Geometry in the Space of Persistence Modules

Vin de Silva and Vidit Nanda

Abstract. Topological persistence is, by now, an established paradigm for constructing robust topological invariants from point-cloud data: the data are converted into a filtered simplicial complex, the complex gives rise to a persistence module, and the module is described by a persistence diagram. In this paper, we study the geometry of the spaces of persistence modules and diagrams, with special attention to Čech and Rips complexes. The metric structures are determined in terms of interleaving maps between modules and matchings between diagrams. We show that the relationship between the Čech and Rips complexes is governed by certain 'coherence' conditions on the corresponding families of interleavings or matchings.

1. Introduction and Motivation

The use of topological techniques to analyze large, complicated, high-dimensional datasets has gained enormous traction over the last few years. The theory of persistence has been at the vanguard of such analysis [**4**, **11**, **9**]: for instance, the persistent homology of a point cloud provides a global multi-scale description of the appearance and disappearance of topological features as the points are thickened into balls of increasing radius. In addition to its efficient computability and representability in the form of a *persistence diagram* [**20**, **16**, **15**], the most important property of persistent homology is the celebrated *stability* theorem [**7**, **6**] which guarantees that the persistence diagram is quantifiably robust to fluctuations in the locations of the data points. Here is a typical workflow in the application of persistence to point cloud data:

$$\text{Point Cloud} \to \text{Filtered Complex} \to \text{Module} \to \text{Diagram}$$

We defer the precise definitions of persistence modules and diagrams to a subsequent section but note here that the success of persistent homology has led inevitably to active investigations of the spaces of persistence modules and diagrams in their own right [**14**, **5**, **2**]. Our work continues this trend but with a geometric and topological flavor and with a particular interest in embedding theorems.

Although persistent homology has been widely used—and continues to be used—for the analysis of static datasets, the progress in terms of using persistence to understand *transformations* of datasets has been comparatively sparse. Let $C^2(X, Y)$ denote the space of all twice continuously differentiable functions between manifolds X and Y. Our work is motivated in part by the following embedding theorem from [**17**].

Theorem 1.1 (Takens). *Let M be a compact manifold of dimension n. Across all pairs (ϕ, f) where $\phi \in C^2(M, M)$ is a diffeomorphism and $f \in C^2(M, \mathbf{R})$ is real-valued, it is generically true that the map $\Theta : M \to \mathbf{R}^{2n+1}$ defined on $x \in M$ by*

$$\Theta(x) = (f(x), f \circ \phi(x), \ldots, f \circ \phi^{2n}(x))$$

is an embedding.

Although M may be unknown, it is generically possible by Takens' theorem to reconstruct it as a subset of $\mathbf{R}^{2n+1}$ from knowledge of the *time series* $\left(f \circ \phi^k : M \to \mathbf{R}\right)_{k\in\mathbf{N}}$.

Let Conf$(m, \mathcal{B})$ be the space of all possible configurations of m point particles in a bounded region $\mathcal{B}$. Assume that an unknown ambient force field acts on m particles in $\mathcal{B}$ recurrently. That is, there exists some unknown *attractor* $M' \subset$ Conf$(m, \mathcal{B})$ and a bijection $\phi' : M' \to M'$ capturing the action of the ambient field on these m points. Although we may not have precise information about the attractor M' or the map ϕ', we can measure discrete snap-shots of the positions of the particles as ϕ' is iterated. Let f' be a map taking each point cloud in M' to the space of persistence modules (or diagrams), and note that $\left\{f' \circ \phi'^k\right\}_{k\in\mathbf{N}}$ generates a time series of persistence modules (or diagrams). We would like to understand to what extent, if any, can the attractor M' be generically reconstructed from such time series as a subset of the space of persistence modules (or diagrams) in a sense analogous to Takens' theorem. A first step in this direction is to develop tools which would enable us to recover geometry and topology in the spaces of persistence modules and diagrams.

In this paper we consider necessary and sufficient compatibility conditions which guarantee the existence of a persistence module (or diagram) within a given distance of a collection of persistence modules (or diagrams). It thus becomes possible to construct Čech complexes in these spaces and hence recover topology of subspaces via the nerve theorem [1]. It turns out that the difference between Rips simplices and Čech simplices is given by a notion of coherence between interleaving maps (of persistence modules) or between matchings (of diagrams).

2. Modules and Diagrams

All vector spaces are taken over a fixed field $\mathbf{k}$. We use $\mathbf{R}$ to denote the real line.

2.1. Persistence modules.

Following [5], a persistence module $\mathbb{U}$ indexed by $\mathbf{R}$—or simply a **module**—is specified by the following data: a vector space U_t for each $t \in \mathbf{R}$, and a linear map $u_t^s : U_s \to U_t$ whenever $s \leqslant t$, such that $u_t^t = 1$ for all t and $u_t^s u_s^r = u_t^r$ whenever $r \leqslant s \leqslant t$.

Certain modules are particularly well behaved. A module $\mathbb{U}$ is **q-tame** (cf. [6]) if rank$(u_t^s) < \infty$ whenever $s < t$. The class of q-tame modules is denoted (**mod**).

Let $\mathbb{U}, \mathbb{V}$ be modules, and let $d \in \mathbf{R}$. A **module map** $\Phi : \mathbb{U} \to \mathbb{V}$ of degree d is specified by a collection of linear maps $\phi_t : U_t \to V_{t+d}$ such that $v_{t+d}^{s+d}\phi_s = \phi_t u_t^s$ whenever $s \leqslant t$. The set of module maps from $\mathbb{U}$ to $\mathbb{V}$ of degree d is denoted Hom$^d(\mathbb{U}, \mathbb{V})$.

