Computational Complexity of the Interleaving Distance

Håvard Bakke Bjerkevik
Department of Mathematical Sciences, Norwegian University of Science and Technology
Trondheim, Norway
havard.bjerkevik@ntnu.no

Magnus Bakke Botnan[1]
Department of Mathematics, TU Munich
Garching bei München, Germany
botnan@ma.tum.de

Abstract

The interleaving distance is arguably the most prominent distance measure in topological data analysis. In this paper, we provide bounds on the computational complexity of determining the interleaving distance in several settings. We show that the interleaving distance is NP-hard to compute for persistence modules valued in the category of vector spaces. In the specific setting of multidimensional persistent homology we show that the problem is at least as hard as a matrix invertibility problem. Furthermore, this allows us to conclude that the interleaving distance of interval decomposable modules depends on the characteristic of the field. Persistence modules valued in the category of sets are also studied. As a corollary, we obtain that the isomorphism problem for Reeb graphs is graph isomorphism complete.

2012 ACM Subject Classification Theory of computation → Problems, reductions and completeness

Keywords and phrases Persistent Homology, Interleavings, NP-hard

Digital Object Identifier 10.4230/LIPIcs.SoCG.2018.13

Related Version A full version of this paper is available at `https://arxiv.org/abs/1712.04281`

Acknowledgements This work was partially carried out while the authors were visitors to the Hausdorff Center for Mathematics, Bonn, during the special Hausdorff program on applied and computational topology. We thank anonymous referees for insightful comments.

1 Introduction

For a category $\mathcal{C}$ and a poset $\mathbf{P}$ we define a $\mathbf{P}$*-indexed (persistence) module valued in* $\mathcal{C}$ to a be a functor $M : \mathbf{P} \to \mathcal{C}$. We will denote the associated functor category by $\mathcal{C}^{\mathbf{P}}$. If $M, N \in \mathcal{C}^{\mathbf{P}}$ then M and N are of the same *type*. Such functors appear naturally in applications, and most commonly when $\mathbf{P} = \mathbf{R}^n$, n-tuples of real numbers under the normal product order, and $\mathcal{C} = \mathbf{Vec}_{\mathbb{K}}$, the category of vector spaces over the field $\mathbb{K}$, or $\mathcal{C} = \mathbf{Set}$, the category of sets. The field $\mathbb{K}$ is assumed to be finite. We suppress notation and simply write $\mathbf{Vec}$ when $\mathbb{K}$ is an arbitrary finite field. The notation $p \in \mathbf{P}$ denotes that p is an object of $\mathbf{P}$.

[1] DFG Collaborative Research Center SFB/TR 109 “Discretization in Geometry and Dynamics”

34th International Symposium on Computational Geometry (SoCG 2018).
Editors: Bettina Speckmann and Csaba D. Tóth; Article No. 13; pp. 13:1–13:15
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

▶ Remark. Throughout the paper we make use of basic concepts from category theory. The reader unfamiliar to such ideas will find the necessary background material in the first few pages of [13].

Assume that $h : X \to \mathbb{R}$ is a continuous function of "Morse type", a generalization of a Morse function on a compact manifold. Roughly, a real-valued function is of Morse type if the homotopy type of the fibers changes at finite set of values; see [16] for a precise definition. We shall now briefly review four different scenarios in which functors of the aforementioned form can be associated to h.

Let $\mathrm{H}_p : \mathbf{Top} \to \mathbf{Vec}_{\mathbb{K}}$ denote the p-th singular homology functor with coefficients in $\mathbb{K}$, and let $\pi_0 : \mathbf{Top} \to \mathbf{Set}$ denote the functor giving the set of path-components. We also associate the two following functors to h whose actions on morphisms are given by inclusions:

$$\mathcal{S}^{\uparrow}(h) : \mathbf{R} \to \mathbf{Top} \qquad\qquad \mathcal{S}(h) : \mathbf{R}^2 \to \mathbf{Top}$$

$$\mathcal{S}^{\uparrow}(h)(t) = \{x \in X \mid h(x) \leq t\} \qquad\qquad \mathcal{S}(h)(-s,t) = \{x \in X \mid s \leq h(x) \leq t\}$$

- *Persistent Homology* studies the evolution of the homology of the sublevel sets of h and is perhaps the most prominent tool in topological data analysis [16]. Specifically, the *p-th sublevel set persistence module associated to h* is the functor $\mathrm{H}_p\mathcal{S}^{\uparrow}(h) : \mathbf{R} \to \mathbf{Vec}$. Importantly, such a module is completely determined by a collection of intervals $\mathcal{B}(\mathrm{H}_p\mathcal{S}^{\uparrow}(h))$ called the *barcode* of $\mathrm{H}_p\mathcal{S}^{\uparrow}(h)$. This collection of intervals is then in turn used to extract topological information from the data at hand. In Fig. 1 we show the associated barcode for $p = 0$ and $p = 1$ for a function of Morse type.
- Upon replacing H_p by π_0 in the above construction we get a *merge tree*. That is, the *merge tree associated to h* is the functor $\tau^h : \pi_0\mathcal{S}^{\uparrow}(h) : \mathbf{R} \to \mathbf{Set}$. A merge tree captures the evolution of the path components of the sublevel sets of h and can be, as the name indicates, be visualized as (a disjoint union of) rooted trees. See Fig. 1 for an example.
- The two aforementioned examples used sublevel sets. A richer invariant is obtained by considering interlevel sets: define the *p-th interlevel set persistence of h* to be the functor $\mathrm{H}_p\mathcal{S}(h) : \mathbf{R}^2 \to \mathbf{Vec}$. Analogously to above, such a module is completely determined by a collection $\mathcal{B}(\mathrm{H}_p\mathcal{S}(h))$ of simple regions in $\mathbb{R}^2$. However, it is often the collection of intervals $\mathcal{L}_p(h)$ obtained by the intersection of these regions with the anti-diagonal $y = -x$ which are used in data analysis. We refer the reader to [9] for an in-depth treatment. In Fig. 1 we show an example of the 0-th interlevel set barcode. Observe how the endpoints of the intervals correspond to different types of features of the Reeb graph.
- Just as interlevel set persistence is a richer invariant than sublevel set persistence, the *Reeb graph* is richer in structure than the merge tree. Specifically, we define the functor $\mathrm{Reeb}^h := \pi_0\mathcal{S}(h) : \mathbf{R}^2 \to \mathbf{Set}$. Just as for Merge trees, Reeb^h admits a visualization of a graph; see Fig. 1. In particular, this appealing representation has made Reeb graphs a popular objects of study in computational geometry and topology, and they have found many applications in data visualization and exploratory data analysis.

These are all examples of topological invariants arising from a single real-valued function. There are many settings for which it is more fruitful to combine a collection of real-valued functions into a single function $g : X \to \mathbb{R}^n$ [14]. By combining them into a single function we not only learn how the data looks from the point of view of each function (i.e. a type of measurement) but how the different functions (measurements) interact. One obvious way to assign a (algebraic) topological invariant to g is to filter it by sublevel sets. That is, define $\mathcal{S}^{\uparrow}(g) : \mathbf{R}^n \to \mathbf{Top}$ by $\mathcal{S}^{\uparrow}(g)(t) = \{x \in X \mid g(x) \leq t\}$. The associated functor $\mathrm{H}_p\mathcal{S}^{\uparrow}(g) : \mathbf{R}^n \to \mathbf{Vec}$ is an example of an *n-dimensional persistence module*. We saw above

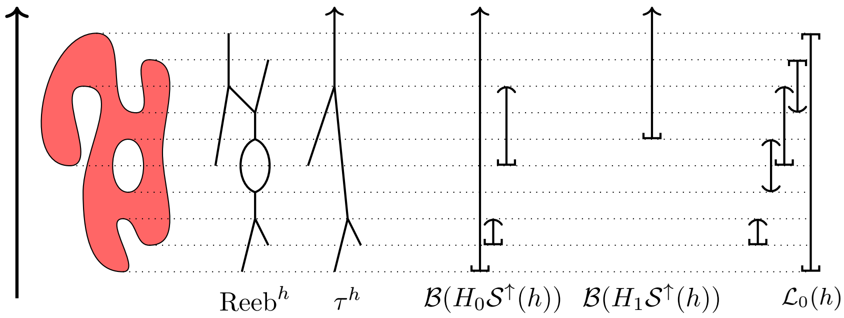


Figure 1 The height function of the solid shape is of Morse type. The associated Reeb graph, merge tree, sublevel set barcodes, and interlevel set barcode are shown to the right.

that for $n = 1$ this functor is completely described by a collection of intervals. This is far from true for $n \geq 2$: there exists no way to describe such functors by interval-like regions in higher-dimensional Euclidean space. Even the task of parameterizing such (indecomposable) modules is known to be a *hopeless* problem (so-called *wild representation type*) [3].

