WEIGHTED PERSISTENT HOMOLOGY

G. BELL, A. LAWSON, J. MARTIN, J. RUDZINSKI, AND C. SMYTH

Abstract. We introduce weighted versions of the classical Čech and Vietoris-Rips complexes. We show that a version of the Vietoris-Rips Lemma holds for these weighted complexes and that they enjoy appropriate stability properties. We also give some preliminary applications of these weighted complexes.

1. Introduction

Topological data analysis (TDA) provides a means for the power of algebraic topology to be used to better understand the shape of a dataset. In the traditional approach to TDA, isometric balls of a fixed radius \( r > 0 \) are centered at each data point in some ambient Euclidean space. One then constructs the nerve of the union of these balls and computes the simplicial homology of this nerve. Computationally, this approach is infeasible, so instead one computes the so-called Vietoris-Rips complex, which is the flag complex over the graph obtained by placing an edge between any pair of vertices that are at distance no more than \( 2r \) from each other.

The key idea of TDA is to allow the radius of these balls to vary and to compute simplicial homology for each value of this radius to create a topological profile of the space. This profile is encoded in either a barcode or a persistence diagram. Topological features such as holes or voids that exist for a relatively large interval of radii are said to persist and are believed to be more important than more transient features that exist for very short intervals of radii. (There are, however, important exceptions to this rule of thumb, see [1]).

In the traditional model, the radius of each ball is the same and can be modeled by the linear function of time \( r(t) = rt \). In this paper, we consider a model of computing persistent homology in which the radius of each ball is allowed to be a different monotonic function \( r_x(t) \) at each point \( x \). In this way we can emphasize certain data points by assigning or weighting them with larger and/or more quickly growing balls and de-emphasize others by weighting them with smaller and/or more slowly growing balls. This is appropriate in the case of a noisy dataset, for instance, as an alternative to throwing away data that fails to meet some threshold of significance. Various other methods of enhancing persistence methods with weights have been considered, (e.g. [3, 7, 13, 14]).

The weighted model we propose fits into the framework of generalized persistence in the sense of [2]. We show that it enjoys many of the properties familiar from the techniques of traditional persistent homology. We prove a weighted Vietoris-Rips Lemma (Theorem 3.2) that relates our weighted Čech and Rips complexes in the

2010 Mathematics Subject Classification. 54F45 (primary), 20F69 (secondary).

Key words and phrases. persistent homology.

Clifford Smyth was supported by NSA MSP Grant H98230-13-1-0222 and by a grant from the Simons Foundation (Grant Number 360486, CS).
same way that they are related in the case of isometric balls. We also show that
the persistent homology computed over weighted complexes is stable with respect
to small perturbations of the rates of growth and/or the points in the dataset
(Theorem 4.1).

As a proof of concept, we apply our methods to the Modified National Institute
of Standards and Technology (MNIST) data set of handwritten digits translated
into pixel information. Our method proves more effective than isometric persistence
in finding the number 8 from among these handwritten digits. (We chose 8 for its
unique 1-dimensional homology among these digits.) We found our methods to be
95.8% accurate as opposed to isometric persistence’s 92.07% accuracy.

In Section 2, we provide the background definitions that are needed for what
follows and describe our weighted persistence model. In Section 3 we prove the
weighted Vietoris-Rips Lemma. In Section 4 we establish our stability results.
Our experiments on MNIST data appear in Section 5. We end with some remarks
and questions for further study.

2. Preliminaries

We begin by defining some terminology and setting our notation. We will assume
some familiarity with simplicial homology and the basic ideas of topological data
analysis. For details, we refer to [6, 15].

In algebraic topology, simplicial homology is a tool that assigns to any simplicial
complex $K$ a collection of $\mathbb{Z}$-modules $H_0(K), H_1(K), \ldots$, called homology groups,
in such a way that the rank of $H_n(K)$ describes the number of “$n$-dimensional
holes” in $K$. For our purposes, we replace the standard definition in terms of
$\mathbb{Z}$-modules with vector spaces (usually over the field with two elements, for ease of
computation). We therefore refer to homology vector spaces instead of homology
groups. We do not attempt to define $H_n(K)$ here, but instead refer to any text in
algebraic topology, such as [15].

Let $U$ be a collection of sets. We define the nerve $\mathcal{N}(U)$ to be the abstract simplicial
complex with vertex set $U$ with the property that the subset $\{U_0, U_1, \ldots, U_n\}$
of $U$ spans an $n$-simplex in $\mathcal{N}$ whenever $\bigcap_{i=0}^n U_i \neq \emptyset$.