The most important module maps are the shift maps: if $d \geqslant 0$ then $I^d = I^d_{\mathbb{U}}$ is the map $\mathbb{U} \to \mathbb{U}$ of degree d defined by $\iota_t = u_{t+d}^t$. We think of these as 'blurred' identity maps, where d is the amount of blur. We can then talk about 'blurred' isomorphisms: a **d-interleaving** between two modules $\mathbb{U}, \mathbb{V}$ is a pair of module maps $\Phi : \mathbb{U} \to \mathbb{V}$ and $\Psi : \mathbb{V} \to \mathbb{U}$ of degree d such that $\Psi\Phi = I^{2d}_{\mathbb{U}}$ and $\Phi\Psi = I^{2d}_{\mathbb{V}}$. When such a pair Φ, Ψ exists, we say that $\mathbb{U}, \mathbb{V}$ are d-interleaved.

We now define a pseudometric on the class of persistence modules, the **interleaving distance**:

$$d_i(\mathbb{U}, \mathbb{V}) = \inf\{d \geqslant 0 \mid \exists \text{ a } d\text{-interleaving between } \mathbb{U}, \mathbb{V}\}$$

If there is no interleaving between $\mathbb{U}, \mathbb{V}$ then $d_i(\mathbb{U}, \mathbb{V}) = \infty$. In general, the infimum need not be attained.

This formalism is designed to support examples of the following type. Let X be a finite polyhedron and let $f : \mathsf{X} \to \mathbf{R}$ be a continuous function. The sublevelsets of f are the subspaces of X

$$\mathsf{X}_f^t = f^{-1}(-\infty, t]$$

and there is an inclusion $\mathsf{X}_f^s \subseteq \mathsf{X}_f^t$ whenever $s \leqslant t$. We can take homology (with coefficients in $\mathbf{k}$) to construct a persistence module $\mathbb{U} = \mathbb{U}_f$, defined by

$$U_t = \mathrm{H}(\mathsf{X}_f^t), \quad u_t^s = \mathrm{H}(\mathsf{X}_f^s \to \mathsf{X}_f^t).$$

It turns out that this module is q-tame ([6], section 2.8; also [3]). Moreover, the map $f \mapsto \mathbb{U}_f$ is 1-Lipschitz with respect to the uniform norm on the space $\mathrm{C}(\mathsf{X})$ of continuous functions on X. Indeed, if $\|f - g\| \leqslant d$ then there are inclusions $\mathsf{X}_f^t \subseteq \mathsf{X}_g^{t+d}$ and $\mathsf{X}_g^t \subseteq \mathsf{X}_f^{t+d}$ which induce a d-interleaving between $\mathbb{U}_f, \mathbb{U}_g$.

2.2. Diagrams. Persistence modules are rather abstract objects. To visualize them concretely, they are converted into persistence diagrams.

For our purposes, a **diagram** is a locally finite multiset in the extended half-plane $\mathcal{H}$ of pairs (p, q) with $p < q$, where $p \in \{-\infty\} \cup \mathbf{R}$ and $q \in \mathbf{R} \cup \{+\infty\}$. The space of all possible diagrams is called (**diag**). We will shortly define a metric on this space.

The general principle—going back to the pioneering work of [10]—is that a sufficiently well-behaved persistence module $\mathbb{U}$ can be described very nearly faithfully by its diagram $\mathsf{dgm}(\mathbb{U})$. The stability theorem of [7] asserts (under some restrictions on the modules) that the map

$$\begin{array}{ccc} (\mathbf{mod}) & \longrightarrow & (\mathbf{diag}) \\ \mathbb{U} & \longmapsto & \mathsf{dgm}(\mathbb{U}) \end{array}$$

is 1-Lipschitz; and in [13] it was shown to be an isometry. For proofs in the case of q-tame modules, see [6].

This immediately implies, for instance, that the composite map,

$$\begin{array}{ccccc} \mathrm{C}(\mathsf{X}) & \longrightarrow & (\mathbf{mod}) & \longrightarrow & (\mathbf{diag}) \\ f & \longmapsto & & & \mathsf{dgm}(\mathrm{H}(\mathsf{X}_f^t)) \end{array}$$

that takes a function on a polyhedron X to its sublevelset persistent homology diagram, is 1-Lipschitz. This is why the persistence diagram is a stable invariant of (X, f).

The metric used in these assertions is the bottleneck distance on diagrams, which is defined in terms of the metric $\mathrm{d}^\infty((p_1, q_1), (p_2, q_2)) = \max(|p_1 - p_2|, |q_1 - q_2|)$ on the extended plane. Let D, E be diagrams. A **d-matching** is a partial bijection $\omega : \mathsf{D} \to \mathsf{E}$ satisfying two conditions:

(1) If $x \in \mathsf{D}$ and $y = \omega(x) \in \mathsf{E}$ then $\mathrm{d}^\infty(x, y) \leqslant d$.
(2) If $x \in \mathsf{D}$ or E is unmatched, then there is a point $(t, t) \in \mathbf{R}^2$ such that $\mathrm{d}^\infty(x, (t, t)) \leqslant d$.

It is convenient to think of a partial matching as a subset of $\mathsf{D} \times \mathsf{E}$, that is a relation between D, E. (One must take some care to navigate around the fact that D, E are multisets rather than sets. This can be done by attaching arbitrary labels to distinguish multiple copies of a point.)

The **bottleneck distance** between diagrams is defined to be

$$\mathrm{d_b}(\mathsf{D}, \mathsf{E}) = \inf\{d \geqslant 0 \mid \exists \text{ a } d\text{-matching between } \mathsf{D}, \mathsf{E}\}.$$

If there is no d-matching for finite d then $\mathrm{d_b}(\mathsf{D},\mathsf{E}) = \infty$. Since we are working with locally finite multisets, a compactness argument can be used to show that the infimum, when finite, is attained ([6], section 4.3). It follows that $\mathrm{d_b}$ is a true metric rather than a pseudometric: if $\mathrm{d_b}(\mathsf{D},\mathsf{E}) = 0$ then $\mathsf{D} = \mathsf{E}$.

