Computing Bottleneck Distance for 2-D Interval Decomposable Modules

Tamal K. Dey
Department of Computer Science and Engineering, The Ohio State University,
Columbus, Ohio, USA
http://web.cse.ohio-state.edu/~dey.8/
tamaldey@cse.ohio-state.edu

Cheng Xin
Department of Computer Science and Engineering, The Ohio State University,
Columbus, Ohio, USA
http://web.cse.ohio-state.edu/~xin.108/
xin.108@buckeyemail.osu.edu

Abstract

Computation of the interleaving distance between persistence modules is a central task in topological data analysis. For 1-D persistence modules, thanks to the isometry theorem, this can be done by computing the bottleneck distance with known efficient algorithms. The question is open for most n-D persistence modules, $n > 1$, because of the well recognized complications of the indecomposables. Here, we consider a reasonably complicated class called *2-D interval decomposable* modules whose indecomposables may have a description of non-constant complexity. We present a polynomial time algorithm to compute the bottleneck distance for these modules from indecomposables, which bounds the interleaving distance from above, and give another algorithm to compute a new distance called *dimension distance* that bounds it from below.

2012 ACM Subject Classification Mathematics of computing → Topology, Theory of computation → Computational geometry

Keywords and phrases Persistence modules, bottleneck distance, interleaving distance

Digital Object Identifier 10.4230/LIPIcs.SoCG.2018.32

Related Version A full version of the paper is available at http://arxiv.org/abs/1803.02869

Funding This research is supported by NSF grants CCF-174061, CCF-1526513, and DMS-1547357

Acknowledgements We thank the anonymous referees for insightful comments on a previous version.

1 Introduction

Persistence modules have become an important object of study in topological data analysis in that they serve as an intermediate between the raw input data and the output summarization with persistence diagrams. The classical persistence theory [19] for $\mathbb{R}$-valued functions produces one dimensional (1-D) persistence modules, which is a sequence of vector spaces (homology groups with a field coefficient) with linear maps over $\mathbb{R}$ seen as a poset. It is known that [16, 26], this sequence can be decomposed uniquely into a set of intervals called *bars* which is also represented as points in $\mathbb{R}^2$ called the persistence diagrams [15]. The space of these diagrams can be equipped with a metric d_B called the *bottleneck distance*. Cohen-Steiner et al. [15] showed that d_B is bounded from above by the input function perturbation measured in infinity norm. Chazal et al. [12] generalized the result by showing that the bottleneck distance is bounded from above by a distance d_I called the *interleaving distance* between two persistence modules; see also [6, 8, 17] for further generalizations. Lesnick [22] (see also [2, 13]) established the isometry theorem which showed that indeed $d_I = d_B$. Consequently, d_I for 1-D persistence modules can be computed exactly by efficient algorithms known for computing d_B; see e.g. [19, 20]. The status however is not so well settled for multidimensional (n-D) persistence modules [9] arising from $\mathbb{R}^n$-valued functions.

Extending the concept from 1-D modules, Lesnick defined the interleaving distance for multidimensional (n-D) persistence modules, and proved its stability and universality [22]. The definition of the bottleneck distance, however, is not readily extensible mainly because the bars for finitely presented n-D modules called *indecomposables* are far more complicated though are guaranteed to be essentially unique by Krull-Schmidt theorem [1]. Nonetheless, one can define d_B as the supremum of the pairwise interleaving distances between indecomposables, which in some sense generalizes the concept in 1-D due to the isometry theorem. Then, straightforwardly, $d_I \leq d_B$ as observed in [7], but the converse is not necessarily true. For some special cases, results in the converse direction have started to appear. Botnan and Lesnick [7] proved that, in 2-D, $d_B \leq \frac{5}{2} d_I$ for what they called block decomposable modules. Bjerkevic [4] improved this result to $d_B \leq d_I$. Furthermore, he extended it by proving that $d_B \leq (2n-1)d_I$ for rectangle decomposable n-D modules and $d_B \leq (n-1)d_I$ for free n-D modules. He gave an example for exactness of this bound when $n = 2$.

Unlike 1-D modules, the question of estimating d_I for n-D modules through efficient algorithms is largely open [5]. Multi-dimensional matching distance introduced in [10] provides a lower bound to interleaving distance [21] and can be approximated within any error threshold by algorithms proposed in [3, 11]. But, it cannot provide an upper bound like d_B. For free, block, rectangle, and triangular decomposable modules, one can compute d_B by computing pairwise interleaving distances between indecomposables in constant time because they have a description of constant complexity. Due to the results mentioned earlier, d_I can be estimated within a constant or dimension-dependent factors by computing d_B for these modules. It is not obvious how to do the same for the larger class of interval decomposable modules mentioned in the literature [4, 7] where indecomposables may not have constant complexity. These are modules whose indecomposables are bounded by “stair-cases". Our main contribution is a polynomial time algorithm that, given indecomposables, computes d_B exactly for 2-D interval decomposable modules. The algorithm draws upon various geometric and algebraic analysis of the interval decomposable modules that may be of independent interest. It is known that no lower bound in terms of d_B for d_I may exist for these modules [7]. To this end, we complement our result by proposing a distance d_0 called *dimension distance* that is efficiently computable and satisfies the condition $d_0 \leq d_I$.

All missing proofs of this article appear in the full version [18].

2 Persistence modules

Our goal is to compute the bottleneck distance between two 2-D interval decomposable modules. The bottleneck distance, originally defined for 1-D persistence modules [15] (also see [2]), and later extended to multi-dimensional persistence modules [7] is known to bound the interleaving distance between two persistence modules from above.

Let $\daleth$ be a field, **Vec** be the category of vector spaces over $\daleth$, and **vec** be the subcategory of finite dimensional vector spaces. In what follows, for simplicity, we assume $\daleth = \mathbb{Z}/2\mathbb{Z}$.

▶ **Definition 1** (Persistence module). Let $\mathbb{P}$ be a poset category. A $\mathbb{P}$-indexed persistence module is a functor $M : \mathbb{P} \to \mathbf{Vec}$. If M takes values in **vec**, we say M is pointwise finite dimensional (p.f.d). The $\mathbb{P}$-indexed persistence modules themselves form another category where the natural transformations between functors constitute the morphisms.

Here we consider the poset category to be $\mathbb{R}^n$ with the standard partial order and all modules to be p.f.d. We call $\mathbb{R}^n$-indexed persistence modules as n-dimensional persistence modules, n-D modules in short. The category of n-D modules is denoted as $\mathbb{R}^n$-**mod**. For an n-D module $M \in \mathbb{R}^n$-**mod**, we use notation $M_x := M(x)$ and $\rho^M_{x\to y} := M(x \leq y)$.

▶ **Definition 2** (Shift). For any $\delta \in \mathbb{R}$, we denote $\vec{\delta} = \delta \cdot \sum e_i$, where $\{e_i\}_{i=1}^n$ is the standard basis of $\mathbb{R}^n$. We define a shift functor $(\cdot)_{\to\delta} : \mathbb{R}^n\text{-}\mathbf{mod} \to \mathbb{R}^n\text{-}\mathbf{mod}$ where $M_{\to\delta} := (\cdot)_{\to\delta}(M)$ is given by $M_{\to\delta}(x) = M(x + \vec{\delta})$ and $M_{\to\delta}(x \leq y) = M(x + \vec{\delta} \leq y + \vec{\delta})$. In words, $M_{\to\delta}$ is the module M shifted diagonally by $\vec{\delta}$.

The following definition of interleaving taken from [24] adapts the original definition designed for 1-D modules in [13] to n-D modules.

▶ **Definition 3** (Interleaving). For two persistence modules M and N, and $\delta \geq 0$, a δ-interleaving between M and N are two families of linear maps $\{\phi_x : M_x \to N_{x+\vec{\delta}}\}_{x\in\mathbb{R}^n}$ and $\{\psi_x : N_x \to M_{x+\vec{\delta}}\}_{x\in\mathbb{R}^n}$ satisfying the following two conditions (see full version [18] for the details.)