Let $(X, d)$ be a metric space. We define $B_r(x) = \{y \in X | d(x, y) \leq r\}$ to be the
closed ball of radius $r$ about $x$. (Note that we’re abusing notation since this is not
necessarily the closure of the open ball, usually denoted $B_r(x)$). We most often
consider examples where $X$ is a subset of $\mathbb{R}^d$ and $d(x, y) = ||x - y||$ is the Euclidean
distance between $x$ and $y$. For a real number $r \geq 0$, we define the Čech complex
of $X$ at scale $r$ by $\check{C}ech(r) = N\{B_r(x) | x \in X\}$.

We generalize this construction by allowing the radius of the ball around each
element $x$ to depend on $x$. Let $r : X \to [0, \infty)$ be any function. We define the weighted $r$-Čech complex $\check{C}ech(r)$ of $X$ by $\check{C}ech(r) = N\{B_{r(x)}(x)\}$.

In practice, it is difficult to determine whether an intersection of balls is nonempty.
A much simpler construction to use is the Vietoris-Rips complex. For a given pa-
rameter $r \geq 0$ the Vietoris-Rips complex is the flag complex of the 1-skeleton
of the Čech complex, i.e. a collection of $n+1$ balls forms an $n$-simplex in the
Vietoris-Rips complex if and only if the balls are pairwise intersecting. For the
Vietoris-Rips complex we identify each ball with its center, so that the Vietoris-
Rips complex at scale $r$ is $VR(r) = \{\sigma \subset X | \text{diam}(\sigma) \leq 2r\}$. Similarly, if
r : X → [0, ∞), the \textbf{weighted r-Vietoris-Rips complex} is \( VR(r) = \{ \sigma \subset X \mid d(x, y) \leq r(x) + r(y), \text{for all } x, y \in \sigma \text{ with } x \neq y \} \).

Fix \( r : X \rightarrow [0, \infty) \) and consider the simplicial complex \( \check{\text{Cech}}(r) \) (or \( VR(r) \)). Using simplicial homology with field coefficients, one can associate homology vector spaces \( H_*(\check{\text{Cech}}(r)) \) to these simplicial complexes. Whenever \( t_0 \leq t_1 \) there is a natural inclusion map of simplicial complexes given by \( t : \check{\text{Cech}}(t_0 r) \rightarrow \check{\text{Cech}}(t_1 r) \) (or the corresponding inclusion of the Vietoris-Rips complexes). By functoriality, there is an induced linear map on homology \( \iota_* : H_*(\check{\text{Cech}}(t_0 r)) \rightarrow H_*(\check{\text{Cech}}(t_1 r)) \).

Let \( X \subset \mathbb{R}^d \) be finite. Although we defined the weighted complexes above for any function \( r : X \rightarrow [0, \infty) \), we want to study the persistence properties of these weighted complexes. For example, in the case of the weighted \( \check{\text{Cech}} \) complex, we want to study the evolution of homology as the radii of the balls grow to infinity.

One straightforward way to do this would be to simply scale our weighted complexes linearly in the same way that one usually scales the isometric balls in persistent homology. We prefer a more flexible approach, which we describe in terms of radius functions.

Let \( \mathcal{C}_r^1 = \mathcal{C}_1^1([0, \infty]) \) denote the collection of differentiable bijective functions \( \phi : [0, \infty) \rightarrow [0, \infty) \) with positive first derivative. By a \textbf{radius function} on \( X \) we mean a function \( r : X \rightarrow \mathcal{C}_r^1 \). We denote the image function \( r(x) \) by \( r_x \).

For \( t \geq 0 \), we define the \textbf{Čech and Vietoris-Rips complexes at scale} \( t \) by
\[
\check{\text{Cech}}_r(t) = \mathcal{N}\{ \check{B}_{r_x(t)}(x) \}
\]
and
\[
VR_r(t) = \{ \sigma \subset X \mid d(x, y) \leq r_x(t) + r_y(t) \text{ for all } x, y \in \sigma \text{ with } x \neq y \},
\]
respectively. We define the \textbf{entry function},
\[
f_{X,r}(x) = \min_{x \in X} \{ r_x^{-1}(d(x, x_i)) \}.
\]
This function captures the scale \( t \) at which the point \( x \in \mathbb{R}^d \) is first captured by some ball \( \check{B}_{r_x(t)}(x) \); we have \( f_{X,r}(y) = t \) if and only if \( y \in \check{B}_{r_x(t)}(x) \) for some \( x \in X \) and \( y \not\in \bigcup_{x \in X} \check{B}_{r_x(t)}(x) \). Thus we have the following proposition.

