\sigma)), & \text{if } \sigma \in S \\ 0 & else \end{cases} \tag{9}$$

Following this procedure, (S^1, W^1) and (S^2, W^2) are converted to fuzzy simplicial complexes (S, w^1) and (S, w^2). Here, S is the $(N-1)$-simplex with vertex set X. By selecting a t-conorm to merge w^1 and w^2, we can equip S with a strength function w to create a fuzzy simplicial complex as follows.

Let T be a t-conorm, and denote S_k as the set of k-simplices of S for $k = 0, \ldots, N-1$. Then, defining $w(\sigma) := T(w^1(\sigma), w^2(\sigma))$ assigns a strength to each simplex σ. This operation, in line with Lemma 2, ensures that (S, w) forms a fuzzy simplicial complex. In fact, $w^i(\sigma)$ (for $i = 1, 2$) for each $\sigma \in S_k$ is at most the minimum of the strengths of its corresponding facets, i.e. $w^i(\sigma) \leq w^i(\partial_j \sigma)$ for $j = 0, \ldots, k-1$. Consequently, Lemma 2 implies that $w(\sigma) = T(w^1(\sigma), w^2(\sigma)) \leq w(\partial_j \sigma) = T(w^1(\partial_j \sigma), w^2(\partial_j \sigma))$.

To convert the resulting fuzzy simplicial set into an admissible simplicial complex $(S^1 \bigcup S^2)$, the strength function w has to be converted into a distance-related weight function W on $S^1 \bigcup S^2$, as outlined in Definition 4, for example with the map Φ^{-1}. The strict monotonicity of the map Φ and the fact that $\Phi(0) = 1$ ensure that conditions 1 and 2 in Definition 4 are satisfied, respectively. The 3rd condition in eq. (3), however, may fail to be satisfied after merging the fuzzy simplicial complexes and converting the resulting probabilistic weight function w back with Φ^{-1}. For instance, even if both w^1 and w^2 satisfy the criteria outlined in the example presented at the end of Section A, it is not necessarily true that the combined strength function $w := T(w^1, w^2)$ for a t-conorm T will still meet the same criteria to make it suitable for conversion through the $-\log$ map that we shall use below. This, however, does not pose any problem, as the metric realization will address this issue, as referenced in Remark B.

D The IsUMap pipeline: A structural overview

This subsection explores the IsUMap pipeline, providing a comprehensive elaboration of the steps from various perspectives, including geometry and combinatorics. It closely aligns with the theoretical framework of merging metric spaces and the metric realization of admissible simplicial complexes, as described in Sections C and B.

1. **Local metrics.** The first step involves constructing local metrics d_i, for $i = 1, \ldots, N$, as

$$\begin{aligned} d_i(x_i, x_{i_j}) = d_i(x_{i_j}, x_i) &= \frac{d(x_i, x_{i_j}) - \rho_i}{\sigma_i} \quad \text{for } j = 1, \ldots k \\ d_i(x, x) &= 0 \quad \text{for all } x \in X_i \\ d_i(x_j, x) &= \infty \quad \text{in all other cases.} \end{aligned} \tag{10}$$

   Here, σ_i and ρ_i are two hyperparameters.

   The hyperparameter σ_i aims to address the issue of non-uniform data distribution on the manifold through a conformal transformation of the metric while ρ_i has the purpose of mitigating the curse of dimensionality. When we have a finite metric space (X, d), derived from a sample drawn from a Riemannian manifold, one can approximate the radial distances of that manifold in the local neighborhood around each point $x_i \in X$. This is achieved by identifying the k closest points $\{x_{i_1}, \ldots, x_{i_k}\}$, and measuring their distances in the metric d to the given point x_i. A suitable choice of k ensures that these

distances approximate the geodesic distances on the manifold, i.e. the distances induced by a Riemannian metric tensor. For the relevant geometric theory, see (Joharinad and Jost 2023; Jost 7th ed., 2017).

These geodesic distances represent the first coordinate in a special coordinate system, the normal coordinate system introduced by Riemann, which we describe in the following.

Let (M, g) be a (finite-dimensional) Riemannian manifold. For every $p \in M$ and every $V \in T_pM$, the tangent space at p, there exists some $\delta > 0$ and a shortest geodesic

$$c_V : [0, \delta] \to M, c_V(0) = p, \frac{d}{dt}c_V(0) = V.$$

This geodesic segment, as it extends from p to an endpoint, represents the distance between these two points.
When c_V is defined on $[0, \delta]$, then $c_{\delta V}$ is defined on $[0, 1]$. These considerations imply that $U_p := \{V \in T_pM : c_V \text{ is defined on } [0, 1]\}$ contains some neighborhood of the origin $0 \in T_pM$.

Definition 9 *Let $p \in M$, where M is a Riemannian manifold.*

$$\begin{aligned} \exp_p : \quad & U_p \to M \\ & V \mapsto c_V(1) \end{aligned} \tag{11}$$

is called the exponential map *of M at p.*

The exponential map thus maps every $V \in U_p \subset T_p(M)$ to the endpoint of a geodesic of length $\|V\|$ that starts at p in the direction V. In particular, the derivative of the exponential map at $0 \in T_pM$ is the identity, as the image of $V \in T_pM$ is the tangent vector of the geodesic starting in the direction V, that is, V itself. It therefore maps some neighborhood of the origin $0 \in T_pM$ diffeomorphically onto a neighborhood of $p \in M$. The coordinates resulting from the inverse of the exponential map are known as *normal coordinates* (or *Riemann normal coordinates*) at p.

If we convert from Cartesian (Euclidean) (i.e. (x^i)) to polar coordinates (i.e. (r, φ^j)), where $r = \sqrt{\sum (x^i)^2}$ represents the distance from the center p and φ are the spherical angles that define points on a distance sphere, we create Riemann polar coordinates. These coordinates provide a way to visualize the manifold from the perspective of a single point. Normal coordinates could be described in an intuitive manner as follows. Given a point p, used as the center of the coordinate system, for any point q in a sufficiently small neighborhood of p, there is a unique shortest geodesic from p to q. The direction of the tangent to this geodesic a p selects a point ϕ in the unit sphere in the tangent space at p, and together with the distance r from p to q, we get the polar coordinates (r, ϕ) of q in the normal polar coordinate system at p.

When using a k-neighborhood graph to model (X, d), with edges weighted according to the distances between the corresponding points, we are effectively capturing the first coordinate in the normal coordinate system at each point x_i. This coordinate r scales when we apply a conformal change to the Riemannian metric tensor g, by scaling the inner-product g_p on the tangent space T_pM with a positive scalar $\lambda(p)$. A conformal transformation involves replacing g with λg, where λ is a positive smooth function on M. This transformation scales the norm of tangent vectors by a positive factor, as the norm is determined by the inner product defined on each tangent space by the Riemannian metric tensor. Importantly, while the magnitudes of vectors are scaled, the angles between any pair of vectors in the tangent space remain unchanged by the conformal transformation. Again we refer to (Jost 7th ed., 2017) for more on this.

The polar coordinate system around p yields a star graph on T_pM, capturing the configuration of the neighboring points in the sample. This configuration magnifies or shrinks by scaling the inner product on T_pM. After applying this pointwise modification, one could map the resulting star graph back to the manifold using the initial geometry, specifically the exponential map $\exp_p$ corresponding to the metric g (i.e. 11). This process, however, moves points along the geodesics connecting them to p, either toward or away from p. This step, which incorporates the geometric configuration around each point on a local scale, suggests that to obtain geodesics under the new metric, a correction considering the global structure is necessary.

By choosing a point-dependent scaling factor $\lambda(x_i) =: \frac{1}{\sigma_i}$ in 10, the metric can be adjusted such that it increases in regions with high data density, where many data points are closely packed, and decreases in regions with lower density. This approach aims to make distances between nearest neighbors more uniform.

Given this scaling factor, the choice of σ_i becomes crucial for effective normalization that aligns with density estimation. One common approach, as used in UMAP (McInnes, Healy, and Melville 2018), sets σ_i such that it equalizes a certain property across all data points. This method involves choosing σ_i with a binary search algorithm so that

$$\sum_{j=1}^{k} \exp\left(-d_i(x_i, x_{i_j}\right) = \log_2(k).$$

This normalization is very similar to, and perhaps inspired by, the one used in t-SNE, cf. (van der Maaten and Hinton 2008).

However, in our case, we point out that our normalization of σ adheres more strictly to the original theoretical idea in UMAP of uniformizing the distribution of the data. We simply set σ_i to be the distance to the kth neighbor, i.e.

$$\sigma_i := d_i(x_i, x_{i_k}).$$

Since this conformal factor σ_i varies as we change the point $x_i \in X$, the corresponding weight $W_{ij} = d_i(x_i, x_j)$ on the k-neighborhood graph could be different from $W_{ji} = d_j(x_i, x_j)$ even if each of them is among the k nearest neighbors of the other.