- $\forall x \in \mathbb{R}^n, \rho^M_{x\to x+2\vec{\delta}} = \psi_{x+\vec{\delta}} \circ \phi_x$ and $\rho^N_{x\to x+2\vec{\delta}} = \phi_{x+\vec{\delta}} \circ \psi_x$
- $\forall x \leq y \in \mathbb{R}^n, \phi_y \circ \rho^M_{x\to y} = \rho^N_{x\to y} \circ \phi_x$ and $\psi_y \circ \rho^N_{x\to y} = \rho^M_{x\to y} \circ \psi_x$ symmetrically

If such a δ-interleaving exists, we say M and N are δ-interleaved. We call the first condition *triangular commutativity* and the second condition *square commutativity*.

▶ **Definition 4** (Interleaving distance). Define the interleaving distance between modules M and N as $d_I(M,N) = \inf_\delta\{M \text{ and } N \text{ are } \delta\text{-}\textit{interleaved}\}$. We say M and N are ∞-interleaved if they are not δ-interleaved for any $\delta \in \mathbb{R}^+$, and assign $d_I(M,N) = \infty$.

▶ **Definition 5** (Matching). A matching $\mu : A \nrightarrow B$ between two multisets A and B is a partial bijection, that is, $\mu : A' \to B'$ for some $A' \subseteq A$ and $B' \subseteq B$. We say $\operatorname{im}\mu = B', \operatorname{coim}\mu = A'$.

For the next definition [7], we call a module δ-*trivial* if $\rho^M_{x\to x+\vec{\delta}} = 0$ for all $x \in \mathbb{R}^n$.

▶ **Definition 6** (Bottleneck distance). Let $M \cong \bigoplus_{i=1}^m M_i$ and $N \cong \bigoplus_{j=1}^n N_j$ be two persistence modules, where M_i and N_j are indecomposable submodules of M and N respectively. Let $I = \{1, \cdots, m\}$ and $J = \{1, \cdots, n\}$. We say M and N are δ-matched for $\delta \geq 0$ if there exists a matching $\mu : I \nrightarrow J$ so that, (i) $i \in I \setminus \operatorname{coim}\mu \implies M_i$ is 2δ-trivial, (ii) $j \in J \setminus \operatorname{im}\mu \implies N_j$ is 2δ-trivial, and (iii) $i \in \operatorname{coim}\mu \implies M_i$ *and* $N_{\mu(i)}$ *are* δ-interleaved.

The bottleneck distance is defined as

$$d_B(M,N) = \inf\{\delta \mid M \text{ and } N \text{ are } \delta\text{-matched}\}.$$

The following fact observed in [7] is straightforward from the definition.

▶ **Fact 7.** $d_I \leq d_B$.

2.1 Interval decomposable modules

Persistence modules whose indecomposables are interval modules (Definition 9) are called *interval decomposable modules*, see for example [7]. To account for the boundaries of free modules, we enrich the poset $\mathbb{R}^n$ by adding points at $\pm\infty$ and consider the poset $\bar{\mathbb{R}}^n = \bar{\mathbb{R}} \times \ldots \times \bar{\mathbb{R}}$ where $\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ with the usual additional rule $a \pm \infty = \pm\infty$.

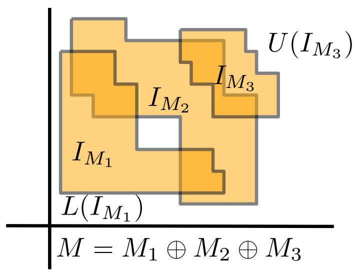
▶ **Definition 8.** An interval is a subset $\emptyset \neq I \subset \bar{\mathbb{R}}^n$ that satisfies the following:
1. If $p, q \in I$ and $p \leq r \leq q$, then $r \in I$;
2. If $p, q \in I$, then there exists a sequence $(p_1, p_2, ..., p_{2m}) \in I$ for some $m \in \mathbb{N}$ such that $p \leq p_1 \geq p_2 \leq p_3 \geq ... \geq p_{2m} \leq q$. We call the sequence $(p = p_0, p_1, p_2, ..., p_{2m}, p_{2m+1} = q)$ a path from p to q (in I).

In what follows, we fix the dimension $n = 2$. Let $\bar{I}$ denote the closure of an interval I in the standard topology of $\bar{\mathbb{R}}^2$. The lower and upper boundaries of I are defined as

$$L(I) = \{x = (x_1, x_2) \in \bar{I} \mid \forall y = (y_1, y_2) \text{ with } y_1 < x_1 \text{ and } y_2 < x_2 \implies y \notin I\}$$
$$U(I) = \{x = (x_1, x_2) \in \bar{I} \mid \forall y = (y_1, y_2) \text{ with } y_1 > x_1 \text{ and } y_2 > x_2 \implies y \notin I\}.$$

See the figure below. Let $B(I) = L(I) \cup U(I)$.

We say an interval I is discretely presented if its boundary consists of a finite set of horizontal and vertical line segments called edges, with end points called vertices, which satisfy the following conditions: (i) every vertex is incident to either a single edge or to a horizontal and a vertical edge, (ii) no vertex appears in the interior of an edge. We denote the set of edges and vertices with $E(I)$ and $V(I)$ respectively.



According to this definition, $\bar{\mathbb{R}}^2$ is an interval with boundary $B(\bar{\mathbb{R}}^2)$ that consists of all the points with at least one coordinate ∞. The vertex set $V(\bar{\mathbb{R}}^2)$ consists of four corners of the infinitely large square $\bar{\mathbb{R}}^2$ with coordinates $(\pm\infty, \pm\infty)$.

▶ **Definition 9** (Interval module)**.** A 2-D **interval persistence module**, or *interval module* in short, is a persistence module M that satisfies the following condition: for some interval $I_M \subseteq \bar{\mathbb{R}}^2$, called the interval of M,

$$M_x = \begin{cases} \mathbb{k} & \text{if } x \in I_M \\ 0 & \textit{otherwise} \end{cases} \qquad \rho^M_{x \to y} = \begin{cases} \mathbb{1} & \text{if } x, y \in I_M \\ 0 & \textit{otherwise} \end{cases}$$

It is known that an interval module is indecomposable [22].

▶ **Definition 10** (Interval decomposable module)**.** A 2-D *interval decomposable module* is a persistence module that can be decomposed into interval modules. We say a 2-D interval decomposable module is finitely presented if it can be decomposed into finitely many interval modules whose intervals are discretely presented.

3 Algorithm to compute d_B

Given the intervals of the indecomposables (interval modules) as input, an approach based on bipartite-graph matching is well known for computing the bottleneck distance $d_B(M, N)$ between two 1-D persistence modules M and N [19]. This approach constructs a bi-partite graph G out of the intervals of M and N and their pairwise interleaving distances including the distances to zero modules. If these distance computations take $O(C)$ time in total, the algorithm for computing d_B takes time $O(m^{\frac{5}{2}} \log m + C)$ if M and N together have m indecomposables altogether. Given indecomposables (say computed by Meat-Axe [23]), this approach is readily extensible to the n-D modules if one can compute the interleaving distance between any pair of indecomposables including the zero modules. To this end, we present an algorithm to compute the interleaving distance between two interval modules M_i and N_j with t_i and t_j vertices respectively on their intervals in $O((t_i + t_j)\log(t_i + t_j))$ time. This gives a total time of $O(m^{\frac{5}{2}} \log m + \sum_{i,j}(t_i + t_j)\log(t_i + t_j)) = O(m^{\frac{5}{2}} \log m + t^2 \log t)$ where t is the number of vertices over all input intervals.