\textbf{Proposition 2.1.} Let \( X \) be a finite subset of some Euclidean space \( \mathbb{R}^d \). Suppose that \( r \) and \( f_{X,r} \) are defined as above. Then,
\[
f_{X,r}^{-1}([0, t]) = \bigcup_{x_i \in X} \check{B}(x_i, r_x(t)).
\]

It follows from the Nerve Lemma (see for example, [8, Corollary 4G.3]), that \( \check{\text{Cech}}_r(t) \) is homotopy equivalent to \( f_{X,r}^{-1}([0, t]) \).

\section{A Weighted Vietoris-Rips Lemma}

The Vietoris-Rips complex is much easier to compute than the \( \check{\text{Cech}} \) complex. To determine whether \( n + 1 \) balls form an \( n \)-simplex in the \( \check{\text{Cech}} \) complex, we must check whether the balls intersect, a computationally complex problem. To determine whether \( n + 1 \) balls \( B_{r_i}(x_i) \) form a simplex in the Vietoris-Rips complex is computationally easy, only \( \binom{n+1}{2} \) condition \( d(x_i, x_j) \leq r_i + r_j \) need be checked. Furthermore, if there are \( m \) points in \( X \), it may be necessary to check all \( 2^m \) sub-collections of balls to determine the \( \check{\text{Cech}} \) complex, whereas determining the Rips complex will only require checking \( \binom{m}{2} \) pairs of points.
The alpha complex, which is similar to the Čech and Vietoris-Rips complexes, and the weighted alpha complex are discussed in [6]. These weighted alpha complexes allow for balls of different radii around each point just as with the weighted Čech and Vietoris-Rips complexes. There are two main advantages to using weighted Čech and Vietoris-Rips complexes instead of weighted alpha complexes. Firstly, alpha complexes may not capture all relevant persistence information in higher dimensions. Secondly, it is clear from the definitions that alpha complexes are necessarily more computationally complex than either Čech or Vietoris-Rips complexes. Indeed our method of finding weighted Vietoris-Rips complexes requires only a marginally more computation than the unweighted Vietoris-Rips complex.

In computational problems it is common to use the Vietoris-Rips complex instead of the Čech complex to simplify the calculational overhead. The following theorem justifies this decision by saying that the Vietoris-Rips complex is “close” enough to the weighted Čech complex.

The classical Vietoris-Rips Lemma can be stated as follows:

**Theorem 3.1.** [5] Let \( X \) be a set of points in \( \mathbb{R}^d \) and let \( t > 0 \). Then

\[
\text{VR}(t') \subseteq \text{Čech}(t) \subseteq \text{VR}(t)
\]

whenever \( 0 < t' \leq t \left( \sqrt{2d/(d+1)} \right)^{-1} \).

The main result of this section is an extension of this result to the weighted case.

**Theorem 3.2** (Weighted Vietoris-Rips Lemma). Let \( X \) be a set of points in \( \mathbb{R}^d \). Let \( r : X \to (0, \infty) \) be the corresponding weight function and let \( t > 0 \). Then

\[
\text{VR}(t'r) \subseteq \text{Čech}(tr) \subseteq \text{VR}(tr)
\]

whenever \( 0 < t' \leq t \left( \sqrt{2d/(d+1)} \right)^{-1} \).

**Proof.** The second containment \( \text{Čech}(tr) \subseteq \text{VR}(tr) \) follows from the fact that the weighted Vietoris-Rips complex is the flag complex of the weighted Čech complex.

To show that \( \text{VR}(t'r) \subseteq \text{Čech}(tr) \), we suppose there is some finite collection \( \sigma = \{ x_k \}_{k=0}^\ell \subseteq \mathbb{R}^d \) with \( k > 0 \) that is a simplex in \( \text{VR}(t'r) \) and show that this is also a simplex in \( \text{Čech}(tr) \). We have \( \| x_i - x_j \|_2 \leq t'(r(x_i) + r(x_j)) \) whenever \( i \neq j \).

