Gromov-Hausdorff Interleaving Stability for Autoencoders

Maria Walch
Matthias Neumann-Brosig

Abstract

We prove an upper bound for the interleaving distance between persistence homology modules for certain simplicial complexes constructed on the input and latent spaces of an Autoencoder. We bound the interleaving distance between the Autoencoder input and output space by the maximum pairwise reconstruction loss. We use our result to investigate whether the topological Autoencoder ([Moo+19]) guarantees small interleaving distances between the input and latent persistence homology modules for all homology dimensions. We prove an upper bound for the interleaving distance between $n+1$ dimensional input and latent persistence homology modules when n dimensional creator and destroyer simplices are preserved via the topological loss term.

1 Introduction

Deep Autoencoders fall into the regime of self-supervised [MM20] deep [CC07] neural network architectures and enjoy a broad range of application. By the architecture of a neural network, we mean the arrangement of *neurons* into *layers* under given *hyperparameters*. A layer in a neural network is a function $\mathbf{l} : \mathbb{R}^d \to \mathbb{R}^k$ which transforms a d-dimensional input vector $\mathbf{i}$ of real numbers to a k-dimensional real output vector $\mathbf{o}$. This transformation is performed by component-wise application of the neurons as the basis of the network's calculation units.

Definition 1.1 (Neuron)**.** Let $\mathbf{l}$ be a given neural network layer whose input vector is given by $\mathbf{i}$ and output vector by $\mathbf{o}$, with components $i_l \in \mathbb{R}$, $l \in \{0, d-1\}$ and $o_s \in \mathbb{R}$, $s \in \{0, k-1\}$ respectively. Let b_s be the components of the bias vector $\mathbf{b}$ for layer $\mathbf{l}$ and $\phi : \mathbb{R} \to \mathbb{R}$ a nonlinear continuous activation function. Let $\mathbf{W} \in \mathbb{R}^d \times \mathbb{R}^k$ be the inter-layer weight matrix with real entries W_{ls} for each component i_l and o_s. A neuron is a function $\mathfrak{n} : \mathbb{R}^d \to \mathbb{R}$, $\mathbf{i} \to o_s$ defined as:

$$o_s^t = \phi\left(\sum_l i_l^t W_{ls} + b_s\right). \tag{1}$$

The input to an Autoencoder architecture denotes the input vector to its very first layer and likewise the output of an Autoencoder architecture is given by the output of its last layer. For a given *error* metric, the *reconstruction error* δ of an Autoencoder is defined as the error metric distance between the input and the output, which both lie in $\mathbb{R}^n$, for some $n \in \mathbb{N}$. An Autoencoder is *trained* [Lar+09] to minimize the reconstruction error between its input and output.

1.1 Autoencoder Bi-Lipschitz Continuity

Let $X \subseteq \mathbb{R}^n$ be a dataset consisting of points $\{x_1, ..., x_N\}$, to be the input to (domain of) a given Autoencoder architecture. For our purposes, we do not work with the above mentioned neuron- and layerwise description, but we consider the map associated to this architecture, $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^n$. The output (image) of $\mathcal{A}|_X$ is a set of points $X' \subseteq \mathbb{R}^n$, $X' = \{x'_1, ..., x'_N\}$ of equal cardinality N. The Autoencoder consists of the consecutive application of the Encoder function $\mathcal{E} : \mathbb{R}^n \to \mathbb{R}^m$ and the Decoder function $\mathcal{D} : \mathbb{R}^m \to \mathbb{R}^n$ to the input X such that, for $x_i \in X$, $x'_i \in X'$, we have that $x'_i = \mathcal{D} \circ \mathcal{E}(x_i)$. The codomain of the Encoder $\mathcal{E}$ (domain of the Decoder $\mathcal{D}$) is called the latent space $Y \subseteq \mathbb{R}^m$, $Y = \{y_1, ..., y_L\}$. Apart from unusual architectures such as overcomplete [VP21] Autoencoders, m is chosen such that $m < n$. As subsets of the Euclidian spaces $\mathbb{R}^n$ or $\mathbb{R}^m$, X, X' and Y canonically inherit the Euclidian metric by restriction. Thus the input, output and latent spaces are metric spaces (X, d_X), $(X', d_{X'})$ and (Y, d_Y).
We make the following assumption:

Assumption 1.1. The Encoder, the Decoder and thus the Autoencoder are bijections, when restricted to X, X' and Y. Thus we have that $N = L$.

Corollary 1.1. From Assumption 1.1 follows that $\mathcal{A} = \mathcal{D} \circ \mathcal{E}$ is continuous when restricted to X and X', $\mathcal{E}$ is continuous when restricted to X and Y.

Corollary 1.2. The Lipschitz bolicities of $\mathcal{E}$, $\mathcal{D}$ and thus $\mathcal{A} = \mathcal{D} \circ \mathcal{E}$ are finite.

Proof. As all the involved sets are finite and the respective maps bijective, for every pairwise distance in X, X' or Y there exists a pairwise distance in the domain of the respective map. By finiteness, there exists a maximal and a minimum pairwise distance in each. Thus the minimum distance in the domain of $\mathcal{A}(X)$ can be maximally dilated to the maximal distance in the codomain of $\mathcal{A}(X)$. Symmetrically, this argument applies to the compression factor of the maximum distances in X and $\mathcal{A}(X)$. □

From Corollaries 1.1 and 1.2 follows that $\mathcal{E}$, $\mathcal{D}$ and $\mathcal{A}$ are bi-Lipschitz homeomorphisms when restricted to X, Y and X'.

1.2 Autoencoder Bi-Lipschitz Constant

Definition 1.2 (Reconstruction Error)**.** Let $\{(x, x') \in X \times X' \mid x' = \mathcal{A}(x)\}$. The pairwise reconstruction error is given by the Euclidian metric distance $d(x, \mathcal{A}(x))$. It has a maximum element $\delta_M = \max(d(x, \mathcal{A}(x)))$.

An estimate for the bi-Lipschitz constant $L_{\mathcal{A}}$ of $\mathcal{A}$ using the maximum reconstruction error δ_M can be obtained using the triangle inequality. Let $x_1, x_2 \in X$ and $x_1', x_2' \in X'$.

$$d(x_1', x_2') \leq d(x_1', x_1) + d(x_1, x_2) + d(x_2, x_2') \leq d(x_1, x_2) + 2 \cdot \delta_M. \tag{2}$$

Let $\eta_m = \min\{d(x_1, x_2) \mid x_1, x_2 \in X, x_1 \neq x_2\}$. As a result, the Lipschitz constant of $\mathcal{A}$ when restricted to X can be estimated as:

$$L_{\mathcal{A}} \geq \left(1 + \frac{2\delta_M}{\eta_m}\right). \tag{3}$$

1.3 Interleavings in the Autoencoder

In the deep learning literature, the *Manifold Hypothesis* [FMN13] is insinuated. For an Autoencoder archtiecture this means that the input data X is assumed to be sampled from a probability distribution that has support on or near a submanifold χ of the Euclidian space $\mathbb{R}^n$. It is hoped that the Autoencoder iteratively *learns* this submanifold structure over the training process. The hypernym *manifold learning* comprises estimating the geometric and/or topological properties of the submanifold under consideration from the given data points. It has been proved by [NSW08] (Theorem 3.1) that it is possible to *algorithmically* learn the homology of the submanifold χ with high confidence also for noisy data (Theorem 7.1) , where the data lies *near* rather than *on* χ. However, this result is not helpful for the question whether the above mentioned hope for the Autoencoder is justified. The problem is, that it is not possible to access χ. Which structure does it carry? - Is it smooth, piecewise linear, analytic etc.? Apart from this, even if we could be sure that the input data exhibits a (sub)manifold structure, the hyperparameters of the Autoencoder are chosen a priori according to expert knowledge, prima facie transferability of already known dataset properties and lucky guessing. Thus the Encoder is not constructed to constitute a manifold embedding into the latent space (Y, d_Y) in the mathematical sense. For example, if χ was smooth, we could yet not guarantee that the Autoencoder architecture fullfills the conditions of the (strong) Whitney embedding theorem.
For this reason, we do not work with the manifold perspective on the data. We construct simplicial complexes on the input data X and the latent data Y and investigate whether and to which extent the simplicial homology is preserved by the Encoder map $\mathcal{E}$ and the Decoder map $\mathcal{D}$. If the Decoder would constitute a homeomorphic reconstruction of a manifold embedding in the latent space, then also the homology groups would be isomorphic by functoriality of homology.

