Journal of Applied and Computational Topology (2025) 9:4
https://doi.org/10.1007/s41468-024-00200-w

Algebraic Wasserstein distances and stable homological invariants of data

Jens Agerberg[1] · **Andrea Guidolin**[1] · **Isaac Ren**[1] · **Martina Scolamiero**[1]

Received: 4 February 2023 / Revised: 29 November 2024 / Accepted: 10 December 2024 /
Published online: 20 January 2025

Abstract
Distances have a ubiquitous role in persistent homology, from the direct comparison of homological representations of data to the definition and optimization of invariants. In this article we introduce a family of parametrized pseudometrics between persistence modules based on the algebraic Wasserstein distance defined by Skraba and Turner, and phrase them in the formalism of noise systems. This is achieved by comparing p-norms of cokernels (resp. kernels) of monomorphisms (resp. epimorphisms) between persistence modules and corresponding bar-to-bar morphisms, a novel notion that allows us to bridge between algebraic and combinatorial aspects of persistence modules. We use algebraic Wasserstein distances to define invariants, called Wasserstein stable ranks, which are 1-Lipschitz stable with respect to such pseudometrics. We prove a low-rank approximation result for persistence modules which allows us to efficiently compute Wasserstein stable ranks, and we propose an efficient algorithm to compute the interleaving distance between them. Importantly, Wasserstein stable ranks depend on interpretable parameters which can be learnt in a machine learning context. Experimental results illustrate the use of Wasserstein stable ranks on real and artificial data and highlight how such pseudometrics could be useful in data analysis tasks.

Keywords Persistent homology · Persistence modules · Stable topological invariants of data · Wasserstein metrics

Mathematics Subject Classification 55N31 · 62R40 · 55U99

✉ Martina Scolamiero
scola@kth.se

Jens Agerberg
jensag@kth.se

Andrea Guidolin
guidolin@kth.se

Isaac Ren
isaacren@kth.se

[1] Department of Mathematics, KTH Royal Institute of Technology, 10044 Stockholm, Sweden

1 Introduction

While Topological Data Analysis (TDA) has historically focused on studying the global shape of data, persistent homology has since grown to provide popular techniques for incorporating both global topological features and local geometry into data analysis pipelines (Adams and Moy 2021). Through the lens of persistent homology, global topological features can be encoded by long bars in a barcode decomposition of the persistence module, while local geometric features are characterized by short bars in the barcode. Indeed, both the information of long bars and short bars in the barcode (Bendich et al. 2016; Hiraoka et al. 2016), as well as their location along the filtration scale (Stolz et al. 2017; Chachólski and Riihimäki 2020; Agerberg et al. 2021), turn out to be relevant in data analysis tasks. Introduced to persistent homology in Cohen-Steiner et al. (2010), Wasserstein distances offer a way to determine a trade-off between global and local features in persistence. Such distances have been widely used in applications and have been studied both from a combinatorial perspective and more recently with an algebraic approach (Bubenik et al. 2023; Skraba and Turner 2020). Wasserstein distances are parametrized by two parameters in $[1, \infty]$, commonly fixed to the values of 1, 2, and ∞. One of the aims of this article is to define a richer family of parametrized Wasserstein distances where, in addition to standard parameters determining sensitivity to short bars globally in the parameter space, a *contour* is introduced to locally weight different parts of the parameter space. We propose that the optimal parameter values for a particular task should be learned in a machine learning context. Our contribution is part of more general efforts of identifying parametrized families of metrics and invariants for persistence (Bubenik et al. 2015; Scolamiero et al. 2017; Hofer et al. 2017; Zhao and Wang 2019; Carrière et al. 2020).

Algebraic Wasserstein distances.
The study of algebraic distances between persistence modules is an active research direction in TDA, as demonstrated by the recent works on amplitudes (Giunti et al. 2024) and exact weights (Bubenik et al. 2023). In this article we provide a new proof that the p-norm of a persistence module, introduced in Skraba and Turner (2020), defines a pseudometric for all $p \in [1, \infty]$. While (Skraba and Turner 2020) constructs a correspondence between the pseudometric induced by the p-norm and the Wasserstein distance between persistence diagrams, our proof shows that the p-norm determines a noise system (Scolamiero et al. 2017) and therefore an induced pseudometric. Our approach easily generalizes to define new pseudometrics on persistence modules, as for example the pseudometrics $d^q_{\mathcal{S}^{p,C}}$ that combine p-norms with contours, effectively used as a reparametrization of the parameter space $[0, \infty)$. From a technical perspective, our framework requires to prove the axioms of noise systems without assuming that the p-norm induces a pseudometric (including the triangular inequality property) but rather studying how the p-norm interacts with monomorphisms, epimorphism, and short exact sequences. Among the axioms of noise systems, the one on short exact sequences (Lemma 4.15) is difficult to prove with our assumptions and to this purpose we introduce bar-to-bar morphisms, explained below.