2.3. Interleavings and matchings. The isometry theorem of [7, 13, 2, 6] amounts to the following specific assertions about q-tame modules and their diagrams:

(1) For every locally finite multiset D there exists a module $\mathbb{U}$ with $\mathsf{dgm}(\mathbb{U}) = \mathsf{D}$.
(2) Given a d-interleaving between $\mathbb{U}, \mathbb{V}$ there exists a d-matching between $\mathsf{dgm}(\mathbb{U})$, $\mathsf{dgm}(\mathbb{V})$.
(3) Given a d-matching between $\mathsf{dgm}(\mathbb{U})$, $\mathsf{dgm}(\mathbb{V})$, there exists a $(\mathsf{d}+\epsilon)$-interleaving between $\mathbb{U}, \mathbb{V}$ for all $\epsilon > 0$.

The construction for assertion 1 uses **interval modules**. Let $A \subseteq \mathbf{R}$ be an interval. Then $\mathbb{J} = \mathbb{J}_A$ is defined to be the persistence module with:

$$J_t = \begin{cases} \mathbf{k} & \text{if } t \in A \\ 0 & \text{if } t \notin A \end{cases} \qquad j_t^s = \begin{cases} 1 & \text{if } s, t \in A \\ 0 & \text{otherwise} \end{cases}$$

To each point (p, q) in the extended half-plane $\mathcal{H}$ we associate the half-open interval $[p, q)$, and therefore the interval module $\mathbb{J}_{[p,q)}$. To any multiset D in $\mathcal{H}$, we can define $\mathbb{U}$ to be the direct sum of the interval modules corresponding to the points in D. If D is locally finite, then $\mathbb{U}$ is q-tame and $\mathsf{dgm}(\mathbb{U}) = \mathsf{D}$. It is equally good to use closed or open intervals, so there are many choices.

Assertion 3 is easy to prove when $\mathbb{U}, \mathbb{V}$ are direct sums of interval modules. The d-matching between the persistence diagrams can be interpreted as a matching between the interval summands of $\mathbb{U}, \mathbb{V}$ so the interleaving can be constructed separately on each matched pair of summands. The d^∞-metric in $\mathbf{R}^2$ precisely governs the interleaving distance between two interval modules (except when both modules are near the diagonal, in which case we still get an upper bound). For the general case, it turns out that every q-tame module can be approximated by modules that decompose into intervals ([6], section 4.5; using a result of [18]). This is good enough to deduce the full result.

Assertion 2 is the most subtle and requires a clever argument, even under the strong hypotheses of [7]. The most general proof is based on the following lemma of [5, 6], which inspires some of our later results:

Lemma 2.1 (The interpolation lemma). *Let $\mathbb{U}$ and $\mathbb{V}$ be a pair of persistence modules which are d-interleaved. Then there is a 1-parameter family $\mathbb{U}_t$, where $t \in [0, \mathsf{d}]$, such that $\mathbb{U}_s, \mathbb{U}_t$ are $|s - t|$-interleaved for all s, t and where $\mathbb{U}_0 = \mathbb{U}$ and $\mathbb{U}_{\mathsf{d}} = \mathbb{V}$.* □

By tracking the persistence diagram as it varies along this 1-parameter family (cf. the theory of vineyards [8]), one shows that the d-matching exists. However, the proof does not uniquely specify a matching, even when the interpolation maps Φ, Ψ are given explicitly. It is important to keep this in mind, to avoid certain natural errors in understanding the relationship between (**mod**) and (**diag**).

The interpolation lemma essentially asserts that (**mod**) is a length space: between any two modules $\mathbb{U}, \mathbb{V}$ there exist connecting paths of length $\mathrm{d_i}(\mathbb{U}, \mathbb{V}) + \epsilon$ for every $\epsilon > 0$.

There is also an interpolation lemma for (**diag**), with a straightforward proof: given two diagrams and a d-matching between them, one constructs an interpolating family of diagrams by moving each point linearly towards its matched partner (or towards the closest point on the diagonal, if not matched).

Since the infimum in the definition of d_b is attained, this gives the stronger conclusion that (**diag**) is a geodesic space: any two diagrams D, E are connected by a path of length $d_b(D, E)$.

3. Čech and Rips complexes

3.1. Definitions. Let M be any metric space. Given a finite data set in M, it is standard practice in topological data analysis to represent the topology of the data by a nested family of simplicial complexes, indexed by a scale parameter d. The persistent homology of this family is then used as a topological descriptor for the data.

Two typical constructions are the Čech and the Vietoris–Rips complexes. Each is defined by a rule which specifies for each abstract simplex $[x_0, x_1, \ldots, x_n]$ the value of d at which it enters the complex:

$$\begin{aligned}
[x_0, \ldots, x_n] \in \mathrm{Rips}(M, d) \quad &\Leftrightarrow d_M(x_i, x_j) \leqslant d \text{ for all } i, j \\
[x_0, \ldots, x_n] \in \check{\mathrm{C}}\mathrm{ech}(M, d) \quad &\Leftrightarrow \text{some } y \in M \text{ satisfies} \\
&\qquad d_M(x_i, y) \leqslant d \text{ for all } i
\end{aligned}$$

In the second definition, we call y a **d-witness** for the simplex $[x_0, x_1, \ldots, x_n]$.

