Coordinatizing Data With Lens Spaces and Persistent Cohomology

Luis Polanco* Jose A. Perea †

Abstract

We introduce here a framework to construct coordinates in *finite* Lens spaces for data with nontrivial 1-dimensional $\mathbb{Z}_q$ persistent cohomology, for $q > 2$ prime. Said coordinates are defined on an open neighborhood of the data, yet constructed with only a small subset of landmarks. We also introduce a dimensionality reduction scheme in $S^{2n-1}/\mathbb{Z}_q$ (Lens-PCA: LPCA), and demonstrate the efficacy of the pipeline $\mathbb{Z}_q$-persistent cohomology $\Rightarrow S^{2n-1}/\mathbb{Z}_q$ coordinates $\Rightarrow$ LPCA, for nonlinear (topological) dimensionality reduction.

1 Introduction

One of the main questions in Topological Data Analysis (TDA) is how to use topological signatures like persistent (co)homology [11] to infer spaces parametrizing a given data set [3, 1, 4]. This is relevant in nonlinear dimensionality reduction since the presence of nontrivial topology—e.g., loops, voids, non-orientability, torsion, etc—can prevent accurate descriptions with low-dimensional Euclidean coordinates.

Here we seek to address this problem motivated by two facts. The first: If G is a topological abelian group, then one can associate to it a contractible space, EG, equipped with a free right G-action. For instance, if $G = \mathbb{Z}$, then $\mathbb{R}$ is a model for $E\mathbb{Z}$, with right $\mathbb{Z}$-action $\mathbb{R} \times \mathbb{Z} \ni (r, n) \mapsto r + n \in \mathbb{R}$. The quotient $BG := EG/G$ is called the classifying space of G [8]. In particular $B\mathbb{Z} \simeq S^1$, $B\mathbb{Z}_2 \simeq \mathbb{R}\mathbf{P}^\infty$, $BS^1 \simeq \mathbb{C}\mathbf{P}^\infty$ and $B\mathbb{Z}_q \simeq S^\infty/\mathbb{Z}_q$; here $\simeq$ denotes homotopy equivalence. The second fact: If B is a topological space and $\mathscr{C}_G$ is the sheaf over B (defined in [9]) sending each $U \subset B$ open to the abelian group of continuous maps from U to G, then $\check{H}^1(B; \mathscr{C}_G)$—the first Čech cohomology group of B with coefficients in $\mathscr{C}_G$—is in bijective correspondence with $[B\,, BG]$—the set of homotopy classes of continuous maps from B to the classifying space BG. This bijection is a manifestation of the Brown representability theorem [2], and implies, in so many words, that Čech cohomology classes can be represented as coordinates with values in a classifying space (like S^1 or $S^\infty/\mathbb{Z}_q$).

For point cloud data—i.e., for a finite subset X of an ambient metric space (M, d)—one does not compute Čech cohomology, but rather *persistent cohomology*. Specifically, the persistent cohomology of the Rips filtration on the data set X (or a subset of landmarks L). The first main result of this paper contends that steps one through three below mimic the bijection $\check{H}^1(B; \mathscr{C}_{\mathbb{Z}_q}) \cong [B, S^\infty/\mathbb{Z}_q]$ for $B \subset M$ an open neighborhood of X:

1. Let (M, d) be a metric space and let $L \subset X \subset M$ be finite. X is the data and L is a set of landmarks.
2. For a prime $q > 2$ compute $PH^1(\mathcal{R}(L); \mathbb{Z}_q)$; the 1-dim $\mathbb{Z}_q$-persistent cohomology of the Rips filtration on L. If the corresponding persistence diagram $\mathsf{dgm}(L)$ has an element (a, b) so that $2a < b$, then let $a \leq \epsilon < b/2$ and choose a representative cocycle $\eta \in Z^1(R_{2\epsilon}(L); \mathbb{Z}_q)$ whose cohomology class has (a, b) as birth-death pair.
3. Let $B_\epsilon(l)$ be the open ball in M of radius ϵ centered at $l \in L = \{l_1, \ldots, l_n\}$, and let $\varphi = \{\varphi_l\}_{l \in L}$ be a partition of unity subordinated to $\mathcal{B} = \{B_\epsilon(l)\}_{l \in L}$. If $\zeta_q \neq 1$ is a q-th root of unity, then the cocycle η yields a map $f : \bigcup \mathcal{B} \longrightarrow L^n_q$ to the Lens space $L^n_q = S^{2n-1}/\mathbb{Z}_q$, given in homogeneous coordinates by the formula
$$B_\epsilon(\ell_j) \ni b\,,\ f(b) = \left[\sqrt{\varphi_1(b)}\zeta_q^{\eta_{j1}} : \cdots : \sqrt{\varphi_n(b)}\zeta_q^{\eta_{jn}}\right]$$
where $\eta_{jk} \in \mathbb{Z}_q$ is the value of the cocycle η on the edge $\{l_j, l_k\} \in R_{2\epsilon}(L)$.

If $X \subset \bigcup \mathcal{B}$, then $f(X) = Y \subset L^n_q$ is the representation of the data—in a potentially high dimensional Lens space—corresponding to the cocycle η. The second contribution of this paper is a dimensionality reduction procedure in L^n_q akin to Principal Component Analysis, called LPCA. This allows us to produce from Y, a family of point clouds $P_k(Y) \subset L^k_q$, $1 \leq k \leq n$, $P_n(Y) = Y$, minimizing an appropriate notion of distortion. These are the Lens coordinates of X induced by the cocycle η.

This work, combined with [10, 12], should be seen as one of the final steps in completing the program of using the classifying space BG, for G abelian and finitely generated, to produce coordinates for data with nontrivial underlying 1^{st} cohomology. Indeed, this follows from the fact that $B(G \oplus G') \simeq BG \times BG'$, and that if G is finitely generated and abelian, then it is ismorphic to $\mathbb{Z}^n \oplus \mathbb{Z}_{n_1} \oplus \cdots \oplus \mathbb{Z}_{n_r}$ for unique integers $n, n_1, \ldots, n_r \geq 0$.

*Department of Computational Mathematics, Science and Engineering, Michigan State University, polanco2@msu.edu

†Department of Computational Mathematics, Science and Engineering, Department of Mathematics, Michigan State University, joperea@msu.edu

2 Preliminaries

2.1 Persistent Cohomology

A family $\mathcal{K} = \{K_\alpha\}_{\alpha\in\mathbb{R}}$ of simplicial complexes is called a filtration if $K_\alpha \subset K_{\alpha'}$ whenever $\alpha \leq \alpha'$. If $\mathbb{F}$ is a field and $i \geq 0$ is an integer, then the direct sum $PH^i(\mathcal{K};\mathbb{F}) := \bigoplus_\alpha H^i(K_\alpha;\mathbb{F})$ of cohomology groups is called the i-th dimensional **$\mathbb{F}$-persistent cohomology** of $\mathcal{K}$. A theorem of Crawley-Boevey [5] contends that if $H^i(K_\alpha;\mathbb{F})$ is finite dimensional for each α, then the isomorphism type of $PH^i(\mathcal{K};\mathbb{F})$—as a persistence module—is uniquely determined by a multiset (i.e., a set whose elements may appear with repetitions)

$$\mathsf{dgm} \subset \{(\alpha,\alpha') \in [-\infty,\infty]^2 : \alpha \leq \alpha'\}$$

called the **persistence diagram** of $PH^i(\mathcal{K};\mathbb{F})$. Pairs (α,α') with large persistence $\alpha'-\alpha$, are indicative of stable topological features throughout the filtration $\mathcal{K}$.