Define a function \( f : \mathbb{R}^d \to \mathbb{R} \) by

\[
f(y) = \max_{0 \leq j \leq \ell} \left\{ \frac{\| x_j - y \|_2}{r(x_j)} \right\}.
\]

Clearly, \( f \) is continuous and \( |f| \to \infty \) as \( \|y\|_2 \to \infty \). Thus \( f \) attains a minimum (say at \( y_0 \)) on some compact set containing \( \text{Conv}(\{x_k\}_{k=0}^\ell) \). There is some positive integer \( n \leq \ell \) so that \( \| x_i - y_0 \|_2 / r(x_i) = f(y_0) \) for exactly \( n \) of these vertices \( x_i \).

By reordering the vertices, we may assume that

\[
f(y_0) = \frac{1}{r(x_j)} \| x_j - y_0 \|_2^2 \quad \text{if } 0 \leq j \leq n
\]

and

\[
f(y_0) > \frac{1}{r(x_j)} \| x_j - y_0 \|_2^2 \quad \text{if } n < j \leq \ell
\]

Let

\[
g(y) = \max_{0 \leq j \leq n} \left\{ \frac{1}{r(x_j)} \| x_j - y \|_2 \right\}
\]
and
\[ h(y) = \max_{n < j \leq \ell} \left\{ \frac{1}{r(x_j)} \|x_j - y\|_2 \right\}. \]

Now we wish to show that \( y_0 \in \text{conv}\{x_j\}_{j=0}^n \). To this end we apply the Separation Theorem \[12\] to obtain: either \( y_0 \in \text{Conv}\{x_j\}_{j=0}^n \) or there is a \( v \in \mathbb{R}^d \) such that \( v \cdot x_j \geq 0 \) for all \( 0 \leq j \leq n \) and \( v \cdot y_0 < 0 \). Thus if \( y_0 \notin \text{Conv}\{x_j\}_{j=0}^n \) there is a \( v \in \mathbb{R}^d \) so that \( v \cdot (x_j - y_0) > 0 \) for \( 0 \leq j \leq n \). We suppose that there is a \( v \) and derive a contraction.

Since
\[ \|x_j - (y_0 + \lambda v)\|_2^2 = \|x_i - y_0\|_2^2 - 2\lambda v \cdot (x_j - y_0) + \lambda^2 \|v\|_2^2 \]
for each \( 0 \leq j \leq n \), it follows that \( g(y_0 - \lambda v) < f(y_0) \) for all \( 0 < \lambda < \lambda_1 \), where
\[ \lambda_1 = \min_{0 \leq j \leq n} \left\{ \frac{2v \cdot (x_j - y_0)}{\|v\|_2^2} \right\}. \]

Since \( h(y) \) is continuous and \( h(y_0) < f(y_0) \), there exists a \( \lambda_2 \) so that \( h(y_0 - \lambda v) < f(y_0) \) for \( 0 < \lambda < \lambda_2 \). Thus, there exists a \( \lambda > 0 \) such that
\[ f(y_0 - \lambda v) = \max \{g(y_0 - \lambda v), h(y_0 - \lambda v)\} < f(y_0), \]
contradicting the minimality of \( y_0 \).

By Carathéodory’s theorem \[12\] and reordering of vertices if necessary, \( y_0 \) is a convex combination of some subcollection of vertices \( \{x_j\}_{j=0}^m \) where \( m \leq \min\{d, n\} \). It is not possible that \( m = 0 \). If so, then \( y_0 = x_0 \) and \( f(y_0) = \frac{1}{r(x_0)} \|x_0 - y_0\|_2 = 0 \) and \( f \) is identically zero. Since \( \sigma \) has dimension at least 1, it contains a vertex \( x_1 \neq x_0 \). It follows that \( f(y_0) = f(x_0) > \frac{1}{r(x_1)} \|x_1 - x_0\|_2 > 0 \), which is a contradiction.

Let \( \hat{x}_j = x_j - y_0 \) for all \( 0 \leq j \leq m \). Note that
\[ (1) \quad \|\hat{x}_j\|_2^2 = r(x_j)^2 f(y_0)^2. \]

Since \( y_0 \in \text{Conv}\{x_j\}_{j=0}^m \), \( y_0 = \sum_{j=0}^m a_j x_j \) for some set of non-negative real numbers \( a_0, \ldots, a_m \) that sum to 1. Thus \( \sum_{j=0}^m a_j \hat{x}_j = 0 \). By relabeling, we may assume that \( a_0 r(x_0) \geq a_j r(x_j) \) when \( j > 0 \). Necessarily \( a_0 > 0 \). (Otherwise \( a_j = 0 \) for all \( 0 \leq j \leq m \), a contradiction.) Then,
\[ \hat{x}_0 = -\sum_{j=0}^m \frac{a_j}{a_0} \hat{x}_j \]
and so
\[ r(x_0)^2 f(y_0)^2 = \|\hat{x}_0\|_2^2 = -\sum_{j=0}^m \frac{a_j}{a_0} \hat{x}_j \cdot \hat{x}_j. \]