It is interesting to see that Wasserstein distances fit in the noise system framework, as they are fundamentally different from noise systems that have been studied from a computational perspective so far. In fact, algorithms for the computation of stable ranks (that can be seen as vectorizations of persistence modules depending on the noise system) were only developed for so called *simple noise systems* (Chachólski and Riihimäki 2020; Gäfvert and Chachólski 2017). These noise systems have the extra property of being closed under direct sums, and can intuitively be thought of as being sensitive only to the longest bars, which leads to L^∞-type distances. The noise systems associated with algebraic Wasserstein distances for $p < \infty$ are of a different nature, and in particular they are not closed under direct sums.

From a practical and computational perspective, combinatorial distances between persistence diagrams are more straightforward to compute than algebraic distances between persistence modules. The combinatorial (p, C)-Wasserstein distances associated to $d^p_{\mathcal{S}^{p,C}}$ in Sect. 4.4 offer a convenient way to compute contour distances and the combination of contour and Wasserstein distances between persistence modules, relying on the already developed computational machinery for Bottleneck ($p = \infty$) and Wasserstein distances between persistence diagrams. In this article, however, our focus is not on the computation of the Wasserstein distance between two given persistence modules, but on invariants called Wasserstein stable ranks defined and computed using the distances.

Bar-to-bar morphisms.
The approach carried out in this article for proving that p-norms of persistence modules satisfy the axioms of noise systems relies on comparing monomorphisms (resp. epimorphisms) between persistence modules and so-called *bar-to-bar* monomorphisms (resp. epimorphisms) between the same persistence modules. Intuitively, in a bar-to-bar morphism (see Definition 3.1) every bar in the barcode decomposition of the domain maps non-trivially to at most one bar in the barcode decomposition of the codomain. Bar-to-bar morphisms are thus much simpler than general morphisms of persistence modules, and we show that they can be used to effectively reduce algebraic problems to easier problems of combinatorial nature. In particular, an important problem related to the definition and construction of algebraic distances is the minimization of kernels and cokernels of morphisms (see e.g. Definition 4.3) with respect to a chosen notion of "size", which in this article is the p-norm of persistence modules (Sect. 2.6) or a more general notion of norm combining p-norms and contours (Definition 4.7). Our main theoretical results Theorem 3.14 and Theorem 3.16 state that for any monomorphism (resp. epimorphism) between two persistence modules there exists a bar-to-bar monomorphism (resp. epimorphism) between the same persistence modules whose cokernel (resp. kernel) has smaller or equal norm.

Various types of bar-to-bar morphisms can be constructed. For example, as we observe in Sect. 3.4, there are bar-to-bar monomorphisms and epimorphisms associated with the induced matchings of Bauer and Lesnick (2015). Bar-to-bar morphisms are however more general then the induced matchings, and are also fundamentally different from other notions of matchings, such as the sub-barcode matchings of Chubet et al. (2022). However, our bar-to-bar morphisms can be used as a tool to prove that the monomorphism (resp. epimorphism) associated to the induced matching has

the cokernel (resp. kernel) with minimal p-norm among all monomorphisms (resp. epimorphisms) with the same domain and codomain. (Corollaries 3.20 and 3.22).

Several other key results of this article are proven leveraging Theorems 3.14 and 3.16. For example, we use these theorems to show that p-norms of persistence modules satisfy the axiom of noise systems on short exact sequences (Lemma 4.15). We also use our main results of Sect. 3 to prove a low-rank approximation result for persistence modules, analogous for example to the Eckart-Young-Mirsky theorem in the context of matrices (compare e.g. with Golub et al. 1987, Sect. 1 and references therein), where the notion of rank we use for persistence modules is the number of bars in the barcode decomposition. Given a persistence module X of rank k, for every $r \leq k$ we identify a persistence module of rank r that is closest to X in algebraic Wasserstein distance, and we express its distance from X (Propositions 4.30 and 4.32). These results and their generalization (Proposition 4.32) allow us to compute Wasserstein stable ranks (see Proposition 5.1), the invariants that we introduce in this article.

Wasserstein stable ranks, a class of learnable vectorizations.
The computation of Wasserstein distances between persistence modules remains expensive despite recent progress (Kerber et al. 2017), and the space of persistent modules is not directly amenable to statistical methods and machine learning. For these reasons, feature maps from persistence modules or diagrams have become an important component of the TDA machine learning pipeline. These techniques introduce a map between the space of persistence modules and a vector space where statistical and machine learning methods are well-developed. We propose a new class of feature maps, directly related to the Wasserstein distances $d^q_{\mathcal{S}^{p,C}}$ between persistence modules, and with interpretable, learnable parameters. Having fixed a pseudometric in the family of Wasserstein distances $d^q_{\mathcal{S}^{p,C}}$, the Wasserstein stable rank of a persistence module with respect to the chosen pseudometric can be explicitly computed with a formula (Proposition 5.1) derived from our results on monomorphisms and epimorphisms. The computational complexity of determining the Wasserstein stable rank is $O(n \log n)$ in the number n of bars of a persistence module.