1.1 The interleaving distance

Different types of distances have been proposed on various types of persistence modules with values in **Vec** [26, 24, 17, 12, 5]. Of all these, the *interleaving distance* is arguably the most prominent for the following reasons: the theory of interleavings lies at the core of the theoretical foundations of 1-dimensional persistence, notably through the Isometry Theorem (Theorem 7). Furthermore, it was shown by Lesnick that when $\mathbb{K}$ is a prime field, the interleaving distance is the *most discriminative* of all *stable* metrics on such modules. We refer to [24] for the precise statement. As we shall see, it is also an immediate consequence of Theorem 7 that the interleaving distance for 1-dimensional persistence modules can be computed in polynomial time.

Lesnick's result generalizes to n-dimensional persistence modules, but the computational complexity of computing the interleaving distance of such modules remains unknown. An efficient algorithm to compute the interleaving distance could carry a profound impact on topological data analysis: the standard pipeline for 1-dimensional persistent homology is to first compute the barcode and then perform analysis on the collection of intervals. However, for multi-dimensional persistence there is no way of defining the barcode. With an efficient algorithm for computing the interleaving distance at hand it would still not be clear how to analyze the persistence modules individually, but we would have a theoretical optimal way of comparing them. This in turn could be used in clustering, kernel methods, and other kinds of data analysis widely applied in the 1-dimensional setting.

Complexity

The purpose of this paper is to determine the computational complexity of computing the interleaving distance. To make this precise, we need to associate a notion of size to the persistence modules.

▶ **Definition 1.** Let $\mathbf{P}$ denote a poset category and $M : \mathbf{P} \to \mathcal{C}$.

- For $\mathcal{C} = \mathbf{Vec}$, define the *total dimension of* M to be $\dim M = \sum_{p\in\mathbf{P}} \dim M_p$.
- For $M : \mathbf{Z} \to \mathbf{Set}$, define the *total cardinality of* M to be $|M| = \sum_{p\in\mathbf{P}} |M_p|$.

The input size will be the total dimension or the total cardinality and for the the remaining of the paper **we shall always assume** that those quantities are finite. The following shows that there exists an algorithm, polynomial in the input size, which determines whether or not two $\mathbf{P}$-indexed modules valued in $\mathbf{Vec}$ are isomorphic.

▶ **Theorem 2** ([10]). *Let $\mathbf{P}$ be a finite poset and $M, M' : \mathbf{P} \to \mathbf{Vec}$. There exists a deterministic algorithm which decides if $M \cong M'$ in $\mathcal{O}\left((\dim M + \dim M')^6\right)$.*

This result will be important to us in what ensues because the strongest of interleavings, the 0-interleaving, is by definition a pair of inverse isomorphisms. Furthermore, by choosing an appropriate basis for each vector space, an isomorphism between M and M' is nothing more than a collection of matrices with entries in a finite field. Likewise a δ-interleaving will be nothing more than a collection of matrices over a finite field satisfying certain constraints. When $\mathcal{C} = \mathbf{Set}$ the morphisms are specified by collections of functions between finite sets. Hence, the decision problems considered in this paper **are trivially in NP**.

Furthermore, it is an immediate property of the Morse type of h, that the modules considered above are *discrete*. Intuitively, we say that an $\mathbf{R}^n$-indexed persistence module M is discrete if there exists a $\mathbf{Z}^n$-indexed persistence module containing all the information of M; see [7]. In practice, persistence modules arising from data will be discrete. Hence, when it comes to algorithmic questions we shall restrict ourselves to the setting in which $\mathbf{P} = \mathbf{Z}^n$ or a slight generalization thereof. Importantly, the modules considered in this paper can be δ-interleaved only for $\delta \in \{0, 1, 2, \ldots\}$.

Contributions

The contributions of this paper are summarized in Table 1. Concretely, a cell in Table 1 gives a complexity bound on the decision problem of deciding if two modules of the given type are δ-interleaved. It is an easy consequence of the definition of the interleaving distance that this is at least as hard as determining the distance itself. The cells with a shaded background indicate that novel contributions to that complexity bound is provided in this paper. Recall that we have defined the input size to be $n = \dim M + \dim M'$ when the modules are valued in $\mathbf{Vec}$, and $n = |M| + |M'|$ when the modules are valued in $\mathbf{Set}$. Observe that any non-trivial functor $M : \mathbf{Z}^m \to \mathbf{Set}$ must have $|M| = \infty$. Hence, when we talk about interleavings of such functors, we shall assume that they are completely determined by a restriction to a finite sub-grid. The input size is then the total cardinalities of the restrictions. We will now give a brief summary of the cells of Table 1.

- $\mathbf{Z} \to \mathbf{Vec}$. $[\delta \geq 0]$ This bound is achieved by first determining the barcodes of the persistence modules and then using Theorem 7 to obtain the interleaving distance. The complexity of this is $\mathcal{O}(\text{FindBarcode} + \text{Match}) = \mathcal{O}(n^\omega + n^{1.5} \log n) = \mathcal{O}(n^\omega)$ where ω is the matrix multiplication exponent[22]. The details can be found in [7]. In [25], the complexity is shown to be $\mathcal{O}(n^\omega + n^2 \log^2 n)$ for essentially the same problem, but with a slightly different input size n.
- $\mathbf{Z} \to \mathbf{Set}$. $[\delta = 0]$ Essentially isomorphism of rooted trees; see [7]. $[\delta \geq 1]$ This follows from arguments in [1].
- $\mathbf{Z}^2 \to \mathbf{Vec}$. $[\delta = 0]$ This is Theorem 2 for $\mathbf{P} = \mathbf{Z}^2$. $[\delta \geq 1]$ A *constrained invertibility* (CI) problem is a triple (P, Q, n) where P and Q are subsets of $\{1, 2, \ldots, n\}^2$. We say that a CI-problem (P, Q, n) is *solvable* if there exists an invertible $n \times n$ matrix M such that $M_{i,j} = 0$ for all $(i, j) \in P$ and $M^{-1}_{i',j'} = 0$ for all $(i', j') \in Q$. We call (M, M^{-1}) a *solution* of (P, Q, n). In Section 4 we show that a CI-problem is solvable if and only if an associated pair of $\mathbf{Z}^2$-indexed modules is 1-interleaved. Thus, the interleaving problem is *constrained invertibility-hard* (CI-hard).