Now we focus on computing the interleaving distance between two given intervals. Given two intervals I_M and I_N with t vertices, this algorithm searches a value δ so that there exists two families of linear maps from M to $N_{\to\delta}$ and from N to $M_{\to\delta}$ respectively which satisfy both triangular and square commutativity. This search is done with a binary probing. For a chosen δ from a candidate set of $O(t)$ values, the algorithm determines the direction of the search by checking two conditions called *trivializability* and *validity* on the intersections of modules M and N.

▶ **Definition 11** (Intersection module). For two interval modules M and N with intervals I_M and I_N respectively let $I_Q = I_M \cap I_N$, which is a disjoint union of intervals, $\coprod I_{Q_i}$. The intersection module Q of M and N is $Q = \bigoplus Q_i$, where Q_i is the interval module with interval I_{Q_i}. That is,

$$Q_x = \begin{cases} \daleth & \text{if } x \in I_M \cap I_N \\ 0 & \textit{otherwise} \end{cases} \quad \text{and for } x \le y,\ \rho^Q_{x\to y} = \begin{cases} \nVdash & \text{if } x, y \in I_M \cap I_N \\ 0 & \textit{otherwise} \end{cases}$$

From the definition we can see that the support of Q, *supp*(Q), is $I_M \cap I_N$. We call each Q_i an intersection component of M and N. Write $I := I_{Q_i}$ and consider $\phi : M \to N$ to be any morphism in the following proposition which says that ϕ is constant on I.

▶ **Proposition 12.** $\phi|_I \equiv a \cdot \nVdash$ for some $a \in \daleth = \mathbb{Z}/2$.

Proof.

$$\begin{tikzcd} M_{p_i} \arrow[r, "\nVdash"] \arrow[d, dashed, "\phi_{p_i}"'] & M_{p_{i+1}} \arrow[d, dashed, "\phi_{p_{i+1}}"] \\ N_{p_i} \arrow[r, "\nVdash"'] & N_{p_{i+1}} \end{tikzcd} \qquad \begin{tikzcd} M_{p_i} \arrow[d, dashed, "\phi_{p_i}"'] & M_{p_{i+1}} \arrow[l, "\nVdash"'] \arrow[d, dashed, "\phi_{p_{i+1}}"] \\ N_{p_i} & N_{p_{i+1}} \arrow[l, "\nVdash"] \end{tikzcd}$$

For any $x, y \in I$, consider a path $(x = p_0, p_1, p_2, ..., p_{2m}, p_{2m+1} = y)$ in I from x *to* y and the commutative diagrams above for $p_i \le p_{i+1}$ (left) and $p_i \ge p_{i+1}$(right) respectively. Observe that $\phi_{p_i} = \phi_{p_{i+1}}$ in both cases due to the commutativity. Inducting on i, we get that $\phi(x) = \phi(y)$. ◀

▶ **Definition 13** (Valid intersection). An intersection component Q_i is (M, N)-*valid* if for each $x \in I_{Q_i}$ the following two conditions hold (see figure below):

$$\text{(i) } y \le x \text{ and } y \in I_M \implies y \in I_N, \text{ and (ii) } z \ge x \text{ and } z \in I_N \implies z \in I_M$$

▶ **Proposition 14.** Let $\{Q_i\}$ be a set of intersection components of M and N with intervals $\{I_{Q_i}\}$. Let $\{\phi_x\} : M \to N$ be the family of linear maps defined as $\phi_x = \nVdash$ for all $x \in I_{Q_i}$ and $\phi_x = 0$ otherwise. Then ϕ is a morphism if and only if every Q_i is (M, N)-*valid*.

For proof see the full version [18].

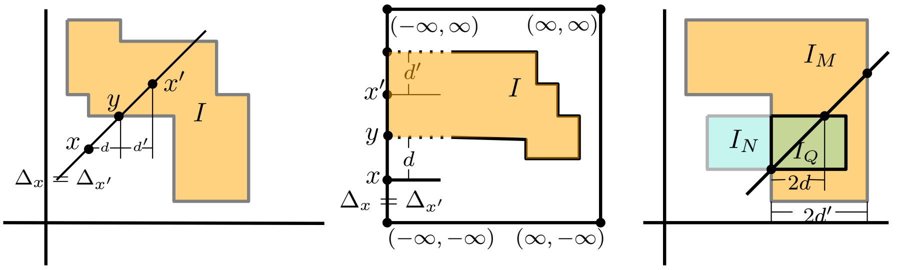
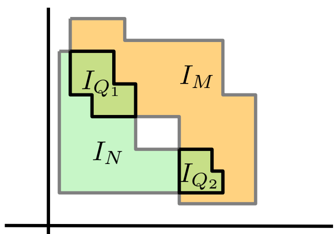


Figure 1 $d = \mathrm{dl}(x, I)$, $y = \pi_I(x)$, $d' = \mathrm{dl}(x', L(I))$ (left); $d = \mathrm{dl}(x, I)$ and $d' = \mathrm{dl}(x', U(I))$ are defined on the left edge of $B(\bar{\mathbb{R}}^2)$ (middle); Q is $d'_{(M,N)}$- and $d_{(N,M)}$-trivializable (right).

We focus on the interval modules with discretely presented intervals (figure on right). They belong to finitely presented persistence modules studied previously in [22]. For an interval module M, let $\overline{M}$ be the interval module defined on the closure $\overline{I_M}$. To avoid complication in this exposition, we assume that the upper and lower boundaries of every interval module meet exactly at two points. We also assume that every interval module has closed intervals which is justified by the following proposition (proof in the full version [18]).



Q_1 (M, N)-valid, Q_2 not

▶ **Proposition 15.** $d_I(M, N) = d_I(\overline{M}, \overline{N})$.

From the definition of boundaries of intervals, the following proposition is immediate.

▶ **Proposition 16.** Given an interval I and any point $x = (x_1, x_2) \in I \setminus (I \cap B(\bar{\mathbb{R}}^2))$, we have $x \in L(I) \iff \forall \epsilon > 0, x - \vec{\epsilon} \notin I$. Similarly, we have $x \in U(I) \iff \forall \epsilon > 0, x + \vec{\epsilon} \notin I$.

▶ **Definition 17** (Diagonal projection and distance)**.** Let I be an interval and $x \in \bar{\mathbb{R}}^2$. For $x \in \mathbb{R}^2 \subseteq \bar{\mathbb{R}}^2$, let Δ_x denote the line called *diagonal* with slope 1 that passes through x. We define (see Figure 1)

$$\mathrm{dl}(x, I) = \begin{cases} \min_{y \in \Delta_x \cap I} \{ d_\infty(x, y) := |x - y|_\infty \} \text{ if } \Delta_x \cap I \neq \emptyset \\ +\infty \text{ otherwise.} \end{cases}$$

In case $\Delta_x \cap I \neq \emptyset$, define $\pi_I(x)$, called the projection point of x on I, to be the point $y \in \Delta_x \cap I$ where $\mathrm{dl}(x, I) = d_\infty(x, y)$. For $x \in B(\bar{\mathbb{R}}^2) \setminus V(\bar{\mathbb{R}}^2)$, Δ_x is defined to be the edge in $E(\bar{\mathbb{R}}^2)$ containing x. Define $\mathrm{dl}(x, I)$ and $\pi_I(x)$ accordingly. For $x \in V(\bar{\mathbb{R}}^2)$, we set $\pi_I(x) = x$ if and only if $x \in I$. Then, $\mathrm{dl}(x, I) = 0$ if $x \in I$ and $\mathrm{dl}(x, I) = +\infty$ otherwise.

Notice that upper and lower boundaries of an interval are also intervals by definition. With this understanding, following properties of dl are obvious from the above definition.

