Homology, Homotopy and Applications, vol. 24(2), 2022, pp.315–326

PERSISTENT HOMOLOGY WITH NON-CONTRACTIBLE PREIMAGES

KONSTANTIN MISCHAIKOW AND CHARLES WEIBEL

(*communicated by John R. Klein*)

Abstract

For a fixed N, we analyze the space of all sequences $z = (z_1, \ldots, z_N)$, approximating a continuous function on the circle, with a given persistence diagram P, and show that the typical components of this space are homotopy equivalent to S^1. We also consider the space of functions on Y-shaped (resp., star-shaped) trees with a 2-point persistence diagram, and show that this space is homotopy equivalent to S^1 (resp., to a bouquet of circles).

Introduction

Topological Data Analysis (TDA) is a rapidly developing body of techniques for the analysis of high dimensional data associated with nonlinear structures. Persistent homology has become one of the primary tools in TDA, for reasons including efficiency of computation, robustness with respect to perturbations in the data, and dramatic data compression. The focus of this paper is on understanding the loss of information due to this compression.

To the best of our knowledge, all applications of persistent homology to experimental data can be characterized as follows. There exists a finite simplicial complex K and a function $\varphi\colon K \to \mathbb{R}$ such that each sublevel set $K_r = \varphi^{-1}((-\infty, r])$ defines a subcomplex of K. With this input, the persistent homology algorithm outputs persistence diagrams $P = \{P_i\}$, one for each dimension of homology; we will view them as point clouds in $\mathbb{R}^2$.

With regard to compression, an obvious question is: given a fixed complex K and persistence diagrams $\{P_i\}$, what is the "preimage" space of functions φ that produce these diagrams? A detailed exposition of the polyhedral structure of these preimages is given in [**LT**]. While this geometric structure is clearly valuable, it does not necessarily translate into an understanding of what information is lost due to compression. In the smooth setting, e.g., K is a manifold, a natural corresponding

Mischaikow was partially supported by NSF grants 1521771, 1622401, 1839294, 1841324, 1934924, by NIH-1R01GM126555-01 as part of the Joint DMS/NIGMS Initiative to Support Research at the Interface of the Biological and Mathematical Science, and a grant from the Simons Foundation. Weibel was supported by NSF grants, and the Simonyi Endowment at IAS.

Received May 17, 2021, revised November 28, 2021; published on September 14, 2022.

2020 Mathematics Subject Classification: 55N31, 18F60, 55P15.

Key words and phrases: persistent homology, persistence diagram, homotopy, poset.

Article available at `http://dx.doi.org/10.4310/HHA.2022.v24.n2.a16`

question is to understand the structure of the space of Morse functions that are associated with a persistence diagram. See [**CCF+**] in the context of the two sphere.

We will consider the slightly more restrictive setting in which $\varphi\colon K \to \mathbb{R}$ is completely determined at σ by its values on the set of adjacent vertices of K by the formula $\varphi(\sigma) = \max_{v \prec \sigma} \varphi(v)$. If K has N vertices, then φ can be characterized by the vector $z = (z_1, \ldots, z_N) \in \mathbb{R}^N$. Therefore, for fixed K we can view persistent homology as inducing a function

$$PH\colon \mathbb{R}^N \to \mathbf{Per},$$

where **Per** denotes the space of persistence diagrams. Given $P \in \mathbf{Per}$, our long term goal is to understand the preimage $\mathbf{data}_P = PH^{-1}(P)$ as a subset of $\mathbb{R}^N$.

In [**CMW**], we considered the case where K is a simplicial complex representing an interval, i.e., K consists of N vertices, $N-1$ edges, and each vertex is the boundary of at most two edges. The primary result is that each component of $\mathbf{data}_P$ is contractible (for the number of components see [**C**]). The work in [**CMW**] was motivated in part to apply topological fixed point theorems to nonlinear dynamics tracked via persistence diagrams. In this setting, contractibility is a sufficiently strong condition, and thus, the collapse of geometry does not result in a loss of information. We hasten to remark that, even in this simple setting, the geometry of $\mathbf{data}_P$ is nontrivial (see [**LT**]). However, the machinery of [**CMW**] that determines the topology of the preimage is fairly simple and thus worth pursuing in and of itself.

We consider two families of examples in this paper. The first is where K is an N-gon, representing a circle, and the second is where K is a simple star-like tree, with at least three branches. In the first case, the typical components of $\mathbf{data}_P$ are homotopic to a circle, reflecting the topology of K; see Corollary 6.4. The reader might find it surprising that in the second case, when K is a tree with only 3 branches, each of length at least two, then the preimage $\mathbf{data}_P$ is homotopic to a circle; adding more branches yields a preimage which is homotopic to a bouquet of many circles (see Sections 7–8).

Though based on the same machinery, the details of the proofs of these two examples are largely independent. As a result it is easy to see that for an arbitrary graph K the topology of $\mathbf{data}_P$ can be quite complicated. In fact, we limit ourselves to presenting these two examples precisely because we do not have a general description of the homotopy type of $\mathbf{data}_P$ based on K and P.

This paper is organized into two parts. The first part of the paper considers the circle, modelled by the N-gon. In Sections 2 and 3, we discuss critical value sequences and the poset **Str** of cellular strings. In Section 4, we show that the connected components of a typical component are homeomorphic to the geometric realization of **Str**, and analyze the homotopy type of various sub-posets of **Str** in Section 5, via a series of simple moves. Finally, we prove the main result in Section 6.

The second family of examples is studied in Section 7 (3 branches) and Section 8 (many branches). The nontrivial topology in the preimage $\mathbf{data}_P$ arises from the fact that there is a larger family of moves.

1. Preliminaries

As indicated in the introduction, we consider a 1-dimensional complex K and fix a labelling of the vertices by $i = 1, \ldots, N$. We consider functions $\varphi\colon K \to \mathbb{R}$ determined by the values of φ on the vertices, i.e., $\varphi(\sigma) = \max_{i \prec \sigma} \varphi(i)$. This allows us to represent φ as a vector $z = (z_1, \dots, z_N) \in \mathbb{R}^N$, where $z_i = \varphi(i)$.

