THE SPACE OF PERSISTENCE DIAGRAMS FAILS TO HAVE YU’S PROPERTY A

Abstract. We define a simple obstruction to Yu’s property A that we call $k$-prisms. This structure allows for a straightforward proof that the space of persistence diagrams fails to have property A in a Wasserstein metric.

1. Introduction

A persistence diagram is one way to visualize the persistent homology of a dataset [3]. Persistent homology allows the power of algebraic topology to be leveraged against problems in diverse disciplines [2, 6].

The space of persistence diagrams can be equipped with several natural metrics, which provide the key feature of persistence diagrams, known as stability: datasets that are close give rise to persistence diagrams that are close. In this brief note, we investigate the coarse geometric properties of persistence diagrams in a family of these natural metrics.

Coarse geometry arose out of the study of metric properties of finitely generated groups. Since Gromov’s seminal paper [4], coarse geometry has established itself as an interesting subject in its own right. Yu defined a simple condition of discrete metric spaces called property A that implies the existence of a uniform embedding in Hilbert space [9]. Nowak provided a simple example of a space that fails to have property A yet still admits a uniform embedding into Hilbert space [7].

In Theorem 2.6 we provide a simple obstruction to property A that we call $k$-prisms. This structure allows for an isometric embedding of the simplest version of Nowak’s example into the metric space in question. We show that the space of persistence diagrams has $k$-prisms, hence it cannot...
have property A. The notion of $k$-prisms was first applied to Cayley graphs of the integers with infinite generating sets [8].

We do not attempt to answer the broader question of whether persistence diagrams admit a uniform embedding into Hilbert space. The authors wish to thank Boris Goldfarb for bringing our attention to possible connections between this question and applications to machine learning.

2. AN OBSTRUCTION TO PROPERTY A

We include the definition of property A (for a discrete metric space) for completeness, but this definition is not used in a substantial way in this paper.

**Definition 2.1 ([9]).** A (discrete) metric space $X$ is said to have property A if for all $R > 0$ and all $\epsilon > 0$, there exists a family $\{A_x\}_{x \in X}$ of finite, non-empty subsets of $X \times \mathbb{N}$ such that

1. for all $x, y \in X$ with $d(x, y) \leq R$, we have $\frac{\#(A_x \Delta A_y)}{\#(A_x \cap A_y)} \leq \epsilon$, and
2. there exists a $B > 0$ such that for every $x \in X$, if $(y, n) \in A_x$, then $d(x, y) \leq B$.

Here $\#A$ is the cardinality of $A$ and $A_x \Delta A_y$ denotes the symmetric difference.

**Example 2.2 ([7, Theorem 5.1]).** Let $\{0, k\}^n$ be the set of vertices of an $n$-dimensional cube at scale $k$ endowed with the $\ell_1$-metric. Endow the disjoint union $\bigcup_{n=1}^{\infty} \{0, k\}^n$ with a metric such that the distance from $\{0, k\}^n$ to $\{0, k\}^{n+1}$ is at least $n + 1$. We denote this union of $k$-scale cubes by $C_k$; it is a locally finite metric space that fails to have property A.

In order to utilize Example 2.2, we define the notion of $k$-prisms. We show that a metric space with $k$-prisms contains an isometric copy of $C_k$.

**Definition 2.3.** Let $k$ be a positive integer. We say that a metric space $(X, d)$ has $k$-prisms if for any finite set $F \subset X$ there exists a function $T: F \to X$ such that

1. $T(F) \cap F = \emptyset$;
2. $d(T(x), T(y)) = d(x, y)$ for all $x, y \in F$; and
3. $d(x, T(y)) = k + d(x, y)$ for all $x, y \in F$.

**Remark 2.4.** Motivated by working with Cayley graphs [8], we take the $k$ in this definition to be an integer, but there is no harm in allowing $k > 0$ to be any real number. We also observe that a metric space with $k$-prisms will have $nk$-prisms for all $n \in \mathbb{N}$.

**Lemma 2.5.** Let $X$ be a metric space with $k$-prisms for some $k \geq 1$. Then,
(1) the space \( X \) contains an isometric copy of \( \{k, 2k, 3k, \ldots \} \) and
(2) for any \( x \in X \) and any \( n \in \mathbb{N} \), the space \( X \) contains an isometric copy of \( \{0, k\}^n \) with \( x \) as a vertex.

Proof. We prove (1). The proof of (2) is similar.