A parametrized family of stable ranks can be obtained by varying the Wasserstein distances, opening up for the possibility to tune parameters for a particular task, resulting in feature maps that focus on the discriminative aspects of the persistence modules in a dataset. Previous learnable feature maps (Hofer et al. 2017; Carrière et al. 2020; Reinauer et al. 2021) make the choice of *expressiveness* (being able to learn any arbitrary function on the space of persistence modules) over *stability* (learning a function under the constraint that it is robust to perturbations of the input). Moreover, since the methods are often parametrized by complex neural networks, it is difficult to compare and interpret parametrizations learned for different tasks. Our Wasserstein stable ranks are stable by construction. More precisely, the interleaving distance between Wasserstein stable ranks is 1-Lipschitz with respect to the corresponding Wasserstein distance used in its construction. Similarly to Wasserstein stable ranks, we also provide a simple formula for computing the interleaving distance between them at the cost of $O(n \log n)$ in the maximum number of bars in the two persistence modules we are comparing.

We use a metric learning framework to learn an optimal parametrization for a problem at hand, observe that a better model can be obtained by jointly optimizing the parameters p and the ones related to the contour C and illustrate that the output can be readily interpreted in terms of the learned parametrization focusing on e.g. global/local features or various parts of the filtration scale. The methods are demonstrated on a synthetic and a real-world datasets.

Outline of the paper.
Section 2 contains background material. In Sect. 3 we prove results on the p-norm of the cokernel of a monomorphism and, dually, of the kernel of an epimorphism of persistence modules. Section 4 is a study of Wasserstein distances and their generalizations involving contours in the framework of noise systems. In Sect. 5 we compute Wasserstein stable ranks and interleaving distances between them, which we use to formulate a metric learning problem. In Sect. 6 we illustrate the use of Wasserstein stable ranks on synthetic and real-world data, learning optimal parameters of algebraic Wasserstein distances.

2 Preliminaries

2.1 Persistence modules and persistent homology

Let $[0, \infty)$ denote the totally ordered set of nonnegative real numbers, regarded as the category induced by the order structure. We consider an arbitrary fixed field K and denote by vect_K the category of finite dimensional vector spaces over K. A **persistence module over** K is a functor $X\colon [0,\infty) \to \mathrm{vect}_K$. Explicitly, X consists of a collection of finite dimensional vector spaces X_t for all t in $[0, \infty)$, together with a collection of linear functions $X_{s\leq t}\colon X_s \to X_t$, called **transition functions**, for all $s \leq t$ in $[0, \infty)$, such that $X_{s\leq t}X_{r\leq s} = X_{r\leq t}$ for all $r \leq s \leq t$, and $X_{t\leq t}$ is the identity function on X_t for all t in $[0, \infty)$. A **morphism** or natural transformation $f\colon X \to Y$ between two persistence modules X and Y is a collection of linear functions $f_t\colon X_t \to Y_t$, for all t in $[0, \infty)$, such that $f_t X_{s\leq t} = Y_{s\leq t} f_s$ for all $s \leq t$ in $[0, \infty)$.

A persistence module X is **tame** if there exist real numbers $0 = t_0 < t_1 < \cdots < t_k$ such that the transition function $X_{s\leq t}$ is a non-isomorphism only if $s < t_i \leq t$ for some $i \in \{1, \ldots, k\}$. We denote by Tame the category of tame persistence modules and morphisms between them. The class of objects of this category will be denoted by Tame as well.

Convention 2.1 In this article we always work in the category of tame persistence modules over a fixed field K. For brevity the term persistence module will be used to refer to tame persistence modules over K.

A morphism $f\colon X \to Y$ in Tame is a monomorphism (respectively, an epimorphism or isomorphism) if the linear functions $f_t\colon X_t \to Y_t$ are monomorphisms (respectively, epimorphisms or isomorphisms) of vector spaces, for all t in $[0, \infty)$. Kernels, cokernels and direct sums in Tame are defined componentwise. For example, for any persistence modules X and Y, the direct sum $X \oplus Y$ is the persistence module

defined by $(X \oplus Y)_t = X_t \oplus Y_t$ and $(X \oplus Y)_{s \leq t} = X_{s \leq t} \oplus Y_{s \leq t}$, for all $s \leq t$ in $[0, \infty)$. The zero persistence module or **zero module**, i.e., the functor identically equal to the zero vector space on objects, will be denoted by 0.