Each vector $z \in \mathbb{R}^N$ defines a sublevel-set filtration of K as follows. For $r \in \mathbb{R}$, let $K_r(z)$ denote the subcomplex of K whose vertices i have $z_i \leqslant r$ and whose edges $[i, j]$ satisfy $\max\{z_i, z_j\} \leqslant r$. As r varies, the homology H_n of the $K_r(z)$ determines a persistence diagram $\mathrm{PH}_n(z)$.

We focus on PH_0 in this paper, because we restrict our attention to complexes that are either an N-gon or a tree, so that each $K_r(z)$ is a subset of either an N-gon or a tree. In these cases $\mathrm{PH}_n(z)$ is empty for $n > 1$, and $\mathrm{PH}_1(z)$ consists of either the single point $(\max\{z_i\}, \infty)$ or the empty set, respectively. We write M for the number of points in $PH_0(z)$.

A persistence diagram P is considered *typical* if the coordinates of its M points are distinct, and we say that $z \in \mathbb{R}^N$ is a *typical point* if its coordinates are distinct. Clearly, typical points have a typical persistence diagram. We leave the proof of the following lemma to the reader.

Lemma 1.1. *Given a complex K of the type discussed in this paper and an associated typical persistence diagram $PH_0(z) = \{(p^b_m, p^d_m)\}_{m=1}^M$, then there is a typical point z' with $PH_0(z) = PH_0(z')$.*

Definition 1.2. If z is a typical point, we say that a coordinate z_i is a *local maximum* (resp., a *local minimum*) if $z_i \geqslant z_j$ (resp., $z_i \leqslant z_j$) for all vertices $[j]$ adjacent to $[i]$. The vector $(z_{i_1}, \dots, z_{i_L})$ in $\mathbb{R}^L$ of the local maxima and minima in z is the *critical value sequence* of the typical point $z \in \mathbb{R}^N$.

Digraphs. Given K and z, we define a (vertex-weighted) directed graph with underlying graph K as follows. If v and w are adjacent vertices in K and $z_v \geqslant z_w$, there is an edge directed from v to w. Note that if $z_v = z_w$ then there are two directed edges between v and w ($v \rightleftharpoons w$); we call this a *two-sided edge.*

For a typical z, the sources and sinks of the directed graph are local maxima and minima of z. In particular, a local minimum occurs at a source if and only if it has out-degree 0 and a local maximum occurs at a source if and only if it has in-degree 0. A vertex of out-degree 1 (and arbitrary in-degree) plays no role in the persistence diagram; this applies in particular to local maxima occurring at leaves of the graph.

For any typical z, there is a unique vertex $i_{\min}$ for which $z_{i_{\min}}$ is a minimum; the corresponding point in the persistence diagram $PH_0(z)$ is $(z_{i_{\min}}, \infty)$. All other persistence points are finite, and each persistence point occurs exactly once.

Lemma 1.3. *Suppose that z is a typical point, and out-degree$(i) \neq 1$. Then there is a persistence point with coordinate z_i.*

Proof. If i has out-degree 0, then z_i is a local minima, and hence there is a persistence point (p^b, p^d) with $p^b = z_i$. If i has out-degree $\geqslant 2$ then at least 2 components of K_r will merge at $r = z_i$, and thus there is a persistence point (p^b, p^d) with $p^d = z_i$. □

Remark 1.4. If z_i is not a typical point then, after identifying vertices connected by 2-sided edges, a similar argument applies. We leave the verification to the reader.

We say that z_i is a *critical* coordinate if the number of out-edges (after the potential identification) from i is not exactly one.

2. Persistence diagrams on an N-gon

Throughout this section K is an N-gon, the vertices are identified with elements of $\mathbb{Z}/N$ (with $0 = N$) and there are N edges, $[i, i+1]$, $0 \leqslant i < N$. Our main result (Corollary 6.4) concerns the preimage $\mathbf{data}_P$ in $\mathbb{R}^N$ of a typical persistence diagram: when $N/2$ is greater than the number M of persistence points of P, the connected components of $\mathbf{data}_P$ are homotopy equivalent to S^1.

The following result is analogous to [**CMW**, Lemma 2.4].

Lemma 2.1. *Let $z \in \mathbb{R}^N$ be a typical point with persistence diagram*

$$\mathrm{PH}_0(z) = \{(p_0^b, \infty)\} \cup \{(p_m^b, p_m^d)\}_{m=1}^{M-1} \quad \textit{and} \quad \mathrm{PH}_1(z) = \{p_M^b, \infty\}.$$

Then z has $2M$ local extrema; each p_m^d is a local maximum, each p_m^b is a local minimum, p_0^b is a global minimum, and p_M^b is a global maximum.

Since there are only finitely many critical value sequences with these local maxima and minima, only finitely many critical value sequences arise for each persistence diagram.

Remark 2.2. Conversely, if $PH_0(z) = \{(p_m^b, p_m^d)\}_{m=1}^{M}$ is a persistence diagram with distinct values, there is a typical point z' with $PH_0(z) = PH_0(z')$.

When $N = 2M$, every z_i is a local extremum, so $\mathbf{data}_P$ is a finite set. We shall assume that $N > 2M$ for the rest of the paper.

Example 2.3. When $N = 2M + 1$, every point $(z_1, \dots, z_N)$ has exactly one non-extremal value, every component of $\mathbf{data}_P$ is homeomorphic to a circle, and the components are indexed by the critical value sequences modulo cyclic rotations.

