Homology through Interleaving

32 Concept of interleaving

A discrete set $P \subset \mathbb{R}^k$ is assumed to be a sample of a set $X \subset \mathbb{R}^k$, if it lies near to it which we can quantify with the Hasudorff distance $d_H(P, X)$. Observe that small $d_H(P, X)$ does not imply that P necessarily lie in X. It can be around X.

Our goal is to examine Čech and Rips complexes built on top of P for inferring the homology of X. We achieve this goal by the following steps:

1. Consider the distance function to X, $d_X : \mathbb{R}^k \to \mathbb{R}$, $x \mapsto d(x, X)$, and the distance function to the sample P, $d_P : \mathbb{R}^k \to \mathbb{R}$, $x \mapsto d(x, P)$.

2. Let $X_\alpha := d_X^{-1}(-\infty, \alpha]$ and $P_\alpha := d_P^{-1}(-\infty, \alpha]$ be the α-offsets of X and P respectively. Observe that P_α is the union of a set of balls with centers in P and radii α.

3. Observe that, for sufficiently small $\alpha < \alpha'$, X_α and X'_α are homotopy equivalent. In fact, α' can be 0 when X is a compact manifold with positive weak feature size.

4. Argue that the sequence of X_α and P_α interleave, that is, for appropriate $\alpha_1 \le \alpha_2 \le \ldots \le \alpha_5$
$$X_{\alpha_1} \subseteq P_{\alpha_2} \subseteq X_{\alpha_3} \subseteq P_{\alpha_4} \subseteq X_{\alpha_5}. \tag{9}$$

5. Use Nerve theorem to establish that the Čech complex $\mathcal{C}^\alpha(P)$ which is the nerve of P_α is homotopy equivalent to P_α. There exists an homotopy equivalence that commutes with the inclusions at the homology level. This will be clear later.

6. Because of 5, we have an interleaving sequence of homomorphisms at the homology level from the sequence in 9:
$$\mathsf{H}(X_{\alpha_1}) \to \mathsf{H}(\mathcal{C}^{\alpha_2}) \to \mathsf{H}(X_{\alpha_3}) \to \mathsf{H}(\mathcal{C}^{\alpha_4}) \to \mathsf{H}(X_{\alpha_5}).$$

7. From the sequence in 6, one can derive that the image of $\mathsf{H}(\mathcal{C}^{\alpha_2}) \to \mathsf{H}(\mathcal{C}^{\alpha_4})$ is isomorphic to $\mathsf{H}(X_{\alpha_3})$.

8. Now use the interleaving between Čech and Rips filtrations to derive that the persistent homology between two Rips complexes is isomorphic to the homology of an offset of X.

33 Data on a compact set

First we consider a point data P that presumably samples a compact subset X of $\mathbb{R}^k$. It is known that an offset X_α for any $\alpha > 0$ may not be homotopy equivalent to X when X is compact. So, in this case we will consider capturing the homology groups of an offset X_α of X. We will need the following definitons for stating the precise results.

Definition 55. Let $X \subset \mathbb{R}^k$ be a compact set. Let M denote the medial axis of X and C be the set of critical points of the distance function $d_X : \mathbb{R}^k \to \mathbb{R}$, $x \mapsto d(x, X)$. The reach $\rho(X)$ and the weak feature size wfs(X) are defined as:

$$\begin{aligned} \rho(X) &= \inf_{x \in X} d(x, M) \\ \mathrm{wfs}(X) &= \inf_{x \in X} d(x, C). \end{aligned}$$

The following result says that the offsets of X remain homotopically equivalent and hence possess isomorphic homology groups as long as the intervals do not contain crtical points of d_X.

Proposition 50. *If $0 < \alpha < \alpha'$ are such that there is no critical value of d_X in the closed interval $[\alpha, \alpha']$, then X_α deformation retracts onto X_α. In particular, $\mathsf{H}(X_\alpha) \simeq \mathsf{H}(X_{\alpha'})$.*

We will need the following useful fact.