Fix a point \( x_0 \in X \), and let \( F = \{x_0\} \). Since \( X \) has \( k \)-prisms, there is a point \( x_1 \in X \) such that \( d(x_0, x_1) = k \). For \( n > 1 \), define \( x_n \) recursively as follows. Let \( F \) be the set \( F = \{x_0, x_1, \ldots, x_{n-1}\} \). Since \( X \) has \( k \)-prisms, use \( T \) from the definition to define \( x_n = T(x_{n-1}) \). We observe that \( d(x_{n-1}, x_n) = k \), and in general \( d(x_i, x_j) = |i - j|k \). The sequence \( \{x_0, x_1, \ldots\} \) is the required isometric copy.

\[ \Box \]

Theorem 2.6. Let \( X \) be a metric space. If \( X \) has \( k \)-prisms for some \( k \geq 1 \), then \( X \) fails to have property \( A \).

Proof. Let \( \{x_0, x_1, \ldots\} \) be an isometric copy of \( \{k, 2k, 3k, \ldots\} \) in \( X \) given by Lemma 2.5(1). Use Lemma 2.5(2) to construct copies of \( \{0, k\}^n \) with vertices along this sequence. Since \( \{x_0, x_1, \ldots\} \) is an isometric copy of \( \{k, 2k, \ldots\} \), we can arrange these cubes in such a way that the distance between \( \{0, k\}^n \) and \( \{0, k\}^{n+1} \) is at least \( n + 1 \). Thus, \( X \) contains an isometrically embedded copy of the space \( C_k \), described in Example 2.2.

\[ \Box \]

3. The space of persistence diagrams fails to have property \( A \)

The notion of a persistence diagram appears in many places. We follow the development given by Chazal, de Silva, Glisse, and Oudot [1] except that we allow more general spaces instead of focusing on the extended half-plane.

For a set \( S \), denote by \( \Delta_S \) the diagonal,

\[ \Delta_S = \{(s, s) \in S^2 \mid s \in S\}. \]

Definition 3.1. Let \( X \) be a set. A diagram on \( X \) is a function \( D: X^2 \to \mathbb{Z}_{\geq 0} \) such that \( D(p) = 0 \) for all but finitely many \( p \in X^2 \), and \( D(p) = 0 \) for all \( p \in \Delta_X \). For \( p \in X^2 \), the value \( D(p) \) is the multiplicity of \( p \). The associated labelled diagram on \( X \) is the set \( \tilde{D} \subseteq X \) given by

\[ \tilde{D} = \{(x, i) \mid i = 1, 2, \ldots, D(x)\}. \]

If \( \rho \) is a metric on \( X^2 \), we write \( \rho(\tilde{x}, \tilde{y}) \) to mean \( \rho(x, y) \), where \( \tilde{x} = (x, i) \) and \( \tilde{y} = (y, j) \) are elements of a labelled diagram on \( X \). We write \( ||\tilde{x}|| \) to mean

\[ ||\tilde{x}|| = ||(x, i)|| = \inf\{\rho(x, z) \mid z \in \Delta_X\}. \]

Definition 3.2. Let \( X \) be a set. A partial matching of labelled diagrams \( \tilde{D}_X \) and \( \tilde{D}_Y \) on \( X \) is a subset \( \tilde{m} \subseteq \tilde{D}_X \times \tilde{D}_Y \) such that
Figure 1. Determining the distance between diagrams.

(a) Two diagrams plotted on the same axes.
(b) A possible partial matching of these diagrams with one unmatched point.

Definition 3.3. Let $\tilde{m}$ be any partial matching of labelled diagrams $\tilde{D}_X$ and $\tilde{D}_Y$ on $X$. Let $\rho$ be a metric on $X^2$. Let $\pi_i(\tilde{m})$ denote the projection to the $i$-th coordinate of the partial matching $\tilde{m}$ ($i \in \{1, 2\}$). The $(\tilde{m}, \rho)$-distance, denoted $W_{\tilde{m}, \rho}(\tilde{D}_X, \tilde{D}_Y)$, is

$$W_{\tilde{m}, \rho}(\tilde{D}_X, \tilde{D}_Y) = \sum_{\tilde{x} \in \tilde{D}_X \setminus \pi_1(\tilde{m})} \|\tilde{x}\| + \sum_{\tilde{y} \in \tilde{D}_Y \setminus \pi_2(\tilde{m})} \|\tilde{y}\| + \sum_{(\tilde{x}, \tilde{y}) \in \tilde{m}} \rho(\tilde{x}, \tilde{y}).$$

The Wasserstein $\rho$-distance, denoted $W_\rho(\tilde{D}_X, \tilde{D}_Y)$, is the minimum of $W_{\tilde{m}, \rho}(\tilde{D}_X, \tilde{D}_Y)$ over the (finite) collection of all partial matchings $\tilde{m}$.