Table 1 The complexity of checking for δ-interleavings between modules M and M'. If the target category is **Vec** then $n = \dim M + \dim M'$, and if the target category is **Set** then $n = |M| + |M'|$. Here ω is the matrix multiplication exponent.

type/δ	$\mathbf{Z} \to \mathbf{Vec}$	$\mathbf{Z} \to \mathbf{Set}$	$\mathbf{Z}^2 \to \mathbf{Vec}$	$\mathbf{Z}^2 \to \mathbf{Set}$	$\mathbf{Z}^{L,C} \to \mathbf{Vec}_{\mathbb{Z}/2\mathbb{Z}}$
$\delta = 0$	$\mathcal{O}(n^\omega)$	$\mathcal{O}(n)$	$\mathcal{O}(n^6)$	GI-complete	$\mathcal{O}(n^6)$
$\delta \geq 1$	$\mathcal{O}(n^\omega)$	NP-complete	CI-hard	NP-complete	NP-complete

- $\mathbf{Z}^2 \to \mathbf{Set}$. [$\delta = 0$] Reeb graphs are a particular type of functors $\mathbf{Z}^2 \to \mathbf{Set}$ and deciding if two Reeb graphs are isomorphic is graph isomorphism-hard (GI-hard) [20]. In [7] we strengthen this result by showing that the isomorphism problem for $\mathbf{Z}^2 \to \mathbf{Set}$ is in fact GI-*complete*. This also implies that Reeb graph isomorphism is GI-complete. [$\delta \geq 1$] This follows from $\mathbf{Z} \to \mathbf{Set}$.
- $\mathbf{Z}^{L,C} \to \mathbf{Vec}_{\mathbb{Z}/2\mathbb{Z}}$. For two sets L and C, define $\mathbf{Z}^{L\to C}$ to be the poset generated by the following disjoint union of posets $\mathbf{Z}^{L\to C} := \bigsqcup_{l\in L, c\in C} \mathbf{Z}$ with the added relation $(l,t) < (c,t)$ for every $l \in L$, $c \in C$ and $t \geq 3$. This poset is a mild generalization of a disjoint union of $\mathbf{Z}$'s. [$\delta = 0$] Immediate from Theorem 2. [$\delta \geq 1$] Follows from a reduction from 3-SAT; see Section 3. This shows that computing the generalized interleaving distance of [12] for **Vec**-valued persistence modules is NP-complete in general.

2 Preliminaries

For $\mathbf{P}$ a poset and $\mathcal{C}$ an arbitrary category, $M : \mathbf{P} \to \mathcal{C}$ a functor, and $a, b \in \mathbf{P}$, let $M_a = M(a)$, and let $\varphi_M(a,b) : M_a \to M_b$ denote the morphism $M(a \leq b)$.

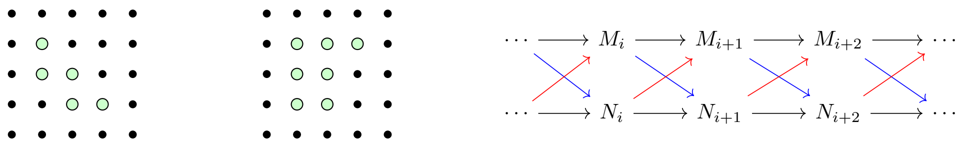
2.1 Interleavings

In this section we review the theory of interleavings for $\mathbf{Z}^n$-indexed modules. For a treatment of the $\mathbf{R}^n$-indexed setting see [24]. For a discussion on interleavings over arbitrary posets see [12].

For $u \in \mathbf{Z}^n$, define the *u-shift functor* $(-)(u) : \mathcal{C}^{\mathbf{Z}^n} \to \mathcal{C}^{\mathbf{Z}^n}$ on objects by $M(u)_a = M_{u+a}$, together with the obvious internal morphisms, and on morphisms $f : M \to N$ by $f(u)_a = f(u+a) : M(u)_a \to N(u)_a$. For $u \in \{0,1,\dots\}^n$, let $\varphi_M^u : M \to M(u)$ be the morphism whose restriction to each M_a is the linear map $\varphi_M(a, a+u)$. For $\delta \in \{0,1,2\dots\}$ we will abuse notation slightly by letting $(-)(\delta)$ denote the $\delta(1,\dots,1)$-shift functor, and letting φ_M^δ denote $\varphi_M^{\delta(1,\dots,1)}$.

▶ **Definition 3.** Given $\delta \in \{0,1,\dots\}$, a δ-interleaving between $M, N : \mathbf{Z}^n \to \mathcal{C}$ is a pair of morphisms $f : M \to N(\delta)$ and $g : N \to M(\delta)$ such that $g(\delta) \circ f = \varphi_M^{2\delta}$ and $f(\delta) \circ g = \varphi_N^{2\delta}$.

We call f and g *δ-interleaving morphisms*. If there exists a δ-interleaving between M and N, we say M and N are δ-interleaved. The *interleaving distance* $d_I : \mathrm{Ob}(\mathcal{C}^{\mathbf{Z}^n}) \times \mathrm{Ob}(\mathcal{C}^{\mathbf{Z}^n}) \to [0,\infty]$ is given by $d_I(M,N) = \min\{\delta \in \{0,1,\dots\} \mid M \text{ and } N \text{ are } \delta\text{-interleaved}\}$. Here we set $d_I(M,N) = \infty$ if there does not exist a δ-interleaving for any δ.



(a) An interval. **(b)** Not an interval. **(c)** A 1-interleaving.

Figure 2 (a) is an interval in $\mathbf{Z}^2$ whereas (b) is not. (c) The persistence modules M and N are 1-interleaved if and only if there exist diagonal morphisms such that the diagram in (c) commutes.

2.2 Interval modules and the Isometry Theorem

Let $\mathcal{C} = \mathbf{Vec}$. An *interval* of a poset $\mathbf{P}$ is a subset $\mathcal{J} \subset \mathbf{P}$ such that

1. $\mathcal{J}$ is non-empty.
2. If $a, c \in \mathcal{J}$ and $a \leq b \leq c$, then $b \in \mathcal{J}$.
3. [connectivity] For any $a, c \in \mathcal{J}$, there is a sequence $a = b_0, b_1, \ldots, b_l = c$ of elements of $\mathcal{J}$ with b_i and b_{i+1} comparable for $0 \leq i \leq l-1$.

We refer to a collection of intervals in $\mathbf{P}$ as a *barcode (over* $\mathbf{P}$*)*.

▶ **Definition 4.** For $\mathcal{J}$ an interval in $\mathbf{P}$, the interval module $I^{\mathcal{J}}$ is the $\mathbf{P}$-indexed module such that

$$I^{\mathcal{J}}_a = \begin{cases} \mathbb{K} & \text{if } a \in \mathcal{J}, \\ 0 & \text{otherwise.} \end{cases} \qquad \varphi_{I^{\mathcal{J}}}(a,b) = \begin{cases} \mathrm{id}_{\mathbb{K}} & \text{if } a \leq b \in I, \\ 0 & \text{otherwise.} \end{cases}$$

We say a persistence module M is *decomposable* if it can be written as $M \cong V \oplus W$ for non-trivial persistence modules V and W; otherwise, we say that M is *indecomposable*.

A $\mathbf{P}$-indexed module M is *interval decomposable* if there exists a collection $\mathcal{B}(M)$ of intervals in $\mathbf{P}$ such that $M \cong \bigoplus_{\mathcal{J} \in \mathcal{B}(M)} I^{\mathcal{J}}$. We call $\mathcal{B}(M)$ the *barcode* of M. This is well-defined by the Azumaya–Krull–Remak–Schmidt theorem [2].

▶ **Theorem 5** (Structure of 1-D Modules [19, 27])**.** *Suppose* $M : \mathbf{P} \to \mathbf{Vec}$ *for* $\mathbf{P} \in \{\mathbf{R}, \mathbf{Z}\}$ *and* $\dim M_p < \infty$ *for all* $p \in \mathbf{P}$*. Then* M *is interval decomposable.*

▶ Remark. Such a decomposition theorem exists only for *very* special choices of $\mathbf{P}$. Two other scenarios appearing in applications are *zigzags* [8, 15] and exact bimodules [18]. The latter is a specific type of $\mathbf{R}^2$-indexed persistence modules.

