Computation of the Barcode of Point Cloud Data with Hodge Decomposition

Dong Du

March 3, 2011

1 Introduction

We call a finite set of points X in $\mathbb{R}^n$ a point cloud data(PCD). A new geometric idea in data analysis is to regard a PCD as a filtration of the simplicial complex Δ^n($n+1$ is the cardinality of the set X) via a familiar construction known as "Rips complex". The first term of the filtration associated to a PCD in the zero-skeleton of Δ^n, the last is Δ^n itself, but the rest give an idea of the qualitative features of the set $X \subseteq \mathbb{R}^n$. The persistent homology of this filtered simplicial complex provides the tool to measure and explain the qualitative patterns of a PCD. The role of Betti numbers, when the homology of a space(simplicial complex) is considered, is taken by **barcodes**, when the persistent homology of a filtered simplicial complex is under consideration. More precisely the invariant "dimension" for a vector space is replaced by the invariant "barcodes" for a persistent vector space. This paper uses elementary Hodge theory to provide algorithms for the calculation of "barcodes" of the persistent homology of a filtered simplicial complex in general but with the case of the filtered simplicial complex provided by PDC in mind.

In section 2, we will discuss persistent linear algebra, including persistent vector spaces and barcodes. In section 3, given a filtered simplicial complex $\mathcal{K}$, we will give the definitions of persistent chain complex $\mathcal{C}(\mathcal{K})$, persistent vector space $\mathcal{H}_r(\mathcal{K})$ and barcode $\mathcal{B}(\mathcal{K})$. In section 4, we will introduce Hodge Decomposition. In section 5, we will use Hodge Decomposition to compute $\mathcal{B}(\mathcal{K})$. In section 6, we will give an algorithm to compute $\mathcal{B}(\mathcal{K})$. In section 7, we will do a numerical experiment.

2 Persistent Linear Algebra

Definition 2.1. 1) A persistent vector space $\mathcal{V}$ is a sequence

$$\{V_n, \varphi_n : V_n \to V_{n+1} | n \in \mathbb{Z}_{\geq 0}\}$$

with V_n vector spaces over a field κ and φ_n linear maps.

A persistent vector space is tame iff each V_n has finite dimension and φ_n is an isomorphisms for n large enough.

2) A linear map of persistent vector spaces $\omega : \mathcal{V} \to \mathcal{W}$, where $\mathcal{V} = \{V_n, \varphi_n : V_n \to V_{n+1} | n \in \mathbb{Z}_{\geq 0}\}$ and $\mathcal{W} = \{W_n, \phi_n : W_n \to W_{n+1} | n \in \mathbb{Z}_{\geq 0}\}$ is a commutative diagram

$$\begin{array}{ccccccccccc}
V_0 & \xrightarrow{\varphi_0} & V_1 & \xrightarrow{\varphi_1} & \cdots & \xrightarrow{\varphi_{n-1}} & V_n & \xrightarrow{\varphi_n} & V_{n+1} & \xrightarrow{\varphi_{n+1}} & \cdots \\
\downarrow{\scriptstyle \omega_0} & & \downarrow{\scriptstyle \omega_1} & & & & \downarrow{\scriptstyle \omega_n} & & \downarrow{\scriptstyle \omega_{n+1}} & & \\
W_0 & \xrightarrow{\phi_0} & W_1 & \xrightarrow{\phi_1} & \cdots & \xrightarrow{\phi_{n-1}} & W_n & \xrightarrow{\phi_n} & W_{n+1} & \xrightarrow{\phi_{n+1}} & \cdots
\end{array}$$

where ω_n is a linear map from V_n to W_n for each $n \geq 0$.

A linear map of persistent vector spaces $\omega : \mathcal{V} \to \mathcal{W}$ is an isomorphism if there exists another linear map of persistent vector spaces $\omega' : \mathcal{W} \to \mathcal{V}$ such that $\omega' \circ \omega : \mathcal{V} \to \mathcal{V}$ and $\omega \circ \omega' : \mathcal{W} \to \mathcal{W}$ are identities. $\mathcal{V}$ and $\mathcal{W}$ are isomorphic if there is an isomorphism between them.

The existence of a linear map $\omega : \mathcal{V} \to \mathcal{W}$ so that each component $\omega_n : \mathcal{V}_n \to \mathcal{W}_n$ is an isomorphism for every $n \geq 0$ implies ω is an isomorphism. One takes $\omega'_n = (w_n)^{-1}$.

3) Let $\mathcal{U} = \{U_n, \psi_n : U_n \to U_{n+1} | n \in \mathbb{Z}_{\geq 0}\}$, $\mathcal{V} = \{V_n, \varphi_n : V_n \to V_{n+1} | n \in \mathbb{Z}_{\geq 0}\}$ and $\mathcal{W} = \{W_n, \phi_n : W_n \to W_{n+1} | n \in \mathbb{Z}_{\geq 0}\}$ be persistent vector spaces. A short exact sequence of persistent vector spaces

$$0 \longrightarrow \mathcal{U} \xrightarrow{\mu} \mathcal{V} \xrightarrow{\nu} \mathcal{W} \longrightarrow 0$$

is a sequence of linear maps of persistent vector spaces such that

$$0 \longrightarrow U_n \xrightarrow{\mu_n} V_n \xrightarrow{\nu_n} W_n \longrightarrow 0$$

is short exact for all $n \geq 0$.

The short exact sequence $0 \longrightarrow \mathcal{U} \xrightarrow{\mu} \mathcal{V} \xrightarrow{\nu} \mathcal{W} \longrightarrow 0$ splits if there exits a linear map $\alpha : \mathcal{V} \to \mathcal{U}$ such that $\alpha \circ \mu = \textit{identity}$ or if there exists a linear map $\beta : \mathcal{W} \to \mathcal{V}$ such that $\nu \circ \beta = \textit{identity}$.

Note. The alternative definitions are equivalent. Given $\alpha : \mathcal{V} \to \mathcal{U}$ such that $\alpha \circ \mu = \textit{identity}$, we can define $\beta : \mathcal{W} \to \mathcal{V}$ such that $\nu \circ \beta = \textit{identity}$ and vice versa.

i)Let $\alpha : \mathcal{V} \to \mathcal{U}$ s.t. $\alpha \circ \mu = \textit{identity}$. For any $x \in W_n$, there exists $x' \in V_n$ such that $\nu_n(x') = x$. Define

$$\begin{aligned} \beta_n : \quad & W_n \to V_n \\ & x \mapsto x' - \mu_n \circ \alpha_n(x'). \end{aligned}$$

To see that β_n is well defined, i.e. independent of the choice of x', consider

$$\begin{aligned} \gamma_n : \quad & V_n \to U_n \oplus W_n \\ & x \mapsto (\alpha_n x, \nu_n x). \end{aligned}$$

γ_n is injective. Indeed if $x \in \ker(\gamma_n)$ hence $\alpha_n x = 0$ and $\nu_n x = 0$. Since $im\mu_n = \ker \nu_n$, there exists $y \in U_n$ such that $x = \mu_n(y)$. Then $y = \alpha_n \circ \mu_n(y) = \alpha_n(x) = 0$, so $x = \mu_n(0) = 0$.

If $x'' \in V_n$ such that $\nu_n(x'') = x$, then γ_n will map both $x' - \mu_n \circ \alpha_n(x')$ and $x'' - \mu_n \circ \alpha_n(x'')$ to $(0, x)$. Since γ_n is injective, $x' - \mu_n \circ \alpha_n(x') = x'' - \mu_n \circ \alpha_n(x'')$ which shows that β_n is well defined.

It is straightforward that $\nu_n \circ \beta_n = \textit{identity}$ and the following diagram is commutative

$$\begin{array}{ccc} W_n & \xrightarrow{\phi_n} & W_{n+1} \\ \downarrow \beta_n & & \downarrow \beta_{n+1} \\ V_n & \xrightarrow{\varphi_n} & V_{n+1} \end{array}$$

ii)Let $\beta : \mathcal{W} \to \mathcal{V}$ with $\nu \circ \beta = \textit{identity}$. For any $x \in V_n$, let $y_x = x - \beta_n \circ \nu_n(x)$, then $\nu_n(y_x) = 0$ and there exist unique $z_x \in \mu_n$ such that $\mu_n(z_x) = y_x$.

Define

$$\begin{aligned} \alpha_n : \quad & V_n \to U_n \\ & x \mapsto z_x \end{aligned}$$

α_n is well defined since for each x, y_x and z_x is unique.

It is straightforward to check $\alpha_n \circ \mu_n = \textit{identity}$ and the following diagram is commutative

$$\begin{array}{ccc} V_n & \xrightarrow{\varphi_n} & V_{n+1} \\ \downarrow \alpha_n & & \downarrow \alpha_{n+1} \\ U_n & \xrightarrow{\psi_n} & U_{n+1}. \end{array}$$

4) Direct sum of a finite collection of persistent vector spaces $\mathcal{V}^i = \{V_n^i, \varphi_n^i : V_n^i \to V_{n+1}^i | n \in \mathbb{Z}_{\geq 0}\}$, $i \in \Lambda(\Lambda$ finite), is defined by $\bigoplus_{i\in\Lambda} \mathcal{V}^i = \{V_n, \varphi_n : V_n \to V_{n+1} | n \in \mathbb{Z}_{\geq 0}\}$, where $V_n = \bigoplus_{i\in\Lambda} V_n^i$, $\varphi_n = \bigoplus_{i\in\Lambda} \varphi_n^i$ for $n \geq 0$. Note that a direct sum of a finite collection of tame persistent vector spaces is tame.

Observation 2.1. *1) Given a split short exact sequence*

$$0 \longrightarrow \mathcal{U} \xrightarrow{\ \mu\ } \mathcal{V} \xrightarrow{\ \nu\ } \mathcal{W} \longrightarrow 0\ ,$$

we have $\mathcal{V} \cong \mathcal{U} \oplus \mathcal{W}$, where $\mathcal{U}$,$\mathcal{V}$,$\mathcal{W}$ and μ,$\nu$are the same as in ***Definition 2.1*** *3).*

2) Not any short exact sequence is split.

Proof. 1) Since the two conditions for a short exact sequence to be split are equivalent, WLOG, we can assume that there exists a linear map $\beta : \mathcal{W} \to \mathcal{V}$ such that $\nu \circ \beta = identity$. Define the linear map

$$\begin{aligned} \delta_n : \quad & U_n \oplus W_n \to V_n \\ & (y, z) \mapsto \mu_n y + \beta_n z \end{aligned}$$

for all $n \geq 0$.

It is easy to check the following diagram is commutative

$$\begin{array}{ccc} U_n \oplus W_n & \xrightarrow{(\psi_n, \phi_n)} & U_{n+1} \oplus W_{n+1} \\ \downarrow{\scriptstyle \delta_n} & & \downarrow{\scriptstyle \delta_{n+1}} \\ V_n & \xrightarrow{\quad \varphi_n \quad} & V_{n+1} \end{array}$$

Given any element $x \in V_n$. Let $z = \nu_n(x)$, then $\nu_n(x - \beta_n z) = z - z = 0$ and there exist unique $y \in U_n$ such that $\mu_n(y) = x - \beta_n z$.

Then $\delta_n(y, z) = \mu_n y + \beta_n z = x - \beta_n z + \beta_n z = x$ and δ_n is surjective.

Let $(y, z) \in \ker \delta_n$, then $\mu_n y + \beta_n z = 0$.

Since $0 = \nu_n(\mu_n y + \beta_n z) = \nu_n \circ \mu_n(y) + \nu_n \circ \beta_n(z) = 0 + z = z$, we have $z = 0$.

Then $\mu_n y = 0$ implies $y = 0$, since μ_n is injective.

Therefore $(y, z) = (0, 0)$ and δ_n is injective.

Hence δ_n is an isomorphism for all $n \geq 0$.

Therefore $\delta : \mathcal{U} \oplus \mathcal{W} \to \mathcal{V}$ is an isomorphism and $\mathcal{V} \cong \mathcal{U} \oplus \mathcal{W}$.

