

# Stability of interval decomposable persistence modules

Håvard Bakke Bjerkevik

January 21, 2020

## Abstract

The algebraic stability theorem for  $\mathbb{R}$ -persistence modules is a fundamental result in topological data analysis. We present a stability theorem for  $n$ -dimensional rectangle decomposable persistence modules up to a constant  $(2n - 1)$  that is a generalization of the algebraic stability theorem, and also has connections to the complexity of calculating the interleaving distance. The proof given reduces to a new proof of the algebraic stability theorem with  $n = 1$ . We give an example to show that the bound cannot be improved for  $n = 2$ . We apply the same technique to prove stability results for zigzag modules and Reeb graphs, reducing the previously known bounds to a constant that cannot be improved, settling these questions.

## 1 Introduction

Persistent homology is a tool in topological data analysis used to determine the structure or shape of data sets. For example, given a point cloud  $X \subset \mathbb{R}^n$  sampled from a subspace  $S$  of  $\mathbb{R}^n$ , we want to guess at the homology of  $S$ , which tells us something about how many “holes”  $S$  has in various dimensions. We can do this by defining  $B(\epsilon)$  to be the union of the (open or closed) balls of radius  $\epsilon$  centered at each point in  $X$ . Calculating homology, we get a group or vector space  $H_n(B(\epsilon))$  for each  $\epsilon \geq 0$ , and the inclusions  $B(\epsilon) \hookrightarrow B(\epsilon')$  induce morphisms  $H_n(B(\epsilon)) \rightarrow H_n(B(\epsilon'))$  for  $\epsilon \leq \epsilon'$ . Such a collection of vector spaces and morphisms is called a *persistence module*, or  $\mathbb{R}$ -module, as the vector spaces are parametrized over  $\mathbb{R}$ . Under certain assumptions, we can decompose an  $\mathbb{R}$ -module into *interval modules* [15], which gives us a set of intervals uniquely determining the persistence module up to isomorphism. This set of intervals is the *barcode* of the persistence module. The intervals in the barcode are interpreted as corresponding to possible features of the space  $S$ , where one might interpret long intervals as more likely to describe actual features of  $S$  and short intervals as more likely to be the result of noise in the input data. In other words, we have an algorithm with a data set as input and a barcode as output. As data sets always carry a certain amount of noise, we would like this algorithm to be *stable* in the sense that a little change in the input data, or in the persistence modules, should not result in a big change in the barcode.

We measure the difference between persistence modules with the *interleaving distance*  $d_I$ , and the difference between barcodes with the *bottleneck distance*  $d_B$ . Proving stability then becomes a question of proving that the bottleneck distance is bounded by the interleaving distance, i.e.  $d_B \leq Cd_I$  for some constant  $C$ . Stability has been proved for persistence modules over  $\mathbb{R}$  in [14, 12, 13, 5] in what is called the *algebraic stability theorem*, which implies the *isometry theorem*  $d_I = d_B$ .

Persistence modules can also be parametrized over other posets. A pair of filtrations  $f, g : S \rightarrow \mathbb{R}$  of a topological space  $S$  gives rise to an  $\mathbb{R}^2$ -module which has a vector space  $V_p$  for each point in  $p \in \mathbb{R}^2$  and linear maps  $V_{(a,b)} \rightarrow V_{(c,d)}$  whenever  $a \leq c$  and  $b \leq d$ , for instance by letting  $V_{(a,b)}$  be  $H_n(f^{-1}(-\infty, a) \cap g^{-1}(-\infty, b))$ . Again, inclusions induce the linear maps on homology. With  $n$  filtrations instead of 2, we get an  $\mathbb{R}^n$ -module. Another example is *zigzag modules*, which are popular objects of study in topological data analysis [10, 20, 19]. These can arise from a sequence of subspaces  $S_i \subset S$ , where we also consider the intersections  $S_i \cap S_{i+1}$  (or unions). In this case, we have

$$\cdots \subseteq S_{i-1} \supseteq S_{i-1} \cap S_i \subseteq S_i \supseteq S_i \cap S_{i+1} \subseteq \cdots,$$

which again gives rise to linear maps on homology. Defining interleavings and thus the interleaving distance is trickier than for  $\mathbb{R}$ -modules, but in fact one can do this by associating  $\mathbb{R}^2$ -modules called *block decomposable*

modules to the zigzag modules. One can also associate block decomposable modules to *Reeb graphs*, which are of interest because of their ability to present geometrical information despite being relatively simple objects. See Section 3.

All these examples serve as motivation for why one would like to talk about stability for multi-parameter modules (that is, persistence modules parametrized over  $\mathbb{R}^n$  for  $n \geq 2$ ). Unfortunately, no isometry theorem is possible even for general  $\mathbb{R}^2$ -modules, because there is no nice decomposition theorem like in the one-parameter cases, meaning that  $d_B$  is not defined. The block decomposable modules, however, decompose nicely, and  $d_B \leq \frac{5}{2}d_I$  has been shown for these [3]. This carries over to stability results for zigzag modules and Reeb graphs.

Our main contribution is a new method of proving stability for interval decomposable modules. We demonstrate several applications of this method. The first is Theorem 4.2:

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

This is a generalization of the algebraic stability theorem for  $\mathbb{R}$ -modules, which is the case  $n = 1$ . For  $n \geq 2$ , the result is new. There already exist several proofs of the algebraic stability theorem, but our approach is different from the ones taken before, which allows this more general theorem, as well as the results below. Our method is combinatorial, which in our opinion reflects the true nature of the problem once some of the algebraic technicalities are stripped away. Also, our proof is fairly short in the case  $n = 1$  compared to earlier proofs of the algebraic stability theorem. In Example 5.2, we construct rectangle decomposable modules  $M$  and  $N$  over  $\mathbb{R}^2$  for which  $d_I(M, N) = 1$  and  $d_B(M, N) = 3$ , disproving a conjecture made in an earlier version of [3] 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. The example also shows that the bound in the theorem cannot be improved for  $n = 2$ . It is an open question if the bound can be improved for  $n \geq 3$ .

We do not know of any examples of rectangle decomposable modules arising naturally from real-world data sets. But as we discuss in Section 6, there is a strong link between the stability of these modules and the recent proof that calculating the interleaving distance between multi-parameter modules is NP-hard [9]. In particular, our way of viewing interleavings as pairs of matrices and our observation in Example 5.2 that the interleaving and bottleneck distances differ for rectangle decomposable modules served as inspiration for the approach used in [9]. The question of whether the hardness results can be strengthened is closely related to the question of whether Theorem 4.2 can be improved. Thus, even if rectangle decomposable modules never arise directly from data sets, the type of questions we consider can have an impact on practical applications.

Another reason why we give the proof in detail for rectangle decomposable modules instead of, say, block decomposable modules, is that this case demonstrates very well exactly when our method works and when it fails. The lesson to take home is that the method gives a bound  $d_B \leq cd_I$  with a  $c$  that increases with the freedom we have in defining the intervals we consider. You need  $2n$  coordinates to define an  $n$ -dimensional (hyper)rectangle, which gives a constant  $2n - 1$  in the theorem.

Another application of our proof method gives Theorem 4.11:

**Theorem.** *Let  $M$  and  $N$  be  $\delta$ -interleaved triangle decomposable modules. Then there is a  $\delta$ -matching between  $B(M)$  and  $B(N)$ .*

This is more immediately connected to practical applications. Theorem 4.11 implies  $d_B \leq d_I$  for block decomposable modules, which is an improvement on the previous best known bound,  $d_B \leq \frac{5}{2}d_I$ . Since the opposite inequality  $d_I \leq d_B$  holds trivially, our bound is the best possible. We discuss how stability results for zigzag modules and Reeb graphs follow in Section 3. The fact that our bound is optimal means that these stability problems are now settled.