This difference could become even larger when subtracting ρ_i, chosen to be the distance to the first nearest neighbor x_{i_1}. Therefore, we consider this multi-graph (or directed graph) as the union of star graphs Γ_i, for $i = 1, \ldots, N$. Each star graph represents the weighted simplicial complex associated with the Vietoris-Rips filtration of the metric space (X, d_i).

2. **Combining local distances.** From a geometric point of view, our aim is to merge the Vietoris-Rips filtrations of metric spaces $\{(X, d_i),\ i = 1, \ldots N\}$, to create a final filtration that corresponds to the Vietoris-Rips filtration of a distance function on X. This final distance function is achieved by gluing the individual metrics d_i along their common parts, which is the original set X.

   To this end, the algorithm employs the canonical merging process utilized by the metric t-conorm, where instead of applying t-conorm to the weights of simplices in the fuzzy setting one directly applies the dual operator to the metric weights, c.f. (Simas, Correia, and Rocha 2021). This method is particularly useful because it can be applied directly to the set of star graphs $\{\Gamma_i :\ i = 1, \ldots N\}$, where each star graph represents the Vietoris-Rips filtration for a local metric d_i. The merging process does not require converting the metric-based weights $d_{ij} = d_i(x_i, x_{i_j})$ to probabilistic ones, allowing for a more straightforward and geometrically meaningful combination of the different metric spaces. However, alternative t-conorms can be applied if desired by transforming the weights, such as by $d_{ij} =: W_{ij} \mapsto w_{ij} := \exp(-W_{ij})$. We will present some results of applying IsUMap with different t-conorms in Section E.

   The canonical t-conorm, also known as the Maximum t-conorm, assigns to each pair $(a, b) \in [0, 1] \times [0, 1]$ their maximum, $(a, b) \mapsto \max\{a, b\}$. Its metric counterpart, used in IsUMap, assigns the minimum to each pair. This can be interpreted in terms of metric gluing, where a simplex is merged across multiple Vietoris-Rips filtrations by choosing the smallest scale at which the simplex appears in at least one of the filtrations. This smallest scale becomes the emerging scale in the final filtration, ensuring that the scale at which a simplex emerges in the final filtration corresponds to its earliest appearance in any of the individual filtrations.

   The final metric is derived through the metric realization (Section B) of the admissible simplicial complex that arises from the combination of these star graphs, as detailed in Section C.

   The merging process via t-conorm yields the symmetrization of the incomplete distance matrix obtained in the IsUMap pipeline by local distortions d_i of d. The metric realization yields the completion of This distance matrix through the Dijkstra algorithm to determine the shortest paths between all pairs of vertices.

3. **Low dimensional embedding.** After constructing the metric realization of the admissible simplicial complex (in our case an admissible graph), and restricting the distance function to the sample X (which results in the completed metric matrix), the next step would be to visualize X in a low dimensional Euclidean space while preserving the mutual distances to the extent possible.

   At this stage, we have two options: One is to follow a methodology akin to UMAP, utilizing techniques like a force-directed graph layout or stochastic gradient descent to construct the embedding based on the information recorded in the fuzzy counterpart graph. In this approach, it is not necessarily required to complete the metric matrix after symmetrization, as the fuzzy graph can serve as a sufficient basis for embedding.

   An alternative approach that is geometrically more resilient can employ multidimensional scaling (MDS) techniques. Both approaches have their advantages. The UMAP-like methodology offers flexibility and is popular for visualizing complex data with minimal computational overhead. MDS, on the other hand, provides a more mathematically grounded way to maintain distance relationships, making it particularly useful when we want to preserve distances.

E Illustrations

In our empirical experiments, great part of which presented in the papre (Section 3), we have consistently observed that our algorithm performs precisely as intended according to its theoretical design. Specifically, it effectively generates a low dimensional representation of a uniformized dataset.
In this section, we provide some experiments and analysis supplementary to those presented in the main note.

E.1 Various factors affecting the outcome of the algorithm

Since some choices can be made within the scheme, it is natural to investigate their effects. In particular, different values of parameter k in constructing the $k-$neighborhood graph defines different concepts of neighborhood, or in the uniformization step, different approaches for defining the scaling factor may be chosen which consequently change the result. Moreover, one can choose different t-conorms for combining local structures, or employ either the probabilistic or the geometric method for the final embedding in low-dimensional space. Here, we explore the impact of some of them.

I) **Different t-conorms.** We probe IsUMap and UMAP on the generated dataset (see image (a) in Figure 3) on the upper hemisphere, varying the t-conorm in both methods. Corresponding visualizations are depicted in Figure 1.

The images in the second row (images (e)-(h)) show the implementation of UMAP using different t-conorms, namely: Algebraic sum (image (e)), Canonical t-conorm (image (f)), Bounded sum (image (g), and Drastic sum (image(h))), where the results are seen to be similar. However, IsUMap in images (a)-(d) produces different results as it creates a spiral in the projection of the hemisphere with Algebraic sum (image (e)) and bounded sum (image (c)), which also fail to project the boundary of the hemisphere. In contrast, the Canonical (image (b)) and Drastic sum (image (d)) operations generate a more uniformly distributed projection.

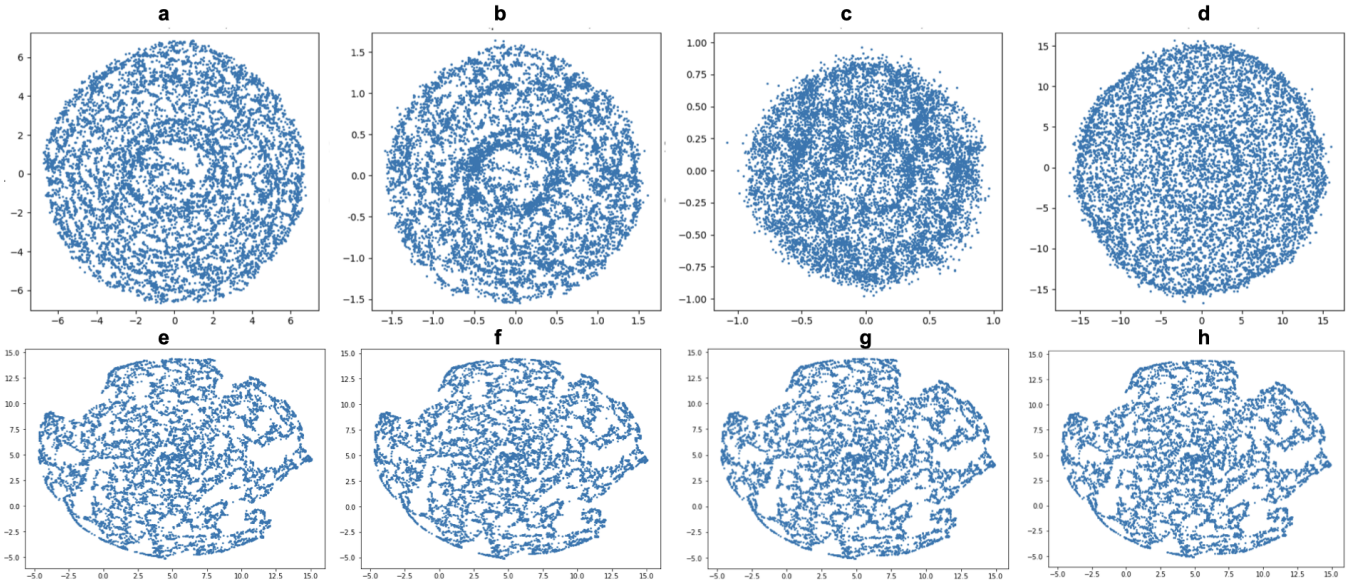

Figure 1: visualization of a sample of size 10000 generated on Hemisphere with non-uniform distribution in dimension 2 by IsUMap (first row) and UMAP (second row) with $k = 30$ and various t-conorms: Algebraic sum (a,e), Canonical (b,f), Bounded sum (c,g), Drastic sum (d,h).

The difference in outcomes between Canonical and Drastic sums might be attributed to the subtraction of ρ_i (the distance to the nearest neighbor). Since this dataset's dimension is reduced only by 1, applying this subtraction may not be necessary. This is depicted in Figure 2.

Comparing the results of UMAP and IsUMap on the dataset generated on the hemisphere, we notice that IsUMap offers a geometrically more robust visualization, whereas UMAP might be more efficient for clustering tasks. This advantage of UMAP over IsUMap for clustering becomes particularly evident in the visualization of the MNIST dataset, as depicted in Figure 7 in the main note. However, IsUMap has an advantage over UMAP when it comes to visualization of lower-dimensional manifolds.