Let $a < b$ in $[0, \infty]$. We denote by $K(a, b)$ the persistence module defined as follows: for any t in $[0, \infty)$,

$$K(a,b)_t := \begin{cases} K & \text{if } a \leq t < b \\ 0 & \text{otherwise,} \end{cases}$$

and for any $s \leq t$ in $[0, \infty)$,

$$K(a,b)_{s \leq t} := \begin{cases} \mathrm{id}_K & \text{if } K(a,b)_s = K = K(a,b)_t \\ 0 & \text{otherwise.} \end{cases}$$

We call $K(a, b)$ the **bar** (or interval module) with **start-point** a and **end-point** b. We say that the bar $K(a, b)$ is **infinite** if $b = \infty$ and **finite** otherwise. We say that the left-closed, right-open interval $[a, b)$ in $[0, \infty)$ is the **support** of the bar $K(a, b)$. As an easy consequence of naturality, a morphism $f : K(a_1, b_1) \to K(a_2, b_2)$ between bars can be nonzero (i.e. have some component f_a different from the zero map) only if $a_2 \leq a_1 < b_2 \leq b_1$. In this case, $\ker f$ is isomorphic to $K(b_2, b_1)$ if $b_2 < b_1$, and is zero otherwise, and $\operatorname{coker} f$ is isomorphic to $K(a_2, a_1)$ if $a_2 < a_1$, and is zero otherwise.

A persistence module is **indecomposable** if, whenever it is isomorphic to a direct sum $Y \oplus Z$ with Y and Z in Tame, either $Y = 0$ or $Z = 0$. Bars are indecomposable and, as the following fundamental result implies, any indecomposable in Tame is isomorphic to a bar. We refer the reader to Chazal et al. (2016) for more details on the algebraic structure of persistence modules.

Theorem 2.2 (Structure of persistence modules) *Any (tame) persistence module X is isomorphic to a finite direct sum of bars of the form $\bigoplus_{i=1}^{k} K(a_i, b_i)$, with $a_i < b_i$ in $[0, \infty]$ for every $i \in \{1, \ldots, k\}$. This decomposition is unique up to permutation: if $X \cong \bigoplus_{i=1}^{k} K(a_i, b_i) \cong \bigoplus_{j=1}^{\ell} K(c_j, d_j)$, then $k = \ell$ and there exists a permutation σ on $\{1, \ldots, k\}$ such that $a_i = c_{\sigma(i)}$ and $b_i = d_{\sigma(i)}$, for every $i \in \{1, \ldots, k\}$.*

A decomposition of a persistence module X as a direct sum of bars as in Theorem 2.2 is called a **barcode decomposition** of X. In this article, we will occasionally denote a barcode decomposition of X by $\bigoplus_{i=1}^{k} X_i$ when we do not need an explicit notation for the bars' start- and end-points. The number k of bars in any barcode decomposition of X is called the **rank** of X, denoted by $\operatorname{rank}(X)$.

Given a persistence module X, consider an element $x \in X_a$ for some a in $[0, \infty)$, and let $b := \sup\{t \in [a, \infty) \mid X_{a \leq t}(x) \neq 0\}$ in $[a, \infty]$. The element x is called a **generator** of X if the morphism $g : K(a, b) \to X$ defined by $g_a(1) = x$ is such that the composition rg with some morphism $r : X \to K(a, b)$ is the identity on $K(a, b)$. We call $K(a, b)$ the **bar generated by** x, and we observe that it is a direct summand of X. We call a collection of elements $\{x_i \in X_{a_i}\}_{i=1}^{k}$ a **set of generators** of X if each

x_i generates a bar $K(a_i, b_i)$ and the morphisms $g_i \colon K(a_i, b_i) \to X$ defined by x_i induce an isomorphism $\bigoplus_{i=1}^{k} K(a_i, b_i) \to X$.

As we will use basic homological algebra methods in Tame, we remark that infinite bars $K(a, \infty)$, for all a in $[0, \infty)$, are free in Tame, and that the notions of free and projective coincide in Tame (see Bubenik and Milićević 2021 for details). Any bar $K(a, b)$ with $b < \infty$ admits a minimal free resolution of the form $0 \to K(b, \infty) \to K(a, \infty) \to K(a, b) \to 0$.

Remark 2.3 We note that $\mathrm{rank}(X)$ can be viewed as a classical homological invariant corresponding to the number of generators in a minimal free resolution of X, which yields an alternative definition of the rank that is applicable to multiparameter persistence modules (Scolamiero et al. 2017).

Lastly, let us briefly comment on a set theoretical detail regarding the category Tame. In Tame, the class of isomorphism classes of objects is a set, as a consequence of Theorem 2.2. In this article, we consider some class functions defined on Tame, and we occasionally refer to them simply as functions for brevity. Since all class functions on Tame we consider are constant on isomorphism classes of objects, they can be regarded as proper functions defined on the set of isomorphism classes of persistence modules.

