



On the Stability of Interval Decomposable Persistence Modules

Håvard Bakke Bjerkevik^{1,2}

Received: 7 December 2018 / Revised: 22 January 2021 / Accepted: 12 March 2021 /

Published online: 11 May 2021

© The Author(s) 2021

Abstract

The algebraic stability theorem for persistence modules is a central result in the theory of stability for persistent homology. We introduce a new proof technique which we use to prove a stability theorem for n -dimensional rectangle decomposable persistence modules up to a constant $2n - 1$ that generalizes the algebraic stability theorem, and give an example showing that the bound cannot be improved for $n = 2$. We then apply the technique to prove stability for block decomposable modules, from which novel results for zigzag modules and Reeb graphs follow. These results are improvements on weaker bounds in previous work, and the bounds we obtain are optimal.

Keywords Persistent homology · Stability · Multiparameter persistence · Reeb graphs · Zigzag modules

Mathematics Subject Classification 55N31

1 Introduction

The main goal of this paper is to study stability of persistence modules and the consequences it has for other common objects studied in topological data analysis (TDA). For us, a persistence module is a functor $M : P \rightarrow \mathbf{vec}$, where P is a poset category, often $P = \mathbb{R}^n$, and \mathbf{vec} is the category of finite dimensional vector spaces over a fixed field k . We might say ‘ P -module’ to emphasize the poset, or just ‘module’. Explicitly, this is a set of finite dimensional vector spaces M_p over a fixed field k and linear transformations $M_p \rightarrow M_q$ for $p \leq q \in P$.

Editor in Charge: János Pach

Håvard Bakke Bjerkevik
bjerkevik@tugraz.at

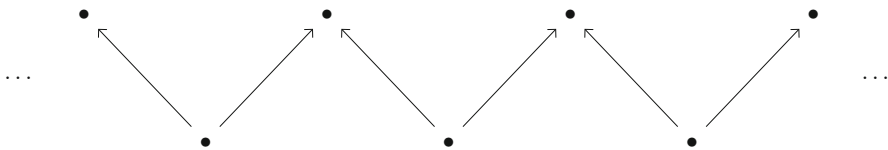
¹ NTNU, Trondheim, Norway

² Present Address: TU Graz, Graz, Austria

One can measure the distance between persistence modules by the *interleaving distance* d_I . Under certain assumptions, persistence modules can be decomposed into direct sums of *interval modules*¹, which are particularly simple and well behaved. For $n = 1$, all persistence modules decompose into interval modules [19], and interval modules are in bijection with the intervals over \mathbb{R} , so a decomposition into interval modules gives an appealing presentation of a module M as a multiset $\mathcal{B}(M)$ of intervals called a *barcode*. The *bottleneck distance* d_B is a distance on barcodes which can be defined combinatorially, as opposed to the algebraic definition of the interleaving distance.

The *algebraic stability theorem* (AST) is a fundamental theorem in TDA stating that $d_I(M, N) = d_B(\mathcal{B}(M), \mathcal{B}(N))$, or $d_I = d_B$ for short, holds for \mathbb{R} -modules M and N [3,15,16,18]. In TDA, data often take the form of a function $f: X \rightarrow \mathbb{R}$, where X is a topological space. For instance, a data set might be a finite set of points $S \subset \mathbb{R}^n$, in which case one can define $f(x)$ as the distance from x to the closest point in S . Let M_f be the \mathbb{R} -module given by $M_f(a) = H_i(f^{-1}((-\infty, a]))$. It is easy to show that $d_I(M_f, M_g) \leq \|f - g\|_\infty$ for $f, g: X \rightarrow \mathbb{R}$. This means that a consequence of the algebraic stability theorem is that small perturbations of a function lead to small changes in the resulting barcode, justifying calling $d_I = d_B$ a stability theorem.

Other commonly studied objects in TDA are *zigzag modules*, *interlevel set persistence modules*, and *Reeb graphs*. Zigzag modules are $\mathbb{Z}\mathbb{Z}$ -modules, where $\mathbb{Z}\mathbb{Z}$ is isomorphic to the infinite poset



These modules have been of interest to the TDA community for more than a decade [5,11,13], and have been used to study flocking behaviour of animals [23]. Zigzag modules are closely related to interlevel set persistence, which is the study of how the topology of $f^{-1}([a, b])$ changes with a and b , where $f: X \rightarrow \mathbb{R}$, as above. This connection is explained in Sect. 3.

Given a function $f: X \rightarrow \mathbb{R}$, the Reeb graph $\bar{f}: (X/\sim) \rightarrow \mathbb{R}$ is defined by letting $x \sim y$ for x and y in the same connected component of $f^{-1}(\{a\})$ for some $a \in \mathbb{R}$. Thus, the Reeb graph keeps track of how the connected components of the level set $f^{-1}(\{a\})$ merge, split, and are born and die as a changes. Introduced by Georges Reeb [26], it has been applied to visualization and graphics [22,24,27,30], and its approximation Mapper [29] has proven very useful; for instance, it has been used to identify a new subgroup of breast cancers [25] and study biomolecular folding pathways [31].

One can define d_I and d_B also for zigzag modules and Reeb graphs, and whether an analogue of AST holds here is a highly relevant question which has been explored in previous work. The best known bounds until now were $d_B \leq 5d_I/2$ for zigzag modules

¹ Our definition of an interval is different from the definition in order theory as a set of the form $\{p \in P \mid a \leq p \leq b\}$ for some $a, b \in P$. If $P = \mathbb{R}^d$, this will instead be a special case of a *rectangle* for us.

and $d_B \leq 5d_I$ for Reeb graphs, where the latter is an improvement on the constant 9 from [4]. These bounds were proven by Botnan and Lesnick [11] through the result $d_B \leq 5d_I/2$ for *block decomposable modules*, which are a class of interval decomposable \mathbb{R}^2 -modules. In Sect. 3, we explain how stability of zigzag modules, interlevel set persistence, and Reeb graphs follows from stability of block decomposable modules.

\mathbb{R}^n -modules with $n \geq 2$, often called multiparameter modules, are important in their own right. Adding a density parameter is helpful when studying data sets with outliers and tendrils [12], and sometimes data are naturally equipped with several parameters. Unfortunately, \mathbb{R}^n -modules do not have a nice invariant like the barcode of an \mathbb{R} -module [14].

We introduce a new method for proving stability that is versatile and simple compared to earlier methods, and that in some cases gives stronger results. We first apply the method in Sect. 4, where we prove Theorem 4.3, which implies

Theorem *Let M and N be rectangle decomposable \mathbb{R}^n -modules. Then we have $d_B(\mathcal{B}(M), \mathcal{B}(N)) \leq (2n - 1)d_I(M, N)$.*

These are modules that decompose into intervals of the form $I_1 \times \cdots \times I_n$, where I_i is an interval over \mathbb{R} for each i . This is a generalization of the algebraic stability theorem for \mathbb{R} -modules, which is the case $n = 1$, while for $n \geq 2$, the result is new. Our proof is more combinatorial in nature than the previous proofs for AST. In Sect. 5, we give an example showing that the bound is optimal for $n = 2$, which also disproves a conjecture made in an earlier version of [11] claiming that $d_B(M, N) = d_I(M, N)$ holds for all n -dimensional interval decomposable modules M and N whose barcodes only contain convex intervals. It is an open question if the bound can be improved for $n \geq 3$.

An arbitrary \mathbb{R}^n -module coming from a data set will usually not be rectangle (or even interval) decomposable, though rectangle decomposable modules do arise from interlevel set persistence in the case $n = 2$ [5,13]. Still, the interleaving distance is certainly of interest, and our work reveals some phenomena that can help our understanding of that subject. Indeed, the ideas in this paper were important ingredients in the recent proof of the NP-hardness of computing d_I [7], which we discuss in Sect. 6. Using the same method, we then prove Theorem 4.12, which gives

Theorem *Let M and N be free modules. Then $d_B(\mathcal{B}(M), \mathcal{B}(N)) \leq (n - 1)d_I(M, N)$.*

Free modules are a special case of rectangle decomposable modules. In practice, multiparameter modules are often given as the cokernel of a morphism $F_1 \rightarrow F_2$ of free modules.

Applying elements from the previous theorems and still using the same method, we show Theorem 4.18, which gives

Theorem *Let M and N be block decomposable modules. Then $d_B(\mathcal{B}(M), \mathcal{B}(N)) = d_I(M, N)$.*

This improves on the mentioned best known bound of $d_B \leq 5d_I/2$ ($d_B \geq d_I$ is trivial), and our proof is simpler and less technical. It follows that $d_B = d_I$ holds for zigzag modules and $d_B \leq 2d_I$ holds for Reeb graphs, which is optimal. These are also improvements on the best known bounds. The precise statements are given in Sect. 3.

There we also show how this theorem is used in the proof of Theorem 3.7 by Cochoy and Oudot [17], which is a generalization of a known stability result for interlevel set persistence.

2 Persistence Modules, Interleavings, and Matchings

In this section we introduce notation and definitions that we will use throughout the paper. Let k be a field that stays fixed throughout the text, and let \mathbf{vec} be the category of finite dimensional vector spaces over k . We identify a poset with its poset category, which has the elements of the poset as objects, a single morphism $p \rightarrow q$ if $p \leq q$ and no morphism if $p \not\leq q$.

Definition 2.1 Let P be a poset category. A P -persistence module is a functor $P \rightarrow \mathbf{vec}$.

If the choice of poset is obvious from the context, we usually write ‘persistence module’ or just ‘module’ instead of ‘ P -persistence module’.

For a persistence module M , $p \leq q \in P$, $M(p)$ is denoted by M_p and $M(p \rightarrow q)$ by $\phi_M(p, q)$. We refer to the morphisms $\phi_M(p, q)$ as the *internal morphisms* of M . M being a functor implies $\phi_M(p, p) = \text{id}_{M_p}$ and $\phi_M(q, r) \circ \phi_M(p, q) = \phi_M(p, r)$. Because persistence modules are defined as functors, they automatically assemble into a category where the morphisms are natural transformations. This category is denoted by $P\text{-mod}$. Let $f: M \rightarrow N$ be a morphism between persistence modules. Such an f consists of a morphism associated to each $p \in P$, and these morphisms are denoted by f_p . Because f is a natural transformation, we have $\phi_N(p, q) \circ f_p = f_q \circ \phi_M(p, q)$ for all $p \leq q$.

Definition 2.2 An *interval* is a non-empty subset $I \subset P$ that satisfies the following:

- If $p, q \in I$ and $p \leq r \leq q$, then $r \in I$.
- If $p, q \in I$, then there exist $p_1, p_2, \dots, p_{2m} \in I$ for some $m \in \mathbb{N}$ such that $p \leq p_1 \geq p_2 \leq \dots \geq p_{2m} \leq q$.

We refer to the last point as the *connectivity axiom* for intervals.

Definition 2.3 An *interval persistence module* or *interval module* is a persistence module M that satisfies the following: for some interval I , $M_p = k$ for $p \in I$ and $M_p = 0$ otherwise, and $\phi_M(p, q) = \text{id}_k$ for points $p \leq q$ in I . We use the notation \mathbb{I}^I for the interval module with I as its underlying interval.

The definitions up to this point have been valid for all posets P , but we need some additional structure on P to get a notion of distance between persistence modules, which is essential to prove stability results. Since we will mostly be working with \mathbb{R}^n -persistence modules, we restrict ourselves to this case from now on. We define the poset structure on \mathbb{R}^n by letting $(a_1, a_2, \dots, a_n) \leq (b_1, b_2, \dots, b_n)$ if and only if $a_i \leq b_i$ for $1 \leq i \leq n$. For $\epsilon \in \mathbb{R}$, we often abuse notation and write ϵ when we mean $(\epsilon, \epsilon, \dots, \epsilon) \in \mathbb{R}^n$. We call an interval $I \subset \mathbb{R}^n$ *bounded* if it is contained in a ball with finite radius.

Definition 2.4 For $\epsilon \in [0, \infty)$, we define the *shift functor* $(\cdot)(\epsilon): \mathbb{R}^n\text{-mod} \rightarrow \mathbb{R}^n\text{-mod}$ by letting $M(\epsilon)$ be the persistence module with $M(\epsilon)_p = M_{p+\epsilon}$ and $\phi_{M(\epsilon)}(p, q) = \phi_M(p + \epsilon, q + \epsilon)$. For morphisms $f: M \rightarrow N$, we define $f(\epsilon): M(\epsilon) \rightarrow N(\epsilon)$ by $f(\epsilon)_p = f_{p+\epsilon}$.

We also define shift on intervals I by letting $I(\epsilon)$ be the interval for which $\mathbb{I}^{I(\epsilon)} = \mathbb{I}^I(\epsilon)$. Define the morphism $\phi_{M,\epsilon}: M \rightarrow M(\epsilon)$ by $(\phi_{M,\epsilon})_p = \phi_M(p, p + \epsilon)$.

Definition 2.5 An ϵ -interleaving between \mathbb{R}^n -modules M and N is a pair of morphisms $f: M \rightarrow N(\epsilon), g: N \rightarrow M(\epsilon)$ such that $g(\epsilon) \circ f = \phi_{M,2\epsilon}$ and $f(\epsilon) \circ g = \phi_{N,2\epsilon}$.