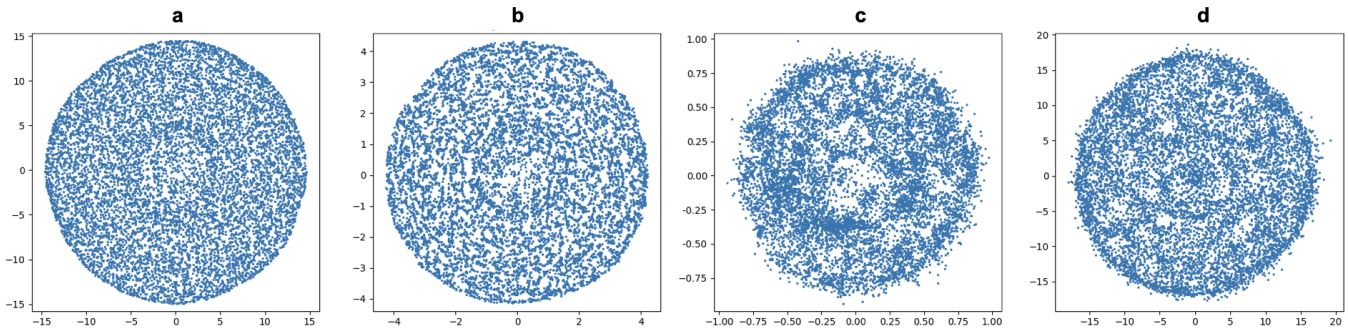

Figure 2: visualization of a sample of size 10000 generated on Hemisphere with non-uniform distribution in dimension 2 by IsUMap without subtracting the distance to the nearest neighbor (ρ_i), with $k = 30$ and varying the t-conorm: Algebraic sum (a), Canonical (b), Bounded sum (c), Drastic sum (d).

II) **Impact of varying function Φ converting metrics to weights.** As mentioned in the main text and in the "Combining metric space" part, the choice of the function ϕ to transform distances to fuzzy weights is arbitrary, as long as it fulfills the requirements on monotonicity and range. While the default IsUMap algorithm uses $\exp(.)$, we generate different such monotone functions ϕ by simply taking $\phi = 1 - F$, where F is a cumulative distribution function. We apply IsUMap with various Φ functions with different t-conorms on the Swiss roll with hole, Möbius strip, and MNIST datasets to study the effect of using different such functions on the the resulting visualizations. Figure 3 shows the results of applying IsUMap with the canonical t-conorm. Since all Φ functions are equivalent in this case, varying the Φ function does not affect the IsUMap results.

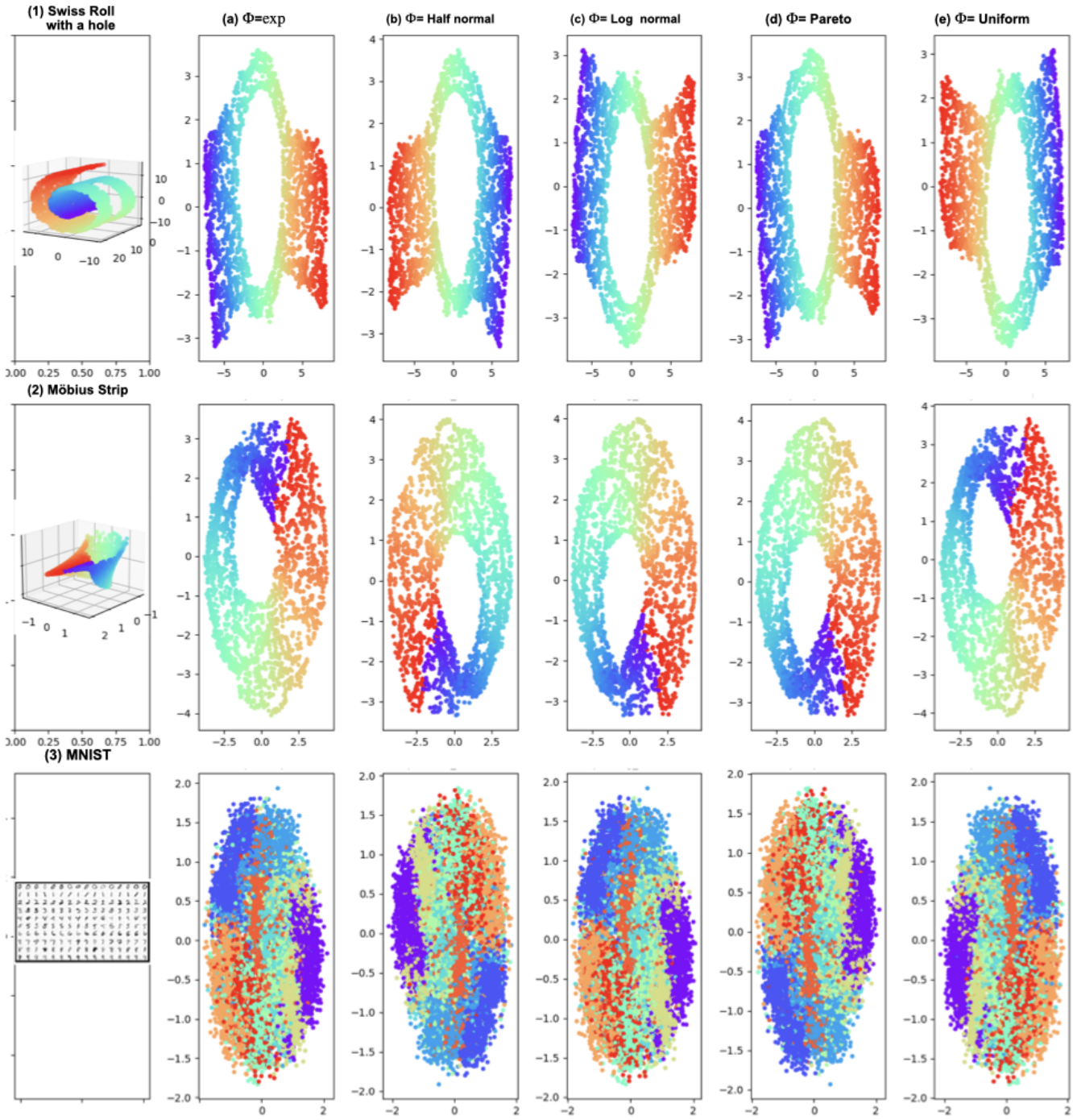

Figure 3: IsUMap with Canonical t-conorm and Varying Φ Functions.

However, as shown in the Figures 5 and 6, using different Φ functions with across Bounded sum and Drastic sum leads to noticeable changes in visualizing the Swiss roll with a hole and Möbius strip. This variation causes the clusters in these datasets to mix. In contrast, Figures 4 and 5 show IsUMap with the Probabilistic sum and Bounded sum, along with $\Phi = \exp$, $\Phi =$ Half-normal and $\Phi =$ Pareto improves the clustring task specifically in the MNIST dataset.