To see this, fix a critical value sequence $v = (a, b, \dots, s, t)$ using the $2M$ values p_m^b, p_m^d in Lemma 2.1; there are only N places to insert a non-extremal element. Let $C_i(v)$, $i = 0, \dots, N-1$ denote the subspace of points in $\mathbf{data}_P$ with critical value sequence v, where z_i is the non-extremal value. Then $C_i(v)$ is an open interval, whose closure $\bar{C}_i(v)$ meets $\bar{C}_{i-1}(v)$ and $\bar{C}_{i+1}(v)$ in an endpoint when $i \neq 0, N-1$. Writing Rv for the cyclic rotation $(t, a, b, \dots, s)$ of v, the closure of $C_0(v)$ meets $\bar{C}_1(v)$ and $\bar{C}_{N-1}(Rv)$, as z_0 is between $z_1 = a$ and $z_{N-1} = d$. Similarly, the closure of $C_{N-1}(v)$ meets $\bar{C}_{N-1}(v)$ and $\bar{C}_0(R^{-1}v)$.

To count the number of components, note that the cyclic order of the local extrema cannot be changed without changing the persistence diagram, all points in $C(z)$ must have the same critical value sequence as z, up to rotation.

3. The poset Str of cellular circular strings

Throughout this section K is an N-gon. A *circular symbol string* is a string of symbols $s = s_1 \cdots s_N$, where each symbol is either $0, 1$ or X. We refer to the symbols 0 and 1 as *bits*. (Cf. [**CMW**].) Any circular symbol string has a canonical representation as the concatenation $s = \gamma_1 \cdots \gamma_J$ of *blocks* γ_i, each block consisting of the same symbol, such that adjacent blocks have different symbols. Because of our wrap-around convention, it is possible that the last block has the same symbol as the first.

Definition 3.1. A circular symbol string $s = s_1 \cdots s_N$ is a *circular cellular string* of *rank* $M > 0$ if for the canonical representation $s = \gamma_1 \cdots \gamma_J$:

(i) exactly M blocks have symbol 0, and
(ii) if γ_j consists of the symbol X and $j \neq 1, J$, then the symbols of γ_{j-1} and γ_{j+1} are different.

The *dimension* of a cellular string s, $\dim(s)$, is the number of symbols X in s; it is at most $N - 2M$.

Fix N and $M < N/2$. The set $\mathbf{Str} = \mathbf{Str}(N, M)$ of circular cellular strings of length N and rank M is a poset, where $s' < s$ if the string s is obtained from s' by replacing some of the bits 0 and 1 in s' by X. For example, in $\mathbf{Str}(3, 1)$ we have

$$X01 > 001 < 0X1 > 011 < 01X.$$

Proposition 3.2. *The maximal elements of* $\mathbf{Str}(N, M)$ *are the strings of dimension* $N - 2M$.

Proof. (Cf. [**CMW**, Prop. 2.8].) If s has smaller dimension, then there is a block of length $\geqslant 2$ of symbols 0 or 1. Replacing the first symbol in the block by X yields a symbol s' with $s < s'$, so s is not maximal. □

Lemma 3.3. *Every string* s' *in* $\mathbf{Str}$ *is the greatest lower bound of the set of maximal strings* s *with* $s' < s$.

Every maximal chain in $\mathbf{Str}$ *has length* $N - 2M$.

Proof. The proof of [**CMW**, Lemma 2.9] goes through. Briefly, we proceed by downward induction on the dimension $d = \dim(s')$. Replacing the two end symbols of a block by X yields two $(d + 1)$-dimensional strings whose greatest lower bound is s'. □

4. The polytopes for the N-gon

Throughout this section we work over an N-gon. Fix a critical value sequence $(z_{n_1}, \ldots, z_{n_{2M}})$ of a typical point z. To each circular cellular string s, represented in block form as $\gamma_1 \cdots \gamma_J$, we associate simplices $T(\gamma_j)$ and their product, the polytope $T(s) = \prod T(\gamma_j)$ as in [**CMW**]:

- if γ_j is the k^{th} block involving 0 or 1 we set $T(\gamma_j) = \{z_k\}^{n_k}$;
- if γ_j involves X and γ_{j-1} is the k^{th} block involving 0 or 1, we define $T(\gamma_j)$ to be the simplex of all monotone sequences $(x_1, \ldots, x_{n_j})$ of length n_j between z_k and z_{k+1}.

Fix a persistence diagram P, and a component C of $\mathbf{data}_P$. Then for any typical point z in C, it is clear from the definition 1.2 of a critical value sequence that C is the union of the simplices $T(s)$, where $s \in \mathbf{Str}(N, M)$. For this, it is convenient to work with the poset of circular cellular strings.

Let s be a circular cellular string with k blocks with symbols X, of lengths $n_1, \ldots, n_k$. Recall from [**CMW**, Example 2.10, Theorem 2.13] that the geometric realization of the sub-poset $\mathbf{Str}/s = \{s' : s' \leqslant s\}$ is homeomorphic to the product $\Delta^{n_1} \times \cdots \Delta^{n_k}$ of simplices, i.e., to $T(s)$.

Theorem 4.1. *If* z *is a typical point, the connected component* $C(z)$ *of* $\mathbf{data}_P$ *is homeomorphic to the geometric realization of* $\mathbf{Str}$.

Proof. The proof in [**CMW**, Theorem 2.13] goes through. The key observation is that for each $s_1, \dots, s_n$, the intersection of the realizations of the $\mathbf{Str}/s_i$ is the realization of $\mathbf{Str}/s'$, where s' is the greatest lower bound of the s_i. □

Example 4.2. In [**LT**], a similar problem is studied with a different filtration, intermediate between the polygon with N vertices and its subdivision, which has $2N$ vertices. The comparison is sketched in Section 5.2 of [**LT**].

5. Homotopy operations

Let $\mathbf{Str}_0$ (resp., $\mathbf{Str}_1$) denote the sub-poset of strings in $\mathbf{Str}$ whose initial bit is 0 (resp., 1), such that s_1 and s_N are not both 0 (resp., 1). Recall that X is not a bit. For example, $0X1X$, $X0101$ and $XXX01$ are elements of $\mathbf{Str}_0$.