▶ **Fact 18.**

(i) For any $x \in I_M$,

$$\mathrm{dl}(x, U(I_M)) = \sup_{\delta \in \bar{\mathbb{R}}} \{ x + \vec{\delta} \in I_M \} \text{ and } \mathrm{dl}(x, L(I_M)) = \sup_{\delta \in \bar{\mathbb{R}}} \{ x - \vec{\delta} \in I_M \}.$$

(ii) Let $L = L(I_M)$ or $U(I_M)$ and let x, x' be two points such that $\pi_L(x), \pi_L(x')$ both exist. If x and x' are on some same horizontal, vertical, or diagonal line, then $|\mathrm{dl}(x, L) - \mathrm{dl}(x', L)| \leq d_\infty(x, x')$.

Set $VL(I) := V(I) \cap L(I)$, $EL(I) := E(I) \cap L(I)$, $VU(I) := V(I) \cap U(I)$, and $EU(I) := E(I) \cap U(I)$. Following proposition is proved in the full version [18].

▶ **Proposition 19.** For an intersection component Q of M and N with interval I, the following conditions are equivalent:

(1) Q is (M, N)-*valid.*

(2) $L(I) \subseteq L(I_M)$ and $U(I) \subseteq U(I_N)$.

(3) $VL(I) \subseteq L(I_M)$ and $VU(I) \subseteq U(I_N)$.

▶ **Definition 20** (Trivializable intersection)**.** Let Q be a connected component of the intersection of two modules M and N. For each point $x \in I_Q$, define

$$d_{triv}^{(M,N)}(x) = \max\{\mathrm{dl}(x, U(I_M))/2, \mathrm{dl}(x, L(I_N))/2)\}.$$

For $\delta \geq 0$, we say a point x is $\delta_{(M,N)}$*-trivializable* if $d_{triv}^{(M,N)}(x) < \delta$. We say an intersection component Q is $\delta_{(M,N)}$*-trivializable* if each point in I_Q is $\delta_{(M,N)}$*-trivializable* (Figure 1).

Following proposition discretizes the search for trivializability (see the full version [18] for a proof).

▶ **Proposition 21.** An intersection component Q is $\delta_{(M,N)}$*-trivializable* if and only if every vertex of Q is $\delta_{(M,N)}$*-trivializable.*

Recall that for two modules to be δ-interleaved, we need two families of linear maps satisfying both triangular commutativity and square commutativity. For a given δ, Theorem 23 below provides criteria which ensure that such linear maps exist. In our algorithm, we make sure that these criteria are verified.

Given an interval module M and the diagonal line Δ_x for any $x \in \bar{\mathbb{R}}^2$, there is a 1-dimensional persistence module $M|_{\Delta_x}$ which is the functor restricted on the poset $\boldsymbol{\Delta}_x$ as a subcategory of $\bar{\mathbb{R}}^2$. We call it a *1-dimensional slice* of M along Δ_x. Define

$$\delta^* = \inf_{\delta \in \bar{\mathbb{R}}} \{\delta : \forall x \in \bar{\mathbb{R}}^2, M|_{\Delta_x} \text{ and } N|_{\Delta_x} \text{ are } \delta\text{-interleaved}\}.$$

Proposition 22 follows from the observation that $\delta^* = \sup_{x \in \bar{\mathbb{R}}^2}\{d_I(M|_{\Delta_x}, N|_{\Delta_x})\}$.

▶ **Proposition 22.** For two interval modules M, N and $\delta \in \mathbb{R}^+$, we have $\delta > \delta^*$ if and only if there exist two families of linear maps $\phi = \{\phi_x : M_x \to N_{(x+\delta)}\}$ and $\psi = \{\psi_x : N_x \to M_{(x+\delta)}\}$ such that for each $x \in \bar{\mathbb{R}}^2$, the 1-dimensional slices $M|_{\Delta_x}$ and $N|_{\Delta_x}$ are δ-interleaved by the linear maps $\phi|_{\Delta_x}$ and $\psi|_{\Delta_x}$.

▶ **Theorem 23.** *Two interval modules M and N are δ-interleaved if and only if*

- $\delta > \delta^*$*, and*
- *each component of $I_M \cap I_{N_{\to\delta}}$ is either $(M, N_{\to\delta})$-valid or $\delta_{(M,N_{\to\delta})}$-trivializable, and each component of $I_{M_{\to\delta}} \cap N$ is either $(N, M_{\to\delta})$-valid or $\delta_{(N,M_{\to\delta})}$-trivializable.*

Proof. $\Longrightarrow$ direction: Suppose M and N are δ-interleaved. By definition, we have two families of linear maps $\{\phi_x\}$ and $\{\psi_x\}$ which satisfy both triangular and square commutativities. Let the morphisms between the two persistence modules constituted by these two families of linear maps be $\phi = \{\phi_x\}$ and $\psi = \{\psi_x\}$ respectively. By Proposition 22, we get the first part of the claim that $\delta > \delta^*$. For each intersection component Q of M and $N_{\to\delta}$ with interval $I := I_Q$, consider the restriction $\phi|_I$. By Proposition 12, $\phi|_I$ is constant, that is, $\phi|_I \equiv 0$ or $\not\Vdash$. If $\phi|_I \equiv \not\Vdash$, by Proposition 14, Q is $(M, N_{\to\delta})$-*valid.* If $\phi|_I \equiv 0$, by the triangular commutativity of ϕ, we have that $\rho^M_{x \to x+2\vec{\delta}} = \psi_{x+\vec{\delta}} \circ \phi_x = 0$ for each point $x \in I$. That means $x + 2\vec{\delta} \notin I_M$. By Fact 18(i), $\mathrm{dl}(x, U(I_M))/2 < \delta$. Similarly, $\rho^N_{x-\vec{\delta} \to x+\vec{\delta}} = \phi_x \circ \psi_{x-\vec{\delta}} = 0 \implies x - \vec{\delta} \notin I_N$, which is the same as to say $x - 2\vec{\delta} \notin I_{N_{\to\delta}}$. By Fact 18(i), $\mathrm{dl}(x, L(I_{N_{\to\delta}}))/2 < \delta$. So $\forall x \in I$, we have $d^{(M,N_{\to\delta})}_{triv}(x) < \delta$. This means Q is $\delta_{(M,N_{\to\delta})}$-*trivializable*. Similar statement holds for intersection components of $M_{\to\delta}$ and N.

$\Longleftarrow$ direction: We construct two families of linear maps $\{\phi_x\}, \{\psi_x\}$ as follows: On the interval $I := I_{Q_i}$ of each intersection component Q_i of M and $N_{\to\delta}$, set $\phi|_I \equiv \mathbb{1}$ if Q_i is $(M, N_{\to\delta})$-*valid* and $\phi|_I \equiv 0$ otherwise. Set $\phi_x \equiv 0$ for all x not in the interval of any intersection component. Similarly, construct $\{\psi_x\}$. Note that, by Proposition 14, $\phi := \{\phi_x\}$ is a morphism between M and $N_{\to\delta}$, and $\psi := \{\psi_x\}$ is a morphism between N and $M_{\to\delta}$. Hence, they satisfy the square commutativity. We show that they also satisfy the triangular commutativity. We claim that $\forall x \in I_M$, $\rho^M_{x\to x+2\vec{\delta}} = \mathbb{1} \implies x + \vec{\delta} \in I_N$ and similar statement holds for I_N. From condition that $\delta > \delta^*$ and by Proposition 22, we know that there exist two families of linear maps satisfying triangular commutativity everywhere, especially on the pair of 1-dimensional persistence modules $M|_{\Delta_x}$ and $N|_{\Delta_x}$. From triangular commutativity we know that $x + \vec{\delta} \in I_N$ since otherwise one cannot construct a δ-interleaving between $M|_{\Delta_x}$ and $N|_{\Delta_x}$. Now for each $x \in I_M$ with $\rho^M_{x\to x+2\vec{\delta}} = \mathbb{1}$, we have $\mathrm{dl}(x, U(I_M))/2 \geq \delta$ by Fact 18, and $x + \vec{\delta} \in I_N$ by our claim. This implies that $x \in I_M \cap I_{N_{\to\delta}}$ is a point in an interval of an intersection component Q_x of $M, N_{\to\delta}$ which is not $\delta_{(M,N_{\to\delta})}$-*trivializable*. Hence, it is $(M, N_{\to\delta})$-*valid* by the assumption. So, by our construction of ϕ on valid intersection components, $\phi_x = \mathbb{1}$. Symmetrically, we have that $x + \vec{\delta} \in I_N \cap I_{M\to\delta}$ is a point in an interval of an intersection component of N and $M_{\to\delta}$ which is not $\delta_{(N,M_{\to\delta})}$-*trivializable* since $\mathrm{dl}(x + \vec{\delta}, L(I_M))/2 \geq \delta$. So by our construction of ψ on valid intersection components, $\psi_{x+\vec{\delta}} = \mathbb{1}$. Then, we have $\rho^M_{x\to x+2\vec{\delta}} = \psi_{x+\vec{\delta}} \circ \phi_x$ for every nonzero linear map $\rho^M_{x\to x+2\vec{\delta}}$. The statement also holds for any nonzero linear map $\rho^N_{x\to x+2\vec{\delta}}$. Therefore, the triangular commutativity holds. ◀