▶ **Corollary 6.** *Let* $\mathbf{P} = \bigsqcup_{i \in \Lambda} \mathbf{Z}$ *be the poset given as a disjoint union of* $\mathbf{Z}$*'s (i.e. elements in different components are incomparable). If* $M : \mathbf{P} \to \mathbf{Vec}$ *satisfies* $\dim M_p < \infty$ *for all* $p \in \mathbf{P}$*, then* M *is interval decomposable.*

Proof. Apply Theorem 5 to each of the components of $\mathbf{P}$ independently. This gives $\mathcal{B}(M) = \bigsqcup_{i \in \Lambda} \mathcal{B}(M|_{(i,\mathbf{Z})})$. ◀

At the very core of topological data analysis are the *isometry theorems*. They say that for certain choices of interval decomposable modules, the interleaving distance coincides with a completely combinatorial distance on their associated barcodes. This combinatorial distance d_B is called the *bottleneck distance* and is defined in [7]. Importantly, for any two barcodes, if the interleaving distance between each pair of interval modules in the barcodes is known, the associated bottleneck distance can be computed by solving a bipartite matching problem. This, in turn, implies that the interleaving distance can be efficiently computed whenever an isometry theorem holds. See [7] for an example.

▶ **Theorem 7** (Isometry Theorem [24, 17, 4, 6]). *Suppose* $M, N : \mathbf{Z} \to \mathbf{Vec}$ *satisfy* $\dim M_i < \infty$ *and* $\dim N_i < \infty$ *for all* $i \in \mathbf{Z}$. *Then* $d_I(M,N) = d_B(\mathcal{B}(M), \mathcal{B}(N))$.

▶ Remark. Continuing on the remark to Theorem 5. An isometry theorem also holds for zigzags and exact bimodules [9, 6]. Although there might be other classes of interval decomposable modules for which an isometry theorem holds, the result is not true in general. See [7] for an example of interval decomposable modules in $\mathbf{Z}^2$ for which $2d_I(M,N) = d_B(\mathcal{B}(M), \mathcal{B}(N))$, and see [9] for a general conjecture. This shows that a matching of the barcodes will not determine the interleaving distance even in the case of very well-behaved modules.

3 NP-completeness

In this section we shall prove that it is NP-hard to decide if two modules $M, N \in \mathbf{Vec}^{\mathbf{Z}^{L\to C}}$ are 1-interleaved. Recall that for two sets L and C, we define $\mathbf{Z}^{L\to C}$ to be the disjoint union $\bigsqcup_{l\in L, c\in C} \mathbf{Z}$ with the added relations $(l,t) < (c,t)$ for all $l \in L, c \in C$, and $t \geq 3$. Define the *u-shift functor* $(-)(u) : \mathcal{C}^{\mathbf{Z}^{L\to C}} \to \mathcal{C}^{\mathbf{Z}^{L\to C}}$ on objects by $M(u)_{(p,t)} = M_{(p,t+u)}$, together with the obvious internal morphisms, and on morphisms $f : M \to N$ by $f(u)_{(p,t)} = f_{(p,t+u)} : M(u)_{(p,t)} \to N(u)_{(p,t)}$. That is, the shift functor simply acts on each of the components independently. With the shift-functor defined, we define a δ-interleaving of $\mathbf{Z}^{L\to C}$-indexed modules precisely as in Section 2.1. Thus, we see that a δ-interleaving is simply a collection of δ-interleavings over each disjoint component of $\mathbf{Z}$ which satisfy the added relations. Indeed, a 1-interleaving is equivalent to the existence of dashed morphisms in the following diagram for all $l \in L$ and $c \in C$:

$$
\begin{tikzcd}
\vdots & \vdots & \vdots & \vdots \\
M_{(l,5)} \arrow[u] \arrow[ur, dashed] \arrow[rr, bend left] & N_{(l,5)} \arrow[u] \arrow[ul, dashed] \arrow[rr, bend left] & M_{(c,5)} \arrow[u] \arrow[ur, dashed] & N_{(c,5)} \arrow[u] \arrow[ul, dashed] \\
M_{(l,4)} \arrow[u] \arrow[ur, dashed] \arrow[rr, bend left] & N_{(l,4)} \arrow[u] \arrow[ul, dashed] \arrow[rr, bend left] & M_{(c,4)} \arrow[u] \arrow[ur, dashed] & N_{(c,4)} \arrow[u] \arrow[ul, dashed] \\
M_{(l,3)} \arrow[u] \arrow[ur, dashed] \arrow[rr, bend left] & N_{(l,3)} \arrow[u] \arrow[ul, dashed] \arrow[rr, bend left] & M_{(c,3)} \arrow[u] \arrow[ur, dashed] & N_{(c,3)} \arrow[u] \arrow[ul, dashed] \\
M_{(l,2)} \arrow[u] \arrow[ur, dashed] & N_{(l,2)} \arrow[u] \arrow[ul, dashed] & M_{(c,2)} \arrow[u] \arrow[ur, dashed] & N_{(c,2)} \arrow[u] \arrow[ul, dashed] \\
\vdots \arrow[u] \arrow[ur, dashed] & \vdots \arrow[u] \arrow[ul, dashed] & \vdots \arrow[u] \arrow[ur, dashed] & \vdots \arrow[u] \arrow[ul, dashed]
\end{tikzcd}
$$

We saw in Theorem 6 that $M : \mathbf{Z}^{L\to\emptyset} \to \mathbf{Vec}$ is interval decomposable. By applying Theorem 7 to each disjoint component independently, the following is easy to show. Here the bottleneck distance is generalized in the obvious way, i.e. matching each component independently.

▶ **Corollary 8** (Isometry Theorem for Disjoint Unions). *Let* L *be any set, and* $M, N : \mathbf{Z}^{L\to\emptyset} \to \mathbf{Vec}$ *such that* $\dim M_p < \infty$ *and* $\dim N_p < \infty$ *for all* $p \in \mathbf{Z}^{L\to\emptyset}$. *Then* $d_I(M,N) = d_B(\mathcal{B}(M), \mathcal{B}(N))$.

In particular, the interleaving distance between M and N can be effectively computed through a bipartite matching. As we shall see, this is not true for $C \neq \emptyset$. The remainder of this section is devoted to proving the following theorem:

▶ **Theorem 9.** *Unless P=NP, there exists no algorithm, polynomial in $n = \dim M + \dim N$, which decides if $M, N : \mathbf{Z}^{L \to C} \to \mathbf{Vec}_{\mathbb{Z}/2\mathbb{Z}}$ are 1-interleaved.*

3.1 The proof

We shall prove Theorem 9 by a reduction from 3-SAT. Let ψ be a boolean formula in 3-CNF defined on literals $L = \{x_1, x_2, \ldots, x_{n_l}\}$ and clauses $C = \{c_1, c_2, \ldots, c_{n_c}\}$. We shall assume that the literals of each clause are distinct and *ordered*. That is, the clause c_i is specified by the three distinct literals $\{x_{i_1}, x_{i_2}, x_{i_3}\}$ wherein $i_1 < i_2 < i_3$. Determining if ψ is satisfiable is well-known to be NP-complete. For the entirety of the proof $\mathbb{K} = \mathbb{Z}/2\mathbb{Z}$.