Proposition 5.1. *The classifying spaces of* $\mathbf{Str}_0$ *and* $\mathbf{Str}_1$ *are contractible.*

Proof. This is the content of Proposition 3.5 and Corollary 3.6 in [**CMW**]. The sub-poset of $\mathbf{Str}$ in *loc. cit.* consisting of strings such that s_1 and s_N are not both 0 is our $\mathbf{Str}_0$, and the poset morphisms used in that proof send $\mathbf{Str}_0$ to itself. The realizations of those poset morphisms, when composed, give a homotopy from $B\mathbf{Str}_0$ to a point. The proof for $\mathbf{Str}_1$ is the same. □

For symbols a, b we write $\mathbf{Str}_{ab}$ for the sub-poset of strings in $\mathbf{Str}$ whose initial and terminal symbols are a and b, respectively; we abbreviate such a string as $a\sigma b$, where σ is a string of length $N-2$.

Let $\overline{\mathbf{Str}}_{00}$ denote the sub-poset of strings in $\mathbf{Str}$ whose initial and terminal symbols are either: both 0; 0 and X; or X and 0. Thus $\overline{\mathbf{Str}}_{00}$ contains $\mathbf{Str}_{00}$ as well as $\mathbf{Str}_{0X}$ and $\mathbf{Str}_{X0}$ and is disjoint from $\mathbf{Str}_1$. Since the initial bit for $s = X\sigma 0 \in \mathbf{Str}_{X0}$ is 1,

$$\overline{\mathbf{Str}}_{00} \cap \mathbf{Str}_0 = \mathbf{Str}_{0X}, \quad \overline{\mathbf{Str}}_{00} \cap \mathbf{Str}_1 = \mathbf{Str}_{X0}. \tag{5.2}$$

We define $\overline{\mathbf{Str}}_{11}$ similarly, by interchanging 0 and 1. Thus:

$$\overline{\mathbf{Str}}_{11} \cap \mathbf{Str}_1 = \mathbf{Str}_{1X}, \quad \overline{\mathbf{Str}}_{11} \cap \mathbf{Str}_0 = \mathbf{Str}_{X1}. \tag{5.3}$$

Lemma 5.4. $B\mathbf{Str}_{00}$ *is a deformation retract of* $B\overline{\mathbf{Str}}_{00}$*. By symmetry,* $B\mathbf{Str}_{11}$ *is a deformation retract of* $B\overline{\mathbf{Str}}_{11}$*.*

Proof. Define $R\colon \overline{\mathbf{Str}}_{00} \to \mathbf{Str}_{00}$ to be the identity on $\mathbf{Str}_{00}$, and $R(0\sigma X) = 0\sigma 0$, $R(X\sigma 0) = 0\sigma 0$. It is easy to see that R is a poset map, and that $R(s) \leqslant s$, i.e., $R \Rightarrow \mathrm{id}$ is a natural transformation. Taking the geometric realization, we see that R is a continuous map, and that R is homotopic to the identity on $B\overline{\mathbf{Str}}_{00}$. □

Definition 5.5. If s is a circular cellular string in $\mathbf{Str}_{00}$, we define $F_1(s)$ as follows (cf. [**CMW**, Def. 3.1]). If $s = 0X\sigma 0$, set $F_1(s) = s$. If not, there are two cases. Case (i): if s has no 00 or 11 preceding the leftmost X, $F_1(s)$ transposes that X with the bit immediately preceding it. Case (ii): if s has the form $\sigma_1 abb\sigma_2$, where $\sigma_1 a$ is an alternating bitstring (beginning with 0) and σ_2 is the remainder of the string, we set $F_1(s) = \sigma_1 aab\sigma_2$. Note that in case (ii), σ_2 is either empty or ends in 0.

Let $\mathbf{Str}_{00}^{(\ell)}$ denote the sub-poset of $\mathbf{Str}_{00}$ consisting in strings beginning $0X\cdots X$ ($\ell-1$ symbols X). The definition of $F_\ell\colon \mathbf{Str}_{00}^{(\ell)} \to \mathbf{Str}_{00}^{(\ell)}$ mimicks that of F_1; if $s = 0\beta\sigma$ (where β is a sequence of ($\ell-1$) symbols X) then $F_\ell(s) = 0\beta F_1(\sigma)$.

Lemma 5.6. $F_1\colon \mathbf{Str}_{00} \to \mathbf{Str}_{00}$ *is a poset morphism and* $F_1^k(\mathbf{Str}_{00}) = \mathbf{Str}_{00}^{(2)}$ *for* $k \gg 0$.

Proof. We proceed by downward induction to show that if $s' < s$ then $F_1(s') \leqslant F_1(s)$. If the initial X in s' is not preceded by a 00 or 11, the same is true for s, and the inequality is evident. Next, suppose that $s' = \sigma_1 abb \dots b\sigma_2$, where $\sigma_1 a$ is an alternating bitstring. If $s = \sigma_1 abb \dots b\sigma_2'$ with $\sigma_2 < \sigma_2'$, we also have $F_1(s') < F_1(s)$. Otherwise, either $s \geqslant s_1$ or $s \geqslant s_2$, where $s_1 = \sigma_1 aXb \cdots b\sigma_2$ and $s_2 = \sigma_1 ab \cdots X\sigma_2$. Since $F_1(s') < F_1(s_1)$ and $F_1(s') < F_1(s_2)$, the result follows by induction. □

Recall that a poset P is said to be contractible if the space BP is.

Proposition 5.7. $\mathbf{Str}_{00}$*,* $\overline{\mathbf{Str}}_{00}$*,* $\mathbf{Str}_{11}$ *and* $\overline{\mathbf{Str}}_{11}$ *are contractible.*

Proof. We give the proof for $\mathbf{Str}_{00}$; it follows by symmetry and Lemma 5.4 that $\overline{\mathbf{Str}}_{00}$, $\mathbf{Str}_{11}$ and $\overline{\mathbf{Str}}_{11}$ are also contractible.