2.2 Contours

Contours can be thought of as describing coherent ways to "flow" across the parameter space $[0, \infty)$ of persistence modules. In this article, we call **contour** a function $C \colon [0, \infty) \times [0, \infty) \to [0, \infty)$ such that, for all a, b, ε, τ in $[0, \infty)$, the following inequalities hold:

1. if $a \leq b$ and $\varepsilon \leq \tau$, then $C(a, \varepsilon) \leq C(b, \tau)$;
2. $a \leq C(a, 0)$;
3. $C(C(a, \varepsilon), \tau) \leq C(a, \varepsilon + \tau)$.

In Gäfvert and Chachólski (2017) contours are defined in the case of n-parameter persistence modules. Contours are further studied for 1-parameter persistence in Chachólski and Riihimäki (2020), where several concrete examples are given. In Chachólski and Riihimäki (2020), the definition of contour is slightly more general than ours; for example, $C(a, \varepsilon)$ can take the value ∞. Similar notions to contours appear in the literature by the name of *superlinear families of tranlations* (Bubenik et al. 2015) and *flows on posets* (de Silva et al. 2018).

A contour C is called an **action** if the inequalities of (2.) and (3.) are equalities, that is, if $a = C(a, 0)$ and $C(C(a, \varepsilon), \tau) = C(a, \varepsilon + \tau)$, for all a, ε, τ. A contour C is **regular** (Chachólski and Riihimäki 2020) if the following conditions hold:

- $C(-, \varepsilon) \colon [0, \infty) \to [0, \infty)$ is a monomorphism for all $\varepsilon \in [0, \infty)$;
- $C(a, -) \colon [0, \infty) \to [0, \infty)$ is a monomorphism whose image is $[a, \infty)$, for all $a \in [0, \infty)$.

The second condition of regular contours ensures that $C(a, 0) = a$, for any a in $[0, \infty)$, and that C is strictly increasing in the second variable: $C(a, \varepsilon) < C(a, \tau)$

whenever $\varepsilon < \tau$, for any a in $[0, \infty)$. For brevity, we call a contour C a **regular action** if it is both regular and an action.

Let C be a regular contour. For all $a \in [0, \infty)$, we define the function $\ell(a, -)$ to be the inverse of the function $C(a, -)\colon [0, \infty) \to [a, \infty)$, that is, $\ell(a, b) = C(a, -)^{-1}(b)$ for any $b \in [a, \infty)$, and we set $\ell(a, \infty) = \infty$. We call ℓ the **lifetime function** associated with C. We observe that, since regular contours are injective functions in the second variable, $\ell(a, b)$ is well-defined for every pair $a \leq b$. Throughout the article, the **lifetime of a bar** $K(a, b)$ **with respect to a contour** C is the value $\ell(a, b)$ of the lifetime function associated with C.

As a first example of contour we consider the **standard contour**, i.e. the function D defined by $D(a, \varepsilon) = a + \varepsilon$, for every $a, \varepsilon \in [0, \infty)$. Informally, the standard contour describes the most uniform way to flow in the parameter space $[0, \infty)$ of a persistence module, linearly with unitary speed. We now introduce a large family of contours, called **integral contours of distance type** (Chachólski and Riihimäki 2020; Agerberg et al. 2021), parametrized by certain real-valued functions. Let $f\colon [0, \infty) \to (0, \infty)$ be a Lebesgue measurable function, called here a **density**. For every $a, \varepsilon \in [0, \infty)$, let $D_f(a, \varepsilon)$ be the real number in $[a, \infty)$ such that

$$\varepsilon = \int_a^{D_f(a,\varepsilon)} f(x)\, dx,$$

which is uniquely defined since f takes strictly positive values. The function $D_f\colon [0, \infty) \times [0, \infty) \to [0, \infty)$ is a contour; moreover, it is regular and an action. We observe that, if the density f is the constant function with value 1, the distance type contour D_1 coincides with the standard contour.

2.3 Noise systems

Noise systems provide a way to quantify the size of persistence modules and to produce pseudometrics on Tame by comparing their sizes (Scolamiero et al. 2017). A **noise system** on Tame is a sequence $\mathcal{S} = \{\mathcal{S}_\varepsilon\}_{\varepsilon \in [0,\infty)}$ of subclasses of Tame such that:

- $0 \in \mathcal{S}_\varepsilon$, for all ε,
- $\mathcal{S}_\tau \subseteq \mathcal{S}_\varepsilon$ whenever $\tau \leq \varepsilon$,
- if $0 \to X_0 \to X_1 \to X_2 \to 0$ is a short exact sequence in Tame, then:
  - if $X_1 \in \mathcal{S}_\varepsilon$, then $X_0, X_2 \in \mathcal{S}_\varepsilon$,
  - if $X_0 \in \mathcal{S}_\varepsilon$ and $X_2 \in \mathcal{S}_\tau$, then $X_1 \in \mathcal{S}_{\varepsilon+\tau}$.