Persistent cohomology is used in TDA to quantify the topology underlying a data set. There are two widely used filtrations associated to a subset X of a metric space (M,d), the **Rips filtration** $\mathcal{R}(X) = \{R_\alpha(X)\}_\alpha$ and the **Čech filtration** $\check{\mathcal{C}}(X) = \{\check{C}_\alpha(X)\}_\alpha$. Specifically, $R_\alpha(X)$ is the set of nonempty finite subsets of X with diameter less than α, and $\check{C}_\alpha(X)$ is the nerve of the collection $\mathcal{B}_\alpha$ of open balls $B_\alpha(x) \subset M$ of radius α, centered at $x \in X$. In other words, $\check{C}_\alpha(X) = \mathcal{N}(\mathcal{B}_\alpha)$. Generally $\mathcal{R}(X)$ is more easily computable, but $\check{\mathcal{C}}(X)$ has better theoretical properties (e.g., the Nerve theorem [6, 4G.3]). Their relative weaknesses are ameliorated by noticing that

$$R_\alpha(X) \subset \mathcal{N}(\mathcal{B}_\alpha) \subset R_{2\alpha}(X)$$

for all α, and using both filtrations in analyses: Rips for computations, and Čech for theoretical inference.

2.2 Lens Spaces

Let $q \in \mathbb{N}$ and let $\zeta_q \in \mathbb{C}$ be a primary q-th root of unity. Fix $n \in \mathbb{N}$ and let $q_1,\ldots,q_n \in \mathbb{N}$ be relatively prime to q. We define the **Lens space** $L^n_q(q_1,\ldots,q_n)$ as the quotient of $S^{2n-1} \subset \mathbb{C}^n$ by the $\mathbb{Z}_q$ right action

$$[z_1,\ldots,z_n]\cdot g := \left[z_1\zeta_q^{q_1 g},\ldots,z_n\zeta_q^{q_n g}\right]$$

with simplified notation $L^n_q := L^n_q(1,\ldots,1)$. Notice that when $q=2$ and $q_1 = \cdots = q_n = 1$, then the right action described above is the antipodal map of S^{2n-1}, and therefore $L^n_2 = \mathbb{R}\mathbf{P}^{2n-1}$. Similarly, the infinite Lens space $L^\infty_q = L^\infty_q(1,1,\ldots)$ is defined as the quotient of the infinite unit sphere $S^\infty \subset \mathbb{C}^\infty$, by the action of $\mathbb{Z}_q$ induced by scalar-vector multiplication by powers of ζ_q.

2.2.1 A Fundamental domain for $L^2_q(1,p)$

In what follows we describe a convenient model for both $L^2_q(1,p)$ and a fundamental domain thereof. This model will allow us to provide visualizations in Lens spaces towards the end of the paper. Let D^3 be the set of points $\mathbf{x} \in \mathbb{R}^3$ with $\|\mathbf{x}\| \leq 1$, and let D_+ (D_-) be the upper (lower) hemisphere of ∂D^3, including the equator. Let $r_{p/q} : D_+ \longrightarrow D_+$ be counterclockwise rotation by $2\pi p/q$ radians around the z-axis, and let $\rho : D_+ \longrightarrow D_-$ be the reflection $\rho(x,y,z) = (x,y,-z)$. Then, $L^2_q(1,p)$ is homeomorphic to $D^3/\sim$, where $\mathbf{x} \sim \mathbf{y}$ if and only if $\mathbf{x} \in D_+$ and $\mathbf{y} = \rho \circ r_{p/q}(\mathbf{x})$.

2.3 Principal Bundles

Let B be a topological space with base point $b_0 \in B$. One of the most transparent methods for producing an explicit bijection between $\check{H}^1(B;\mathscr{C}_{\mathbb{Z}_q})$ and $[B, L^\infty_q]$ is via the theory of Principal bundles. We present a terse introduction here, but direct the interested reader to [7] for details. A continuous map $\pi : P \longrightarrow B$ is said to be a **fiber bundle** with fiber $F = \pi^{-1}(b_0)$ and total space P, if π is surjective, and every $b \in B$ has an open neighborhood $U \subset B$ as well as a homeomorphism $\rho_U : U \times F \longrightarrow \pi^{-1}(U)$, so that $\pi \circ \rho_U(x,e) = x$ for every $(x,e) \in U \times F$.

Let $(G,+)$ be an abelian topological group. A fiber bundle $\pi : P \longrightarrow B$ is said to be a **principal G-bundle** over B, if P comes equipped with a free right G-action $P \times G \ni (e,g) \mapsto e\cdot g \in P$ which is transitive in $\pi^{-1}(b)$ for every $b \in B$. Moreover, two principal G-bundles $\pi : P \longrightarrow B$ and $\pi' : P' \longrightarrow B$ are isomorphic, if there exits a homeomorphism $\Phi : P \longrightarrow P'$, with $\pi' \circ \Phi = \pi$ and so that $\Phi(e\cdot g) = \Phi(e)\cdot g$ for all $(e,g) \in P \times G$. Given an open cover $\mathcal{U} = \{U_j\}_{j\in J}$ of B, a **Čech cocycle**

$$\eta = \{\eta_{jk}\} \in \check{Z}^1(\mathcal{U};\mathscr{C}_G)$$

is a collection of continuous maps $\eta_{jk} : U_j \cap U_k \longrightarrow G$ so that $\eta_{jk}(b) + \eta_{kl}(b) = \eta_{jl}(b)$ for every $b \in U_j \cap U_k \cap U_l$. Given such a cocycle, one can construct a principal G-bundle over B with total space

$$P_\eta = \left(\bigcup_{j\in J} U_j \times \{j\} \times G\right)/\sim$$

where $(b,j,g) \sim (b,k,g+\eta_{jk}(b))$ for every $b \in U_j \cap U_k$, and $\pi : P_\eta \longrightarrow B$ sends the class of (b,j,g) to $b \in B$.

Theorem 1 *If $\mathsf{Prin}_G(B)$ denotes the set of isomorphism classes of principal G-bundles over B, then*

$$\begin{array}{ccc} \check{H}^1(B;\mathscr{C}_G) & \longrightarrow & \mathsf{Prin}_G(B) \\ [\eta] & \mapsto & [P_\eta] \end{array}$$

is a bijection.

Proof. See 2.4 and 2.5 in [10] $\square$

Now, let us see describe the relation between principal G-bundles over B, and maps from B to the classifying space BG. Indeed, let $\jmath : EG \longrightarrow BG = EG/G$ be the quotient map. Given $h : B \longrightarrow BG$ continuous, the pullback h^*EG is the principal G-bundle over B with total space $\{(b,e) \in B \times EG : h(b) = \jmath(e)\}$, and projection map $(b,e) \mapsto b$. Moreover,

Theorem 2 *Let $[B, BG]$ denote the set of homotopy class of maps from B to the classifying space BG. Then, the function*

$$\begin{array}{ccc} [B, BG] & \longrightarrow & \mathsf{Prin}_G(B) \\ [h] & \mapsto & [h^*EG] \end{array}$$

is a bijection.

Proof. See [7], Chapter 4: Theorems 12.2 and 12.4. $\square$

In summary, given a principal G-bundle $\pi : P \longrightarrow B$, or its corresponding Čech cocycle η, there exists a continuous map $h : B \longrightarrow BG$ so that h^*EG is isomorphic to (π, P), and the choice of h is unique up to homotopy. Any such choice is called a classifying map for $\pi : P \longrightarrow B$.

3 Main Theorem: Explicit Classifying Maps for L_q^∞

The goal of this section is to show how one can go from a singular cocycle $\eta \in Z^1(\mathcal{N}(\mathcal{U}); \mathbb{Z}_q)$ to an explicit map $f : \bigcup \mathcal{U} \longrightarrow L_q^\infty$. All proofs are included in the Apendix. Let $J = \{1, \ldots, n\}$, let $\mathcal{U} = \{U_j\}_{j\in J}$ be an open cover for B, and let $\{\varphi_j\}_{j\in J}$ be a partition of unity dominated by $\mathcal{U}$. If $\eta = Z^1(\mathcal{N}(\mathcal{U}); \mathbb{Z}_q)$ and ζ_q is a primitive q-th root of unity, let $f_j : U_j \times \{j\} \times \mathbb{Z}_q \longrightarrow S^{2n-1} \subset \mathbb{C}^n$ be

$$f_j(b, j, g) = \left[\sqrt{\varphi_1(b)}\zeta_q^{(g+\eta_{j1})}, \ldots, \sqrt{\varphi_n(b)}\zeta_q^{(g+\eta_{jn})}\right]$$

If $b \in U_j \cap U_k$, then $f_j(b, j, g) = f_k(b, k, g + \eta_{jk})$ and we get an induced map $\Phi : P_\eta \longrightarrow S^{2n-1} \subset S^\infty$ taking the class of (b, j, g) in the quotient P_η to $f_j(b, j, g)$.