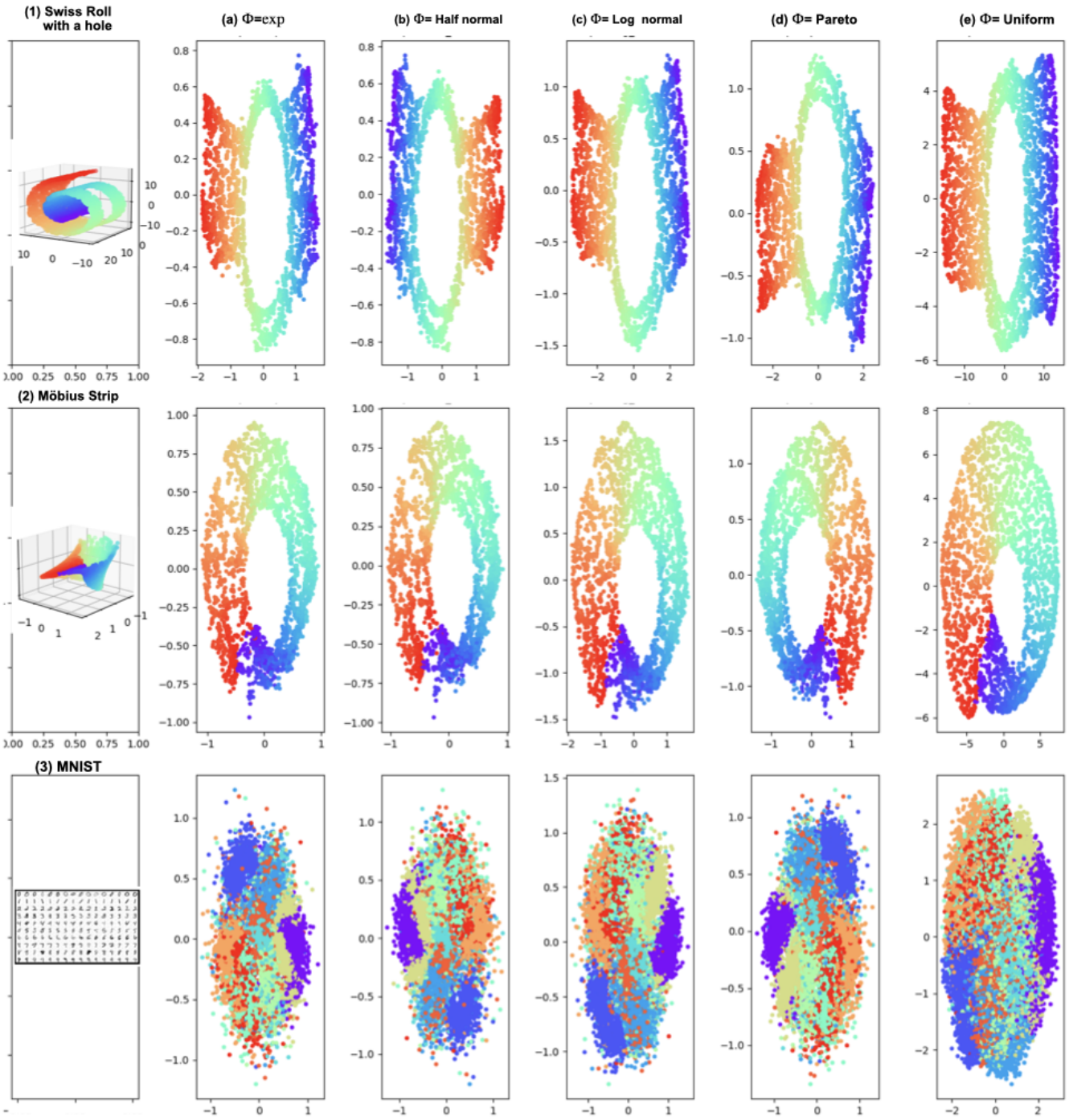

Figure 4: IsUMap with Probabilistic sum and Varying Φ Functions.

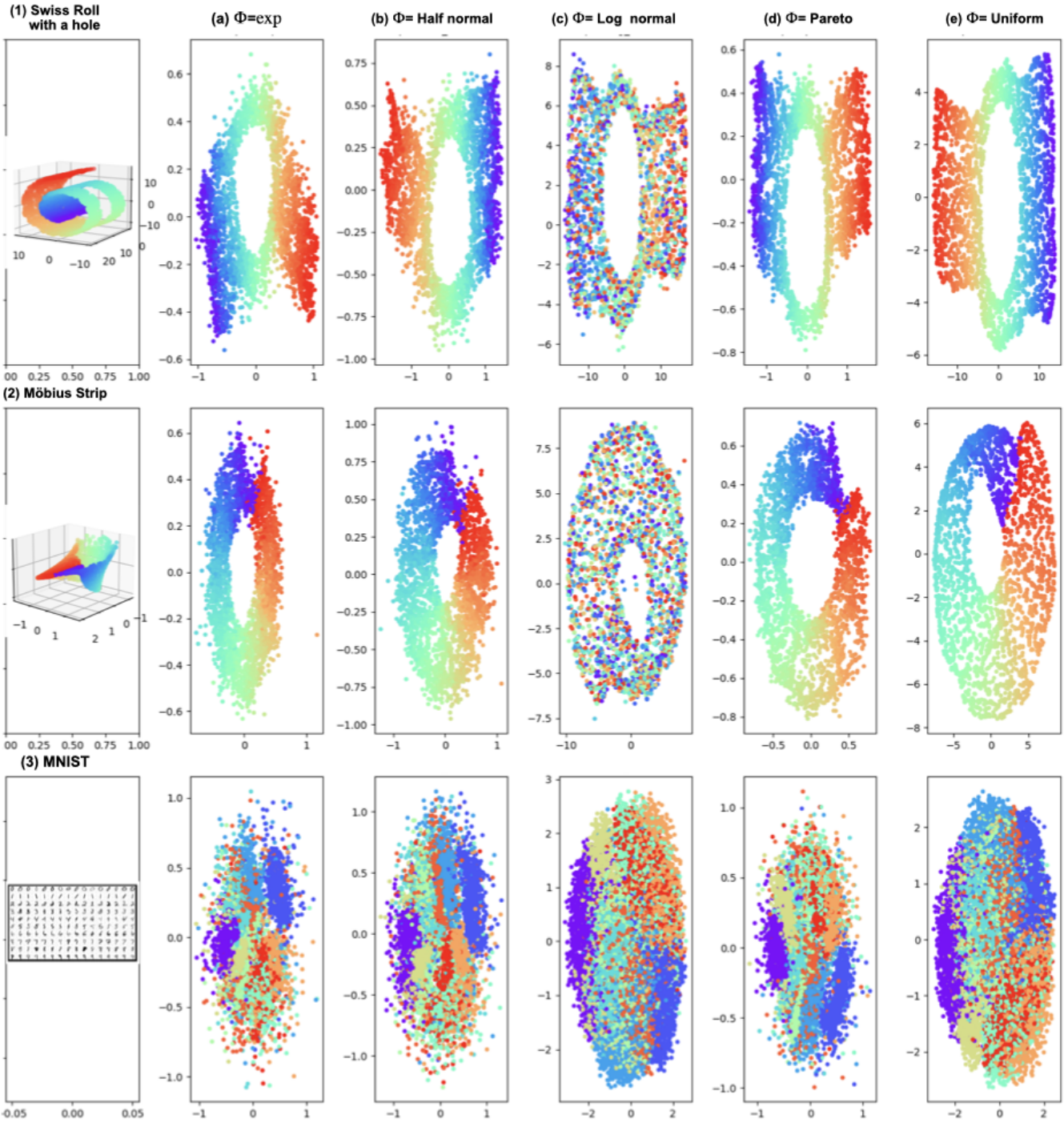

Figure 5: IsUMap with Bounded sum and Varying Φ Functions.

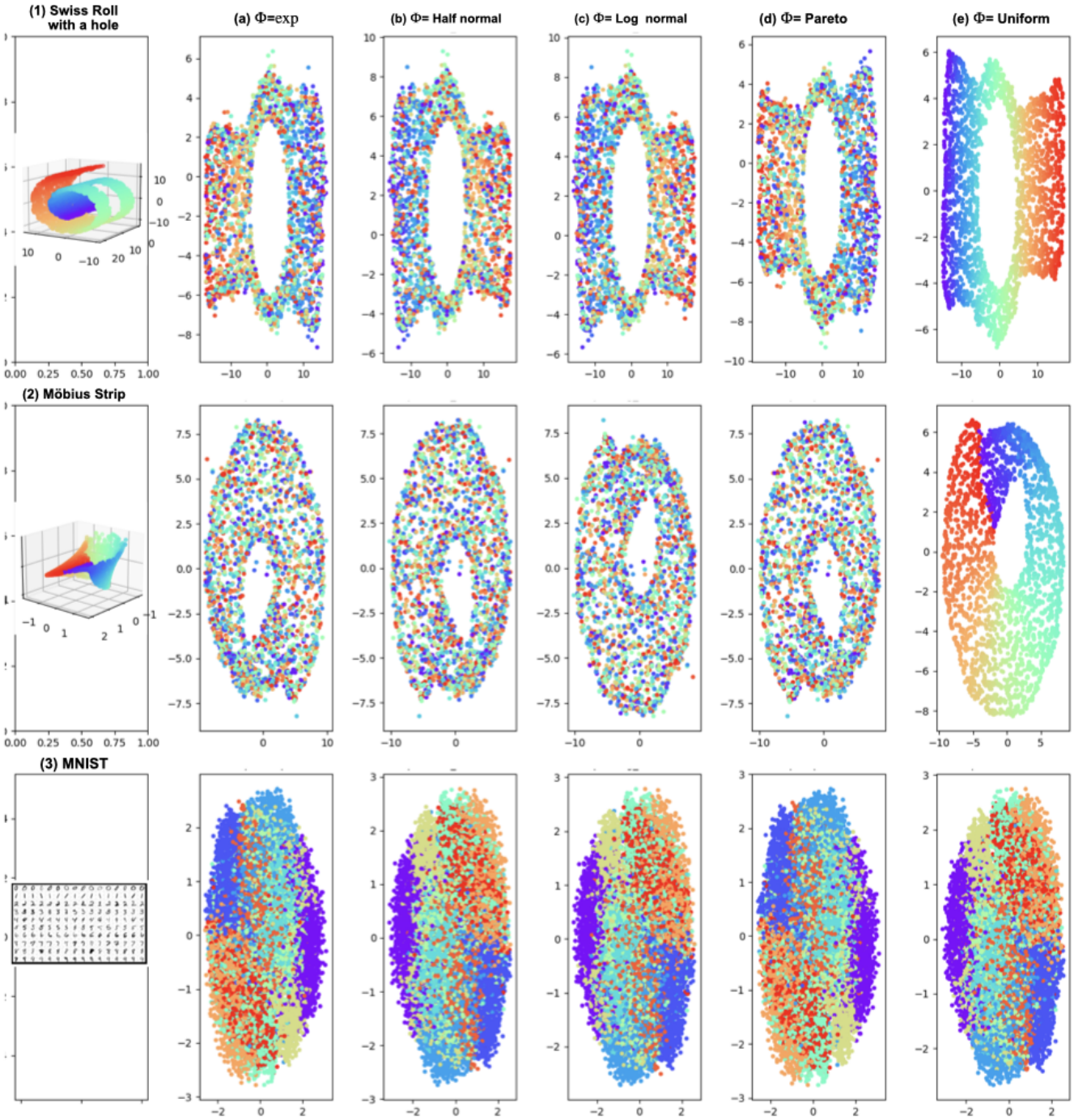

Figure 6: IsUMap with Drastic sum and Varying Φ Functions.