We finish off Section 4 by showing stability for free modules.

We assume that all modules are pointwise finite dimensional (p.f.d.). In a previous version of this paper [8], we strengthened the theorems by removing this assumption.

## 2 Persistence modules, interleavings, and matchings

In this section we introduce some basic 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  $M$  and  $p \leq q \in P$ ,  $M(p)$  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 that  $\phi_M(p, p) = \text{id}_{M_p}$ , and that  $\phi_M(q, r) \circ \phi_M(p, q) = \phi_M(p, r)$ . Because the 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 subset  $\emptyset \neq I \subseteq P$  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  $\mathbb{I}^J$  for the interval module with  $J$  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  $\epsilon$  when we mean  $(\epsilon, \epsilon, \dots, \epsilon) \in \mathbb{R}^n$ . We call an interval  $I \subset \mathbb{R}^n$  *bounded* if it is bounded as a subset of  $\mathbb{R}^n$  in the usual sense. That is, 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  **$\epsilon$ -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  $\epsilon$ -interleaving between  $M$  and  $N$ , then  $M$  and  $N$  are said to be  $\epsilon$ -interleaved. An interleaving can be viewed as an ‘approximate isomorphism’, and a 0-interleaving is in fact an isomorphism. We call a module  $M$   $\epsilon$ -significant if  $\phi_M(p, p + \epsilon) \neq 0$  for some  $p$ , and  $\epsilon$ -trivial otherwise.  $M$  is  $2\epsilon$ -trivial if and only if it is  $\epsilon$ -interleaved with the zero module. We call an interval  $I$   $\epsilon$ -significant if  $\mathbb{I}^I$  is  $\epsilon$ -significant, and  $\epsilon$ -trivial otherwise.

**Definition 2.6.** *We define the **interleaving distance**  $d_I$  on persistence modules  $M$  and  $N$  by*

$$d_I(M, N) = \inf\{\epsilon \mid M \text{ and } N \text{ are } \epsilon\text{-interleaved}\}. \quad (1)$$

The interleaving distance intuitively 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,  $d_I$  satisfies the axioms for a metric, so  $d_I$  is an extended pseudometric.

**Definition 2.7.** Suppose  $M \cong \bigoplus_{I \in B} \mathbb{I}^I$  for a multiset<sup>1</sup>  $B$  of intervals. Then we call  $B$  the **barcode** of  $M$ , and write  $B = B(M)$ . We say that  $M$  is **interval decomposable**.

Since the endomorphism ring of any interval module is isomorphic to  $k$ , it follows from Theorem 1 in [2] that if a persistence module  $M$  is interval decomposable, the decomposition is unique up to isomorphism. Thus the barcode is well-defined, even if we let  $M$  be a  $P$ -module for an arbitrary poset  $P$ . If  $M$  is a p.f.d.  $\mathbb{R}$ -module, it is interval decomposable [15], but this is not true for  $\mathbb{R}$ -modules or p.f.d.  $\mathbb{R}^n$ -modules in general. [21] gives an example showing the former, and the following is an example of a  $P$ -module for a poset  $P$  with four points that is not interval decomposable.

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

A corresponding  $\mathbb{R}^2$ -module that is not interval decomposable and is at most two-dimensional at each point can be constructed.

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  $\sigma : A \rightarrowtail B$ . We write  $\text{coim } \sigma = A'$  and  $\text{im } \sigma = B'$ .

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

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

If there is an  $\epsilon$ -matching between  $B(M)$  and  $B(N)$  for persistence modules  $M$  and  $N$ , we say that  $M$  and  $N$  are  $\epsilon$ -matched.

We have adopted this definition of  $\epsilon$ -matching from [3], which differs from e.g. the one in [13], which allows two intervals  $I$  and  $J$  to be matched if  $d_I(\mathbb{I}^I, \mathbb{I}^J) \leq \epsilon$  (or rather, this is equivalent to their definition). Conveniently, with the definition we have chosen, an  $\epsilon$ -interleaving is easily constructed given an  $\epsilon$ -matching. We feel that this is the more natural definition for this paper, as several of our results are phrased as statements about matchings and interleavings, and the interleaving distance might not come into the picture at all. The other definition is perhaps more natural in the context of ‘persistence diagrams’, where intervals are identified with points in a diagram, and the interleaving distance between the corresponding modules is simply the distance between the points. This is irrelevant to us, however, as we never consider persistence diagrams.

We can also define  $\epsilon$ -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  $G_\epsilon$  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  $\epsilon$ -interleaved. Then an  $\epsilon$ -matching between  $A$  and  $B$  is a matching in  $G_\epsilon$  such that the set of  $2\epsilon$ -significant intervals in  $A \sqcup B$  is covered.

<sup>1</sup>We will not be rigorous in our treatment of multisets. A multiset may contain multiple copies of one element, but we will assume that we have some way of separating the copies, so that we can treat the multiset as a set. If e.g.  $I$  and  $J$  are intervals in a multiset and we say that  $I \neq J$ , we mean that they are “different” elements of the multiset, not that they are different intervals.

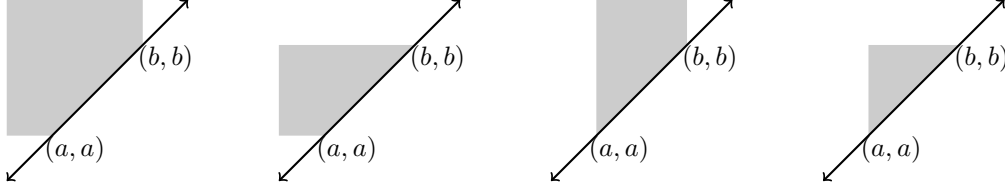


Figure 1: The intervals  $[a, b]_{BL}$ ,  $[a, b)_{BL}$ ,  $(a, b]_{BL}$  and  $(a, b)_{BL}$ .

**Definition 2.9.** The **bottleneck distance**  $d_B$  is defined by

$$d_B(M, N) = \inf\{\epsilon \mid M \text{ and } N \text{ are } \epsilon\text{-matched}\} \quad (3)$$

for any interval decomposable  $M$  and  $N$ .

We might abuse notation and talk about  $d_B(C, D)$ , where  $C$  and  $D$  are barcodes.

### 3 Zigzag modules and Reeb graphs

In this section we will give some intuition for how block decomposable modules relate to Reeb graphs and zigzag modules. We refer to [3] for a more detailed and rigorous treatment.

When explaining the connection to Reeb graphs and zigzag modules, it is more convenient to flip one of the axes in  $\mathbb{R}^2$ , so that we work with  $\mathbb{R}^{\text{op}} \times \mathbb{R}$  instead. This way,  $(a, b) \leq (c, d)$  iff  $c \leq a$  and  $b \leq d$ , or, equivalently, if  $(a, b) \subseteq (c, d)$  as intervals, assuming  $a < b$ . Let  $\mathbb{U} = \{(a, b) \in \mathbb{R}^{\text{op}} \times \mathbb{R} \mid a \leq b\}$ .

**Definition 3.1.** An interval decomposable  $\mathbb{R}^{\text{op}} \times \mathbb{R}$ -module is called **block decomposable** if its barcode only contains intervals of the following types:

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

We call these intervals *blocks*. Each interval intersects the diagonal in an  $\mathbb{R}$ -interval that is open, closed or half-open one way or the other depending on the type of the block.

#### 3.1 Reeb graphs

There have been proposed several distances on Reeb graphs; see [4] for a summary, as well as references to various applications. The interleaving distance we consider was introduced in [16].