Proposition 3 *Φ is well defined and $\mathbb{Z}_q$-equivariant.*

Let $p : S^{2n-1} \longrightarrow L_q^n$ be the quiotient map. Since $\Phi : P_\eta \longrightarrow S^{2n-1} \subset S^\infty$ is $\mathbb{Z}_q$-equivariant, it induces a map $f : B \longrightarrow L_q^n \subset L_q^\infty$ such that $p \circ \Phi = f \circ \pi$. By construction of $\pi : P_\eta \longrightarrow B$, $f(\pi([b, j, g])) = f(b)$ for any $g \in \mathbb{Z}_q$. In particular for $0 \in \mathbb{Z}_q$

$$U_j \ni b \ \ , \ \ f(b) = \left[\sqrt{\varphi_1(b)}\zeta_q^{\eta_{j1}} : \cdots : \sqrt{\varphi_n(b)}\zeta_q^{\eta_{jn}}\right] \quad (1)$$

Remark 4 *The notation $[a_1 : \cdots : a_n]$ corresponds to homogeneous coordinates in $S^{2n-1}/\mathbb{Z}_q$. In other words, $[a_1 : \cdots : a_n] = \{[a_1 \cdot \alpha, \ldots, a_n \cdot \alpha] \in S^{2n-1} : \alpha \in \mathbb{Z}_q\}$.*

Theorem 5 *The map f classifies the $\mathbb{Z}_q$-principal bundle P_η associated to the cocycle $\eta \in Z^1(\mathcal{N}(\mathcal{U}); \mathbb{Z}_q)$.*

4 Lens coordinates for data

Let (M, d) be a metric space and let $L \subset M$ be a finite subset. We will use the following notation from now on: $B_\epsilon(l) = \{y \in M : d(y, l) < \epsilon\}$, $\mathcal{B}_\epsilon = \{B_\epsilon(l)\}_{l\in L}$, and $L^\epsilon = \bigcup \mathcal{B}_\epsilon$. Given a data set $X \subset M$, our goal will be to choose $L \subset X$, a suitable ϵ such that $X \subset L^\epsilon$, and a cocycle $\eta \in Z^1(\mathcal{N}(\mathcal{B}_\epsilon); \mathbb{Z}_q)$. Equation (1) yields a map $f : L^\epsilon \to L_q^\infty$ defined for every point in X, but constructed from a much smaller subset of landmarks. Next we describe the details of this construction.

4.1 Landmark selection

We select the landmark set $L \subset X$ either at random or through `maxmin` sampling. The latter proceeds inductively as follows: Fix $n \leq |X|$, and let $l_1 \in X$ be chosen at random. Given $l_1, \ldots, l_j \in X$ for $j < n$, we let $l_{j+1} = \underset{x\in X}{\operatorname{argmax}} \min\{d(x, l_1), \ldots, d(x, l_j)\}$.

4.2 A Partition of Unity subordinated to $\mathcal{B}_\epsilon$

Defining f requires a partition of unity subordinated to $\mathcal{B}_\epsilon$. Since $\mathcal{B}_\epsilon$ is an open cover composed of metric balls, then we can provide an explicit construction. Indeed, for $r \in \mathbb{R}$ let $|r|_+ := \max\{r, 0\}$, then

$$\varphi_l(x) := |\epsilon - d(x, l)|_+ \Big/ \sum_{l' \in L} |\epsilon - d(x, l')|_+ \quad (2)$$

is a partition of unity subordinated to $\mathcal{B}_\epsilon$.

4.3 From Rips to Čech to Rips

As we alluded to in the introduction, a persistent cohomology calculation is an appropriate vehicle to select a scale ϵ and a candidate cocycle η. That said, determining $\eta \in Z^1(\mathcal{N}(\mathcal{B}_\epsilon), \mathbb{Z}_q)$ would require computing $\mathcal{N}(\mathcal{B}_\epsilon)$ for all ϵ, which in general is an expensive procedure. Instead we will use the homomorphisms

$$\begin{tikzcd} H^1(\mathcal{R}_{2\epsilon}(L)) \arrow[r, "i^*"] \arrow[rr, bend right, "\iota"'] & H^1(\mathcal{N}(\mathcal{B}_\epsilon)) \arrow[r] & H^1(\mathcal{R}_\epsilon(L)) \end{tikzcd}$$

induced by the appropriate inclusions. Indeed, let $\tilde{\eta} \in Z^1(\mathcal{R}_{2\epsilon}(L); \mathbb{Z}_q)$ be such that $[\tilde{\eta}] \notin \ker(\iota)$. This is where we use the persistent cohomology of $\mathcal{R}(L)$. Since the previous diagram commutes, then $[\tilde{\eta}] \notin \ker(i^*)$, so $i^*([\tilde{\eta}]) \neq 0$ in $H^1(\mathcal{N}(\mathcal{B}_\epsilon); \mathbb{Z}_q)$. We will let $[\eta] = i^*([\tilde{\eta}])$ be the class that we use in Theorem 5. However,

Proposition 6 *If $b \in \mathcal{B}_\epsilon(l_j)$ and $1 \leq k \leq n$, then*

$$\sqrt{\varphi_k(b)}\zeta_q^{\eta_{jk}} = \sqrt{\varphi_k(b)}\zeta_q^{\tilde{\eta}_{jk}}.$$

That is, we can compute Lens coordinates using only the Rips filtration on the landmark set.

5 Dimensionality Reduction in L_q^n via Principal Lens Components

Equation (1) gives an explicit formula for the classifying map $f : B \longrightarrow L_q^n$. By construction, the dimension of L_q^n depends on the number n of landmarks selected, which in general can be large. The main goal of this section is to construct a dimensionality reduction procedure in L_q^n to address this shortcoming. To this end, we define the distance $d_L : L_q^n \times L_q^n \longrightarrow [0,\infty)$ as

$$d_L([x],[y]) := d_H(x\cdot\mathbb{Z}_q\,,\, y\cdot\mathbb{Z}_q) \tag{3}$$

where d_H id the Hausdorff distance for subsets of S^{2n-1}.

We will now describe a notion of **projection in L_q^n** onto lower-dimensional Lens spaces. Indeed, let $u \in S^{2n-1}$. Since $\zeta_q^k w \in \mathsf{span}_{\mathbb{C}}(u)^\perp$ for any $k \in \mathbb{Z}_q$ and $w \in \mathsf{span}_{\mathbb{C}}(u)^\perp$, then

$$L_q^{n-1}(u) := (\mathsf{span}_{\mathbb{C}}(u)^\perp \cap S^{2n-1})/\mathbb{Z}_q$$

is isometric to L_q^{n-1}. Let $P_u^\perp(v) = v - \langle v,u\rangle_{\mathbb{C}} u$ for $v \in \mathbb{C}^n$, and if $v \notin \mathsf{span}_{\mathbb{C}}(u)$, then we let

$$\mathcal{P}_u([v]) := \left[P_u^\perp(v)/\|P_u^\perp(v)\|\right] \in L_q^{n-1}(u)$$

It readily follows that $\mathcal{P}_u$ is well defined, and that

Lemma 7 *For $u \in S^{2n-1}$ and $v \notin \mathsf{span}_{\mathbb{C}}(u)$, we have*

$$d_L([v],\mathcal{P}_u([v])) = d\left(v\ ,\ P_u^\perp(v)/\|P_u^\perp(v)\|\right)$$

where d is the distance on S^{2n-1}. Furthermore, $\mathcal{P}_u([v])$ is the point in $L_q^{n-1}(u)$ closest to $[v]$ with respect to d_L.