Step 1: Defining the representations. Associate to ψ two functors $M, N : \mathbf{Z}^{L \to C} \to \mathbf{Vec}_{\mathbb{Z}/2\mathbb{Z}}$ in the following way: For all literals $x_j \in L$ define

$$M_{(x_j,1)} \longrightarrow M_{(x_j,2)} \longrightarrow M_{(x_j,3)} \longrightarrow M_{(x_j,4)} = \mathbb{K} \xrightarrow{1} \mathbb{K} \xrightarrow{1} \mathbb{K} \xrightarrow{1} \mathbb{K}$$

$$N_{(x_j,1)} \longrightarrow N_{(x_j,2)} \longrightarrow N_{(x_j,3)} \longrightarrow N_{(x_j,4)} = \mathbb{K} \xrightarrow{1} \mathbb{K} \xrightarrow{(1;1)} \mathbb{K}^2 \longrightarrow 0$$

and for every clause c_i in ψ define

$$M_{(c_i,1)} \longrightarrow M_{(c_i,2)} \longrightarrow M_{(c_i,3)} \longrightarrow M_{(c_i,4)} = 0 \longrightarrow \mathbb{K} \xrightarrow{1} \mathbb{K} \xrightarrow{1} \mathbb{K}$$

$$N_{(c_i,1)} \longrightarrow N_{(c_i,2)} \longrightarrow N_{(c_i,3)} \longrightarrow N_{(c_i,4)} = 0 \longrightarrow \mathbb{K}^3 \xrightarrow{1} \mathbb{K}^3 \longrightarrow 0$$

For any other $p \in \mathbf{Z}^{L \to C}$, $M_p = N_p = 0$. Next we specify the remaining non-trivial morphisms: let $c_i = z_{i_1} \vee z_{i_2} \vee z_{i_3}$ be a clause in ψ, where $z_{j_l} = x_{j_l}$ or $z_{j_l} = \neg x_{j_l}$, and for which $i_1 < i_2 < i_3$. For $s = 1, 2, 3$ define $H_s : \mathbb{K}^2 \to \mathbb{K}^3$ by $e_1 \to u \cdot e_s$ and $e_2 \to (1 - u) \cdot e_s$, where $u = 1$ if $z_{i_s} = x_{i_s}$ and $u = 0$ if $z_{i_s} = \neg x_{i_s}$. Here e_d is the d-th standard basis vector of $\mathbb{K}^3$. Given this we define the following for $s = 1, 2, 3$:

$$\begin{array}{ccc} \begin{array}{ccc} M_{(x_{i_s},4)} & \longrightarrow & M_{(c_i,4)} \\ \uparrow & & \uparrow \\ M_{(x_{i_s},3)} & \longrightarrow & M_{(c_i,3)} \end{array} & = & \begin{array}{ccc} \mathbb{K} & \xrightarrow{1} & \mathbb{K} \\ {\scriptstyle 1}\uparrow & & {\scriptstyle 1}\uparrow \\ \mathbb{K} & \xrightarrow{1} & \mathbb{K} \end{array} \end{array}$$

and

$$N_{(x_{i_s},3)} \longrightarrow N_{(c_i,3)} = \mathbb{K}^2 \xrightarrow{H_s} \mathbb{K}^3.$$

Clearly $\dim M + \dim N = \mathcal{O}(n_c + n_l)$. Thus, the total dimension is polynomial in the input size of 3-SAT.

Step 2: Showing the reduction. Observe that M and N are 1-interleaved if and only if there exist dashed morphisms such that the below diagram is commutative for every literal x_{i_s} and for every clause c_i containing x_{i_s}:

$$\begin{tikzcd}
M_{(x_{i_s},5)} = 0 & 0 = N_{(x_{i_s},5)} & M_{(c_i,5)} = 0 & 0 = N_{(c_i,5)} \\
M_{(x_{i_s},4)} = \mathbb{K} \arrow[u] \arrow[ur, dashed] \arrow[rr, bend left] & 0 = N_{(x_{i_s},4)} \arrow[u] \arrow[ul, dashed] \arrow[rr, bend left] & M_{(c_i,4)} = \mathbb{K} \arrow[u] \arrow[ur, dashed] & 0 = N_{(c_i,4)} \arrow[u] \arrow[ul, dashed] \\
M_{(x_{i_s},3)} = \mathbb{K} \arrow[u] \arrow[ur, dashed] \arrow[rr, bend left] & \mathbb{K}^2 = N_{(x_{i_s},3)} \arrow[u] \arrow[ul, dashed] \arrow[rr, bend left, "H_s"] & M_{(c_i,3)} = \mathbb{K} \arrow[u] \arrow[ur, dashed] & \mathbb{K}^3 = N_{(c_i,3)} \arrow[u] \arrow[ul, dashed] \\
M_{(x_{i_s},2)} = \mathbb{K} \arrow[u] \arrow[ur, dashed] & \mathbb{K} = N_{(x_{i_s},2)} \arrow[u, "(1;1)"'] \arrow[ul, dashed] & M_{(c_i,2)} = \mathbb{K} \arrow[u] \arrow[ur, dashed] & \mathbb{K}^3 = N_{(c_i,2)} \arrow[u, "I_3"'] \arrow[ul, dashed] \\
M_{(x_{i_s},1)} = \mathbb{K} \arrow[u] \arrow[ur, dashed] & \mathbb{K} = N_{(x_{i_s},1)} \arrow[u] \arrow[ul, dashed] & M_{(c_i,1)} = 0 \arrow[u] & 0 = N_{(c_i,1)} \arrow[u]
\end{tikzcd} \tag{1}$$

We shall see there are few degrees of freedom in the choice of interleaving morphisms. Indeed, consider the left part of the above diagram:

$$\begin{tikzcd}
M & 0 \arrow[r] & \mathbb{K} \arrow[r, "1"] \arrow[dr] & \mathbb{K} \arrow[r, "1"] \arrow[dr] & \mathbb{K} \arrow[r, "1"] \arrow[dr] & \mathbb{K} \arrow[r, "1"] \arrow[dr] & 0 \\
N & 0 \arrow[r] & \mathbb{K} \arrow[r, "1"'] \arrow[ur] & \mathbb{K} \arrow[r, "(1;1)"'] \arrow[ur] & \mathbb{K}^2 \arrow[r] \arrow[ur, dashed] & 0 \arrow[r] \arrow[ur] & 0
\end{tikzcd}$$

We leave it to the reader to verify that if M and N are 1-interleaved, then all the solid diagonal morphisms in the above diagram are completely determined by commutativity. For the dashed morphism $(\varphi_{x_{i_s}}, \varphi_{\neg x_{i_s}}) : \mathbb{K}^2 \to \mathbb{K}$ there are two choices: by commutativity it must satisfy $(\varphi_{x_{i_s}}, \varphi_{\neg x_{i_s}}) \cdot (1;1) = 1$ and thus $\varphi_{x_{i_s}} + \varphi_{\neg x_{i_s}} = 1$. As $\mathbb{K} = \mathbb{Z}/2\mathbb{Z}$, this implies that precisely one of $\varphi_{x_{i_s}}$ and $\varphi_{\neg x_{i_s}}$ is multiplication by 1. This corresponds to a choice of truth value for x_{i_s}: $\varphi_{x_{i_s}} = 1 \iff x_{i_s} = \text{True}$ and $\varphi_{\neg x_{i_s}} = 1 \iff x_{i_s} = \text{False}$. Next, consider the right part of 1:

$$\begin{tikzcd}
M & 0 \arrow[r] & \mathbb{K} \arrow[r, "1"] \arrow[dr, dashed] & \mathbb{K} \arrow[r, "1"] \arrow[dr] & \mathbb{K} \arrow[r] \arrow[dr] & 0 \\
N & 0 \arrow[r] & \mathbb{K}^3 \arrow[r, "1"'] \arrow[ur, dashed] & \mathbb{K}^3 \arrow[r] \arrow[ur, dashed] & 0 \arrow[r] \arrow[ur] & 0
\end{tikzcd}$$