A Reeb graph is a topological graph  $G$  together with a continuous function  $\gamma : G \rightarrow \mathbb{R}$  such that the level sets of  $\gamma$  are discrete. Let  $S(\gamma) = \mathbb{R}^{\text{op}} \times \mathbb{R} \rightarrow \mathbf{Set}$  be the functor sending  $(a, b)$  to the set of connected components of  $\gamma^{-1}(a, b)$  and  $S((a, b) \rightarrow (c, d))$  be induced by the inclusion  $\gamma^{-1}(a, b) \subseteq \gamma^{-1}(c, d)$  for  $c \leq a \leq b \leq d$ . In Figure 2,  $\gamma$  is projection to a horizontal axis. Above the graph, the functor  $S(\gamma)$  is shown, the shade of grey at  $(a, b)$  determined by the size of  $S(\gamma)_{(a, b)}$ .

Given two Reeb graphs  $(G_1, \gamma_1)$  and  $(G_2, \gamma_2)$ , we get two functors  $S(\gamma_1)$  and  $S(\gamma_2)$ , and we can talk about interleavings and interleaving distance by adjusting the definitions in the previous section. It turns out that this interleaving distance is at least as big as the one we get by replacing  $S(\gamma_1)$  and  $S(\gamma_2)$  by corresponding block decomposable modules  $M_{\gamma_1}$  and  $M_{\gamma_2}$ . In Figure 2, the blocks comprising this block decomposable module are exactly what you would guess by looking at the figure. In [3],  $d_B(M, N) \leq \frac{5}{2} d_I(M, N)$  is proved for such modules; with Theorem 4.14, we have  $d_B(M, N) = d_I(M, N)$ .

There is also a barcode  $L_0(\gamma)$  of  $\mathbb{R}$ -intervals (the level set persistence diagram [11]) associated to a Reeb graph  $(G, \gamma)$ , which we can think of as arising from the intersection of  $S(\gamma)$  with the diagonal  $x = y$ . This

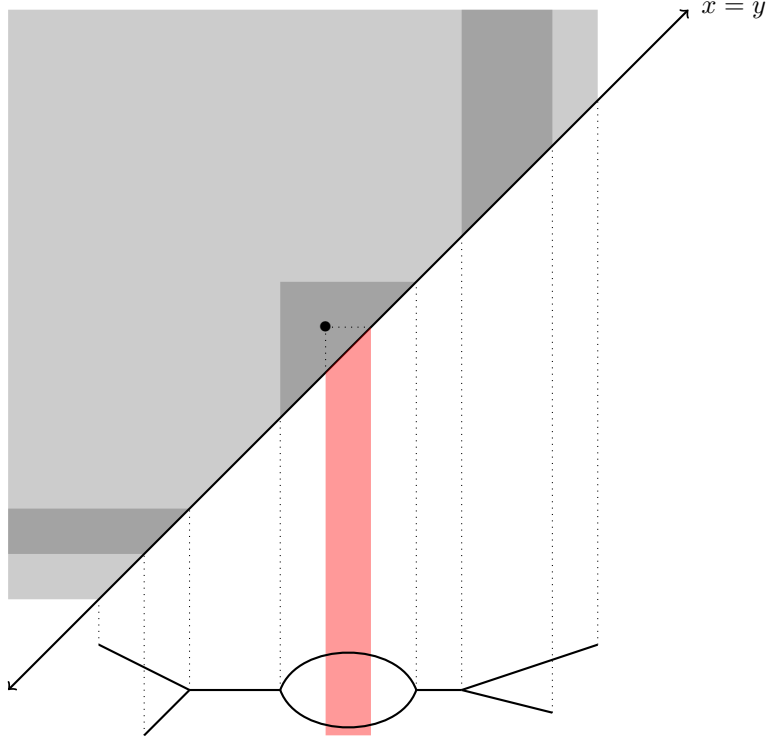


Figure 2: A Reeb graph  $(G, \gamma)$  with  $S(\gamma)$  above. Evaluating  $S(\gamma)$  at the point shown, we get the intersection of  $G$  with the red strip, which has two connected components.

barcode is the same as  $B(M_\gamma)$ , except that  $(a, b]_{BL}$  is replaced by  $(a, b]$ , and so on. It is not too hard to see that  $d_B(L_0(\gamma_1), L_0(\gamma_2)) \leq 2d_B(M_{\gamma_1}, M_{\gamma_2})$ .<sup>2</sup> Altogether, this gives

$$\begin{aligned} d_B(L_0(\gamma_1), L_0(\gamma_2)) &\leq 2d_B(M_{\gamma_1}, M_{\gamma_2}) \\ &= 2d_I(M_{\gamma_1}, M_{\gamma_2}) \\ &\leq 2d_I(S(\gamma_1), S(\gamma_2)). \end{aligned}$$

In other words:

**Theorem 3.2.** *For Reeb graphs  $\gamma_1, \gamma_2$ , the inequality  $d_B(L_0(\gamma_1), L_0(\gamma_2)) \leq 2d_I(S(\gamma_1), S(\gamma_2))$  holds.*

Thus an easily computed bottleneck distance gives a lower bound for the interleaving distance between Reeb graphs. This improves the result in [3], which was again an improvement on [6], by lowering the constant in the inequality from 5 to 2, and this cannot be improved.

### 3.2 Zigzag modules

A zigzag module is a module over  $\mathbb{ZZ} = \{(a, b) \in \mathbb{Z}^2 \mid a = b \vee a = b + 1\}$  taken as a sub-poset of  $\mathbb{R}^{\text{op}} \times \mathbb{R}$ . Let  $\mathbb{ZZ}|_{(a, b)}$  be the sub-poset of  $\mathbb{ZZ}$  containing the elements  $\{(c, d) \in \mathbb{ZZ} \mid a \leq c, d \leq b\}$ . A zigzag module  $M$  gives rise to a block decomposable module  $M_{BL}$  defined by letting  $M_{BL}(a, b)$  be the colimit of the restriction of  $M$  to  $\mathbb{ZZ}|_{(a, b)}$ .  $M_{BL}((a, b) \rightarrow (c, d))$  is defined to be the induced morphism we get by the universal property of colimits for  $(a, b) \leq (c, d)$ . (This definition is given in [3], but something very similar is described in the discussions of pyramids in [11] and [7].) This way, we can define interleaving and bottleneck distance between zigzag modules by letting  $d_I(M, N) = d_I(M_{BL}, N_{BL})$  and  $d_B(M, N) = d_B(M_{BL}, N_{BL})$ . Thus Theorem 4.14 holds if we replace ‘block decomposable modules’ by ‘zigzag modules’:

<sup>2</sup>The reason for the constant 2 is that  $(a, b)_{BL}$  is  $(b - a)/4$ -trivial, while  $(a, b)$  is not  $\epsilon$ -trivial for  $\epsilon < (b - a)/2$ .

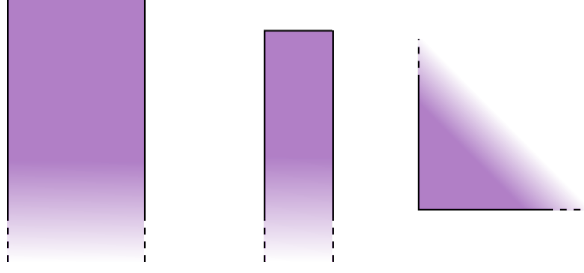


Figure 3: Three rectangles, where the left and middle rectangles are of the same type (unbounded downwards), while the last is of a different type (unbounded upwards and to the right). Assuming that it contains its boundary, the rightmost rectangle is also an example of a free interval, which we will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

**Theorem 3.3.** *Let  $M$  and  $N$  be zigzag modules. If  $M$  and  $N$  are  $\delta$ -interleaved, there exists a  $\delta$ -matching between  $B(M)$  and  $B(N)$ .*

This implies an isometry theorem for zigzag modules:  $d_I(M, N) = d_B(M, N)$ .