This last result suggests that a PCA-like approach is possible for dimensionality reduction in Lens spaces. Specifically, for $Y = \{[y_1],\dots,[y_N]\} \subset L_q^n$, the goal is to find $u \in S^{2n-1}$ such that $L_q^{n-1}(u)$ is the best $(n-1)$-Lens space approximation to Y, then project Y onto $L_q^{n-1}(u)$ using $\mathcal{P}_u$, and repeat the process iteratively reducing the dimension by 1 each time. At each stage, the appropriate constrained optimization problem is

$$\begin{aligned} u^* &= \underset{u\in\mathbb{C}^n,\|u\|=1}{\operatorname{argmin}} \sum_{j=1}^N d_L([y_j],\mathcal{P}_u([y_i]))^2 \\ &= \underset{u\in\mathbb{C}^n,\|u\|=1}{\operatorname{argmin}} \sum_{j=1}^N \left(\frac{\pi}{2} - \arccos(|\langle y_i,u\rangle|)\right)^2 \end{aligned}$$

which can be linearized using the Taylor series expansion of $\arccos(\theta)$ around 0. Indeed, $|\frac{\pi}{2} - \arccos(\theta)| \approx |\theta|$ to third order, and thus

$$u^* \approx \underset{u\in\mathbb{C}^n,\|u\|=1}{\operatorname{argmin}} \sum_{j=1}^N |\langle y_i,u\rangle|^2.$$

This approximation is a linear least square problem whose solution is given by the eigenvector corresponding to the smallest eigenvalue of the covariance matrix

$$\mathsf{Cov}\,(y_1,\dots,y_N) = \begin{bmatrix} | & & | \\ y_1 & \cdots & y_N \\ | & & | \end{bmatrix}\begin{bmatrix} - & \overline{y}_1 & - \\ & \vdots & \\ - & \overline{y}_N & - \end{bmatrix}.$$

Moreover, for any $\alpha_1,\dots,\alpha_N \in S^1 \subset \mathbb{C}$ we have that $\mathsf{Cov}\,(\alpha_1 y_1,\dots,\alpha_N y_N) = \mathsf{Cov}\,(y_1,\dots,y_N)$, so $\mathsf{Cov}(Y)$ is well defined for $Y \subset L_q^n$.

5.1 Inductive construction of LPCA

Let $v_n = \mathsf{LastLensComp}(Y)$ be the eigenvector of $\mathsf{Cov}(Y)$ corresponding to the smallest eigenvalue. Assume that we have constructed $v_{k+1},\dots,v_n \in S^{2n-1}$ for $1 < k < n$, and let $\{u_1,\dots,u_k\}$ be an orthonormal basis for $\mathsf{span}_{\mathbb{C}}(v_{k+1},\dots,v_n)^\perp$. Let $U_k \in \mathbb{C}^{n\times k}$ be the matrix with columns $u_1,\dots,u_k$, and let $U_k^\dagger$ be its conjugate transpose. We define the **k-th Lens Principal component** of Y as the vector

$$v_k \ := \ U_k \cdot \mathsf{LastLensComp}\left(\frac{U_k^\dagger y_1}{\|U_k^\dagger y_1\|},\dots,\frac{U_k^\dagger y_N}{\|U_k^\dagger y_N\|}\right)$$

This inductive procedure yields a collection $[v_2],\dots,[v_n] \in L_q^n$, and we let $v_1 \in S^{2n-1}$ be such that $\mathsf{span}_{\mathbb{C}}\{v_1\} = \mathsf{span}_{\mathbb{C}}\{v_2,\dots,v_n\}^\perp$. Finally

$$\mathsf{LPCA}(Y) := \{[v_1],\dots,[v_n]\}$$

are the **Lens Principal Components** of Y. Let $V_k \in \mathbb{C}^{n\times k}$ be the n-by-k matrix with columns $v_1,\dots,v_k$, and let $P_k(Y) \subset L_q^k$ be the set of classes $\left[\frac{V_k^\dagger y_j}{\|V_k^\dagger y_j\|}\right]$, $1 \le j \le N$. The point clouds $P_k(Y)$, $k = 1,\dots,n$, are the **Lens Principal Coordinates** of Y.

5.2 Choosing a target dimension.

The **variance recovered** by the first k Lens Principal Components $[v_1],\dots,[v_k] \in L_q^n$ is defined as

$$\mathsf{var}_k(Y) := \frac{1}{N}\sum_{l=2}^k\sum_{j=1}^N d_L\left(\left[\frac{V_l^\dagger y_j}{\|V_l^\dagger y_j\|}\right], L_q^{l-1}(e_{l-1})\right)^2$$

where V_l is the n-by-l matrix with columns $v_1,\dots,v_l$, $1 < l \le k$, and $e_{l-1} \in \mathbb{C}^l$ is the vector $[0,\dots,0,1,0]$.

Therefore, the **percentage of cumulative variance** $p.\mathsf{var}(k) := \mathsf{var}_k(Y)/\mathsf{var}_n(Y)$, can be interpreted as the portion of total variance of Y along $\mathsf{LPCA}(Y)$, explained by the first k components.

Thus we can select the target dimension as the smallest k for which $p.\mathsf{var}_k(Y)$ is greater than a predetermined value. In other words, we select the dimension that recovers a significant portion of the total variance. Another possible guideline to choose the target dimension is as the minimum value of k for which $p.\mathsf{var}(k) - p.\mathsf{var}(k+1) < \gamma$ for a small $\gamma > 0$.

5.3 Independence of the cocycle representative.

Let $\eta \in Z^1(\mathcal{N}(\mathcal{B}_\epsilon); \mathbb{Z}_q)$ be such that $[\eta] \neq 0$ in $H^1(\mathcal{N}(\mathcal{B}_\epsilon); \mathbb{Z}_q)$, and let $\eta' = \eta + \delta^0(\alpha)$ with $\alpha \in C^0(\mathcal{N}(\mathcal{B}_\epsilon); \mathbb{Z}_q)$. If $b \in U_j$, then

$$f_{\eta'}(b) = [\sqrt{\phi_1(b)}\zeta_q^{\eta_{j1}+\alpha_1} : \cdots : \sqrt{\phi_n(b)}\zeta_q^{\eta_{jn}+\alpha_n}]$$

If Z_α is the square diagonal matrix with entries $\zeta_q^{\alpha_1}, \zeta_q^{\alpha_2}, \ldots, \zeta_q^{\alpha_n}$, then $f_{\eta'}(b) = Z_\alpha \cdot f(b)$. Moreover, after taking classes in L_q^n, this implies that $f_{\eta'}(X) = Z_\alpha \cdot f(X)$. Since $\mathsf{Cov}(Z_\alpha \cdot f(X)) = Z_\alpha \mathsf{Cov}(f(X)) Z_\alpha^\dagger$ and Z_α is orthonormal, then if v is an eigenvector of $\mathsf{Cov}(f(X))$ with eigenvalue σ, we also have that $Z_\alpha v$ is an eigenvector of $\mathsf{Cov}(Z_\alpha \cdot f(X))$ with the same eigenvalue. Therefore

$$\mathsf{LastLensComp}(f_{\eta'}(X)) = Z_\alpha \mathsf{LastLensComp}(f(X)).$$

Since each component in LPCA is obtained in the same manner, we have that $\mathsf{LPCA}(f_{\eta'}(X)) = Z_\alpha \mathsf{LPCA}(f(X))$. Thus, the lens coordinates from two cohomologous cocycles η and $\eta + \delta^0(\alpha)$ (i.e., representing the same cohomology class) only differ by the isometry of L_q^n induced by the linear map Z_α.

6 Examples

6.1 The Circle S^1

Let $S^1 \subset \mathbb{C}$ be the unit circle, and let X a random sample around S^1, with $10,000$ points and Gaussian noise in the normal direction. $L \subset X$ is a landmark set with 10 points obtained as described in Section 4.1.