We alter this slightly in the case of (**mod**) because the interleaving distance d_i is defined as an infimum which is not always attained. For a simplex $\sigma = [\mathbb{U}_0, \mathbb{U}_1, \ldots, \mathbb{U}_n]$ of persistence modules:

$$\begin{aligned}
\sigma \in \mathrm{Rips}((\mathbf{mod}), d) &\Leftrightarrow \text{every pair } \mathbb{U}_i, \mathbb{U}_j \text{ is } d\text{-interleaved} \\
\sigma \in \check{\mathrm{C}}\mathrm{ech}((\mathbf{mod}), d) &\Leftrightarrow \text{some } \mathbb{V} \text{ is } d\text{-interleaved with all } \mathbb{U}_i
\end{aligned}$$

Remark. When a metric is defined as an infimum over some family of comparison objects (such as interleavings), it makes sense to define Čech and Vietoris–Rips complexes in terms of the comparison objects rather than the infimum. When the infimum is guaranteed to be attained, such as for d_b on (**diag**), then it makes no difference which we use.

For topological data analysis, a finite data set $S \subset M$ may be represented by the restriction of $\mathrm{Rips}(M, d)$ or $\check{\mathrm{C}}\mathrm{ech}(M, d)$ to the vertex set S.

3.2. Sandwiching ratio. It is well known that Vietoris–Rips and Čech complexes are nested in the following way:

$$\check{\mathrm{C}}\mathrm{ech}(M, d) \subseteq \mathrm{Rips}(M, 2d) \subseteq \check{\mathrm{C}}\mathrm{ech}(M, 2d)$$

For the first inclusion, if y is a d-witness then $d_X(x_i, x_j) \leqslant d_X(x_i, y) + d_X(y, x_j) \leqslant 2d$. For the second inclusion, if every $d_X(x_i, x_j) \leqslant 2d$ then any x_i is a $2d$-witness for the simplex.

In general, the ratio 2 between the parameters d and $2d$, for the Čech complexes that sandwich the Vietoris–Rips complex, is best possible. We will show that under certain conditions this ratio can be tightened to 1.

If M is a geodesic space then the left-hand inclusion cannot be tightened: if $[x_0, x_1] \in \mathrm{Rips}(M, 2d)$ then the midpoint of a minimal geodesic from x_0 to x_1 serves as a d-witness for $[x_0, x_1]$. Thus the 1-skeleta of $\check{\mathrm{C}}\mathrm{ech}(M, d)$ and $\mathrm{Rips}(M, 2d)$ are equal so one cannot increase the Čech parameter d and still get an inclusion. The same conclusion holds more generally for length spaces, using the same argument with added ϵ.

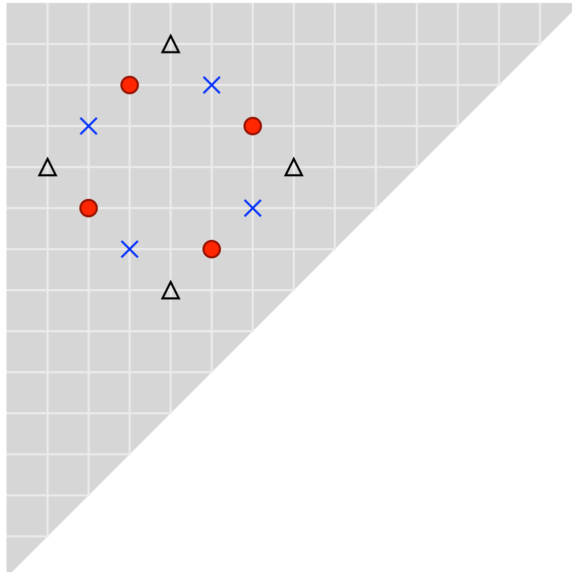

FIGURE 1. Three overlaid persistence diagrams D_1 (●), D_2 (△), and D_3 (×). There exist incoherent pairwise 1-matchings obtained by associating each point in one diagram to the nearest point in any other diagram. However, there is no coherent 1-matching across all three diagrams. Note that there is no $\frac{1}{2}$-witness for the collection $\{D_1, D_2, D_3\}$ and indeed, the best one obtains is a 1-witness.

Thus, in order to tighten the ratio for a length space—such as a normed vector space, or (**mod**) or (**diag**)—we seek the smallest e such that $\text{Rips}(M, 2d) \subseteq \check{\text{C}}\text{ech}(M, e)$.

It is well known that if $M = \mathbf{R}^n$ with the ℓ^∞-norm, then $\text{Rips}(M, 2d) = \check{\text{C}}\text{ech}(M, d)$, so in that case the sandwiching ratio is 1. More generally, this is true for function spaces with the supremum norm:

PROPOSITION 3.1. *For any topological space* X*, we have* $\text{Rips}(C(X), 2d) = \check{\text{C}}\text{ech}(C(X), d)$.

PROOF. Indeed, if $f_0, f_1, \ldots, f_n$ satisfy $\|f_i - f_j\| \leqslant 2d$ for all i, j then

$$g(x) = \frac{\max_i f_i(x) + \min_i f_i(x)}{2}$$

is continuous and satisfies $\|g - f_i\| \leqslant d$ for all i. □

In contrast, for (**mod**) and (**diag**) the ratio 2 cannot be improved without further hypotheses. See Figure 1 for three diagrams spanning a triangle in $\text{Rips}((\textbf{diag}), 1)$ which is not in $\check{\text{C}}\text{ech}((\textbf{diag}), e)$ for any $e < 1$.

The same example can be lifted to (**mod**) by virtue of the isometry theorem.

3.3. Coherent interleavings of modules. Suppose we are given a collection $\mathbb{U}_1, \ldots, \mathbb{U}_n$ of modules which are pairwise $2d$-interleaved; that is, suppose we are given a simplex in $\text{Rips}((\textbf{mod}), 2d)$. What is its Čech radius? That is, for what e does there exist an e-witness?