## 4 Higher-dimensional stability

The algebraic stability theorem for  $\mathbb{R}$ -modules states that an  $\epsilon$ -interleaving between  $\mathbb{R}$ -modules  $M$  and  $N$  induces an  $\epsilon$ -matching between  $B(M)$  and  $B(N)$ , implying  $d_I(M, N) = d_B(M, N)$ , the isometry theorem. The main purpose of this paper is to find out when similar results for  $\mathbb{R}^n$ -modules hold. Our first result, Theorem 4.2, is a generalization of the algebraic stability theorem for  $\mathbb{R}$ -modules. Variations of the algebraic stability theorem have been proved several times already [14, 12, 13, 5], but this is a new proof with ideas that are applicable to more than just  $\mathbb{R}$ -modules.

### 4.1 Rectangle decomposable modules

For any interval  $I \subset \mathbb{R}^n$ , we let its projection on the  $i$ 'th coordinate be denoted by  $I_i$ .

**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  $(a, \infty)$  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 Figure 3.

In [13], *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  $a \in \overline{\mathbb{R}}$ .<sup>3</sup> The notation is as follows for  $a, b \in \overline{\mathbb{R}}$ :

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

<sup>3</sup>The decorated numbers  $-\infty^-$  and  $\infty^+$  are never used, as no interval contains points at infinity, but it does not matter whether we include these two points in the definition.

- $I = (a^-, b^+)$  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 unknown '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  $a, b \in \overline{\mathbb{R}}$ . 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  $a \in \overline{\mathbb{R}}$ ,  $x \in \mathbb{R}$ . We add  $n$ -dimensional decorated points and  $n$ -tuples of real numbers coordinatewise.

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

Our goal is to prove the following theorem:

**Theorem 4.2.** *Let  $M = \bigoplus_{I \in B(M)} \mathbb{I}^I$  and  $N = \bigoplus_{J \in B(N)} \mathbb{I}^J$  be rectangle decomposable  $\mathbb{R}^n$ -modules. If  $M$  and  $N$  are  $\delta$ -interleaved, there exists a  $(2n - 1)\delta$ -matching between  $B(M)$  and  $B(N)$ .*

The inequality  $d_B(M, N) \leq (2n - 1)d_I(M, N)$  for rectangle decomposable modules  $M$  and  $N$  immediately follows.

Fix  $0 \leq \delta \in \mathbb{R}$ . Assume that  $M$  and  $N$  are  $\delta$ -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  $I \in B(M)$ , 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  $J \in B(N)$ . We define

$$\begin{aligned} f_{I,J} &= \pi_J(\delta) \circ f \circ \iota_I : \mathbb{I}^I \rightarrow \mathbb{I}^J(\delta) \\ g_{J,I} &= \pi_I(\delta) \circ g \circ \iota_J : \mathbb{I}^J \rightarrow \mathbb{I}^I(\delta). \end{aligned} \tag{4}$$

We prove the theorem by a mix of combinatorial and geometric arguments. First we show that it is enough to prove the theorem under the assumption that all the rectangles in  $B(M)$  and  $B(N)$  are of the same type. Then we define a real-valued function  $\alpha$  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  $\leq_\alpha$  associated to  $\alpha$ . The idea behind  $\leq_\alpha$  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  $B(M)$  and  $B(N)$  that are close is exactly what we need to construct a  $(2n - 1)\delta$ -matching. Lemmas 4.6 and 4.7 say that such morphisms behave nicely in a precise sense that we will exploit when we prove Lemma 4.8. If we remove the conditions mentioning  $\leq_\alpha$ , Lemmas 4.6 and 4.7 are not even close to being true, so one of the main points in the proof of Lemma 4.8 is that we must exclude the cases that are not covered by Lemmas 4.6 and 4.7. 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 4.8 is what ties together the geometric and combinatorial parts of the proof of Theorem 4.2. While we prove Lemma 4.8 by geometric arguments, by Hall's marriage theorem the lemma is equivalent to a statement about matchings between  $B(M)$  and  $B(N)$ . We have to do some combinatorics to get exactly the statement we need, namely that there is a  $(2n - 1)\delta$ -matching between  $B(M)$  and  $B(N)$ , and we do this after stating Lemma 4.8.

We begin by describing morphisms between rectangle modules.

**Lemma 4.3.** *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  $\phi$ -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. Note that we could have relaxed the assumptions in the proof above and assumed that  $a$  is in  $I$  instead of in  $A$ , and still have gotten  $\chi_a = \chi_b$ . In particular, this means that if  $0 \neq \chi : \mathbb{I}^I \rightarrow \mathbb{I}^J$ , and  $I$  and  $J$  are rectangles, then  $\min_{J_i} \leq \min_{I_i}$  for all  $i$ , which gives  $\min_J \leq \min_I$ . Similarly,  $\max_J \leq \max_I$ , and one can also see that  $\min_I < \max_J$  must hold, or else  $I \cap J = \emptyset$ . We summarize these observations as a corollary of Lemma 4.3:

**Corollary 4.4.** *Let  $R$  and  $S$  be rectangles, and let  $\chi : \mathbb{I}^R \rightarrow \mathbb{I}^S$  be a nonzero morphism. Then  $\min_S \leq \min_R$  and  $\max_S \leq \max_R$ .*

This will come in handy when we prove Lemmas 4.5, 4.6, and 4.7.

We define a function  $w : (B(M) \times B(N)) \sqcup (B(N) \times B(M)) \rightarrow k$  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  $B(M) \sqcup B(N)$  by letting  $w(I, J)$  be the weight of the edge from  $I$  to  $J$ . The reader is invited 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 4.5.** *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 Corollary 4.4,  $\min_R \geq \min_S \geq \min_T$  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  $I \in B(M)$  and  $J \in B(N)$  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  $I, I' \in B(M)$ . Then we have

$$\sum_{J \in B(N)} g_{J,I'}(\delta) f_{I,J} = \sum_{J \in B(N)} g'_{J,I'}(\delta) f'_{I,J}. \quad (5)$$

When  $I$  and  $I'$  are of different types, the left side is zero because  $f$  and  $g$  are  $\delta$ -interleaving morphisms, and all the summands on the right side are zero by definition of  $f'$  and  $g'$ . When  $I$  and  $I'$  are of the same type, the equality follows from Lemma 4.5. This means that  $g'(\delta)f' = g(\delta)f$ . We also have  $f'(\delta)g' = f(\delta)g$ , so  $f'$  and  $g'$  are  $\delta$ -interleaving morphisms. In particular,  $f'$  and  $g'$  are  $\delta$ -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 4.2. In other words, we have reduced the problem to the case where all the intervals in  $B(M)$  and  $B(N)$  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  $R$ . Define an order  $\leq_\alpha$  on decorated points given by  $a \leq_\alpha b$  if either

- $\alpha(a) < \alpha(b)$ , or
- $\alpha(a) = \alpha(b)$  and  $P(a) \leq P(b)$

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  $\leq_\alpha$  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  $\epsilon = 0$ ,  $R$

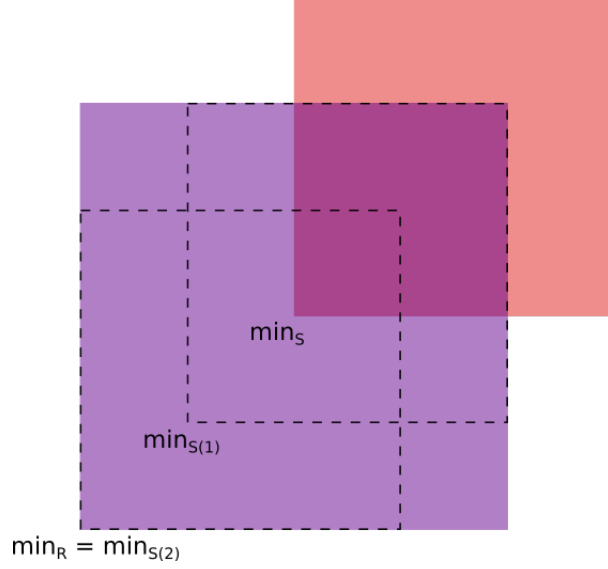


Figure 4: Rectangles  $R = (0, 4) \times (0, 4)$  (purple),  $S = (2, 5) \times (2, 5)$  (pink),  $S(1) = (1, 4) \times (1, 4)$  (dotted border), and  $S(2) = (0, 3) \times (0, 3)$  (dotted border).

and  $S$  actually have to be equal. This ‘closeness property’ is expressed in Lemma 4.6, and is also exploited in Lemma 4.7. Finally, in the proof of Lemma 4.8, 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 Figure 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  $P(\alpha)$  in the definition of  $\alpha$  is that e.g.  $(a, b]$  is a tiny bit ‘further to the right’ than  $[a, b)$ , and this is a subtlety that  $P$  recognizes, and that matters in the proofs of Lemmas 4.6 and 4.7.