Theorem 3.4. Let $\tilde{D}_X$ be the set of all diagrams on a set $X$. If $\rho$ is a metric on $X^2$, then $W_\rho$ is a metric on $\tilde{D}_X$.

Proof. It is clear that $W_\rho$ is symmetric. The fact that $W_\rho$ is positive definite follows from the requirement that $D(p) = 0$ for all points $p \in \Delta_X$. The triangle inequality follows from Proposition 3.6. □

Definition 3.5. Let $\tilde{D}_X$, $\tilde{D}_Y$, and $\tilde{D}_Z$ be labelled diagrams. Let $\tilde{m}_{X,Z}$ be a partial matching of $\tilde{D}_X$ and $\tilde{D}_Z$, and let $\tilde{m}_{Z,Y}$ be a partial matching of $\tilde{D}_Z$ and $\tilde{D}_Y$. The composition of $\tilde{m}_{X,Z}$ and $\tilde{m}_{Z,Y}$ is the subset $\tilde{m}_{X,Y} \subseteq \tilde{D}_X \times \tilde{D}_Y$ consisting of elements $(\tilde{x}, \tilde{y})$ such that there exists $\tilde{z} \in \tilde{D}_Z$ such that $(\tilde{x}, \tilde{z}) \in \tilde{m}_{X,Z}$ and $(\tilde{z}, \tilde{y}) \in \tilde{m}_{Z,Y}$. 
It is clear that the composition of partial matchings is a partial matching.

**Proposition 3.6.** Let $X$ be a set and let $(X^2, \rho)$ be a metric space. Let $D_X$, $D_Y$, and $D_Z$ be diagrams on $X$. Then

$$W_\rho(D_X, D_Y) \leq W_\rho(D_X, D_Z) + W_\rho(D_Z, D_Y).$$

**Proof.** By definition, there exist a partial matching $\tilde{m}_{X,Z}$ of labelled diagrams $\tilde{D}_X$ and $\tilde{D}_Z$ associated to diagrams $D_X$ and $D_Z$ that realizes $W_\rho(D_X, D_Z)$ and a partial matching $\tilde{m}_{Z,Y}$ of labelled diagrams $\tilde{D}_Z$ and $\tilde{D}_Y$ associated to diagrams $D_Z$ and $D_Y$ that realizes $W_\rho(D_Z, D_Y)$. Let $\tilde{m}$ be the composition of $\tilde{m}_{X,Z}$ and $\tilde{m}_{Z,Y}$. Then,

$$W_{\tilde{m}, \rho}(D_X, D_Y) = \sum_{\tilde{x} \in \tilde{D}_X \setminus \pi_1(\tilde{m})} ||\tilde{x}|| + \sum_{\tilde{y} \in \tilde{D}_Y \setminus \pi_2(\tilde{m})} ||\tilde{y}|| + \sum_{(\tilde{x}, \tilde{y}) \in \tilde{m}} \rho(\tilde{x}, \tilde{y}).$$

We examine more closely the terms in each sum. Suppose $(\tilde{x}, \tilde{y}) \in \tilde{m}$. Then there exists $\tilde{z} \in \tilde{D}_Z$ such that $(\tilde{x}, \tilde{z}) \in \tilde{m}_{X,Z}$ and $(\tilde{z}, \tilde{y}) \in \tilde{m}_{Z,Y}$. By the triangle inequality for $\rho$, we have

$$\rho(\tilde{x}, \tilde{y}) \leq \rho(\tilde{x}, \tilde{z}) + \rho(\tilde{z}, \tilde{y}).$$

Thus

1. \[ \sum_{(\tilde{x}, \tilde{y}) \in \tilde{m}} \rho(\tilde{x}, \tilde{y}) \leq \sum_{(\tilde{x}, \tilde{z}) \in \tilde{m}_{X,Z}} \rho(\tilde{x}, \tilde{z}) + \sum_{(\tilde{z}, \tilde{y}) \in \tilde{m}_{Z,Y}} \rho(\tilde{z}, \tilde{y}). \]