If there exists an ϵ -interleaving between M and N , then M and N are said to be ϵ -interleaved. An interleaving can be viewed as an ‘approximate isomorphism’, and a 0-interleaving is in fact a pair of isomorphisms. We call a module M ϵ -significant if $\phi_M(p, p + \epsilon) \neq 0$ for some p , and ϵ -trivial otherwise. M is 2ϵ -trivial if and only if it is ϵ -interleaved with the zero module. We call an interval I ϵ -significant if \mathbb{I}^I is ϵ -significant, and ϵ -trivial otherwise.

Definition 2.6 We define the *interleaving distance* d_1 on persistence modules M and N by

$$d_1(M, N) = \inf \{ \epsilon \mid M \text{ and } N \text{ are } \epsilon \text{-interleaved} \}.$$

Intuitively, the interleaving distance measures how close the modules are to being isomorphic. The interleaving distance between two modules might be infinite, and the interleaving distance between two different, even non-isomorphic modules, might be zero. Apart from this, d_1 satisfies the axioms for a metric, so d_1 is an extended pseudometric.

Definition 2.7 Suppose $M \simeq \bigoplus_{I \in B} \mathbb{I}^I$ for a multiset B of intervals. Then we call B the *barcode* of M , and write $B = \mathcal{B}(M)$. We say that M is *interval decomposable*.

When we write $I \neq I'$ for intervals I and I' in a barcode, we mean that they are separate elements of the multiset, not that they necessarily represent different subsets of \mathbb{R}^n .

For \mathbb{R} -modules, barcodes and *persistence diagrams* are equivalent and both commonly used; the only difference is that they refer to two different ways of presenting the intervals: either as multisets of intervals, or as multisets of points in \mathbb{R}^2 . For higher-dimensional intervals, however, persistence diagrams become impossible to draw, while one can still draw pictures analogous to the one-dimensional barcodes. For this reason we will use the term ‘barcode’.

Since the endomorphism ring of any interval module is isomorphic to k , it follows from [2, Thm. 1] that for any poset P , if a P -module M is interval decomposable, the decomposition is unique up to isomorphism. Thus, the barcode is well defined. Any \mathbb{R} -module is interval decomposable [19], but this is not true for \mathbb{R}^n -modules in general. The following is an example of a P -module for a poset P with four points that is not interval decomposable. A corresponding \mathbb{R}^2 -module that is not interval decomposable and is at most two-dimensional at each point can be constructed.

$$\begin{array}{ccccc}
 k & \xrightarrow{(1 \ 0)} & k^2 & \xrightarrow{\begin{pmatrix} 1 \\ 1 \end{pmatrix}} & k \\
 & & \uparrow & & \\
 & & (0 \ 1) & & \\
 & & k & &
 \end{array}$$

For multisets A, B , we define a *partial bijection* as a bijection $\sigma: A' \rightarrow B'$ for some subsets $A' \subset A$ and $B' \subset B$, and we write $\sigma: A \dashrightarrow B$. We write $\text{coim } \sigma = A'$ and $\text{im } \sigma = B'$.

Definition 2.8 Let A and B be multisets of intervals. An ϵ -*matching* between A and B is a partial bijection $\sigma: A \dashrightarrow B$ such that

- all $I \in A \setminus \text{coim } \sigma$ are 2ϵ -trivial,
- all $J \in B \setminus \text{im } \sigma$ are 2ϵ -trivial,
- for all $I \in \text{coim } \sigma$, \mathbb{I}^I and $\mathbb{I}^{\sigma(I)}$ are ϵ -interleaved.

We can interpret ϵ -matchings in the context of graph theory. A *matching in a graph* is a set of edges in the graph without common vertices, and a matching is said to *cover* a set S of vertices if all elements in S are adjacent to an edge in the matching. Let G_ϵ be the bipartite graph on $A \sqcup B$ with an edge between $I \in A$ and $J \in B$ if \mathbb{I}^I and \mathbb{I}^J are ϵ -interleaved. Then an ϵ -matching between A and B is a matching in G_ϵ such that the set of 2ϵ -significant intervals in $A \sqcup B$ is covered.

Definition 2.9 The *bottleneck distance* d_B is defined by

$$d_B(C, D) = \inf \{ \epsilon \mid \text{there is an } \epsilon\text{-matching between } C \text{ and } D \}$$

for multisets C and D of intervals.

We will often abuse notation and write $d_B(M, N)$ meaning $d_B(\mathcal{B}(M), \mathcal{B}(N))$ for persistence modules M and N . The inequality $d_I(M, N) \leq d_B(M, N)$ always holds when M and N are interval decomposable, as one can easily construct an ϵ -interleaving between M and N given an ϵ -matching between $\mathcal{B}(M)$ and $\mathcal{B}(N)$. For this reason, we will treat the statements $d_B \leq d_I$ and $d_B = d_I$ as the same.

3 Zigzag Modules, Interlevel Set Persistence, and Reeb Graphs

In this section we define block modules and explain how they relate to the stability of zigzag modules, interlevel set persistence, and Reeb graphs. In particular, we explain the consequences of our stability result for block decomposable modules.

In the following, \mathbb{R}^{op} is the poset with \mathbb{R} as the underlying set and the opposite order of the standard one. Sending $a \in \mathbb{R}$ to $-a$ gives an isomorphism between \mathbb{R}^{op} and \mathbb{R} . \mathbb{Z}^{op} is defined similarly. Let $\mathbb{U} = \{(a, b) \in \mathbb{R}^{\text{op}} \times \mathbb{R} \mid a \leq b\}$. This is a subposet of \mathbb{R}^2 with a flipped x -axis, but it is also useful to think of it as the poset of closed intervals (over \mathbb{R}) with $I \leq J$ if $I \subset J$. There is an obvious one-to-one

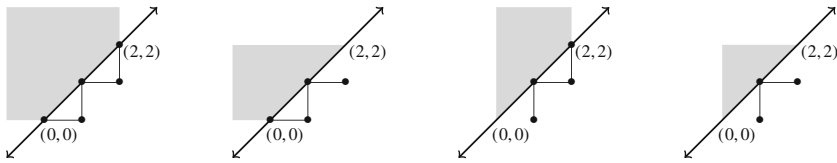


Fig. 1 From left to right: the intervals $[0, 2]_{BL}$, $[0, 2)_{BL}$, $(0, 2]_{BL}$, and $(0, 2)_{BL}$ together with the $\mathbb{Z}\mathbb{Z}$ -intervals $[0, 2]$, $[0, 2)$, $(0, 2]$, and $(0, 2)$, respectively

correspondence between \mathbb{U} -modules and $(\mathbb{R}^{op} \times \mathbb{R})$ -modules that are zero at all points (a, b) with $b < a$. Since $\mathbb{R}^{op} \times \mathbb{R} \simeq \mathbb{R}^2$, we can define interleavings on \mathbb{U} -modules via the usual definition on \mathbb{R}^2 -modules. Explicitly, the shift functor on \mathbb{U} -modules is given by $M(\epsilon)_{(a,b)} = M_{(a-\epsilon, b+\epsilon)}$, from which the definition of ϵ -interleaving follows.

We are interested in the following particular subset of \mathbb{U} -modules.

Definition 3.1 A *block* is an interval of one of the following forms, where $a, b \in \mathbb{R} \cup \{\pm\infty\}$:

$$[a, b]_{BL} = \{(c, d) \in \mathbb{U} \mid c \leq b, d \geq a\}, \quad [a, b)_{BL} = \{(c, d) \in \mathbb{U} \mid a \leq d < b\},$$

$$(a, b]_{BL} = \{(c, d) \in \mathbb{U} \mid a < c \leq b\}, \quad (a, b)_{BL} = \{(c, d) \in \mathbb{U} \mid c > a, d < b\}.$$

A *block module* is an interval module over \mathbb{U} supported on a block. A \mathbb{U} -module is *block decomposable* if it is isomorphic to a direct sum of block modules.

In all but the first case, $a \leq b$ is necessary for the block to be nonempty. Blocks of the form $[a, b]_{BL}$ with $a > b$ arise from interlevel set persistence, but not from zigzag modules or Reeb graphs by the constructions we will describe.

In Sect. 4.3, we prove Theorem 4.18:

Theorem *Let M and N be block decomposable modules. If M and N are δ -interleaved, there exists a δ -matching between $\mathcal{B}(M)$ and $\mathcal{B}(N)$. Thus, $d_1(M, N) = d_B(M, N)$.*

The previous best known bound was $d_1(M, N) \leq 5d_B(M, N)/2$, proved in [11].

In this section, we will explain how stability results for zigzag modules, interlevel set persistence, and Reeb graphs follow from this result. In the case of zigzag modules and Reeb graphs, our result strengthens the previously known bounds to where they cannot be improved. For interlevel set persistence, Theorem 4.18 was an ingredient in the proof in [17] of a generalization of a known stability result.

3.1 Zigzag Modules

A zigzag module is a module over the poset $\mathbb{Z}\mathbb{Z} = \{(a, b) \in \mathbb{Z}^{op} \times \mathbb{Z} \mid a = b \text{ or } a = b + 1\}$. Like in ordinary 1-parameter persistence, zigzag modules are interval decomposable and therefore have well-defined barcodes [9]. It is not hard to see that any interval over $\mathbb{Z}\mathbb{Z}$ must be of one of the following forms, where $a, b \in \mathbb{Z} \cup \{\pm\infty\}$.

$$[a, b] = \{(c, d) \in \mathbb{Z}\mathbb{Z} \mid c \leq b, d \geq a\}, \quad [a, b) = \{(c, d) \in \mathbb{Z}\mathbb{Z} \mid a \leq d < b\},$$

$$(a, b] = \{(c, d) \in \mathbb{Z}\mathbb{Z} \mid a < c \leq b\}, \quad (a, b) = \{(c, d) \in \mathbb{Z}\mathbb{Z} \mid c > a, d < b\}.$$

For $(a, b) \in \mathbb{U}$, let $\mathbb{Z}\mathbb{Z}|_{(a,b)}$ be the sub-poset of $\mathbb{Z}\mathbb{Z}$ containing the elements $\{(c, d) \in \mathbb{Z}\mathbb{Z} \mid c \geq a, d \leq b\}$. A zigzag module M gives rise to a \mathbb{U} -module $U(M)$ by letting

$$U(M)_{(a,b)} = \text{colim } (\mathbb{Z}\mathbb{Z}|_{(a,b)}),$$

and defining the internal morphism $\phi_{U(M)}((a, b), (c, d))$ as the induced morphism we get by the universal property of colimits. For any multiset B of intervals over $\mathbb{Z}\mathbb{Z}$, let $U(B)$ be the multiset of blocks where each interval I in B has been replaced by the block I_{BL} with the same multiplicity. That is, $[a, b]$ is replaced by $[a, b]_{BL}$, $[a, b)$ by $[a, b)_{BL}$ and so on. Figure 1 illustrates how intervals over $\mathbb{Z}\mathbb{Z}$ give rise to blocks. One can show the following theorem by observing that $U(\cdot)$ preserves direct sums and that $U(\mathbb{I}^I) = \mathbb{I}^{I_{BL}}$ for all $\mathbb{Z}\mathbb{Z}$ -intervals I .

Theorem 3.2 *If M is a zigzag module, then $U(M)$ is block decomposable and $\mathcal{B}(U(M)) = U(\mathcal{B}(M))$.*

Thus, we can define the interleaving and bottleneck distances between zigzag modules by letting $d_I(M, N) = d_I(U(M), U(N))$ and $d_B(M, N) = d_B(U(M), U(N))$. Then Theorem 4.18 immediately implies stability for zigzag modules:

Theorem 3.3 *Let M and N be zigzag modules. If $U(M)$ and $U(N)$ are δ -interleaved, there is a δ -matching between their barcodes. Thus, $d_I(M, N) = d_B(M, N)$.*

Like for block decomposable modules, this is an improvement on the previous bound of $d_B(M, N) \leq 5d_I(M, N)/2$ [11].

3.2 Interlevel Set Persistence

Given a topological space X and a continuous function $f: X \rightarrow \mathbb{R}$, we define a \mathbb{U} -module M_n^f by $M_n^f(a, b) = H_n(f^{-1}([a, b]))$, where the internal morphisms are induced by inclusions. When f is sufficiently nice, we can study the stability of the interlevel set persistence of f via M_n^f . Let $\mathbb{X}_I = f^{-1}(I)$ for subsets $I \subset \mathbb{R}$.

Definition 3.4 We say that (X, f) is of *Morse type* if the homology of $\mathbb{X}_{\{t\}}$ is finitely generated for every $t \in \mathbb{R}$ and there are real numbers $a_1 < a_2 < \dots < a_m$ such that for all $I = (a_i, a_{i+1})$, there is a Y and a homeomorphism $h: Y \times I \rightarrow \mathbb{X}_I$ such that $f \circ h$ is the projection onto I , and h extends to a continuous function $\bar{h}: Y \times \bar{I} \rightarrow \mathbb{X}_{\bar{I}}$. Here, \bar{I} is the closure of I , i ranges from 0 to m , and we let $a_0 = -\infty$ and $a_{m+1} = \infty$.

