A Topological Regularizer for Classifiers via Persistent Homology – Supplemental Material –

Chao Chen[1] **Xiuyan Ni**[2] **Qinxun Bai**[3] **Yusu Wang**[4]

[1]Stony Brook University, Stony Brook, NY [2]City University of New York, New York, NY

[3]Hikvision Research America, Santa Clara, CA [4]Ohio State University, Columbus, OH

1 Background: Persistent Homology

Persistent homology (Edelsbrunner et al., 2002; Zomorodian and Carlsson, 2005; Carlsson and de Silva, 2010; Carlsson et al., 2009) is a fundamental recent development in the field of computational topology, underlying many topological data analysis methods. Below, we provide an an intuitive description to help explain its role in measuring the robustness of topological features in the zero-th level set (the separation boundary) of classifier function f.

Suppose we are given a space Y and a continuous function $f : Y \to \mathbb{R}$ defined on it. To characterize f and Y, imagine we now sweep the domain Y in increasing f values. This gives rise to the following growing sequence of sublevel sets:

$$Y_{\leq t_1} \subseteq Y_{\leq t_2} \subseteq \cdots \subseteq Y_{\leq t_m}, \text{with} \;\; t_1 < t_2 < \cdots t_m,$$

where $Y_{\leq t} := \{x \in Y \mid f(x) \leq t\}$ is the *sublevel set of f at t*. We call it the *sublevel set filtration of Y w.r.t. f*, which intuitively inspects Y from the point of view of function f. During the sweeping process, sometimes, new topological features (homology classes), say, a new component or a handle, will be created. Sometimes an existing one will be killed, say a component either disappear or merged into another one, or a void is filled. It turns out that these changes will only happen when we sweep through a critical points of the function f. The persistent homology tracks these topological changes, and pair up critical points into a collection of *persistence pairings* $\Pi(f) = \{(p_b, q_d)\}$. Each pair (p_b, q_d) are the critical points where certain topological feature is created and killed. Their function values $f(p_b)$ and $f(q_d)$ are referred to as the *birth time* and *death time* of this feature. The corresponding collection of pairs of (birth-time, death-time) is called the *persistence diagram*, formally, $\mathrm{dgm}(f) = \{(f(b), f(d)) \mid (p_b, q_d) \in \Pi\}$. We use Π_k and dgm_k to denote those pairings corresponding to k-dimensional topological features. For each persistent pairing (p_b, q_d), its *persistence* is defined to be $|f(p_d) - f(q_b)|$, which measures the life-time (and thus importance) of the corresponding topological feature w.r.t. f. A simple 1D example is given in Figure 1.

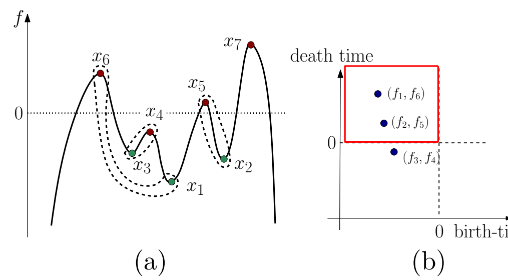
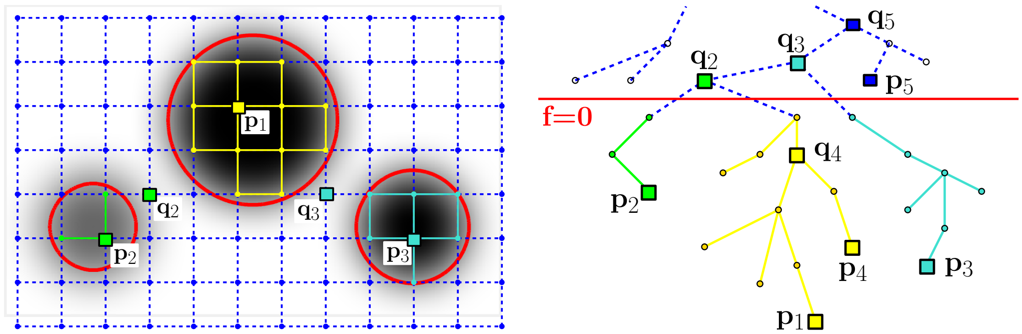


Figure 1: (a) A function $f : \mathbb{R} \to \mathbb{R}$. Its persistence pairings (of critical points) are marked by the dotted curves: $\Pi = \Pi_0 = \{(x_1, x_6), (x_2, x_5), (x_3, x_4), \ldots\}$. The corresponding persistence diagram is shown in (b), with $\mathrm{dgm}(f) = \mathrm{dgm}_0(f) = \{(f_1, f_6), (f_2, f_5), (f_3, f_4), \ldots\}$, where $f_i = f(x_i)$ for each $i \in [1, 6]$. For example, as we sweep pass minimum x_3, a new component is created in the sub-level set. This component is merged to an older component (created at x_1) when we sweep past critical point (maximum) x_4. This gives rise to a persistence pairing (x_3, x_4) corresponding to the point (f_3, f_4) in the persistence diagram.