III) **Varying k.** The parameter k, which controls the construction of the graph where each sample point is connected to its k nearest neighbors, requires careful tuning to achieve a satisfactory embedding. As shown in Figure 7, IsUMap is sensitive to the hyperparameter k when the dataset has a geometric structure, such as a Swiss Roll. By increasing the number of nearest neighbors, IsUMap attempts to fold the Swiss Roll and Swiss Roll with a hole, thereby preserving the extrinsic geometry. This sensitivity does not affect the results for the Möbius strip dataset; however, for the MNIST dataset, increasing the number of neighbors improves the clustering performance.

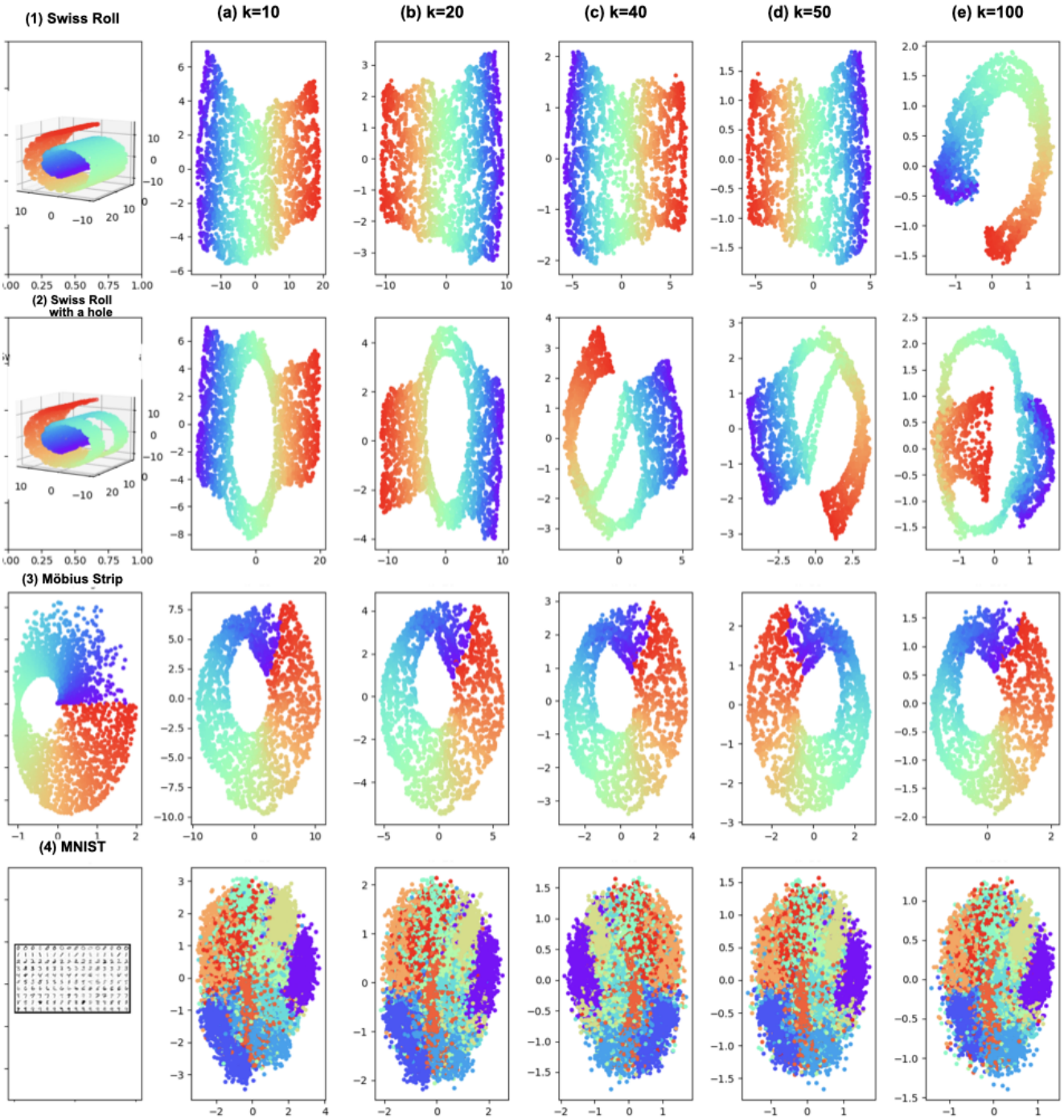

Figure 7: IsUMap with different k

VI) **Impact of initial geometry.** To consider how local distortion changes when altering the geometry of a shape, we demonstrate a Swiss Roll by further rolling up the 2D plane. Figure 8, shows the results of IsUMap with varying numbers of nearest neighbors for each column. We observe that for up to 30 neighbors, IsUMap attempts to unroll the Swiss roll, preserving the intrinsic geometry. By increasing the number of nearest neighbors, it rolls the Swiss roll up again to preserve the extrinsic geometry. IsUMap performs well in preserving the clusters.
In contrast, Figure 9, shows the results of UMAP. With a lower number of neighbors, UMAP tries to unroll the Swiss roll but fails due to the discontinuous nature of the underlying structure, making it ineffective in representing the data in low

dimensions. For a higher number of neighbors, UMAP tends to mix points together, thereby rolling the Swiss roll and failing to preserve clusters and underlying structures. Additionally, we observe a hole in the UMAP results, indicating a change in the topology of the dataset.

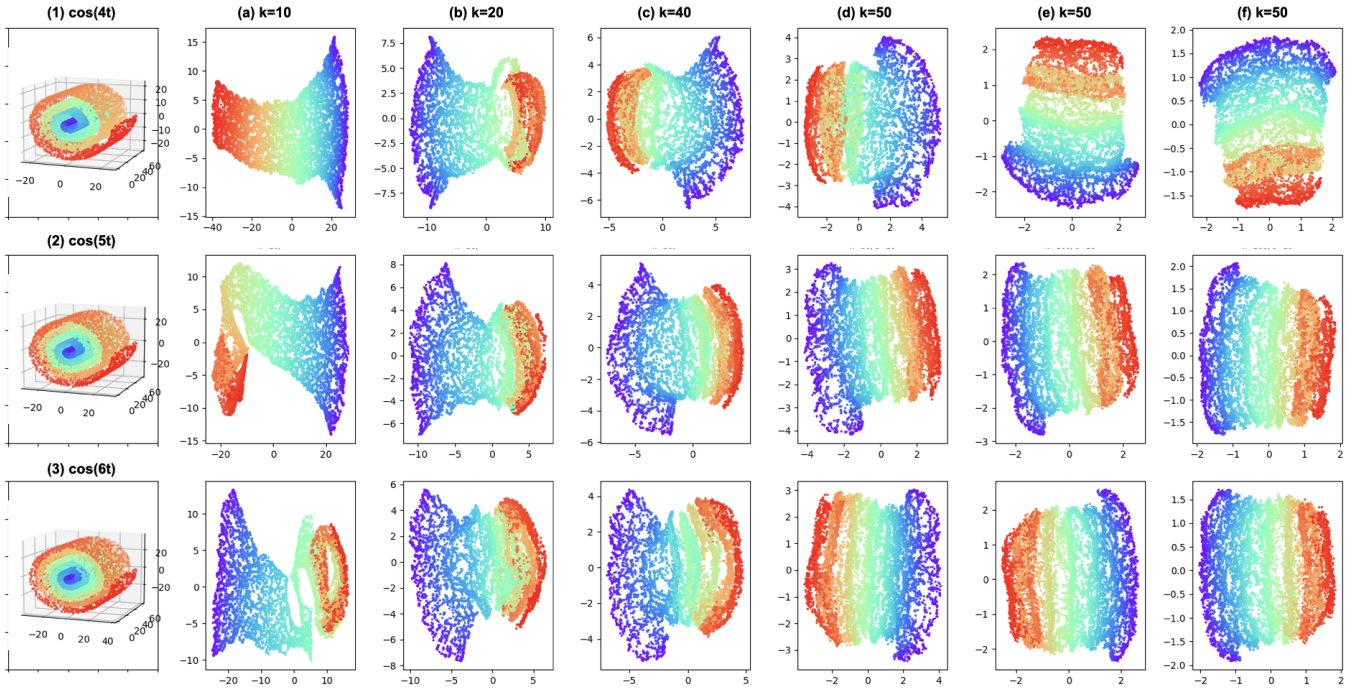

Figure 8: Swiss Roll IsUMap embeddings with varying degrees of tightness in the roll and different k.