We first show that $\mathbf{Str}_{00}^{(2)} \to \mathbf{Str}_{00}$ is a homotopy equivalence. As in [**CMW**, Proposition 3.5], we filter $\mathbf{Str}_{00}$ by sub-posets Fil_i, where $Fil_0 = \mathbf{Str}_{00}^{(2)}$ and Fil_i is the full sub-poset on the strings s with $F_1^i(s) \in \mathbf{Str}_{00}^{(2)}$. Since F_1 maps Fil_i to Fil_{i-1}, the geometric realization BF_1 of F_1 restricts to a continuous map from $BFil_i$ to $BFil_{i-1}$.

To see that $BFil_{i-1} \subseteq BFil_i$ is a homotopy equivalence, we define a poset endomorphism h on Fil_i as follows. If $s \in Fil_{i-1}$ then $h(s) = s$. Otherwise, define $h(s)$ to be the greatest lower bound of s and $F_1(s)$. Thus Bh is a retract of $BFil_i$ onto $BFil_{i-1}$. For $s \in Fil_i$, the inequalities $F_1(s) \leqslant h(s) \geqslant s$ yield natural transformations from h to F_1 and to the identity, and hence homotopies between the identity map, Bh, and BF_1. These homotopies show that $BFil_{i-1} \simeq BFil_i$. Composing these homotopies gives a homotopy equivalence between $B\mathbf{Str}_{00}$ and $BFil_0 = B\mathbf{Str}_{00}^{(2)}$.

The same argument, *mutatis mutandis*, then shows that each inclusion of the form $\mathbf{Str}_{00}^{(\ell-1)} \to \mathbf{Str}_{00}^{(\ell)}$ is a homotopy equivalence. Since $B\mathbf{Str}_{00}^{(\ell)}$ is only the point $\{0X \cdots X10 \cdots 10\}$ when $\ell = N - 2m + 1$, $B\mathbf{Str}_{00}$ is homotopy equivalent to a point, as claimed. □

Proposition 5.8. $\mathbf{Str}_{0X}$ *and* $\mathbf{Str}_{X0}$ *are contractible. By symmetry,* $\mathbf{Str}_{1X}$ *and* $\mathbf{Str}_{X1}$ *are also contractible.*

Proof. Since the posets $\mathbf{Str}_{0X}$ and $\mathbf{Str}_{X0}$ are isomorphic (by the front-to-back permutation of strings), it suffices to give the proof for $\mathbf{Str}_{0X}$. By Example 2.3, we may assume that $N > 2M + 1$. Definition 5.5 goes through word for word in this setting to yield a poset endomorphism F_1 on $\mathbf{Str}_{0X}$, with the image of F_1^K being $\mathbf{Str}_{0X}^{(2)}$ for $K \gg 0$. Now the proof of Proposition 5.7 goes through to show that $\mathbf{Str}_{0X}$ is contractible. □

6. Circular components

Let Q denote the 8-element poset on the left of diagram (6.1); the 4 corners are minimal elements, and the 4 side-vertices are maximal. The geometric realization BQ of Q has a vertex for each element of Q and an edge for each strict inequality; there are no higher simplices because the poset Q has no chains $q_0 < q_1 < q_2$. Thus BQ is an octagon, homeomorphic to a circle. The sub-posets of **Str** we have described fit into the right-hand diagram below, where the arrows indicate inclusion.

$$\begin{array}{ccccc} 0X & \longrightarrow & 0 & \longleftarrow & X1 \\ \downarrow & & & & \downarrow \\ 00 & & Q & & 11 \\ \uparrow & & & & \uparrow \\ X0 & \longrightarrow & 1 & \longleftarrow & X1 \end{array} \qquad \begin{array}{ccccc} \mathbf{Str}_{0X} & \longrightarrow & \mathbf{Str}_0 & \longleftarrow & \mathbf{Str}_{X1} \\ \downarrow & & & & \downarrow \\ \overline{\mathbf{Str}}_{00} & & & & \overline{\mathbf{Str}}_{11} \\ \uparrow & & & & \uparrow \\ \mathbf{Str}_{X0} & \longrightarrow & \mathbf{Str}_1 & \longleftarrow & \mathbf{Str}_{1X} \end{array} \tag{6.1}$$

Define $f\colon \mathbf{Str} \to Q$ by sending elements of $\mathbf{Str}_{0X}$, $\mathbf{Str}_{0X}$, $\mathbf{Str}_{1X}$ and $\mathbf{Str}_{X1}$ to the corresponding minimal vertices of Q, as indicated by (6.1); strings in $\mathbf{Str}_{00}$ and $\mathbf{Str}_{11}$ are sent to the vertices indicated by $\overline{\mathbf{Str}}_{00}$ and $\overline{\mathbf{Str}}_{11}$, respectively. The strings in $\mathbf{Str}_0$ not in $\mathbf{Str}_{0X}$ or $\mathbf{Str}_{X1}$ (resp., in $\mathbf{Str}_1$ not in $\mathbf{Str}_{X0}$ or $\mathbf{Str}_{1X}$) are sent to the other maximal vertices of Q, as indicated. It is clear that f is a poset morphism.

Recall from [**WK**, IV.3.2.3] that for $q \in Q$, the comma category f/q has objects the pairs (s, q), where $f(s) \leqslant q$, i.e., $s \in \mathbf{Str}$ such that $f(s) \leqslant q$; and there is a morphism from (s', q) to (s, q) if and only if $s' \leqslant s$ in **Str**.