2 Related Work

The work most closely related to ours is [NL22], which also uses the results of [CSO12]. In [NL22], it was shown that the interleaving distance between an original data set X and a low dimensional embedding Y can be used to optimize the embedding of Y. Furthermore, a scale at which topological features in X and Y are in correspondence was determined using the interleaving distance. The bottleneck distance between persistence diagrams of X and its low dimensionaly embedding Y is also bounded via the Gromov-Hausdorff distance in [NL22]. The idea of bounding the bottleneck distance between persistence diagrams of Rips filtrations built on top of finite metric spaces was also used in [Cha+09].

3 Contributions

This work presents a novel approach to topologically quantifying the generalization capability of Autoencoder architectures via the interleaving distance.

1. In Section 7.1 we give an upper bound for the interleaving distance between the input and latent persistence homology modules depending on the Autoencoder bi-Lipschitz constant and the diameter of the input space X.

2. We show that the bottleneck and thus the interleaving distance between input and output persistence diagrams / persistence homology modules can be bounded by the maximum pairwise reconstruction error in Section 7.2.

3. Our results indicate that the distortion of a given correspondence has to be controlled in order to obtain a low interleaving distance between input and latent persistence homology modules. We use this insight to investigate whether the topological Autoencoder, a novel Autoencoder architecture introduced by [Moo+19] guarantees small interleaving distances between input and latent persistence homology modules for all homology dimensions in Section 8.

4 Filtered Input/Output/Latent Complexes

To the set of points X, X' and Y we associate objects K_X, $K_{X'}$ and K_Y of the category **SCpx** of finite abstract simplicial complexes such that the vertex sets of K_X, $K_{X'}$ and K_Y are given by the elements of X, X' and Y respectively. Given a feature scale parameter ρ, the Cěch- and the Rips-construction are two natural ways to construct such an abstract simplicial complex K on a given metric space (M, d_M).

Definition 4.1 (Cěch-Complex)**.** Let $\overline{B}_{\rho/2}(p)$ be the closed ball around $p \in M$ with radius $\rho/2$. Then the Cěch complex $\mathcal{C}(M, \rho)$ is the abstract simplicial

complex whose r-simplices σ_r are given by the sets of $r+1$ points $\{p_k\}_{0\leq k\leq r} \in M$ such that:

$$\cap_{k=0}^{r} \overline{B}_{\rho/2}(p_k) \neq \emptyset. \tag{4}$$

Definition 4.2 (Rips-Complex)**.** The Rips complex $\mathcal{R}(M,\rho)$ is defined as the abstract simplicial complex whose r-simplices correspond to the power sets $P_r(M) = \{M_r \subseteq M : |M_r| = r+1\}$ of $r+1$ points in $M \subset \mathbb{R}^d$ with diameter at most ρ :

$$\mathcal{R}(M,\rho) = \{\sigma | \emptyset \neq \sigma \subset M, \mathbf{d}(\sigma) \leq \rho\}, \tag{5}$$

where $\mathbf{d} : M_r \subseteq \mathbb{R}^d \to \mathbb{R}$ denotes the diameter function of M restricted to M_r.

The category of filtered simplicial complexes, **FSCpx**, is given by the subcategory of the functor category $\mathbf{SCpx}^{\mathbb{I}}$, such that the simplicial morphism corresponding to every morphism $(i,j) \in \mathbb{I}$ is a simplicial injection. In the following, we will consider $\mathbb{I}$ to equal the real numbers $\mathbb{R}$, regarded as a poset category.
An object of **FSCpx** is called a filtered simplicial complex. This is also called a filtration $\mathcal{F}$ and is given by a nested family of simplices built on top of a set $S \subset M$. That is, a filtered simplicial complex is a family $(K_a)_{a\in\mathbb{R}}$ of subcomplexes of some fixed simplicial complex K such that $K_a \subseteq K_b$ whenever $a \leq b$. To come back to the above mentioned examples of the Rips- and the Cěch complex, the *Rips filtration* for $\rho \in \mathbb{R}$ is given by:

$$[p_1, ..., p_m] \in \mathcal{R}(M,\rho) \Longleftrightarrow d_M(p_i, p_j) \leq \rho \quad \forall i,j. \tag{6}$$

Herein, $[p_1, ..., p_m]$ denotes the subsimplex spanned by $\{p_1, ..., p_m\} \in M$, while $\mathrm{Rips}(M,\rho)$ is the associated Rips subcomplex. Analogously, the radii of the minimum enclosing balls around a point $p \in M$ can be used as a filtration parameter for the Cěch complex.

5 Simplicial Homology and Interleavings

We use simplicial homology with coefficients in a field $\mathbf{k}$. The following presentation follows Section 2 in [CSO12] and Sections 3 and 4 in [YG20a] as well as Chapter 3 in [Oud15].

5.1 Simplicial Homology and Persistence Modules

In general, a persistence module is a functor of a poset category $\mathbb{I}$ into an abelian category $\mathcal{C}$. In the following, we will only consider a special type of persistence modules obtained by applying the homology functors of dimension n to the sequence of subcomplexes in an $\mathbb{R}$-filtered simplicial complex K whose vertices are in $\mathbb{R}^d$. Thus we will work with the following definition:

Definition 5.1 (Persistence Module)**.** A persistence module $\mathbb{V}$ over the real numbers is an $\mathbb{R}$-indexed family of vector spaces $(V_\mathfrak{r})_{\mathfrak{r}\in\mathbb{R}}$ together with a doubly-indexed family of linear maps $(v_a^b : V_a \to V_b |\ a \leq b, a, b \in \mathbb{R})$ satisfying the

composition law $v_b^c \circ v_a^b = v_a^c$ whenever $a \leq b \leq c \in \mathfrak{r} \subset \mathbb{R}$ and where v_a^a is the identity map on V_a.

Let $K_\mathfrak{r}$ be the subcomplexes corresponding to filtration parameters $\mathfrak{r} \in \mathbb{R}$ and $\iota_\mathfrak{r}$ be the inclusion maps between them:

$$K_a \xhookrightarrow{\iota_a} K_b \xhookrightarrow{\iota_b} ... \qquad K_g \xhookrightarrow{\iota_g} K_t, \tag{7}$$

for $a, b, ..., g, t$ some real filtration parameters in $\mathbb{R}$. Let $\epsilon_{\text{functor}}$ denote the shift functor in our preorder category $\mathbb{R}$ used for filtration: $\epsilon_{\text{functor}} : \mathbb{I} \to \mathbb{I}$, $a \to a + \epsilon$ for $a, \epsilon \in \mathbb{R}$ and $\epsilon > 0$. From this we get a shift functor $\Sigma^\epsilon : \mathbf{FSCpx} \to \mathbf{FSCpx}$ on $\mathbf{FSCpx}$ via $\Sigma^\epsilon = \text{Hom}(\epsilon_{\text{functor}}, -)$. Then Σ^ϵ sends an element K_a in $\mathbf{SCpx}$ to $K_{a+\epsilon}$. Application of the (n-dimensional) simplicial homology functor H_* to Eq. (7) leads to the following persistence module:

$$H_*(K_a) \xhookrightarrow{\iota_{a,*}} H_*(K_b) \xhookrightarrow{\iota_{b,*}} ... \qquad H_*(K_g) \xhookrightarrow{\iota_{g,*}} H_*(K_t). \tag{8}$$

By functoriality, a shift functor is defined for the terms $H_*(K_\mathfrak{r})$ appearing in the above sequence via composition $H \circ \Sigma^\epsilon$. The homology of a subcomplex $K_\mathfrak{r}$ is graded, namely $H(K_\mathfrak{r}) = \bigoplus_n H_n(K_\mathfrak{r})$.