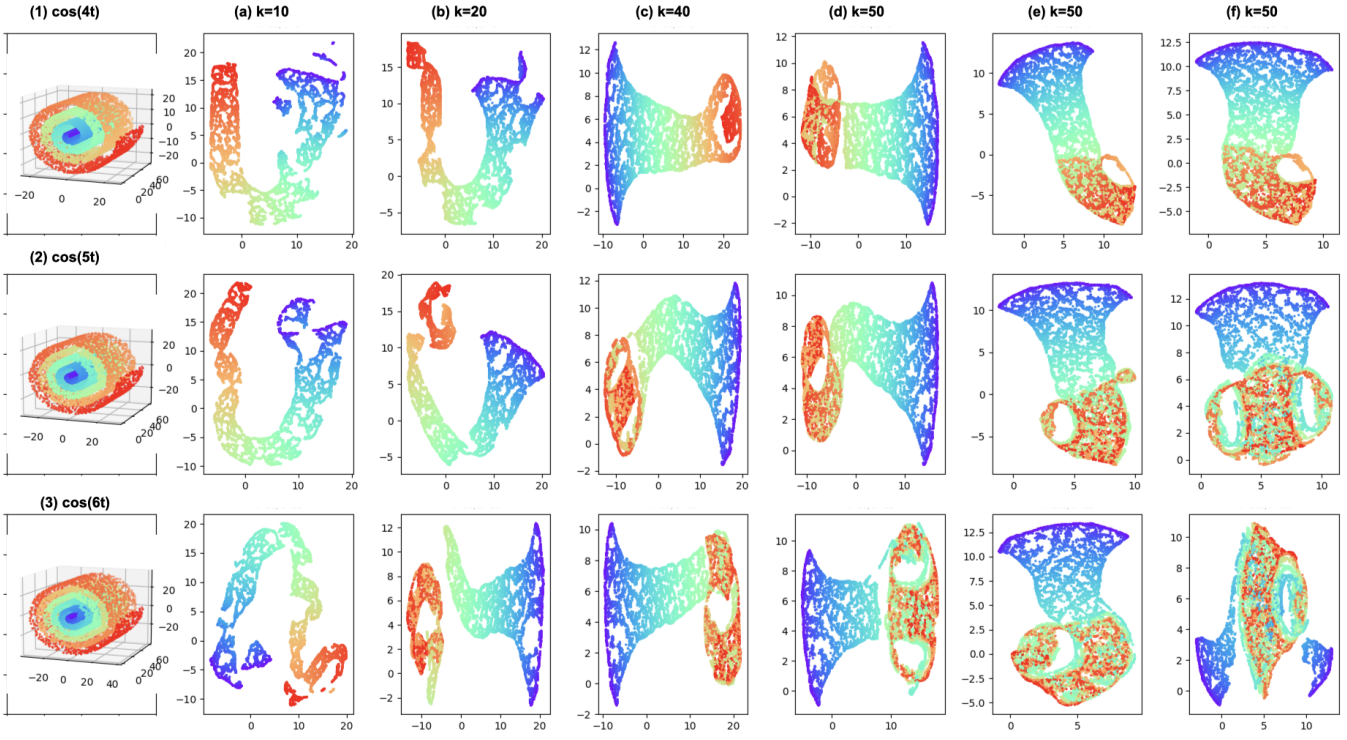

Figure 9: Swiss Roll UMAP embeddings with varying degrees of tightness in the roll and different values of k.

E.2 Empirical examples

I) **Trefoil-knotted protein chains.** We are including the remaining illustrations from the analysis on trefoil-knotted protein dataset, that is, dataset labeled by homology classes employing Wasserstein and labeled by depth category using L_1 distance on their persistence diagrams and persistence landscapes, respectively. In Figure 10, we show the visualization of the dataset with labels specifying the structural homology classes and both distances, after applying IsUMap and UMAP. As emphasized in the main text, we observe that each of UMAP and IsUMap yield similar visualizations across both initial representations, persistence diagrams (using the Wasserstein distance) and persistence landscapes (using the L_1 distance). The same fact is observable in the visualizations of the dataset labeled with depth category using UMAP and IsUMap, as depicted in Figure 11.

Notably, in both cases, whether the dataset is labeled by homology class or depth category, UMAP tends to map points on top of each other as it aggressively seeks to separate clusters. In contrast, IsUMap not only preserves the integrity of clusters but also effectively visualizes certain geometric phenomena that we believe to be retained from the initial data representation through to the final visualization.

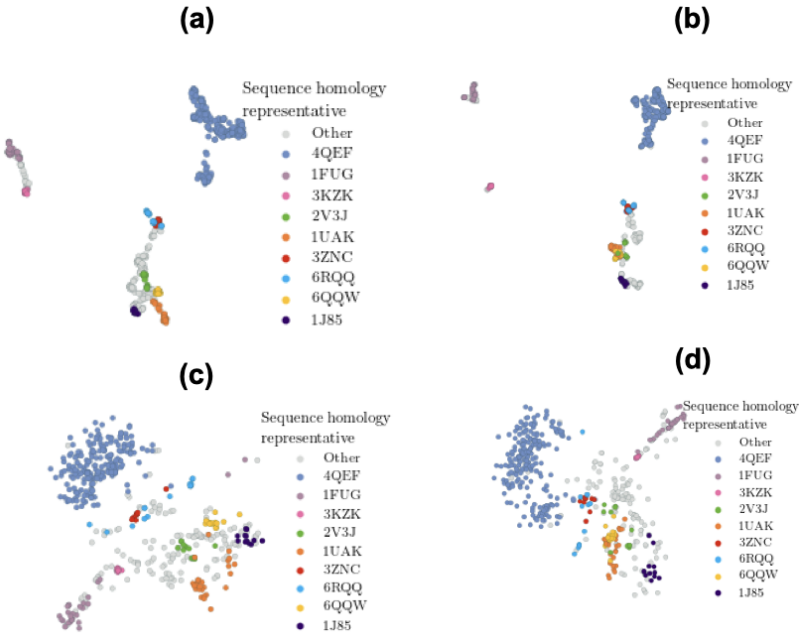

Figure 10: Representations of Trefoil-knotted protein chains, labeled by sequence homology classes, using Wasserstein distances (left) and landscape distance (right). Top: UMAP embedding, Bottom: IsUMap embedding, both with k=15 in neighborhood graph

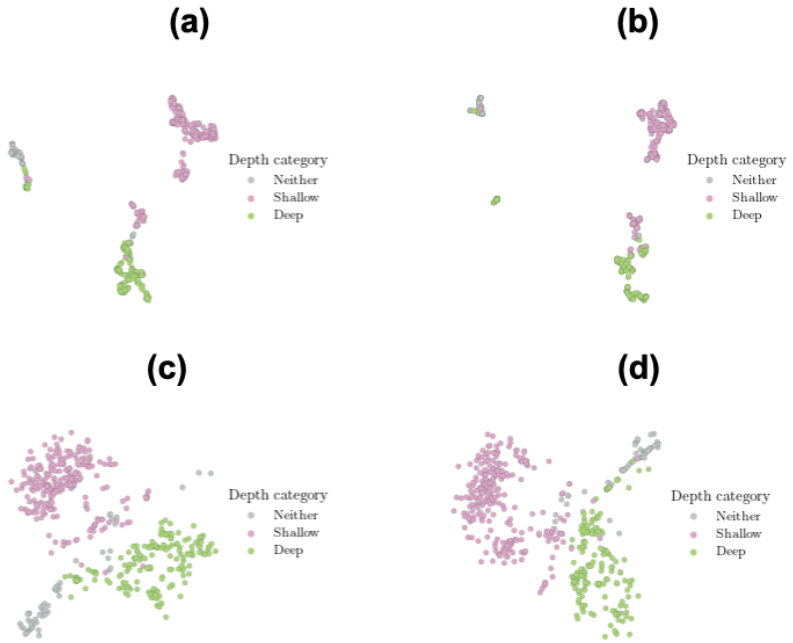

Figure 11: Representations of Trefoil-knotted protein chains, labeled by depth category , using Wasserstein distances (left) and landscape distance (right). Top: UMAP embedding, Bottom: IsUMap embedding, both with k=15 in neighborhood graph

II) **Trajectory inference from single-cell RNA data.** In the example of the single-cell RNA dataset of neural progenitor cells, as highlighted in (Chari and Pachter 2023) and confirmed by our comparison in Subsection 3.1 of the main text, UMAP exhibits limitations in representing continuous relationships, thereby failing to provide a reliable foundation for trajectory inference as a downstream task. Contrary to the explanation proposed in (Chari and Pachter 2023), we attribute this shortcoming primarily to UMAP's reliance on negative sampling rather than local distortions (Damrich and Hamprecht 2021).