Among the indices \( 1, 2, \ldots, m \), there is some \( j_0 \) such that
\[ (2) \quad \frac{1}{d} r(x_0)^2 f(y_0)^2 \leq \frac{1}{m} r(x_0) f(y_0)^2 \leq -\frac{a_{j_0}}{a_0} \hat{x}_0 \cdot \hat{x}_{j_0}. \]

We must have \( a_{j_0} > 0 \). (Otherwise, \( f(y_0) = 0 \), which, as shown earlier, is a contradiction.) By reordering, we may assume \( j_0 = 1 \). Putting (1) and (2) together,
we find
\[ f(y_0)^2 \left( r(x_0)^2 + \frac{2a_0 r(x_0)^2}{a_1 d} + r(x_1)^2 \right) \leq \| \tilde{x}_0 \|_2^2 - 2 \tilde{x}_0 \cdot \tilde{x}_1 + \| \tilde{x}_1 \|_2^2 \]
\[ = \| \tilde{x}_0 - \tilde{x}_1 \|_2^2 \]
\[ = \| x_0 - x_1 \|_2^2 \]
\[ \leq (t'(r(x_0) + r(x_1)))^2. \]

We will now show that
\[ \frac{(r(x_0))^2 + r(x_1)^2)}{r(x_0)^2 + \frac{2a_0 r(x_0)^2}{a_1 d} + r(x_1)^2} \leq \frac{2d}{d + 1}. \]
It suffices to show \((d - 1 + 4 \frac{a_0}{a_1}) r(x_0)^2 - 2(d + 1)r(x_0)r(x_1) + (d - 1)r(x_1)^2 \geq 0.\)

Since \(\frac{a_0}{a_1} \geq \frac{r(x_1)}{r(x_0)}\) we get
\[ \left( d - 1 + 4 \frac{a_0}{a_1} \right) r(x_0)^2 - 2(d + 1)r(x_0)r(x_1) + (d - 1)r(x_1)^2 \]
\[ \geq \left( d - 1 + 4 \frac{r(x_1)}{r(x_0)} \right) r(x_0)^2 - 2(d + 1)r(x_0)r(x_1) + (d - 1)r(x_1)^2 \]
\[ = (d - 1)(r(x_0) - r(x_1))^2 \]
\[ \geq 0 \]
as desired. Our assumption that \(t' \leq t(\sqrt{2d/(d + 1)} - 1)\) implies \(f(y_0) \leq t\) and thus
\[ y_0 \in \bigcap_{i=0}^t \tilde{B}_{t r(x_i)}(x_i). \]
Therefore \(\sigma \in \tilde{C}(tr)\) and we are done. \(\square\)

4. Stability

In this section we discuss the stability of our weighted persistence. Informally, we prove results show that small changes to \(X\) and \(r\) result only in small changes to the corresponding persistence diagram or barcode.

Let \(X\) and \(Y\) be finite subsets of \(R^d\) with the same cardinality. Let \(\eta : X \rightarrow Y\) be a bijection. For notational convenience, we will take \(X = \{x_1, \ldots, x_n\}, Y = \{y_1, \ldots, y_n\}\) and \(\eta(x_i) = y_i\).

Suppose \(K\) is a compact subset of \(R^d\) that contains \(X \cup X'\). Given a continuous function \(f : [0, \text{diam}(K)] \rightarrow \mathbb{R}\) let \(||f|| = \max_{t \in [0, \text{diam}(K)]} |f(t)|\). Given a continuous function \(F : K \rightarrow \mathbb{R}\) let \(||F||_K = \max_{x \in K} |F(x)|\). Let \(||\eta|| = \max_{x \in X} d(x, \eta(x)) = \max_j d(x_j, y_j)\).

Let \(r : X \rightarrow \mathcal{C}^1_+\) and \(s : Y \rightarrow \mathcal{C}^1_+\).