The numbers a_i are to be understood as the critical values of f . Picking values $-\infty < s_0 < a_1 < s_1 < \dots < s_{m-1} < a_m < s_m < \infty$ and defining $\mathbb{X}_i^j = f^{-1}([s_i, s_j])$, Carlsson et al. [13] consider the sequence

$$\begin{aligned} H_n(\mathbb{X}) : H_n(\mathbb{X}_0^0) &\rightarrow H_n(\mathbb{X}_0^1) \leftarrow H_n(\mathbb{X}_1^1) \rightarrow \dots \\ &\leftarrow H_n(\mathbb{X}_{m-1}^{m-1}) \rightarrow H_n(\mathbb{X}_{m-1}^m) \leftarrow H_n(\mathbb{X}_m^m). \end{aligned}$$

We will view this as a persistence module living over the poset

$$S = \{(s_i, s_j) \mid (i = j \text{ or } i = j - 1) \text{ and } 0 \leq i, j \leq m\} \subset \mathbb{R}^{\text{op}} \times \mathbb{R}.$$

S is isomorphic to a (finite, connected) subposet P of $\mathbb{Z}\mathbb{Z}$ via an isomorphism ι . Thus, we can consider a zigzag module Z supported on P such that $H_n(\mathbb{X}) = Z \circ \iota$, decompose Z into interval modules, and pull the decomposition back along ι to get an interval decomposition of $H_n(\mathbb{X})$. This means that $H_n(\mathbb{X})$ is interval decomposable, so the barcode $\mathcal{B}(H_n(\mathbb{X}))$ is well defined. We borrow the following notation from Carlsson et al., which covers all the intervals that can appear in $\mathcal{B}(H_n(\mathbb{X}))$:

$$\begin{aligned} [a_i, a_j] &= \{(s_c, s_d) \mid c < j, d \geq i\}, & [a_i, a_j] &= \{(s_c, s_d) \mid i \leq d < j\}, \\ (a_i, a_j] &= \{(s_c, s_d) \mid i \leq c < j\}, & (a_i, a_j) &= \{(s_c, s_d) \mid c \geq i, d < j\}. \end{aligned}$$

By definition, $H_n(\mathbb{X})$ is simply the restriction of M_n^f to the subposet S . Carlsson et al. also consider a larger diagram which consists of the vector spaces $H_n(\mathbb{X}_i^j)$ for $0 \leq i \leq j \leq m$ with maps $H_n(\mathbb{X}_i^j) \rightarrow H_n(\mathbb{X}_k^l)$ induced by the inclusions $\mathbb{X}_i^j \hookrightarrow \mathbb{X}_k^l$ whenever $k \leq i$ and $j \leq l$. This is in fact a discrete version of M_n^f , and carries exactly the same information as long as we remember the a_i : Let M be the \mathbb{U} -module given by $M_{(a,b)} = H_n(\mathbb{X}_i^j)$ for $a_i < a \leq a_{i+1}$ and $a_j \leq b < a_{j+1}$, and internal morphisms equal to the relevant morphisms $H_n(\mathbb{X}_i^j) \rightarrow H_n(\mathbb{X}_k^l)$. Then $M \cong M_n^f$.

Using what they call the Mayer–Vietoris Diamond Principle, Carlsson et al. were able to prove part (i) of the following theorem, determining M_n^f (or rather the discrete version just mentioned) using only $H_n(\mathbb{X})$ and $H_{n-1}(\mathbb{X})$. The principle is very similar to how a zigzag module M gives rise to a block decomposable module $U(M)$, though here we also get blocks of the form $[a, b]_{\text{BL}}$ with $a > b$. These blocks are not detected by $H_n(\mathbb{X})$, but arise from $H_{n-1}(\mathbb{X})$ through a dimension shift. Cochoy and Oudot [17] proved part (ii) with help from a result by Botnan and Crawley-Boevey [10], generalizing (i) by relaxing the assumptions on f . A function $f: X \rightarrow \mathbb{R}$ is *pf*d (pointwise finite dimensional) if it is continuous (but not necessarily of Morse type) and $H_n(f^{-1}((a, b)))$ is finite dimensional for all a, b, n .

Theorem 3.5 [10,13,17]

(i) *If f is of Morse type, then M_n^f is block decomposable and its barcode consists of the following:*

- a block $[a_i, a_j]_{\text{BL}}$ for each $[a_i, a_j]$ in $\mathcal{B}(H_n(\mathbb{X}))$,
- a block $[a_j, a_i]_{\text{BL}}$ for each (a_i, a_j) in $\mathcal{B}(H_{n-1}(\mathbb{X}))$,
- a block $[a_i, a_j]_{\text{BL}}$ for each $[a_i, a_j]$ in $\mathcal{B}(H_n(\mathbb{X}))$,
- a block $(a_i, a_j]_{\text{BL}}$ for each (a_i, a_j) in $\mathcal{B}(H_n(\mathbb{X}))$,
- a block $(a_i, a_j)_{\text{BL}}$ for each (a_i, a_j) in $\mathcal{B}(H_n(\mathbb{X}))$.

(ii) *If f is pf*d, *there is a block decomposable module M such that $d_1(M, M_n^f) = 0$.*

The reason for the indirect phrasing of part (ii) is that Cochoy and Oudot work with inverse images of open intervals instead of closed, which gives a module that is not

isomorphic to M_n^f , though the interleaving distance from it is 0. This is not a significant problem, however, as we can define $\mathcal{B}(M_n^f)$ as $\mathcal{B}(M)$ for M as in (ii).

If M is a block decomposable module, we will find it convenient to use the notation $\mathcal{B}_{\text{diag}}(M)$ for the barcode we get by intersecting the intervals in $\mathcal{B}(M)$ with the diagonal and projecting this to \mathbb{R} . We call $\mathcal{B}_{\text{diag}}(M_n^f)$ the n^{th} level set barcode. This has the following stability property:

Theorem 3.6 [13] *Let $f, g: X \rightarrow \mathbb{R}$ be of Morse type. Then*

$$d_B(\mathcal{B}_{\text{diag}}(M_n^f), \mathcal{B}_{\text{diag}}(M_n^g)) \leq \|f - g\|_\infty,$$

where $\|f - g\|_\infty = \sup_{x \in X} |f(x) - g(x)|$.

Theorem 3.5 allows us to ask if there is an analogue of level set stability for interlevel set persistence by considering block decomposable modules. Botnan and Lesnick give an affirmative answer for f of Morse type. The result was generalized to pfd functions in [17], applying a result of [10] and our Theorem 4.18:

Theorem 3.7 [17] *Let $f, g: X \rightarrow \mathbb{R}$ be pfd. Then*

$$d_B(M_n^f, M_n^g) \leq \|f - g\|_\infty.$$

Proof If $\epsilon > \|f - g\|_\infty$, we have inclusions

$$f^{-1}([a, b]) \hookrightarrow g^{-1}([a - \epsilon, b + \epsilon]) \hookrightarrow f^{-1}([a - 2\epsilon, b + 2\epsilon])$$

and vice versa. These induce an ϵ -interleaving between M_n^f and M_n^g , so $d_1(M_n^f, M_n^g) \leq \|f - g\|_\infty$. By results from [10,17], M_n^f and M_n^g are block decomposable, and by Theorem 4.18, $d_B(M_n^f, M_n^g) = d_1(M_n^f, M_n^g)$. □

3.3 Reeb Graphs

Given a continuous function $f: X \rightarrow \mathbb{R}$ of Morse type with locally path-connected level sets $f^{-1}(x)$, let R_f be the quotient space X/\sim , where $x \sim y$ if x and y are in the same connected component of $f^{-1}(r)$ for some $r \in \mathbb{R}$. We get a function $\tilde{f}: R_f \rightarrow \mathbb{R}$ induced by f . The *Reeb graph* of (X, f) is the pair (R_f, \tilde{f}) . Alternatively, one can define a Reeb graph directly as a pair (R, f) such that R is a finite topological graph (a compact triangulable space of dimension at most 1), $f: R \rightarrow \mathbb{R}$ is continuous and of Morse type, and the level sets are discrete. By this definition, the Reeb graph of (X, f) as above is indeed a Reeb graph. We sometimes abuse notation and refer to R as a Reeb graph without explicitly mentioning f .

De Silva et al. [28] show that an equivalent way of viewing a Reeb graph (R, f) is as a functor $C^R: \mathbb{U} \rightarrow \mathbf{Set}$ given by $C^R(a, b) = \pi_0(f^{-1}([a, b]))$ and morphisms $C^R(a, b) \rightarrow C^R(c, d)$ induced by inclusions of sets, which in their language is a cosheaf. Strictly speaking, they work with open intervals instead of closed, but this does not matter for the distances we will discuss. We think of C^R as a set-valued persistence

module over \mathbb{U} . Shift, interleaving, and the interleaving distance are defined in exactly the same way as for persistence modules with values in \mathbf{vec} . Using these cosheaves, de Silva et al. are also able to describe interleavings in a geometrically intuitive way as level-preserving maps between Reeb graphs. We will not need this description, so we skip the details and invite the interested reader to study [28].

Though the interleaving distance is well defined, we cannot talk about decompositions into direct sums of interval modules when we are working over \mathbf{Set} . To avoid this problem, let $C_{\mathbf{vec}}^R$ be the \mathbb{U} -module (with values in \mathbf{vec}) such that $C_{\mathbf{vec}}^R(a, b)$ has $C^R(a, b)$ as a basis, and the internal morphisms are determined by the basis maps given by C^R . The level sets are finite since $H_0(X)$ is finitely generated, so the vector spaces are finite dimensional. $C_{\mathbf{vec}}^R$ is simply the zeroth homology interlevel set persistence module of (R, f) , so as above, it decomposes into block modules; see Fig. 2. Applying Theorem 4.18 and observing that an ϵ -interleaving between C^R and $C^{R'}$ representing Reeb graphs (R, f) and (R', f') induces an ϵ -interleaving between $C_{\mathbf{vec}}^R$ and $C_{\mathbf{vec}}^{R'}$, we get

$$d_B(C_{\mathbf{vec}}^R, C_{\mathbf{vec}}^{R'}) = d_I(C_{\mathbf{vec}}^R, C_{\mathbf{vec}}^{R'}) \leq d_I(C^R, C^{R'}).$$

We would like to have a similar statement for level set barcodes, though we have to be a little careful, as for $2\epsilon < b - a < 4\epsilon$, $\mathbb{I}^{(a,b)_{BL}}$ is ϵ -interleaved with the zero module, while the \mathbb{R} -module $\mathbb{I}^{(a,b)}$ is not. This causes $d_B(\mathcal{B}(C_{\mathbf{vec}}^R), \mathcal{B}(C_{\mathbf{vec}}^{R'}))$ and $d_B(\mathcal{B}_{\text{diag}}(C_{\mathbf{vec}}^R), \mathcal{B}_{\text{diag}}(C_{\mathbf{vec}}^{R'}))$ to disagree by a factor of up to two. Still, it is straightforward to check that an ϵ -matching between $\mathcal{B}(C_{\mathbf{vec}}^R)$ and $\mathcal{B}(C_{\mathbf{vec}}^{R'})$ induces a 2ϵ -matching between $\mathcal{B}_{\text{diag}}(C_{\mathbf{vec}}^R)$ and $\mathcal{B}_{\text{diag}}(C_{\mathbf{vec}}^{R'})$. Thus,

$$d_B(\mathcal{B}_{\text{diag}}(C_{\mathbf{vec}}^R), \mathcal{B}_{\text{diag}}(C_{\mathbf{vec}}^{R'})) \leq 2d_B(C_{\mathbf{vec}}^R, C_{\mathbf{vec}}^{R'}).$$

Summing up the preceding paragraphs, we get

Theorem 3.8 *Let (R, f) and (R', f') be Reeb graphs. Then the following inequalities hold:*

$$\begin{aligned} d_B(C_{\mathbf{vec}}^R, C_{\mathbf{vec}}^{R'}) &\leq d_I(C^R, C^{R'}), \\ d_B(\mathcal{B}_{\text{diag}}(C_{\mathbf{vec}}^R), \mathcal{B}_{\text{diag}}(C_{\mathbf{vec}}^{R'})) &\leq 2d_I(C^R, C^{R'}). \end{aligned}$$

The previous best known bounds were $5/2$ for the first inequality and 5 for the second [11]. De Silva et al. observe that if $f, f': X \rightarrow \mathbb{R}$ give rise to Reeb graphs R and R' , then $d_I(C^R, C^{R'}) \leq \|f - f'\|_\infty$ follows almost immediately from the definitions. This justifies calling Theorem 3.8 a stability result in the same way as for the classic algebraic stability theorem.

The following simple example shows that neither inequality in the theorem can be improved:

Example 3.9 Let R be the line segment from $(0, -2)$ to $(0, 2)$ in \mathbb{R}^2 , R' the circle of radius 2 centered at the origin, and f and f' the projections to the second coordinate.