2) Counterexample:

Consider the following short exact sequence:

$$\begin{array}{ccccccccccccc} \mathcal{U} & & \mathbb{R} & \xrightarrow{0} & \mathbb{R} & \xrightarrow{0} & 0 & \xrightarrow{0} & 0 & \xrightarrow{0} & 0 & \xrightarrow{0} & \cdots \\ {\scriptstyle \mu}\downarrow \ \nwarrow {\scriptstyle \alpha} & & \downarrow{\scriptstyle i_1} & & \downarrow{\scriptstyle i_1} & & \downarrow{\scriptstyle 0} & & \downarrow{\scriptstyle 0} & & \downarrow{\scriptstyle 0} & & \\ \mathcal{V} & & \mathbb{R} \oplus \mathbb{R} & \xrightarrow{c} & \mathbb{R} \oplus \mathbb{R} & \xrightarrow{0} & 0 & \xrightarrow{0} & 0 & \xrightarrow{0} & 0 & \xrightarrow{0} & \cdots \\ {\scriptstyle \nu}\downarrow & & \downarrow{\scriptstyle p_2} & & \downarrow{\scriptstyle p_2} & & \downarrow{\scriptstyle 0} & & \downarrow{\scriptstyle 0} & & \downarrow{\scriptstyle 0} & & \\ \mathcal{W} & & \mathbb{R} & \xrightarrow{0} & \mathbb{R} & \xrightarrow{0} & 0 & \xrightarrow{0} & 0 & \xrightarrow{0} & 0 & \xrightarrow{0} & \cdots \end{array}$$

where

$$\begin{aligned} i_1 : \quad & \mathbb{R} \to \mathbb{R} \oplus \mathbb{R}, \\ & x \mapsto (x, 0) \end{aligned} \quad \begin{aligned} p_2 : \quad & \mathbb{R} \oplus \mathbb{R} \to \mathbb{R}, \\ & (x, y) \mapsto y \end{aligned} \quad \begin{aligned} c : \quad & \mathbb{R} \oplus \mathbb{R} \to \mathbb{R} \oplus \mathbb{R} \\ & (x, y) \mapsto (y, 0) \end{aligned}$$

If the sequence splits, in view of $\alpha \circ \mu = identity$, $\alpha_1(x, 0) = x$, which makes the commutativity of the diagram

$$\begin{array}{ccc} \mathbb{R} & \xrightarrow{\quad 0 \quad} & \mathbb{R} \\ \uparrow{\scriptstyle \alpha_0} & & \uparrow{\scriptstyle \alpha_1} \\ \mathbb{R} \oplus \mathbb{R} & \xrightarrow{\quad c \quad} & \mathbb{R} \oplus \mathbb{R} \end{array}$$

impossible.

□

Notation 2.2. Define the following tame persistent vector spaces as **basic tame persistent vector spaces**.

1) $\kappa[t]$ is the tame persistent vector space over a field κ

$$\{V_n, \varphi_n : V_n \to V_{n+1} | n \in \mathbb{Z}_{\geq 0}\}$$

where $V_n = \kappa$ and $\varphi_n = identity$ for all $n \geq 0$.

2) $\sum^r \kappa[t]$ is the tame persistent vector space over a field κ

$$\{V_n, \varphi_n : V_n \to V_{n+1} | n \in \mathbb{Z}_{\geq 0}\}$$

where $V_n = 0$ for $0 \leq n < r$, $V_n = \kappa$ for $n \geq r$, $\varphi_n = 0$ for $0 \leq n < r$ and $\varphi_n = identity$ for $n \geq r$.

3) $\kappa[t]/(t^{r+1})$ is the tame persistent vector space over a field κ

$$\{V_n, \varphi_n : V_n \to V_{n+1} | n \in \mathbb{Z}_{\geq 0}\}$$

where $V_n = \kappa$ for $0 \leq n \leq r$, $V_n = 0$ for $n > r$, $\varphi_n = identity$ for $0 \leq n < r$ and $\varphi_n = 0$ for $n \geq r$.

4) $\sum^r (\kappa[t]/(t^{p+1}))$ is the tame persistent vector space over a field κ

$$\{V_n, \varphi_n : V_n \to V_{n+1} | n \in \mathbb{Z}_{\geq 0}\}$$

where $V_n = \kappa$ for $r \leq i \leq r + p$ and $V_n = 0$ otherwise; $\varphi_n = identity$ for $r \leq i \leq (r + p - 1)$ and $\varphi_n = 0$ otherwise.

The above notations will be justified later.

Notation 2.3. For tame persistent vector space $\mathcal{V} = \{V_n, \varphi_n : V_n \to V_{n+1} | n \in \mathbb{Z}_{\geq 0}\}$

1) Denote $\varphi_{i,j} = \varphi_{j-1} \circ \cdots \circ \varphi_i : V_i \to V_j$ for $i < j$ and $\varphi_{i,i} = identity : V_i \to V_i$, with $i, j \in \mathbb{Z}_{\geq 0}$.

2) Denote $\beta(i, j) = \dim(im(\varphi_{i,j} : V_i \to V_j))$ with $i \leq j \in \mathbb{Z}_{\geq 0}$.

Note $\beta(i, i) = \dim V_i$ and $\beta(i, j) = \beta(i, j + 1)$ for j large enough. If φ_n is an isomorphism for $n \geq N$, denote $\beta(i, \infty) = \beta(i, m)$, where m is any integer larger than i and N.

Definition 2.4. Define an order of all basic tame persistent vector spaces:

1) $\sum^r \kappa[t] < \sum^{r'} \kappa[t]$ if $r < r'$;

2) $\sum^r \kappa[t]/(t^{p+1}) < \sum^{r'} \kappa[t]/(t^{p'+1})$ if $r < r'$ or ($r = r'$ and $p < p'$);

3) $\sum^r \kappa[t]/(t^{p+1}) < \sum^{r'} \kappa[t]$.

Clearly, this order is a strict total order.

Lemma 2.2. *If $\mathcal{V}$ and $\mathcal{W}$ are two basic tame persistent vector spaces such that $\mathcal{V} < \mathcal{W}$, then any map from $\mathcal{V}$ to $\mathcal{W}$ is trivial.*

Proof. We check situation 2) of **Definition 2.4** first.

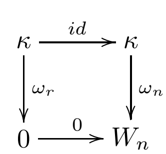
Suppose $\mathcal{V} = \sum^r (\kappa[t]/(t^{p+1}))$ and $\mathcal{W} = \sum^{r'} (\kappa[t]/(t^{p'+1}))$ and there is a linear map $\omega : \mathcal{V} \to \mathcal{W}$. We want to show $\omega = 0$.

We have 2 cases:

Case 1: $r < r'$

In this case, $\omega_n = 0$ for all $n < r$ or $n > r + p$, since $V_n = 0$.

For $r \leq n \leq r + p$, consider the following commutative diagram:



We have $\omega_n = \omega_n \circ id = 0 \circ \omega_r = 0$.

Hence, $\omega = 0$, if $r < r'$.

Case 2: $r = r'$ and $p < p'$

Consider the commutative diagram:

$$\begin{array}{ccc} \kappa & \xrightarrow{0} & 0 \\ \downarrow{\scriptstyle \omega_n} & & \downarrow{\scriptstyle \omega_{r'+p'}} \\ \kappa & \xrightarrow{id} & \kappa \end{array}$$

for $r \leq n \leq r + p$.

The proofs of situations 1) and 3) of **Definition 2.4** are similar.

$\square$

Proposition 2.3. *Any tame persistent vector space* $\mathcal{V} = \{V_n, \varphi_n : V_n \to V_{n+1} | n \in \mathbb{Z}_{\geq 0}\}$ *over a field* κ *is isomorphic to*

$$\bigoplus_{1 \leq i \leq p} \sum^{r_i} \kappa[t] \oplus \bigoplus_{1 \leq j \leq q} \sum^{m_j} (\kappa[t]/(t^{n_j+1}))$$

where p, q, r_i, m_j *and* $n_j \in \mathbb{Z}_{\geq 0}$.

So, any tame persistent vector space can be decomposed in to a direct sum of a finite collection of basic tame persistent vector spaces.

Proof. Since $\mathcal{V}$ is a tame persistent vector space $\dim(V_n)$ is finite for all $n \geq 0$ and there exists $N \geq 0$ such that φ_n is an isomorphisms for $n \geq N$.

Define $T(\mathcal{V}) = \sum_{0 \leq n \leq N} \dim V_n$.

We prove by induction on $T(\mathcal{V})$.

If $T(\mathcal{V}) = 0$, then $\mathcal{V} = 0$. We are done.

If $T(\mathcal{V}) \neq 0$, then $V_k \neq 0$ for some $0 \leq k \leq N$, with $V_i = 0$ if $i < k$.

WLOG, suppose $V_0 \neq 0$ and choose a nonzero element v_0 from V_0.

Define $v_n = \varphi_{0,n}(v_0)$, $n \geq 0$.

Case 1: $v_n \neq 0$ for $0 \leq n < j$ and $v_n = 0$ for $n \geq j$, where $j \geq 1$.

Note. We must have $j \leq N$, otherwise $v_n \neq 0$ for all $n \geq 0$, a contradiction.

Consider the following "short exact sequence" of tame persistent vector spaces

$$\begin{array}{lccccccccccccccc}
\mathcal{W}_0 & \kappa & \xrightarrow{\alpha_0} & \kappa & \xrightarrow{\alpha_1} & \cdots & \xrightarrow{\alpha_{j-1}} & \kappa & \xrightarrow{\alpha_{j-2}} & \kappa & \xrightarrow{\alpha_{j-1}} & 0 & \xrightarrow{\alpha_j} & 0 & \xrightarrow{\alpha_{j+1}} & \cdots \\
 & {\scriptstyle f_0}\downarrow\uparrow{\scriptstyle h_0} & & {\scriptstyle f_1}\downarrow\uparrow{\scriptstyle h_1} & & & & {\scriptstyle f_{j-2}}\downarrow\uparrow{\scriptstyle h_{j-2}} & & {\scriptstyle f_{j-1}}\downarrow\uparrow{\scriptstyle h_{j-1}} & & {\scriptstyle f_j}\downarrow\uparrow{\scriptstyle h_j} & & {\scriptstyle f_{j+1}}\downarrow\uparrow{\scriptstyle h_{j+1}} & & \\
\mathcal{V} & V_0 & \xrightarrow{\varphi_0} & V_1 & \xrightarrow{\varphi_1} & \cdots & \xrightarrow{\varphi_{j-1}} & V_{j-2} & \xrightarrow{\varphi_{j-2}} & V_{j-1} & \xrightarrow{\varphi_{j-1}} & V_j & \xrightarrow{\varphi_j} & V_{j+1} & \xrightarrow{\varphi_{j+1}} & \cdots \\
 & {\scriptstyle g_0}\downarrow & & {\scriptstyle g_1}\downarrow & & & & {\scriptstyle g_{j-2}}\downarrow & & {\scriptstyle g_{j-1}}\downarrow & & {\scriptstyle g_j}\downarrow & & {\scriptstyle g_{j+1}}\downarrow & & \\
\mathcal{V}' & V_0/\kappa v_0 & \xrightarrow{\widetilde{\varphi}_0} & V_1/\kappa v_1 & \xrightarrow{\widetilde{\varphi}_1} & \cdots & \xrightarrow{\widetilde{\varphi}_{j-1}} & V_{j-2}/\kappa v_{j-2} & \xrightarrow{\widetilde{\varphi}_{j-2}} & V_{j-1}/\kappa v_{j-1} & \xrightarrow{\widetilde{\varphi}_{j-1}} & V_j & \xrightarrow{\widetilde{\varphi}_j} & V_{j+1} & \xrightarrow{\widetilde{\varphi}_{j+1}} & \cdots
\end{array}$$

where

$\alpha_n = identity$ for $0 \leq n \leq j-2$ and $\alpha_n = 0$ for $n \geq j-1$;

$$\begin{array}{ll} f_n : & \kappa \to V_n \quad for \ \ 0 \leq n \leq j-1 \\ & \lambda \mapsto \lambda v_n \end{array}$$

and $f_n = 0$ for $n \geq j$;

$$\begin{array}{ll} g_n : & V_n \to V_n/\kappa v_n \quad for \ \ 0 \leq n \leq j-1 \\ & x \mapsto x + \kappa v_n \end{array}$$

and $g_n = identity$ for $n \geq j$;

$$\begin{array}{rll} \widetilde{\varphi}_n : & V_n/\kappa v_n \to V_{n+1}/\kappa v_{n+1} & for \;\; 0 \leq n \leq j-1 \\ & x + \kappa v_n \mapsto \varphi_n(x) + \kappa v_{n+1} & \end{array}$$

and $\widetilde{\varphi}_n = \varphi_n$ for $n \geq j$.