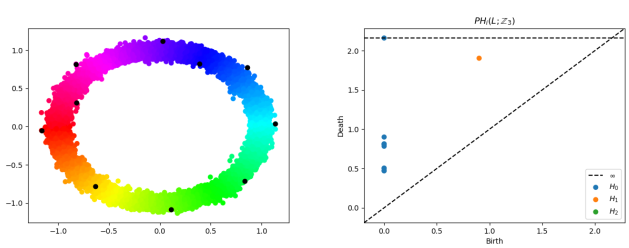


Figure 1: **Left:** Sample X, in black landmark set $L \subset X$. **Right:** $PH^i(\mathcal{R}(L); \mathbb{Z}_3)$ for $i = 0, 1, 2$.

Let a be the cohomological death of the most persistent class $PH^1(\mathcal{R}(L); \mathbb{Z}_q)$. For $\epsilon := a + 10^{-5}$ and $\eta = i^*(\eta') \in Z^1(\mathcal{N}(\mathcal{B}_\epsilon); \mathbb{Z}_q)$ we define the map $f : B_\epsilon \to L_3^{10}$ as in Equation (1).

After computing LPCA for $f(X) \subset L_3^{10}$ and the percentage of cumulative variance $p.\mathsf{var}_Y(k)$ we obtain the row in Table 1 with label S^1 (see Figure 7 for more details). We see that dimension 1 recovers $\sim 60\%$ of the variance. Moreover, Figure 2 shows $P_2(f(X)) \subset L_3^2$ in the fundamental domain described in Section 2.2.1 trough the map in Equation (4).

Dim. (n)	**1**	**2**	**3**	**4**	**5**
S^1	0.62	0.75	0.81	0.86	0.89
$M(\mathbb{Z}_3, 1)$	0.56	0.7	0.76	0.8	0.83
L_3^2	0.47	0.62	0.67	0.71	0.73

Table 1: Percentage of recovered variance in L_3^n.

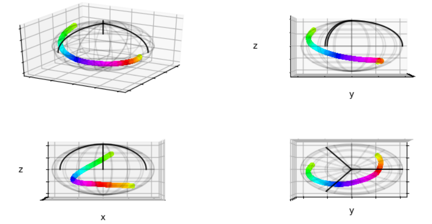


Figure 2: Visualization $P_2(f(X)) \subset L_3^2$.

One key aspect of LC (Lens coordinates) is that it is designed to highlight the cohomology class η used on Equation (1). This is easily observed in this example; we selected the most persistent class in $PH^1(\mathcal{R}(L); \mathbb{Z}_3)$ and as a consequence in Figure 2 we see how this class is preserved while all the information in the normal direction is lost in the process.

6.2 The Moore space $M(\mathbb{Z}_3, 1)$.

Let G be an abelian group and $n \in \mathbb{N}$. The Moore space $M(G, n)$ is a CW-complex such that $H_n(M(G, n), \mathbb{Z}) = G$ and $\tilde{H}_i(M(G, n), \mathbb{Z}) = 0$ for all $i \neq n$. A well known construction for $M(\mathbb{Z}_3, 1)$ can be found in [6]. Equation (5) defines a metric on $M(\mathbb{Z}_3, 1)$.

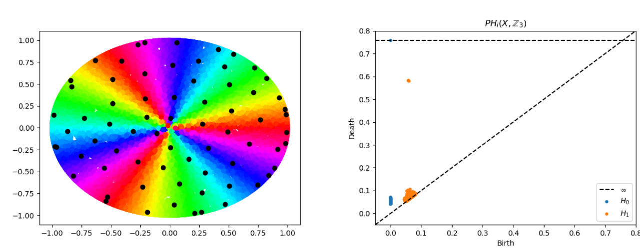


Figure 3: **Left:** $X \subset M(\mathbb{Z}_3, 1)$ with landmarks in black. **Right:** $PH^i(\mathcal{R}(L); \mathbb{Z}_3)$ for $i = 0, 1$.

Figure 3, on the left, shows a sample $X \subset M(\mathbb{Z}_3, 1)$ with $|X| = 15,000$ and 70 landmarks. The landmarks were obtained by minmax sampling after feeding the algorithm with an initial set of 10 point on the boundary on the disc. Figure 4 shows the persistent cohomology of $\mathcal{R}(L)$ with coefficients in $\mathbb{Z}_2$ and $\mathbb{Z}_3$ side-by-side.

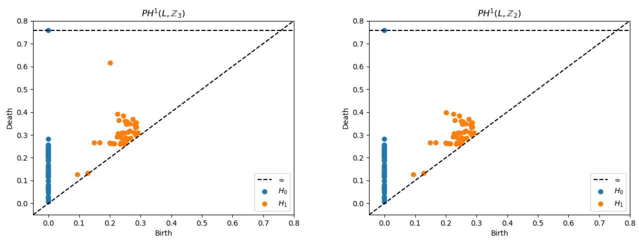


Figure 4: $PH^i(R(L);\mathbb{F})$ for $i = 0, 1$ and $\mathbb{F} = \mathbb{Z}_2, \mathbb{Z}_3$.

We compute $f : M(\mathbb{Z}_3, 1) \longrightarrow L_3^{70}$ analogously to the previous example and obtain a point cloud $f(X) \subset L_3^{70}$. The profile of recovered variance is shown in Table 1. Dimension 2 provides a low dimensional representation of $f(X)$ inside L_3^2 with 70% of recovered variance (Figure 8).

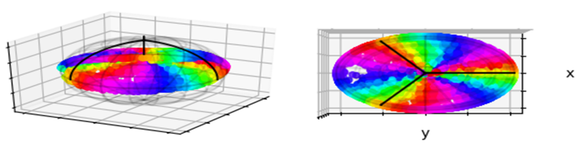


Figure 5: Visualization of the resulting $P_2(f(X)) \subset L_3^2$.

Since f classifies the principal $\mathbb{Z}_3$-bundle P_η over $M(\mathbb{Z}_3, 1)$, then f must be homotopic to the inclusion of $M(\mathbb{Z}_q, 1)$ in L_q^∞. Figure 5 shows $X \subset M(\mathbb{Z}_3, 1)$ mapped by f in L_3^2. Notice the identifications on X are handled by the identification on $S^1 \times \{0\} \subset D^3$ from the fundamental domain on Section 2.2.1. See https://youtu.be/_Ic730_xFkw for a more complete visualization.

6.3 The Lens space $L_3^2 = S^3/\mathbb{Z}_3$.

We use the metric defined in Equation (3) on L_3^2 and randomly sample $15,000$ points to create $X \subset L_3^2$. Figure 6(left) shows the sample set using the fundamental domain from section 2.2.1.

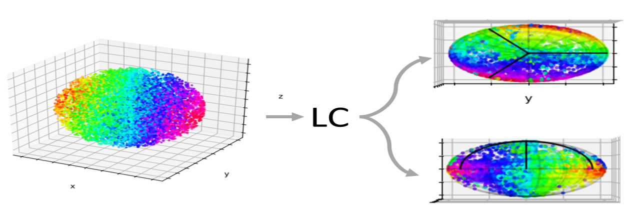


Figure 6: **Left:** $X \subset L_3^2$. **Right:** Lens coordinates.

We can use $PH^i(\mathcal{R}(X);\mathbb{Z}_2)$ and $PH^i(\mathcal{R}(X);\mathbb{Z}_3)$ to verify that the sampled metric space has the expected topological features. Figure 10 contains the corresponding persistent diagrams.

Just as in the previous examples define $f : L_3^2 \to L_3^\infty$ using the most persistent class in $PH^1(\mathcal{R}(L);\mathbb{Z}_3)$. The homotopy class of f must be the same as that of the inclusion $L_3^2 \subset L_3^\infty$, since f classifies the $\mathbb{Z}_3$-principal bundle P_η. Thus we expect L_3^2 to be preserved up to homotopy under LPCA. Figure 6 offers a side and top view of $P_2(f(X)) \subset L_3^2$. Here we clearly see how the original data set X is transformed while preserving the identifications on the boundary of the fundamental domain. Finally in Table 1 we show the variance profile for the dimensionality reduction problem. We see that for dimension 4 we have recovered more than 70% of the total variance as seen in Table 1 and Figure 9.

6.4 Isomap dimensionality reduction

We conclude this section by providing evidence that Lens coordinates (LC) preserve topological features when compared to other dimensionality reduction algorithms. For this purpose we use Isomap ([13]) as our point of comparison.