**Theorem 4.1.** In the above notation,
\[ ||f_{X,r} - f_{Y,s}||_K \leq \max_j ||r_{x_j}^{-1} - s_{y_j}^{-1}|| + ||\eta|| \max_j \left( ||(r_{x_j}^{-1})'||, ||(s_{y_j}^{-1})'|| \right). \]

**Proof.** There is some point \(x\) in the compact set \(K\) and some pair of indices \(i\) and \(j\) so that
\[ ||f_{X,r} - f_{Y,s}||_K = |f_{X,r}(x) - f_{Y,s}(x)| = |r_{x_i}^{-1}(d(x, x_i)) - s_{y_j}^{-1}(d(x, y_j))|. \]
We first suppose \( r_{x_i}^{-1}(d(x, x_i)) \geq s_{y_j}^{-1}(d(x, y_j)) \). Since \( r_{x_j}^{-1}(d(x, x_j)) \geq r_{x_i}^{-1}(d(x, x_i)) \) we have
\[
\|f_{X,r} - f_{Y,s}\|_K \leq \|r_{x_j}^{-1}(d(x, x_j)) - s_{y_j}^{-1}(d(x, y_j))\|
\leq |r_{x_j}^{-1}(d(x, x_j)) - s_{y_j}^{-1}(d(x, x_j))| + |s_{y_j}^{-1}(d(x, x_j)) - s_{y_j}^{-1}(d(x, y_j))|
\]
Since \( d(x, x_j) \in [0, \text{diam}(K)] \),
\[
|r_{x_j}^{-1}(d(x, x_j)) - s_{y_j}^{-1}(d(x, x_j))| \leq \|r_{x_j}^{-1} - s_{y_j}^{-1}\| \leq \max_j \|r_{x_j}^{-1} - s_{y_j}^{-1}\|.
\]
We apply the mean value theorem to obtain the bound
\[
|s_{y_j}^{-1}(d(x, x_j)) - s_{y_j}^{-1}(d(x, y_j))| \leq \|\eta\| \cdot \max_j \|l_{r_{x_j}^{-1}}(\eta)\| \leq \|\eta\| \max_j \|l_{r_{x_j}^{-1}}(\eta)\| \cdot \|l_{s_{y_j}^{-1}}(\eta)\|.
\]
Together, these last two bounds give the bound of the theorem. A similar argument gives the same bound if \( r_{x_i}^{-1}(d(x, x_i)) \leq s_{y_j}^{-1}(d(x, y_j)) \). \(\square\)

We have the following immediate corollaries of Theorem 4.2.

**Corollary 4.2.** If the radii functions are all linear, i.e. if for all \( j \) there are positive constants \( r_{x_j} \) and \( s_{y_j} \) such that \( r_{x_j}(t) = r_{x_j} t \) and \( s_{y_j}(t) = s_{y_j} t \), then
\[
\|f_{X,r} - f_{Y,s}\|_K \leq \text{diam}(K) \max_j \left| \frac{1}{r_{x_j}} - \frac{1}{s_{y_j}} \right| + \|\eta\| \max_j \left( \frac{1}{r_{x_j}}, \frac{1}{s_{y_j}} \right).
\]

The next corollary is a point stability result concerning the case in which the points are perturbed but the weight functions stay the same.

**Corollary 4.3.** (Point-stability) If only the locations of the points are perturbed and the radius functions stay the same, i.e. \( s_{y_j}(t) = r_{x_j}(t) \) for all \( j \), then
\[
\|f_{X,r} - f_{Y,s}\|_K \leq \|\eta\| \max_j \left\| \left( r_{x_j}^{-1} \right)' \right\|.
\]

The last corollary is a weight-function stability result concerning a case in which the points stay the same but the weight functions are perturbed.

**Corollary 4.4.** (Weight-function stability) In the notation of this section,
\[
\|f_{X,r} - f_{X,s}\|_K \leq \max_j \left\{ \|r_{x_j}^{-1} - s_{x_j}^{-1}\| \right\}.
\]