If $\tilde{x} \in \tilde{D}_X \setminus \pi_1(\tilde{m})$, then $\tilde{x}$ is unmatched in $\tilde{m}$. Then either

1. $\tilde{x}$ is unmatched in $\tilde{m}_{X,Z}$ so that $x \in \tilde{D}_X \setminus \pi_1(\tilde{m}_{X,Z})$; or
2. $\tilde{x}$ is matched in $\tilde{m}_{X,Z}$ so there exists $\tilde{z} \in \tilde{D}_Z$ with $(\tilde{x}, \tilde{z}) \in \tilde{m}_{X,Z}$, but $\tilde{z}$ is unmatched in $\tilde{m}_{Z,Y}$ so that $\tilde{z} \notin \pi_1(\tilde{m}_{Z,Y})$.

For every $\tilde{x}$ and $\tilde{z}$ in a labelled diagram on $X$, the triangle inequality implies

2. \[ ||\tilde{x}|| \leq \rho(\tilde{x}, \tilde{z}) + ||\tilde{z}||. \]

Thus

3. \[ \sum_{\tilde{x} \in \tilde{D}_X \setminus \pi_1(\tilde{m}_{X,Z})} ||\tilde{x}|| \leq \sum_{\tilde{x} \in \tilde{D}_X \setminus \pi_1(\tilde{m}_{X,Z})} ||\tilde{x}|| + \sum_{(\tilde{x}, \tilde{z}) \in \tilde{m}_{X,Z}} \rho(\tilde{x}, \tilde{z}) + \sum_{\tilde{z} \in \tilde{D}_Z \setminus \pi_1(\tilde{m}_{Z,Y})} ||\tilde{z}||. \]

Similarly, if $\tilde{y} \in \tilde{D}_Y \setminus \pi_2(\tilde{m})$, then $\tilde{y}$ is unmatched in $\tilde{m}$. Then either

1. $\tilde{y}$ is unmatched in $\tilde{m}_{Z,Y}$ so that $\tilde{y} \in \tilde{D}_Y \setminus \pi_2(\tilde{m}_{Z,Y})$; or
(2) \( \tilde{y} \) is matched in \( \tilde{m}_{Z,Y} \) so there exists \( \tilde{z} \in \tilde{D}_Z \) with \( (\tilde{z}, \tilde{y}) \in \tilde{m}_{Z,Y} \), but \( \tilde{z} \) is unmatched in \( \tilde{m}_{X,Z} \) so that \( \tilde{z} \in \tilde{D}_Z \setminus \pi_2(\tilde{m}_{X,Z}) \).

Thus,

\[
(4) \sum_{\tilde{y} \in \tilde{D}_Y \setminus \pi_2(\tilde{m})} \|\tilde{y}\| \leq \sum_{\tilde{y} \in \tilde{D}_Y \setminus \pi_1(\tilde{m}_{X,Z})} \|\tilde{y}\| + \sum_{(\tilde{x},\tilde{z}) \in \tilde{m}_{X,Z}} \rho(\tilde{x}, \tilde{z}) + \sum_{(\tilde{x},\tilde{z}) \in \tilde{m}_{Z,Y}} \|\tilde{z}\|.
\]

Combining the inequalities (1), (3), and (4), we have

\[
W_{\tilde{m},\rho}(D_X, D_Y) \leq W_{\tilde{m},\rho}(D_{X,Z}) + W_{\tilde{m},\rho}(D_{Z,Y}),
\]

and the result follows. \( \square \)

**Definition 3.7.** Let \( k \geq 1 \) be an integer. A set \( X \) is \( k \)-diagrammable if there exists a metric \( \rho \) on \( X^2 \) in which the \( k \)-shell around the diagonal, \( \{ x \in X^2 \mid \rho(x, \Delta_X) = k \} \), is unbounded. Such a metric is called a diagram metric. We call a set \( X \) diagrammable if it is \( k \)-diagrammable for some \( k \).

**Lemma 3.8.** Let \( D_X \) be the set of all diagrams on a \( k \)-diagrammable set \( X \) with diagram metric \( \rho \). Then the space \( (D_X, W_\rho) \) has \( k \)-prisms.