Definition 5.2 (n-persistent Homology; Definition 2.6 in [Cas19])**.** Let K be a filtered simplicial complex and $n \geq 0$. The n-th persistent homology of K is defined as the composed functor $H_n(K) : \mathbb{R} \to \mathbf{Vect}$. This will also be denoted by $PH_n(K)$.

In general, the homology of K is approximated by the homology of its subcomplexes $K_\mathfrak{r}$ and their intersections. Using a generalized Mayer-Vietoris sequence, $H_*(K)$ can be seen as the abutment of the respective spectral sequence, i.e. $E^\infty \approx \text{Gr}H_*(K)$ relative to some filtration on $H_*(K)$. Alternatively, the topological information contained in the subcomplexes $K_\mathfrak{r}$ can be glued using the Mayer-Vietoris blowup complex, where the intersections of subcomplexes are explicitely calculated and thence the chain complex of the blowup is evaluated, [LM15]. Since all the spaces involved in the Autoencoder setting contain a finite number of points, a maximum filtration parameter $\bar{\rho}$ exists for which the connectivity stabilizes. Hence, if we choose the respective filtration parameter ρ large enough, the sequence in Eq. (7) will attain its final term $K_t = K$ and thus we get the homology of the whole complex K.

5.2 Metrics for Persistence Modules

To compare two persistence modules, we need to introduce the notion of a *shift map* or *homomorphism of degree* ϵ between persistence modules. For two persistence modules $\mathbb{U}$ and $\mathbb{V}$ over $\mathbb{R}$, a homomorphism of degree ϵ for $\epsilon > 0$ is a

collection Φ of linear maps between the underlying vector spaces, see Definition 5.1:

$$(\phi_a : U_a \to V_{a+\epsilon})_{a\in\mathbb{R}}, \tag{9}$$

such that $v_{a+\epsilon}^{b+\epsilon} \circ \phi_a = \phi_b \circ u_a^b$ for all $a \leq b \in \mathbb{R}$. The collection of all homomorphisms of degree ϵ between $\mathbb{U}$ and $\mathbb{V}$ is defined as $\mathrm{Hom}^\epsilon(\mathbb{U}, \mathbb{V})$.

Definition 5.3 (ϵ-Interleaving)**.** Two persistence modules $\mathbb{U}$ and $\mathbb{V}$ are ϵ-interleaved for a real number ϵ if there are maps

$$\Phi \in \mathrm{Hom}^\epsilon(\mathbb{U}, \mathbb{V}), \quad \Psi \in \mathrm{Hom}^\epsilon(\mathbb{V}, \mathbb{U}) \tag{10}$$

such that $\Psi\Phi = 1_{\mathbb{U}}^{2\epsilon}$ and $\Phi\Psi = 1_{\mathbb{V}}^{2\epsilon}$; where $1_{\mathbb{V}}^{\epsilon} \in \mathrm{Hom}^\epsilon(\mathbb{V}, \mathbb{V})$ is the shift map defined to be the collection of maps $(v_a^{a+\epsilon})$ from the persistence structure (Definition 5.1) on $\mathbb{V}$ and analogously for $1_{\mathbb{U}}^{\epsilon}$.

When we say that two persistence modules $\mathbb{U}$ and $\mathbb{V}$ are ϵ-interleaved, this means the following diagrams are commutative

$$\begin{array}{ccc} U_a & \longrightarrow & U_b \\ & {\scriptstyle \phi_a} \searrow \quad \nearrow {\scriptstyle \phi_b} & \\ & V_{a+\epsilon} \longrightarrow V_{b+\epsilon} & \end{array} \qquad \begin{array}{ccc} & U_{a+\epsilon} \longrightarrow U_{b+\epsilon} & \\ {\scriptstyle \psi_a} \nearrow & & \nearrow {\scriptstyle \psi_b} \\ V_a & \longrightarrow & V_b \end{array} \tag{11}$$

$$\begin{array}{ccc} U_a & \longrightarrow & U_{a+2\epsilon} \\ & {\scriptstyle \phi_a} \searrow \quad \nearrow {\scriptstyle \psi_{a+\epsilon}} & \\ & V_{a+\epsilon} & \end{array} \qquad \begin{array}{ccc} & U_{a+\epsilon} & \\ {\scriptstyle \psi_a} \nearrow & & \searrow {\scriptstyle \phi_{a+\epsilon}} \\ V_a & \longrightarrow & V_{a+2\epsilon} \end{array} \tag{12}$$

For simplicial homology as discussed in Section 5.1, the persistence modules $H_*(K)$ and $H_*(W)$ being ϵ-interleaved requires the diagrams in Eq. (11) and Eq. (12) to be commutative at every place in the sequences Eq. (8) for some filtered complexes K and W and homology dimension n respectively.

Definition 5.4 (Interleaving Distance)**.** The interleaving distance between two persistence modules $\mathbb{V}$ and $\mathbb{U}$ is the infimum of ϵ for which $\mathbb{V}$ and $\mathbb{U}$ are interleaved in the sense of Definition 5.3.

It has the properties of an extended pseudometric on the category of persistence modules [BSS14], and will be denoted as d_I in the following.

5.3 The Bottleneck Distance

The persistence module of Eq. 8, obtained by applying the n-dimensional simplicial Homology functor to the sequence of filtered subcomplexes in Eq. 7, constitutes a one dimensional persistence module by definition. One would obtain

a higher dimensional persistence module if one would consider multiparameter persistence, e.g. by using the density of data points as a second filtration parameter (besides the subcomplex relation over the metric ρ as for the Rips- or Cěch construction). For dimension 1, all persistence modules decompose into interval modules [Cra12], which follows from the Krull-Schmidt Theorem for representations of A_n. Thus persistence modules indexed over $\mathbb{R}$ (such as our one-parameter persistence homology modules) are in correspondence with intervals (a, b) in $\mathbb{R}$ whose collection is called the Barcode diagram dgm($\mathbb{V}$) of a persistence module $\mathbb{V}$. The bars stand for features reflected in the module which "appear" at stage a (birth) and "persist" until stage b and then "disappear" (death). These intervals can also be depicted as points in a diagram whose x-axis corresponds to the birth time of the feature and whose y-axis corresponds to the death time. This diagram is called *persistence diagram* and is the quantity that is computed up to the n-th homology dimension in applications of topological data analysis. Most often, n is not larger than 3. Associated to the persistence diagram diagram is a combinatorically defined metric called the bottleneck distance d_B, ([BL15], [CEH05]).

Definition 5.5 (Bottleneck Distance, Definition 3.1 in [CEH05])**.** Let P and Q be two multisets of points. Points of P with no correspondence in Q (and vice versa) are allowed to be matched to their diagonal projection. Then the bottleneck distance between P and Q is defined as:

$$d_B(P, Q) = \inf_\gamma \sup_p ||p - \gamma(p)||_\infty, \tag{13}$$

where $p \in P$ ranges over all points in P and γ ranges over all bijections from P to Q. Each point with multiplicity k is interpreted as k individual points and the bijection is between the resulting sets.

Definition 5.6 (Q-Tameness)**.** A persistence module $\mathbb{V}$ is called q-tame if

$$r_a^b = \text{rank}(v_a^b) < \infty \quad \text{whenever } a < b \in \mathbb{R}. \tag{14}$$

Interval decomposable persistence modules are q-tame by definition. However, not every q-tame persistence module is also interval decomposable [CCS14]. For one dimensional persistence modules, the bottleneck distance and the interleaving distance are equivalent in the following sense:

Theorem 5.1 (Isometry Theorem; Theorem 4.11 in [Cha+12])**.** Let $\mathbb{U}$ and $\mathbb{V}$ be q-tame persistence modules. Then

$$d_I(\mathbb{U}, \mathbb{V}) = d_B(\text{dgm}(\mathbb{U}), \text{dgm}(\mathbb{V})). \tag{15}$$

6 Interleavings through Correspondences

For simplicial complexes with a filtration dependent on the metric d_M of a metric space (M, d_M), [CSO12] proved that ϵ-*simplicial* ([CSO12], Definition

3.2) *correspondences* between subsets S_1 and S_2 of M, with S_1 and S_2 being the vertex sets of two filtered simplicial complexes K_{S_1} and K_{S_2}, induce a canonical ϵ-interleaving between the persistence modules $H_*(K_{S_1})$ and $H_*(K_{S_2})$ (in the sense of subsection 5.2).