The above diagram is commutative and is a short exact sequence of tame persistent vector spaces. Next we will show the above short exact sequence splits.

Extend v_{j-1} to a basis $\{x_1 = v_{j-1}, x_2, \cdots, x_p\}$ of V_{j-1}.

Given any element $x \in V_{j-1}$, we have a unique expression

$$x = \sum_{i=1}^{p} a_i x_i$$

Define a linear map

$$\begin{array}{rl} h_{j-1} : & V_{j-1} \to \kappa \\ & \sum_{i=1}^{p} a_i x_i \mapsto a_1 \end{array}$$

For $0 \leq n \leq j-1$, define a linear map

$$\begin{array}{rl} h_n : & V_n \to \kappa \\ & x \mapsto h_{j-1} \circ \varphi_{n,j-1}(x) \end{array}$$

For $n \geq j$, define $h_n = 0$.

Clearly $h : \mathcal{V} \to \mathcal{W}_0$ is well defined and $h \circ f = identity$.

Therefore the above short exact sequence splits, and

$$\begin{array}{rl} \mathcal{V} & \cong \mathcal{W}_0 \oplus \mathcal{V}' \\ & = \kappa[t]/(t^j) \oplus \mathcal{V}' \end{array}$$

Hence $T(\mathcal{V}') = T(\mathcal{V}) - j < T(\mathcal{V})$.

Case 2: $v_n \neq 0$ for all $n \geq 0$.

Consider the following "short exact sequence" of tame persistent vector spaces

$$\begin{array}{lcccccccccccc}
\mathcal{W}_1 & \kappa & \xrightarrow{\alpha_0} & \kappa & \xrightarrow{\alpha_1} \cdots \xrightarrow{\alpha_{N-2}} & \kappa & \xrightarrow{\alpha_{N-1}} & \kappa & \xrightarrow{\alpha_N} & \kappa & \xrightarrow{\alpha_{N+1}} \cdots \\
& f_0 \downarrow \uparrow h_0 & & f_1 \downarrow \uparrow h_1 & & f_{N-1} \downarrow \uparrow h_{N-1} & & f_N \downarrow \uparrow h_N & & f_{N+1} \downarrow \uparrow h_{N+1} & \\
\mathcal{V} & V_0 & \xrightarrow{\varphi_0} & V_1 & \xrightarrow{\varphi_1} \cdots \xrightarrow{\varphi_{N-2}} & V_{N-1} & \xrightarrow{\varphi_{N-1}} & V_N & \xrightarrow{\varphi_N} & V_{N+1} & \xrightarrow{\varphi_{N+1}} \cdots \\
& g_0 \downarrow & & g_1 \downarrow & & g_{N-1} \downarrow & & g_N \downarrow & & g_{N+1} \downarrow & \\
\mathcal{V}' & V_0/\kappa v_0 & \xrightarrow{\widetilde{\varphi}_0} & V_1/\kappa v_1 & \xrightarrow{\widetilde{\varphi}_1} \cdots \xrightarrow{\widetilde{\varphi}_{N-2}} & V_{N-1}/\kappa v_{N-1} & \xrightarrow{\widetilde{\varphi}_{N-1}} & V_N/\kappa v_N & \xrightarrow{\widetilde{\varphi}_N} & V_{N+1}/\kappa v_{N+1} & \xrightarrow{\widetilde{\varphi}_{N+1}} \cdots
\end{array}$$

$$\begin{array}{llllllll} where & \alpha_n = identity \; , & f_n : & \kappa \to V_n \; , & g_n : & V_n \to V_n/\kappa v_n & \widetilde{\varphi}_n : & V_n/\kappa v_n \to V_{n+1}/\kappa v_{n+1} \\ & & & \lambda \mapsto \lambda v_n & & x \mapsto x + \kappa v_n & & x + \kappa v_n \mapsto \varphi_n(x) + \kappa v_{n+1} \end{array}$$

for all $n \geq 0$.

We will show the above short exact sequence splits.

Extend v_N to a basis $\{y_1 = v_N, y_2, \cdots, y_q\}$ of V_N.

Given any element $y \in V_N$, we have a unique expression

$$y = \sum_{j=1}^{q} b_j y_j$$

Define a linear map

$$\begin{array}{rl} h_N : & V_N \to \kappa \\ & \sum_{j=1}^{q} b_j y_j \mapsto b_1 \end{array}$$

For $0 \leq n \leq N$, define a linear map

$$\begin{aligned} h_n: \quad & V_n \to \kappa \\ & x \mapsto h_N \circ \varphi_{n,N}(x) \end{aligned}$$

For $n > N$, define a linear map

$$\begin{aligned} h_n: \quad & V_n \to \kappa \\ & x \mapsto h_N \circ \varphi_{N,n}^{-1}(x) \end{aligned}$$

Clearly $h_n \circ f_n = identity$ and

$$\begin{array}{ccc} \kappa & \xrightarrow{\alpha_n} & \kappa \\ \uparrow h_n & & \uparrow h_{n+1} \\ V_n & \xrightarrow{\varphi_n} & V_{n+1} \end{array}$$

is commutative for all $n \geq 0$.

Therefore the short exact sequence in **Case 2** splits, and

$$\begin{aligned} \mathcal{V} & \cong \mathcal{W}_1 \oplus \mathcal{V}' \\ & = \kappa[t] \oplus \mathcal{V}' \end{aligned}$$

Hence $T(\mathcal{V}') = T(\mathcal{V}) - (N+1) < T(\mathcal{V})$.

In both cases $T(\mathcal{V}') < T(\mathcal{V})$.

By induction on $T(\mathcal{V})$,

$$\bigoplus_{1 \leq i \leq s} \sum^{r_i} \kappa[t] \oplus \bigoplus_{1 \leq j \leq t} \sum^{m_j} (\kappa[t]/(t^{n_j+1}))$$

□

Proposition 2.4. *Let*

$$\mathcal{V} = \bigoplus_{1 \leq i \leq p} \sum^{r_i} \kappa[t] \oplus \bigoplus_{1 \leq j \leq q} \sum^{m_j} (\kappa[t]/(t^{n_j+1}))$$

and

$$\mathcal{V}' = \bigoplus_{1 \leq i \leq p'} \sum^{r'_i} \kappa[t] \oplus \bigoplus_{1 \leq j \leq q'} \sum^{m'_j} (\kappa[t]/(t^{n'_j+1}))$$

.

If $\mathcal{V} \cong \mathcal{V}'$, then $p = p'$, $q = q'$ and $r_i = r'_i$, $m_j = m'_j$, $n_j = n'_j$ after a suitable permutation.

Proof. Reorder components in $\mathcal{V}$ and $\mathcal{V}'$ in increasing order (See **Definition 2.4**) and group all copies of the same basic tame persistent vector space together into isotypical components, so we have

$$\mathcal{W} = \bigoplus_{1 \leq n \leq a} W_n$$

and

$$\mathcal{W}' = \bigoplus_{1 \leq n \leq b} W'_n$$

Precisely each isotypical component of $\mathcal{W}$ or $\mathcal{W}'$ is a direct sum of isomorphic basic tame persistent vector spaces.

We only need to prove that $a = b$ and $W_n = W'_n$ for $1 \leq n \leq a = b$.

Define $S(\mathcal{W}) = a$, the cardinality of the isotypical components of $\mathcal{W}$. We will prove the statement by induction on $S(\mathcal{W})$.

If $S(\mathcal{W}) = 0$, clearly $\mathcal{W} = 0 = \mathcal{W}'$.

If $S(\mathcal{W}) > 0$, write $\mathcal{W} = W_1 \oplus R$ and $\mathcal{W}' = P_1 \oplus R'$,

where $R = \oplus_{2 \leq n \leq a} W_n$, $P_1 = \oplus_{1 \leq n \leq b_1} W'_n$, $R' = \oplus_{b_1+1 \leq n \leq b} W'_n$,

b_1 is an integer such that $W'_n \leq W_1$ for $1 \leq n \leq b_1$ and $W'_n > W_1$ for $n > b_1$.

Here the order of W_n and W'_n is determined by the order of their basic components.

Since $\mathcal{W} \cong \mathcal{W}'$, there is a pair of isomorphisms $\omega : \mathcal{W} \to \mathcal{W}'$ and $\omega' : \mathcal{W}' \to \mathcal{W}$ such that $\omega \circ \omega' = id$ and $\omega' \circ \omega = id$.

Write $\omega : \mathcal{W} \to \mathcal{W}'$ as a matrix

$$\begin{array}{cc} & \begin{array}{cc} W_1 & R \end{array} \\ \begin{array}{c} P_1 \\ R' \end{array} & \begin{pmatrix} A & C \\ B & D \end{pmatrix} \end{array}$$

Since any component of $W_1 <$ any component of R', by **Lemma 2.2**, $B = 0$.

So matrix form of ω is

$$\begin{array}{cc} & \begin{array}{cc} W_1 & R \end{array} \\ \begin{array}{c} P_1 \\ R' \end{array} & \begin{pmatrix} A & C \\ 0 & D \end{pmatrix} \end{array}$$

Similarly, ω' has matrix form

$$\begin{array}{cc} & \begin{array}{cc} P_1 & R' \end{array} \\ \begin{array}{c} W_1 \\ R \end{array} & \begin{pmatrix} A' & C' \\ 0 & D' \end{pmatrix} \end{array}$$

Since $\omega \circ \omega' = id$ and $\omega' \circ \omega = id$, we have

$$\begin{pmatrix} A & C \\ 0 & D \end{pmatrix} \begin{pmatrix} A' & C' \\ 0 & D' \end{pmatrix} = I \ \ and \begin{pmatrix} A' & C' \\ 0 & D' \end{pmatrix} \begin{pmatrix} A & C \\ 0 & D \end{pmatrix} = I.$$

So $AA' = I$, $A'A = I$, $DD' = I$, $D'D = I$.

Then $W_1 \cong P_1$ and $R \cong R'$.

Since W_1 and P_1 are isomorphic, each basic component of P_1 must be isomorphic to the basic component of W_1. Their number in W_1 and P_1 should be the same.

So $W_1 = P_1$.

Since $R \cong R'$ and $S(R) = a - 1 < a$, by induction we finish the proof.

□

Definition 2.5. Barcode is a finite collection of intervals

$$[i, j]$$

with $i \in \mathbb{Z}_{\geq 0}$,$j \in \mathbb{Z}_{\geq 0} \bigcup \{\infty\}$ and $i \leq j$.

Given a tame persistent vector space $\mathcal{V}$, there exist a decomposition

$$\mathcal{V} \cong \bigoplus_{1 \leq i \leq p} \sum\nolimits^{r_i} \kappa[t] \oplus \bigoplus_{1 \leq j \leq q} \sum\nolimits^{m_j} (\kappa[t]/(t^{n_j+1}))$$

by **Proposition 2.3**.

Then assign barcode

$$\mathcal{B}(\mathcal{V}) = \{[r_i, \infty], [m_j, m_j + n_j] | 1 \leq i \leq p, 1 \leq j \leq q\}$$

to $\mathcal{V}$. Call it the barcode of the tame persistent vector space $\mathcal{V}$.

This barcode $\mathcal{B}(\mathcal{V})$ is unique by **Proposition 2.4**.

Theorem 2.5. ***Two tame persistent vector spaces are isomorphic iff their barcodes are the same.***

Proof. **Theorem 2.5** is obtained directly from **Proposition 2.3** and **Proposition 2.4**. □

Observation 2.6. *$\beta(i,j)$=number of intervals in $\mathcal{B}(\mathcal{V})$ which contain $[i,j]$ for $i \in \mathbb{Z}_{\geq 0}$,$j \in \mathbb{Z}_{\geq 0} \bigcup \{\infty\}$ and $i \leq j$. In particular,* $\dim(V_i)$*=number of intervals in $\mathcal{B}(\mathcal{V})$ which contain $\{i\}$ for $i \geq 0$.*

Proof. Suppose

$$\mathcal{V} \cong \bigoplus_{1 \leq i \leq p} \sum^{r_i} \kappa[t] \oplus \bigoplus_{1 \leq j \leq q} \sum^{m_j} (\kappa[t]/(t^{n_j+1})) \triangleq \mathcal{W}$$

.