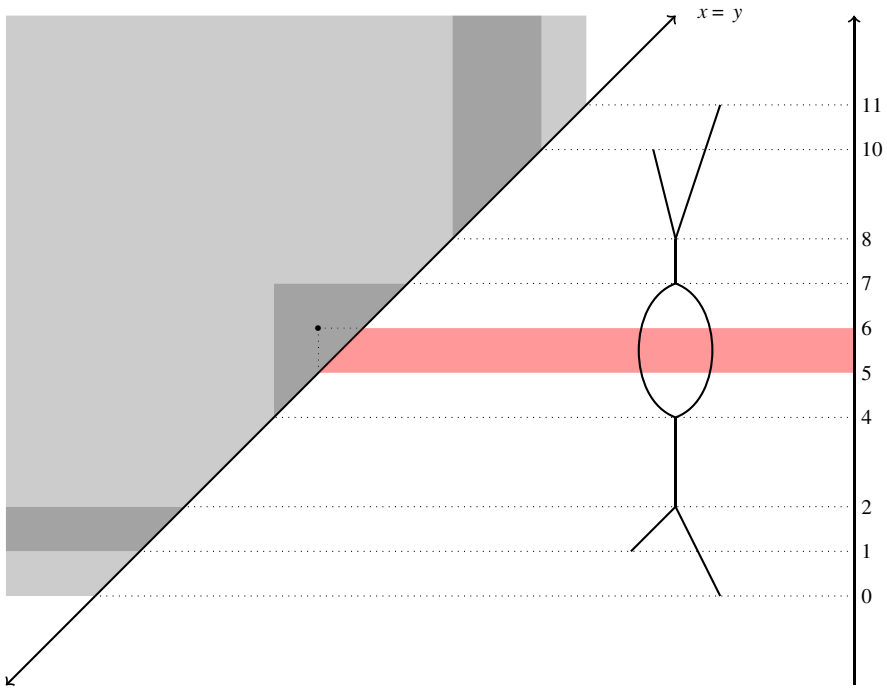


Fig. 2 A Reeb graph (R, f) with f illustrated as a height function on the right, C^R on the left, darker shades of grey meaning larger cardinality. The intersection of R and the red strip is $f^{-1}([5, 6])$, which has two connected components. Thus, $|C^R_{(5,6)}| = 2$, shown by the dot. We can easily read off $\mathcal{B}(C^R_{\text{vec}}) = \{[0, 11]_{\text{BL}}, [1, 2]_{\text{BL}}, (4, 7)_{\text{BL}}, (8, 10]_{\text{BL}}\}$ from the figure, which also gives us the level set barcode $\mathcal{B}_{\text{diag}}(C^R_{\text{vec}}) = \{[0, 11], [1, 2], (4, 7), (8, 10]\}$

Then

$$d_1(C^R, C^{R'}) = 1, \quad d_B(C^R_{\text{vec}}, C^{R'}_{\text{vec}}) = 1, \quad d_B(\mathcal{B}_{\text{diag}}(C^R_{\text{vec}}), \mathcal{B}_{\text{diag}}(C^{R'}_{\text{vec}})) = 2.$$

We leave the verification of the equalities to the reader.

4 Higher-Dimensional Stability

The algebraic stability theorem states that an ϵ -interleaving between \mathbb{R} -modules M and N induces an ϵ -matching between $\mathcal{B}(M)$ and $\mathcal{B}(N)$, implying $d_1(M, N) = d_B(M, N)$. The purpose of this section is to prove similar results for \mathbb{R}^n -modules. Our first result, Theorem 4.3, is a generalization of AST. There already exist several proofs of AST [3,15,16,18], but our approach has ideas that are applicable to more than just \mathbb{R} -modules.

Theorem 4.3 is the most technically challenging application of our proof method, and the proof demonstrates very well exactly when our method works and when it fails. The lesson to take home is that the method gives a bound $d_B \leq cd_1$ with some

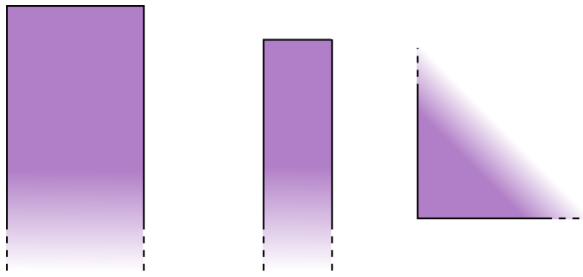


Fig. 3 Three rectangles, where the left and middle rectangles are of the same type (unbounded downwards), while the last is of a different type (unbounded upwards and to the right). Assuming that it contains its boundary, the rightmost rectangle is also an example of a free interval, which we will define in Sect. 4.2

c that increases with the freedom we have in defining the intervals we consider. One needs $2n$ coordinates to define an n -dimensional (hyper)rectangle, which gives the constant $2n - 1$ in the theorem.

4.1 Rectangle Decomposable Modules

For any interval $I \subset \mathbb{R}^n$, let I_i denote its projection onto the i^{th} coordinate.

Definition 4.1 A *rectangle* is an interval of the form $R = R_1 \times R_2 \times \cdots \times R_n$.

Two rectangles R and S are of the *same type* if $R_i \setminus S_i$ and $S_i \setminus R_i$ are bounded for every i . For $n = 1$, we have four types of rectangles:

- intervals of finite length,
- intervals of the form (a, ∞) or $[a, \infty)$,
- intervals of the form $(-\infty, a)$ or $(-\infty, a]$,
- $(-\infty, \infty)$,

for some $a \in \mathbb{R}$. We see that for $n \geq 1$, rectangles R and S are of the same type if R_i and S_i are of the same type for all $1 \leq i \leq n$. Examples of 2-dimensional rectangles are given in Fig. 3.

In [16], *decorated numbers* were introduced. These are endpoints of intervals ‘decorated’ with a plus or minus sign depending on whether the endpoints are included in the interval or not. Let $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$. A decorated number is of the form a^+ or a^- , where $a \in \overline{\mathbb{R}}$. (The decorated numbers $-\infty^-$ and ∞^+ are never used, as no interval contains points at infinity, but it does not matter whether we include these two points in the definition.) The notation is as follows for $a, b \in \overline{\mathbb{R}}$:

- $I = (a^+, b^+)$ if $I = (a, b)$,
- $I = (a^+, b^-)$ if $I = (a, b)$,
- $I = (a^-, b^+)$ if $I = [a, b)$,
- $I = (a^-, b^-)$ if $I = [a, b)$.

We define decorated points in n dimensions for $n \geq 1$ as tuples $a = (a_1, a_2, \dots, a_n)$, where all the a_i ’s are decorated numbers. For an n -dimensional rectangle R and decorated points (a_1, a_2, \dots, a_n) and (b_1, b_2, \dots, b_n) , we write $R = ((a_1, a_2, \dots, a_n),$

(b_1, b_2, \dots, b_n)) if $R_i = (a_i, b_i)$ for all i . We define \min_R and \max_R as the decorated points for which $R = (\min_R, \max_R)$. We write a^* for decorated numbers with unspecified ‘decoration’, so a^* is either a^+ or a^- .

There is a total order on the decorated numbers given by $a^* < b^*$ for $a < b$, and $a^- < a^+$ for all $a, b \in \mathbb{R}$. This induces a poset structure on decorated n -dimensional points given by $(a_1, a_2, \dots, a_n) \leq (b_1, b_2, \dots, b_n)$ if $a_i \leq b_i$ for all i . We can also add decorated numbers and real numbers by letting $a^+ + x = (a + x)^+$ and $a^- + x = (a + x)^-$ for $a \in \mathbb{R}, x \in \mathbb{R}$. We add n -dimensional decorated points and n -tuples of real numbers coordinatewise.

Definition 4.2 Let M be an \mathbb{R}^n -module. If R is a rectangle and $M \simeq \mathbb{I}^R$, M is a *rectangle module*. If M is interval decomposable and all $I \in \mathcal{B}(M)$ are rectangles, M is *rectangle decomposable*.

Our goal is to prove the following theorem:

Theorem 4.3 Let $M = \bigoplus_{I \in \mathcal{B}(M)} \mathbb{I}^I$ and $N = \bigoplus_{J \in \mathcal{B}(N)} \mathbb{I}^J$ be rectangle decomposable \mathbb{R}^n -modules. If M and N are δ -interleaved, there exists a $(2n - 1)\delta$ -matching between $\mathcal{B}(M)$ and $\mathcal{B}(N)$. Thus, $d_B(M, N) \leq (2n - 1)d_1(M, N)$.

We prove the theorem by a mix of combinatorial, geometric, and algebraic arguments. First we show that it is enough to prove the theorem under the assumption that all the rectangles in $\mathcal{B}(M)$ and $\mathcal{B}(N)$ are of the same type. Then we define a real-valued function α on the set of rectangles which in a sense measures, in the case $n = 2$, how far ‘up and to the right’ a rectangle is. There is a preorder \leq_α associated to α . The idea behind \leq_α is that if there is a nonzero morphism $\chi: \mathbb{I}^I \rightarrow \mathbb{I}^J(\epsilon)$ and $I \leq_\alpha J$, then I and J have to be close to each other. Finding pairs of intervals in $\mathcal{B}(M)$ and $\mathcal{B}(N)$ that are close is exactly what we need to construct a $(2n - 1)\delta$ -matching. Lemmas 4.7 and 4.8 say that such morphisms behave nicely in a precise sense that we will exploit when we prove Lemma 4.9. If we remove the conditions mentioning \leq_α , Lemmas 4.7 and 4.8 are not even close to being true, so one of the main points in the proof of Lemma 4.9 is that we must exclude the cases that are not covered by Lemmas 4.7 and 4.8. We do this by proving that a certain matrix is upper triangular, where the ‘bad cases’ correspond to the elements above the diagonal and the ‘good cases’ correspond to elements on and below the diagonal.

Lemma 4.9 is what ties together the geometric and combinatorial parts of the proof of Theorem 4.3. While we prove Lemma 4.9 by geometric arguments, by Hall’s marriage theorem the lemma is equivalent to a statement about matchings between $\mathcal{B}(M)$ and $\mathcal{B}(N)$. We have to do some combinatorics to get exactly the statement we need, namely that there is a $(2n - 1)\delta$ -matching between $\mathcal{B}(M)$ and $\mathcal{B}(N)$, and we do this after stating Lemma 4.9.

Fix $0 \leq \delta \in \mathbb{R}$. Assume that M and N are δ -interleaved, with interleaving morphisms $f: M \rightarrow N(\delta)$ and $g: N \rightarrow M(\delta)$. Recall that this means that $g(\delta) \circ f = \phi_{M,2\delta}$ and $f(\delta) \circ g = \phi_{N,2\delta}$. For any $I \in \mathcal{B}(M)$, we have a canonical injection $\mathbb{I}^I \xrightarrow{\iota_I} M$ and projection $M \xrightarrow{\pi_I} \mathbb{I}^I$, and likewise, we have canonical morphisms $\mathbb{I}^J \xrightarrow{\iota_J} N$ and $N \xrightarrow{\pi_J} \mathbb{I}^J$ for $J \in \mathcal{B}(N)$. We define

$$f_{I,J} = \pi_J(\delta) \circ f \circ \iota_I: \mathbb{I}^I \rightarrow \mathbb{I}^J(\delta),$$

$$g_{J,I} = \pi_I(\delta) \circ g \circ \iota_J: \mathbb{I}^J \rightarrow \mathbb{I}^I(\delta).$$

Lemma 4.4 *Let $\chi: \mathbb{I}^I \rightarrow \mathbb{I}^J$ be a morphism between interval modules. Suppose $A = I \cap J$ is an interval. Then, for all $a, b \in A$, $\chi_a = \chi_b$ as k -endomorphisms.*

Proof Suppose $a \leq b$ and $a, b \in A$. Then $\chi_b \circ \phi_{\mathbb{I}^I}(a, b) = \phi_{\mathbb{I}^J}(a, b) \circ \chi_a$. Since the ϕ -morphisms are identities, we get $\chi_a = \chi_b$ as k -endomorphisms. By the connectivity axiom for intervals, the equality extends to all elements in A . \square

Since the intersection of two rectangles is either empty or a rectangle, we can describe a morphism between two rectangle modules uniquely as a k -endomorphism if their underlying rectangles intersect. A k -endomorphism, in turn, is simply multiplication by a constant.

Lemma 4.5 *Let R and S be rectangles. Then there exists a nonzero morphism $\chi: \mathbb{I}^R \rightarrow \mathbb{I}^S$ if and only if $\min_S \leq \min_R$, $\max_S \leq \max_R$, and $R \cap S \neq \emptyset$.*

Proof Suppose $\chi_a \neq 0$, which immediately gives $a \in R \cap S$, and let $a \geq x \in R$ and $a \leq y \in S$. We have the following commutative diagram.

$$\begin{array}{ccccc}
 k & \xrightarrow{\text{id}} & k & \xrightarrow{\phi_{\mathbb{I}^R}(a,y)} & \mathbb{I}_y^R \\
 \downarrow \chi_x & & \downarrow \chi_a \neq 0 & & \downarrow \chi_y \\
 \mathbb{I}_x^S & \xrightarrow{\phi_{\mathbb{I}^S}(x,a)} & k & \xrightarrow{\text{id}} & k
 \end{array}$$

\mathbb{I}_x^S is not zero, as there is a nonzero morphism factoring through it. Thus, $\min_S \leq x$ and $\min_S \leq \min_R$. An analogous argument using the square on the right-hand side gives $\max_R \geq \max_S$.

To prove the converse, define χ by $\chi_a = \text{id}$ for $a \in R \cap S$ and $\chi_a = 0$ otherwise. This is well defined unless there are $a \leq b$ such that $a \in R \setminus S$ and $b \in R \cap S$, or $a \in R \cap S$ and $b \in S \setminus R$, which are impossible by the assumptions $\min_S \leq \min_R$ and $\max_S \leq \max_R$, respectively. \square

This will come in handy when we prove Lemmas 4.6, 4.7, and 4.8.

We define a function $w: (\mathcal{B}(M) \times \mathcal{B}(N)) \sqcup (\mathcal{B}(N) \times \mathcal{B}(M)) \rightarrow k$ by letting $w(I, J) = x$ if $f_{I,J}$ is given by multiplication by x , and $w(I, J) = 0$ if $f_{I,J}$ is the zero morphism. $w(J, I)$ is given by $g_{J,I}$ in the same way. With the definition of w , it is starting to become clear how combinatorics comes into the picture. We can now construct a bipartite weighted directed graph on $\mathcal{B}(M) \sqcup \mathcal{B}(N)$ by letting $w(I, J)$ be the weight of the edge from I to J . The reader is encouraged to keep this picture in mind, as a lot of what we do in the rest of the proof can be interpreted as statements about the structure of this graph.