THEOREM 3.2. *Let* $\mathbb{U}_1, \ldots, \mathbb{U}_n$ *be modules. Then the following are equivalent:*

(i) There exists a module $\mathbb{V}$ *which is* d*-interleaved with each* $\mathbb{U}_i$*.*

(ii) There exists a collection of module maps

$$\Phi^i_j \in \mathrm{Hom}^{2d}(\mathbb{U}_i, \mathbb{U}_j) \quad \textit{all } i, j \textit{ distinct.}$$

such that

$$\Phi^j_i \Phi^i_j = I^{4d}_i \qquad \textit{all } i, j \textit{ distinct} \tag{†}$$
$$\Phi^j_k \Phi^i_j = \Phi^i_k I^{2d}_i \quad \textit{all } i, j, k \textit{ distinct} \tag{‡}$$

(Here I^r_i denotes the shift map on $\mathbb{U}_i$ of degree r.)

Statement (i) asserts that the simplex $[\mathbb{U}_1, \dots, \mathbb{U}_n]$ belongs to Čech$((\mathbf{mod}), d)$.

Statement (ii) without condition (‡) asserts that the simplex $[\mathbb{U}_1, \dots, \mathbb{U}_n]$ belongs to Rips$((\mathbf{mod}), 2d)$. The full statement, including condition (‡), asserts that the interleaving maps between the modules can be chosen to commute with each other (up to shift operators, which are needed to make the degrees of the maps agree).

In other words, a $2d$-Rips simplex is d-Čech if and only if the interleaving maps between the modules can be chosen to be a commuting family (up to shift operators). This suggests the following definition and restatement of the theorem.

Definition 3.3. A **coherent $2d$-interleaving** between modules $\mathbb{U}_1, \dots, \mathbb{U}_n$ is a family of maps Φ^j_i satisfying the conditions in statement (ii) of Theorem 3.2.

Theorem 3.2′. *A collection $\mathbb{U}_1, \dots, \mathbb{U}_n$ of modules has a d-witness if and only if it has a coherent $2d$-interleaving.*

The proof of the theorem is similar in spirit to the proof of the interpolation lemma given in [6]. We adopt the notation used in that paper for suspensions (i.e. shifts): given a module $\mathbb{U}$ and a real number r, let $\mathbb{U}[r]$ denote the module with $U[r]_t = U_{t+r}$ together with the natural maps. In other words $\mathbb{U}[r]$ is $\mathbb{U}$ 'shifted down' by r.

of Theorem 3.2. Suppose some persistence module $\mathbb{V}$ admits d-interleaving maps $\Psi_i \in \mathrm{Hom}^d(\mathbb{V}, \mathbb{U}_i)$ and $\Upsilon_i \in \mathrm{Hom}^d(\mathbb{U}_i, \mathbb{V})$ for each i. It is readily seen that the desired commuting family of d-interleaving maps is obtained by setting $\Phi^i_j = \Psi_j \Upsilon_i \in \mathrm{Hom}^{2d}(\mathbb{U}_i, \mathbb{U}_j)$.

Conversely, suppose there exists a coherent family of maps $\Phi^i_j \in \mathrm{Hom}^{2d}(\mathbb{U}_i, \mathbb{U}_j)$. To define $\mathbb{V}$, we consider the following two persistence modules:

$$\mathbb{V}^- = \mathbb{U}_1[-d] \oplus \mathbb{U}_2[-d] \oplus \dots \oplus \mathbb{U}_n[-d]$$
$$\mathbb{V}^+ = \mathbb{U}_1[+d] \oplus \mathbb{U}_2[+d] \oplus \dots \oplus \mathbb{U}_n[+d]$$

There is a map $\Gamma \in \mathrm{Hom}^0(\mathbb{V}^-, \mathbb{V}^+)$ defined by the following matrix:

$$\Gamma = \begin{bmatrix} I^{2d}_1 & \Phi^2_1 & \cdots & \Phi^n_1 \\ \Phi^1_2 & I^{2d}_2 & \cdots & \Phi^n_2 \\ \vdots & \vdots & & \vdots \\ \Phi^1_n & \Phi^2_n & \cdots & I^{2d}_n \end{bmatrix}$$

Since the image of a module map is itself a module, we can define $\mathbb{V} = \mathrm{im}(\Gamma)$.

We now show that $\mathbb{V}$ is d-interleaved with $\mathbb{U}_1$ (and hence, by symmetry, with each $\mathbb{U}_i$). To do this, we represent $\mathbb{U}_1$ in an unusual way. We consider the modules

$$\mathbb{W}^- = \mathbb{U}_1 \oplus \mathbb{U}_2[-2\mathrm{d}] \oplus \cdots \oplus \mathbb{U}_n[-2\mathrm{d}]$$
$$\mathbb{W}^+ = \mathbb{U}_1 \oplus \mathbb{U}_2[+2\mathrm{d}] \oplus \cdots \oplus \mathbb{U}_n[+2\mathrm{d}]$$

and the map $\Delta \in \mathrm{Hom}^0(\mathbb{W}^-, \mathbb{W}^+)$ defined by the matrix

$$\Delta = \begin{bmatrix} \mathrm{I}_1^0 & \Phi_1^2 & \cdots & \Phi_1^n \\ \Phi_2^1 & \mathrm{I}_2^{4\mathrm{d}} & \cdots & \mathrm{I}_n^{2\mathrm{d}}\Phi_2^n \\ \vdots & \vdots & & \vdots \\ \Phi_n^1 & \mathrm{I}_2^{2\mathrm{d}}\Phi_n^2 & \cdots & \mathrm{I}_n^{4\mathrm{d}} \end{bmatrix}$$

and set $\mathbb{W} = \mathrm{im}(\Delta)$.

The proof is completed by establishing two claims:

Claim 1: $\mathbb{V}, \mathbb{W}$ are d-interleaved.