Fact 11. *Given a sequence $A \to B \to C \to D \to E \to F$ of homomorphisms between finite-dimensional vector spaces, if* rank $(A \to F) =$ rank $(C \to D)$*, then this quantity also equals the rank of $B \to E$. Similarly, if $A \to B \to C \to E \to F$ is a sequence of homomorphisms such that* rank $(A \to F) = \dim C$*, then* rank $(B \to E) = \dim C$.

Proposition 51. *Let P be finite set in $\mathbb{R}^k$, such that $d_H(X, P) < \varepsilon$ for some $\varepsilon < \frac{1}{4}\mathrm{wfs}(X)$. Then, for all $\alpha, \alpha' \in [\varepsilon, \mathrm{wfs}(X) - \varepsilon]$ such that $\alpha' - \alpha \geq 2\varepsilon$, and for all $\in (0, \mathrm{wfs}(X))$, we have $\mathsf{H}(X_\lambda) \simeq$ i*mage* i_**, where $i_* : \mathsf{H}(P_\alpha) \to \mathsf{H}(P_{\alpha'})$ is the homomorphism between homology groups induced by the canonical inclusion $i : P_\alpha \to P_{\alpha'}$.*

P . Assume without loss of generality that $\varepsilon < \alpha < \alpha' - 2\varepsilon < \mathrm{wfs}(X) - 3\varepsilon$, since otherwise we can replace ε by any $\varepsilon' \in (d_H(X, P), \varepsilon)$. From the hypothesis we deduce the following sequence of inclusions:

$$X_{\alpha-\varepsilon} \subseteq P_\alpha \subseteq X_{\alpha+\varepsilon} \subseteq P_{\alpha'} \subseteq X_{\alpha'+\varepsilon} \tag{10}$$

By Proposition 50, for all $0 < \beta < \beta' < \mathrm{wfs}(X)$, the canonical inclusion $X_\beta \to X'_\beta$ is a homotopy equivalence. As a consequence, Eq.(10) induces a sequence of homomorphisms between homology groups, such that all homomorphisms between homology groups of $X_{\alpha-\varepsilon}, X_{\alpha+\varepsilon}, X_{\alpha'+\varepsilon}$ are isomorphisms. It follows then from Fact 11 that $i_* : \mathsf{H}(P_\alpha) \to \mathsf{H}(P_{\alpha'})$ has same rank as these isomorphisms. Now, this rank is equal to the dimension of $\mathsf{H}(X_\lambda)$, since the X_β are homotopy equivalent to X for all $0 < \beta < \mathrm{wfs}(X)$. It follows that i*mage* $i_* \simeq \mathsf{H}(X_\lambda)$, since our ring of coefficients is a field. □

The above proposition relates the homology of X_λ with the persistent homology between two union of balls. We can go to the nerve of the union of balls, that is, the Čech complexes if we know that the folliowing diagram commutes. The downward vertical arrows are isomorphisms induced by the homotopy equivalence due to the nerve theorem. The horizontal arrows are induced by inclusions.

$$\begin{array}{ccc} \mathsf{H}(P_\alpha) & \xrightarrow{i_*} & \mathsf{H}(P_{\alpha'}) \\ \downarrow h_* & & \downarrow h_* \\ \mathsf{H}(\mathcal{C}^\alpha(P)) & \xrightarrow{i_*} & \mathsf{H}(\mathcal{C}^{\alpha'}(P)) \end{array}$$

Chazal and Oudot [1] showed that the above diagram commutes. Then, we have the following Proposition.

Proposition 52. *Let P be finite set in $\mathbb{R}^k$, such that $d_H(X,P) < \varepsilon$ for some $\varepsilon < \frac{1}{4}\mathrm{wfs}(X)$. Then, for all $\alpha, \alpha' \in [\varepsilon, \mathrm{wfs}(X) - \varepsilon]$ such that $\alpha' - \alpha \geq 2\varepsilon$, and for all $\in (0, \mathrm{wfs}(X))$, we have $\mathsf{H}(X_\lambda) \simeq$ image i_*, where $i_* : \mathsf{H}(\mathcal{C}^\alpha(P)) \to \mathsf{H}(\mathcal{C}^{\alpha'}(P))$ is the homomorphism between homology groups induced by the canonical inclusion $i : \mathcal{C}^\alpha(P) \to \mathcal{C}^{\alpha'}(P)$.*