Note that the above proof provides a construction of the interleaving maps for a specific δ if it exists. Furthermore, the interleaving distance $d_I(M, N)$ is the infimum of all δ satisfying the two conditions in the theorem, which means $d_I(M, N)$ is the infimum of all $\delta > \delta^*$ satisfying condition 2 in Theorem 23. Based on this observation, we propose a search algorithm for computing the interleaving distance $d_I(M, N)$ for interval modules M and N.

▶ **Definition 24** (Candidate set)**.** For two interval modules M and N, and for each point x in $I_M \cup I_N$, let

$$\begin{aligned} D(x) &= \{\mathrm{dl}(x, L(I_M)), \mathrm{dl}(x, L(I_N)), \mathrm{dl}(x, U(I_M)), \mathrm{dl}(x, U(I_N))\} \text{ and} \\ S &= \{d \mid d \in D(x) \text{ or } 2d \in D(x) \text{ for some vertex } x \in V(I_M) \cup V(I_N)\} \text{ and} \\ S_{\geq\delta} &:= \{d \mid d \geq \delta, d \in S\}. \end{aligned}$$

Algorithm Interleaving (output: $d_I(M, N)$, input: I_M and I_N with t vertices in total)

1. Compute the candidate set S and let ϵ be the smallest difference between any two numbers in S. /* $O(t)$ time */
2. Compute δ^*; Let $\delta = \delta^*$. /* $O(t)$ time */
3. Output δ after a binary search in $S_{\geq\delta^*}$ by following steps /* $O(\log t)$ probes */
   - let $\delta' = \delta + \epsilon$
   - Compute intersections $I_M \cap I_{N_{\to\delta'}}$ and $I_N \cap I_{M_{\to\delta'}}$. /* $O(t)$ time */
   - For each intersection component, check if it is valid or trivializable according to Theorem 23. /* $O(t)$ time */

In the above algorithm, the following generic task of computing *diagonal span* is performed for several steps. Let L and U be any two chains of vertical and horizontal edges that are both x- and y-monotone. Assume that L and U have at most t vertices. Then, for a set X of $O(t)$ points in L, one can compute the intersection of Δ_x with U for every $x \in X$ in $O(t)$ total time. The idea is to first compute by a binary search a point x in X so that Δ_x intersects U if at all. Then, for other points in X, traverse from x in both directions while searching for the intersections of the diagonal line with U in lock steps.

Now we analyze the complexity of the algorithm Interleaving. The candidate set, by definition, has only $2t$ values which can be computed in $O(t)$ time by the diagonal span procedure. Proposition 25 shows that δ^* is in S and can be determined by computing the one dimensional interleaving distances $d_I(M|_{\Delta_x}, N|_{\Delta_x})$ for diagonal lines passing through $O(t)$ vertices of I_M and I_N. This can be done in $O(t)$ time by diagonal span procedure. Once we determine δ^*, we search for $\delta = d_I(M,N)$ in the truncated set $S_{\delta \geq \delta^*}$ to satisfy the first condition of Theorem 23. Intersections between two polygons I_M and I_N bounded by x- and y-monotone chains can be computed in $O(t)$ time by a simple traversal of the boundaries. The validity and trivializability of each intersection component can be determined in time linear in the number of its vertices due to Proposition 19 and Proposition 21 respectively. Since the total number of intersection points is $O(t)$, validity check takes $O(t)$ time in total. The check for trivializability also takes $O(t)$ time if one uses the diagonal span procedure.

Proposition 25 below says that δ^* is determined by a vertex in I_M or I_N and $\delta^* \in S$. Its proof appears in the full version [18].

▶ **Proposition 25.** (i) $\delta^* = \max_{x \in V(I_M) \cup V(I_N)} \{d_I(M|_{\Delta_x}, N|_{\Delta_x})\}$, (ii) $\delta^* \in S$.

The correctness of the algorithm Interleaving already follows from Theorem 23 as long as the candidate set contains the distance $d_I(M,N)$. The following concept of stable intersections helps us to establish this result.

▶ **Definition 26** (Stable intersection)**.** Let Q be an intersection component of M and N. We say Q is stable if every intersection point $x \in I_Q \cap B(I_M) \cap B(I_N)$ is non-degenerate, that is, x is in the interior of two edges $e_1 \in E(I_M)$ and $e_2 \in E(I_N)$, and $e_1 \perp e_2$ at x.

From Proposition 42 and Corollary 43 in Appendix A of the full version [18], we have the following claim.

▶ **Proposition 27.** $d \notin S$ if and only if each intersection component of $M, N_{\to d}$, and $N_{\to d}, M$ is stable.

The main property of a stable intersection component Q of M and N is that if we shift one of the interval module, say N, to $N_{\to\epsilon}$ continuously for some small value $\epsilon \in \mathbb{R}^+$, the interval I_{Q^ϵ} of the intersection component Q^ϵ of M and $N_{\to\epsilon}$ changes continuously. Next proposition follows directly from the stability of intersection components.

▶ **Proposition 28.** For a stable intersection component Q of M and N, there exists a positive real $\delta \in \mathbb{R}^+$ so that the following holds:

For each $\epsilon \in (-\delta, +\delta)$, there exists a unique intersection component Q^ϵ of M and $N_{\to\epsilon}$ so that it is still stable and $I_{Q^\epsilon} \cap I_Q \neq \emptyset$. Furthermore, there is a bijection $\mu_\epsilon : V(I_Q) \to V(I_{Q^\epsilon})$ so that $\forall x \in V(I_Q)$, x and $\mu_\epsilon(x)$ are on the same horizontal, vertical, or diagonal line, and $d_\infty(\mu_\epsilon(x), x) = \epsilon$. We call the set $\{Q^\epsilon \mid \epsilon \in (-\delta, +\delta)\}$ a stable neighborhood of Q.

▶ **Corollary 29.** *For a stable intersection component Q, we have:*

(i) Q is (M,N)-valid iff each Q^ϵ in the stable neighborhood is $(M, N_{\to\epsilon})$-valid.

(ii) If Q is $d_{(M,N)}$-trivializable, then Q^ϵ is $(d+2\epsilon)_{(M,N_{\to\epsilon})}$-trivializable.

Proof. (i): Let Q^ϵ be any intersection component in a stable neighborhood of Q. We know that if Q is (M,N)-valid, then $VL(I_Q) \subseteq L(I_M)$ and $VU(I_Q) \subseteq U(I_N)$. By Proposition 28, $\mu_\epsilon(VL(I_Q)) = VL(I_{Q^\epsilon}) \subseteq L(I_M)$ and $\mu_\epsilon(UL(I_Q)) = UL(I_{Q^\epsilon}) \subseteq L(I_{N\to\epsilon})$. So Q^ϵ is $(M, N_{\to\epsilon})$-valid. Other direction of the implication can be proved by switching the roles of Q and Q^ϵ in the above argument.