The above description is the standard persistence *induced by the sub-level set filtratoin of f* originally introduced in Edelsbrunner et al. (2002); we refer to this as *ordinary persistence* in what follows. To capture the topological features in the *levelsets* (instead of sublevel sets) of all different threshold values, we use an extension of the aforementioned sublevel set persistence, called the *levelset zigzag persistence* (Carlsson and de Silva, 2010; Carlsson et al., 2009). Intuitively, we sweep the domain Y in increasing function values and now track topological features of the levelsets, instead of the sublevel sets. The resulting set of persistence pairings $\Pi^Z(f)$ and persistence diagram $\mathrm{dgm}^Z(f)$ have analogous meanings: each pair of criti-



(a) (b)

Figure 2: Illustration of the merging algorithm to compute 0th persistent pairing. (a) Grid graph on a given function f. Red curves are the classification boundary (zero-valued level set). (b) The corresponding merging tree. the green tree is created at its minimum p_2 and is merged to the yellow tree at the saddle q_2. The turquoise subtree is created at its minimum p_3 and is merged to the yellow tree at saddle q_3. Their corresponding pairings are (p_2, q_2) and (p_3, q_3), respectively. Persistence pairings like (p_4, q_4) and (p_5, q_5) are in $\Pi_f \subseteq \Pi_0$, but not in Π_{S_f}, as they are either created after 0 or killed before 0.

cal points $(b, d) \in \Pi^Z(f)$ corresponds to the creation and killing of some homological feature (e.g, connected components if we look at 0-th dimensional homological features) in the level set, and the corresponding pair $(f(b), f(d)) \in \text{dgm}^Z(f)$ are the birth / death times of this feature. Again, we use dgm_0^Z and Π_0^Z to represent 0-th dimensional (levelset zigzag) persistence diagram, and its corresponding persistence pairings.

In Figure 2, we give a simple illustration of the merge tree computed via the algorithm described in Section 2 of the main paper.

2 Sketch of proof for Theorem 2.1

For a classifier function $f : \mathcal{X} \to \mathbb{R}$, given its 0th levelset zigzag persistence diagram $\text{dgm}_0^Z(f)$ and its corresponding set of persitsence pairings $\Pi_0^Z(f)$ w.r.t. f, we collect $\Pi(S_f) := \{(p, q) \in \Pi_0^Z(f) \mid f(p) \leq 0, f(q) \geq 0\}$ as defined in the main text. Intuitively, each $(p, q) \in \Pi$ corresponds to a 0-D homological feature (connected component) that first appeared in a level set $f^{-1}(a)$ with $a = f(p) \leq 0$ below the zero level set $f^{-1}(0)$, and it persists through the zero level set and dies only in the level set w.r.t. value $f(q) \geq 0$. Thus intuitively, this set of persistence pairings Π maps to the set of connected components in the separation boundary S_f *bijectively* as claimed in Theorem 2.1.

Indeed, this follows from the decomposition of zigzag persistence module as introduced and studied in Carlsson and de Silva (2010). We briefly sketch the reasoning: Consider the 0th-dimensional levelset zigzag persistence module induced by the following zigzag sequence:

$$\mathcal{X}_{t_1} \subseteq \mathcal{X}_{[t_1,t_2]} \supseteq X_{t_2} \subseteq X_{[t_2,t_3]} \supseteq X_3 \subseteq \cdots$$
$$\cdots \supseteq X_{t_{m-1}} \subseteq X_{[t_{m-1},t_m]} \supseteq X_{t_m},$$

where (i) $X_t := f^{-1}(t)$ is the levelset at t, and $X_{[t,t']} := f^{-1}([t, t']) = \{x \in \mathcal{X} \mid t \leq f(x) \leq t'\}$ is the inerval levelest w.r.t. $[t, t']$; and (ii) $t_1 < t_2 < \cdots < t_m$ is a set of function values of f containing all critical values, and where each interval $[t_i, t_{i+1}]$ can contain at most one critical value. (Recall that a critical value is a function value that some critical point of f takes.)

For simplicity, denoting X_{t_i} by X_i and $X_{[t_i,t_{i+1}]}$ by X_i^{i+1}. Applying homology functor (using $\mathbb{Z}_2$ coefficient ring) to the above sequence, we have the folling levelset zigzag persistence module,

$$\mathbb{P} : H_0(X_1) \to H_0(X_1^2) \leftarrow H_0(X_2) \to H_0(X_2^3) \leftarrow$$
$$\cdots H_0(X_{m-1}^m) \leftarrow H_0(X_m).$$