Lemma 6.2. *The right-hand side of diagram* (6.1) *is the diagram of the comma categories f/q for $q \in Q$.*

Proof. For the minimal elements $q = ab$ of Q, it is a tautology that $f/q = f^{-1}(q) = \mathbf{Str}_q$. Since $\overline{\mathbf{Str}}_{00}$ is the union of $f^{-1}(00)$, $f^{-1}(0X)$ and $f^{-1}(X1)$, we see that $f/0 = \overline{\mathbf{Str}}_{00}$; by symmetry we also have $f/1 = \overline{\mathbf{Str}}_{11}$. The definition of f on $\mathbf{Str}_0$ and $\mathbf{Str}_1$ ensures that we also have $f/0 = \mathbf{Str}_0$ and $f/1 = \mathbf{Str}_1$. ☐

Theorem 6.3. *$B\mathbf{Str}$ is homotopic to the circle S^1.*

Proof. Quillen's Theorem A says that if the geometric realization of every f/q is contractible, then $Bf\colon B\mathbf{Str} \to BQ \simeq S^1$ is a homotopy equivalence (see [**WK**, IV.3.7]). By (5.2), (5.3), Propositions 5.7 and 5.8, the geometric realizations of all the f/q are contractible. ☐

Combining Lemma 1.1 with Theorems 4.1 and 6.3, we obtain:

Corollary 6.4. *If P is a typical persistence diagram, every connected component of $\mathbf{data}_P$ is homotopy equivalent to S^1.*

7. Y-shaped configurations

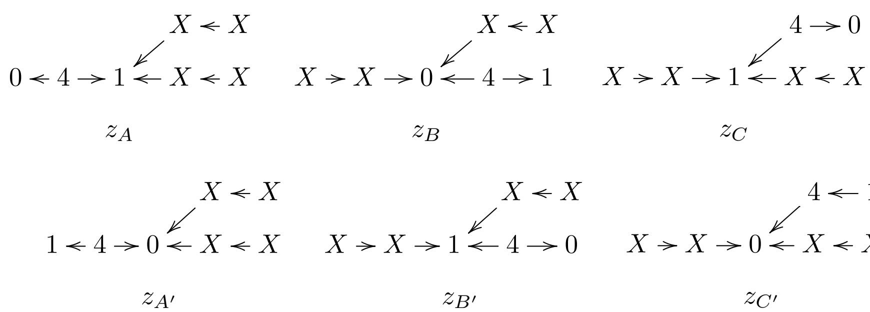
In this section we show that $\mathbf{data}_P$ can still be homotopic to a circle, even for rooted trees with three branches and persistence diagram $P = \{(0, \infty), (1, 4)\}$ for H_0. (The choice of $0 < 1 < 4$ is for concreteness.)

For simplicity, we focus on the case where the tree K has

$$\text{vertex set } V = \{i \mid i = 1, \ldots, 7\} \text{ and edges } E = \{[1, 2], [2, 3], [3, 4], [4, 5], [3, 6], [6, 7]\}.$$

That is, the central vertex 3 has degree 3, and the endpoints are vertices 1, 5, and 7. The three branches (α, β, and γ) are generated by the vertices $\{1, 2, 3\}$, $\{3, 4, 5\}$, and $\{3, 6, 7\}$, respectively.

Figure (7.1) illustrates six points $z_A, z_B, \ldots$ in $\mathbf{data}_P$ with their directed graphs $\Gamma(z_A), \Gamma(z_B), \ldots$ (as in Section 1). The critical coordinates are marked by their values 0, 1, and 4. The other z_i are marked by X; they are not critical coordinates and their exact value is unimportant.

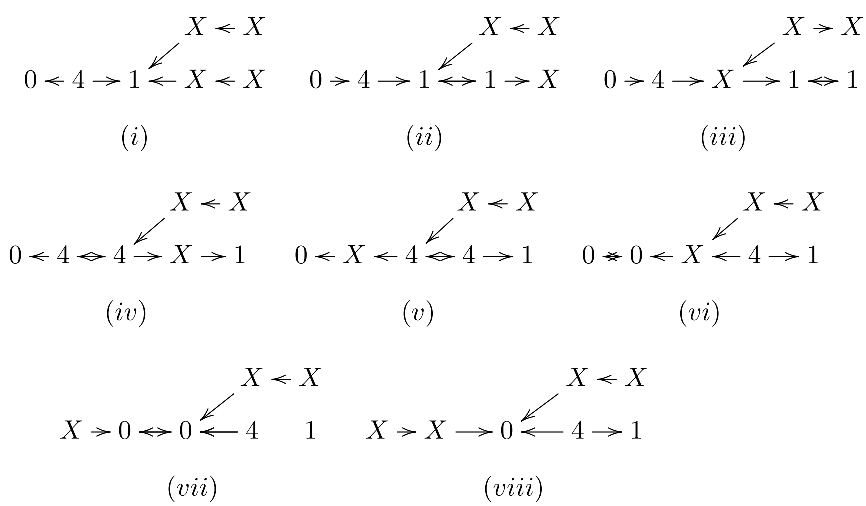


(7.1)

Construction 7.2. Consider the point $z_A \in \mathbb{R}^7$, illustrated on the left of (7.1). The extremal points 0, 4 and 1 all lie on branch α.

It is possible to continuously transform z_A to z_B by sliding the 041 two places to the right, from branch α to branch β, without changing the directed edges on branch γ. This sliding process is explained in [**CMW**], and is illustrated by steps (i) through (viii) in Figure (7.3). (Recall that double edges do not impact the persistence diagram.)

In analogy with [**CMW**] and as shown in Figure (7.3) in steps (i) through (viii), it is possible to slide the 041 two places to the right (to z_B on branch β) without changing the directed edges (recall that double edges do not impact the persistence diagram) on branch γ.



(7.3)

Beginning with z_B, we can slide 041 clockwise up (to z_C on branch γ) without changing the directed edges on branch α. We now slide 140 two places to the left (to $z_{A'}$ on branch α) without changing the directed edges on branch β.

This maneuver (which resembles a clockwise 'K-turn' in a car) results in the critical coordinates 0 and 1 switching places. Following this with a second K-turn of the form ($z_{A'}$ to $z_{B'}$ to $z_{C'}$ to z_A) returns us to the starting configuration.