Since $\mathcal{V} \cong \mathcal{W}$, there exist an isomorphism $f : \mathcal{V} \to \mathcal{W}$ and following commutative diagram:

$$\begin{array}{ccccccccccc}
V_0 & \xrightarrow{\varphi_0} & V_1 & \xrightarrow{\varphi_1} & \cdots & \xrightarrow{\varphi_{n-1}} & V_n & \xrightarrow{\varphi_n} & V_{n+1} & \xrightarrow{\varphi_{n+1}} & \cdots \\
\downarrow{\scriptstyle f_0} & & \downarrow{\scriptstyle f_1} & & & & \downarrow{\scriptstyle f_n} & & \downarrow{\scriptstyle f_{n+1}} & & \\
W_0 & \xrightarrow{\phi_0} & W_1 & \xrightarrow{\phi_1} & \cdots & \xrightarrow{\phi_{n-1}} & W_n & \xrightarrow{\phi_n} & W_{n+1} & \xrightarrow{\phi_{n+1}} & \cdots
\end{array}$$

Since $f_j : \mathrm{i}m(\varphi_{i,j} : V_i \to V_j) \to \mathrm{i}m(\phi_{i,j} : W_i \to W_j)$ is an isomorphism, $\beta(i,j) = \dim(\mathrm{i}m(\phi_{i,j} : W_i \to W_j))$.

Observe that $\phi_{i,j}$ is a direct sum of linear maps $\kappa \xrightarrow{id} \kappa$ or $0 \longrightarrow \kappa$ or $\kappa \longrightarrow 0$.$\beta(i,j) = \dim(\mathrm{i}m\phi_{i,j})$ is the number of linear maps $\kappa \xrightarrow{id} \kappa$. Each $\kappa \xrightarrow{id} \kappa$ corresponds to an interval in the barcode $\mathcal{B}(\mathcal{W})$ that contains $[i,j]$.

Therefore $\beta(i,j) =$ the number of intervals in barcode $\mathcal{B}(\mathcal{V})$ that contains $[i,j]$

□

Definition 2.6. Given a tame persistent vector space $\mathcal{V}$, define $\mu(i,j) =$ number of intervals in $\mathcal{B}(\mathcal{V})$ which equal to $[i,j]$.

Observation 2.7. *1)*

$$\mu_(i,j) = \begin{cases} \beta(i,j) - \beta(i-1,j) - \beta(i,j+1) + \beta(i-1,j+1) & 0 < i \leq j < \infty \\ \beta(0,j) - \beta(0,j+1) & i = 0, 0 \leq j < \infty \\ \beta(i,\infty) - \beta(i-1,\infty) & 0 < i < \infty, j = \infty \\ \beta(0,\infty) & i = 0, j = \infty \end{cases}$$

2) $\beta(i,j) = \sum_{l \leq i,\ m \geq j} \mu(l,m)$

Proof. 1) In the case $0 < i \leq j < \infty$, we have

$$\begin{aligned} \mu(i,j) = & \ \#\{\tau \in \mathcal{B}(\mathcal{V}) | \tau = [i,j]\} \\ = & \ \#\{\tau \in \mathcal{B}(\mathcal{V}) | \tau \supseteq [i,j]\} - \#\{\tau \in \mathcal{B}(\mathcal{V}) | \tau \supseteq [i-1,j]\} \\ & -\#\{\tau \in \mathcal{B}(\mathcal{V}) | \tau \supseteq [i,j+1]\} + \#\{\tau \in \mathcal{B}(\mathcal{V}) | \tau \supseteq [i-1,j+1]\} \\ = & \ \beta(i,j) - \beta(i-1,j) - \beta(i,j+1) + \beta(i-1,j+1) \end{aligned}$$

The second identity in the above identities holds because

$$\begin{aligned} & \{\tau \in \mathcal{B}(\mathcal{V}) | \tau = [i,j]\} \\ = & \ \{\tau \in \mathcal{B}(\mathcal{V}) | \tau \supseteq [i,j]\} - \{\tau \in \mathcal{B}(\mathcal{V}) | \tau \supseteq [i-1,j]\} - \{\tau \in \mathcal{B}(\mathcal{V}) | \tau \supseteq [i,j+1]\} \end{aligned}$$

The other three cases are easier to prove.
2) Follows directly from definition.

□

Edelsbrunner-Letscher-Zomorodian's Interpretation [4]

One says that the nonzero element $x \in V_j$ is born in $V_i, i \leq j$ if it is in the image of $\varphi_{i,j}$ and not in the image of $\varphi_{i-1,j}$ and dies in $V_k, k \geq j+1$, if $\varphi_{j,k}(x) = 0$ but $\varphi_{j,k-1} \neq 0$.

A subset $S \subseteq V$, V vector space, is called linearly independent if its elements form a collection of linearly independent vectors in V. If $S_i \subseteq V_i$ is linearly independent in V_i and $\varphi_{i,j}(S_i)$ is linearly independent in V_j, then we call S_i linearly independent on interval $[i, j]$.

Then $\beta(i, j) =$ the maximal cardinality of linearly independent sets on the interval $[i, j]$. $\mu(i, j) =$ the maximal cardinality of linearly independent sets on the interval $[i, j]$, whose elements are born in V_i and die in V_{j+1}.

Each interval $[i, j]$ in $\mathcal{B}(\mathcal{V})$ represents an element in V_i, born in V_i, survives in each $V_r (i \leq r \leq j)$, which dies in V_{j+1}.

Zomorodian - Carlsson's Interpretation [7]

Definition 2.7. 1) A $\kappa[t]-$module V is the vector space over κ equipped with a linear map $A : V \to V$. The module action is defined as

$$\begin{array}{ll} \kappa[t] \times V & \to V \\ (\sum_{i=0}^{n} a_i t^i, v) & \mapsto \sum_{i=0}^{n} a_i A^i(v) \end{array}$$

2) A $\kappa[t]-$module V with a linear map $A : V \to V$ is finitely generated if there exists $\{v_1, v_2, \cdots, v_r\}$ such that any $v \in V$ is a linear combination of

$$\{v_1, v_2, \cdots, v_r, A(v_1), A(v_2), \cdots, A(v_r), A^2(v_1), A^2(v_2), \cdots, A^2(v_r), \cdots\}.$$

3) A graded $\kappa[t]-$module V is a $\kappa[t]-$module together with a decomposition of the vector space $V = \bigoplus_{n \geq 0} V_n$ and a linear map $A : V \to V$ such that $A(v_n) \in V_{n+1},\ \forall v_n \in V_n$.

4) Let $V = \bigoplus_{n \geq 0} V_n$ with the linear map $A : V \to V$ and $W = \bigoplus_{n \geq 0} W_n$ with the linear map $B : V \to V$ be two graded $\kappa[t]-$modules.

A morphism of graded $\kappa[t]-$modules $f : V \to W$ is a linear map of vector spaces such that $f(V_n) \subseteq W_n$ and $f \circ A = B \circ f$.

A morphism of graded $\kappa[t]-$module $f : V \to W$ is an isomorphism if there exists another morphism of graded $\kappa[t]-$module $g : W \to V$ such that $g \circ f : V \to V$ and $f \circ g : W \to W$ are identities. V and W are isomorphic if there is an isomorphism between them.

One of the main results of [7] is the equivalence of the category of finitely generated $\kappa[t]-$modules and the category of tame persistent vector spaces.

Proposition 2.8. *Finitely generated graded $\kappa[t]-$modules identify to tame persistent vector spaces and so do their morphisms.*

Proof. 1) Let $V = \bigoplus_{n \geq 0} V_n$ be a finitely generated $\kappa[t]-$module with a linear map $A : V \to V$ such that $A(v_n) \in V_{n+1},\ \forall v_n \in V_n$.

There exist $\{v_1, v_2, \cdots, v_r\}$ such that any $v \in V$ is a linear combination of

$$(*) \quad \{v_1, v_2, \cdots, v_r, A(v_1), A(v_2), \cdots, A(v_r), A^2(v_1), A^2(v_2), \cdots, A^2(v_r), \cdots\}.$$

Since each v_i is a sum of homogeneous components, WLOG, we can assume $\{v_1, v_2, \cdots, v_r\}$ themselves are homogeneous, i.e., $v_i \in V_{n_i}$, $1 \leq i \leq r$.

Let n_i be the degree of v_i , $1 \leq i \leq r$. WLOG, we can assume $n_i \leq n_{i+1}$.

Let $\varphi_n = A|_{V_n} : V_n \to V_{n+1}$, then

$$\mathcal{V} = \{V_n, \varphi_n : V_n \to V_{n+1} | n \in Z_{\geq 0}\}$$

is a persistent vector space.

To see that $\mathcal{V}$ is tame we denote by $N_0 = \max_{1\leq i\leq r} n_i$. Clearly $\dim V_n (n \leq N_0)$ is finite and in view of $(*)$ the linear maps $\varphi_{n,n+k} = \varphi_{n+k-1} \circ \varphi_{n+k-2} \circ \cdots \circ \varphi_n$ are surjective hence $\dim V_n$ is finite for any $n > N_0$. Since φ_n is surjective for any $n > N_0$ and $\dim V_{N_0}$ is finite, there exists N so that $\dim V_n$ is constant for $n > N$. Since any surjective map between vector spaces of the same finite dimension is an isomorphism, the linear map φ_n is an isomorphism for $n > N$.

Hence, $\mathcal{V}$ is a tame persistent vector space.

If $\mathcal{V} = \{V_n, \varphi_n : V_n \to V_{n+1} | n \in Z_{\geq 0}\}$ is a tame persistent vector space,define the graded $\kappa[t]-$module $V = \bigoplus_{n\geq 0} V_n$ with linear map $A = \bigoplus_{n\geq 0} \varphi_n : V \to V$.

There exist $N \geq 0$ such that φ_n is an isomorphism for $n \geq 0$. Let $\{v_1, v_2, \cdots, v_r\}$ be the set of all generators of $V_1, V_2, \cdots, V_N$, then any $v \in V$ is a linear combination of

$$\{v_1, v_2, \cdots, v_r, A(v_1), A(v_2), \cdots, A(v_r), A^2(v_1), A^2(v_2), \cdots, A^2(v_r), \cdots\}.$$

2) From definitions of morphisms of finitely generated graded $\kappa[t]-$modules and linear maps of tame persistent vector spaces, we can see that two finitely graded $\kappa[t]-$modules are isomorphic iff the tame persistent vector spaces associated with them are isomorphic.

□

Recall that the ring $\kappa[t]$ is a principal ideal domain and a basic theorem in algebra [5] claims that any finitely generated modules over a principal ideal domain decompose uniquely. In particular, any finitely generated modules over $\kappa[t]$ decomposes uniquely as a finite direct sum of $\kappa[t]$'s and $(\kappa[t]/t^{d_i})$'s. As noticed by [7], the above result extends to finitely generated graded modules where the free module $\kappa[t]$ is to be replaced by $\sum^r \kappa[t]$ for some r and torsion module $\kappa[t]/(t^d)$ by $\sum^r \kappa[t]/(t^d)$ for some r. Notice that the module $\sum^r A$ has the component $(\sum^r A)_p = A_{p-r}$ for $p \geq r$ and equal to zero for $p < r$. Note that each component $\sum^r \kappa[t]$ corresponds to a barcode $[r, \infty)$ and each component $\sum^r \kappa[t]/(t^d)$ corresponds to a barcode $[r, r+d]$(cf [7]).