**Lemma 4.6.** *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.* Since  $\chi \neq 0$ , we have

- $\min_S \leq \min_R + \epsilon$
- $\max_S \leq \max_R + \epsilon$ .

This follows from Corollary 4.4.

Suppose  $\mathbb{I}^R$  and  $\mathbb{I}^S$  are not  $(2n - 1)\epsilon$ -interleaved. Then either  $\min_S + (2n - 1)\epsilon \not\leq \min_R$  or  $\max_S + (2n - 1)\epsilon \not\leq \max_R$ ; let us assume the latter. (The former is similar.) In this case, there is an  $m$  such that  $\max_{S_m} < \max_{R_m} - (2n - 1)\epsilon$ . For  $i \neq m$ , we have  $\max_{S_i} \leq \max_{R_i} + \epsilon$  by the second bullet point. We get

$$\begin{aligned} \sum_{1 \leq i \leq n} u(\max_{S_i}) &\leq \left( \sum_{1 \leq i \leq n} u(\max_{R_i}) \right) - (2n - 1)\epsilon + (n - 1)\epsilon \\ &= \left( \sum_{1 \leq i \leq n} u(\max_{R_i}) \right) - n\epsilon. \end{aligned} \tag{6}$$

The first bullet point gives us

$$\sum_{1 \leq i \leq n} u(\min_{S_i}) \leq \left( \sum_{1 \leq i \leq n} u(\min_{R_i}) \right) + n\epsilon, \tag{7}$$

so we get  $\alpha(S) \leq \alpha(R)$ . If the inequality is strict, we have  $S <_\alpha R$ . If not, we have

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

Because of the inequalities  $\min_S \leq \min_R + \epsilon$  and  $\max_S \leq \max_R + \epsilon$  (recall that these are inequalities of decorated points with the poset structure we defined earlier), we have  $P(\min_{S_i}) \leq P(\min_{R_i})$  for all  $i$  and  $P(\max_{S_i}) \leq P(\max_{R_i})$  for  $i \neq m$ . But since  $\max_{S_m} < \max_{R_m} - (2n - 1)\epsilon$ , we have  $P(\max_{S_m}) < P(\max_{R_m})$ , so  $S <_\alpha R$ . Similarly, we can prove  $T <_\alpha S$  if  $\mathbb{I}^S$  and  $\mathbb{I}^T$  are not  $(2n - 1)\epsilon$ -interleaved, so we have  $T <_\alpha R$ , which is a contradiction.  $\square$

**Lemma 4.7.** *Let  $R$ ,  $S$ , and  $T$  be rectangles of the same type with  $R$  and  $T$   $(4n - 2)\epsilon$ -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  $\psi(\epsilon) \circ \chi \neq 0$ .*

The constant  $(4n - 2)$  can be improved on for  $n > 1$ , but since the constant  $(2n - 1)$  in Lemma 4.6 is optimal, strengthening Lemma 4.7 will not help us get a better constant in Theorem 4.2.

*Proof.* Suppose that  $\chi$  and  $\psi$  are nonzero, but  $\psi(\epsilon) \circ \chi = 0$ . We have

- $\min_R + 2\epsilon \geq \min_T$
- $\min_{R_m} + 2\epsilon \geq \max_{T_m}$  for some  $m$
- $\max_R + 2\epsilon \geq \max_T$
- $\max_{R_m} \geq \max_{T_m} + (4n - 4)\epsilon$ .

The first and third statements hold because  $\chi, \psi \neq 0$ . (See Corollary 4.4.) The second is equivalent to  $\min_R \not\leq \max_{T(2\epsilon)}$ . If this did not hold,  $R$  and  $T(2\epsilon)$  would intersect, and  $\psi(\epsilon) \circ \chi$  would be nonzero in this intersection, which is a contradiction. The fourth statement follows from the second and the fact that  $R$  is  $(4n - 2)\epsilon$ -significant.

Since  $T$  is  $(4n - 2)\epsilon$ -significant,  $\min_T + (4n - 2)\epsilon < \max_T$ . Thus the second bullet point implies that  $\min_{R_m} + 2\epsilon > \min_{T_m} + (4n - 2)\epsilon$ . The first point gives  $\min_{R_i} \geq \min_{T_i} - 2\epsilon$  for  $i \neq m$ . In a similar fashion, we get from the last two points that  $\max_{R_m} \geq \max_T + (4n - 4)\epsilon$  and  $\max_{R_i} \geq \max_{T_i} - 2\epsilon$  for  $i \neq m$ . From all this, we get

$$\begin{aligned} \alpha(R) &= \sum_{1 \leq i \leq n} u(\min_{R_i}) + u(\max_{R_i}) \\ &\geq u(\min_{T_m}) + u(\max_{T_m}) + 2(4n - 4)\epsilon + \sum_{i \neq m} (u(\min_{T_i}) + u(\max_{T_i}) - 4\epsilon) \\ &= \alpha(T) + (4n - 4)\epsilon \\ &\geq \alpha(T). \end{aligned} \tag{8}$$

Equality only holds if  $u(\min_{T_m}) + (4n - 2)\epsilon = u(\max_{T_m})$ ,  $u(\min_{R_m}) + (4n - 2)\epsilon = u(\max_{R_m})$ , and  $n = 1$ . This means that  $R = R_1 = T = T_1 = [u(\min_R), u(\min_R) + 2\epsilon]$ . As we see,  $R \cap T(2\epsilon) = [u(\min_R), u(\min_R)] \neq \emptyset$ , so  $\psi(\epsilon) \circ \chi \neq 0$ .  $\square$

We define a function  $\mu$  by

$$\mu(I) = \{J \in B(N) \mid I \text{ and } J \text{ are } (2n - 1)\delta\text{-interleaved}\} \tag{9}$$

for  $I$  in  $B(M)$ . In other words,  $\mu(I)$  contains all the intervals that can be matched with  $I$  in a  $(2n - 1)\delta$ -matching. Let  $I \in B(M)$  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 p.f.d., this means that  $\mu(I)$  is a finite set. For  $A \subset B(M)$ , we write  $\mu(A) = \bigcup_{I \in A} \mu(I)$ .

**Lemma 4.8.** *Let  $A$  be a finite subset of  $B(M)$  containing no  $(4n-2)\delta$ -trivial elements. Then  $|A| \leq |\mu(A)|$ .*

Before we prove Lemma 4.8, we show that it implies that there is a  $(2n-1)\delta$ -matching between  $B(M)$  and  $B(N)$  and thus completes the proof of Theorem 4.2.

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

For a subset  $S$  of a graph  $G$ , let  $A_G(S)$  be the neighbourhood 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 [18] to bridge the gap between Lemma 4.8 and the statement we want to prove about matchings.

**Theorem 4.9** (Hall's theorem). *Let  $G$  be a bipartite graph on bipartite sets  $X$  and  $Y$  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 is sufficient for a matching in  $G$  covering  $X$  to exist.