Definition 6.1 (Correspondence between Sets, Definition 7.3.17 in [BBI01])**.** Let S_1 and S_2 be two sets. A *correspondence* between them is defined as a set $R \subset S_1 \times S_2$ satisfying the following condition: For every $q \in S_1$ there exists at least one $v \in S_2$ such that $(q, v) \in R$ and similarly for each $v \in S_2$ there exists a $q \in S_1$ such that $(q, v) \in R$.

Between metric spaces, the notion of a distortion can be defined from the respective correspondences between sets:

Definition 6.2 (Distortion, Definition 7.3.21 in [BBI01])**.** Let R be a correspondence between two metric spaces (M_1, d_{M_1}) and (M_2, d_{M_2}). The distortion of R is defined by:

$$\mathrm{dis}(R) = \sup\{|d_{M_1}(q, q') - d_{M_2}(v, v')| : (q, v), (q', v') \in R\}, \tag{16}$$

with d_{M_1} and d_{M_2} being the metrics of (M_1, d_{M_1}) and (M_2, d_{M_2}) respectively.

This can be used for an alternative definition of the *Gromov-Hausdorff* distance, an extended pseudo-metric for the class of (finite) metric spaces **Met**. The standard definition of the Gromov-Hausdorff distance is the following:

Definition 6.3 (Gromov-Hausdorff Distance, Definition 7.3.10 in [BBI01])**.** Let (M_1, d_{M_1}) and (M_2, d_{M_2}) be metric spaces. Let $r > 0$, then the Gromov-Hausdorff distance $d_{GH}(M_1, M_2) < r$ if and only if there exists a metric space M_3 and subspaces M_1' and M_2' of it, which are isometric to M_1 and M_2 respectively and such that $d_H(M_1', M_2') < r$. In other words, the Gromov-Hausdorff distance is defined as the infimum of positive r for which the above M_3, M_1' and M_2' exist. The distance d_H denotes the ordinary Hausdorff distance between subsets of M_3.

Using correspondences, the Gromov-Hausdorff distance can be defined in the following way:

Definition 6.4 (Gromov-Hausdorff Distance, Theorem 7.3.25 in [BBI01])**.** For any two metric spaces (M_1, d_{M_1}) and (M_2, d_{M_2}),

$$d_{GH}(M_1, M_2) = \frac{1}{2}\mathrm{inf}_R(\mathrm{dis}(R)), \tag{17}$$

where the infimum is taken over all correspondences R between M_1 and M_2. Thus, the Gromov-Hausdorff distance is equal to the infimum of $r > 0$ for which there exists a correspondence between M_1 and M_2 with $\mathrm{dis}(R) < 2r$.

For the Vietoris-Rips- and the Cěch complex, [CSO12] deduce the following two Lemmata 6.1 and 6.2, which depend on the Gromov-Hausdorff distance d_{GH}.

Lemma 6.1 (Vietoris-Rips Interleaving; Lemma 4.3 in [CSO12])**.** Let (M_1, d_{M_1}) and (M_2, d_{M_2}) be two metric spaces. For any $\epsilon > 2d_{GH}(M_1, M_2)$ the persistence modules $H_*(\mathcal{R}(M_1, \rho))$ and $H_*(\mathcal{R}(M_2, \rho))$ are ϵ-interleaved.

Lemma 6.2 (Cěch Interleaving; Lemma 4.4 in [CSO12])**.** Let (M_1, d_{M_1}) and (M_2, d_{M_2}) be two metric spaces. For any $\epsilon > 2d_{GH}(M_1, M_2)$ the persistence modules $H_*(\mathcal{C}(M_1, \rho))$ and $H_*(\mathcal{C}(M_2, \rho))$ are ϵ-interleaved.

7 Autoencoder Interleaving Stability

7.1 Input and Latent Space Interleaving

The Gromov-Hausdorff distance between isometric spaces is zero. This case will only be attained for zero reconstruction error and Lipschitz-1 Encoder/Decoder maps. In the following, we use the Encoder bi-Lipschitz constant $L_{\mathcal{E}}$ on X and Y to deduce an upper bound for the interleaving between the persistence homology modules constructed as in Eq. 8 on the input space (X, d_X) and the latent space (Y, d_Y).

Theorem 7.1 (Gromov-Hausdorff Interleaving Stability)**.** Let $L_{\mathcal{E}}$ be the Encoder bi-Lipschitz constant. We consider filtered Rips- or Cěch complexes on X and Y respectively. Then, for the interleaving distance between the persistence modules $H_*(K_X)$ and $H_*(K_Y)$ we obtain the following upper bound:

$$d_I(H_*(K_X), H_*(K_Y)) \leq ((L_{\mathcal{E}} - 1) \cdot \mathrm{diam}(X)) \,. \tag{18}$$

Proof. The Encoder constitutes a surjective map. This defines a correspondence given by

$$R = \{(x_i, \mathcal{E}(x_i)) : x_i \in X\}. \tag{19}$$

This is called the correspondence associated with $\mathcal{E}$. By ([BBI01], Exercise 7.3.22), the distortion of a correspondence R associated with a map f is given by $\mathrm{dis}(R) = \mathrm{dis}(f)$. Thus, to obtain an upper bound for the Gromov-Hausdorff distance between the input space (X, d_X) and the latent space (Y, d_Y), we need to evaluate the distortion of the Encoder map. From the bi-Lipschitz continuity assumption follows that

$$\frac{1}{L_{\mathcal{E}}} d_X(x_i, x_j) \leq d_Y(\mathcal{E}(x_i), \mathcal{E}(x_j)) \leq L_{\mathcal{E}} d_X(x_i, x_j), \tag{20}$$

for all $x_i, x_j \in X$ and for a bi-Lipschitz constant $L_{\mathcal{E}} > 1$. Thus, using definition 6.2, we obtain:

$$\begin{aligned} \mathrm{dis}(R) = \mathrm{dis}(\mathcal{E}) &= \sup\{|d_X(x_i, x_j) - d_Y(\mathcal{E}(x_i), \mathcal{E}(x_j)|\} \\ &\leq \max\{\frac{(L_{\mathcal{E}} - 1) \cdot \mathrm{diam}(X)}{L_{\mathcal{E}}}; (L_{\mathcal{E}} - 1) \cdot \mathrm{diam}(X)\} \\ &= (L_{\mathcal{E}} - 1) \cdot \mathrm{diam}(X). \end{aligned} \tag{21}$$

By Definition 6.4, we get that

$$d_{GH}(X,Y) \leq \frac{(L_{\mathcal{E}}-1)\cdot \mathrm{diam}(X)}{2}. \tag{22}$$

And thus, using Lemma 6.1 and 6.2, for the interleaving distance d_I between $H_*(K_X)$ and $H_*(K_Y)$ we get

$$d_I(H_*(K_X), H_*(K_Y)) \leq (L_{\mathcal{E}}-1)\mathrm{diam}(X). \tag{23}$$

. □

7.2 Input and Output Space Interleaving

Our result in Theorem 7.1 applied to the Autoencoder map $\mathcal{A}$ would indicate that the persistence homology modules on the input and output space are interleaved by

$$d_I(H_*(K_X), H_*(K_{X'})) \leq (L_{\mathcal{A}}-1)\mathrm{diam}(X), \tag{24}$$

with the Autoencoder bi-Lipschitz constant $L_{\mathcal{A}}$ being estimated as in Eq. (3). In the following, we will give a closer bound in terms of the maximum pairwise reconstruction error. This will be constructed using the bottleneck distance, Definition 5.5. The following two Theorems from [Bot22] will be needed to evaluate the bottleneck distance between the persistence diagrams calculated on X and Y respectively.