Quiver Representation Perspective

As noticed by G. Carlsson and Vin de Silva [2], persistent vector spaces which stabilize for $n \geq N$ can be regarded as representation of the oriented graph

$$\bullet_1 \longrightarrow \bullet_2 \longrightarrow \bullet_3 \longrightarrow \cdots \longrightarrow \bullet_{N-1} \longrightarrow \bullet_N$$

Any such representation is a sum of indecomposable representations which are classified by the intervals $[i, j], 1 \leq i \leq j \leq N$. The interval $[i, j]$ with $j < N$ correspond to barcode $[i, j]$ while the interval $[i, N]$ to the barcode $[i, \infty)$. The interval $[i, j]$ corresponds to the representation

$$0_1 \longrightarrow 0_2 \longrightarrow \cdots \longrightarrow 0_{i-1} \longrightarrow \kappa_i \xrightarrow{id} \kappa_{i+1} \xrightarrow{id} \cdots \xrightarrow{id} \kappa_j \longrightarrow 0_{j+1} \longrightarrow \cdots$$

This is a well known result in the theory of quiver representation.

3 The Persistent Homology and the Barcodes of a Filtered Simplicial Complex

Definition 3.1. 1) A filtered simplicial complex $\mathcal{K}$ is a finite simplicial complex K together with a filtration

$$(*) \qquad K_0 \subseteq K_1 \subseteq K_2 \subseteq \cdots \subseteq K_N = K.$$

It is convenient to regard this filtration as an "infinite filtration" indexed by $s \in \mathbb{Z}_{\geq 0}$ by taking $K_N = K_{N+1} = K_{N+2} = \cdots$.

2) Given a filtered simplicial complex as above, denote by C_r^s the $\mathbb{R}-$ vector space $C_r(K_s)$ generated by the set of $r-$simplexes of the $s-$th simplicial complex K_s, $\partial_r^s : C_r^s \to C_{r-1}^s$ the boundary map from $C_r(K_s)$ to $C_{r-1}(K_s)$, $i_r^s : C_r^s \to C_r^{s+1}$ the linear map induced by the inclusion from K_s to K_{s+1} (clearly i_r^s is one to one) and obtain the commutative diagram

$$
\begin{array}{ccccccccccccccccc}
\vdots & & \vdots & & & & \vdots & & \vdots & & & & \vdots & & \vdots & & \\
\Big\downarrow{\scriptstyle\partial^0_{M+1}} & & \Big\downarrow{\scriptstyle\partial^1_{M+1}} & & & & \Big\downarrow{\scriptstyle\partial^s_{M+1}} & & \Big\downarrow{\scriptstyle\partial^{s+1}_{M+1}} & & & & \Big\downarrow{\scriptstyle\partial^N_{M+1}} & & \Big\downarrow{\scriptstyle\partial^{N+1}_{M+1}} & & \\
C^0_M & \xrightarrow{i^0_M} & C^1_M & \xrightarrow{i^1_M} & \cdots & \xrightarrow{i^{s-1}_M} & C^s_M & \xrightarrow{i^s_M} & C^{s+1}_M & \xrightarrow{i^{s+1}_M} & \cdots & \xrightarrow{i^{N-1}_M} & C^N_M & \xrightarrow[\cong]{i^N_M} & C^{N+1}_M & \xrightarrow[\cong]{i^{N+1}_M} & \cdots \\
\Big\downarrow{\scriptstyle\partial^0_{M}} & & \Big\downarrow{\scriptstyle\partial^1_{M}} & & & & \Big\downarrow{\scriptstyle\partial^s_{M}} & & \Big\downarrow{\scriptstyle\partial^{s+1}_{M}} & & & & \Big\downarrow{\scriptstyle\partial^N_{M}} & & \Big\downarrow{\scriptstyle\partial^{N+1}_{M}} & & \\
\vdots & & \vdots & & & & \vdots & & \vdots & & & & \vdots & & \vdots & & \\
\Big\downarrow{\scriptstyle\partial^0_{r+1}} & & \Big\downarrow{\scriptstyle\partial^1_{r+1}} & & & & \Big\downarrow{\scriptstyle\partial^s_{r+1}} & & \Big\downarrow{\scriptstyle\partial^{s+1}_{r+1}} & & & & \Big\downarrow{\scriptstyle\partial^N_{r+1}} & & \Big\downarrow{\scriptstyle\partial^{N+1}_{r+1}} & & \\
C^0_r & \xrightarrow{i^0_r} & C^1_r & \xrightarrow{i^1_r} & \cdots & \xrightarrow{i^{s-1}_r} & C^s_r & \xrightarrow{i^s_r} & C^{s+1}_r & \xrightarrow{i^{s+1}_r} & \cdots & \xrightarrow{i^{N-1}_r} & C^N_r & \xrightarrow[\cong]{i^N_r} & C^{N+1}_r & \xrightarrow[\cong]{i^{N+1}_r} & \cdots \\
\Big\downarrow{\scriptstyle\partial^0_{r}} & & \Big\downarrow{\scriptstyle\partial^1_{r}} & & & & \Big\downarrow{\scriptstyle\partial^s_{r}} & & \Big\downarrow{\scriptstyle\partial^{s+1}_{r}} & & & & \Big\downarrow{\scriptstyle\partial^N_{r}} & & \Big\downarrow{\scriptstyle\partial^{N+1}_{r}} & & \\
C^0_{r-1} & \xrightarrow{i^0_{r-1}} & C^1_{r-1} & \xrightarrow{i^1_{r-1}} & \cdots & \xrightarrow{i^{s-1}_{r-1}} & C^s_{r-1} & \xrightarrow{i^s_{r-1}} & C^{s+1}_{r-1} & \xrightarrow{i^{s+1}_{r-1}} & \cdots & \xrightarrow{i^{N-1}_{r-1}} & C^N_{r-1} & \xrightarrow[\cong]{i^N_{r-1}} & C^{N+1}_{r-1} & \xrightarrow[\cong]{i^{N+1}_{r-1}} & \cdots \\
\Big\downarrow{\scriptstyle\partial^0_{r-1}} & & \Big\downarrow{\scriptstyle\partial^1_{r-1}} & & & & \Big\downarrow{\scriptstyle\partial^s_{r-1}} & & \Big\downarrow{\scriptstyle\partial^{s+1}_{r-1}} & & & & \Big\downarrow{\scriptstyle\partial^N_{r-1}} & & \Big\downarrow{\scriptstyle\partial^{N+1}_{r-1}} & & \\
\vdots & & \vdots & & & & \vdots & & \vdots & & & & \vdots & & \vdots & & \\
\Big\downarrow{\scriptstyle\partial^0_{2}} & & \Big\downarrow{\scriptstyle\partial^1_{2}} & & & & \Big\downarrow{\scriptstyle\partial^s_{2}} & & \Big\downarrow{\scriptstyle\partial^{s+1}_{2}} & & & & \Big\downarrow{\scriptstyle\partial^N_{2}} & & \Big\downarrow{\scriptstyle\partial^{N+1}_{2}} & & \\
C^0_1 & \xrightarrow{i^0_1} & C^1_1 & \xrightarrow{i^1_1} & \cdots & \xrightarrow{i^{s-1}_1} & C^s_1 & \xrightarrow{i^s_1} & C^{s+1}_1 & \xrightarrow{i^{s+1}_1} & \cdots & \xrightarrow{i^{N-1}_1} & C^N_1 & \xrightarrow[\cong]{i^N_1} & C^{N+1}_1 & \xrightarrow[\cong]{i^{N+1}_1} & \cdots \\
\Big\downarrow{\scriptstyle\partial^0_{1}} & & \Big\downarrow{\scriptstyle\partial^1_{1}} & & & & \Big\downarrow{\scriptstyle\partial^s_{1}} & & \Big\downarrow{\scriptstyle\partial^{s+1}_{1}} & & & & \Big\downarrow{\scriptstyle\partial^N_{1}} & & \Big\downarrow{\scriptstyle\partial^{N+1}_{1}} & & \\
C^0_0 & \xrightarrow{i^0_0} & C^1_0 & \xrightarrow{i^1_0} & \cdots & \xrightarrow{i^{s-1}_0} & C^s_0 & \xrightarrow{i^s_0} & C^{s+1}_0 & \xrightarrow{i^{s+1}_0} & \cdots & \xrightarrow{i^{N-1}_0} & C^N_0 & \xrightarrow[\cong]{i^N_0} & C^{N+1}_0 & \xrightarrow[\cong]{i^{N+1}_0} & \cdots \\
\Big\downarrow{\scriptstyle\partial^0_{0}} & & \Big\downarrow{\scriptstyle\partial^1_{0}} & & & & \Big\downarrow{\scriptstyle\partial^s_{0}} & & \Big\downarrow{\scriptstyle\partial^{s+1}_{0}} & & & & \Big\downarrow{\scriptstyle\partial^N_{0}} & & \Big\downarrow{\scriptstyle\partial^{N+1}_{0}} & & \\
0 & & 0 & & & & 0 & & 0 & & & & 0 & & 0 & &
\end{array}
\tag{3.1}
$$

Note that since K is finite dimensional the rows become eventually 0, i.e. $C_r^s = 0$ if $r > \dim K$, each of the vector spaces C_r^s is finite dimensional and each row is a persistent vector space.

Definition 3.2. 1) Passing to homology, i.e. consider $\mathcal{H}_r^s = \ker(\partial_r^s)/\mathrm{im}(\partial_{r+1}^s)$, i_r^s induce the linear maps $\widetilde{i}_r^s : \mathcal{H}_r^s \to \mathcal{H}_r^{s+1}$. For each r

$$
\mathcal{H}_r(\mathcal{K}) := \{\mathcal{H}_r^s, \widetilde{i}_r^s : \mathcal{H}_r^s \to \mathcal{H}_r^{s+1} | s \in \mathbb{Z}_{\geq 0}\}
$$

is a persistent vector space with barcode $\mathcal{B}(\mathcal{H}_r(\mathcal{K}))$.

Following [ELZ], the vector space $\mathcal{H}_r^{s,p} = \mathrm{im}(\widetilde{i_r}^{s,s+p} : \mathcal{H}_r^s \to \mathcal{H}_r^{s+p})$ represent the persistent homology.

2) The collection of barcodes $\mathcal{B}(\mathcal{H}_r(\mathcal{K}))$ for all r is denoted by $\mathcal{B}(\mathcal{K})$ and referred to as the barcode of $\mathcal{K}$.

4 Hodge Decomposition of a Chain Complex

The following consideration follow the standard "Hodge decomposition" familiar in Riemannian geometry. This finite dimensional elementary formulation was first used by B. Eckmann [3]. One start with a complex of finite dimensional vector spaces

$$\cdots \xrightarrow{\partial_{r+2}} C_{r+1} \xrightarrow{\partial_{r+1}} C_r \xrightarrow{\partial_r} C_{r-1} \xrightarrow{\partial_{r-1}} \cdots \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\partial_0} C_{-1} = 0 ,$$

each of them equipped with a positive definite inner product.

In our situation, each complex comes from a finite simplicial complex L, with C_r the $\mathbb{R}-$vector space generated by the $r-$simplices of L(hence C_r is a vector space equipped with a base indexed by the oriented $r-$simplices) , and ∂_r the linear maps known as the boundary maps.

We consider the inner product on C_r which makes the standard basis orthonormal, so we can view $\partial_{r+1} : C_{r+1} \to C_r$ as an linear operator between two inner product spaces. Let $\delta_r = \partial^*_{r+1}$ be the adjoint operator of ∂_{r+1}.

We begin with:

Lemma 4.1. *If A and B are two inner product spaces, $f : A \to B$ a linear map and $f^* : B \to A$ its adjoint, i.e. $\langle b, f(a)\rangle = \langle f^*(b), a\rangle$ for any $a \in A, b \in B$, then*
i) $\ker(f) = (im(f^*))^\perp$;
ii) For any $a \in A$, if $f \circ f^* \circ f(a) = 0$ then $f(a) = 0$.

Define $\Delta_r = \partial_{r+1} \circ \delta_r + \delta_{r-1} \circ \partial_r : C_r \to C_r$ for $r \in \mathbb{Z}_{\geq 0}$, $(C_r)_+ = \mathrm{im}(\partial_{r+1})$, $(C_r)_- = \mathrm{im}(\delta_{r-1})$ and $H_r = \ker(\Delta_r)$.