We now show the stability of the persistence diagrams of \( f_{X,r} \) under perturbations of \( X \) and \( r \). Let \( f : K \to [0, \infty) \) be a real valued function on a compact set \( K \subseteq \mathbb{R}^d \). The **persistence diagram** of \( f \), \( \text{dgm}(f) \), is a multi-set of points in \([0, +\infty)^2 \) recording the appearance and disappearance of homological features in \( f^{-1}([0, t]) \) as \( t \) increases. Each point \((b, d)\) in the diagram tracks a single homological feature, recording the scale \( t = b \) at which the feature first appears and the scale \( t = d \) at which it disappears \([6]\). It should also be noted that if one considers the birth-death pair as an interval we obtain the **barcode** as seen in \([17]\) (see Figures 2 and 3). Given two persistence diagrams \( P \) and \( Q \) (where as usual we include all points along the diagonal) we let \( N \) denote the set of all bijections from \( P \) to \( Q \). We recall that the **bottleneck distance** between the diagrams \([6]\) is given by
\[
\inf_{\eta \in N} \sup_{x \in P} \|x - \eta(x)\|_\infty.
\]
We have
Theorem 4.5. \cite[Theorem 6.9]{4} Suppose $\mathcal{X}$ is a triangulable space and that $f : \mathcal{X} \to \mathbb{R}$ and $g : \mathcal{X} \to \mathbb{R}$ are tame, continuous functions. If $(f - g)$ is bounded, then for each $n$,

$$d_B(\text{dgm}_n(f), \text{dgm}_n(g)) \leq \|f - g\|_{\infty}$$

where $d_B$ denotes the bottleneck distance and $\text{dgm}_n(f)$ denotes the $n$-th persistence diagram of the filtration of $f$.

We refer to \cite{6} for the technical definitions of tame and triangulable. Note that as our spaces are nerves of balls around finite collections of points, they are finite simplicial complexes. Hence they are triangulable and only admit tame functions. Thus for our setting we get the following corollary.

Corollary 4.6. Let $X$ and $Y$ be finite subsets of $\mathcal{X}$ with equal size. Suppose that $\eta : X \to Y$ is a one-to-one correspondence and that $r : X \to C^1_+$ and $s : Y \to C^1_+$. Let $K$ be a compact subset containing $X \cup Y$. Then, for each $n$,

$$d_B(\text{dgm}_n(f_{X,r}), \text{dgm}_n(f_{Y,s})) < \max_j \|r^{-1}_x - s^{-1}_y\| + \|\eta\| \max_j \left(\|r^{-1}_x\|', \|s^{-1}_y\|'\right).$$

5. MNIST eights recognition

In this section, we give an application of weighted persistence to a simple computer vision problem. We apply our methods to the Modified National Institute of Standards and Technology (MNIST) data set of handwritten digits \cite{10}. The set consists of handwritten digits (0 through 9) translated into pixel information. Each row contains a label and 784 other values ranging from 0 to 255 that correspond to a 28 by 28 grid of pixels. The values 0 through 255 correspond to the intensity of the pixels in gray-scale with 0 meaning completely black and 255 meaning completely white. Considering the digits from zero through nine, unweighted persistence would easily be able to classify these numbers as having zero, one, or two holes, provided they are well written. However, handwriting is not perfect. Consider an eight as in figure 1. Unweighted persistence would pick up on two holes, but one of those holes might be slightly too small and ultimately considered insignificant, see figure 3. Our methods are able to pick up on both holes and would count them as significant, see figure 2. We chose to work with the digit eight due to its unique homology.

To begin, we convert each 28 by 28 grid of pixels to a 28 by 28 matrix whose entries are the pixel intensities. We treat the location of a value in the matrix as a location in the plane. That is, the value in the $i$th row, $j$th column corresponds to the point $(i,j)$. We then use Javaplex \cite{16} to compute the persistent homology.
Figure 2. Weighted persistence produces a barcode clearly with two long bars.

Figure 3. We see unweighted persistence produce a barcode that has one long bar. The second longest bar is hard to distinguish (in length) from the rest.

using a weighted distance matrix for weighted persistence and the analogous matrix for unweighted persistence. We compare weighted persistence to unweighted persistence by measuring accuracy of classifying eights. Notice in the barcodes that the deciding factor in determining an eight is the ability to distinguish the length of the second longest bar from the length of the third longest and smaller bars. For this reason, we consider the ratio of the third longest bar to the second longest bar. We will say (arbitrarily) that a barcode represents an eight if this ratio is less than $\frac{1}{2}$. For each of the 42000 handwritten digits in the MNIST data set, we compute both weighted and unweighted persistence and collect the predictions. We obtain the confusion matrices as in Figure 4.

Notice that the weighted persistence has an accuracy rate of 95.8% whereas unweighted persistence had an accuracy of 92.07%. A full summary can be seen
Figure 4. The confusion matrices show that weighted persistence outperforms its unweighted counterpart.