Theorem 7.2. Let P and Q be finite sets of points in $\mathbb{R}^d$, and $\beta : P \to Q$ a bijection such that $||p-\beta(p)|| \leq \epsilon$ for all $p \in P$. Then

$$d_B(\mathrm{dgm}(H_*(\mathcal{R}(P,\rho))), \mathrm{dgm}(H_*(\mathcal{R}(Q,\rho))) \leq \epsilon. \tag{25}$$

Theorem 7.3. Let P and Q be finite sets of points in $\mathbb{R}^d$. If there exists an $\epsilon \geq 0$ such that $P \subseteq Q_\epsilon$ and $Q \subseteq P_\epsilon$, then

$$d_B(\mathrm{dgm}(H_*(\mathcal{C}(P,\rho))), \mathrm{dgm}(H_*(\mathcal{C}(Q,\rho)))) \leq \epsilon. \tag{26}$$

The notion P_ϵ denotes the points in $\mathbb{R}^d$ that are closer than ϵ to the point cloud P and likewise for Q.

For a bijection $\beta : P \to Q$ between finite sets of points P and Q the Euclidian distance $||p-\beta(p)||$ induces an ϵ-matching $Q \subseteq P_\epsilon$ and $P \subseteq Q_\epsilon$ with $\epsilon = \max||p-\beta(p)||$, $\forall p \in P$ and thus Theorem 7.2 and Theorem 7.3 are equivalent.

Theorem 7.4 (Bottleneck Distance Stability)**.** We consider filtered Rips- or Cěch complexes on X and X' respectively. For the interleaving distance between persistence homology modules on the input space X and output space X' we get:

$$d_H(X,X') \leq d_I(H_*(K_X), H_*(K_{X'})) \leq \delta_M. \tag{27}$$

Proof. The Hausdorff distance between X and X' is $d_H(X, X') \leq \delta_M$. As the bottleneck distance satisfies one more constraint than the Hausdorff distance, namely a bijection between points, for two metric spaces (M_1, d_{M_1}) and (M_2, d_{M_2}) we have that $d_H(M_1, M_2) \leq d_B(M_1, M_2)$, [CEH05]. Furthermore, from Theorem 7.2 and 7.3 follows that

$$d_B(\mathrm{dgm}(H_*(K_X), \mathrm{dgm}(H_* K_{X'}))) \leq \delta_M. \tag{28}$$

Hence, for the bottleneck distance between the barcode diagrams on the input space X and the output space X', we have that:

$$\delta_M \geq d_B(\mathrm{dgm}(H_*(K_X), \mathrm{dgm}(H_*(K_{X'}))) \geq d_H(X, X'), \tag{29}$$

with $d_H(X, X') \leq \delta_M$. The result for the interleaving distance follows from the Isometry Theorem 5.1. □

8 Topological Autoencoder

From Theorem 7.1 and Definition 6.4 follows that after training, a low distortion between the input space (X, d_X) and (Y, d_Y) is desirable in order to have a small interleaving distance between the respective persistence homology modules. In [Moo+19] the authors design a *topological* Autoencoder via addition of a specific loss term $\mathcal{L}_t$ (Eq.(2) and Eq.(3) in [Moo+19]) to the standard loss $\mathcal{L}_r(X, \mathcal{A}(X))$. This loss term $\mathcal{L}_t$ is designed to preserve *destroyer* [BDW10] edges in homology dimension 0. The authors propose a generalisation to higher dimensions via selection of relevant (i.e. *creator* [BDW10] or destroyer) edges from weighted simplices. Thus by design, the topological Autoencoder is trained to keep parts of the distortion low.

Theorem 8.1. If $\mathcal{L}_t = 0$, then the n-dimensional input and latent persistence homology modules have interleaving distance zero.

Proof. The n-dimensional persistence diagrams of X and Y are fully determined by the n-dimensional creator and destroyer simplices, which are assumed to be unique. Thus the two parts of the loss term $\mathcal{L}_t$, namely (Equation (2) and (3) in [Moo+19]):

$$\mathcal{L}_{X \to Y} := \frac{1}{2} ||\mathbf{A}^X[\pi^X] - \mathbf{A}^Y[\pi^X]||^2 \tag{30}$$

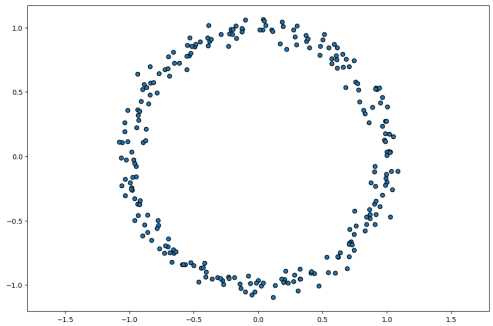
$$\mathcal{L}_{Y \to X} := \frac{1}{2} ||\mathbf{A}^Y[\pi^Y] - \mathbf{A}^Y[\pi^X]||^2 \tag{31}$$

represent bijections of the input and latent creator and destroyer simplices with bi-Lipschitz constant 1. □

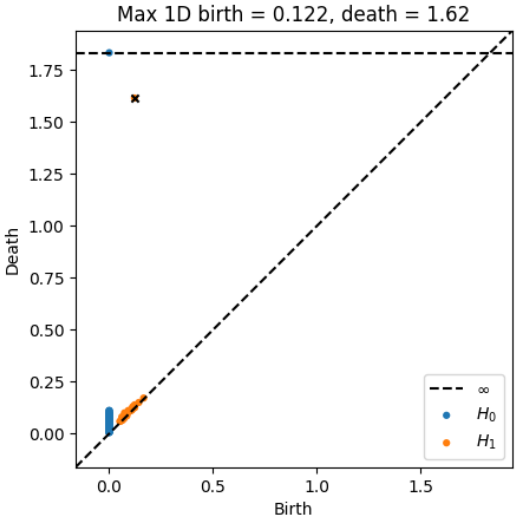
The authors of [Moo+19] perform experiments using the preservation of destroyer edges only for homology dimension 0 as they observed, that "using 1-dimensional topological features merely increases runtime" [Moo+19]. In the following, we construct a 2 dimensional counter example to demonstrate that this

is not correct in general and thus the authors observed a data-, hyperparameter- or optimisation specific effect. Furthermore, we will give an estimate for the interleaving distance between input and latent persistent homology modules for homology dimension $n+1$ if only creator or destroyer simplices for homology dimension n are preserved.

8.1 $2D$ Noisy Circle



(a) Noisy Circle

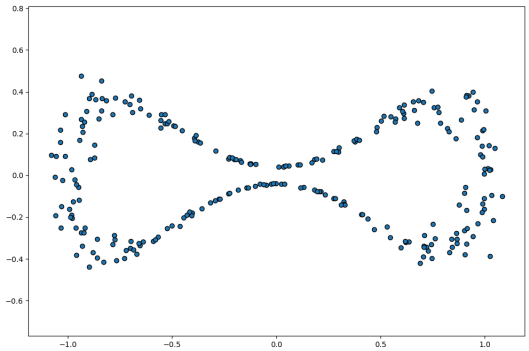


(b) Persistence Diagram

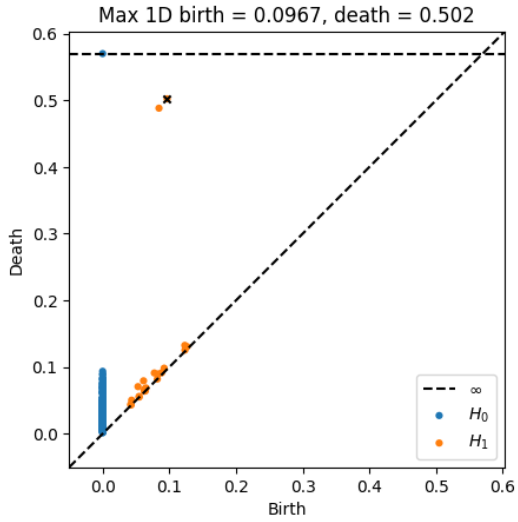
Figure 1: $2D$ Persistence of a Noisy Circle.

We consider a noisy circle in $2D$ Euclidian space, Figure 1 (a), and its corresponding persistence diagram in Figure 1 (b). The persistence diagram was calculated using ripser.py [Bau21]. Ripser is cohomology based which means that cohomological birth is actually homological death and vice versa. Yet the

homology birth/death convention is still used for plotting. In the persistence diagrams (Figure 1 (b) and Figure 2 (b)), the maximum birth/death interval is marked with a cross and its coordinates are denoted in the caption.