The following lemma allows us to break up the problem and focus on the components of M and N with the same types separately.

Lemma 4.6 *Let R and T be rectangles of the same type, and S be a rectangle of a different type. Then $\psi\chi = 0$ for any pair $\chi: \mathbb{I}^R \rightarrow \mathbb{I}^S$, $\psi: \mathbb{I}^S \rightarrow \mathbb{I}^T$ of morphisms.*

Proof Suppose $\psi, \chi \neq 0$. By Lemma 4.5, $\min_R \geq \min_S \geq \min_T$ and $\max_R \geq \max_S \geq \max_T$. We get $\min_{R_i} \geq \min_{S_i} \geq \min_{T_i}$ and $\max_{R_i} \geq \max_{S_i} \geq \max_{T_i}$ for all i , and it follows that if R and T are of the same type, then S is of the same type as R and T . \square

Let $f': M \rightarrow N(\delta)$ be defined by $f'_{I,J} = f_{I,J}$ for $I \in \mathcal{B}(M)$ and $J \in \mathcal{B}(N)$ if I and J are of the same type, and $f'_{I,J} = 0$ if they are not, and let $g': N \rightarrow M(\delta)$ be defined analogously. Here f' and g' are assembled from $f'_{I,J}$ and $g'_{J,I}$ the same way f and g are from $f_{I,J}$ and $g_{J,I}$. Suppose $I, I' \in \mathcal{B}(M)$. Then we have

$$\sum_{J \in \mathcal{B}(N)} g_{J,I'}(\delta) f_{I,J} = \sum_{J \in \mathcal{B}(N)} g'_{J,I'}(\delta) f'_{I,J}.$$

Since (f, g) is an interleaving, the left-hand side is equal to the composition $\mathbb{I}^I \xrightarrow{f'} M \xrightarrow{\phi_{M,2\delta}} M \xrightarrow{\pi_{I'}} \mathbb{I}^{I'}$, which is zero if $I \neq I'$. Thus, if I and I' are of different types, the left-hand side is zero, while all the summands on the right-hand side are zero by definition of f' and g' . If I and I' are of the same type, the equality follows from Lemma 4.6. This means that $g'(\delta)f' = g(\delta)f$. We also have $f'(\delta)g' = f(\delta)g$, so f' and g' are δ -interleaving morphisms. In particular, f' and g' are δ -interleaving morphisms when restricted to the components of M and N of a fixed type. If we can show that f' and g' induce a $(2n - 1)\delta$ -matching on each of the mentioned components, we will have proved Theorem 4.3. In other words, we have reduced the problem to the case where all the intervals in $\mathcal{B}(M)$ and $\mathcal{B}(N)$ are of the same type.

For a decorated number a^* , let $u(a^*) = a$ if $a \neq \pm\infty$ and $u(a^*) = 0$ otherwise. Let $a = (a_1, a_2, \dots, a_n)$ be a decorated point. We define $P(a)$ to be the number of the decorated numbers a_i decorated with $+$, and we also define $\alpha(a) = \sum_{1 \leq i \leq n} u(a_i)$. What we really want to look at are rectangles and not decorated points by themselves, so we define $P(R) = P(\min_R) + P(\max_R)$ and $\alpha(R) = \alpha(\min_R) + \alpha(\max_R)$ for any rectangle R . Let the order \leq_α on rectangles be given by $R \leq_\alpha S$ if either

- $\alpha(R) < \alpha(S)$, or
- $\alpha(R) = \alpha(S)$ and $P(R) \leq P(S)$.

This defines a preorder. In other words, it is transitive ($R \leq_\alpha S \leq_\alpha T$ implies $R \leq_\alpha T$) and reflexive ($R \leq_\alpha R$ for all R). We write $R <_\alpha S$ if $R \leq_\alpha S$ and not $R \geq_\alpha S$.

The order \leq_α is one of the most important ingredients in the proof. The point is that if there is a nonzero morphism from \mathbb{I}^R to $\mathbb{I}^S(\epsilon)$ and $R \leq_\alpha S$, then R and S have to be close to each other. If $\epsilon = 0$, R and S actually have to be equal. This ‘closeness property’ is expressed in Lemma 4.7, and is also exploited in Lemma 4.8. Finally, in the proof of Lemma 4.9, we make sure that we only have to deal with morphisms $g_{J,I'}(\delta) \circ f_{I,J}$ for $I \leq_\alpha I'$ and not $I >_\alpha I'$, so that our lemmas can be applied.

In Fig. 4 we see two rectangles $R = (0, 4) \times (0, 4)$ and $S = (2, 5) \times (2, 5)$. There is no nonzero morphism from \mathbb{I}^R to \mathbb{I}^S or $\mathbb{I}^{S(1)}$, because $\min_R < \min_{S(\epsilon)}$ for all $\epsilon < 2$. This is connected to the fact that $\alpha(R) = 8 < 14 = \alpha(S)$, which can be interpreted to mean that R is ‘further down and to the left’ than S . The point of including $P(R)$ in the definition of α is that e.g. $(a, b]$ is a tiny bit ‘further to the right’ than $[a, b)$, and

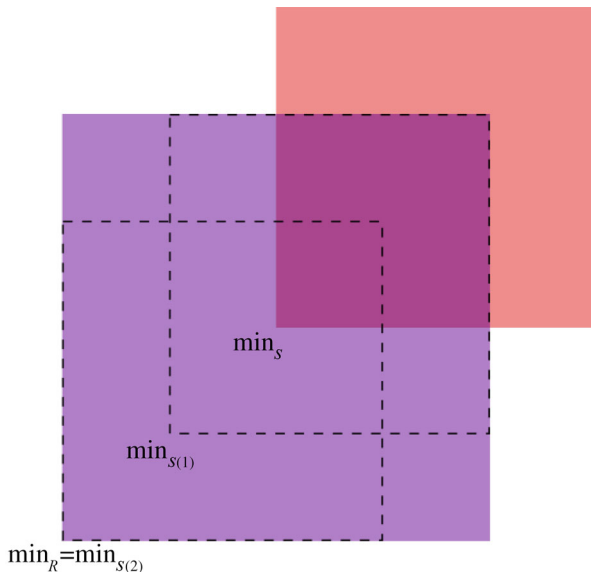


Fig. 4 Rectangles $R = (0, 4) \times (0, 4)$ (purple), $S = (2, 5) \times (2, 5)$ (pink), $S(1) = (1, 4) \times (1, 4)$ (dotted border), and $S(2) = (0, 3) \times (0, 3)$ (dotted border)

this is a subtlety that P recognizes, and that matters in the proofs of Lemmas 4.7 and 4.8.

Lemma 4.7 *Let R, S , and T be rectangles of the same type with $R \leq_\alpha T$. Suppose there are nonzero morphisms $\chi: \mathbb{I}^R \rightarrow \mathbb{I}^S(\epsilon)$ and $\psi: \mathbb{I}^S \rightarrow \mathbb{I}^T(\epsilon)$. Then \mathbb{I}^S is $(2n - 1)\epsilon$ -interleaved with either \mathbb{I}^R or \mathbb{I}^T .*

Proof We must have either $R \leq_\alpha S$ or $S \leq_\alpha T$. Let us assume the former; the latter is the same. Since $\chi \neq 0$, it follows from Lemma 4.5 that

- (i) $\min_S \leq \min_R + \epsilon$,
- (ii) $\max_S \leq \max_R + \epsilon$.

We can assume that at least one of R and S is $(4n - 2)\epsilon$ -significant, since if not, they would be $(2n - 1)\epsilon$ -interleaved by zero morphisms. Suppose also

- (iii) $\min_R \leq \min_S + (2n - 1)\epsilon$,
- (iv) $\max_R \leq \max_S + (2n - 1)\epsilon$

hold. Assume R is $(4n - 2)\epsilon$ -significant (replacing R by S changes nothing), so there exists an a such that $a, a + (4n - 2)\epsilon \in R$. Then $a + (2n - 1)\epsilon \in S$, so by Lemma 4.5 and (i)–(iv), there are nonzero morphisms $\mathbb{I}^R \rightarrow \mathbb{I}^S((2n - 1)\epsilon)$ and $\mathbb{I}^S \rightarrow \mathbb{I}^R((2n - 1)\epsilon)$. These can be chosen to be the identity where they are nonzero, and then they are interleaving morphisms.

Therefore, we only need to show (iii) and (iv). It suffices to show $\min_{R_m} \leq \min_{S_m} + (2n - 1)\epsilon$ for an arbitrary m ; the max case follows from the same arguments. Suppose

$u(\min_{R_m}) \geq u(\min_{S_m}) + (2n - 1)\epsilon$. This, (i), and (ii) give

$$\begin{aligned} \alpha(S) &= \sum_{i=1}^n u(\min_{S_i}) + \sum_{i=1}^n u(\max_{S_i}) \\ &\leq u(\min_{R_m}) - (2n - 1)\epsilon + \sum_{i \neq m} (u(\min_{R_i}) + \epsilon) + \sum_{i=1}^n (u(\max_{R_i}) + \epsilon) \\ &= \sum_{i=1}^n u(\min_{R_i}) + \sum_{i=1}^n u(\max_{R_i}) = \alpha(R). \end{aligned}$$

Since $R \leq_\alpha S$, we have equality, which means that

- $u(\min_{S_m}) = u(\min_{R_m}) - (2n - 1)\epsilon$,
- $u(\min_{S_i}) = u(\min_{R_i}) + \epsilon$ for $i \neq m$,
- $u(\max_{S_i}) = u(\max_{R_i}) + \epsilon$ for all i .

But then $P(\min_{S_i}) \leq P(\min_{R_i})$ for $i \neq m$ and $P(\max_{S_i}) \leq P(\max_{R_i})$ for all i by (i) and (ii), so $P(\min_{R_m}) \leq P(\min_{S_m})$ follows from $R \leq_\alpha S$. Thus, $\min_{R_m} \leq \min_{S_m} + (2n - 1)\epsilon$. □

Let $C \geq 4n - 2$,

$$\begin{aligned} R &= [0, C]^n, & S &= [\epsilon, C + \epsilon]^{n-1} \times [-(2n - 1)\epsilon, C + \epsilon], \\ T &= [2\epsilon, C + 2\epsilon]^{n-1} \times [-(4n - 2)\epsilon, C + 2\epsilon]. \end{aligned}$$

These satisfy the hypotheses of Lemma 4.7 and $d_I(\mathbb{I}^R, \mathbb{I}^S) = d_I(\mathbb{I}^S, \mathbb{I}^T) = (2n - 1)\epsilon$. This example shows that the constant $2n - 1$ cannot be improved in Lemma 4.7, and this is the source of the same constant in Theorem 4.3.

Lemma 4.8 *Let R, S , and T be rectangles of the same type with R and T $2n\epsilon$ -significant and $\alpha(R) \leq \alpha(T)$. Suppose there are nonzero morphisms $\chi: \mathbb{I}^R \rightarrow \mathbb{I}^S(\epsilon)$ and $\psi: \mathbb{I}^S \rightarrow \mathbb{I}^T(\epsilon)$. Then $\psi(\epsilon) \circ \chi \neq 0$.*

Proof It suffices to show that for any m , $\max_{T_m} > \min_{R_m} + 2\epsilon$. If this holds, then R and $T(2\epsilon)$ intersect, and $\psi(\epsilon) \circ \chi$ is nonzero in this intersection.

Assume that χ and ψ are nonzero and that R and T are $2n\epsilon$ -significant. For all i , we have

- (i) $\min_{R_i} + 2\epsilon \geq \min_{T_i}$,
- (ii) $\max_{R_i} + 2\epsilon \geq \max_{T_i}$,
- (iii) $\max_{R_i} > \min_{R_i} + 2n\epsilon$,
- (iv) $\max_{T_i} > \min_{T_i} + 2n\epsilon$.

(i) and (ii) follow from $\chi, \psi \neq 0$, and Lemma 4.5. (iii) and (iv) are consequences of R and T being $2n\epsilon$ -significant. Pick $1 \leq m \leq n$. We get

$$2u(\max_{T_m}) \geq 2n\epsilon + u(\min_{T_m}) + u(\max_{T_m})$$

$$\begin{aligned}
 &= 2n\epsilon + \alpha(T) - \sum_{i \neq m} (u(\min_{T_i}) + u(\max_{T_i})) \\
 &\geq 2n\epsilon + \alpha(R) - \sum_{i \neq m} (u(\min_{R_i}) + u(\max_{R_i}) + 4\epsilon) \\
 &\geq 2n\epsilon - (4n - 4)\epsilon + u(\min_{R_m}) + u(\max_{R_m}) \\
 &\geq 2n\epsilon - (4n - 4)\epsilon + 2u(\min_{R_m}) + 2n\epsilon = 2u(\min_{R_m}) + 4\epsilon.
 \end{aligned}$$

Either the inequality $u(\max_{T_m}) \geq u(\min_{R_m}) + 2\epsilon$ is strict, or we get $R_m = T_m = [u(\min_{R_m}), u(\min_{R_m}) + 2\epsilon]$. In both cases we have proven what we wanted. \square

For $I \in \mathcal{B}(M)$, let

$$\mu(I) = \{J \in \mathcal{B}(N) \mid \mathbb{I}^I \text{ and } \mathbb{I}^J \text{ are } (2n - 1)\delta\text{-interleaved}\}.$$

In other words, $\mu(I)$ contains all the intervals that can be matched with I in a $(2n - 1)\delta$ -matching. Let $I \in \mathcal{B}(M)$ be $(4n - 2)\delta$ -significant, and pick $p \in \mathbb{R}^n$ such that $p, p + (4n - 2)\delta \in I$. Then, $p + (2n - 1)\delta \in J$ for every $J \in \mu(I)$. Since M and N are pointwise finite dimensional, this means that $\mu(I)$ is a finite set. For $A \subset \mathcal{B}(M)$, we write $\mu(A) = \bigcup_{I \in A} \mu(I)$.