Letting  $X$  be the set of  $(4n-2)\delta$ -significant intervals in  $B(M)$  and  $Y$  be  $B(N)$ , Hall's theorem and Lemma 4.8 give us a matching  $\sigma$  in the graph  $G_\mu$  covering all the  $(4n-2)\delta$ -significant elements in  $B(M)$ .<sup>4</sup> By symmetry, we also have a matching  $\tau$  in  $G_\mu$  covering all the  $(4n-2)\delta$ -significant elements in  $B(N)$ . 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  $\sigma$  and  $\tau$  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  $B(M) \sqcup B(N)$  for which the set of edges is the union of the edges in the matchings  $\sigma$  and  $\tau$ . Let  $C$  be a connected component of  $H$ . Suppose the submatching of  $\sigma$  in  $C$  does not cover all the  $(4n-2)\delta$ -significant elements of  $C$ . Then there is a  $(4n-2)\delta$ -significant  $J \in C \cap B(N)$  that is not matched by  $\sigma$ . If we view  $\sigma$  and  $\tau$  as partial bijections  $\sigma : B(M) \rightarrow B(N)$  and  $\tau : B(N) \rightarrow B(M)$ , we can write the connected component of  $J$ , which is  $C$ , 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  $\tau$  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  $B(M) \sqcup B(N)$ , so Lemma 4.8 completes the proof of Theorem 4.2.

*Proof of Lemma 4.8.* Because  $\leq_\alpha$  is a preorder, 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  $I \in B(M)$ , we have

$$\begin{aligned}
\phi_{\mathbb{I}I, 2\delta} &= \pi_I(2\delta)g(\delta)f|_I \\
&= \pi_I(2\delta) \left( \sum_{J \in B(N)} g|_J \pi_J \right) (\delta)f|_I \\
&= \sum_{J \in B(N)} \pi_I(2\delta)g|_J(\delta)\pi_J(\delta)f|_I \\
&= \sum_{J \in B(N)} g_{J,I}(\delta)f_{I,J}.
\end{aligned} \tag{10}$$

<sup>4</sup>Strictly speaking, Lemma 4.8 says nothing about infinite  $A$ , but the case with  $A$  countably infinite follows from the finite cases. Each interval in  $A$  contains a rational point, so since  $M$  is p.f.d., the cardinality of  $A$  is at most finite times countably infinite, which is countable. Thus we have covered all the possible cases.

Also,  $\sum_{J \in B(N)} g_{J,I'}(\delta) f_{I,J} = 0$  for  $I \neq I' \in B(M)$ , since  $\phi_{M,2\delta}$  is zero between different components of  $M$ . Lemma 4.6 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

$$\begin{aligned} 0 &= \sum_{J \in B(N)} g_{J,I_{i'}}(\delta) f_{I_i,J} \\ &= \sum_{J \in \mu(A)} g_{J,I_{i'}}(\delta) f_{I_i,J}, \end{aligned} \tag{11}$$

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

$$\begin{aligned} \phi_{1I_i,2\delta} &= \sum_{J \in B(N)} g_{J,I_i}(\delta) f_{I_i,J} \\ &= \sum_{J \in \mu(A)} g_{J,I_i}(\delta) f_{I_i,J}. \end{aligned} \tag{12}$$

Writing this in matrix form, we get

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

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  $\psi$  between rectangle modules, we introduce the notation  $w(\psi) = c$  if  $\psi$  is given by multiplication by  $c$ , and 0 otherwise. A consequence of Lemma 4.7 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  $i \leq i'$ . We get

$$\begin{aligned} 1 &= w(\phi_{1I_i,2\delta}) \\ &= w\left(\sum_{J \in \mu(A)} g_{J,I_i}(\delta) f_{I_i,J}\right) \\ &= \sum_{J \in \mu(A)} w(g_{J,I_i}(\delta) f_{I_i,J}) \\ &= \sum_{J \in \mu(A)} w(J,I_i)w(I_i,J), \end{aligned} \tag{13}$$

and similarly  $0 = \sum_{J \in \mu(A)} w(J,I_{i'})w(I_i,J)$  for  $i \leq i'$ . 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.  $\square$

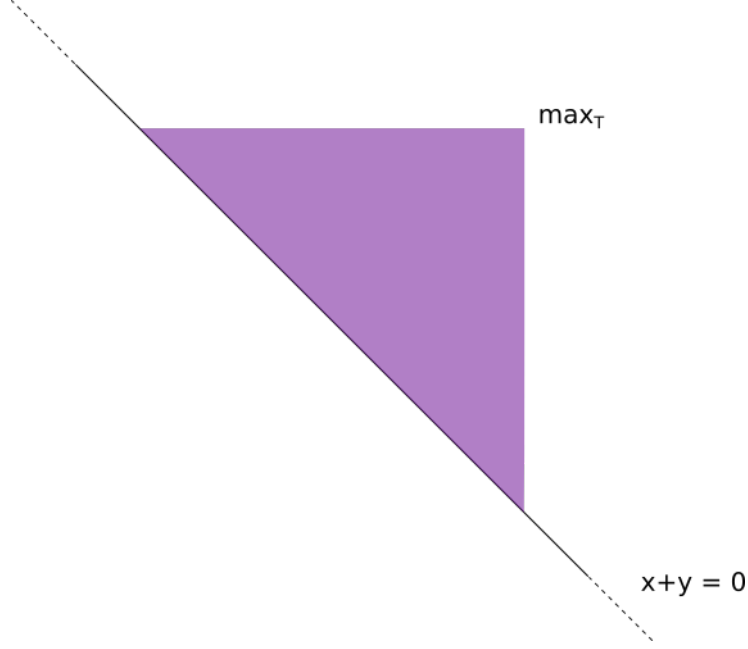


Figure 5: A bounded triangle  $T$ .

## 4.2 Block decomposable modules

Next, we prove stability for block decomposable modules, which, as explained in Section 3, implies stability for zigzag modules and Reeb graphs. Let  $\mathbb{R}_+^2 = \{(x, y) \in \mathbb{R}^2 \mid x + y \geq 0\}$ .

**Definition 4.10.** A **triangle** is a nonempty set  $T$  of the form  $\{(x, y) \in \mathbb{R}^2 \mid x < a, y < b\} \cap \mathbb{R}_+^2$  for some  $(a, b) \in (\mathbb{R} \cup \{\infty\})^2$  with  $a + b > 0$ .

It follows that triangles are intervals. For a triangle  $T = \{(x, y) \in \mathbb{R}^2 \mid x < a, y < b\} \cap \mathbb{R}_+^2$ , we write  $\max_T = (a, b) \in (\mathbb{R} \cup \{\infty\})^2$ . If  $T$  is bounded,  $\max_T$  is the maximal element in the closure of  $T$ , as illustrated in Figure 5. A *triangle decomposable module* is an interval decomposable  $\mathbb{R}^2$ -module whose barcode only contains triangles. Observe that triangles correspond exactly to blocks of the form  $(a, b)_{BL}$  under the poset isomorphism between  $\mathbb{R}^{\text{op}} \times \mathbb{R}$  and  $\mathbb{R}^2$  flipping the  $x$ -axis.

**Theorem 4.11.** Let  $M$  and  $N$  be  $\delta$ -interleaved triangle decomposable modules. Then there is a  $\delta$ -matching between  $B(M)$  and  $B(N)$ .