Theorem 53. *Let P be a finite point set such that $d_H(X,P) < \varepsilon$ for some $\varepsilon < 1/9\mathrm{wfs}(X)$. Then, for all $\alpha \in [2\varepsilon, \frac{1}{4}(\mathrm{wfs}(X) - \varepsilon)]$ and all $\lambda \in (0, \mathrm{wfs}(X))$, we have $\mathsf{H}(X_\lambda) \simeq$ image j_*, where j_* is the homomorphism between homology groups induced by the canonical inclusion $j : \mathcal{R}^\alpha(P) \to \mathcal{R}^{4\alpha}(P)$.*

P . We have already seen the following sequence:

$$\mathcal{C}^{\alpha/2}(P) \to \mathcal{R}^\alpha(P) \to \mathcal{C}^\alpha(P) \to \mathcal{C}^{2\alpha}(P) \to \mathcal{R}^{4\alpha}(P) \to \mathcal{C}^{4\alpha}(P). \tag{11}$$

Since $2\varepsilon \leq \alpha \leq \frac{1}{4}(\mathrm{wfs} - \varepsilon)$, by Proposition 52 this sequence of inclusions induces a sequence of homomorphisms between homology groups, such that $\mathsf{H}(\mathcal{C}^{\alpha/2}(P)) \to \mathsf{H}(\mathcal{C}^{4\alpha}(P))$ and $\mathsf{H}(\mathcal{C}^\alpha(P)) \to \mathsf{H}(\mathcal{C}^{2\alpha}(P))$ have ranks equal to dim $\mathsf{H}(X_\lambda)$. Hence, by Proposition 11, rank j_* is also equal to dim $\mathsf{H}(X_\lambda)$. It follows that i*mage* $j_* \simeq \mathsf{H}(X_\lambda)$. □

34 Data on manifold

When X is a smooth manifold, the above results can be slightly improved. The main observation is that, for manifolds, the homology of union balls indeed become isomorphic to that of the manifold. Therefore, one does not need to go through the persistent homology between two Čech complexes to capture the homology of X. Instead, one can compute the homology of a single Čech complex to obtain that of X. The following result due to Niyogi, Smale, Weinberger [4] is key for this observation.

Proposition 54. *Let $P \subset X$ be such that $d_H(X,P) \leq \varepsilon$ where $X \subset \mathbb{R}^k$ is a smooth manifold. If $2\varepsilon \leq \alpha \leq \sqrt{\frac{3}{5}}\rho(X)$, there is a deformation retraction from P_α to X which implies that $H(\mathcal{C}^\alpha(P))$ is isomorphic to $H(X)$.*

Now we can state a result similar to Theorem 53 where we use Proposition 54 instead of Proposition 52.

Theorem 55. *Let P be a finite point set such that $d_H(X,P) < \varepsilon$ for some $\varepsilon < 1/9\mathrm{wfs}(X)$. Then, for all $\alpha \in [2\varepsilon, \frac{1}{2}(\mathrm{wfs}(X) - \varepsilon)]$ and all $\lambda \in (0, \mathrm{wfs}(X))$, we have $\mathsf{H}(X_\lambda) \simeq$ image j_*, where j_* is the homomorphism between homology groups induced by the canonical inclusion $j : \mathcal{R}^\alpha(P) \to \mathcal{R}^{2\alpha}(P)$.*

P . The proof is exactly same as the proof of Theorem 53 except that the sequence in (11) is shrunk by one Čech complex in the middle:

$$\mathcal{C}^{\alpha/2}(P) \to \mathcal{R}^\alpha(P) \to \mathcal{C}^\alpha(P) \to\to \mathcal{R}^{2\alpha}(P) \to \mathcal{C}^{2\alpha}(P).$$