**Proof.** Consider a finite set of diagrams \( \mathcal{F} \subseteq D_X \). Fix a non-diagonal point \( p \in X^2 \) that is not in any of the diagrams,

\[
p \in X^2 \setminus \left( \bigcup_{D \in \mathcal{F}} \{ x \mid D(x) \neq 0 \} \cup \Delta_X \right).
\]

Since \( X \) is \( k \)-diagrammable, we may assume \( p \) to have been chosen such that \( \rho(p, \Delta_X) = k \), and

\[
(5) \min \{ \rho(p, x) \mid \tilde{x} \in D, D \in \mathcal{F} \} > \max_{D, D' \in \mathcal{F}} \{ k + W_\rho(D, D') \}.
\]
Let \( T: \mathcal{F} \to \mathcal{D}_X \) be given by \( D \mapsto D + \mathbb{1}_p \). We show that \( T \) satisfies the conditions of Definition 2.3. It is clear that the partial matching \( \tilde{\rho} \) yields \( m \) defines a partial matching between \( W \) as prescribed above is

\[
\rho_{p,q} > \rho_{p,q} \quad \text{each}
\]

the collection of these points

\[
\text{take a partial matching } \tilde{m} \text{ such that } \tilde{W}_{\tilde{m},\rho}(D,D') = \tilde{W}_\rho(D,D').
\]

We claim that for every pair of diagrams \( D, D' \) in \( \mathcal{F} \), \( W_\rho(D,T(D')) = k + W_\rho(D,D') \).

Take a partial matching \( \tilde{m} \) such that \( \tilde{W}_{\tilde{m},\rho}(D,D') = W_\rho(D,D') \). Then \( \tilde{m} \) defines a partial matching between \( D \) and \( T(D') \).

Thus,

\[
W_\rho(D,T(D')) \leq \tilde{W}_{\tilde{m},\rho}(D,T(D')) = \tilde{W}_{\tilde{m},\rho}(D,D') + \|\tilde{p}\| = W_\rho(D,D') + k.
\]

If \( \tilde{m}' \) is any partial matching between \( D \) and \( T(D') \) such that \( (\tilde{x},\tilde{p}) \in \tilde{m}' \), then

\[
W_{\tilde{m}'',\rho}(D,T(D')) \geq \rho(\tilde{x},\tilde{p}) \geq W_\rho(D,D') + k,
\]

where the second inequality follows from (5). Thus, \( W_\rho(D,T(D')) = W_\rho(D,D') + k \), as required. \( \square \)

Let \( p, q > 1 \). We recall that for persistence diagrams \( D \) and \( D' \) we can calculate the Wasserstein \( p,q \)-metric as

\[
W_p^q(D,D') = \inf_{\tilde{m}} \left\{ \sum_{(\tilde{x},\tilde{y}) \in \tilde{m}} \|\tilde{x} - \tilde{y}\|_p^q + \sum_{(x_1,x_2) \in D \setminus \pi_1(\tilde{m})} |x_1 - x_2|^q + \sum_{(y_1,y_2) \in D' \setminus \pi_2(\tilde{m})} |y_1 - y_2|^q \right\}^{1/q}.
\]

Hence, we see by taking \( \rho(x,y) = \|x - y\|_p^q \), we can realize \( W_p^q \) as \( (W_\rho)^{1/q} \). Notice this function is a metric on diagrams. Moreover, for any \( k > 0 \) we see

\[
\rho(x,x + k^{1/q}, \Delta_k) = \rho((x,x + k^{1/q}), (x,x)) = (k^{1/q})^q = k.
\]

The collection of these points \( \{(x,x + k^{1/q})\} \) is unbounded. Hence, for each \( p,q > 0 \) the collection of persistence diagrams with diagram metric \( \rho \) as prescribed above is \( k \)-diagrammable for any \( k > 0 \). Thus we obtain the following.
Theorem 3.9. The space of persistence diagrams in the Wasserstein $p,q$-metric does not have property A.

There is another common metric on the space of persistence diagrams called the bottleneck distance [1]. We remark that Theorem 3.9 does not cover this case and so the following question remains open.

Question 3.10. Does the space of persistence diagrams over $\mathbb{R}_{\geq 0}$ with the bottleneck distance have property A?

Indeed, we are not even able to answer the simpler question (see [5]).

Question 3.11. Does the space of persistence diagrams over $\mathbb{R}_{\geq 0}$ with the bottleneck distance have infinite asymptotic dimension?

Finally, because the space $C_k$ does embed uniformly in Hilbert space, the existence of $k$-prisms does not seem to prevent a uniform embedding in Hilbert space. Thus, the following question remains open.

Question 3.12. Does the space of persistence diagrams (in a Wasserstein or Bottleneck metric) embed uniformly in Hilbert space?

References