Proposition 4.2. *i)* $H_r = \ker(\delta_r) \cap \ker(\partial_r)$;
ii) ***(Hodge Decomposition)*** $C_r = (C_r)_+ \oplus H_r \oplus (C_r)_-$*, where $(C_r)_+$, H_r and $(C_r)_-$ are pairwise orthogonal;*
iii) There is an isomorphism $j_r : H_r = \ker(\delta_r) \cap \ker(\partial_r) \to \mathcal{H}_r = \ker(\partial_r)/im(\partial_{r+1})$.

We will give an algorithm for calculating Hodge decomposition, i.e., given a chain complex $\mathcal{C}(L)$, we will calculate the three orthogonal projections:

$$(p_r)_+ : C_r \to C_r \quad with \ (p_r)_+(C_r) = (C_r)_+$$

$$(p_r)_- : C_r \to C_r \quad with \ (p_r)_-(C_r) = (C_r)_-$$

and

$$(p_r)_H : C_r \to C_r \quad with \ (p_r)_H(C_r) = H_r$$

for each $r \geq 0$ respectively.

Lemma 4.3. *Given any $m \times n$ matrix A over $\mathbb{R}$, if the rank of A is k then there exists a $m \times k$ matrix $[A]$ of the form*

$$[A] = (v_1, v_2, \cdots , v_k)_{m\times k} \tag{4.1}$$

where $\{v_1, v_2, \cdots , v_k\}$ is a set of orthonormal column vectors which is equivalent to the set of column vectors of A. Two sets of vectors are equivalent if they generate the same subspace. Moreover, there is a canonical construction of such orthonormal column vectors known as Gram-Schmidt Orthonormalization.

Proof: Choose k linearly independent column vectors of A, then apply the Gram-Schmidt orthonormalization.

Note. 1. Given a linear map $A : \mathbb{R}^n \to \mathbb{R}^m$ and view A as a $m \times n$ matrix with respect to the standard basis of $\mathbb{R}^n$ and $\mathbb{R}^m$, the set of column vectors of $[A]$ represents an orthonormal basis of $\mathrm{im}(A)$.

2. Although $[A]$ is not unique, $[A][A]^T$ is (See **Lemma 4.4**).

Lemma 4.4. *Given a linear map $A : \mathbb{R}^n \to \mathbb{R}^m$, the linear map*

$$\begin{aligned} p_A : \quad & \mathbb{R}^m \to \mathbb{R}^m \\ & y \mapsto [A][A]^T y \end{aligned}$$

is the orthogonal projection on $im(A)$.

Proposition 4.5. *Given a chain complex $\mathcal{C}(L)$ of a finite simplicial complex L over $\mathbb{R}$*

$$\cdots \xrightarrow{\partial_{r+2}} C_{r+1} \xrightarrow{\partial_{r+1}} C_r \xrightarrow{\partial_r} C_{r-1} \xrightarrow{\partial_{r-1}} \cdots \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\partial_0} C_{-1} = 0$$

Each ∂_r can be regarded as an $n_{r-1} \times n_r$ matrix with respect to the standard basis of C_r and C_{r-1}. The following linear maps are orthogonal projections onto $(C_r)_+$,$(C_r)_-$ and H_r, respectively.

$$\begin{aligned} (p_r)_+ : \quad & C_r && \to C_r \\ & y && \mapsto [\partial_{r+1}][\partial_{r+1}]^T y \\ (p_r)_- : \quad & C_r && \to C_r \\ & y && \mapsto [(\partial_r)^T][(\partial_r)^T]^T y \\ (p_r)_H : \quad & C_r && \to C_r \\ & y && \mapsto (I_{n_r} - [\partial_{r+1}][\partial_{r+1}]^T - [(\partial_r)^T][(\partial_r)^T]^T) y \end{aligned}$$

The linear map

$$\begin{aligned} k_r : \quad \mathcal{H}_r = \ker(\partial_r)/\mathrm{im}(\partial_{r+1}) &\to H_r = \ker(\delta_r) \cap \ker(\partial_r) \\ y + \mathrm{im}(\partial_{r+1}) &\mapsto (p_r)_H(y) \end{aligned}$$

is the inverse of j_r, which verifies **Proposition 4.2** iii).

Suppose now $\mathcal{K}$ is a filtered simplicial complex as in the previous section. For each s consider the chain complex $\mathcal{C}(K_s)$

$$\cdots \xrightarrow{\partial^s_{r+2}} C^s_{r+1} \xrightarrow{\partial^s_{r+1}} C^s_r \xrightarrow{\partial^s_r} C^s_{r-1} \xrightarrow{\partial^s_{r-1}} \cdots \xrightarrow{\partial^s_2} C^s_1 \xrightarrow{\partial^s_1} C^s_0 \xrightarrow{\partial^s_0} = 0$$

equipped with the scalar products defined by the standard basis.

Let

(1) $\delta^s_{r-1} = (\partial^s_r)^*$ the adjoint operator of ∂^s_r for $r \in \mathbb{Z}_{\geq 0}$,

(2) $\Delta^s_r = \partial^s_{r+1} \circ \delta^s_r + \delta^s_{r-1} \circ \partial^s_r : C^s_r \to C^s_r$ for $r \in \mathbb{Z}_{\geq 0}$ and

(3) $(C^s_r)_+ = \mathrm{im}(\partial^s_{r+1})$, $(C^s_r)_- = \mathrm{im}(\delta^s_{r-1})$ and $H^s_r = \ker(\Delta^s_r)$, with

$$H^s_r = \ker(\delta^s_r) \cap \ker(\partial^s_r)$$

and

$$C^s_r = (C^s_r)_+ \oplus H^s_r \oplus (C^s_r)_-$$

where $(C^s_r)_+$, H^s_r and $(C^s_r)_-$ are pairwise orthogonal.

With respect to the standard basis, ∂^s_r can be regarded as an $n^s_{r-1} \times n^s_r$ matrix. In view of the above considerations, the following linear maps are orthogonal projections onto $(C^s_r)_+$,$(C^s_r)_-$ and H^s_r, respectively.

$$\begin{aligned} (p^s_r)_+ : \quad & C^s_r && \to C^s_r \\ & y && \mapsto [\partial^s_{r+1}][\partial^s_{r+1}]^T y \\ (p^s_r)_- : \quad & C^s_r && \to C^s_r \\ & y && \mapsto [(\partial^s_r)^T][(\partial^s_r)^T]^T y \\ (p^s_r)_H : \quad & C^s_r && \to C^s_r \\ & y && \mapsto (I_{n^s_r} - [\partial^s_{r+1}][\partial^s_{r+1}]^T - [(\partial^s_r)^T][(\partial^s_r)^T]^T) y \end{aligned} \tag{4.2}$$

with

$$j_r^s: \quad H_r^s = \ker(\delta_r^s) \cap \ker(\partial_r^s) \quad \to \mathcal{H}_r^s = \ker(\partial_r^s)/\mathrm{im}(\partial_{r+1}^s)$$
$$x \quad \mapsto x + \mathrm{im}(\partial_{r+1}^s)$$

and

$$k_r^s: \quad \mathcal{H}_r^s = \ker(\partial_r^s)/\mathrm{im}(\partial_{r+1}^s) \quad \to H_r^s = \ker(\delta_r^s) \cap \ker(\partial_r^s)$$
$$y + \mathrm{im}(\partial_{r+1}^s) \quad \mapsto (p_r^s)_H(y)$$

isomorphisms between H_r^s and $\mathcal{H}_r^s$.

5 Computation of the Barcode of a Filtered Simplicial Complex

Given a filtered simplicial complex $\mathcal{K}$, we are interested in the persistent vector space

$$\mathcal{H}_r(\mathcal{K}) = \{\mathcal{H}_r^s, \widetilde{i}_r^s : \mathcal{H}_r^s \to \mathcal{H}_r^{s+1} | s \in \mathbb{Z}_{\geq 0}\}$$

(Refer to **Definition 3.2**).

Since $\mathcal{H}_r^s$ identifies to H_s^r(we have the pair of isomorphisms $j_r^s : H_r^s \to \mathcal{H}_r^s$ and $k_r^s : \mathcal{H}_r^s \to H_r^s$), we replace the persistent vector space $\mathcal{H}_r(\mathcal{K})$ by the isomorphic persistent vector space

$$H_r(\mathcal{K}) = \{H_r^s, g_r^s : H_r^s \to H_r^{s+1} | s \in \mathbb{Z}_{\geq 0}\},$$

where $g_r^s = k_r^{s+1} \circ \widetilde{i}_r^s \circ j_r^s$.

Notation 5.1. 1) Denote

$$g_r^{s,t} = g_r^{t-1} \circ \cdots \circ g_r^s : H_r^s \to H_r^t$$

for $s < t$ and

$$g_r^{s,s} = identity : H_r^s \to H_r^s$$

with $s, t \in \mathbb{Z}_{\geq 0}$.

Define $i_r^{s,t} : C_r^s \to C_r^t$ and $\widetilde{i}_r^{s,t} : \mathcal{H}_r^s \to \mathcal{H}_r^t$ in similar way.

2) Denote $\beta_r(s,t) = \dim(im(g_r^{s,t} : H_r^s \to H_r^t))$ with $s \leq t \in \mathbb{Z}_{\geq 0}$.

Note. $\beta_r(s,s) = \dim H_r^s$ and $\beta_r(s,t) = \beta_r(s,t+1)$ for $t \geq N$.

Hence $\beta_r(s,\infty) = \beta_r(s,m)$ for m larger than s and N.

Notice that $\beta_r(s,t)$=number of intervals in $\mathcal{B}(H_r(\mathcal{K}))$ which contain $[s,t]$ for $0 \leq s \leq t \leq \infty$. (See **Observation 2.6**)

3) Denote $\mu_r(s,t)$=number of intervals in $\mathcal{B}(H_r(\mathcal{K}))$ which equal to $[s,t]$ for $0 \leq s \leq t \leq \infty$. We have

$$\mu_r(s,t) = \begin{cases} \beta_r(s,t) - \beta_r(s-1,t) - \beta_r(s,t+1) + \beta_r(s-1,t+1) & 0 < s \leq t < \infty \\ \beta_r(0,t) - \beta_r(0,t+1) & s = 0, 0 \leq t < \infty \\ \beta_r(s,\infty) - \beta_r(s-1,\infty) & 0 < s < \infty, t = \infty \\ \beta_r(0,\infty) & s = 0, t = \infty \end{cases} \quad (5.1)$$

Note. If we know $\beta_r(s,t)$ for all $0 \leq s \leq t \leq N$ then from the formula above we know $\mu_r(s,t)$ and the barcode $\mathcal{B}(H_r(\mathcal{K}))$.

Theorem 5.1. *$\beta_r(s,s) = rank((p_r^s)_H)$ and $\beta_r(s,t) = rank((p_r^t)_H \circ i_r^{s,t} \circ (p_r^s)_H)$ for $s < t$, where*

$$i_r^{s,t} = \begin{pmatrix} I_{n_r^s} \\ 0_{(n_r^t - n_r^s) \times n_r^s} \end{pmatrix}.$$

For the definition of $(p_r^s)_H$, refer to (4.1) *and* (4.2).

We will use MATLAB and in order to avoid the use of the function "rank" which is not very reliable for large matrices we will need the following observations to complement **Theorem 5.1**.

Observation 5.2. *The rank of a real matrix A equals to the number of positive eigenvalues of AA^T or A^TA.*

Observation 5.3. $\beta_r(s,s) = \dim(H_r^s) = \dim(C_r^s) - rank(\partial_{r+1}^s) - rank(\partial_r^s)$.

Observation 5.4. *If $\beta_r(s,t) \neq 0, s < t$, then*
i) H_r^i is nonzero for all $s \leq i \leq t$;
ii) $\beta_r(s,i)$ is nonzero for all $s \leq i \leq t$.

6 Algorithm to Compute the Barcode of Point Cloud Data

1) Point Cloud Data and Rips Complex of PCD.

We call a finite set of points X in $\mathbb{R}^n$ a point cloud data, PCD for short.(See Example1.1, [7])

Given $\epsilon \geqslant 0$, the Rips complex $R_\epsilon(X)$ of PCD X has X as the vertex set. We declare a set of vertices $\sigma = [x_0, x_1, \ldots, x_k]$ a $k-$simplex of $R_\epsilon(X)$ iff $d(x_i, x_j) \leq \epsilon$ for all pairs $x_i, x_j \in \sigma$. (See Example1.1, [7])

Notice that Rips complex will be determined by its one skeleton, so it is a flag complex.