Given a noise system $\mathcal{S} = \{\mathcal{S}_\varepsilon\}_{\varepsilon \in [0,\infty)}$ it is natural to associate to each persistence module X the smallest ε such that $X \in \mathcal{S}_\varepsilon$. This defines a class function $\alpha_{\mathcal{S}}\colon \mathsf{Tame} \to [0, \infty]$ called in Giunti et al. (2024) the **amplitude** associated to $\mathcal{S}$.

A noise system $\mathcal{S} = \{\mathcal{S}_\varepsilon\}_{\varepsilon \in [0,\infty)}$ is **closed under direct sums** if $X \oplus Y \in \mathcal{S}_\varepsilon$ whenever $X, Y \in \mathcal{S}_\varepsilon$, for every $\varepsilon \in [0, \infty)$. Contours (Sect. 2.2) provide examples of noise systems satisfying this property. Given a contour C and any $\varepsilon \in [0, \infty)$, let

$$\mathcal{S}_\varepsilon := \{X \in \mathsf{Tame} \mid X_{a \leq C(a,\varepsilon)} = 0 \text{ for all } a \in [0, \infty)\}.$$

It is proved in (Gäfvert and Chachólski 2017, Prop. 9.4) that the sequence $\{\mathcal{S}_\varepsilon\}_{\varepsilon\in[0,\infty)}$ defined in this way is a noise system closed under direct sums. In particular, the noise system induced by the standard contour has components

$$\mathcal{S}_\varepsilon := \{X \in \mathsf{Tame} \mid X_{a\leq a+\varepsilon} = 0 \text{ for all } a \in [0,\infty)\},$$

and coincides with the **standard noise system** introduced in Scolamiero et al. (2017).

2.4 Pseudometrics between persistence modules

In this article, we call (extended) **pseudometric** on Tame a class function d assigning to any pair of persistence modules X, Y in Tame an element $d(X,Y) \in [0,\infty]$ such that the following conditions hold for any X, Y, Z:

- $d(X,Y) = d(Y,X)$,
- $d(X,Y) = 0$ whenever X is isomorphic to Y,
- $d(X,Z) \leq d(X,Y) + d(Y,Z)$.

The third condition, known as the triangle inequality, combined with the second one yields $d(X,Y) = d(X',Y')$ whenever $X \cong X'$ and $Y \cong Y'$. This definition of pseudometric coincides with Definition 3.3 in Bubenik et al. (2023) when considering the category Tame.

We now briefly explain how noise systems yield pseudometrics on Tame. Let $\mathcal{S}$ be a noise system on Tame. For any $\varepsilon \in [0,\infty)$, we say that two persistence modules X and Y are ε-**close** if there exists a persistence module Z and a pair of morphisms $X \xleftarrow{f} Z \xrightarrow{g} Y$ such that

$$\ker f \in \mathcal{S}_{\varepsilon_1}, \quad \operatorname{coker} f \in \mathcal{S}_{\varepsilon_2}, \quad \ker g \in \mathcal{S}_{\varepsilon_3}, \quad \operatorname{coker} g \in \mathcal{S}_{\varepsilon_4},$$

for some $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4 \in [0,\infty)$ such that $\varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \varepsilon_4 \leq \varepsilon$. Define

$$d_{\mathcal{S}}(X,Y) := \inf\{\varepsilon \in [0,\infty) \mid X \text{ and } Y \text{ are } \varepsilon\text{-close}\},$$

adopting the convention $\inf \varnothing = \infty$. As shown in (Scolamiero et al. 2017, Prop. 8.7), $d_{\mathcal{S}}$ is a pseudometric on Tame.

We remark that the pseudometric $d_{\mathcal{S}}$ associated with the standard noise system is equivalent to the interleaving distance (Lesnick 2015), as proved by (Gäfvert and Chachólski (2017), Prop. 12.2).

2.5 Hierarchical stabilization and stable rank

In the context of TDA, **hierarchical stabilization** is a method to convert a discrete invariant of persistence modules into a stable invariant suitable for data analysis. This technique has been studied in Scolamiero et al. (2017), Gäfvert and Chachólski (2017) in the case of multiparameter persistence modules, and has been further investigated

in Chachólski and Riihimäki (2020) in the case of one-parameter persistence. Hierarchical stabilization has a very general formulation, which allows for several choices of discrete invariants, and in principle is not restricted to categories of persistence modules. For the hierarchical stabilization of the rank, also called stable rank, some computational methods have been developed (Chachólski and Riihimäki 2020; Gäfvert and Chachólski 2017). In this article we will restrict our attention to the stable rank and further develop its computation.

Besides choosing a discrete invariant, hierarchical stabilization requires the choice of a pseudometric between persistence modules, which plays an active role in calculating the corresponding stable invariant. Consider the rank of a persistence module (Sect. 2.1) as a class function $\mathrm{rank}\colon \mathsf{Tame} \to \mathbb{N}$ mapping any persistence module X to the natural number $\mathrm{rank}(X)$.