Proof. We first observe that $\mathbb{V}^-, \mathbb{W}^-$ are d-interleaved through the maps:

$$\Psi^- = \mathrm{I}_1^{2\mathrm{d}} \oplus \mathrm{I}_2^0 \oplus \cdots \oplus \mathrm{I}_n^0 \ \in \mathrm{Hom}^{\mathrm{d}}(\mathbb{V}^-, \mathbb{W}^-)$$
$$\Upsilon^- = \mathrm{I}_1^0 \oplus \mathrm{I}_2^{2\mathrm{d}} \oplus \cdots \oplus \mathrm{I}_n^{2\mathrm{d}} \in \mathrm{Hom}^{\mathrm{d}}(\mathbb{W}^-, \mathbb{V}^-)$$

and $\mathbb{V}^+, \mathbb{W}^+$ are d-interleaved through the maps:

$$\Psi^+ = \mathrm{I}_1^0 \oplus \mathrm{I}_2^{2\mathrm{d}} \oplus \cdots \oplus \mathrm{I}_n^{2\mathrm{d}} \in \mathrm{Hom}^{\mathrm{d}}(\mathbb{V}^+, \mathbb{W}^+)$$
$$\Upsilon^+ = \mathrm{I}_1^{2\mathrm{d}} \oplus \mathrm{I}_2^0 \oplus \cdots \oplus \mathrm{I}_n^0 \ \in \mathrm{Hom}^{\mathrm{d}}(\mathbb{W}^+, \mathbb{V}^+)$$

One can readily verify that $\Psi^+\Gamma = \Delta\Psi^-$ and $\Upsilon^+\Delta = \Gamma\Upsilon^-$, and hence $\Psi^\pm$ and $\Upsilon^\pm$ determine maps:

$$\Psi \in \mathrm{Hom}^{\mathrm{d}}(\mathrm{im}(\Gamma), \mathrm{im}(\Delta)) = \mathrm{Hom}^{\mathrm{d}}(\mathbb{V}, \mathbb{W})$$
$$\Upsilon \in \mathrm{Hom}^{\mathrm{d}}(\mathrm{im}(\Delta), \mathrm{im}(\Gamma)) = \mathrm{Hom}^{\mathrm{d}}(\mathbb{W}, \mathbb{V})$$

Moreover, $\Upsilon^+\Psi^+ = \mathrm{I}^{2\mathrm{d}}$ and $\Psi^+\Upsilon^+ = \mathrm{I}^{2\mathrm{d}}$ imply that $\Upsilon\Psi = \mathrm{I}^{2\mathrm{d}}$ and $\Psi\Upsilon = \mathrm{I}^{2\mathrm{d}}$ so we have a d-interleaving. □

Claim 2: $\mathbb{W}$ is isomorphic to $\mathbb{U}_1$.

Proof. For this we observe that there is a matrix factorization

$$\Delta = \begin{bmatrix} \mathrm{I}_1^0 \\ \Phi_2^1 \\ \vdots \\ \Phi_n^1 \end{bmatrix} \begin{bmatrix} \mathrm{I}_1^0 & \Phi_1^2 & \cdots & \Phi_n^2 \end{bmatrix}$$

because of the properties (†) and (‡) of the coherent family. This means that Δ can be written as a composite

$$\mathbb{W}^- \longrightarrow \mathbb{U}_1 \longrightarrow \mathbb{W}^+$$

where the first map is surjective, thanks to its I_1^0 term; and the second map is injective, thanks to its I_1^0 term. (We are using the fact that I_1^0 is the identity map of $\mathbb{U}_1$.) It follows that $\mathrm{im}(\Delta)$ is isomorphic to $\mathbb{U}_1$. □

The two claims complete the proof of the theorem. □ □

3.4. Coherent matchings of diagrams. Let $\mathcal{D} = (D_1, \ldots, D_n)$ be a collection of n persistence diagrams. Given $d \geqslant 0$, we say that a persistence diagram E is a **d-witness** of $\mathcal{D}$ if there exist d-matchings $\gamma_i : E \to D_i$ for each i. The construction of Čech complexes in the space of persistence diagrams reduces to solving the following fundamental problem: *what is the smallest $d \geqslant 0$ such that there exists an d-witness of $\mathcal{D}$?* We answer this question with a necessary and sufficient condition intrinsic to $\mathcal{D}$ which guarantees the existence of a witness diagram.

Definition 3.4. *A subset $\Omega \subset D_1 \times \cdots \times D_n$ is called a* **coherent d-matching** *of the family $\mathcal{D} = (D_i)$ if the following conditions hold:*

(1) **diameter:** *each $P = (p_1, \ldots, p_n) \in \Omega$ has d^∞ diameter at most d;*
(2) **partition:** *each point $p \in D_i$ appears in the i-th component of at most one $P \in \Omega$; and*
(3) **diagonal:** *each $p \in D_i$ not appearing in any $P \in \Omega$ lies within d^∞-distance d of the diagonal.*

Here is our main result for Čech constructions in the space of persistence diagrams.

Theorem 3.5. *$\mathcal{D}$ has a d-witness E if and only if it admits a coherent $2d$-matching Ω.*

Proof. Assume the existence of a witness E along with d-matchings $\omega_i : E \to D_i$. For each point $e \in E$, the n-tuple $P_e = (\omega_1(e), \ldots, \omega_n(e))$ of points in the upper half plane has diameter at most $2d$. Either P_e contains a point on the diagonal, or it does not. Let Ω consist of all P_e which do not contain a diagonal point as e ranges over the points of E. It is easily seen that this constructs a coherent $2d$-matching.