(ii): From Proposition 28, we have that $\forall x' \in V(I_{Q^\epsilon})$, there exists a point $x \in V(I_Q)$ so that x and x' are on some horizontal, vertical, or diagonal line (Δ_x), and $d_\infty(x,x') \le \epsilon$. Then, by Fact 18(ii), one observes

$$d_{triv}^{(M,N_{\to\epsilon})}(x) \le d_{triv}^{(M,N_{\to\epsilon})}(x') + \epsilon \le d_{triv}^{(M,N)}(x) + 2\epsilon < d + 2\epsilon.$$

Therefore, Q^ϵ is $(d+2\epsilon)_{(M,N_{\to\epsilon})}$-trivializable. ◀

▶ **Theorem 30.** *$d_I(M,N) \in S$.*

Proof. Suppose that $d = d_I(M,N) \notin S$. Let d^* be the largest value in S satisfying $d^* \le d$. Note that $d \in S$ if and only if $d = d^*$. Then, $d^* < d$ by our assumption that $d \notin S$.

By definition of interleaving distance, we have $\forall d' > d$, there is a d'-interleaving between M and N, and $\forall d'' < d$, there is no d''-interleaving between M and N. By Proposition 25(ii), one can see that $\delta^* \le d^* < d$. So, to get a contradiction, we just need to show that there exists d'', $d^* < d'' < d$, satisfying the condition 2 in Theorem 23.

Let Q be any intersection component of $M, N_{\to d}$ or $N, M_{\to d}$. Without loss of generality, assume Q is an intersection component of M and $N_{\to d}$. By Proposition 27, Q is stable. We claim that there exists some $\epsilon > 0$ such that $Q^{-\epsilon}$ is an intersection component of M and $N_{\to d-\epsilon}$ in a stable neighborhood of Q, and $Q^{-\epsilon}$ is either $(M, N_{\to d-\epsilon})$-valid or $(d-\epsilon)_{(M,N_{\to d-\epsilon})}$-trivializable.

Let $\epsilon > 0$ be small enough so that $Q^{+\epsilon}$ is a stable intersection component of M and $N_{\to d+\epsilon}$ in a stable neighborhood of Q. By Theorem 23, $Q^{+\epsilon}$ is either $(M, N_{\to(d+\epsilon)})$-valid or $(d+\epsilon)_{(M,N_{\to(d+\epsilon)})}$-trivializable. If $Q^{+\epsilon}$ is $(M, N_{\to(d+\epsilon)})$-valid, then by Corollary 29(i), any intersection component in a stable neighborhood of Q is valid, which means there exists $Q^{-\epsilon}$ that is $(M, N_{\to d-\epsilon})$-valid for some $\epsilon > 0$. Now assume $Q^{+\epsilon}$ is not $(M, N_{\to(d+\epsilon)})$-valid. Then, $\forall \epsilon > 0$, $Q^{+\epsilon}$ is $(M, N_{\to(d+\epsilon)})$-trivializable, By Proposition 21 and 29(ii), we have $\forall x \in V(I_Q)$, $d_{triv}^{(M,N_{\to d+\epsilon})}(x) < d + 3\epsilon$, $\forall \epsilon > 0$. Taking $\epsilon \to 0$, we get $\forall x \in V(I_Q)$, $d_{triv}^{(M,N_{\to d})}(x) \le d$. We claim that, actually, $\forall x \in V(I_Q)$, $d_{triv}^{(M,N_{\to d})}(x) < d$. If the claim were not true, some point $x \in V(I_Q)$ would exist so that $d_{triv}^{(M,N_{\to d})}(x) = d$. There are two cases. If $x \in V(I_M) \cup V(I_N)$, then obviously $d = d_{triv}^{(M,N_{\to d})}(x) \in S$ contradicting $d \neq d^*$. The other case is that x is the intersection point of two perpendicular edges $e_1 \in E(I_M)$ and $e_2 \in E(I_N)$ since Q is a stable intersection component. But, then x and $\pi_L(x)$ are always on two parallel edges where L is either $U(I_M)$ or $L(I_N)$. By Proposition 41(ii) in [18], we have $d = d^*$, reaching a contradiction. Now by our claim and Proposition 21, Q is $d_{(M,N_{\to d})}$-trivializable where $d > d^* \ge \max_{x\in V(I_Q)}\{d_{triv}^{(M,N_{\to d})}(x)\}$. Let $\delta = d - d^*$ and $\epsilon = \delta/4$. Since $d - \epsilon = d - \delta/4 > d - \delta/2 = d - \delta + 2\cdot\delta/4 = d^* + 2\epsilon$ and $d^* \ge \max_{x\in V(I_Q)}\{d_{triv}^{(M,N_{\to d})}(x)\}$, we have $d > d^*$ and $d - \epsilon > \max_{x\in V(I_Q)}\{d_{triv}^{(M,N_{\to d})}(x)\} + 2\epsilon$. Therefore, by Corollary 21, $Q^{-\epsilon}$ is $(d-\epsilon)_{(M,N_{\to d-\epsilon})}$-trivializable.

The above argument shows that there exists a d''-interleaving where $d'' = d - \epsilon < d$, reaching a contradiction. ◀

4 A lower bound on d_I

In this section we propose a distance between two persistence modules that bounds the interleaving distance from below. This distance is defined for n-D modules and not necessarily only for 2-D modules. It is based on dimensions of the vectors involved with the two modules and is efficiently computable.

Let $[n] = \{1, 2, \ldots, n\}$ be the set of all the integers from 1 to n. Let $\binom{[n]}{k} = \{s \subseteq [n] : |s| = k\}$ be the set of all subset in $[n]$ with cardinality k.

▶ **Definition 31.** For a right continuous function $f : \mathbb{R}^n \to \mathbb{Z}$, define the *differential* of f to be $\Delta f : \mathbb{R}^n \to \mathbb{Z}$ where

$$\Delta f(x) = \sum_{k=0}^{n} (-1)^k \cdot \sum_{s \in \binom{[n]}{k}} \lim_{\epsilon \to 0_+} f(x - \epsilon \cdot \sum_{i \in s} e_i)$$

Note that for $k = 0$, $\sum_{s \in \binom{[n]}{k}} \lim_{\epsilon \to 0_+} f(x - \epsilon \cdot \sum_{i \in s} e_i) = f(x)$. We say f is *nice* if the support $supp(\Delta f)$ is finite and $supp(f) \subseteq \{x \,|\, x \geq \vec{a}\}$ for some $a \in \mathbb{R}$.

The differential Δf is a function recording the change of function values of f at each point, especially at 'jump points'. For $n = 1$, $\Delta f(x) = f(x) - \lim_{\epsilon \to 0_+} f(x - \epsilon)$. For $n = 2$, which is the case we deal with, we have

$$\Delta f(x) = f(x) - \lim_{\epsilon \to 0_+} f(x - (\epsilon, 0)) - \lim_{\epsilon \to 0_+} f(x - (0, \epsilon)) + \lim_{\epsilon \to 0_+} f(x - (\epsilon, \epsilon)).$$

See Figure 2 and 3 for illustrations in 1- and 2-D cases respectively.

▶ **Proposition 32.** For a nice function f, $f(x) = \sum_{y \leq x} \Delta f(y)$.

For a proof see the full version [18].

We also define $\Delta f_+ = \max\{\Delta f, 0\}$, $\Delta f_- = \min\{\Delta f, 0\}$ and $f_{\Sigma+}(x) = \sum_{y \leq x} \Delta f_+(y)$, $f_{\Sigma-}(x) = \sum_{y \leq x} \Delta f_-(y)$. Note that $f_{\Sigma+} \geq 0$, $f_{\Sigma-} \leq 0$, and are both monotonic functions. By definition and property of Δf, we have $f = f_{\Sigma+} + f_{\Sigma-}$.