Definition 2.4 Given a pseudometric d on Tame (Sect. 2.4), the **stable rank** of a persistence module X with respect to the pseudometric d is the function $\widehat{\mathrm{rank}}_d(X)\colon [0,\infty) \to [0,\infty)$ defined, for all $t \in [0,\infty)$, by

$$\widehat{\mathrm{rank}}_d(X)(t) := \min\{\mathrm{rank}(Y) \mid Y \in \mathsf{Tame} \text{ and } d(X,Y) \leq t\}.$$

We observe that the function $\widehat{\mathrm{rank}}_d(X)$ is non-increasing and takes values in $\mathbb{N}$, so it belongs to the set $\mathcal{M}$ of Lebesgue measurable functions $[0,\infty) \to [0,\infty)$.

To illustrate the stability of the invariant $\widehat{\mathrm{rank}}_d$, we consider a pseudometric $d_{\bowtie}$ on $\mathcal{M}$, called the **interleaving distance**, defined for all $f, g \in \mathcal{M}$ by

$$d_{\bowtie}(f,g) := \inf\{\varepsilon \in [0,\infty) \mid f(t) \geq g(t+\varepsilon) \text{ and } g(t) \geq f(t+\varepsilon), \text{ for all } t \in [0,\infty)\},$$

setting by convention $\inf \varnothing = \infty$. The stable rank then satisfies the following Lipschitz condition.

Proposition 2.5 (Scolamiero et al. 2017) *Let d be a pseudometric on* Tame*, and let X, Y be persistence modules. Then $d(X,Y) \geq d_{\bowtie}(\widehat{\mathrm{rank}}_d(X), \widehat{\mathrm{rank}}_d(Y))$.*

2.6 *p*-norms

In this subsection, we briefly review properties of p-norms that are useful for our work. For $p \in [1,\infty]$, the p-norm (also called L^p-norm) on $\mathbb{R}^n$ is the function $\|\cdot\|_p : \mathbb{R}^n \to [0,\infty)$ defined, for each $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, by

$$\|x\|_p := \begin{cases} \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}} & \text{for } p \in [1,\infty) \\ \max\{|x_i|\}_{i\in\{1,\ldots,n\}} & \text{for } p = \infty. \end{cases}$$

We note that $\|x\|_\infty = \lim_{p\to\infty} \|x\|_p$, for all $x \in \mathbb{R}^n$. The triangle inequality (or subadditivity condition) $\|x+y\|_p \leq \|x\|_p + \|y\|_p$, for all $x, y \in \mathbb{R}^n$, is also referred to as Minkowski inequality.

A fundamental property of p-norms on $\mathbb{R}^n$ is the following: for $x \in \mathbb{R}^n$ and for $1 \leq p \leq q \leq \infty$, the inequalities

$$\|x\|_q \leq \|x\|_p \leq n^{\left(\frac{1}{p}-\frac{1}{q}\right)} \|x\|_q \tag{1}$$

hold and are sharp, where by convention we set $\frac{1}{\infty} = 0$. We refer to the first inequality as the monotonicity property of p-norms.

The following elementary property of p-norms is useful in this work: for $p \in [1, \infty]$, if $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$ and $z = (x_1, \ldots, x_n, y_1, \ldots, y_m) \in \mathbb{R}^{n+m}$, then

$$\left\| \left(\|x\|_p , \|y\|_p \right) \right\|_p = \|z\|_p . \tag{2}$$

Finally, let us also observe that p-norms are permutation invariant, and that they preserve the order on $[0, \infty)^n$, meaning that if $x \leq y$ in $[0, \infty)^n$ according to the coordinate-wise order, then $\|x\|_p \leq \|y\|_p$.

In this article, we generally consider p-norms as functions from $[0, \infty]^n$ to $[0, \infty]$, extending the usual definition by setting $\|x\|_p = \infty$ whenever x has some coordinate $x_i = \infty$. All properties stated above still hold with this definition.

Following (Skraba and Turner 2020), we will consider p-norms of persistence modules, whose definition relies on the barcode decomposition (Sect. 2.1). Such norms are already introduced in (Chen and Edelsbrunner 2011) for persistence diagrams. For $p \in [1, \infty]$, the p-norm of a persistence module X having barcode decomposition $X \cong \bigoplus_{i=1}^{k} K(a_i, b_i)$ is defined by

$$\|X\|_p := \begin{cases} \left(\sum_{i=1}^{k} |b_i - a_i|^p \right)^{\frac{1}{p}} & \text{for } p \in [1, \infty) \\ \max\{|b_i - a_i|\}_{i \in \{1,\ldots,k\}} & \text{for } p = \infty. \end{cases}$$

For $p \in [1, \infty]$ and $\varepsilon \in [0, \infty)$, the class of tame persistence modules with p-norm smaller or equal to ε is denoted by:

$$\mathcal{S}_\varepsilon^p := \{X \in \mathsf{Tame} \mid \|X\|_p \leq \varepsilon\},$$

and we set $\mathcal{S}^p := \{\mathcal{S}_\varepsilon^p\}_{\varepsilon \in [0,\infty)}$.