The Isomap algorithm consist of 3 main steps. The first step determines neighborhoods of each point using k-th nearest neighbors. The second step estimates the geodesic distances between all pairs of points using shortest distance path, and the final step applies classical MDS to the matrix of graph distances.

Let dgm be a persistent diagram. Define per_1 to be the largest persistence of an element in dgm, and let per_2 be the second largest persistence of an element dgm.

$\mathsf{per}_1/\mathsf{per}_2$		$\mathbb{Z}_2$	$\mathbb{Z}_3$
$M(\mathbb{Z}_q, 1)$	Isomap	1.0105	1.0105
	LC	1.7171	3.6789
L_3^2	Isomap	1.0080	1.0080
	LC	1.1592	2.8072

Table 2: In green we highlight the fraction that indicates which method better identifies the topological features.

For both $M(\mathbb{Z}_3, 1)$ and L_3^2 it is clear that the Isomap projection fails to preserve the difference between the cohomology groups with coefficients in $\mathbb{Z}_2$ and $\mathbb{Z}_3$. On the other hand the LC projections maintains this difference in both examples (see Table 4 for more details).

Acknowledgements

This work was partially supported by the NSF under grant DMS-1622301.

References

[1] J. A. Perea and G. Carlsson. A klein-bottle-based dictionary for texture representation. *International Journal of Computer Vision*, 107:75–97, 03 2014. doi: 10.1007/s11263-013-0676-2.

[2] E. H. Brown. Cohomology theories. *Annals of Mathematics*, pages 467–484, 1962.

[3] G. Carlsson. Topological pattern recognition for point cloud data. *Acta Numerica*, 23:289368, 2014. doi: 10.1017/S0962492914000051.

[4] G. Carlsson, T. Ishkhanov, V. Silva, and A. Zomorodian. On the local behavior of spaces of natural images. *Int. J. Comput. Vision*, 76(1): 1–12, Jan. 2008. ISSN 0920-5691. doi: 10.1007/s11263-007-0056-x. URL http://dx.doi.org/10.1007/s11263-007-0056-x.

[5] W. Crawley-Boevey. Decomposition of pointwise finite-dimensional persistence modules. *Journal of Algebra and its Applications*, 14(05):1550066, 2015.

[6] A. Hatcher. *Algebraic topology*. Cambridge University Press, 2002.

[7] D. Husemoller and D. Husemöller. *Fibre Bundles*. Graduate Texts in Mathematics. Springer, 1994. ISBN 9780387940878. URL https://books.google.com/books?id=DPr_BSH89cAC.

[8] J. Milnor. Construction of universal bundles, ii. *Annals of Mathematics*, pages 430–436, 1956.

[9] R. Miranda. *Algebraic curves and Riemann surfaces*, volume 5. American Mathematical Soc., 1995.

[10] J. A. Perea. Sparse Circular Coordinates via Principal $\mathbb{Z}$-Bundles. *arXiv e-prints*, art. arXiv:1809.09269, Sep 2018.

[11] J. A. Perea. A brief history of persistence. *preprint arXiv:1809.03624*, 2018. https://arxiv.org/abs/1809.03624.

[12] J. A. Perea. Multiscale projective coordinates via persistent cohomology of sparse filtrations. *Discrete & Computational Geometry*, 59(1):175–225, 2018.

[13] J. B. Tenenbaum, V. de Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319, 2000.

Appendix

Proof. [of Proposition 3] Take $[b, j, g] \in P_\eta$ and consider a different representative of the class. Namely, an element $(b, k, g + \eta_{jk})$ such that $b \in U_j \cap U_k$. By definition of Φ, we have $\Phi([b, j, g]) = f_j(b, j, g)$ and $\Phi([b, k, g + \eta_{jk}]) = f_k(b, k, g + \eta_{jk})$. And since $f_j(b, j, g) = f_k(b, k, g + \eta_{jk})$, we have that

$$\Phi([b, j, g]) = \Phi([b, k, g + \eta_{jk}]),$$

which shows that Φ is well defined.

To see that Φ is $\mathbb{Z}_q$-equivariant, take $m \in \mathbb{Z}_q$ for any $m = 0, \ldots, q-1$ and compute

$$\begin{aligned}
&\Phi([b, j, g]) \cdot m \\
&\quad = \left[\sqrt{\varphi_1(b)}\zeta_q^{(g+m+\eta_{j1})}, \ldots, \sqrt{\varphi_n(b)}\zeta_q^{(g+m+\eta_{jn})}\right] \\
&\quad = f_j(b, j, g + m) = \Phi([b, j, g + m]) \\
&\quad = \Phi([b, j, g] \cdot m).
\end{aligned}$$

□

Proof. [of Theorem 5] First we need to see that f is well defined. Let $b \in U_j \cap U_k$, therefore

$$\begin{aligned}
p(\Phi([b, j, 0])) &= \left[\sqrt{\varphi_1(b)}\zeta_q^{\eta_{j1}} : \cdots : \sqrt{\varphi_n(b)}\zeta_q^{\eta_{jn}}\right] \\
&= p(\Phi([b, k, 0])).
\end{aligned}$$

This shows that $f(b)$ is independent of the open set containing b.

Hence $(\Phi, f) : (P_\eta, \pi, B) \to (S^{2n-1}, \pi, L_q^n)$ is a morphism of principal $\mathbb{Z}_q$-bundles, and by [[7], Chapter 4: Theorem 4.2] we conclude that P_η and $f^*(S^{2n-1})$ are isomorphic principal $\mathbb{Z}_q$-bundles over B. □

Proof. [of Proposition 6] First of all, $\mathcal{R}_{2\epsilon}(L)^{(0)} = \mathcal{N}(\mathcal{B}_\epsilon)^{(0)} = L$. If $b \notin B_\epsilon(l_k)$, then $\varphi_k(b) = 0$ and therefore the equality holds. If on the other hand $b \in B_\epsilon(l_k) \cap B_\epsilon(l_j)$, then $\{j, k\} \in \mathcal{N}(\mathcal{B}_\epsilon)^{(1)} \subset \mathcal{R}_{2\epsilon}(L)^{(1)}$. In which case, by definition of i^*, we have $\tilde{\eta}_{jk} = \eta_{jk}$. □

Proposition 8 *Let $[x], [y] \in L_q^n$, then*

$$d_L([x], [y]) = d(x, y \cdot \mathbb{Z}_q) = \min_{g \in \mathbb{Z}_q} d(x, y \cdot g).$$

Proof. For $x, y \in \mathbb{C}^n$ let $\langle x, y\rangle_\mathbb{R} := \mathsf{real}(\langle x, y\rangle_\mathbb{C})$. By definition of Hausdorff distance, we have that

$$\begin{aligned}
d_L([x], [y]) = \max \Big\{ &\max_{g \in \mathbb{Z}_q} \min_{h \in \mathbb{Z}_q} \arccos(\langle x \cdot g, y \cdot h\rangle_\mathbb{R}), \\
&\max_{h \in \mathbb{Z}_q} \min_{g \in \mathbb{Z}_q} \arccos(\langle x \cdot g, y \cdot h\rangle_\mathbb{R}) \Big\}.
\end{aligned}$$

Notice that

$$\begin{aligned}
\langle x \cdot g, y \cdot h\rangle_\mathbb{R} &= \mathsf{real}\left(\langle \zeta_q^g x, \zeta_q^h y\rangle_\mathbb{C}\right) \\
&= \mathsf{real}\left(\left\langle x, \zeta_q^{(h-g)} y\right\rangle_\mathbb{C}\right) \\
&= \langle x, y \cdot (h - g)\rangle_\mathbb{R}
\end{aligned}$$