Figure 5

6. Concluding remarks and open questions

The method of weighted persistence satisfies the appropriate Veitoris-Rips Lemma, is stable under small perturbations of the points, or the weights, or both, and can be successfully applied to data such as the MNIST data set to improve upon usual persistence. We conclude the paper with some further observations and questions.

One can imagine weighted persistence as interpolating between two extreme approaches to a data set that is partitioned into data $D$ and noise $N$. More precisely, we consider a noisy data set $X$. Various methods exist to filter $X$ into data $D$ and noise $N$. Traditional persistence can be applied to $D \cup N$ in two ways. We can either assign the same radius to every point of $D \cup N$ or we can throw the points of $N$ out entirely and compute persistence on $D$ alone. Using weighted persistence, we can assign the radius 0 to each point of $N$ and compute weighted persistence of $D \cup N$. It is easy to see that this will differ from persistence of $D$ itself only in dimension 0. By gradually increasing the $N$-radii from 0 to 1, our stability results can be interpreted as producing a continuum of barcodes/persistence diagrams that interpolate between the usual persistence applied to $D$ and the usual persistence applied to $D \cup N$ (in dimensions above 0), see [9].

As mentioned in the introduction, weighted persistence fits into the framework of generalized persistence in the sense of [2]. This direction was explored in detail in [11].

Finally, it would be interesting to apply weighted persistence to the MNIST data set to determine its effectiveness in distinguishing the 1-homology of the other nine digits. One complication is that the number 4 presents an interesting challenge since it is appropriate to write it both as a simply connected space and as a space...
with non-trivial $\pi_1$. Distinguishing 1-homology creates 3 clusters of digits from which we could use other machine learning techniques to create an ensemble and make accurate predictions.

References

[1] P. Bendich, J. S. Marron, E. Miller, A. Pieloch, and S. Skwerer. Persistent homology analysis of brain artery trees. *Ann. Appl. Stat.*, 10(1):198–218, 03 2016.

[2] P. Bubenik, V. de Silva, and J. Scott. Metrics for generalized persistence modules. *Found. Comput. Math.*, 15(6):1501–1531, 2015.

[3] M. Buchet, F. Chazal, S. Y. Oudot, and D. R. Sheehy. Efficient and robust persistent homology for measures. *Comput. Geom.*, 58:70–96, 2016.

[4] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer. Stability of persistence diagrams. *Discrete Comput. Geom.*, 37(1):103–120, 2007.

[5] V. de Silva and R. Ghrist. Coverage in sensor networks via persistent homology. *Algebr. Geom. Topol.*, 7:339–358, 2007.

[6] H. Edelsbrunner and J. L. Harer. *Computational topology*. American Mathematical Society, Providence, RI, 2010. An introduction.

[7] H. Edelsbrunner and D. Morozov. Persistent homology: theory and practice. In *European Congress of Mathematics*, pages 31–50. Eur. Math. Soc., Zürich, 2013.

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

[9] A. Lawson. Multiradial (multi)filtrations and persistent homology. Master’s thesis, University of North Carolina at Greensboro, 2016.

[10] Y. LeCun, C. Cortes, and C. Burges. The MNIST database of handwritten digits. This dataset was retrieved from [https://www.kaggle.com/c/digit-recognizer/data](https://www.kaggle.com/c/digit-recognizer/data).

[11] J. Martin. Multiradial (multi)filtrations and persistent homology. Master’s thesis, University of North Carolina at Greensboro, 2016.

[12] J. Matoušek. *Lectures on Discrete Geometry*. Springer, 2002.

[13] G. Petri, M. Scolamiero, I. Donato, and F. Vaccarino. Topological Strata of Weighted Complex Networks. *PLoS ONE*, 8:e66506, June 2013.

[14] S. Ren, C. Wu, and J. Wu. Weighted Persistent Homology. *ArXiv e-prints*, Aug. 2017.

[15] J. J. Rotman. *An introduction to algebraic topology*, volume 119 of Graduate Texts in Mathematics. Springer-Verlag, New York, 1988.

[16] A. Tausz, M. Vejdemo-Johansson, and H. Adams. JavaPlex: A research software package for persistent (co)homology. In H. Hong and C. Yap, editors, *Proceedings of ICMS 2014*, Lecture Notes in Computer Science 8592, pages 129–136, 2014. Software available at [http://appliedtopology.github.io/javaplex/](http://appliedtopology.github.io/javaplex/).

[17] A. Zomorodian and G. Carlsson. Computing persistent homology. *Discrete Comput. Geom.*, 33(2):249–274, 2005.