Lemma 4.9 *Let A be a subset of $\mathcal{B}(M)$ containing no $(4n - 2)\delta$ -trivial elements. Then $|A| \leq |\mu(A)|$.*

Before we prove Lemma 4.9, we show that it implies that there is a $(2n - 1)\delta$ -matching between $\mathcal{B}(M)$ and $\mathcal{B}(N)$, and thus completes the proof of Theorem 4.3.

Let G_μ be the undirected bipartite graph on $\mathcal{B}(M) \sqcup \mathcal{B}(N)$ with an edge between I and J if $J \in \mu(I)$. Observe that G_μ is the same as the graph $G_{(2n-1)\delta}$ we defined when we gave the graph theoretical definition of an ϵ -matching (in this case, $(2n - 1)\delta$ -matching) in Sect. 2. Following that definition, a $(2n - 1)\delta$ -matching is a matching in G_μ that covers the set of all $(4n - 2)\delta$ -significant elements in $\mathcal{B}(M)$ and $\mathcal{B}(N)$.

For a subset S of a graph G , let $A_G(S)$ be the neighborhood of S in G , that is, the set of vertices in G that are adjacent to at least one vertex in S . We now apply Hall’s marriage theorem [21] to bridge the gap between Lemma 4.9 and the statement we want to prove about matchings.

Theorem 4.10 (Hall’s marriage theorem) *Let G be a bipartite graph on $X \sqcup Y$ such that $A_G(\{x\})$ is finite for all $x \in X$. Then the following are equivalent:*

- for all $X' \subset X$, $|X'| \leq |A_G(X')|$,
- there exists a matching in G covering X .

One of the two implications is easy, since if $|X'| > |A_G(X')|$ for some $X' \subset X$, then there is no matching in G covering X' . It is the other implication we will use, namely that the first statement implies the existence of a matching in G covering X .

Letting X be the set of $(4n - 2)\delta$ -significant intervals in $\mathcal{B}(M)$ and Y be $\mathcal{B}(N)$, Hall’s theorem and Lemma 4.9 give us a matching σ in the graph G_μ covering all the $(4n - 2)\delta$ -significant elements in $\mathcal{B}(M)$. By symmetry, we also have a matching τ

in G_μ covering all the $(4n - 2)\delta$ -significant elements in $\mathcal{B}(N)$. Neither of these is necessarily a $(2n - 1)\delta$ -matching, however, as each of them only guarantees that all the $(4n - 2)$ -significant intervals in one of the barcodes are matched. We will use σ and τ to construct a $(2n - 1)\delta$ -matching. This construction is similar to one used to prove the Cantor–Bernstein theorem [1, pp. 110–111].

Let H be the undirected bipartite graph on $\mathcal{B}(M) \sqcup \mathcal{B}(N)$ for which the set of edges is the union of the edges in the matchings σ and τ . Let C be a connected component of H . Suppose the submatching of σ in C does not cover all the $(4n - 2)\delta$ -significant elements of C . Then there is a $(4n - 2)\delta$ -significant $J \in C \cap \mathcal{B}(N)$ that is not matched by σ . If we view σ and τ as partial bijections $\sigma : \mathcal{B}(M) \rightarrow \mathcal{B}(N)$ and $\tau : \mathcal{B}(N) \rightarrow \mathcal{B}(M)$, we can write C as $\{J, \tau(J), \sigma(\tau(J)), \tau(\sigma(\tau(J))), \dots\}$. Either this sequence is infinite, or it is finite, in which case the last element is $(4n - 2)\delta$ -trivial. In either case, we get that the submatching of τ in C covers all $(4n - 2)\delta$ -significant elements in C .

By this argument, there is a $(2n - 1)\delta$ -matching in each connected component of H . We can piece these together to get a $(2n - 1)\delta$ -matching in $\mathcal{B}(M) \sqcup \mathcal{B}(N)$, so Lemma 4.9 completes the proof of Theorem 4.3.

Proof of Lemma 4.9 Suppose A is infinite. Each interval in A contains a rational point, so since M is pointwise finite dimensional, the cardinality of A is at most finite times countably infinite, which is countable. If $|A'| \leq |\mu(A')|$ for all finite $A' \subset A$, $|\mu(A)|$ has no finite upper bound, so it is infinite, which implies $|A| \leq |\mu(A)|$. Thus, it is enough to prove the lemma for finite A , and we will assume this for the rest of the proof.

Because \leq_α is a preorder such that all elements are comparable, we can order $A = \{I_1, I_2, \dots, I_r\}$ so that $I_i \leq_\alpha I_{i'}$ for all $i \leq i'$. Write $\mu(A) = \{J_1, J_2, \dots, J_s\}$. For $I \in \mathcal{B}(M)$, we have

$$\begin{aligned} \phi_{\mathbb{1}^I, 2\delta} &= \pi_I(2\delta)g(\delta)f_{\mathbb{1}^I} = \pi_I(2\delta) \left(\sum_{J \in \mathcal{B}(N)} g_{IJ}\pi_J \right) (\delta)f_{\mathbb{1}^I} \\ &= \sum_{J \in \mathcal{B}(N)} \pi_I(2\delta)g(\delta)\iota_J(\delta)\pi_J(\delta)f_{\mathbb{1}^I} = \sum_{J \in \mathcal{B}(N)} g_{J,I}(\delta)f_{I,J}. \end{aligned}$$

Also, $\sum_{J \in \mathcal{B}(N)} g_{J,I'}(\delta)f_{I,J} = 0$ for $I \neq I' \in \mathcal{B}(M)$, since $\phi_{M, 2\delta}$ is zero between different components of M . Lemma 4.7 says that if $g_{J,I'}(\delta)f_{I,J} \neq 0$ and $I \leq_\alpha I'$, then J is $(2n - 1)\delta$ -interleaved with either I or I' . This means that if $i < i'$, then

$$0 = \sum_{J \in \mathcal{B}(N)} g_{J,I'}(\delta)f_{I,J} = \sum_{J \in \mu(A)} g_{J,I'}(\delta)f_{I,J}, \tag{1}$$

as $g_{J,I'}(\delta)f_{I,J} = 0$ for all J that are not $(2n - 1)\delta$ -interleaved with either I_i or $I_{i'}$. Similarly,

$$\phi_{\mathbb{1}^{I_i}, 2\delta} = \sum_{J \in \mathcal{B}(N)} g_{J,I_i}(\delta)f_{I_i,J} = \sum_{J \in \mu(A)} g_{J,I_i}(\delta)f_{I_i,J}. \tag{2}$$

Writing this in matrix form, we get

$$\begin{bmatrix} g_{J_1, I_1}(\delta) & \dots & g_{J_s, I_1}(\delta) \\ \vdots & \ddots & \vdots \\ g_{J_1, I_r}(\delta) & \dots & g_{J_s, I_r}(\delta) \end{bmatrix} \begin{bmatrix} f_{I_1, J_1} & \dots & f_{I_r, J_1} \\ \vdots & \ddots & \vdots \\ f_{I_1, J_s} & \dots & f_{I_r, J_s} \end{bmatrix} = \begin{bmatrix} \phi_{\mathbb{1}^{I_1, 2\delta}} & ? & \dots & ? \\ 0 & \phi_{\mathbb{1}^{I_2, 2\delta}} & \dots & ? \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \phi_{\mathbb{1}^{I_r, 2\delta}} \end{bmatrix}.$$

That is, on the right-hand side we have the internal morphisms of the I_i on the diagonal, and 0 below the diagonal.

Recall that a morphism between rectangle modules can be identified with a k -endomorphism, and that in our notation, $f_{I, J}$ and $g_{J, I}$ are given by multiplication by $w(I, J)$ and $w(J, I)$, respectively. For an arbitrary morphism ψ between rectangle modules, we introduce the notation $w(\psi) = c$ if ψ is given by multiplication by c , and 0 otherwise. A consequence of Lemma 4.8 is that $w(g_{J, I_{i'}}(\delta) f_{I_i, J}) = w(g_{J, I_{i'}})w(f_{I_i, J}) = w(J, I_i)w(I_{i'}, J)$ whenever $I_i \leq_\alpha I_{i'}$, in particular if $i \leq i'$. Using (2), we get

$$\begin{aligned} 1 &= w(\phi_{\mathbb{1}^{I', 2\delta}}) = w\left(\sum_{J \in \mu(A)} g_{J, I_i}(\delta) f_{I_i, J}\right) = \sum_{J \in \mu(A)} w(g_{J, I_i}(\delta) f_{I_i, J}) \\ &= \sum_{J \in \mu(A)} w(J, I_i)w(I_i, J), \end{aligned}$$

and similarly $0 = \sum_{J \in \mu(A)} w(J, I_{i'})w(I_i, J)$ for $i \leq i'$ by (1). Again we can interpret this as a matrix equation:

$$\begin{bmatrix} w(J_1, I_1) & \dots & w(J_s, I_1) \\ \vdots & \ddots & \vdots \\ w(J_1, I_r) & \dots & w(J_s, I_r) \end{bmatrix} \begin{bmatrix} w(I_1, J_1) & \dots & w(I_r, J_1) \\ \vdots & \ddots & \vdots \\ w(I_1, J_s) & \dots & w(I_r, J_s) \end{bmatrix} = \begin{bmatrix} 1 & ? & \dots & ? \\ 0 & 1 & \dots & ? \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}.$$

That is, the right-hand side is an $r \times r$ upper triangular matrix with 1’s on the diagonal. The right-hand side has rank $|A|$ and the left-hand side has rank at most $|\mu(A)|$, so the lemma follows immediately from this equation. □

4.2 Free Modules

Definition 4.11 A *free interval* is an interval of the form $\langle p \rangle := \{q \mid q \geq p\} \subset \mathbb{R}^n$.

For a free interval R , we define \min_R by $R = \langle \min_R \rangle$. This makes \min_R an undecorated point, while we have previously defined \min_I as decorated points, but this does not matter, as we will not need decorated points in this subsection. We define a *free \mathbb{R}^n -module* as an interval decomposable module whose barcode only contains free

intervals. It is easy to see that free intervals are rectangles, so it follows from Theorem 4.3 that $d_B(M, N) \leq (2n - 1)d_1(M, N)$ for free modules M, N . But because of the geometry of free modules, this result can be strengthened.

Theorem 4.12 *Let M and N be free δ -interleaved \mathbb{R}^n -modules with $n \geq 2$. Then there is a $(n - 1)\delta$ -matching between $\mathcal{B}(M)$ and $\mathcal{B}(N)$. Thus, $d_B(M, N) \leq (n - 1) d_1(M, N)$.*

Using a different method, Botnan and Lesnick [11] proved this for $n = 2$. For $n \geq 3$, the result is new.

We already did most of the work while proving Theorem 4.3, and there are some obvious simplifications. Firstly, free intervals are ϵ -significant for all $\epsilon \geq 0$. Secondly, for all nonzero $f: \mathbb{I}^R \rightarrow \mathbb{I}^S$ and $g: \mathbb{I}^S \rightarrow \mathbb{I}^T$ with R, S, T free, gf is nonzero. For $I \in \mathcal{B}(M)$, define $v(I) = \{J \in \mathcal{B}(N) \mid I \text{ and } J \text{ are } (n - 1)\delta\text{-interleaved}\}$. By the arguments in the proof of Theorem 4.3, we only need to prove Lemma 4.9 with μ replaced by v . Lemmas 4.7 and 4.8 still hold for free modules, but we need to sharpen Lemma 4.7.

Lemma 4.13 *Let R, S , and T be free intervals with $R \leq_\alpha T$. Suppose there are morphisms $0 \neq f: \mathbb{I}^R \rightarrow \mathbb{I}^S(\epsilon)$ and $0 \neq g: \mathbb{I}^S \rightarrow \mathbb{I}^T(\epsilon)$. Then \mathbb{I}^S is $(n - 1)\epsilon$ -interleaved with either \mathbb{I}^R or \mathbb{I}^T .*

Proof We have $\min_S \leq \min_R + \epsilon$. Suppose \mathbb{I}^R and \mathbb{I}^S are not $(n - 1)\epsilon$ -interleaved. Then $\min_S + (n - 1)\epsilon \not\leq \min_R$ so, for some m , we must have $\min_{S_m} < \min_{R_m} - (n - 1)\epsilon$. We get

$$\alpha(S) = \sum_{i=1}^n \min_{S_i} < \min_{R_m} - (n - 1)\epsilon + \sum_{i \neq m} (\min_{R_i} + \epsilon) = \sum_{i=1}^n \min_{R_i} = \alpha(R).$$

We can also prove that $\alpha(T) < \alpha(S)$ if \mathbb{I}^S and \mathbb{I}^T are not $(n - 1)\epsilon$ -interleaved, so we have $\alpha(T) < \alpha(R)$, a contradiction. □

4.3 Block Decomposable Modules

Next, we consider block decomposable modules, which we are interested in because of their relation to zigzag modules, interlevel set stability, and Reeb graphs, explained in Sect. 3. Let $\mathbb{R}^2_+ = \{(x, y) \in \mathbb{R}^2 \mid x + y \geq 0\}$.