Now, apply the second part of Proposition 11 to obtain the stated result. □

35 Interleaving of towers and stability

We have considered filtration of simplicial complexes so far for illustrating persistence and stability of its diagram. In a filtration, the connecting maps between consecutive complexes are inclusions. Assuming a discrete subset of reals, $I : a_0 \leq a_1 \leq \cdots \leq a_n$, as index set, we can write a filtration as:

$$\{\mathcal{K}_a\}_{a\in I} : \mathcal{K}_{a_0} \hookrightarrow \mathcal{K}_{a_1} \hookrightarrow \cdots \hookrightarrow \mathcal{K}_{a_n}$$

A more generalized scenario is when the inclusions are replaced with simplicial maps: $f_{ij} : \mathcal{K}_{a_i} \to \mathcal{K}_{a_j}$. In that case, we call the sequence a *tower* of simplicial complexes:

$$\{\mathcal{K}_a\}_{a\in I} : \mathcal{K}_{a_0} \xrightarrow{f_{01}} \mathcal{K}_{a_1} \xrightarrow{f_{12}} \cdots \xrightarrow{f_{(n-1)n}} \mathcal{K}_{a_n}$$

Considering the homology group of each complex in the sequence, we obtain a sequence of vector spaces connected with linear maps, which we have seen before. Specifically, we obtain the following *tower* of vector spaces:

$$H(\{\mathcal{K}_a\}_{a\in I}) : H_k(\mathcal{K}_{a_0}) \xrightarrow{f_{01*}} H_k(\mathcal{K}_{a_1}) \xrightarrow{f_{12*}} \cdots \xrightarrow{f_{(n-1)n*}} H_k(\mathcal{K}_{a_n})$$

In the above sequence each linear map f_{ij*} is the homomorphism induced by the simplicial map f_{ij}. We have already seen that persistent homology of such a sequence of vector spaces and linear maps are well defined. However, since the linear maps here are induced by simplicial maps rather than inclusions, the original persistence algorithm as described in the previous chapter does not work. A new algorithm to compute the persistence diagram of towers of simplicial complexes has been presented in [3]. Here, we generalize the notion of stability for this general case. The result is due to [2].

36 Stability of towers

In the previous chapter, we described the stability of the persistence diagrams of towers of vector spaces with respect to the perturbation of the functions whose sublevel sets generate the tower. Now we will define the stability with respect to the perturbation of the towers themselves forgetting the functions who generate them. This requires a definition of a distance between towers both at simplicial and homology levels.

It turns out that it is convenient and sometimes appropriate if the objects (simplicial complexes or vector spaces) in a tower are indexed over the positive real axis instead of a discrete subset of it. This, in turn, requires to spell out the connecting map between every pair of objects.

Definition 56 (Tower)**.** A *tower* is any collection $\mathsf{T} = \{\mathcal{T}_a\}_{a\geq 0}$ of objects $\mathcal{T}_a$ together with maps $t_{a,a'} : \mathcal{T}_a \to \mathcal{T}_{a'}$ so that $t_{a,a} = \mathrm{i}d$ and $t_{a',a''} \circ t_{a,a'} = t_{a,a''}$ for all $0 \leq a \leq a' \leq a''$. Sometimes we write $\mathsf{T} = \{\mathcal{T}_a \xrightarrow{t_{a,a'}} \mathcal{T}_{a'}\}_{0\leq a\leq a'}$ to denote the collection with the maps.

When T is a collection of *vector spaces* equipped with linear maps between them, we call it *a tower of vector spaces*. When T is a collection of *finite* simplicial complexes equipped with simplicial maps between them, we call it *a tower of simplicial complexes*. When T is a collection of *vector spaces* equipped with linear maps between them, we call it *a tower of vector spaces*.

Definition 57 (Interleaving of simplicial towers)**.** Let $\mathsf{K} = \{\mathcal{K}_a \xrightarrow{f_{a,b}} \mathcal{K}_b\}_{a\leq b}$ and $\mathsf{L} = \{\mathcal{L}_a \xrightarrow{g_{a,b}} \mathcal{L}_b\}_{a\leq b}$ be two towers of simplicial complexes. For any real $\varepsilon \geq 0$, we say that they are ε-interleaved if for each $a \geq 0$ one can find simplicial maps $\varphi_a : \mathcal{K}_a \to \mathcal{L}_{a+\varepsilon}$ and $\psi_a : \mathcal{L}_a \to \mathcal{K}_{a+\varepsilon}$ so that:

(i) for all $a \geq 0$, $\psi_{a+\varepsilon} \circ \varphi_a$ and $f_{a,a+2\varepsilon}$ are contiguous,

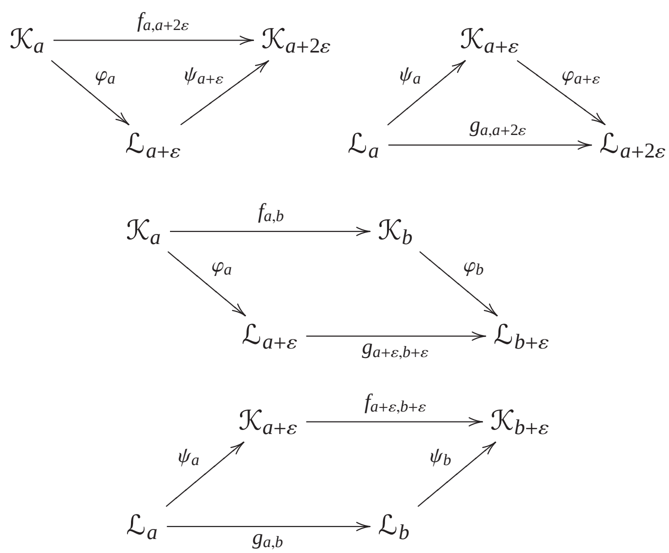
(ii) for all $a \geq 0$, $\varphi_{a+\varepsilon} \circ \psi_a$ and $g_{a,a+2\varepsilon}$ are contiguous.

(iii) for all $b \geq a \geq 0$, $\varphi_b \circ f_{a,b}$ and $g_{a+\varepsilon,b+\varepsilon} \circ \varphi_a$ are contiguous,

(iv) for all $b \geq a \geq 0$, $f_{a+\varepsilon,b+\varepsilon} \circ \psi_a$ and $\psi_b \circ g_{a,b}$ are contiguous.

If no such finite ε exists, we say the two towers are ∞-interleaved.

These four conditions are summarized by requiring that the four diagrams below commute up to contiguity:



(12)

Definition 58 (Interleaving distance of towers of simplicial complexes)**.** The interleaving distance between two towers of simplicial complexes K and L is:

$$d_I(\mathsf{K},\mathsf{L}) = \inf_{\varepsilon}\{\mathsf{K} \text{ and } \mathsf{L} \text{ are } \varepsilon\textit{-interleaved}\}.$$

Similar to the simplicial towers, we can define interleaving of towers of vector spaces. But, in that case, we replace contiguity with equality in conditions (i) through (iv).

Definition 59 (Interleaving of towers of vector spaces)**.** Let $\mathbb{U} = \{\mathrm{U}_a \xrightarrow{f_{a,b}} \mathrm{U}_b\}_{a\leq b}$ and $\mathsf{V} = \{\mathrm{V}_a \xrightarrow{g_{a,b}} \mathrm{V}_b\}_{a\leq b}$ be two towers of vector spaces. For any real $\varepsilon \geq 0$, we say that they are ε-interleaved if for each $a \geq 0$ one can find linear maps $\varphi_a : \mathrm{U}_a \to \mathrm{V}_{a+\varepsilon}$ and $\psi_a : \mathrm{V}_a \to \mathrm{U}_{a+\varepsilon}$ so that:

(i) for all $a \geq 0$, $\psi_{a+\varepsilon} \circ \varphi_a = f_{a,a+2\varepsilon}$,

(ii) for all $a \geq 0$, $\varphi_{a+\varepsilon} \circ \psi_a = g_{a,a+2\varepsilon}$.

(iii) for all $b \geq a \geq 0$, $\varphi_b \circ f_{a,b} = g_{a+\varepsilon,b+\varepsilon} \circ \varphi_a$,

(iv) for all $b \geq a \geq 0$, $f_{a+\varepsilon,b+\varepsilon} \circ \psi_a = \psi_b \circ g_{a,b}$.

If no such finite ε exists, we say the two towers are ∞-interleaved.