There are three non-trivial morphisms, out of which two are equal by commutativity. Let $Z_1^i : \mathbb{K} \to \mathbb{K}^3$ and $Z_2^i : \mathbb{K}^3 \to \mathbb{K}$ denote the two unspecified morphisms. Returning to (1), we see that Z_2^i must satisfy the following for $s \in \{1, 2, 3\}$:

$$\begin{tikzcd}
\mathbb{K} \arrow[rr, "1"] & & \mathbb{K} \\
& \mathbb{K}^2 \arrow[ul, "(\varphi_{x_{i_s}}, \varphi_{\neg x_{i_s}})"] \arrow[r, "H_s"] & \mathbb{K}^3 \arrow[u, dashed, "Z_2^i"']
\end{tikzcd}$$

Thus, Z_2^i restricted to its s-th component equals either $\varphi_{x_{i_s}}$ or $\varphi_{\neg x_{i_s}}$, depending on whether x_{i_s} or its negation $\neg x_{i_s}$ appears in the clause c_i. This implies that Z_2^i is given by

$$Z_2^i = \begin{bmatrix} \varphi_{z_{i_1}} & \varphi_{z_{i_2}} & \varphi_{z_{i_3}} \end{bmatrix}$$

Hence, if M and N are to be 1-interleaved, then there are no degrees of freedom in choosing Z_2^i after the $\varphi_{x_{i_s}}$ are specified. However, Z_1^i only needs to satisfy $Z_2^i \circ Z_1^i = 1$. As this is the sole restriction imposed on Z_1^i, we see that this can be satisfied if and only if $Z_2^i \neq 0$, which is true if and only if $z_{i_s} = \text{True}$ for at least one $s \in \{1, 2, 3\}$.

▶ **Theorem 10.** *Let ψ be a boolean formula as above. Then ψ is satisfiable if and only if the associated persistence modules $M, N : \mathbf{Z}^{L \to C} \to \mathbf{vec}$ are 1-interleaved.*

Proof. Summarizing the above: we have that M and N are 1-interleaved if and only if we can choose morphisms $(\varphi_{x_{i_s}}, \varphi_{\neg x_{i_s}})$ such that $Z_2^i \neq 0$ for all clauses c_i. This means precisely that we can choose truth values for each x_{i_s} such that every clause $c_i = z_{i_1} \vee z_{i_2} \vee z_{i_3}$ evaluates to true. This shows that a 1-interleaving implies that ψ is satisfiable. Conversely, if ψ is satisfiable, then we see that the morphisms defined by $\varphi_{x_{i_s}} = 1 \iff x_{i_s} = \text{True}$ and $\varphi_{\neg x_{i_s}} = 1 \iff x_i = \text{False}$ satisfy $Z_2^i \neq 0$ for every clause c_i. Thus, M and N are 1-interleaved. ◀

▶ Remark. Let $i : \mathbf{P} \hookrightarrow \mathbf{Q}$ be an inclusion of posets and $M : \mathbf{P} \to \mathbf{Vec}$. There are multiple functorial ways of extending M to a representation $E(M) : \mathbf{Q} \to \mathbf{Vec}$, e.g. by means of left or right Kan extensions. This is a key ingredient in one of the more recent proofs of Theorem 7; see [11] for details. However, if we impose the condition that $E(M) \circ i \cong M$ then such an extension need not exist. Indeed, Theorem 10 implies that the associated decision problem is NP-complete.

4 Interleavings of multidimensional persistence modules

Recall that a constrained invertibility (CI) problem is a triple (P, Q, n) where P and Q are subsets of $\{1, 2, \dots, n\}^2$, and that a CI-problem is solvable if there exists an invertible $n \times n$ matrix M such that $M_{(i,j)} = 0$ for all $(i, j) \in P$ and $M^{-1}_{(i',j')} = 0$ for all $(i', j') \in Q$. We shall show that a CI-problem is solvable if and only if a pair of associated persistence modules $\mathbf{Z}^2 \to \mathbf{Vec}$ is 1-interleaved. Hence, if deciding solvability is NP-hard, then so is computing the interleaving distance for multidimensional persistence modules.

▶ **Example 11.** Let $P = \{(2, 2), (3, 3)\}, Q = \{(2, 3), (3, 2)\} \subset \{1, 2, 3\}^2$. Then $(P, Q, 3)$ is solvable by

$$M = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \quad M^{-1} = \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}.$$

▶ **Example 12.** Let $P = \{(1, 1), (1, 3)\}, Q = \{(2, 1)\} \subset \{1, 2, 3\}^2$. Then $(P, Q, 3)$ is not solvable, as $(MN)_{(1,1)} = 0$ for all 3×3-matrices M, N with $M_{(1,1)} = M_{(1,3)} = N_{(2,1)} = 0$. Note that it matters that we view P and Q as subsets of $\{1, 2, 3\}^2$ and not of $\{1, \dots, n\}^2$ for some $n > 3$, in which case (P, Q) would be solvable.

▶ **Example 13.** Observe that a CI-problem $(P, \emptyset, n)$ reduces to a bipartite matching problem. Build a graph G on $2n$ vertices $\{v_1, \dots, v_n, u_1, \dots, u_n\}$ with an edge from v_i to u_j if $(i, j) \notin P$. Then the CI-problem is solvable if and only if there exists a perfect matching of G.

A CI-problem can be seen as a problem of choosing weights for the edges in a directed simple graph: Given (P, Q, n), let G be the bipartite directed simple graph with vertices $\{u_1, \dots, u_n, v_1, \dots, v_n\}$, an edge from u_i to v_j if $(i, j) \notin P$, and an edge from v_j to u_i if $(j, i) \notin Q$. Solving (P, Q, n) is then equivalent to weighting the edges in G with elements from $\mathbb{K}$ so that

$$\sum_{j=1}^{n} w(u_i, v_j) w(v_j, u_i) = 1$$

for all i, and

$$\sum_{j=1}^{n} w(u_i, v_j) w(v_j, u_{i'}) = 0$$

for all $i \neq i'$, where $w(u, v)$ is the weight of the edge from u to v if there is one, and 0 if not. If the weights are elements of $\mathbb{Z}/2\mathbb{Z}$, this is equivalent to picking a subset of the edges such that there is an odd number of paths of length two from any vertex to itself and an even number of paths of length two from any vertex to any other vertex.

Fix a CI-problem (P, Q, n) and let $m = |P| + |Q|$. We will construct $\mathbf{Z}^2$-indexed modules M and N that are 1-interleaved if and only if (P, Q, n) is solvable, and that are zero outside a grid of size $(2m+3) \times (2m+3)$ in $\mathbf{Z}^2$. The dimension of each vector space $M_{(a,b)}$ or $N_{(a,b)}$ is bounded by n, so the total dimensions of M and N are polynomial in n.

For $p \in \mathbf{Z}^2$, let $\langle p \rangle = \{q \in \mathbf{Z}^2 \mid p \leq q \leq (2m+2, 2m+2)\}$. Let $\mathcal{W}$ be the interval $\bigcup_{k=0}^{m} \langle (2m-2k, 2k) \rangle$, and for $i \in \{1, 2, \ldots, m\}$, let $x_i = (2m - 2i + 1, 2i - 1)$; see Fig. 3.

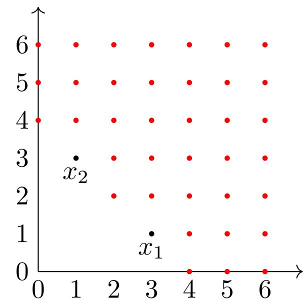


Figure 3 The interval $\mathcal{W}$ for $m = 2$ along with $x_1 = (3, 1)$ and $x_2 = (1, 3)$.