Definition 4.14 A *triangle* is a nonempty set of the form $\{(x, y) \in \mathbb{R}^2_+ \mid x < a, y < b\}$ for some $(a, b) \in (\mathbb{R} \cup \{\infty\})^2$ with $a + b > 0$.

Thus, a triangle by this definition is a right triangle with a vertical and a horizontal edge and the hypotenuse on the diagonal $x = -y$ as its boundary. For a triangle T as above, we write $\max_T = (a, b)$. Note that we treat \max_T as an undecorated point, like \min_R in the previous subsection. If T is bounded, \max_T is the maximal element in the closure of T , as illustrated in Fig. 5. A *triangle decomposable module* is an interval decomposable \mathbb{R}^2 -module whose barcode only contains triangles.

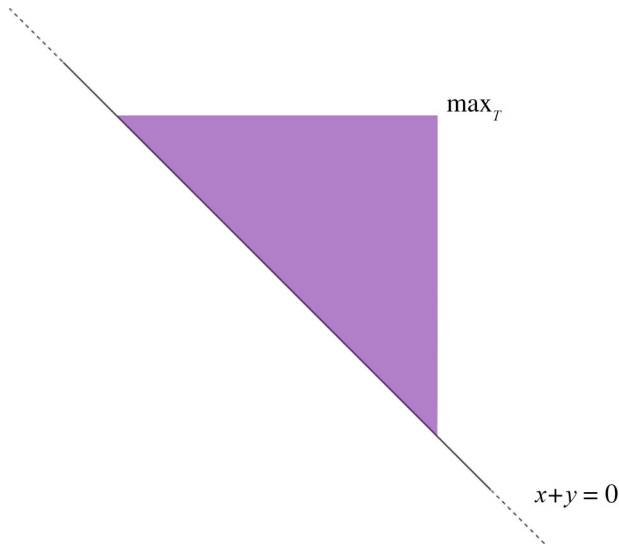


Fig. 5 A bounded triangle T

We will focus on proving stability for open blocks, i.e., blocks of the form $(a, b)_{\text{BL}}$, as this is the only case where an optimal constant has not been obtained in previous work. If we consider blocks to be subsets of $\mathbb{R}^{\text{op}} \times \mathbb{R}$ instead of \mathbb{U} , open blocks correspond to triangles under the isomorphism of posets $\mathbb{R}^{\text{op}} \times \mathbb{R} \rightarrow \mathbb{R}^2$ flipping the x -axis. We prefer to consider triangles, as \mathbb{R}^2 fits better into the framework we have built than $\mathbb{R}^{\text{op}} \times \mathbb{R}$.

Theorem 4.15 *Let M and N be triangle decomposable modules. If M and N are δ -interleaved, there exists a δ -matching between $\mathcal{B}(M)$ and $\mathcal{B}(N)$. Thus, $d_1(M, N) = d_{\text{B}}(M, N)$.*

As we did with the rectangles, we can split the triangles into sets of various ‘types’. We get four different types of triangles T , depending on whether \max_T is of the form (a, b) , (∞, b) , (a, ∞) , or (∞, ∞) for $a, b \in \mathbb{R}$. Now a result analogous to Lemma 4.6 holds, implying that it is enough to show Theorem 4.15 under the assumption that the barcodes only contain intervals of a single type. The case in which the triangles are bounded is the hardest one, and the only one we will prove. So from now on, we assume all triangles to be bounded.

Again, we reuse parts of the proof of Theorem 4.3. For $I \in \mathcal{B}(M)$, we define

$$v(I) = \{J \in \mathcal{B}(N) \mid \mathbb{I}^I \text{ and } \mathbb{I}^J \text{ are } \delta\text{-interleaved}\}.$$

The discussion about Hall’s theorem is still valid, so we only need to prove the analogue of Lemma 4.9 for v . For $\max_T = (a, b)$, define $\alpha(T) = a + b$. The only things we need to complete for the proof of the analogue of Lemma 4.9 for triangle decomposable modules, are the following analogues of Lemmas 4.7 and 4.8:

Lemma 4.16 *Let $R, S,$ and T be triangles with $\alpha(R) \leq \alpha(T)$. Suppose there are morphisms $f: \mathbb{I}^R \rightarrow \mathbb{I}^S(\epsilon)$ and $g: \mathbb{I}^S \rightarrow \mathbb{I}^T(\epsilon)$ such that $g(\epsilon) \circ f \neq 0$. Then \mathbb{I}^S is ϵ -interleaved with either \mathbb{I}^R or \mathbb{I}^T .*

Lemma 4.17 *Let $R, S,$ and T be triangles with T 2ϵ -significant and $\alpha(R) \leq \alpha(T)$. Suppose there are nonzero morphisms $f: \mathbb{I}^R \rightarrow \mathbb{I}^S(\epsilon)$ and $g: \mathbb{I}^S \rightarrow \mathbb{I}^T(\epsilon)$. Then $g(\epsilon) \circ f \neq 0$.*

Proof of Lemma 4.16 Suppose \mathbb{I}^R and \mathbb{I}^S are not ϵ -interleaved, and write $\max_R = (r_1, r_2)$ and $\max_S = (s_1, s_2)$. Then $\max_S \not\geq \max_R - \epsilon$, so $s_1 < r_1 - \epsilon$ or $s_2 < r_2 - \epsilon$. At the same time, $\max_R \geq \max_S - \epsilon$ since f is nonzero, which gives $s_i \leq r_i + \epsilon$ for each i . Thus,

$$\alpha(S) = s_1 + s_2 < r_1 + r_2 - \epsilon + \epsilon = \alpha(R).$$

Assuming that \mathbb{I}^S and \mathbb{I}^T are not ϵ -interleaved, either, we also get $\alpha(S) > \alpha(T)$. Thus $\alpha(R) > \alpha(T)$, a contradiction. \square

Proof of Lemma 4.17 We have $\max_T - \epsilon \leq \max_S$ and $\max_S - \epsilon \leq \max_R$, so $\max_T - 2\epsilon \leq \max_R$. Because T is 2ϵ -significant, $\max_T - 2\epsilon - \epsilon' \in \mathbb{R}_+^2$ for some $\epsilon' > 0$. Combining these facts, we get $\max_T - 2\epsilon - \epsilon' \in R$, so $(g(\epsilon) \circ f)_{\max_T - 2\epsilon - \epsilon'} \neq 0$. \square

Theorem 4.15 implies $d_B(M, N) = d_I(M, N)$ for block decomposable M and N such that $\mathcal{B}(M)$ and $\mathcal{B}(N)$ only have blocks of the form $(a, b)_{BL}$ (so no closed or half-closed blocks). Our proof technique extends easily to prove the same equality for all block decomposable M and N . In fact, $d_B(M, N) \leq d_I(M, N)$ in the case where all the intervals in the barcodes are of the form $[a, b]_{BL}$ follows from Theorem 4.12 with $n = 2$ by the correspondence $[a, b]_{BL} \leftrightarrow \langle (-a, b) \rangle$, while the two cases with half-open blocks are both essentially the algebraic stability theorem. A variant of Lemma 4.6 and the discussion following it would allow us to handle all the cases simultaneously. We omit the details, and anyway the closed and half-open cases are taken care of in [11]. Thus, either by appealing to previous work for the other cases or using our own methods, we get

Theorem 4.18 *Let M and N be block decomposable modules. If M and N are δ -interleaved, there exists a δ -matching between $\mathcal{B}(M)$ and $\mathcal{B}(N)$. Thus, $d_I(M, N) = d_B(M, N)$.*

5 Counterexamples to a General Algebraic Stability Theorem

Theorem 4.3 gives an upper bound of $2n - 1$ on d_B/d_I for rectangle decomposable modules, which increases with the dimension. An obvious question is whether it is possible to improve this, or if the constant $2n - 1$ in Theorem 4.3 is optimal. We know that $d_B(M, N) \geq d_I(M, N)$ for any M and N whenever the bottleneck distance is defined, so for $n = 1$, the constant is optimal. For $n > 1$, however, it turns out that the equality $d_B(M, N) = d_I(M, N)$ does not always hold, and the geometry becomes more complicated when n increases. In Example 5.2, we construct rectangle

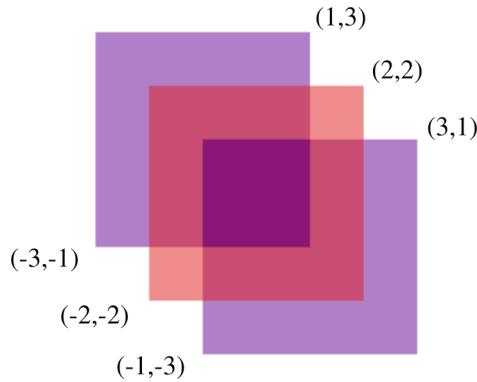


Fig. 6 M and N . I_1 and I_2 are the light purple squares, I_3 is deep purple, and J is pink

decomposable \mathbb{R}^2 -modules M and N with $d_B(M, N) = 3d_1(M, N)$, showing that the bound is optimal for $n = 2$, as well. This is a counterexample to a conjecture made in a previous version of [11] (Conjecture 8.4 in Paper II of [8]) which claims that interval decomposable \mathbb{R}^n -modules M and N such that $\mathcal{B}(M)$ and $\mathcal{B}(N)$ only contain convex intervals are ϵ -matched if they are ϵ -interleaved. As a consequence of this counterexample, they weakened the conjecture to saying that there are constants c_n depending on n such that $d_B \leq c_n d_1$ for such modules. This problem is still open, and there are no known counterexamples even if we put all the c_n equal to 3.

Before the example with $d_B(M, N) = 3d_1(M, N)$, we give a simpler example with $d_B(M, N) = 2d_1(M, N)$. These easily generalize to examples with $d_B(M, N) = 2d_1(M, N)$ and $d_B(M, N) = 3d_1(M, N)$ in n dimensions for any $n \geq 2$. For instance, one can replace M and N with $M \times I^{n-2}$ and $N \times I^{n-2}$ for a sufficiently large interval I to get such examples in n dimensions.

Example 5.1 Let $\mathcal{B}(M) = \{I_1, I_2, I_3\}^2$ and $\mathcal{B}(N) = \{J\}$, where

$$\begin{aligned} I_1 &= (-3, 1) \times (-1, 3), & I_2 &= (-1, 3) \times (-3, 1), \\ I_3 &= (-1, 1) \times (-1, 1), & J &= (-2, 2) \times (-2, 2). \end{aligned}$$

See Fig. 6. We can define 1-interleaving morphisms $f: M \rightarrow N(1)$ and $g: N \rightarrow M(1)$ by letting $w(I_1, J) = w(I_2, J) = w(I_3, J) = w(J, I_1) = w(J, I_2) = 1$ and $w(J, I_3) = -1$, where w is defined as in the proof of Theorem 4.3. On the other hand, in any matching between $\mathcal{B}(M)$ and $\mathcal{B}(N)$ we have to leave either I_1 or I_2 unmatched, and they are ϵ -significant for all $\epsilon < 4$. In fact, any possible matching between $\mathcal{B}(M)$ and $\mathcal{B}(N)$ is a 2-matching. Thus $d_1(M, N) = 1$ and $d_B(M, N) = 2$.

A crucial point is that even though $w(I_1, J)$, $w(J, I_2)$, $w(I_2, J)$, and $w(J, I_1)$ are all nonzero, both $g_{J,I_2} \circ f_{I_1,J}$ and $g_{J,I_1} \circ f_{I_2,J}$ are zero. To do the same with one-dimensional intervals, we would have to shrink I_1 and I_2 so much that they were no longer 2-significant (see Lemma 4.8), and then they would not need to be matched

² Here we use subscripts to index different intervals, not to indicate projections, as we did earlier.

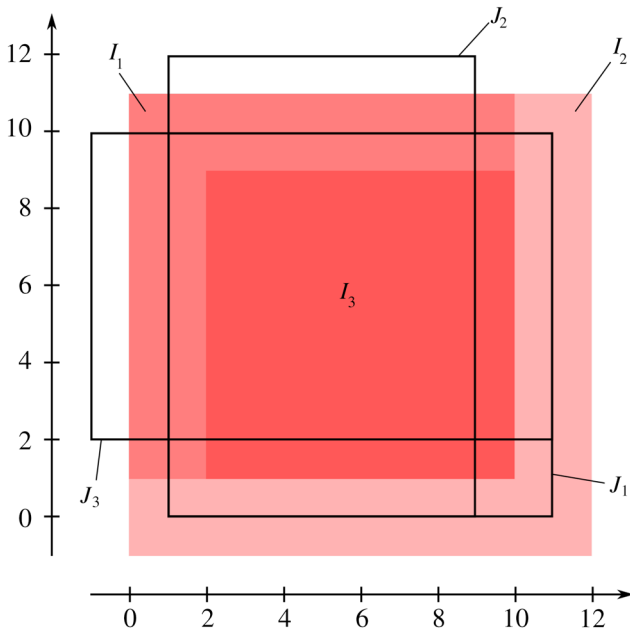


Fig. 7 $I_1, I_2,$ and I_3 are the filled pink rectangles, and $J_1, J_2,$ and J_3 are the black rectangles without fill

in a 1-matching. This shows how the geometry of higher dimensions can allow us to construct examples that would not work in lower dimensions.