Definition 60 (Interleaving distance of towers of vector spaces)**.** The interleaving distance between two towers of vector spaces $\mathbb{U}$ and $\mathbb{V}$ is:

$$d_I(\mathbb{U}, \mathbb{V}) = \inf_{\varepsilon}\{\mathbb{U} \text{ and } \mathbb{V} \text{ are } \varepsilon\textit{-interleaved}\}.$$

Suppose that we have two simplicial towers $\mathsf{K} = \{\mathcal{K}_a \xrightarrow{f_{a,b}} \mathcal{K}_b)\}$ and $\mathsf{L} = \{\mathcal{L}_a \xrightarrow{g_{a,b}} \mathcal{L}_b)\}$. Consider the two towers of vector spaces obtained by taking the homology groups of the complexes, that is,

$$\mathbb{U} = \{\mathsf{H}_k(\mathcal{K}_a) \xrightarrow{f_{(a,b)*}} \mathsf{H}_(\mathcal{K}_b)\} \text{ and } \mathbb{V} = \{\mathsf{H}_k(\mathcal{L}_a) \xrightarrow{g_{(a,b)*}} \mathsf{H}_k(\mathcal{L}_b)\}.$$

The following should be obvious because simplicial maps become linear maps and contiguous maps become equal at the homology level.

Proposition 56. $d_I(\mathsf{K}, \mathsf{L}) = d_I(\mathbb{U}, \mathbb{V})$.

Let $\mathrm{Dgm}\,\mathbb{U}$ denote the persistence diagram of the tower $\mathbb{U}$ of vector spaces. Recall that d_b denotes the bottleneck distance between persistence diagrams.

Theorem 57. $d_b(\mathrm{Dgm}\,(\mathbb{U}), \mathrm{Dgm}\,(\mathbb{V})) \leq d_I(\mathbb{U}, \mathbb{V})$.

Combining Proposition 56 and Theorem 57, we obtain the following result.

Theorem 58. *Let K and L be two simplicial towers and $\mathbb{U}$ and $\mathbb{V}$ be their homology towers respectively. Then,* $d_b(\mathrm{Dgm}\,(\mathbb{U}), \mathrm{Dgm}\,(\mathbb{V})) \leq d_I(\mathsf{K}, \mathsf{L})$.

37 Examples

We show two examples where we can use the stability result in Theorem 58. Let $P \subseteq M$ be a finite subset of a metric space (M, d). Consider the Rips and Čech-filtrations:

$$\mathsf{R} : \{\mathcal{R}^{\varepsilon}(P) \hookrightarrow \mathcal{R}^{\varepsilon'}(P)\}_{0 \leq \varepsilon \leq \varepsilon'} \text{ and } \mathsf{C} : \{\mathcal{C}^{\varepsilon}(P) \hookrightarrow \mathcal{C}^{\varepsilon'}(P)\}_{0 \leq \varepsilon \leq \varepsilon'}.$$

We know that the following inclusions hold.

$$\mathcal{C}^{\alpha}(P) \subseteq \mathcal{R}^{\alpha}(P) \subseteq \mathcal{C}^{2\alpha}(P).$$

Then, one can verify that Čech- and Rips-filtrations interleave for any $\varepsilon \geq 0$, but in the log-scale. The log-scale enters into the picture to convert a multiplicative interleaving into an additive interleaving...to be continued.

References

[1] F. Chazal and S. Oudot. Towards persistence-based reconstruction in Euclidean spaces. *Proc. 24th Ann. Sympos. Comput. Geom.* (2008), 232–241.

[2] F. Chazal, D. Cohen-Steiner, M. Glisse, L. Guibas, and S. Oudot. Proximity of persistence modules and their diagrams. *Proc. 25th Annu. Sympos. Comput. Geom., SoCG* (2009), 237–246.

[3] T. K. Dey, F. Fan, and Y. Wang. Computing topological persistence for simplicial maps. *Proc. 13th Annu. Sympos. Comput. Geom., SoCG* (2014), 345–354.

[4] P. Niyogi, S. Smale, and S. Weinberger. Finding the homology of submanifolds with high confidence from random samples. *Discrete Comput. Geom.* **39** (2008), 419–441.