And since $\mathbb{Z}_q$ is Abelian, then

$$\begin{aligned}
&\max_{h\in\mathbb{Z}_q}\min_{g\in\mathbb{Z}_q}\arccos(\langle x\cdot g, y\cdot h\rangle_{\mathbb{R}})\\
&\quad=\max_{h\in\mathbb{Z}_q}\min_{g\in\mathbb{Z}_q}\arccos(\langle x\cdot(g-h), y\rangle_{\mathbb{R}})\\
&\quad=\max_{h\in\mathbb{Z}_q}\min_{g\in\mathbb{Z}_q}\arccos(\langle x\cdot(-h), y\cdot(-g)\rangle_{\mathbb{R}})\\
&\quad=\max_{h'\in\mathbb{Z}_q}\min_{g'\in\mathbb{Z}_q}\arccos(\langle x\cdot h', y\cdot g'\rangle_{\mathbb{R}}).
\end{aligned}$$

Thus

$$d_L([x],[y]) = \max_{g\in\mathbb{Z}_q}\min_{h\in\mathbb{Z}_q}\arccos(\langle x\cdot g, y\cdot h\rangle_{\mathbb{R}}).$$

Furthermore $d_L([x],[y]) = \max_{g\in\mathbb{Z}_q} d(x\cdot g, y\cdot\mathbb{Z}_q) = \max_{g\in\mathbb{Z}_q} d(x, y\cdot(-g)\mathbb{Z}_q)$. Since $y\cdot\big((-g)\mathbb{Z}_q\big) = y\cdot\mathbb{Z}_q$ for any $g\in\mathbb{Z}_q$, we obtain $d_L([x],[y]) = \max_{g\in\mathbb{Z}_q} d(x, y\cdot\mathbb{Z}_q) = d(x, y\cdot\mathbb{Z}_q) = \min_{h\in\mathbb{Z}_q} d(x, y\cdot h)$. □

Proof. [of Lemma 7] From Theorem 8 we know that

$$\begin{aligned}
d_L([v], P_u^{\perp}([v])) &= \min_{g\in\mathbb{Z}_q} d(v, P_u^{\perp}([v])\cdot g)\\
&= \min_{g\in\mathbb{Z}_q} d\left(v, \frac{P_u^{\perp}(v)}{\|P_u^{\perp}(v)\|}\cdot g\right).
\end{aligned}$$

Let $g^* := \operatorname*{argmin}_{g\in\mathbb{Z}_q} d\left(v, \frac{P_u^{\perp}(v)}{\|P_u^{\perp}(v)\|}\cdot g\right)$, so we have

$$d_L([v], P_u^{\perp}([v])) = \arccos\left(\left\langle v, \frac{P_u^{\perp}(v)}{\|P_u^{\perp}(v)\|}\cdot g^*\right\rangle_{\mathbb{R}}\right).$$

Notice that the argument of the arccos can be simplified as follows

$$\begin{aligned}
\left\langle v, \frac{P_u^{\perp}(v)}{\|P_u^{\perp}(v)\|}\cdot g^*\right\rangle_{\mathbb{R}} &= \left\langle \langle v,u\rangle_{\mathbb{C}}u + P_u^{\perp}(v), \frac{P_u^{\perp}(v)}{\|P_u^{\perp}(v)\|}\cdot g^*\right\rangle_{\mathbb{R}}\\
&= \left\langle \langle v,u\rangle_{\mathbb{C}}u, \frac{P_u^{\perp}(v)}{\|P_u^{\perp}(v)\|}\cdot g^*\right\rangle_{\mathbb{R}}\\
&\quad+\left\langle P_u^{\perp}(v), \frac{P_u^{\perp}(v)}{\|P_u^{\perp}(v)\|}\cdot g^*\right\rangle_{\mathbb{R}}.
\end{aligned}$$

since u and $P_u^{\perp}(v)$ are orthogonal in $\mathbb{C}^n$ then they are also orthogonal in $\mathbb{R}^{2n}$, making the then the firs summand on the right hand side equal to zero. Additionally since arccos as a real valued function is monotonically decreasing we have

$$g^* = \operatorname*{argmax}_{g\in\mathbb{Z}_q}\frac{1}{\|P_u^{\perp}(v)\|}\left\langle P_u^{\perp}(v), P_u^{\perp}(v)\cdot g\right\rangle_{\mathbb{R}}.$$

Using the fact that the action of $\mathbb{Z}_q$ is an isometry (and therefore an operator of norm one) as well as the Cauchy-Schwartz inequality we obtain

$$\begin{aligned}
\frac{\langle P_u^{\perp}(v), P_u^{\perp}(v)\cdot g\rangle_{\mathbb{R}}}{\|P_u^{\perp}(v)\|} &\le \left|\frac{1}{\|P_u^{\perp}(v)\|}\left\langle P_u^{\perp}(v), P_u^{\perp}(v)\cdot g\right\rangle_{\mathbb{R}}\right|\\
&\le \frac{1}{\|P_u^{\perp}(v)\|}\|P_u^{\perp}(v)\|\|P_u^{\perp}(v)\cdot g\|\\
&= \|P_u^{\perp}(v)\cdot g\| = \|P_u^{\perp}(v)\|.
\end{aligned}$$

And the equality holds whenever $g = e\in\mathbb{Z}_q$, so we must have $g^* = e$.

Let $[w]\in L_q^{n-1}(u)$, so $w\in \mathsf{span}_{\mathbb{C}}^{\perp}(u)$ which implies that for any $h\in\mathbb{Z}_q$

$$\langle u, w{\cdot}h\rangle_{\mathbb{C}} = \sum_k u_k(\overline{\zeta_q^h w_k}) = \zeta_q^{-h}\sum_k u_k\overline{w_k} = \zeta_q^{-h}\langle u,w\rangle = 0.$$

In other words $w\cdot h\in\mathsf{span}_{\mathbb{C}}^{\perp}(u)$ for any $h\in\mathbb{Z}_q$.

Thus by the Cauchy-Schwartz inequality

$$\begin{aligned}
\langle v, w\cdot h\rangle_{\mathbb{R}} &= \langle\langle v,u\rangle_{\mathbb{C}}u + P_u^{\perp}(v), w\cdot h\rangle_{\mathbb{R}} = \langle P_u^{\perp}(v), w\cdot h\rangle_{\mathbb{R}}\\
&\le |\langle P_u^{\perp}(v), w\cdot h\rangle_{\mathbb{R}}| \le \|P_u^{\perp}(v)\|\|w\cdot h\|\\
&= \|P_u^{\perp}(v)\|\|w\| = \|P_u^{\perp}(v)\|,
\end{aligned}$$

since the action of $\mathbb{Z}_q$ is an isometry and $w\in S^{2n-1}$.

Finally since arccos is decreasing

$$d_L([v], P_u^{\perp}([v])) = \arccos(\|P_u^{\perp}(v)\|) \le \arccos(\langle v, w{\cdot}h\rangle_{\mathbb{R}})$$

for all $h\in\mathbb{Z}_q$, thus $d_L([v], P_u^{\perp}([v])) \le d_L([v],[w])$. □

Visualization map for L_3^2. Given $v_1,\ldots,v_n\in S^{2n-1}$ representatives for the classes in $\mathsf{LPCA}(Y)$. We want to visualize $P_2(Y)\subset L_3^2$ in the fundamental domain described in Section 2.2.1. Let

$$P_2(Y) = \left\{\left[\langle y_i, v_1\rangle_{\mathbb{C}}, \langle y_i, v_2\rangle_{\mathbb{C}}\right]\in S^3\subset\mathbb{C}^2 : [y_i]\in Y\right\}$$

and define $G : P_2(Y)\longrightarrow S^3\subset\mathbb{C}^2$ to be

$$G(z,w) := \left(\zeta_3^{-k}z, \left(\arg(w) - \frac{\pi}{3}\right)\sqrt{1-|z|^2}\right) \tag{4}$$

where $\arg(w)\in\left[0,\frac{2\pi}{3}\right)$, and k an integer such that

$$\arg(z) = k\frac{2\pi}{3} + \theta,$$

where θ is the remainder after division by $\frac{2\pi}{3}$.