On the other hand, assume that $\mathcal{D}$ admits a coherent $2d$-matching Ω. For each $P \in \Omega$ of diameter $2d$, there exists some point u_P in the upper half plane so that $\|u_P - p\|_\infty \leqslant d$ for each $p \in P$. Let Q be the multiset of all points in all the diagrams in $\mathcal{D}$ which do not appear in any $P \in \Omega$. Each $q \in Q$ lies within distance $2d$ of the diagonal, so there exists a point v_q within distance d of both the diagonal and of q. Let E consist of the union of all points u_P and v_q as P ranges over the elements of Ω and q ranges over the points in Q. Construct the matching $\omega_i : E \to D_i$ as follows: each point of D_i belonging to some $P \in \Omega$ is matched with u_P, and each point q of D_i within $2d$ of the diagonal is matched with v_q. By construction, all matched pairs have d^∞-distance smaller than d. ☐

Note that a coherent d-matching Ω of $\mathcal{D} = (D_i)$ induces pairwise matchings between each pair D_i, D_j via projection. In particular, one can define $\omega_{ij} : D_i \to D_j$ as follows. For each $P \in \Omega$, match the point of D_i in the i-th component of P with the corresponding point of D_j in the j-th component of P. Any points not matched in this fashion are d-close to the diagonal. If one restricts attention to the set of points not matched to the diagonal, one obtains a coherence condition among these pairwise matchings. That is, $\omega_{jk} \circ \omega_{ij}(p) = \omega_{ik}(p)$ for any $p \in D_i$ which appears inside an element of Ω. The reader is encouraged to compare this to the definition of coherent interleavings in Section 3.3.

3.5. Computational complexity. Given a collection of N points in a metric space and a scale $d > 0$, a naïve construction of the 1-skeleton of the associated Vietoris–Rips complex incurs a quadratic cost in N: one must test whether each possible pair of points is distance d apart or not. Depending on whether this 1-skeleton is almost disconnected or almost the complete graph on N vertices, the total number of simplices in the Rips complex can vary from N to 2^N. Thus, the complexity of constructing a Rips complex may vary dramatically even for a fixed number of vertices. On the other hand, various good approximations may be used for constructing the 1-skeleton, typically involving landmark or nearest-neighbor estimates. The reader is invited to consult

[19] for a survey of the most common approximation techniques as well as three practical algorithms for constructing Vietoris–Rips complexes from a given 1-skeleton.

Although interleaving maps between persistence modules form an integral part of stability analysis, there is little occasion in practical situations to actually compute optimal interleaving maps (coherent or otherwise). In sharp contrast, measuring the proximity of persistence diagrams is crucial to topological data analysis and consequently some care must be taken in establishing the worst-case algorithmic burden of computing distances in (**diag**). With this in mind, we provide a brief description of the computational cost of finding the smallest d for which there exists a d-witness to a collection of persistence diagrams.

Let $\mathcal{D} = (D_1, \ldots, D_n)$ be the finite collection of persistence diagrams in question. The following notions from graph theory will be useful in our complexity analysis. Recall that a graph Γ is n-partite if its vertex set may be divided into n non-empty subsets called *bins* so that there are no edges between two vertices in the same bin. Such a graph is called *complete* if in addition there exists an edge between any pair of vertices in different bins. A *clique* of the complete n-partite graph Γ is a collection of n vertices, one from each bin. If the edges of Γ are weighted then the weight of a clique is defined to be the maximum weight encountered between any pair of vertices in that clique. Finally, a *clique-decomposition* of Γ is a partition of the vertex set into cliques.

Note that one may construct a complete weighted n-partite graph $\Gamma_{\mathcal{D}}$ associated to $\mathcal{D}$ as follows. For each $i \in \{1, \ldots, n\}$ the i-th vertex bin $\Gamma_{\mathcal{D}}(i)$ consists of all the points in the persistence diagram D_i. The edge from a vertex in $\Gamma_D(i)$ to a vertex in $\Gamma_D(j)$ for $i \neq j$ is weighted by the ∞-norm distance between the corresponding points.

In light of Theorem 3.5, it suffices to find the smallest $d \geqslant 0$ such that $\mathcal{D}$ admits a coherent d-matching. But this is equivalent to the following graph theoretic problem: *what is the smallest* $d > 0$ *such that* $\Gamma_{\mathcal{D}}$ *admits a decomposition by cliques of weight less than* d*?* This problem has been well-studied in graph theory and is known to be NP-complete. Fortunately, it is not hard to find reasonable polynomial-time approximations. For instance, see [12] for a cubic-time algorithm which approximates the optimal clique decomposition problem.

4. A Lipschitz extension theorem

The Interpolation Lemma 2.1 and Theorem 3.2 turn out to be special cases of a very general result, Theorem 4.1, about Lipschitz extensions of maps into (**mod**).

Let $X = (X, d)$ be a metric space, and let $f : X \to (\mathbf{mod})$. We say that f is a **coherent embedding** if f is 1-Lipschitz and there exists a family of module maps

$$\Phi_y^x \in \mathrm{Hom}^{d(x,y)}(f(x), f(y)) \quad \text{for } x, y \in X$$

such that Φ_x^x is the identity on $f(x)$ for every $x \in X$, and $\Phi_z^y \Phi_y^x = \Phi_z^x I_x^{e(x,y,z)}$ for every $x, y, z \in X$. Here $e(x, y, z) = d(x, y) + d(y, z) - d(x, z)$, the shift required to get the degrees to match.

Theorem 4.1 (Lipschitz extension). *Let* X *be a metric space and let* A *be a subpace. Then any coherent embedding of* A *in* (**mod**) *extends to a coherent embedding of* X.

The proof involves a category theory argument that is beyond the scope of the present paper. Coherent embeddings can be thought of as functors on a category with objects $X \times \mathbf{R}$, and the extension problem then has a known solution. The details will appear in a future paper with Peter Bubenik.

We finish with some easy corollaries of the theorem.