Construction 7.2 shows that the six points in (7.1) lie on a non-trivial loop; it will turn out to be a generator of the fundamental group of $\mathbf{data}_P$. While obvious, it is perhaps worth emphasizing that these sequences of slides are not possible on the interval or N-gon. These new moves make it possible for the topology of $\mathbf{data}_P$ to be more interesting.

Let $\mathbf{Br}_\alpha$ be the subspace of $\mathbf{data}_P$ consisting of all z with $z_i = 0$ for some vertex i on branch α, and $z_j \neq 4$ for all vertices j on branch α further from the central vertex than vertex i. In Figure (7.1), z_A, z_B and $z_{C'}$ are in $\mathbf{Br}_\alpha$.

The subspaces $\mathbf{Br}_\beta$ and $\mathbf{Br}_\gamma$ are defined similarly.

Lemma 7.4. *$\mathbf{Br}_\alpha$, $\mathbf{Br}_\beta$ and $\mathbf{Br}_\gamma$ are contractible.*

Figure (7.5) illustrates the steps in the proof, starting from $z_{C'}$.

Proof. By symmetry, it suffices to consider $\mathbf{Br}_\alpha$. We shall use three steps to construct a deformation retraction of $\mathbf{Br}_\alpha$ to a point. By definition there exists a vertex i_0 closest to the endpoint $i_\alpha = 1$ of branch α with $z_{i_0} = 0$. For the first step, continuously decrease the value of z_j to 0 for every vertex between i_α and i_0 (and do nothing if $i_0 = i_\alpha$). This is a deformation retraction onto the subspace of all z' which are 0 at i_α.

For the second step of the homotopy, given a point with $z_{i_\alpha} = 0$, consider the set of vertices j such that the path from vertex i_α to vertex j does not contain a vertex k with $z_k = 1$. We can continuously change the values of z_j at all these vertices to 4. This is a deformation retraction to the subspace of all z'' where $z'' = 4$ at vertex 2 (the vertex adjacent to $i_\alpha = 1$).

Finally, given z'', we can continuously decrease the value of z''_j to 1 for all vertices j other than i_α and its neighbor. The result is a deformation retraction to the point $z = (0, 4, 1, \dots, 1)$, showing that $\mathbf{Br}_\alpha$ is contractible. □

$$\begin{array}{ccc} \quad\quad 4 \to 1 & \quad\quad 4 \to 1 & \quad\quad 1 \leftarrow 1 \\ \quad\swarrow & \quad\swarrow & \quad\swarrow \\ 0 \leftarrow 0 \to 0 \leftarrow X \leftarrow X & 0 \leftarrow 4 \to 4 \leftarrow 4 \leftarrow 4 & 0 \leftarrow 4 \to 1 \leftarrow 1 \leftarrow 1 \end{array} \tag{7.5}$$

We write $\mathbf{Br}_{\alpha\gamma}$ for $\mathbf{Br}_\alpha \cap \mathbf{Br}_\gamma$. It is the subspace of all z where $z = 0$ at the central vertex, while the vertex with $z = 1$ (and hence the vertex with $z = 4$) lies on branch β. For example, z_B is in $\mathbf{Br}_{\alpha\gamma}$. The subspaces $\mathbf{Br}_{\alpha\beta}$ and $\mathbf{Br}_{\beta\gamma}$ are defined similarly.

Lemma 7.6. *$\mathbf{Br}_{\alpha\gamma}$, $\mathbf{Br}_{\alpha\beta}$ and $\mathbf{Br}_{\beta\gamma}$ are contractible.*

Proof. By symmetry, it suffices to consider $\mathbf{Br}_{\alpha\gamma}$. For the 7-vertex tree, $\mathbf{Br}_{\alpha\gamma}$ consists of just the points of the form z_B, illustrated by the second diagram of (7.1). In particular, $\mathbf{Br}_{\alpha\gamma}$ is contractible. □

Remark 7.7. The proofs of Lemmas 7.4 and 7.6 go through for longer Y-shaped trees, i.e., rooted trees with a central vertex of degree 3 with 3 linear branches of length $\geqslant 2$ attached to it. (The vertices $1, 2, 3$ are at the end of branch $\mathbf{Br}_\alpha$.)

By inspection, every point in the preimage $\mathbf{data}_P$ lies in one of the subspaces $\mathbf{Br}_\alpha$, $\mathbf{Br}_\beta$, or $\mathbf{Br}_\gamma$. Since the intersection of any two branches is contractible, we see that the preimage $\mathbf{data}_P$ is path-connected.

Let Q denote the 6-element poset on the left of diagram (7.8); the elements α, β, γ are maximal and the others are minimal. Thus BQ is a hexagon, homeomorphic to the circle S^1.

Consider the topological poset $\mathbf{Br}$ of pairs (x, q) with $x \in \mathbf{Br}_q$, illustrated by the right of (7.8). It is clear that there is a poset morphism $f \colon \mathbf{Br} \to Q$ sending elements (x, q) to q.

$$\begin{array}{ccccc} \alpha & \longrightarrow & \alpha\gamma & \longleftarrow & \gamma \\ \downarrow & & & & \downarrow \\ \alpha\beta & \longleftarrow & \beta & \longrightarrow & \beta\gamma \end{array} \qquad \begin{array}{ccccc} \mathbf{Br}_\alpha & \longrightarrow & \mathbf{Br}_{\alpha\gamma} & \longleftarrow & \mathbf{Br}_\gamma \\ \downarrow & & & & \downarrow \\ \mathbf{Br}_{\alpha\beta} & \longleftarrow & \mathbf{Br}_\beta & \longrightarrow & \mathbf{Br}_{\beta\gamma} \end{array} \tag{7.8}$$

Lemma 7.9. *The geometric realization of* $\mathbf{Br}$ *is homotopy equivalent to* $\mathbf{data}_P$*.*