Write $P = \{(p_1, q_1), \ldots, (p_r, q_r)\}$ and $Q = \{(p_{r+1}, q_{r+1}), \ldots, (p_m, q_m)\}$. We define $M = \bigoplus_{i=1}^{n} I^{\mathcal{I}_i}$ and $N = \bigoplus_{i=1}^{n} I^{\mathcal{J}_i}$, where $\mathcal{I}_i$ and $\mathcal{J}_i$ are constructed as follows: let $\mathcal{I}_i^0 = \mathcal{J}_i^0 = \mathcal{W}$ for all i. For $k = 1, 2, \ldots, r$, let

$$\mathcal{I}_i^k = \begin{cases} \mathcal{I}_i^{k-1} \cup \langle x_k - (1,1) \rangle, & \text{if } i = p_k \\ \mathcal{I}_i^{k-1} \cup \langle x_k \rangle, & \text{if } i \neq p_k \end{cases}, \qquad \mathcal{J}_i^k = \begin{cases} \mathcal{J}_i^{k-1}, & \text{if } i = q_k \\ \mathcal{J}_i^{k-1} \cup \langle x_k \rangle, & \text{if } i \neq q_k \end{cases}$$

and for $k = r+1, \ldots, m$, let

$$\mathcal{I}_i^k = \begin{cases} \mathcal{I}_i^{k-1}, & \text{if } i = q_k \\ \mathcal{I}_i^{k-1} \cup \langle x_k \rangle, & \text{if } i \neq q_k \end{cases}, \qquad \mathcal{J}_i^k = \begin{cases} \mathcal{J}_i^{k-1} \cup \langle x_k - (1,1) \rangle, & \text{if } i = p_k \\ \mathcal{J}_i^{k-1} \cup \langle x_k \rangle, & \text{if } i \neq p_k \end{cases}$$

and let $\mathcal{I}_i = \mathcal{I}_i^m$ and $\mathcal{J}_i = \mathcal{J}_i^m$. This way, we ensure that there is no nonzero morphism from $I^{\mathcal{I}_i}$ to $I^{\mathcal{J}_j}(1)$ when $(i, j) \in P$, and no nonzero morphism from $I^{\mathcal{J}_j}$ to $I^{\mathcal{I}_i}(1)$ when $(j, i) \in Q$. In all other cases, there exist nonzero morphisms.

▶ **Lemma 14.** *Suppose $(i, j) \notin P$. Then there is an isomorphism $\mathrm{Hom}(I^{I_i}, I^{\mathcal{J}_j}(1)) \cong \mathbb{K}$. In particular, any morphism $f \in \mathrm{Hom}(I^{I_i}, I^{\mathcal{J}_j}(1))$ is completely determined by $f_{(2m+1,2m+1)}$: if f_p is nonzero, then $f_p = f_{(2m+1,2m+1)}$.*

The same holds if $(j, i) \notin Q$ instead of $(i, j) \notin P$, and $\mathcal{I}_i$ and $\mathcal{J}_j$ are interchanged. As f_p is a $\mathbb{K}$-endomorphism, this implies that any f can be identified with an element of $\mathbb{K}$.

Proof. Let $f : I^{\mathcal{I}_i} \to I^{\mathcal{J}_j}(1)$ be nonzero. If $p \notin \mathcal{I}_i$ or $p \not\leq (2m+1, 2m+1)$, $f_p = 0$. For $(2m+1, 2m+1) \geq p \in \mathcal{I}_i$, we have $p + (1,1) \in \mathcal{J}_i$ by construction and the fact that $(i, j) \notin P$, so $\varphi_{I^{\mathcal{J}_j}}(p + (1,1), (2m+2, 2m+2))$ is nonzero and hence the identity. We get

$$\begin{aligned} f_p &= \varphi_{I^{\mathcal{J}_j}(1)}(p, (2m+1, 2m+1)) \circ f_p \\ &= f_{(2m+1,2m+1)} \circ \varphi_{I^{\mathcal{I}_i}}(p, (2m+1, 2m+1)) = f_{(2m+1,2m+1)}. \end{aligned}$$

◀

Describing a morphism from $M = \bigoplus_{i=1}^n I^{\mathcal{I}_i}$ to $N(1) = \bigoplus_{j=1}^n I^{\mathcal{J}_j}(1)$ is the same as describing morphisms from $I^{\mathcal{I}_i}$ to $I^{\mathcal{J}_j}(1)$ for all i and j[2]. We have just proved that these can be identified with elements of $\mathbb{K}$, so we conclude that any $f : M \to N(1)$ is uniquely defined by an $n \times n$-matrix A_f where the entry (i, j) is the element in $\mathbb{K}$ corresponding to the morphism $I^{\mathcal{I}_i} \to I^{\mathcal{J}_j}(1)$ given by f. Note that we get the same result by writing $f_{(2m,2m)} = f_{(2m+1,2m+1)} : \mathbb{K}^n \to \mathbb{K}^n$ as a matrix, where each copy of $\mathbb{K}$ in the domain and codomain comes from one of the interval modules $I^{\mathcal{I}_i}$ or $I^{\mathcal{J}_j}(1)$, respectively.

If we also have a morphism $g : N \to M(1)$, we can define a matrix A_g symmetrically, and similarly A_g is $g_{(2m,2m)} = g_{(2m+1,2m+1)} : \mathbb{K}^n \to \mathbb{K}^n$ in matrix form.

▶ **Theorem 15.** *With f and g as above, (f, g) is a 1-interleaving if and only if A_f and A_g are inverse matrices.*

Proof. Suppose (f, g) is a 1-interleaving. The internal morphism $\varphi_M((2m, 2m),(2m+2, 2m+2))$ is the identity on $\mathbb{K}^n$, and is by definition of interleaving the same as

$$g(1)_{(2m,2m)} \circ f_{(2m,2m)} = g_{(2m+1,2m+1)} \circ f_{(2m,2m)} = A_g A_f.$$

Thus $A_g A_f$ is the identity matrix, and so A_f and A_g are inverses of each other, as both are $n \times n$-matrices.

Suppose A_f and A_g are inverse matrices. We must check that at every point $p \in \mathbb{Z}^2$, $\varphi_M(p, p + (2, 2)) = g(1)_p \circ f_p$. If $p \not\leq (2m, 2m)$ or $M(p) = 0$, both sides are zero. If $p \leq (2m, 2m)$ and $M(p) \neq 0$,

$$g(1)_p \circ f_p = \varphi_M(p + (2, 2),(2m + 2, 2m + 2)) \circ g(1)_p \circ f_p,$$

since $\varphi_M(p + (2, 2),(2m + 2, 2m + 2))$ must be the identity by construction of M and the fact that $M(p) \neq 0$. This is equal to

$$\begin{aligned}
&\varphi_M(p + (2, 2),(2m + 2, 2m + 2)) \circ g_{p+(1,1)} \circ f_p \\
&= g_{(2m+1,2m+1)} \circ \varphi_N(p + (1, 1),(2m + 1, 2m + 1)) \circ f_p \\
&= g_{(2m+1,2m+1)} \circ f_{(2m,2m)} \circ \varphi_M(p,(2m, 2m)) \\
&= A_g A_f \circ \varphi_M(p,(2m, 2m)) = \varphi_M(p,(2m, 2m)) = \varphi_M(p, p + (2, 2)).
\end{aligned}$$ ◀

We have proved that defining morphisms $f : M \to N(1)$ and $g : N \to M(1)$ is the same as choosing $n \times n$-matrices A_f and A_g such that the entries corresponding to the elements of P and Q are zero, and that (f, g) is a 1-interleaving if and only if A_f and A_g are inverse matrices. Thus M and N are 1-interleaved if and only if the CI-problem (P, Q, n) is solvable.