▶ **Definition 33.** For any $\delta > 0$, we define the δ-extension of f as $f^{+\delta} = f_{\Sigma+}(x+\delta) + f_{\Sigma-}(x-\delta)$. Similarly we define the δ-shrinking of f as $f^{-\delta} = f_{\Sigma-}(x + \delta) + f_{\Sigma+}(x - \delta)$ (see Figure 2).

Proposition 34 below follows from the definition.

▶ **Proposition 34.** For any $\delta > 0 \in \mathbb{R}$, we have $f^{\pm\delta}(x) = f(x \mp \delta) + \sum_{y \leq x \pm \delta, y \not\leq x \mp \delta} \Delta f_{\pm}(y)$.

That is to say, for any $\delta \in \mathbb{R}$, the extended (shrunk) function f^{δ}, can be computed by adding to $f(x - |\delta|)$ the positive (negative) difference values of Δf in $(x - |\delta|, x + |\delta|]$. From this, it follows:

▶ **Corollary 35.** *Given* $0 \leq \delta \leq \delta' \in \mathbb{R}$, *we have* $f^{+\delta} \leq f^{+\delta'}$ *and* $f^{-\delta} \geq f^{-\delta'}$.

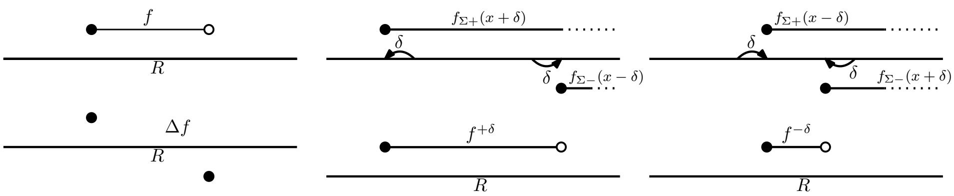


Figure 2 A nice function and its differential (left), its δ-extension (middle), δ-shrinking (right).

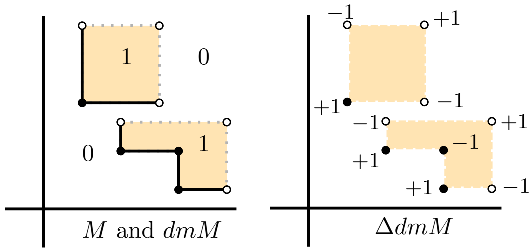


Figure 3 Dimension function (left), its differential is non-zero only at vertices (right).

▶ **Definition 36.** For any two nice functions $f, g : \mathbb{R}^n \to \mathbb{Z}$ and $\delta \geq 0$, we say f, g are within δ-extension, denoted as $f_{\leftarrow\delta\rightarrow}g$, if $f \leq g^{+\delta}$ and $g \leq f^{+\delta}$. Similarly, we say f, g are within δ-shrinking, denoted as $f_{\rightarrow\delta\leftarrow}g$, if $f \geq g^{-\delta}$ and $g \geq f^{-\delta}$.

Let d_+, d_-, d_0 be defined as follows on the space of all nice real-valued functions on $\mathbb{R}^n$:

$$d_-(f,g) = \inf_\delta \{\delta \mid f_{\rightarrow\delta\leftarrow}g\}, \quad d_+(f,g) = \inf_\delta \{\delta \mid f_{\leftarrow\delta\rightarrow}g\}, \quad d_0(f,g) = \min(d_-, d_+)$$

One can verify that d_0 is indeed a distance function. Also, note that when $f, g \geq 0$ (for example, f, g are dimension functions as defined below), we have $d_- \leq d_+$, hence $d_0 = d_-$. It seems that the definition of d_- has a similar connotation as the erosion distance defined by Patel [25] in 1-D case.

4.1 Dimension distance

Given a persistence module M, let the *dimension function* $\mathrm{dm}M : \mathbb{R}^n \to Z$ be defined as $\mathrm{dm}M(x) = dim(M_x)$. The distance $d_0(dmM, dmN)$ for two modules M and N is called the *dimension distance*. Our main result in theorem 38 is that this distance is stable with respect to the interleaving distance and thus provides a lower bound for it.

▶ **Definition 37.** A persistence module M is nice if there exists a value $\epsilon_0 \in \mathbb{R}^+$ so that for every $\epsilon < \epsilon_0$, each linear map $\rho^M_{x\rightarrow x+\vec{\epsilon}} : M_x \to M_{x+\vec{\epsilon}}$ is either injective or surjective (or both).

For example, a finitely presented persistence module generated by a simplicial filtration defined on a grid with at most one additional simplex being introduced between two adjacent grid points satisfies this nice condition above.

▶ **Theorem 38.** *For nice persistence modules M and N, $d_0(\mathrm{dm}M, \mathrm{dm}N) \leq d_I(M, N)$.*

Proof. Let $d_I(M, N) = \delta$. There exists δ-interleaving, $\phi = \{\phi_x\}, \psi = \{\psi_x\}$ which satisfy both triangular and square commutativity. We claim $(\mathrm{dm}M)^{-\delta} \leq \mathrm{dm}N$ and $(\mathrm{dm}N)^{-\delta} \leq \mathrm{dm}M$.

Let $x \in \mathbb{R}^n$ be any point. By Proposition 34, we know that $(\mathrm{dm}M)^{-\delta}(x) = \mathrm{dm}M(x - \delta) + \sum_{y \leq x+\delta, y \nleq x-\delta}(\Delta\mathrm{dm}N_-)(y)$. If $\mathrm{dm}M(x-\delta) \leq \mathrm{dm}N(x)$, then we get $(\mathrm{dm}M)^{-\delta}(x) \leq \mathrm{dm}M(x-\delta) \leq \mathrm{dm}N(x)$, because $\sum_{y \leq x+\delta, y \nleq x-\delta}(\Delta\mathrm{dm}N_-)(y) \leq 0$.

Now assume $\mathrm{dm}M(x-\delta) > \mathrm{dm}N(x)$. From triangular commutativity, we have $rank(\psi_x \circ \phi_{x-\vec{\delta}}) = rank(\rho^M_{x-\vec{\delta}\rightarrow x+\vec{\delta}})$, which gives $\dim(im(\rho^M_{x-\vec{\delta}\rightarrow x+\vec{\delta}})) \leq \dim(im(\phi_{x-\vec{\delta}})) \leq \mathrm{dm}N(x)$.

There exists a collection of linear maps $\{\rho_i : M_{x_i} \to M_{x_{i+1}}\}_{i=0}^k$ such that $\rho^M_{x-\vec{\delta}\rightarrow x+\vec{\delta}} = \rho_k \circ \rho_{k-1} \circ \ldots \circ \rho_1 \circ \rho_0$ and each ρ_i is either injective or surjective. Let $\mathrm{im}_i = im(\rho_i \circ \ldots \circ \rho_0)$. Note that $\mathrm{im}_k = im(\rho^M_{x-\vec{\delta}\rightarrow x+\vec{\delta}})$. Let $\epsilon_i = \dim(\mathrm{im}_i) - \dim(\mathrm{im}_{i-1})$. Then note that $\epsilon_i = 0$ if ρ_i is injective and $\dim(\mathrm{im}_k) - \dim(M_{x_0}) = \sum_{i=1}^k \epsilon_i$. Since $\dim(\mathrm{im}_k) - \dim(M_{x_0}) < 0$, there exists a collection of ρ_{i_j}'s such that $\epsilon_{i_j} < 0$. This means these ρ_{i_j}'s are non-isomorphic surjective linear maps with $\dim(M_{x_{i_j}}) - \dim(M_{x_{i_j-1}}) < 0$. By definition of $\Delta\mathrm{dm}$, this means that, for each pair (x_{i_j-1}, x_{i_j}), there exists a collection $y_1, y_2, \ldots$ such that $y_l \leq x_{i_j}, y_l \not\leq x_{i_j-1}$ and $\sum_l (\Delta\mathrm{dm}M)_-(y_l) \leq \epsilon_{i_j}$. All these y's also satisfy that $y \leq x + \vec{\delta}, y \not\leq x - \vec{\delta}$. So,

$$\sum_{y \leq x+\delta, y \not\leq x-\delta} (\Delta\mathrm{dm}M_-)(y) \leq \sum_j \epsilon_j = \dim(\mathrm{im}_{\,k}) - \dim(M_{x_0}) \leq \dim(N_x) - \dim(M_{x-\vec{\delta}}),$$

which gives $(\mathrm{dm}M)^{-\delta}(x) \leq \mathrm{dm}N(x)$. Similarly, we can show $(\mathrm{dm}N)^{-\delta}(x) \leq \mathrm{dm}M(x)$. ◀