EXAMPLE 4.2. Let X be the real interval $[0, d]$ and let A be the two-point subspace $\{0, d\}$. We recover the Interpolation Lemma.

EXAMPLE 4.3. Let $e_1, e_2, \dots, e_n$ denote the standard basis of $\mathbf{R}^n$ with the ℓ^1-metric. Let $X = \{0, e_1, \dots, e_n\}$ and let $A = \{e_1, \dots, e_n\}$. We recover Theorem 3.2 (after scaling by d).

EXAMPLE 4.4. Let $\mathbb{U}, \mathbb{V}, \mathbb{W}$ be modules which are pairwise a-, b- and c-interleaved (in the natural cyclic order). Suppose that a, b, c satisfy the triangle inequality, so that the quantities

$$\alpha = \frac{b+c-a}{2}, \quad \beta = \frac{c+a-b}{2}, \quad \gamma = \frac{a+b-c}{2}$$

are all nonnegative. Then there exists a module $\mathbb{G}$ which is α-interleaved with $\mathbb{U}$, β-interleaved with $\mathbb{V}$, and γ-interleaved with $\mathbb{W}$ if and only if the interleavings between $\mathbb{U}, \mathbb{V}, \mathbb{W}$ can be chosen to be a coherent family.

PROOF. In one direction, one can construct the coherent family of interleavings directly from the interleaving maps of $\mathbb{G}$ with $\mathbb{U}, \mathbb{V}, \mathbb{W}$. In the other direction, let X be a tree consisting of three edges of lengths α, β, γ meeting at a point, and let A be the three outer endpoints of these edges; then apply the theorem. □

The theorem indicates that algebraic coherence of interleaving maps is closely related to the geometric question of finding 1-Lipschitz extensions. More generally, we conjecture that there are versions of this last example which relate approximate algebraic coherence to δ-hyperbolicity.

5. Acknowledgements

The authors thank Peter Bubenik and Konstantin Mischaikow for valuable discussions. The first author held a Digiteo Chair at INRIA Saclay Île-de-France, where some of this work was carried out. The second author is grateful to Amit Patel for various insightful conversations.

References

[1] A. Björner, *Nerves, Fibers and Homotopy Groups*, Journal of Combinatorial Theory, 102: 88–93 (2003).
[2] P. Bubenik, J. Scott, *Categorification of persistent homology*, preprint, arXiv:1205.3669 (2012).
[3] F. Cagliari, C. Landi, *Finiteness of rank invariants of multidimensional persistent homology groups*, Applied Mathematics Letters, 24(4): 516–518 (2011).
[4] G. Carlsson, *Topology and Data*, Bulletin of the American Mathematical Society, 46(2): 255–308 (2009).
[5] F. Chazal, D. Cohen-Steiner M. Glisse, L. Guibas, S. Oudot, *Proximity of persistence modules and their diagrams*, Proceedings of the 25th annual Symposium on Computational Geometry (SoCG), 237–246 (2009).
[6] F. Chazal, V. de Silva, M. Glisse and S. Oudot, *The structure and stability of persistence modules*, preprint, arXiv:1207.3674 (2012).
[7] D. Cohen-Steiner, H. Edelsbrunner and J. Harer, *Stability of persistence diagrams*, Discrete and Computational Geometry, 37(1): 103–120 (2007).
[8] D. Cohen-Steiner, H. Edelsbrunner and D. Morozov, *Vines and vineyards by updating persistence in linear time*, Proceedings of the 22nd annual Symposium on Computational Geometry (SoCG), 119–126 (2006).
[9] H. Edelsbrunner and J. Harer, *Persistent homology—a survey*, in *Surveys on discrete and computational geometry*, Contemporary Mathematics, 453: 257–282 (2008).
[10] H. Edelsbrunner, D. Letscher and A. Zomorodian, *Topological persistence and simplification*, Discrete and Computational Geometry, 28: 511–533 (2002).
[11] R. Ghrist, *Barcodes: the persistent topology of data*, Bulletin of the American Mathematical Society, 45(1): 61–75 (2008).

[12] G. He, J. Liu, C. Zhao, *Approximation algorithms for some graph partitioning problems*, Journal of Graph Algorithms and Applications, 4(2): 1–11 (2000).
[13] M. Lesnick, *The theory of the interleaving distance on multidimensional persistence modules*, preprint, arXiv:1106.5305 (2011).
[14] Y. Mileyko, S. Mukherjee and J. Harer, *Probability measures on the space of persistence diagrams*, Inverse Problems, 27(12): 124007 (2012).
[15] N. Milosavljevic, D. Morozov and P. Skraba, *Zigzag persistent homology in matrix multiplication time*, Proceedings of the 27th annual Symposium on Computational Geometry (SoCG), 216–225 (2011).
[16] K. Mischaikow and V. Nanda, *Morse theory for filtrations and efficient computation of persistent homology*, to appear in Discrete and Computational Geometry (2013).
[17] F. Takens, *Detecting strange attractors in turbulence*, in *Dynamical Systems and Turbulence*, Springer Lecture Notes in Mathematics, 898 (1981).
[18] C. Webb, *Decomposition of graded modules*, Proceedings of the American Mathematical Society, 94(4): 565–571 (1985).
[19] A. Zomorodian, *Fast construction of the Vietoris–Rips complex*, Computers and Graphics Archive, 34(3): 263–271 (2010).
[20] A. Zomorodian and G. Carlsson, *Computing persistent homology*, Discrete and Computational Geometry, 33(2): 249–274 (2005).

Department of Mathematics, Pomona College, Claremont, California

E-mail address: vin.desilva@pomona.edu

Department of Mathematics, The University of Pennsylvania, Philadelphia, Pennsylvania

E-mail address: vnanda@sas.upenn.edu