Proof. For each q, the realization $|\mathbf{Br}|$ contains a subspace homeomorphic to $\mathbf{Br}_q$, and for each $q' < q$ the realization contains the mapping cylinder of the inclusion $\mathbf{Br}_{q'} \subset \mathbf{Br}_q$. Thus there is a natural map from $|\mathbf{Br}|$ onto $\mathbf{data}_P$. Since $\mathbf{Br}_{q'}$ is a subspace of two subspaces $\mathbf{Br}_q$, it is easy to see that $|\mathbf{Br}| \to \mathbf{data}_P$ is a homotopy equivalence. □

Theorem 7.10. *The preimage* $\mathbf{data}_P$ *is homotopic to* S^1*.*

Proof. By Lemma 7.9, it suffices to show that f induces a homotopy equivalence $|\mathbf{Br}| \to BQ \simeq S^1$. By Quillen's Theorem A, it suffices to show that (the realization of) each comma category f/q is contractible.

The comma category f/q is the poset of all pairs $(x, q' \leqslant q)$ with $x \in \mathbf{Br}_{q'}$. If q is minimal in Q, $f/q = f^{-1}(q) = \mathbf{Br}_q$, which is contractible by Lemma 7.6. If q is maximal, we still have $f^{-1}(q) = \mathbf{Br}_q$ but f/q contains elements $(x, q' < q)$. The geometric realization of the natural transformation $\eta \colon (x, q' \leqslant q) \Rightarrow (x, q)$ is a deformation retraction from $B(f/q)$ to the subspace $B(f^{-1}q)$, which is contractible by Lemma 7.4. □

8. Star-like configurations

In this section, we generalize from Y-shaped trees to star-like trees, i.e., trees with a central vertex of degree $n \geqslant 3$ and n linear branches of length at least 2 attached to it. (The Y-shaped trees of Section 7 form the case $n = 3$.)

For $q = 1, \ldots, n$, let $\mathbf{Br}_q$ be the subspace of all z with $z_i = 0$ for some vertex on branch q, and $z_j \neq 4$ for all vertices j on branch q further from the central vertex than i. The proof of Lemma 7.4 goes through to show that $\mathbf{Br}_q$ is contractible for each $q = 1, \ldots, n$.

Set $\mathbf{Br}'_q = \bigcap_{p \neq q} \mathbf{Br}_p$; it is the subspace of all z where $z = 0$ at the central vertex, while there is a vertex with $z = 1$ (and hence a vertex with $z = 4$) lying on branch q. The proof of Lemma 7.6 goes through to show that each $\mathbf{Br}'_q$ is contractible.

Since each point of $\mathbf{data}_P$ lies on one of the branches, which are contractible, and each $\mathbf{Br}'_q$ is contractible, $\mathbf{data}_P$ is path-connected.

Theorem 8.1. $\mathbf{data}_P$ *is homotopy equivalent to a bouquet* $\bigvee S^1$ *of* $(n^2 - 3n + 1)$ *circles.*

When $n = 3$, this yields 1 circle, as in Theorem 7.10; for $n = 4$ branches, $\mathbf{data}_P$ is homotopy equivalent to a bouquet of 5 circles.

Proof. Consider the poset Q whose elements are the $2n$ branches $\mathbf{Br}_q$ and $\mathbf{Br}'_q$, with $\mathbf{Br}'_q < \mathbf{Br}_p$ for every $p \neq q$. The realization of this poset is a bipartite graph Γ such that every vertex of Γ has degree $n - 1$. Since Γ has $2n$ vertices and $n^2 - n$ edges, its Euler characteristic is

$$\chi = V - E = 3n - n^2.$$

Since Γ is connected, and $\chi = \dim H_0(\Gamma) - \dim H_1(\Gamma)$, Γ is homotopy equivalent to a bouquet of $1 - \chi = (n^2 - 3n + 1)$ circles.

Consider the topological poset $\mathbf{Br}$ of pairs (x, s), where $x \in \mathbf{Br}_s$, and pairs (x, s'), where $x \in \mathbf{Br}'_s$; there is an obvious poset morphism $f\colon \mathbf{Br} \to Q$, and hence a map $|\mathbf{Br}| \to \Gamma$. The proof of Lemma 7.9 goes through (with 'two' replaced by $n - 1$) to show that $|\mathbf{Br}|$ is homotopy equivalent to $\mathbf{data}_P$. Finally, the proof of Theorem 7.10 goes through to show that $|\mathbf{Br}| \to \Gamma$ is a homotopy equivalence. (One uses the version of Quillen's Theorem A for the realization of topological categories; see [**WK**, IV.3.9].) The homotopy equivalence of the theorem follows. □

Remark 8.2. The proof of Theorem 8.1 breaks down when $n = 2$ because $\mathbf{Br}'_p = \mathbf{Br}_q$ for $p \neq q$, so Q only has 2 elements.

References

[**CCF+**] M. Catanzaro, J. Curry, B.T. Fasy, J.Lazovskis, G. Malen, H. Riess, B. Wang, and M. Zabka, Moduli spaces of Morse functions for persistence, *J. Appl. Comp. Top.* **4** (2020), 1–33.

[**CMW**] J. Cyranka, K. Mischaikow, and C. Weibel, Contractibility of a persistence map preimage, *J. Appl. Comp. Top.* **4** (2020), 509–523.

[**C**] J. Curry, The fiber of the persistence map for functions on the interval, *J. Appl. Comp. Top.* **2** (2018), 301–321.

[**LT**] J. Leygonie and U. Tillmann, The fibre of the persistence homology map for simplicial complexes, preprint, 2021. Posted at arXiv:2104.01372.

[**WK**] C. Weibel, *The K-book*, Grad. Studies in Math. 145, AMS, 2013.

Konstantin Mischaikow mischaik@math.rutgers.edu
Charles Weibel weibel@math.rutgers.edu

Department of Mathematics, Rutgers University, New Brunswick, NJ 08901, USA