There is an obvious inclusion $R_\epsilon(X) \hookrightarrow R_{\epsilon'}(X)$ when $\epsilon < \epsilon'$.

Since a PCD X is a finite set, Rips complex $R_\epsilon(X)$ will change at only finitely many epsilons

$$0 = \epsilon_0 < \epsilon_1 < \cdots < \epsilon_N. \tag{6.0}$$

We define the filtered simplicial complex of Rips complexes of PCD X as

$$R_{\epsilon_0}(X) \subseteq R_{\epsilon_1}(X) \subseteq \cdots \subseteq R_{\epsilon_N}(X) = R_{\epsilon_N}(X) = \cdots . \tag{6.1}$$

(See Example1.1, [7])

In order to restrict the size of the Rips complex $R_\epsilon(X)$, we will work with $R_\epsilon(X, m)$, the $m-$skeleton of $R_\epsilon(X)$, when concerned with the persistent homology up to dimension $m - 1$.

If we consider the $m-$skeleton of Rips complexes, we still get the restricted filtered simplicial complex

$$R_{\epsilon_0}(X,m) \subseteq R_{\epsilon_1}(X,m) \subseteq \cdots \subseteq R_{\epsilon_N}(X,m) = R_{\epsilon_N}(X,m) = \cdots . \tag{6.2}$$

The restriction will have the same persistent homology and barcodes up to dimension $m - 1$.

Given a positive integer number P, consider the restricted filtered simplicial complex

$$R_{\epsilon_0}(X,m) \subseteq R_{\epsilon_1}(X,m) \subseteq \cdots \subseteq R_{\epsilon_P}(X,m) = R_{\epsilon_P}(X,m) = \cdots . \tag{6.3}$$

The barcode of PCD X is the barcode of the filtered simplicial complex provided by the sequence (6.1).

The barcode of PCD X up to dimension $m - 1$ is the barcode of the restricted filtered simplicial complex (6.2).

The **barcode of PCD X up to dimension $m - 1$ and step P** is the barcode of the restricted filtered simplicial complex (6.3).

Next, we will describe the algorithm to compute the barcode of PCD X up to dimension $m - 1$ and step P.

Given a point cloud data $X = \{x_1, x_2, \cdots, x_p\} \in \mathbb{R}^n$(suppose all points in X are distinct), we will first consider its distance matrix

$$D = \begin{pmatrix} d(x_1,x_1) & d(x_1,x_2) & \cdots & d(x_1,x_p) \\ d(x_2,x_1) & d(x_2,x_2) & \cdots & d(x_2,x_p) \\ \vdots & \vdots & \ddots & \vdots \\ d(x_p,x_1) & d(x_p,x_2) & \cdots & d(x_p,x_p) \end{pmatrix}$$

Notice that the D is a symmetric matrix, the diagonal elements of D are zeros and all the other elements are positive real numbers.

Let S be the set of the entries in D and order these numbers increasingly (since they will be our epsilons), so we have

$$\mathcal{E} : 0 = \epsilon_0 < \epsilon_1 < \epsilon_2 < \cdots < \epsilon_N.$$

$\mathcal{E}$ is exactly (6.0).

Given a positive integer m and P, let's compute the barcode of PCD X up to dimension $m-1$ and step P.

To simplify the notation, from now on we will write K_i instead of $R_{\epsilon_i}(X, m)$. K_0 is just a collection of vertices $\{x_1, x_2, \cdots, x_p\}$.

We will abbreviate the vertex x_{i_0} simply by the number i_0, the edge $[x_{i_0}, x_{i_1}]$ by $[i_0, i_1]$ and $k-$simplex $[x_{i_0}, \cdots, x_{i_k}]$ is represented by $[i_0, i_1, \cdots, i_k]$.

We want first to have the information about the filtered simplicial complex stored as

(a) An $(m+1) \times (P+1)$ matrix named **Dimension** where the entry $\mathbf{Dimension}_{r+1,s+1}$ is the integer $n_r^s =$ dimension of C_r^s (See **Definition 2.1**).

(b) An $(m+1)$ array of matrices named **Simplex**. The component $\mathbf{Simplex}_{r+1}$ is a matrix describing $r-$simplices of K_P; precisely is a matrix of $r+1$ columns, each row representing an $r-$simplex $[i_0, i_1, \cdots, i_k]$. These rows are ordered consistently with the filtration. The order for simplices of $K_s \backslash K_{s-1}$ is however arbitrary.

There is an obvious way to create these data, which is not very efficient but worth to mention. Given K_s, we can easily find its $1-$simplices from the distance matrix D. All the nonzero numbers in D which are less than ϵ_s correspond to the edges in K_s. In order to find the $r-$simplices, we scan all possible combination of $r+1$ numbers of $\{1, 2, \cdots, p\}$. $[i_0, i_1, \cdots, i_r]$ $(i_0 < i_1 < \cdots < i_r)$belongs to the $r-$simplices of K_s iff every pair $[i_j, i_k]$ belongs to the $1-$simplices of K_s. As s increases, we have more $r-$ simplices in K_s. Store each $r-$ simplex as a row in $\mathbf{Simplex}_{r+1}$ in filtration order. In another word, if $[i_0, i_1, \cdots, i_r]$ belongs to K_s but not in K_{s-1}, then it is "new" and can be added as a new row to the rows provided by the $r-$simplices of K_{s-1}.

We use however the package **JPlex** [1] [6] to compute **Dimension** and **Simplex** , which can be download from http://comptop.stanford.edu/programs/jplex/.

Once we get **Simplex**, we get the standard basis of C_r^P, which is stored in $\mathbf{Simplex}_{r+1}$, for $0 \le r \le m$. Given $\mathbf{Simplex}_r$ and $\mathbf{Simplex}_{r+1}$, we compute ∂_r^P with respect to the standard basis:

Write $\mathbf{Simplex}_{r+1}$ as

$$\begin{pmatrix} a_{11} & \cdots & a_{1,r+1} \\ \vdots & \ddots & \vdots \\ a_{n_r^P,1} & \cdots & a_{n_r^P,r+1} \end{pmatrix}$$

and $\mathbf{Simplex}_r$ as

$$\begin{pmatrix} b_{11} & \cdots & b_{1,r} \\ \vdots & \ddots & \vdots \\ b_{n_{r-1}^P,1} & \cdots & a_{n_{r-1}^P,r} \end{pmatrix},$$

where n_r^P and n_{r-1}^P can be obtained from **Dimension**.

∂_r^P is an $n_{r-1}^P \times n_r^P$ matrix with element $(\partial_r^P)_{i,j}$ equals to the incidence number of the simplices $[b_{i,1}, \cdots, b_{i,r}]$ and $[a_{j,1}, \cdots, a_{j,r+1}]$. If $[b_{i,1}, \cdots, b_{i,r}]$ is not contained in $[a_{j,1}, \cdots, a_{j,r+1}]$, $(\partial_r^P)_{i,j} = 0$. Otherwise, $[b_{i,1}, \cdots, b_{i,r}]$ equals to $[a_{j,1}, \cdots, \hat{a_{j,k}}, \cdots, a_{j,r+1}]$ for some k, where " $\hat{}$ " means deletion. If k is odd, $(\partial_r^P)_{i,j} = 1$; if k is even, $(\partial_r^P)_{i,j} = -1$.

∂_r^s is the $n_{r-1}^s \times n_r^s$ upper-left block of ∂_r^P.

Once **Dimension** and **Simplex** are determined, we apply **Theorem 5.1** and the observations **5.2, 5.3 and 5.4** to get $\beta_r(s,t)$ and then $\mu_r(s,t)$ according to(5.1).

Here are the brief description of functions, all of them are written in Matlab(for details see my program). Go to `http://www.math.ohio-state.edu/~dudong`, download `"Computing Barcode of PCD.zip"`. Then unzip it and put the folder in **Matlab** direction.

Initial Data:

A $p \times n$ matrix **X**, which stores p points $x_1, \cdots, x_p$ in $\mathbb{R}^n$ as rows, two positive integers **m** and **S**. **m** is the given upper bound of dimension of restricted Rips Complex. **S** is the given upper bound of steps of filtration. $\mathbf{P} = \min(\mathbf{S}, \mathbf{N})$. See (6.3).

1. **function getDistance**

Input: **X**.

Output: A p-by-p upper triangular matrix **D**. $\mathbf{D}_{ij} = d(x_i, x_j), 1 \le i \le p-1, i+1 \le j \le p$.

2. **function scaleX**

Input: **X**, **D**.

Output: **X**.

Given a point cloud data **X**, we find all the distances between two points and sort them as $\epsilon_0 < \epsilon_1 < \cdots < \epsilon_N$. **Note:** If $\epsilon_{i+1} - \epsilon_i$ is very small(less than $3 \cdot$ **diam(X)** $\cdot$ **eps**, the upper bound of round off errors, where **diam(X)** is the maximal distance between any two points of **X**, **eps** $= 2.2204 \times 10^{-16}$ is floating-point relative accuracy in Matlab), we delete ϵ_{i+1}. We scale **X** such that the minimum difference between two consecutive distances is larger than 3×10^{-4}. For details and more explanation, read the comments in my program.

3. **function getEpsilon**

Input: **D**, **S**.

Output: Two positive integers $\mathbf{P} = \min(\mathbf{S}, \mathbf{N})$ and $\mathbf{e} = \epsilon_P$. Two row vectors **epsiorg**, which stores different distances $\epsilon_0 < \epsilon_1 < \cdots < \epsilon_N$ of **D** in increasing order and **epsiavg**, which stores

$$\frac{\epsilon_0 + \epsilon_1}{2}, \cdots, \frac{\epsilon_k + \epsilon_{k+1}}{2}, \cdots, \frac{\epsilon_{N-1} + \epsilon_N}{2}, \epsilon_N + \frac{1}{2}.$$

4. **function getDimensionSimplex**

Input: **X**, **P**, **epsiavg**, **m** and **e**.

Output:

a) An integer $\mathbf{m_a} = \dim(R_{\epsilon_P}(X, m))$.

b) An $(\mathbf{m_a} + 1) \times (\mathbf{P} + 1)$ matrix **Dimension**. $\mathbf{Dimension}_{r+1,s+1} = n_r^s$, the dimension of C_r^s.

c) An $\mathbf{m_a} + 1$ array **Simplex**.

The component $\mathbf{Simplex}_{r+1}$ is a matrix with $r+1$ columns, which stores $r-$simplex of $R_{\epsilon_P}(X, m)$ as rows in filtration order.

Note: Here we use the package **edu.stanford.math.plex.*** of **JPlex**[1][6].

5. **function getDeltaP**

Input: **Dimension**, **Simplex** and $\mathbf{m_a}$.

Output: An $\mathbf{m_a}$ array **DeltaP**. $\mathbf{DeltaP}_r =$ the matrix of $\partial_r^P : C_r^P \to C_{r-1}^P$ with respect to the the standard basis of C_r^P and C_{r-1}^P.

6. **function getDelta**

Input: **DeltaP**, **Dimension**, r, s

Output: The matrix $\partial_r^s : C_r^s \to C_{r-1}^s$ with respect to standard basis of C_r^s and C_{r-1}^s.

7. **function getRank**
Input: **P**, **m**, and $\mathbf{m_a}$.
Output: An $(\mathbf{m}+1)\times(\mathbf{P}+1)$ matrix **R**. $\mathbf{R}_{r+1,s+1}$ = rank of ∂_r^s.
Algorithm: $\mathbf{R}_{r+1,s+1}$ = number of eigenvalues of $\partial_r^s \times {\partial_r^s}'$ which is greater than 10^{-8}. Here we use **Observation 5.2**.

8. **function getDimensionHrs**
Input: **Dimension**, **R**, **P**, **m** and $\mathbf{m_a}$.
Output: An integer $\mathbf{m_c}$ and an $(\mathbf{m_c}+1)\times(\mathbf{P}+1)$ matrix **dimHrs**. $\mathbf{dimHrs}_{r+1,s+1} = \dim(H_r^s)$.
Algorithm: $\mathbf{dimHrs}_{r+1,s+1} = \mathbf{Dimension}_{r+1,s+1} - \mathbf{R}_{r+2,s+1} - \mathbf{R}_{r+1,s+1}$.
Here we use use **Observation 5.3**.