Next, we give an example of rectangle decomposable \mathbb{R}^2 -modules M and N such that $d_B(M, N) = 3d_1(M, N)$, proving that our upper bound of $2n - 1$ is the best possible for $n = 2$.

Example 5.2 Let $\mathcal{B}(M) = \{I_1, I_2, I_3\}$ and $\mathcal{B}(N) = \{J_1, J_2, J_3\}$, where

$$\begin{aligned}
 I_1 &= (0, 10) \times (1, 11), & I_2 &= (0, 12) \times (-1, 11), & I_3 &= (2, 10) \times (1, 9), \\
 J_1 &= (1, 11) \times (0, 10), & J_2 &= (1, 9) \times (0, 12), & J_3 &= (-1, 11) \times (2, 10).
 \end{aligned}$$

The rectangles in $\mathcal{B}(M)$ and $\mathcal{B}(N)$ are shown in Fig. 7. We give an example of 1-interleaving morphisms f and g that we write in matrix form. In the first matrix, $w(I_i, J_j)$ is in row i , column j . In the second, $w(J_j, I_i)$ is in row j , column i .

$$f: \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad g: \begin{bmatrix} -1 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}.$$

This means that M and N are 1-interleaved, but they are not ϵ -interleaved for any $\epsilon < 1$, so $d_1(M, N) = 1$. The following matrix shows $d_1(\mathbb{I}^{I_i}, \mathbb{I}^{J_j})$ in position (i, j) :

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 3 \\ 1 & 3 & 3 \end{bmatrix}$$

This is straightforward to check. For instance, the difference between $\max_{I_2} = (12, 11)$ and $\max_{J_2} = (9, 12)$ is 3 in the first coordinate, so I_2 and J_2 are not ϵ -interleaved for any $\epsilon < 3$, and thus, $d_1(\mathbb{I}^{I_2}, \mathbb{I}^{J_2}) \geq 3$. We see from the table that any ϵ -matching with $\epsilon < 3$ between $\mathcal{B}(M)$ and $\mathcal{B}(N)$ has to match both I_2 and I_3 with J_1 , which is not possible. Therefore, $d_B(M, N) \geq 3$. On the other hand, any bijection between $\mathcal{B}(M)$ and $\mathcal{B}(N)$ is a 3-matching, so $d_B(M, N) = 3$.

There is a strong connection between n -dimensional rectangle decomposable modules and $2n$ -dimensional free modules. This is related to the fact that we need $2n$ coordinates to determine an n -dimensional rectangle, and also $2n$ coordinates to determine a $2n$ -dimensional free interval. The following example illustrates this connection, as we simply rearrange the coordinates of \min_R, \max_R for all rectangles R involved in Example 5.2 to get 4-dimensional free modules with similar properties as in Example 5.2.

Example 5.3 Let $\mathcal{B}(M) = \{I_1, I_2, I_3\}$ and $\mathcal{B}(N) = \{J_1, J_2, J_3\}$, where

$$\begin{aligned} I_1 &= \langle(0, 1, 10, 11)\rangle, & I_2 &= \langle(0, -1, 12, 11)\rangle, & I_3 &= \langle(2, 1, 10, 9)\rangle, \\ J_1 &= \langle(1, 0, 11, 10)\rangle, & J_2 &= \langle(1, 0, 9, 12)\rangle, & J_3 &= \langle(-1, 2, 11, 10)\rangle. \end{aligned}$$

Compare with the intervals I_i and J_j in Example 5.2. We have 1-interleaving morphisms defined the same way as in Example 5.2. Just as in that example, we can deduce that there is nothing better than a 3-matching between $\mathcal{B}(M)$ and $\mathcal{B}(N)$, so $d_B(M, N) = 3$ and $d_1(M, N) = 1$.

As a consequence of this example, we get that our upper bound of $d_B/d_1 \leq n - 1$ for free n -dimensional modules cannot be improved for $n = 4$.

6 Relation to the Complexity of Calculating Interleaving Distance

The interleaving distance between arbitrary persistence modules is on the surface not easy to compute, as naively trying to construct interleaving morphisms can quickly lead to a complicated set of equations for which it is not clear that one can decide if there is a solution in polynomial time. For \mathbb{R} -modules, however, the interval decomposition theorem plus the algebraic stability theorem gives us a polynomial time algorithm to compute d_1 : decompose the modules into interval modules and find the bottleneck distance. Since $d_I = d_B$, this gives us the interleaving distance. When the barcodes exist, one can compute the bottleneck distance in polynomial time also in two dimensions [20], but the approach fails for general \mathbb{R}^n -modules already at the first step, as

we do not have a nice decomposition theorem. But in the recent proof that calculating interleaving distance is NP-hard [7], it is the failure of the second step that is exploited. Specifically, a set of modules that decompose nicely into interval modules (*staircase modules*, to be precise) is constructed, but for these, d_I and d_B are different. It turns out that calculating d_I for these corresponds to deciding whether certain matrix invertibility problems called *constrained invertibility problems* (CI problems) defined in [6] are solvable, which is shown to be NP-hard.

Though rectangle modules are not considered in the NP-hardness proof, they have similar properties to staircase modules. The only significant difference in this setting is that in a fixed dimension, rectangle modules are defined by a limited number of coordinates, or “degrees of freedom”, while there is no such restriction on staircase modules even in dimension 2. Example 5.2 is essentially a CI problem with a corresponding pair of modules. Importantly, it shows that $d_I = d_B$ does not hold in general for modules corresponding to CI problems. This crucial observation, which appeared first in a preprint of this paper, opened the door to proving NP-hardness of calculating d_I by the approach used in [7].

In [7], it is also shown that *c*-approximating d_I is NP-hard for $c < 3$, where an algorithm is said to *c*-approximate d_I if it returns a number in the interval $[d_I(M, N), cd_I(M, N)]$ for any input pair M, N of modules. Whether the approach by CI problems can be used to prove hardness of *c*-approximation for $c \geq 3$ is closely related to the question of whether Theorem 4.3 can be strengthened: It can be shown that if $d_B(M, N) \leq cd_I(M, N)$ holds for all pairs M, N of rectangle decomposable modules, the same holds for staircase modules, and therefore there is a polynomial time algorithm *c*-approximating d_I for these, meaning that the strategy of going through CI problems will not give a proof that *c*-approximation of d_I is NP-hard. On the other hand, if one can find an example of rectangle decomposable modules M and N such that $d_B(M, N) = cd_I(M, N)$ for $c > 3$, one might be able to use that to increase the constant 3 in the approximation hardness result. Thus, there is a strong link between stability of rectangle decomposable modules and the only successful method so far known to the author of determining the complexity of computing or approximating multiparameter interleaving distance.

Acknowledgements I would like to thank my PhD supervisors Gereon Quick and Nils Baas for invaluable support and help. I would also like to thank Peter Landweber, Steve Oudot, and the referees for valuable feedback, and Magnus Bakke Botnan for interesting discussions.

Funding Open access funding provided by NTNU Norwegian University of Science and Technology (incl. St. Olavs Hospital – Trondheim University Hospital).

Data availability Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted

by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Aigner, M., Ziegler, G.M.: *Proofs from The Book*. Springer, Berlin (2010)
2. Azumaya, G.: Corrections and supplementaries to my paper concerning Krull–Remak–Schmidt’s theorem. *Nagoya Math. J.* **1**, 117–124 (1950)
3. Bauer, U., Lesnick, M.: Induced matchings and the algebraic stability of persistence barcodes. *J. Comput. Geom.* **6**(2), 162–191 (2015)
4. Bauer, U., Munch, E., Wang, Y.: Strong equivalence of the interleaving and functional distortion metrics for Reeb graphs. In: 31st International Symposium on Computational Geometry. *Leibniz Int. Proc. Inform.*, vol. 34, pp. 461–475. Leibniz-Zent. Inform., Wadern (2015)
5. Bendich, P., Edelsbrunner, H., Morozov, D., Patel, A.: Homology and robustness of level and interlevel sets. *Homol. Homotopy Appl.* **15**(1), 51–72 (2013)
6. Bjerkevik, H.B., Botnan, M.B.: Computational complexity of the interleaving distance. In: 34th International Symposium on Computational Geometry. *Leibniz Int. Proc. Inform.*, vol. 99, # 13. Leibniz-Zent. Inform., Wadern (2018)
7. Bjerkevik, H.B., Botnan, M.B., Kerber, M.: Computing the interleaving distance is NP-hard. *Found. Comput. Math.* **20**(5), 1237–1271 (2020)
8. Botnan, M.B.: *Applications and Generalizations of the Algebraic Stability Theorem*. PhD thesis, Norwegian University of Science and Technology (2015)
9. Botnan, M.B.: Interval decomposition of infinite zigzag persistence modules. *Proc. Am. Math. Soc.* **145**(8), 3571–3577 (2017)
10. Botnan, M.B., Crawley-Boevey, W.: Decomposition of persistence modules. *Proc. Am. Math. Soc.* **148**(11), 4581–4596 (2020)
11. Botnan, M.B., Lesnick, M.: Algebraic stability of zigzag persistence modules. *Algebr. Geom. Topol.* **18**(6), 3133–3204 (2018)
12. Carlsson, G., Mémoli, F.: *Multiparameter hierarchical clustering methods. Classification as a Tool for Research* (Dresden 2009). *Studies in Classification Data Analysis and Knowledge Organization*, pp. 63–70. Springer, Berlin (2010)
13. Carlsson, G., de Silva, V., Morozov, D.: Zigzag persistent homology and real-valued functions. In: 25th Annual Symposium on Computational Geometry (Aarhus 2009), pp. 247–256. ACM, New York (2009)
14. Carlsson, G., Zomorodian, A.: The theory of multidimensional persistence. *Discret. Comput. Geom.* **42**(1), 71–93 (2009)
15. Chazal, F., Cohen-Steiner, D., Glisse, M., Guibas, L.J., Oudot, S.Y.: Proximity of persistence modules and their diagrams. In: 25th Annual Symposium on Computational Geometry (Aarhus 2009), pp. 237–246. ACM, New York (2009)
16. Chazal, F., de Silva, V., Glisse, M., Oudot, S.: *The Structure and Stability of Persistence Modules*. SpringerBriefs in Mathematics. Springer, Cham (2016)
17. Cochoy, J., Oudot, S.: Decomposition of exact pfd persistence bimodules. *Discret. Comput. Geom.* **63**(2), 255–293 (2020)
18. Cohen-Steiner, D., Edelsbrunner, H., Harer, J.: Stability of persistence diagrams. *Discret. Comput. Geom.* **37**(1), 103–120 (2007)
19. Crawley-Boevey, W.: Decomposition of pointwise finite-dimensional persistence modules. *J. Algebra Appl.* **14**(5), # 1550066 (2015)
20. Dey, T.K., Xin, C.: Computing bottleneck distance for 2-D interval decomposable modules. In: 34th International Symposium on Computational Geometry. *Leibniz Int. Proc. Inform.*, vol. 99, # 32. Leibniz-Zent. Inform., Wadern (2018)
21. Hall, P.: On representatives of subsets. *J. Lond. Math. Soc.* **10**, 26–30 (1935)
22. Hilaga, M., Shinagawa, Y., Kohmura, T., Kunii, T.L.: Topology matching for fully automatic similarity estimation of 3D shapes. In: 28th Annual Conference on Computer Graphics and Interactive Techniques (Los Angeles 2001), pp. 203–212. ACM, New York (2001)

23. Kim, W., Mémoli, F.: Stable signatures for dynamic metric spaces via zigzag persistent homology (2017). [arXiv:1712.04064](https://arxiv.org/abs/1712.04064)
24. Natali, M., Biasotti, S., Patanè, G., Falcidieno, B.: Graph-based representations of point clouds. *Graph. Models* **73**(5), 151–164 (2011)
25. Nicolau, M., Levine, A.J., Carlsson, G.: Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival. *Proc. Natl. Acad. Sci. USA* **108**(17), 7265–7270 (2011)
26. Reeb, G.: Sur les points singuliers d'une forme de Pfaff complètement intégrable ou d'une fonction numérique. *C. R. Acad. Sci. Paris* **222**, 847–849 (1946)
27. Shinagawa, Y., Kunii, T.L., Kergosien, Y.L.: Surface coding based on Morse theory. *IEEE Comput. Graph. Appl.* **11**(5), 66–78 (1991)
28. de Silva, V., Munch, E., Patel, A.: Categorized Reeb graphs. *Discret. Comput. Geom.* **55**(4), 854–906 (2016)
29. Singh, G., Mémoli, F., Carlsson, G.: Topological methods for the analysis of high dimensional data sets and 3D object recognition. In: *Eurographics Symposium on Point-Based Graphics (Prague 2007)*, pp. 91–100. Eurographics Association (2007)
30. Wood, Z., Hoppe, H., Desbrun, M., Schröder, P.: Removing excess topology from isosurfaces. *ACM Trans. Graph.* **23**(2), 190–208 (2004)
31. Yao, Y., Sun, J., Huang, X., Bowman, G.R., Singh, G., Lesnick, M., Guibas, L.J., Pande, V.S., Carlsson, G.: Topological methods for exploring low-density states in biomolecular folding pathways. *J. Chem. Phys.* **130**(14), # 144115 (2009)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.