To prove this, we can split the triangles into sets of different ‘types’, as we did with the rectangles. We get four different types of triangles  $T$ , depending on whether  $\max_T$  is of the form  $(a, b)$ ,  $(\infty, b)$ ,  $(a, \infty)$ , or  $(\infty, \infty)$  for  $a, b \in \mathbb{R}$ . Now a result analogous to Lemma 4.5 holds, implying that it is enough to show Theorem 4.11 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 4.2. For  $I \in B(M)$ , we define  $\nu(I) = \{J \in B(N) \mid I \text{ and } J \text{ are } \delta\text{-interleaved}\}$ . The discussion about Hall’s theorem is still valid, so we only need to prove the analogue of Lemma 4.8 for  $\nu$ . Define  $\alpha(T) = m_1 + m_2$ , where  $\max_T = (m_1, m_2)$ . The only things we need to complete the proof of the analogue of Lemma 4.8 for triangle decomposable modules are the following analogues of Lemmas 4.6 and 4.7:

**Lemma 4.12.** 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  $\epsilon$ -interleaved with either  $\mathbb{I}^R$  or  $\mathbb{I}^T$ .

**Lemma 4.13.** Let  $R, S$ , and  $T$  be triangles with  $T$   $2\epsilon$ -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 4.12.* Suppose  $\mathbb{I}^R$  and  $\mathbb{I}^S$  are not  $\epsilon$ -interleaved. Then  $\max_S \not\geq \max_R - \epsilon$ . But at the same time,  $\max_R \geq \max_S - \epsilon$ , which gives  $\alpha(R) > \alpha(S)$ . Assuming that  $\mathbb{I}^S$  and  $\mathbb{I}^T$  are not  $\epsilon$ -interleaved, either, we also get  $\alpha(S) > \alpha(T)$ . Thus  $\alpha(R) > \alpha(T)$ , a contradiction.  $\square$

*Proof of Lemma 4.13.* For all triangles  $I$ , we treat  $\min_I$  and  $\max_I$  as undecorated points. 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  $2\epsilon$ -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 4.11 implies  $d_B(M, N) = d_I(M, N)$  for block decomposable  $M$  and  $N$  such that  $B(M)$  and  $B(N)$  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 4.16 below 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. In the end we could stitch the cases together by something similar to Lemma 4.5 and the discussion following it. We omit the details, and anyway the closed and half-open cases are taken care of in [3]. Thus, either by appealing to previous work for the other cases or using our own methods, we get

**Theorem 4.14.** *Let  $M$  and  $N$  be block decomposable modules. If  $M$  and  $N$  are  $\delta$ -interleaved, there exists a  $\delta$ -matching between  $B(M)$  and  $B(N)$ .*

### 4.3 Free modules

**Definition 4.15.** *We define a **free interval** as 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  $R = \langle \min_R \rangle$ .<sup>5</sup> 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 4.2 that  $d_B(M, N) \leq (2n - 1)d_I(M, N)$  for free modules  $M, N$ . But because of the geometry of free modules, this result can be strengthened.

**Theorem 4.16.** *Let  $M$  and  $N$  be free  $\delta$ -interleaved  $\mathbb{R}^n$ -modules with  $n \geq 2$ . Then there is a  $(n - 1)\delta$ -matching between  $B(M)$  and  $B(N)$ .*

We already did most of the work while proving Theorem 4.2, and there are some obvious simplifications. Firstly, free intervals are  $\epsilon$ -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  $I \in B(M)$ , define  $\nu(I) = \{J \in B(N) \mid I \text{ and } J \text{ are } (n - 1)\delta\text{-interleaved}\}$ . By the arguments in the proof of Theorem 4.2, we only need to prove Lemma 4.8 with  $\mu$  replaced by  $\nu$ . Lemmas 4.6 and 4.7 still hold for free modules, but we need to sharpen Lemma 4.6.

**Lemma 4.17.** *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.* In this proof, we treat  $\min_I$  and  $\max_I$  as undecorated points for all free intervals  $I$ , so that we can add them. 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

$$\begin{aligned} \alpha(S) &= \sum_{1 \leq i \leq n} \min_{S_i} \\ &< \min_{R_m} - (n - 1)\epsilon + \sum_{i \neq m} (\min_{R_i} + \epsilon) \\ &= \sum_{1 \leq i \leq n} \min_{R_i} \\ &= \alpha(R). \end{aligned} \tag{14}$$

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.  $\square$

<sup>5</sup>This makes  $\min_R$  an undecorated point, while we have previously defined  $\min_-$  as decorated points, but this does not matter, as we will not need decorated points in this subsection.

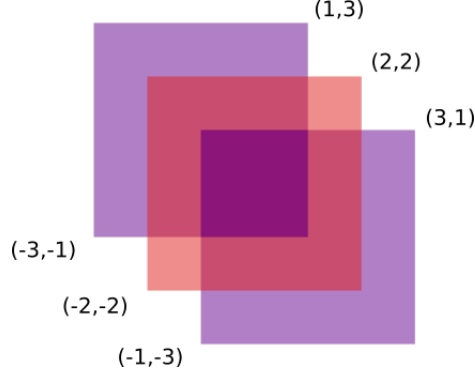


Figure 6:  $M$  and  $N$ .  $I_1$  and  $I_2$  are the light purple squares,  $I_3$  is deep purple, and  $J$  is pink.

## 5 Counterexamples to a general algebraic stability theorem

Theorem 4.2 gives an upper bound of  $(2n - 1)$  on  $d_B/d_I$  for rectangle decomposable modules that increases with the dimension. An obvious question is whether it is possible to improve this constant, or if for each  $C < 2(n - 1)$  there exist pairs  $M, N$  of modules for which  $d_B(M, N) > Cd_I(M, N)$ , in which case the bound 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 confusing when  $n$  increases. In dimension 2, we give an example of rectangle decomposable modules  $M$  and  $N$  with  $d_B(M, N) = 3d_I(M, N)$  in Example 5.2, which means that the bound is optimal for  $n = 2$ , as well. This is a counterexample to a conjecture made in a previous version of [3] which claims that interval decomposable  $\mathbb{R}^n$ -modules  $M$  and  $N$  such that  $B(M)$  and  $B(N)$  only contain convex intervals are  $\epsilon$ -matched if they are  $\epsilon$ -interleaved.

**Example 5.1.** Let  $B(M) = \{I_1, I_2, I_3\}$ <sup>6</sup> and  $B(N) = \{J\}$ , where

- $I_1 = (-3, 1) \times (-1, 3)$
- $I_2 = (-1, 3) \times (-3, 1)$
- $I_3 = (-1, 1) \times (-1, 1)$
- $J = (-2, 2) \times (-2, 2)$ .

See Figure 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 4.2. On the other hand, in any matching between  $B(M)$  and  $B(N)$  we have to leave either  $I_1$  or  $I_2$  unmatched, and they are  $\epsilon$ -significant for all  $\epsilon < 4$ . In fact, any possible matching between  $B(M)$  and  $B(N)$  is a 2-matching. Thus  $d_I(M, N) = 1$  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 no longer would be 2-significant (see Lemma 4.7), and then they would not need to be matched 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  $d_B(M, N) = 3d_I(M, N)$ , proving that our upper bound of  $2(n - 1)$  is the best possible for  $n = 2$ .

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

- $I_1 = (0, 10) \times (1, 11)$

<sup>6</sup>Here we use subscripts to index different intervals, not to indicate projections, as we did earlier.



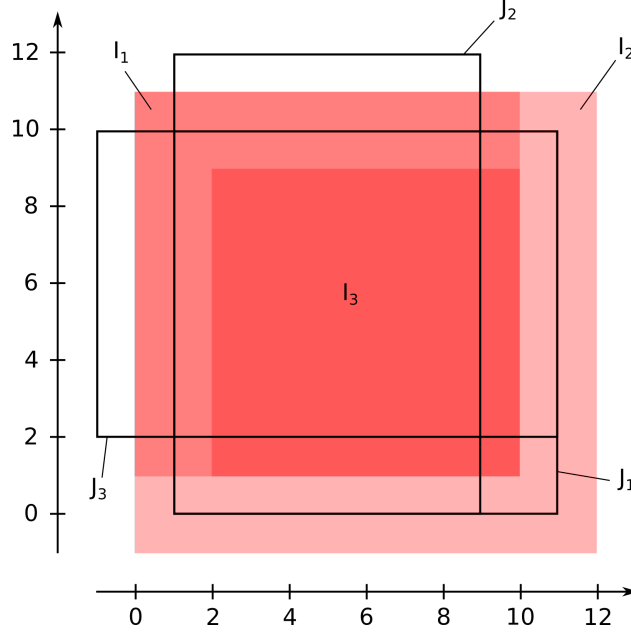


Figure 7:  $I_1$ ,  $I_2$ , and  $I_3$  are the filled pink rectangles, and  $J_1$ ,  $J_2$ , and  $J_3$  are the black rectangles without fill.