9. **function getHarmonicProjection**
Input: **Dimension**, r and s.
Output: The harmonic projection matrix $(p_r^s)_H$ from C_r^s to itself with respect to the standard basis of C_r^s.
Algorithm: See (4.2).

10. **function getBeta**
Input: **Dimension**, **P**, **dimHrs** and $\mathbf{m_c}$. $\mathbf{m_c}$ is the maximal dimension of nonzero H_r^s's.
Output: A $(\mathbf{P}+1)\times(\mathbf{P}+1)\times(\mathbf{m_c}+1)$ matrix **Beta**. $\mathbf{Beta}(s+1,t+1,r+1) = \beta_r(s,t)$.
Algorithm:

$$\beta_0(\cdot,\cdot) = \begin{pmatrix} \dim(H_0^0) & \dim(H_0^1) & \cdots & \dim(H_0^P) \\ & \dim(H_0^1) & \cdots & \dim(H_0^P) \\ & & \ddots & \vdots \\ & & & \dim(H_0^P) \end{pmatrix};$$

$\beta_r(s,s) = rank((p_r^s)_H) = \dim(H_r^s)$(see **Observation 5.3**);
$\beta_r(s,t) = rank((p_r^t)_H \circ i_r^{s,t} \circ (p_r^s)_H)$ for $s<t$ (see **Theorem 5.1**).
Apply **Observation 5.2** when computing rank as in **function getRank**.
Apply **Observation 5.4** to avoid unnecessary computing.

11. **function getMu**
Input: **Beta**, **P** and $\mathbf{m_c}$.
Output: A $(\mathbf{P}+1)\times(\mathbf{P}+1)\times(\mathbf{m_c}+1)$ matrix **Mu**. $\mathbf{Mu}(s+1,t+1,r+1) = \mu_r(s,t)$
Algorithm: Apply (5.1).

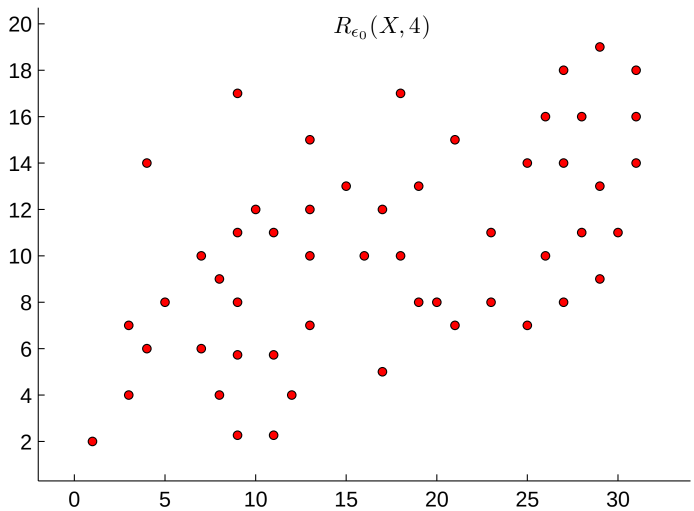
12. **function getBarcodeMatrix**
Input: **Mu**, **P**, and $\mathbf{m_c}$.
Output: **A**, a $q\times 4$ matrix which stores the barcode in increasing order according to dimension, left endpoint and right endpoint. M intervals $[a,b]$ in barcode $\mathcal{B}(\mathcal{H}_r(\mathcal{K}))$ will be represented by [r a b M] in **A**.

13. **drawBarcode**
Input: **A** and **P**.
Output: A picture of barcode stored in **barcode.eps**.

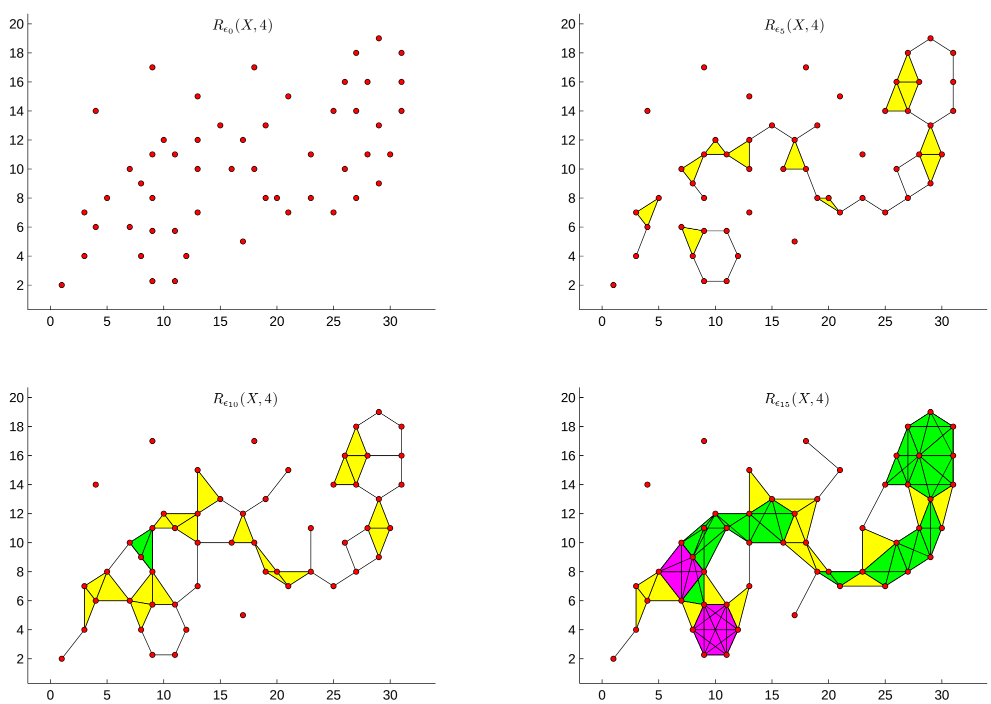
14. **function main**
Input: **X**, **S** and **m**.
Output: A matrix **A** stored in **A.mat** and a picture of barcode stored in **barcode.eps**.
Algorithm: Run above functions except 6. and 9. in order.

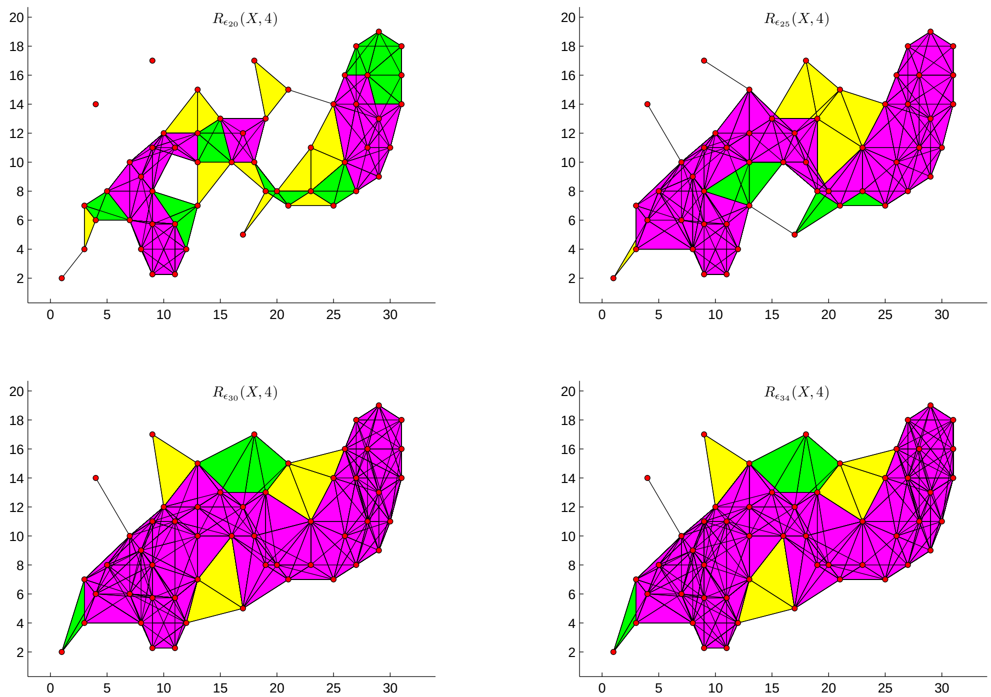
7 Numerical Experiment

We have PCD in $\mathbb{R}^2$ with 53 points:

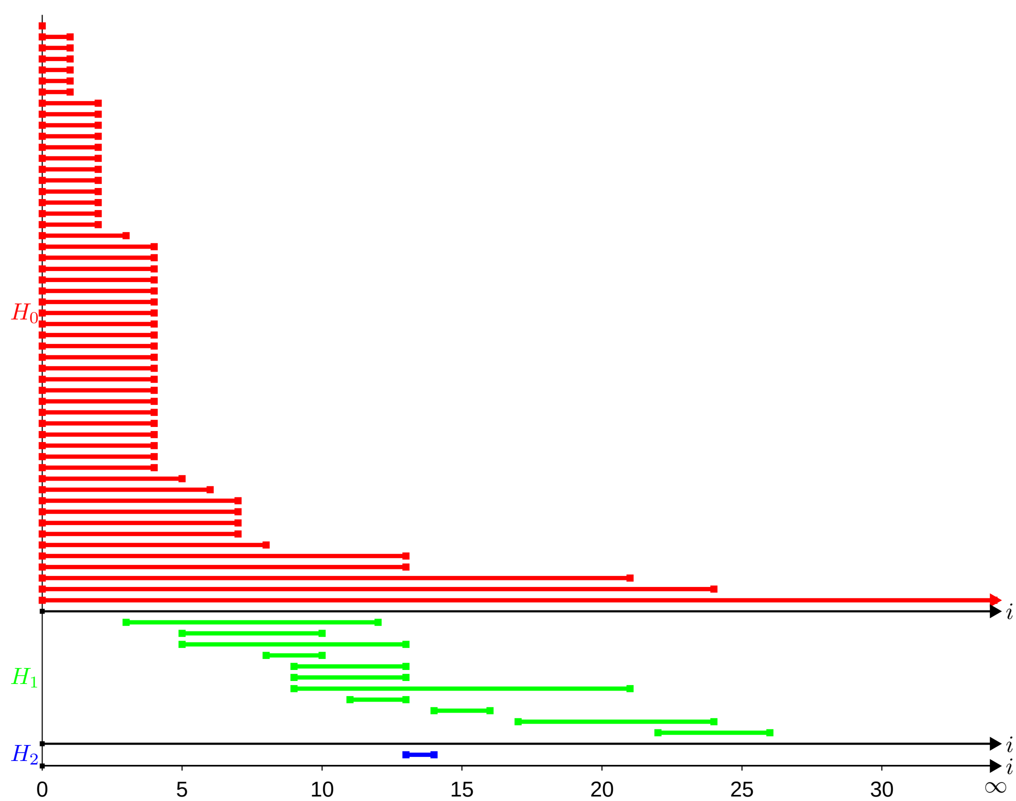


The restricted filtered simplicial complex of this PCD ($m = 4$, $P = 34$) is





And the barcode of above filtered simplicial complex is



References

[1] Henry Adams. JPlex with Matlab Tutorial.

[2] G. Carlsson and V. D. Silva. Zigzag Persistence. *Foundations of Computational Mathematics*, 10(4): 367 - 405, 2010.

[3] B. Eckmann. Harmonische Funktionen und Randvertanfgaben in einem Komplex. *Commentarii Math. Helvetici*, 17, 240-245, (1944-45).

[4] H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological persistence and simplification. *Discrete Comput. Geom.*, 28: 511-533, 2002.

[5] S. Lang. Chapter III Modules, Section 7 Modules over principal rings. *Algebra(GTM 211) Revised Third Edition*, 146-155, 2002.

[6] Harlan Sexton and Mikael Vejdemo Johansson. JPlex, a Java software package for computing the persistent homology of filtered simplicial complexes.

[7] A. Zomorodian and G. Carlsson. Computing Persistent Homology. *Discrete Comput. Geom.*, 33: 249 - 274, 2005.