It is known (see Carlsson and de Silva (2010)) that this zigzag sequence of vector spaces connected by linear maps can be written uniquely (up to isomorphism) as the direct sum of a set of indecomosible interval modules

$$\mathbb{P} \approx \mathbb{I}(b_1, d_1) \oplus \cdots \oplus \mathbb{I}(b_\ell, d_\ell),$$

where each interval module $\mathbb{I}(b_i, d_i)$ can be thought of as an interval $[b_i, d_i] \subseteq \mathbb{R}$ (or, equivalently, a point $(b_i, d_i) \in \mathbb{R}^2$). The collection of $\{(b_i, d_i)\}$ derived from the above decomposition gives rise to the 0-th levelset zigzag persistence diagram dgm_0^Z (or, using the language of Carlsson and de Silva (2010), the persistence barcode). Furthermore, for any i, restricting the decomposition to X_i, we have that $H_0(X_i)$ is isomorphic to the direct sum of all interval modules $\mathbb{I}(b_i, d_i)$ having non-trivial restriction to $H_0(X_i)$. This implies that

$$rank(H_0(X_i)) = |\{(b, d) \in \text{dgm}_0^Z \mid b \leq t_i, d \geq t_i\}|.$$

Choosing i such that $t_i = 0$ (i.e, $X_i = S_f = f^{-1}(0)$, we thus have that $rank(H_0(S_f)) = |\Pi(S_f)|$, which establishes the bijection as claimed in Theorem 2.1.

In general, the level set zigzag persistence takes $O(n^3)$ time to compute Carlsson et al. (2009), where n is the total complexity of the discretized representation of the domain $\mathcal{X}$. However, first, we only need the 0th dimensional levelset zigzag persistence. Furthermore, our domain $\mathcal{X}$ is a hypercube (thus simply connected). Using Theorem 2 of Bendich et al. (2013), and combined with the EP Symmetry Corollary of Carlsson et al. (2009), one can then show the following:

Let $\widehat{\mathrm{dgm}}(f)$ and $\widehat{\mathrm{dgm}}(-f)$ denote the ordinary 0-dimensional persistence diagrams w.r.t. the sublevel set filtrations of f and of $-f$, respectively. Let $\widehat{\Pi}_f$ and $\widehat{\Pi}_{-f}$ denote their corresponding set of persistence pairings. Set $\widehat{\Pi}_f(S_f) := \{(p,q) \in \widehat{\Pi}_f \mid f(p) \leq 0, f(q) \geq 0\}$ and $\widehat{\Pi}_{-f}(S_f) := \{(p,q) \in \widehat{\Pi}_{-f} \mid -f(p) \leq 0, -f(q) \geq 0\}$. (For example, in Figure 1(b), points in the red box correspond to $\widehat{\Pi}_f(S_f)$. Given a persistence pair (p,q), we say that the range of this pair covers 0 if $f(p) \leq 0 \leq f(q)$. Then by Theorem 2 of Bendich et al. (2013), and combined with the EP Symmetry Corollary of Carlsson et al. (2009), we have that

$$\Pi(S_f) = \widehat{\Pi}_f(S_f) \cup \widehat{\Pi}_{-f}(S_f) \cup \widehat{\Pi}_E,$$

where $\widehat{\Pi}_E$ consists certain persistence pairs whose range covers 0 from the 0-th and 1-st dimensional *extended subdigrams* induced by the extended persistent homology Cohen-Steiner et al. (2009). However, since $\mathcal{X}$ is simply connected, $H_1(\mathcal{X})$ is trivial. Hence there is no point in the 1-st extended subdigram. As $\mathcal{X}$ is connected, there is only one point $\{(v_1, v_n)\}$ in the 0-th extended subdigram, where v_1 and v_n are the global minimum and global maximum of the function f, respectively. (If the domain $\mathcal{X}$ has multiple connected component, then we can apply the formula below to each component separately.) It then follows that

$$\Pi(S_f) = \widehat{\Pi}_f(S_f) \cup \widehat{\Pi}_{-f}(S_f) \cup \{(v_1, v_n)\}. \qquad (2.1)$$

Hence one can compute $\Pi(S_f)$ by computing the 0-th ordinary persistence homology induced by the sublevel set filtration of f, and of $-f$, respectively. This finishes the proof of Theorem 2.1.

Remark 1. *Finally, we can naturally extend the above definition by considering persistent pairs and the diagram corresponding to the birth and death of high dimensional topological features, e.g., handles, voids, etc.*

References

Bendich, P., Edelsbrunner, H., Morozov, D., Patel, A., et al. (2013). Homology and robustness of level and interlevel sets. *Homology, Homotopy and Applications*, 15(1):51–72.

Carlsson, G. and de Silva, V. (2010). Zigzag persistence. *Foundations of Computational Mathematics*, 10(4):367–405.

Carlsson, G., de Silva, V., and Morozov, D. (2009). Zigzag persistent homology and real-valued functions. In *Proc. 25th Annu. ACM Sympos. Comput. Geom.*, pages 247–256.

Cohen-Steiner, D., Edelsbrunner, H., and Harer, J. (2009). Extending persistence using Poincaré and Lefschetz duality. *Foundations of Computational Mathematics*, 9(1):79–103.

Edelsbrunner, H., Letscher, D., and Zomorodian, A. (2002). Topological persistence and simplification. *Discrete Comput. Geom.*, 28:511–533.

Zomorodian, A. and Carlsson, G. (2005). Computing persistent homology. *Discrete & Computational Geometry*, 33(2):249–274.