- $I_2 = (0, 12) \times (-1, 11)$
- $I_3 = (2, 10) \times (1, 9)$
- $J_1 = (1, 11) \times (0, 10)$
- $J_2 = (1, 9) \times (0, 12)$
- $J_3 = (-1, 11) \times (2, 10)$ .

The rectangles in  $B(M)$  and  $B(N)$  are shown in Figure 7.

We give an example of 1-interleaving morphisms  $f$  and  $g$  that we write on 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  $i$ .

$$f : \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad g : \begin{bmatrix} -1 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}. \quad (15)$$

This means that  $M$  and  $N$  are 1-interleaved, but they are not  $\epsilon$ -interleaved for any  $\epsilon < 1$ , so  $d_I(M, N) = 1$ .

Let  $\epsilon < 3$ . We see that 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  $\epsilon$ -interleaved, and thus they cannot be matched in an  $\epsilon$ -matching. In fact,  $I_i$  and  $J_j$  cannot be matched in an  $\epsilon$ -matching for any  $i, j \in \{2, 3\}$  by similar arguments. Since  $I_2$  and  $I_3$  cannot both be matched with  $J_1$ , one of them has to be left unmatched, but since both  $I_2$  and  $I_3$  are 6-significant, this means that there is no  $\epsilon$ -matching between  $B(M)$  and  $B(N)$ . On the other hand, any bijection between  $B(M)$  and  $B(N)$  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  $B(M) = \{I_1, I_2, I_3\}$  and  $B(N) = \{J_1, J_2, J_3\}$ , where

- $I_1 = \langle(0, 1, 10, 11)\rangle$
- $I_2 = \langle(0, -1, 12, 11)\rangle$
- $I_3 = \langle(2, 1, 10, 9)\rangle$
- $J_1 = \langle(1, 0, 11, 10)\rangle$
- $J_2 = \langle(1, 0, 9, 12)\rangle$
- $J_3 = \langle(-1, 2, 11, 10)\rangle$ .

(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  $B(M)$  and  $B(N)$ , so  $d_B(M, N) = 3$  and  $d_I(M, N) = 1$ .

As a consequence of this example, we get that our upper bound of  $d_B/d_I \leq n - 1$  for free  $n$ -dimensional modules cannot be improved on 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 find, as naively trying to construct interleaving morphisms can quickly lead you to consider 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  $d_I$ : decompose the modules into intervals and find the bottleneck distance. Since  $d_I = d_B$ , this gives us the interleaving distance. When it exists, one can compute the bottleneck distance in polynomial time also in two dimensions [17], 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 [9], 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 *CI problems* 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,<sup>7</sup> and 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 [8], opened the door to proving NP-hardness of calculating  $d_I$  by the approach used in [9].

In [9], 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 Theorem 4.2. It can be shown that if  $d_B(M, N) \leq cd_I(M, N)$  for any pair  $M, N$  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.

<sup>7</sup>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.

## 7 Acknowledgements

I would like to thank my supervisors Gereon Quick and Nils Baas for invaluable support and help. I would also like to thank Peter Landweber for detailed comments on several drafts of this text, Steve Oudot for feedback on the first arXiv version and Magnus Bakke Botnan for interesting discussions.

## References

- [1] Martin Aigner and Günter M. Ziegler. *Proofs from THE BOOK*. Springer, Berlin, 4th edition, 2010.
- [2] Gorô Azumaya. Corrections and supplementaries to my paper concerning Krull-Remak-Schmidt's theorem. *Nagoya Mathematical Journal*, 1:117–124, 1950.
- [3] Magnus Bakke Botnan and Michael Lesnick. Algebraic stability of zigzag persistence modules. *Algebraic & Geometric Topology*, 18, 04 2016.
- [4] Ulrich Bauer, Xiaoyin Ge, and Yusu Wang. Measuring distance between reeb graphs. In *Proceedings of the thirtieth annual symposium on Computational geometry*, page 464. ACM, 2014.
- [5] Ulrich Bauer and Michael Lesnick. Induced matchings and the algebraic stability of persistence barcodes. *Journal of Computational Geometry*, 6(2):162–191, 2015.
- [6] Ulrich Bauer, Elizabeth Munch, and Yusu Wang. Strong Equivalence of the Interleaving and Functional Distortion Metrics for Reeb Graphs. In *31st International Symposium on Computational Geometry (SoCG 2015)*, volume 34 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 461–475. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2015.
- [7] Paul Bendich, Herbert Edelsbrunner, Dmitriy Morozov, Amit Patel, et al. Homology and robustness of level and interlevel sets. *Homology, Homotopy and Applications*, 15(1):51–72, 2013.
- [8] Håvard Bakke Bjerkevik. Stability of higher-dimensional interval decomposable persistence modules. *arXiv preprint arXiv:1609.02086v2*, 2016.
- [9] Håvard Bakke Bjerkevik, Magnus Bakke Botnan, and Michael Kerber. Computing the interleaving distance is np-hard. *arXiv preprint arXiv:1811.09165*, 2018.
- [10] Gunnar Carlsson and Vin De Silva. Zigzag persistence. *Foundations of computational mathematics*, 10(4):367–405, 2010.
- [11] Gunnar Carlsson, Vin De Silva, and Dmitriy Morozov. Zigzag persistent homology and real-valued functions. In *Proceedings of the twenty-fifth annual symposium on Computational geometry*, pages 247–256. ACM, 2009.
- [12] Frédéric Chazal, David Cohen-Steiner, Marc Glisse, Leonidas J Guibas, and Steve Y Oudot. Proximity of persistence modules and their diagrams. In *Proceedings of the twenty-fifth annual symposium on Computational geometry*, pages 237–246. ACM, 2009.
- [13] Frédéric Chazal, Vin De Silva, Marc Glisse, and Steve Oudot. *The structure and stability of persistence modules*. Springer, 2016.
- [14] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. *Discrete & Computational Geometry*, 37(1):103–120, 2007.
- [15] William Crawley-Boevey. Decomposition of pointwise finite-dimensional persistence modules. *Journal of Algebra and Its Applications*, 14(05):1550066, 2015.
- [16] Vin De Silva, Elizabeth Munch, and Amit Patel. Categorized reeb graphs. *Discrete & Computational Geometry*, 55(4):854–906, 2016.

- [17] Tamal K. Dey and Cheng Xin. Computing Bottleneck Distance for 2-D Interval Decomposable Modules. In *34th International Symposium on Computational Geometry (SoCG 2018)*, volume 99 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 32:1–32:15. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2018.
- [18] Philip Hall. On representatives of subsets. *J. London Math. Soc*, 10(1):26–30, 1935.
- [19] Woojin Kim and Facundo Memoli. Stable signatures for dynamic metric spaces via zigzag persistent homology. *arXiv preprint arXiv:1712.04064*, 2017.
- [20] Steve Y Oudot and Donald R Sheehy. Zigzag zoology: Rips zigzags for homology inference. *Foundations of Computational Mathematics*, 15(5):1151–1186, 2015.
- [21] Cary Webb. Decomposition of graded modules. *Proceedings of the American Mathematical Society*, 94(4):565–571, 1985.