Metric on the Moore space $M(\mathbb{Z}_3,1)$. For $x,y\in\mathbb{C}$ with $|x|,|y|\le 1$, we let

$$d(x,y) = \begin{cases}\sqrt{|\langle x,y\rangle_{\mathbb{R}}|} & \text{if } |x|, |w| < 1\\ \min_{\zeta\in\mathbb{Z}_3}\sqrt{|\langle x,\zeta y\rangle_{\mathbb{R}}|} & \text{if } |x| = 1 \text{ or } |w| = 1\\ \min_{\zeta\in\mathbb{Z}_3}\arccos(|\langle x,\zeta y\rangle_{\mathbb{R}}|) & \text{if } |x| = 1 \text{ and } |w| = 1\end{cases}. \tag{5}$$

Profiles of recovered variance.

Recovered variance of LPCA on S^1.

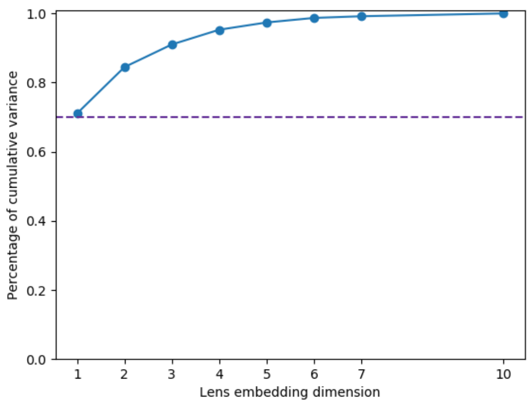


Figure 7: Profile of recovered variance on S^1.

Recovered variance of LPCA on $M(\mathbb{Z}_3, 1)$.

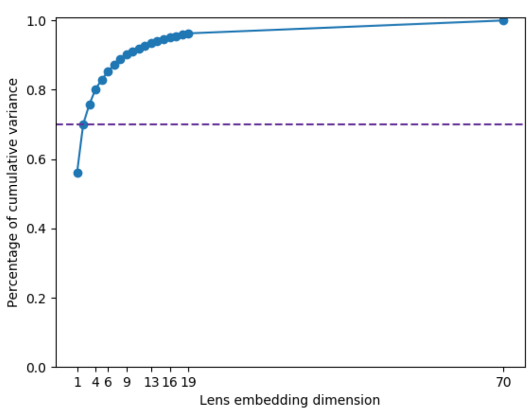


Figure 8: Profile of recovered variance on $M(\mathbb{Z}_3, 1)$.

Recovered variance of LPCA on L_3^2.

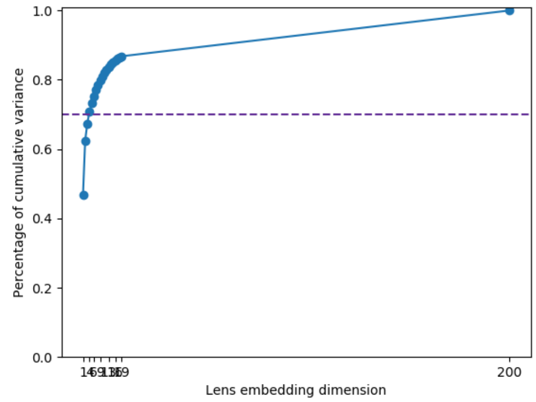


Figure 9: Profile of recovered variance on L_3^2.

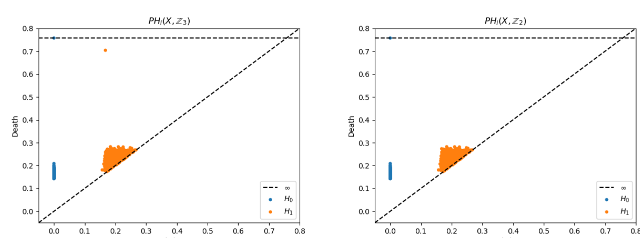


Figure 10: $PH^i(\mathcal{R}(L); \mathbb{Z}_3)$ for $i = 0, 1$. $PH^i(\mathcal{R}(L); \mathbb{Z}_2)$ for $i = 0, 1$.

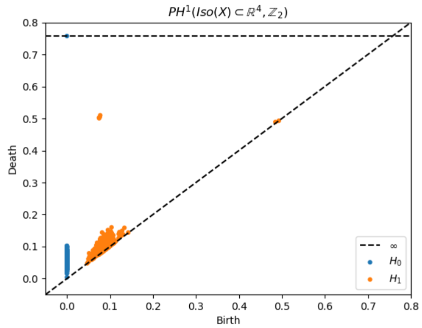
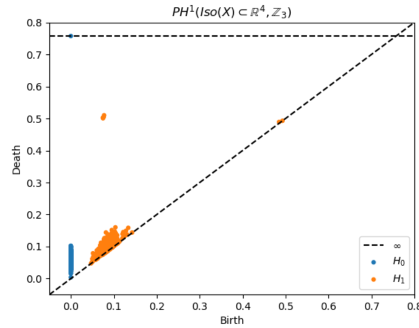
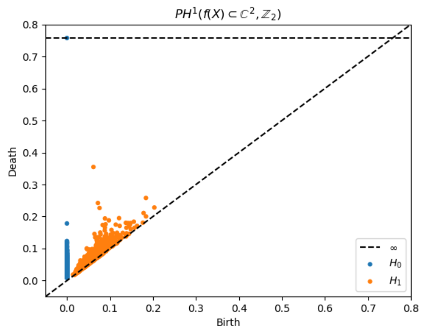
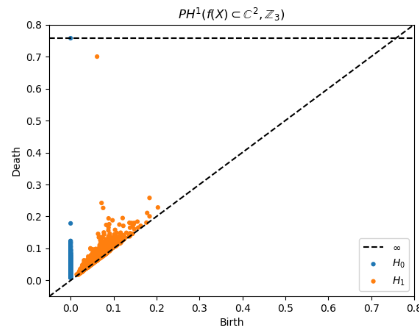
	Coefficients $\mathbb{Z}_2$	Coefficients $\mathbb{Z}_3$
Isomap		
LC		

Table 3: Persistent homology of the Isomap vs. LPCA for $M(\mathbb{Z}_3, 1)$ into a 4 dimensional space.

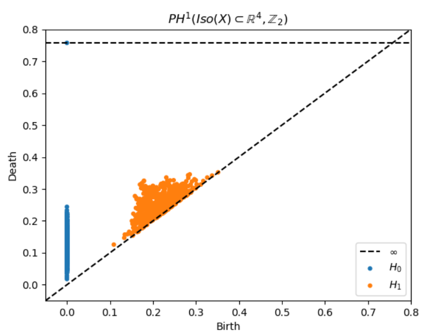
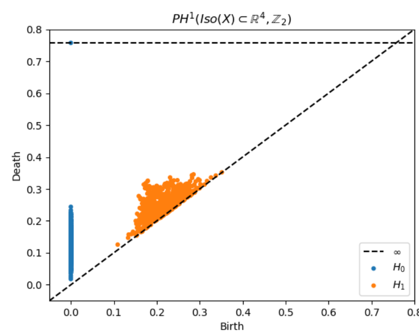
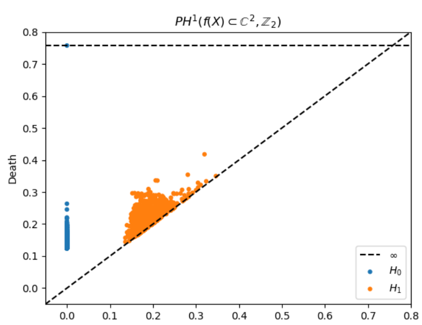
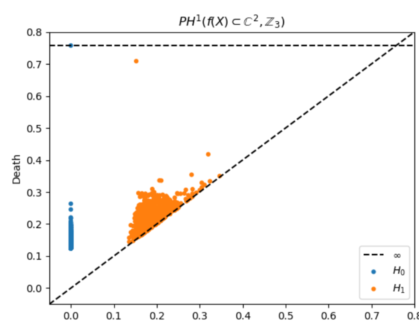
	Coefficients $\mathbb{Z}_2$	Coefficients $\mathbb{Z}_3$
Isomap		
LC		

Table 4: Persistent homology of the Isomap vs. LPCA for L_3^2 into a 4 dimensional space.