We constructed M and N by setting all the interval modules comprising M and N equal to $I^{\mathcal{W}}$, then modifying them in m steps each, where the complexity of each step is clearly polynomial in n. Thus the complexity of constructing M and N is polynomial in n, and so are the total dimensions of M and N. Taking n^2 as the input complexity of solving a CI-problem (P, Q, n), we have proved a reduction implying the following theorem:

▶ **Theorem 16.** *Determining the interleaving distance for modules $\mathbf{Z}^2 \to \mathbf{Vec}$ is CI-hard.*

▶ Remark. We give an example in [7] of a CI-problem whose associated matrices M and N satisfy $d_I(M, N) = 1$ and $d_B(\mathcal{B}(M), \mathcal{B}(N)) = 2$. This shows that it is not enough to find the bottleneck distance of the barcodes of M and N to decide whether M and N are 1-interleaved and thus whether the CI-problem is solvable. In fact, recent work shows that d_B can be efficiently computed [21].

---

[2] $\mathrm{Hom}(\oplus_i M_i, \oplus_j N_j) \cong \oplus_i \oplus_j \mathrm{Hom}(M_i, N_j)$.

We end this paper with the somewhat surprising observation that the interleaving distance of the above interval decomposable modules depends the characteristic $\mathrm{char}(\mathbb{K})$ of the underlying field $\mathbb{K}$. That is, let $M, N : \mathbf{Z}^2 \to \mathbf{Vec}_{\mathbb{K}}$, $M', N' : \mathbf{Z}^2 \to \mathbf{Vec}_{\mathbb{K}'}$, $\mathbb{K} \neq \mathbb{K}'$, and for which $\mathcal{B}(M) = \mathcal{B}(M')$ and $\mathcal{B}(N) = \mathcal{B}(N')$. Clearly, any matching distance d would satisfy $d(\mathcal{B}(M), \mathcal{B}(N)) = d(\mathcal{B}(M'), \mathcal{B}(N'))$, but it is not always true that $d_I(M, N) = d_I(M', N')$.

For a fixed $n \geq 2$, let $Q = \{(2,2), \ldots, (n+2, n+2)\}$ and $P = \{(1,1)\} \cup \{2, \ldots, n+2\}^2 \setminus Q$. Then the CI-problem $(P, Q, n+2)$ is solvable if and only if the characteristic of $\mathbb{K}$ divides n. We will only prove this for $n = 2$ for clarity, but the argument easily generalizes to all n.

Assume that (M, M^{-1}) is a solution to $(P, Q, 4)$:

$$M = \begin{bmatrix} 0 & ? & ? & ? \\ a & ? & 0 & 0 \\ b & 0 & ? & 0 \\ c & 0 & 0 & ? \end{bmatrix}, \quad M^{-1} = \begin{bmatrix} ? & d & e & f \\ ? & 0 & ? & ? \\ ? & ? & 0 & ? \\ ? & ? & ? & 0 \end{bmatrix}$$

Here we have put the entries corresponding to the elements of P and Q equal to 0, and left the rest as unknown. The entries we will use in the calculations that follow are labeled a, b, c, d, e, f. We see that $(MM^{-1})_{(2,2)} = ad$, $(MM^{-1})_{(3,3)} = be$, $(MM^{-1})_{(4,4)} = cf$, that is, $ad = be = cf = 1$. At the same time, $(M^{-1}M)_{(1,1)} = ad + be + cf$, so we get $1 = 1+1+1$, or $2 = 0$. Thus $\mathrm{char}(\mathbb{K}) = 2$, and in this case we can put all the unknowns in M and M^{-1} above equal to 1 to obtain a solution. (For $n > 2$, we put the nonzero elements on the diagonal of M equal to -1.)

Our motivation for introducing CI-problems was working towards determining the computational complexity of calculating the interleaving distance. While the last examples say little about complexity, they illustrate the underlying philosophy of our approach: By considering CI-problems, we can avoid the confusing geometric aspects of persistence modules and interleavings. E.g., in the case above, working with persistence modules over a 23×23 size grid is reduced to looking at a pair of 4×4-matrices.

5 Discussion

The problem of determining the computational complexity of computing the interleaving distance for multidimensional persistence modules (valued in **Vec**) was first brought up in Lesnick's thesis [23]. Although it has been an important open question for several years, a non-trivial lower bound on the complexity class has not yet been given. In light of Theorem 2, one might hope that tools from computational algebra can be efficiently extended to the setting of interleavings. Theorem 9 is an argument against this, as it shows that the problem of computing the interleaving distance is NP-hard in general. This leads us to conjecture that the problem of computing the interleaving distance for multidimensional persistence modules is also NP-hard. Unfortunately, writing down the conditions for an interleaving becomes intractable already for small grids. To make the decision problem more accessible to researchers in other fields of mathematics and computer science, we have shown that the problem is at least as hard as an easy to state matrix invertibility problem. We speculate that this problem is also NP-hard. If that is not the case, then an algorithm would provide valuable insight into the interleaving problem for interval decomposable modules.

References

1 Pankaj K Agarwal, Kyle Fox, Abhinandan Nath, Anastasios Sidiropoulos, and Yusu Wang. Computing the gromov-hausdorff distance for metric trees. In *International Symposium on*

Algorithms and Computation, pages 529–540. Springer, 2015.

2 Gorô Azumaya. Corrections and supplementaries to my paper concerning Krull–Remak–Schmidt's theorem. *Nagoya Mathematical Journal*, 1:117–124, 1950.

3 Michael Barot. *Introduction to the representation theory of algebras.* Springer, 2014.

4 Ulrich Bauer and Michael Lesnick. Induced matchings and the algebraic stability of persistence barcodes. *Journal of Computational Geometry*, 6(2):162–191, 2015.

5 Silvia Biasotti, Andrea Cerri, Patrizio Frosini, Daniela Giorgi, and Claudia Landi. Multidimensional size functions for shape comparison. *Journal of Mathematical Imaging and Vision*, 32(2):161–179, 2008.

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

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

8 Magnus Bakke Botnan. Interval decomposition of infinite zigzag persistence modules. *Proceedings of the American Mathematical Society*, 145(8):3571—-3577, 2017.

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

10 Peter A Brooksbank and Eugene M Luks. Testing isomorphism of modules. *Journal of Algebra*, 320(11):4020–4029, 2008.

11 Peter Bubenik, Vin de Silva, and Vidit Nanda. Higher interpolation and extension for persistence modules. *SIAM Journal on Applied Algebra and Geometry*, 1(1):272–284, 2017.

12 Peter Bubenik, Vin de Silva, and Jonathan Scott. Metrics for generalized persistence modules. *Foundations of Computational Mathematics*, 15(6):1501–1531, 2015.

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

14 G. Carlsson and A. Zomorodian. The theory of multidimensional persistence. *Discrete & Computational Geometry*, 42(1):71–93, 2009.

15 Gunnar Carlsson and Vin de Silva. Zigzag persistence. *Foundations of computational mathematics*, 10(4):367–405, 2010.

16 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.

17 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.

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

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

20 Vin de Silva, Elizabeth Munch, and Amit Patel. Categorified Reeb graphs. *arXiv preprint arXiv:1501.04147*, 2015.

21 Tamal K Dey and Cheng Xin. Computing bottleneck distance for 2-d interval decomposable modules. *arXiv preprint arXiv:1803.02869*, 2018.

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

23 Michael Lesnick. Multidimensional interleavings and applications to topological inference. *arXiv preprint arXiv:1206.1365*, 2012.

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

25 Nikola Milosavljević, Dmitriy Morozov, and Primoz Skraba. Zigzag persistent homology in matrix multiplication time. In *Proceedings of the twenty-seventh annual symposium on Computational geometry*, pages 216–225. ACM, 2011.

26 Martina Scolamiero, Wojciech Chachólski, Anders Lundman, Ryan Ramanujam, and Sebastian Öberg. Multidimensional persistence and noise. *Foundations of Computational Mathematics*, 17(6):1367–1406, 2017.

27 Cary Webb. Decomposition of graded modules. *Proceedings of the American Mathematical Society*, 94(4):565–571, 1985.