To investigate how neighborhood definition impacts the continuity of trajectories, we increased the number of neighbors k in both UMAP and IsUMapand compared the results of applying RNA velocity to the resulting representations. The outcomes, depicted in Figure 12, show that this adjustment leads to improved continuity in the underlying structures and more accurate trajectories after performing IsUMap, making it a stable pre-processing step for this task. However, even with adjusted hyperparameters like the number of neighbors, UMAP continues to struggle with preserving trajectory inference and continuous structures.

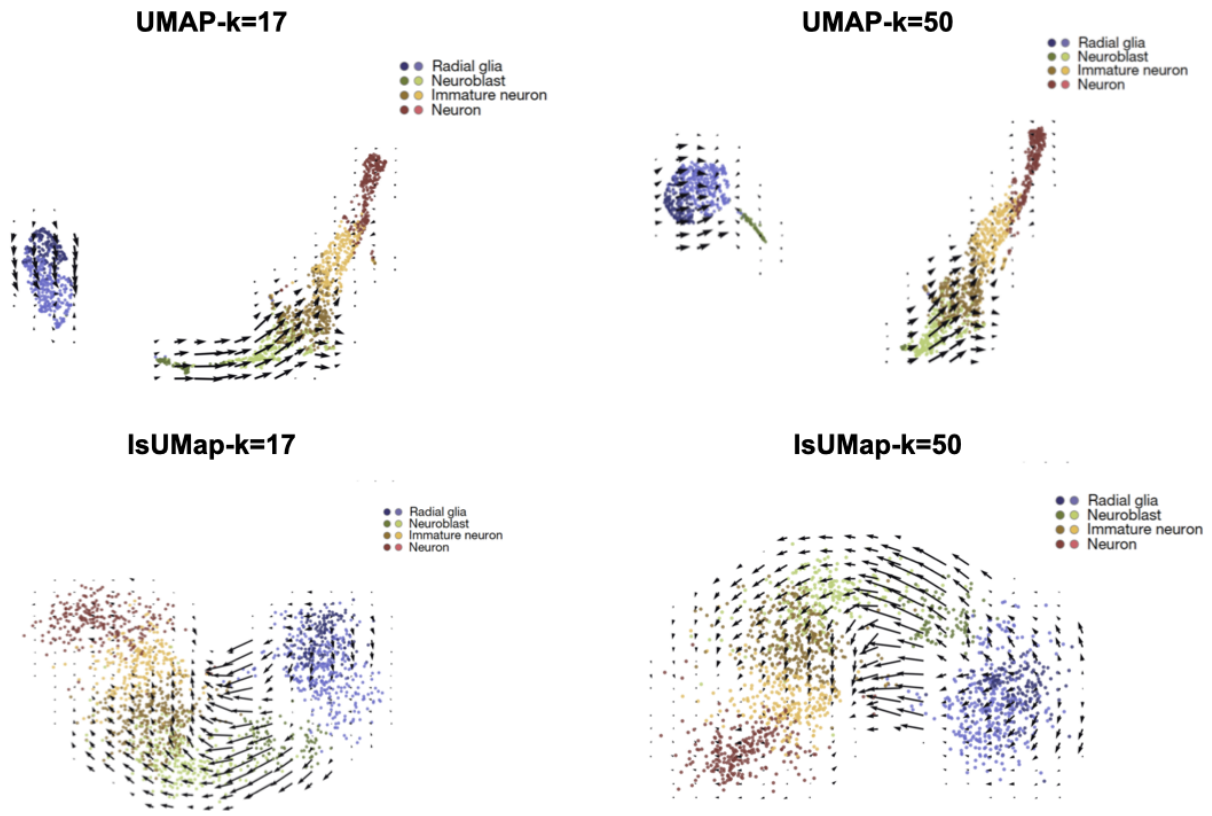

Figure 12: Trajectory inference by RNA velocity method on neural progenitor cells after dimensionality reduction performed by UMAP and IsUMap, with $k = 17$ and $k = 50$.

III) **Topology inference in gridcells dataset.** In Subsection 3.2 of the main text, we extended the work of (Gardner et al. 2022) by analyzing neural recordings of grid cells to identify the topology of population codes. This dataset is interesting because on the one hand, it consists of high dimensional, real biological data, while on the other hand there is a clear expectation to observe a low dimensional manifold. Indeed, grid cells are neurons which fire at regularly spaced positions, arranged on a hexagonal lattice. Two different grid cells (from the same module) will fire with a similar pattern, only offset by a phase. This periodicity suggests that the population activity resides on a toroidal manifold. Indeed, with advancements allowing for the simultaneous recording of firing patterns from a greater number of neurons, (Gardner et al. 2022) were able to demonstrate this toroidal structure clearly for the first time. Their method employs persistent (co)homology techniques in TDA to identify two 1-dimensional cocycles, which can be used to parametrize the torus. The more accurately these cocycles are identified, the more efficient the resulting parametrization will be. Our investigation in Subsection 3.2 of the main text focused on how different non-linear dimensionality reduction methods, specifically UMAP and IsUMap, influence the application of persistent homology. Our findings demonstrated that IsUMap, by preserving the intrinsic geometry of the data, provides a low-dimensional representation that effectively supports the application of persistent homology. Notably, the toroidal topology of the data is clearly inferred from the IsUMap representation in three dimensions. In contrast, the dimensionality reduction achieved by UMAP fails to detect or visualize the one-dimensional holes, as they are neither apparent in the persistence diagram nor observable in the 3D visualization.

In this section, we explore the impact of linear dimensionality reduction, such as PCA, as a preliminary step, following the approach of (Gardner et al. 2022). Our objective is to examine the effects of both linear versus nonlinear dimensionality reductions and the uniformization of data distribution. To achieve this, we apply PCA to reduce the dimensionality to 6, as done in (Gardner et al. 2022). Subsequently, we apply UMAP and IsUMap to further reduce the dimension to 3 in one instance, and in another instance, retain the representation in the same dimension (6) to solely investigate the impact of uniformization. We then apply persistent homology to the resulting representations using cosine distance, consistent with the distance metric used for the input data. The corresponding persistence diagrams of the embeddings in dimensions 3 and 6, along with their visualizations in 3D, are shown in Figure 13.

Applying PCA as an initial step before IsUMap results in a less distinct inference of topological features, as observed in the persistence diagram on the second row of Figure 13. This suggests that the linear projection step negatively impacts the preservation of underlying structures, both in the 3D visualization and in the persistence diagrams. Conversely, this

preliminary step improves the outcome of UMAP for visualization and for the downstream task of topology inference when UMAP is further applied to reduce the dimension to 3, as shown in the first row of the figure. However, when the dimension remains at 6, the corresponding persistence diagram fails to indicate the expected toroidal topology.
Therefore, while linear projection of high-dimensional data enhances UMAP's results, IsUMap offers superior preservation of the toroidal structure in a single step. This advantage stems from IsUMap's ability to avoid the negative sampling step, which can disrupt continuous structures after uniformization. Instead, IsUMap preserves these structures by computing the geodesic distance between each pair of points.