(a) Noisy 8-Shape



(b) Persistence Diagram

Figure 2: $2D$ Persistence of a Noisy 8-Shape.

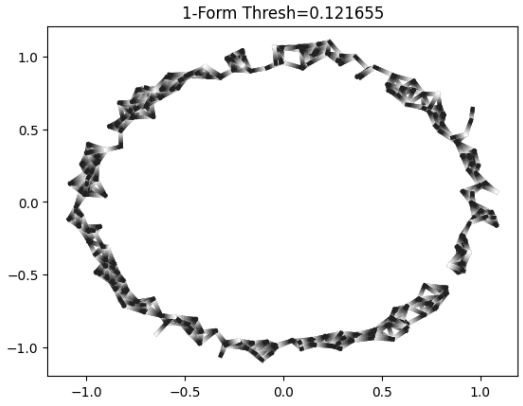
The $2D$ noisy circle is transformed to the noisy eight-shaped structure shown in Figure 2 (a) via the following diffeomorphism:

$$f : x \to \tanh x^2 + 0.04. \tag{32}$$

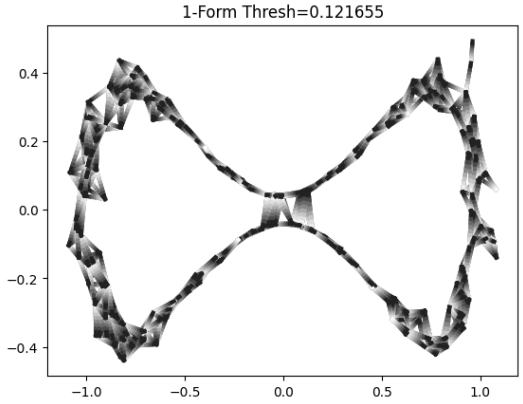
Apart from the noise, which clusters along the diagonal, the persistence diagram in Figure 2 (b) exhibits two significant features for homology dimension 1 in contrast to the persistence diagram in Figure 1 (b), which exhibits only one.

In Figure 3, the representative cocycles for the maximum birth/death intervals marked in the persistence diagrams in Figure 1 (b) and Figure 2 (b) are

shown for coefficients in $\mathbb{Z}/51\mathbb{Z}$. In [BM19] persistence diagrams with various prime field $\mathbb{Z}/p\mathbb{Z}$ coefficients were investigated and the authors experimentally observed that the differences are very small and appear for small prime numbers. As the cocycle class only exists in the range $[b_i, d_i)$, a small threshold denoted as "tresh" in the caption is subtracted from the death time of the feature marked in Figure 1 (b) and Figure 2 (b) respectively. The grey scale pattern in Figure 3 is used to highlight edge orientations (which are needed when working not with coefficients in $\mathbb{Z}/2\mathbb{Z}$). Edges are oriented from black to white.



(a) Noisy 8-Shape



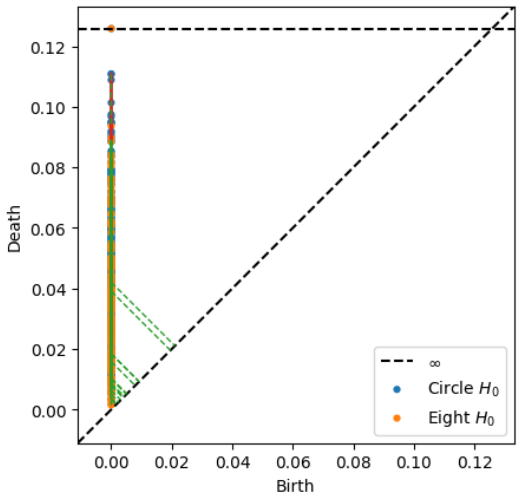
(b) Persistence Diagram

Figure 3: Representative cocycles of the noisy circle and the noisy 8-shape.

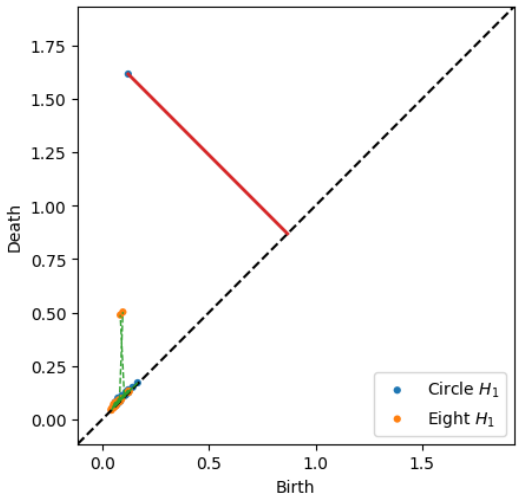
From Figure 3 becomes apparent that the 1 dimensional homology classes for high filtration parameters are quite distinct for the noisy circle and the noisy eight-shape, although they are connected via a diffeomomrphism. This is also clearly shown by the bottleneck distances between the respective persistence diagrams in Figure 1 (b) and Figure 2 (b) for homology dimensions 0 and 1. The results are shown in Table 1 and illustrated in Figure 4.

Homology Dimension	Bottleneck Distance
0	0.0210
1	0.7470

Table 1: Bottleneck Distances



(a) 0 Dimension Bottleneck Distance



(b) 1 Dimension Bottleneck Distance

Figure 4: Bottleneck Distances

In Figure 4, the red lines denote the bottleneck distances for dimension 0

and 1 respectively. The other pairs in the perfect matching which are less than the diagonal are shown as green lines. For homology dimension 0 we get a very small bottleneck and thus also interleaving distance of 0.0210 by Theorem 5.1.

This means, the diffeomorphism of Eq. (32) preserves 0 dimensional (co)homology classes. For homology dimension 1, this is not the case. Hence, preserving only the destroyer edges for 0 dimensional homology and thus 0 dimensional homology classes is not enough to guarantee a small interleaving distance for higher dimensional homology. This two-dimensional example could also be generalized to higher dimensions by using a noisy sphere instead of a noisy circle and deforming it likewise. One has to take deformations such as the one we constructed in Eq. (32) into account when training an Autoencoder, as the Encoder does not a priori give an isometry of homology groups.

Theorem 8.2. We assume an input data set with cardinality greater than 1. For metric simplicial complex constructions (such as in Eq. (4) and Eq. (5)), a one layer-operation $\mathbf{l}$ with weights initialized i.i.d. randomly ([He+15], [GB10]) in an interval $[-\mathfrak{w}, \mathfrak{w}]$, $\mathfrak{w} \in \mathbb{R}$, and application of a standard activation function **relu**, **tanh** or **sigmoid**, does not induce an isomorphism of the homology groups $H_*(K_i)$ and $H_*(K_o)$ almost surely. The simplicial complex K_i is constructed metrically on the input to the layer $\mathbf{l}$, and likewise a simplicial complex K_o is constructed on the output thereof.

Proof. Although in general there exist non-isometric maps that produce isomorphic complexes, for metric complex constructions we only need to show that a one layer operation (see Eq. (1)) does not constitute an isometry. As can be seen from Eq. (1), via the singular value decomposition of the weight matrix $\mathbf{W} = U\tilde{\Sigma}V^T$, the input values might be anisotropically scaled by $\tilde{\Sigma}$. As the entries of $\mathbf{W}$ are initialized randomly, $\mathbf{W}$ cannot be expected to be orthogonal and thus multiplication of the input vector with $\mathbf{W}$ does not constitute an isometry for filtration parameters $|\delta| > |\mathfrak{w}|$. To the thence weighted sum of input vector components, the nonlinear activation function gets applied as shown in Eq. 1. Among the most popular activation functions are **relu**, which is not even a bijection for domain values smaller than 0, **tanh** and **sigmoid**. Both tanh and sigmoid squeeze the input domain to a small codomain and hence will also not constitute an isometry. Thus, unless explicitly constructed otherwise, the one-layer operation of a neural network is not isometric and the filtration, Eq. 7, will not remain stable after a one-layer operation and even more for a composition thereof. □

8.2 Interleaving Distance in Dimension $n + 1$

In the following, we will give an upper bound for the interleaving distance between persistence input and latent homology modules in dimension $n+1$ when creator or destroyer simplices are only preserved up to dimension n. This bound depends on data inherent parameters. Following [NL22], topological errors when **selecting** (see [NL22], Section 3.1) corresponding topological features for the

comparison of the input space (X, d_X) and the latent space (Y, d_Y) can be divided into two classes defined as follows:

Definition 8.1 (Topological Type-I Error, Definition 3.1 in [NL22])**.** A topological type-I error in (Y, d_Y) is the selection of a feature in (Y, d_Y) which has no corresponding feature in (X, d_X).