Notice that for dimension functions which are always non-negative, we have $d_0 = d_-$. It may seem that we could have avoided introducing d_+ altogether. But, since nice functions also include negative valued functions, one can verify that d_+ plays the same role for such functions as d_- does for non-negative ones. Then to make d_0 a distance on the space of all nice functions, one needs to define it as the minimum of both d_+ and d_-. For dimension functions, d_+ is not necessarily bounded above by d_I.

4.2 Computation of d_0

For computational purpose, assume that two input persistence modules M and N are finite in that they are functors on the subcategory $\{1, \ldots, k\}^n \subset \mathbb{R}^n$ and the dimension functions $f := \mathrm{dm}M$, $g := \mathrm{dm}N$ have been given as input on an n-dimensional k-ary grid.

First, for the dimension functions f, g, we compute $\Delta f, \Delta g, \Delta f_\pm, \Delta g_\pm, f_\pm, g_\pm$ in $O(k^2)$ time. By Proposition 34, for any $\delta \in \mathbb{Z}^+$, we can also compute $f^{\pm\delta}$, $g^{\pm\delta}$ in $O(k^2)$ time. Then we can apply the binary search to find the minimal value δ within a bounded region such that f, g are within δ-extension or δ-shrinking. This takes $O(\log k)$ time. So the entire computation takes $O(k^2 \log k)$ time.

5 Conclusions

In this paper, we presented an efficient algorithm to compute the bottleneck distance of two 2-D persistence modules given by indecomposables that may have non-constant complexity. No such algorithm for such case is known. Making the algorithm more efficient will be one of our future goals. Extending the algorithm or its modification to larger classes of modules such as the n-D modules or exact pfd bi-modules considered in [14] will be interesting. The definition of valid and trivializable intersection component and Theorem 21 can be extended easily to n-D modules. So is the algorithm– possibly with sacrificing some of the efficiencies. But, further work is necessary to establish the correctness of the algorithm for this general case.

The assumption of nice modules for dimension distance d_0 is needed so that the dimension function, which is a weaker invariant compared to the rank invariants or barcodes in one dimensional case, provides meaningful information without ambiguity. There are cases where the dimension distance can be larger than interleaving distance if the assumption of nice modules is dropped. Of course, one can adjust the definition of dimension distance to incorporate more information so that it remains bounded from above by the interleaving distance.

References

1 Michael Atiyah. On the krull-schmidt theorem with application to sheaves. *Bulletin de la Société Mathématique de France*, 84:307–317, 1956. URL: http://eudml.org/doc/86907.

2 Ulrich Bauer and Michael Lesnick. Induced matchings of barcodes and the algebraic stability of persistence. In *Proceedings of the Thirtieth Annual Symposium on Computational Geometry*, SOCG'14, pages 355:355–355:364, 2014.

3 Silvia Biasotti, Andrea Cerri, Patrizio Frosini, and Daniela Giorgi. A new algorithm for computing the 2-dimensional matching distance between size functions. *Pattern Recognition Letters*, 32(14):1735–1746, 2011.

4 Håvard Bjerkevik. Stability of higher-dimensional interval decomposable persistence modules. *arXiv preprint arXiv:1609.02086*, 2016.

5 Håvard Bjerkevik and Magnus Botnan. Computational complexity of the interleaving distance. *arXiv preprint arXiv:1712.04281*, 2017.

6 Magnus Botnan, Justin Curry, and Elizabeth Munch. The poset interleaving distance, 2016. URL: https://jointmathematicsmeetings.org/amsmtgs/2180_abstracts/1125-55-1151.pdf.

7 Magnus Botnan and Michael Lesnick. Algebraic stability of zigzag persistence modules. *arXiv preprint arXiv:1604.00655*, 2016.

8 Peter Bubenik and Jonathan Scott. Categorification of persistent homology. *Discrete & Computational Geometry*, 51(3):600–627, 2014.

9 Gunnar Carlsson and Afra Zomorodian. The theory of multidimensional persistence. *Discrete & Computational Geometry*, 42(1):71–93, Jul 2009.

10 Andrea Cerri, Barbara Di Fabio, Massimo Ferri, Patrizio Frosini, and Claudia Landi. Betti numbers in multidimensional persistent homology are stable functions. *Mathematical Methods in the Applied Sciences*, 36(12):1543–1557, 2013.

11 Andrea Cerri and Patrizio Frosini. A new approximation algorithm for the matching distance in multidimensional persistence. Technical report, February 2011. URL: http://amsacta.unibo.it/2971.

12 Frédéric Chazal, David Cohen-Steiner, Marc Glisse, Leonidas Guibas, and Steve Oudot. Proximity of persistence modules and their diagrams. In *Proceedings of the Twenty-fifth Annual Symposium on Computational Geometry*, SCG '09, pages 237–246, 2009.

13 Frédéric Chazal, Vin de Silva, Marc Glisse, and Steve Oudot. The structure and stability of persistence modules. *arXiv preprint arXiv:1207.3674*, 2012.

14 Jérémy Cochoy and Steve Oudot. Decomposition of exact pfd persistence bimodules. *arXiv preprint arXiv:1605.09726*, 2016.

15 David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. *Discrete & Computational Geometry*, 37(1):103–120, 2007.

16 William Crawley-Boevey. Decomposition of pointwise finite-dimensional persistence modules. *Journal of Algebra and Its Applications*, 14(05):1550066, 2015. doi:10.1142/S0219498815500668.

17 Vin de Silva, Elizabeth Munch, and Amit Patel. Categorified reeb graphs. *Discrete & Computational Geometry*, 55(4):854–906, Jun 2016.

18 Tamal Dey and Cheng Xin. Computing bottleneck distance for 2-d interval decomposable modules, 2018. URL: http://arxiv.org/abs/1803.02869.

19 Herbert Edelsbrunner and John Harer. *Computational Topology: An Introduction*. Applied Mathematics. American Mathematical Society, 2010.

20 Michael Kerber, Dmitriy Morozov, and Arnur Nigmetov. Geometry helps to compare persistence diagrams. *Journal of Experimental Algorithmics (JEA)*, 22(1):1–4, 2017.

21 Claudia Landi. The rank invariant stability via interleavings. *arXiv preprint arXiv:1412.3374*, 2014.

22 Michael Lesnick. The theory of the interleaving distance on multidimensional persistence modules. *Foundations of Computational Mathematics*, 15(3):613–650, 2015.

23 Klaus Lux and Magdolna Sźoke. Computing decompositions of modules over finite-dimensional algebras. *Experimental Mathematics*, 16(1):1–6, 2007.

24 Steve Oudot. *Persistence theory: from quiver representations to data analysis*, volume 209. American Mathematical Society, 2015.

25 Amit Patel. Generalized persistence diagrams. *arXiv preprint arXiv:1601.03107*, 2016.

26 Carry Webb. Decomposition of graded modules. *Proc. American Math. Soc.*, 94(4):565–571, 1985.