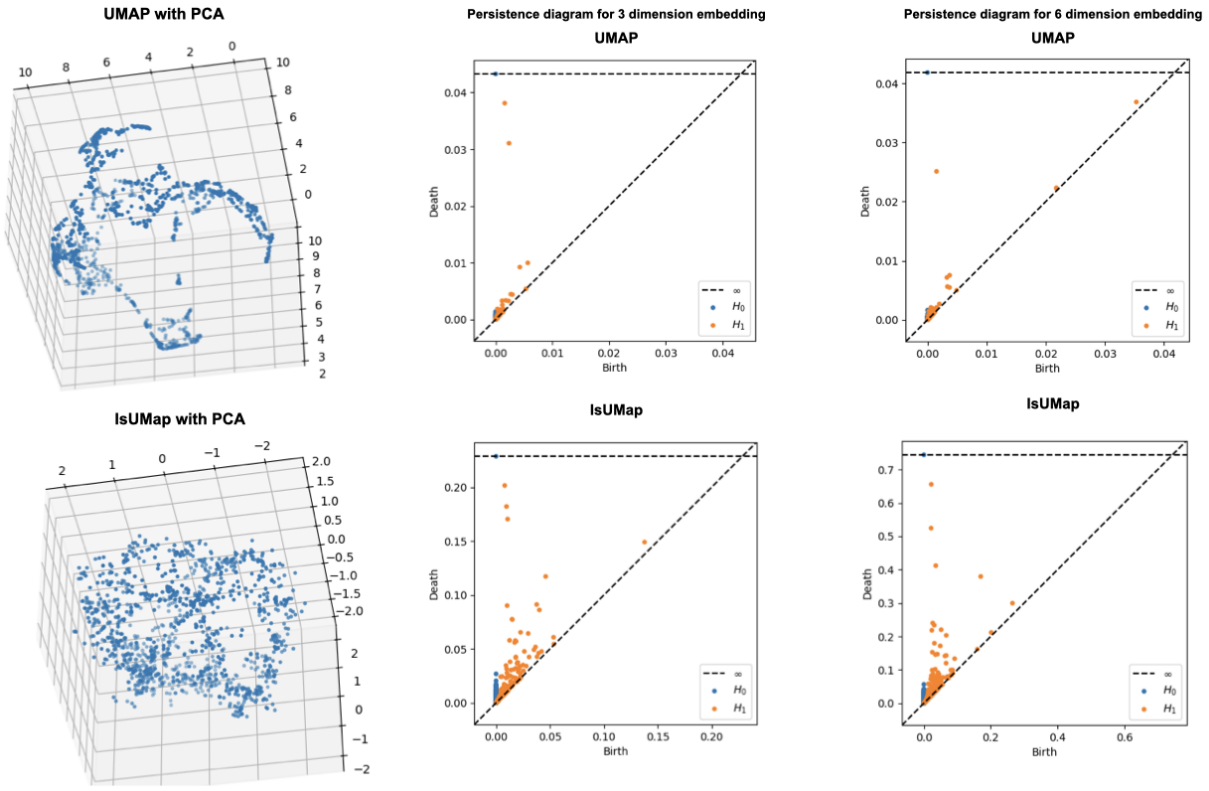

Figure 13: Grid cells dataset, after PCA to 6D and then IsUMap and UMAP to 3D and 6D. The second and third columns display the corresponding persistence diagrams in 3D and 6D using the cosine metric, respectively.

To further explore the sensibility of outcomes to the first linear dimensionality reduction, we change the dimension of PCA from 6 to 15, and repeat the rest of the process as above. That is, after performing PCA to map this dataset to 15 dimensions, we further reduce the dimension to 3 by UMAP and IsUMap and then apply persistent homology to the resulting embedding equipped with cosine distance function. Results, presented in Figure 14, show that as the PCA dimensions increase, UMAP starts to produce more than one connected component. However, IsUMAP maintains consistent results across different PCA dimensions in revealing topological features. Interestingly, UMAP's performance seems to align more with its behavior in the absence of PCA, highlighting the importance of the initial dimensional reduction step for UMAP results.

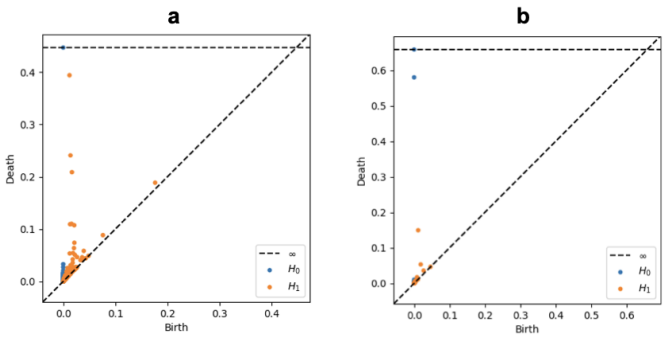

Figure 14: Persistence diagrams of grid cells dataset, after PCA to 15D and then IsUMap (a) and UMAP (b) to 3D.

We further explore whether nonlinear dimensionality reduction offers a more accurate embedding for topology inference compared to linear projection via PCA. Figure 15 presents the persistence diagrams of the grid cells dataset after dimensionality reduction using PCA at three different target dimensions: $3, 6$ and 15. As shown in image (a), the $3-$dimensional PCA embedding produces multiple connected components, making it unsuitable for effective visualization. In 6 dimensions (image b), PCA successfully reveals the toroidal topology. However, in 15 dimensions, although PCA preserves a single connected component (a key topological feature of dimension 0), it also introduces spurious topological features, specifically additional $1-$dimensional holes. As seen in Figure 14, applying UMAP after PCA exacerbates the problem by increasing the number of connected components, thereby worsening the results. In contrast, IsUMap, when applied to the $15-$dimensional PCA embedding, more effectively recovers the toroidal topology, demonstrating its superiority in maintaining the underlying structure of the data.

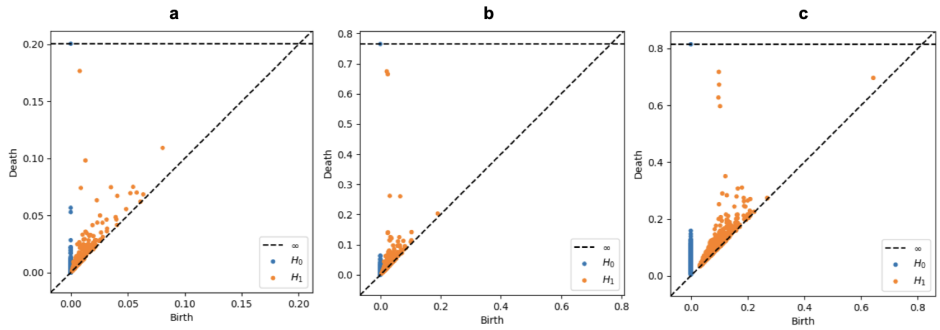

Figure 15: Persistence diagrams of grid cells dataset, after PCA to 3D (a), 6D (b), and 15D(c).

References

Battiston, F.; Cencetti, G.; Iacopini, I.; Latora, V.; Lucas, M.; Patania, A.; Young, J.-G.; and Petri, G. 2020. Networks beyond pairwise interactions: Structure and dynamics. *Physics Reports*, 874: 1–92.

Burago, D.; Burago, Y.; and Ivanov, S. 2001. *A course in metric geometry*. AMS.

Chari, T.; and Pachter, L. 2023. The specious art of single-cell genomics. *PLOS Computational Biology*, 19(8): e1011288.

Damrich, S.; and Hamprecht, F. A. 2021. On UMAP's true loss function. *Advances in Neural Information Processing Systems*, 34: 5798–5809.

Gardner, R. J.; Hermansen, E.; Pachitariu, M.; Burak, Y.; Baas, N. A.; Dunn, B. A.; Moser, M.-B.; and Moser, E. I. 2022. Toroidal topology of population activity in grid cells. *Nature*, 602(7895): 123–128.

Hausmann, J. C. 1995. On the Vietoris-Rips complexes and a cohomology theory for metric spaces. *Annals of Mathematics Studies*, 138: 175–188.

Horak, D.; Maletić, S.; and Rajković, M. 2009. Persistent homology of complex networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2009(03): P03034.

Joharinad, P.; and Jost, J. 2023. *Mathematical Principles of Topological and Geometric Data Analysis*. Mathematics of Data, Springer.

Jost, J. 7th ed., 2017. *Riemannian geometry and geometric analysis*. Springer.

Latschev, J. 2001. Vietoris-Rips complexes of metric spaces near a closed Riemannian manifold. *Archiv der Mathematik*, 77(6): 522–528.

McInnes, L.; Healy, J.; and Melville, J. 2018. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. Https://arxiv.org/abs/1802.03426.

Menger, K.; and Menger, K. 1942. Statisticals. *Proc. Natl. Acad. Sci. USA*, 28, 535.

Myers, A.; Munch, E.; and Khasawneh, F. A. 2019. Persistent homology of complex networks for dynamic state detection. *Physical Review E*, 100(2): 022314.

Simas, T.; Correia, R. B.; and Rocha, L. M. 2021. The distance backbone of complex networks. *Journal of Complex Networks*, 9(6): cnab021.

Spivak, D. I. 2009. Metric realization of fuzzy simplicial sets. *N.A.* Https://math.mit.edu/ dspivak/files/metric_realization.pdf.

van der Maaten, L.; and Hinton, G. 2008. Visualizing Data using t-SNE. *Journal of Machine Learning Research*, 9(86): 2579–2605.

Wallace, A. H. 2007. *Algebraic topology: homology and cohomology*. Courier Corporation.

Zadeh, L. A. 1965. Fuzzy sets and systems. *System Theory (J. Fox ed.). Brooklyn, NY: Polytechnic Press*, 29–37.