Definition 8.2 (Topological Type-II Error, Definition 3.2 in [NL22])**.** A topological type-II error in (Y, d_Y) occurs when a structure which corresponds to a structure in (X, d_X) is not selected as significant.

Furthermore, for $\epsilon = d_I(H_n(\mathcal{R}(Y, \bar{\rho}'), H_n(\mathcal{R}(X, \bar{\rho}))$ [NL22] proved that selecting homology classes $(b, d) \in H_n(\mathcal{R}(Y, \bar{\rho}'))$ with $|d - b| > 2\epsilon$ causes no type-I errors ([NL22], Proposition 3.3). Herein, the notation $H_n(\mathcal{R}(X, \bar{\rho}))$ and $H_n(\mathcal{R}(Y, \bar{\rho}'))$ signifies that the persistence homology modules are considered for all filtration parameters up to the maximum filtration parameters $\bar{\rho}$ and $\bar{\rho}'$ for K_X and K_Y respectively. If one wants to preserve all significant n-dimensional topological features from the input space into the latent space, the interleaving distance ϵ between their persistence homology modules has to be significantly smaller than half of the maximum birth death interval in the n-th persistence diagram of (X, d_X). In order to also prevent type-II topological errors, the interleaving distance between n-dimensional input and latent persistence homology modules has to be significantly smaller than one fourth of the maximum birth death interval in the n-th persistence diagram of (X, d_X), ([NL22], Proposition 3.4). We will prove that restricting the topological loss to the preservation of n-dimensional significant simplices might cause the interleaving distance for homology dimension $n + 1$ to be larger than both of these bounds.

Theorem 8.3. Let $\bar{\rho} := \max(\mathbf{A}^X)$, with $\mathbf{A}^X$ being the distance matrix of the input metric space (X, d_X). Let $\mathcal{R}(X, \bar{\rho})$ be the Vietoris-Rips complex defined on X up to the maximum filtration parameter $\bar{\rho}$ such as described in [Moo+19]. Likewise, let $\mathbf{A}^Y$ be the distance matrix of the latent space with $\mathcal{R}(Y, \bar{\rho}')$ being the Vietoris-Rips complex defined up to the latent maximum filtration parameter $\bar{\rho}'$. Furthermore, let η_m denote the positive definite minimum distance in $\mathbf{A}^X$. Training an Autoencoder in a way such that creator and destroyer simplices are preserved up to homology dimension n will result in the following upper bound for the interleaving distance between input and latent persistence homology modules of dimension $(n + 1)$ or larger :

$$d_I(H_{d\geq n+1}(\mathcal{R}(X, \bar{\rho})), H_{d\geq n+1}(\mathcal{R}(Y, \bar{\rho}'))) < \frac{\bar{\rho} - \eta_m}{2}. \tag{33}$$

Proof. In [Moo+19] the authors assume, according to common persistent homology literature ([Hof+19], [PSO18]), that each persistence diagram entry corresponds to a unique distance between two data points. Thus there exists a unique positive definite minimum η_m and a unique pairwise maximum distance $\bar{\rho}$ for X. The latter is equal to the maximum filtration parameter of the Vietoris-Rips complex. Let the entries in $\mathbf{A}^X$ be sorted in an ascending way from η_m to $\bar{\rho}$.

By definition, η_m and also the next ascending entry in $\mathbf{A}^X$ have to be destroyer edges in homology dimension 0 and hence are preserved for all homology dimensions. Let κ' be the smallest entry in $\mathbf{A}^x$ which corresponds to neither a creator nor a destroyer simplex in homology dimension n but a creator simplex in homology dimension $n+1$. Let this simplex appear at some filtration parameter $(\eta_m+\kappa')$. In the worst case, κ' is small and the thence created $n+1$ dimensional topological feature persists in the input data until the end of the filtration interval $\bar{\rho}$ for homology dimension $n+1$. In this case, this feature, let it be ξ, in the persistence diagram of X has no correspondence in the $n+1$ dimensional persistence diagram of Y. Let ξ have birth-death coordinates $(\eta_m+\kappa', \bar{\rho})$. Due to the absence of a matching in the latent persistence diagram, ξ is matched to its diagonal projection $(\frac{\bar{\rho}+\eta_m+\kappa'}{2}, \frac{\bar{\rho}+\eta_m+\kappa'}{2})$.
The bottleneck distance as defined in Definition 5.5 measures the similarity between two persistence diagrams. It is the shortest distance b for which there exists a perfect matching between the points of two persistence diagrams, with unmatched points being completed with their diagonal projections, such that any couple of matched points are at distance at most b. The distance for a matched couple is measured via the supremum norm.
By definition of the topological loss, the training aims to preserve n-dimensional features. The training would fail if the n-dimensional features in the input and latent persistence diagram would be matched with a matching distance larger than the one of ξ, an unmatched point with maximum persistence in homology dimension $n+1$. This would indicate that training is not robust to the appearance of random artefacts, see also Theorem 8.1. Thus we can assume that the bottleneck distance for homology dimension n is smaller than the bottleneck distance for homology dimension $n+1$ whose persistence diagram exhibits the unmatched feature ξ. Furthermore, by assumption, ξ is the most persistent feature in our $n+1$ dimensional persistence diagram. Hence we can assume that the matching distance of the unmatched point ξ is the supremum of matching distances between the input and latent persistence diagrams for homology dimension $n+1$. The bottleneck distance between the input and latent space persistence diagram with ξ appearing in the input persistence diagram as above thus amounts to to the supremum distance of ξ with birth death coordinates $(\eta_m+\kappa', \bar{\rho})$ to its diagonal projection. The case $\kappa'=0$ cannot be attained, as the first two edges in $\mathbf{A}^X$ will always be preserved by definition. Thus, with ξ defined as above, we obtain for the bottleneck distance in homology dimension $n+1$:

$$d_B(\mathrm{dgm}H_*(\mathcal{R}(X,\bar{\rho})), \mathrm{dgm}H_*(\mathcal{R}(Y,\bar{\rho}'))) < \frac{\bar{\rho}-\eta_m}{2} \tag{34}$$

Our result follows by application of the Isometry Theorem 5.1. □

When implying the additional topological loss term developed by [Moo+19], the upper bound for the interleaving distance between $n+1$-dimensional persistence homology modules depends on the difference between the maximum filtration parameter $\bar{\rho}$ and the minimum positive definite pairwise distance η_m of the input data complex. Following the assumption dominant in the persis-

tent homology literature, that features at small filtration parameters are not relevant and can be considered as noise artefacts, the impact of an unmatched feature in a homology dimension larger than n is the more significant the larger the difference $|\bar{\rho} - \eta_m|$. Our noisy circle and noisy eight-shape in Figure 1 and Figure 2 illustrate such an example for $2D$. The radius of the circle in Figure 1 (a) is much larger than the inter-point distances. Due to the diffeomorphic transformation to Figure 2 the respective $1d$ homological feature in Figure 1 (b) remains unmatched in Figure 2 (b).

On the other hand, if $|\bar{\rho} - \eta_m|$ is small, the data points are rather homogeneously distributed and hence training the Autoencoder to preserve 0-dimensional homology will also preserve most of its higher dimensional homology, as features such as ξ in the proof of Theorem 8.3 will not appear. However, with regard to the above mentioned assumption about small scale topological features, this case does not exhibit topologically interesting features. Thus the topological loss in [Moo+19] might rather be used for dimension reductive clustering of homogeneously distributed data.

9 Discussion

In this paper, we prove an upper bound for the interleaving distance between the input and the latent persistence homology modules using the bi-Lipschitz constant of the map associated to a given Autoencoder architecture and the diameter of the input space X. We use the results of [CSO12] and elucidate that the distortion between the input and latent space has to be controlled in order to keep the interleaving distance low. This is partly realized in [Moo+19], however, we prove that one indeed has to control *all* pairwise distances in order to guarantee upper bounds on the interleaving distance for *all* homology dimensions. This is computationally not realizable. For further work, we thus pursue a local approach to persistence homology as proposed in [YG20b] and also [Cha+09]. One possible experimental streamline could be to investigate how well an Autoencoder can be trained to low distorsion. In addition, it would be interesting to see for which datasets it is sufficient to train the Encoder part just with the distorsion as loss. To refine our bound in Theorem 7.1 one could investigate further correspondences compatible with given Autoencoder architectures. One interesting question connected to this is also whether it is possible to refine this bound only partially, for low dimensional homology (0-, 1- up to x-dimensional homology).

References

[Bau21] Ulrich Bauer. "Ripser: efficient computation of Vietoris-Rips persistence barcodes". In: *J. Appl. Comput. Topol.* 5.3 (2021), pp. 391–

423. ISSN: 2367-1726. DOI: 10.1007/s41468-021-00071-5. URL: https://doi.org/10.1007/s41468-021-00071-5.

[BBI01] Dimitri Burago, Yuri Burago, and Sergei Ivanov. *A Course in Metric Geometry*. The American Mathematical Society, 2001.

[BDW10] Oleksiy Busaryev, Tamal K. Dey, and Yusu Wang. "Tracking a Generator by Persistence". In: *Discret. Math. Algorithms Appl.* 2010.

[BL15] Ulrich Bauer and Michael Lesnick. "Induced matchings and the algebraic stability of persistence barcodes". en. In: *Journal of Computational Geometry* (2015), Vol. 6 No. 2 (2015): Special issue of Selected Papers from SoCG 2014. DOI: 10.20382/JOCG.V6I2A9. URL: https://jocg.org/index.php/jocg/article/view/2983.

[BM19] Jean-Daniel Boissonnat and Clément Maria. "Computing persistent homology with various coefficient fields in a single pass". In: *Journal of Applied and Computational Topology* 3.1-2 (Apr. 2019), pp. 59–84. DOI: 10.1007/s41468-019-00025-y. URL: https://doi.org/10.1007%2Fs41468-019-00025-y.

[Bot22] Magnus Bakke Botnan. *Topological Data Analysis*. 2022. URL: https://www.few.vu.nl/~botnan/lecture_notes.pdf.

[BSS14] Peter Bubenik, Vin de Silva, and Jonathan Scott. "Metrics for Generalized Persistence Modules". In: *Foundations of Computational Mathematics* 15.6 (Oct. 2014), pp. 1501–1531. DOI: 10.1007/s10208-014-9229-5. URL: https://doi.org/10.1007%2Fs10208-014-9229-5.

[Cas19] Álvaro Torras Casas. *Distributing Persistent Homology via Spectral Sequences*. 2019. DOI: 10.48550/ARXIV.1907.05228. URL: https://arxiv.org/abs/1907.05228.

[CC07] Anthony L. Caterini and Dong Eui Chang. *Deep Neural Networks in a Mathematical Framework*. Springer, 2007.

[CCS14] Frederic Chazal, William Crawley-Boevey, and Vin de Silva. *The observable structure of persistence modules*. 2014. DOI: 10.48550/ARXIV.1405.5644. URL: https://arxiv.org/abs/1405.5644.

[CEH05] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. "Stability of Persistence Diagrams". In: vol. 37. Jan. 2005, pp. 263–271. DOI: 10.1007/s00454-006-1276-5.

[Cha+09] Frédéric Chazal et al. "Gromov-Hausdorff Stable Signatures for Shapes using Persistence". In: *Comput. Graph. Forum* 28 (July 2009), pp. 1393–1403. DOI: 10.1111/j.1467-8659.2009.01516.x.

[Cha+12] Frederic Chazal et al. *The structure and stability of persistence modules*. 2012. DOI: 10.48550/ARXIV.1207.3674. URL: https://arxiv.org/abs/1207.3674.

[Cra12] William Crawley-Boevey. *Decomposition of pointwise finite-dimensional persistence modules*. 2012. DOI: 10.48550/ARXIV.1210.0819. URL: https://arxiv.org/abs/1210.0819.

[CSO12] Frederic Chazal, Vin de Silva, and Steve Oudot. *Persistence stability for geometric complexes*. 2012. DOI: 10.48550/ARXIV.1207.3885. URL: https://arxiv.org/abs/1207.3885.

[FMN13] Charles Fefferman, Sanjoy Mitter, and Hariharan Narayanan. "Testing the Manifold Hypothesis". In: *Journal of the American Mathematical Society* 29 (Oct. 2013). DOI: 10.1090/jams/852.

[GB10] Xavier Glorot and Yoshua Bengio. "Understanding the difficulty of training deep feedforward neural networks". In: *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*. Ed. by Yee Whye Teh and Mike Titterington. Vol. 9. Proceedings of Machine Learning Research. Chia Laguna Resort, Sardinia, Italy: PMLR, 13–15 May 2010, pp. 249–256. URL: https://proceedings.mlr.press/v9/glorot10a.html.

[He+15] Kaiming He et al. "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification". In: *CoRR* abs/1502.01852 (2015). arXiv: 1502.01852. URL: http://arxiv.org/abs/1502.01852.

[Hof+19] Christoph Hofer et al. "Connectivity-Optimized Representation Learning via Persistent Homology". In: *Proceedings of the 36th International Conference on Machine Learning*. Ed. by Kamalika Chaudhuri and Ruslan Salakhutdinov. Vol. 97. Proceedings of Machine Learning Research. PMLR, Sept. 2019, pp. 2751–2760. URL: https://proceedings.mlr.press/v97/hofer19a.html.

[Lar+09] Hugo Larochelle et al. "Exploring Strategies for Training Deep Neural Networks". In: *Journal of Machine Learning Research* 10.1 (2009), pp. 1–40. URL: http://jmlr.org/papers/v10/larochelle09a.html.

[LM15] Ryan Lewis and Dmitriy Morozov. "Parallel Computation of Persistent Homology using the Blowup Complex". In: June 2015, pp. 323–331. DOI: 10.1145/2755573.2755587.

[MM20] Ishan Misra and Laurens van der Maaten. "Self-Supervised Learning of Pretext-Invariant Representations". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2020.

[Moo+19] Michael Moor et al. *Topological Autoencoders*. 2019. DOI: 10.48550/ARXIV.1906.00722. URL: https://arxiv.org/abs/1906.00722.

[NL22] Bradley J. Nelson and Yuan Luo. *Topology-Preserving Dimensionality Reduction via Interleaving Optimization*. 2022. DOI: 10.48550/ARXIV.2201.13012. URL: https://arxiv.org/abs/2201.13012.

[NSW08] Partha Niyogi, Stephen Smale, and Shmuel Weinberger. "Finding the Homology of Submanifolds with High Confidence from Random Samples". In: *Discrete & Computational Geometry* 39 (Mar. 2008), pp. 419–441. DOI: 10.1007/s00454-008-9053-2.

[Oud15] Steve Oudot. *Persistence Theory: From Quiver Representations to Data Analysis*. Jan. 2015.

[PSO18] Adrien Poulenard, Primoz Skraba, and Maks Ovsjanikov. "Topological Function Optimization for Continuous Shape Matching". In: *Computer Graphics Forum* 37 (2018).

[VP21] Jeya Maria Jose Valanarasu and Vishal M. Patel. "Overcomplete Deep Subspace Clustering Networks". In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. Jan. 2021, pp. 746–755.

[YG20a] Hee Rhang Yoon and Robert Ghrist. *Persistence by Parts: Multiscale Feature Detection via Distributed Persistent Homology*. 2020. DOI: 10.48550/ARXIV.2001.01623. URL: https://arxiv.org/abs/2001.01623.

[YG20b] Hee Rhang Yoon and Robert Ghrist. *Persistence by Parts: Multiscale Feature Detection via Distributed Persistent Homology*. 2020. DOI: 10.48550/ARXIV.2001.01623. URL: https://arxiv.org/abs/2001.01623.