3 Monomorphisms, epimorphisms, and their p-norms

In this section we introduce bar-to-bar morphisms between persistence modules (Definition 3.1), which can informally be described as morphisms such that every bar in the barcode decomposition of the domain maps non-trivially to at most one bar in the barcode decomposition of the codomain. Our aim is proving results (Theorems 3.14 and 3.16) which compare monomorphisms and epimorphisms between two persistence modules to bar-to-bar monomorphisms and epimorphisms between the same

persistence modules. These results allow us to reduce algebraic problems to much simpler combinatorial problems, as shown for example in Corollaries 3.20 and 3.22.

3.1 Free presentations of monomorphisms

Given a monomorphism $f\colon Z \hookrightarrow X$ between persistence modules, we want to determine the barcode decomposition of coker f. We briefly describe a method that uses free resolutions of the persistence modules Z and X.

Consider the diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & R_Z & \xrightarrow{i_Z} & G_Z & \xrightarrow{p_Z} & Z & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle f_R} & & \downarrow{\scriptstyle f_G} & & \downarrow{\scriptstyle f} & & \\ 0 & \longrightarrow & R_X & \xrightarrow{i_X} & G_X & \xrightarrow{p_X} & X & \longrightarrow & 0 \\ & & & & & & \downarrow{\scriptstyle q} & & \\ & & & & & & \operatorname{coker} f & & \end{array}$$

where the rows are (minimal) free resolutions of the persistence modules Z and X respectively, and q denotes the canonical epimorphism. The given morphism f induces a morphism $f_G\colon G_Z \to G_X$ between the modules of generators and a morphism $f_R\colon R_Z \to R_X$ between the modules of relations that make the diagram commutative (see e.g. Rotman 2009, Thm. 6.16). We have coker $f \cong \operatorname{coker}([f_G\ i_X]\colon G_Z \oplus R_X \to G_X)$, where the morphism $[f_G\ i_X]$ sends $(z, r) \in G_Z \oplus R_X$ to $f_G(z) + i_X(r)$. The isomorphism of cokernels is easy to prove, for example observing that the image of the composition qp_X is coker f given that both q and p_X are surjective and verifying via diagram chasing that its kernel coincides with the image of $[f_G\ i_X]\colon G_Z \oplus R_X \to G_X$.

In other words, we have a free presentation of coker f

$$G_Z \oplus R_X \xrightarrow{[f_G\ i_X]} G_X \twoheadrightarrow \operatorname{coker} f,$$

and we can use it to determine the barcode decomposition of coker f. More precisely, observing that coker f is isomorphic to the homology at the middle term of the free chain complex

$$G_Z \oplus R_X \xrightarrow{[f_G\ i_X]} G_X \longrightarrow 0,$$

we can compute the barcode decomposition of coker f by using the persistent homology algorithm on a matrix M_f representing the morphism $[f_G\ i_X]$, as we detail in Sect. 3.2. The persistent homology algorithm determines "pairings" of the basis elements of $G_Z \oplus R_X$ with the basis elements of G_X, which corresponds to the start- and end-point pairs of the bars of coker f.

In this section, we are interested in particular morphisms between persistence modules, which we call bar-to-bar morphisms.

Definition 3.1 A morphism $f\colon Z \to X$ of persistence modules is **bar-to-bar** if there are barcode decompositions $Z = \bigoplus_{i=1}^{m} Z_i$ and $X = \bigoplus_{j=1}^{n} X_j$ and there exist a subset $I \subseteq \{1, \dots, m\}$ and an injective function $\alpha\colon I \to \{1, \dots, n\}$ such that

$$f = \bigoplus_{i \in I} f_i \oplus \bigoplus_{i \in \{1,\dots,m\} \setminus I} g_i \oplus \bigoplus_{j \in \{1,\dots,n\} \setminus \alpha(I)} h_j, \tag{3}$$

where each $f_i := f|_{Z_i}$ is a nonzero morphism $Z_i \to X_{\alpha(i)}$, and where g_i denotes the zero morphism $Z_i \to 0$ and h_j denotes the zero morphism $0 \to X_j$.

Remark 3.2 If f is a bar-to-bar morphism as in (3), then $\ker f$ and $\operatorname{coker} f$ are easily determined recalling the case of a morphism between two bars (see Sect. 2.1), namely:

$$\ker f = \bigoplus_{i \in I} \ker f_i \oplus \bigoplus_{i \in \{1,\dots,m\} \setminus I} Z_i, \qquad \operatorname{coker} f = \bigoplus_{i \in I} \operatorname{coker} f_i \oplus \bigoplus_{j \in \{1,\dots,n\} \setminus \alpha(I)} X_j.$$

Furthermore, if f is a monomorphism, the fact that $\ker f$ vanishes implies that $I = \{1, \dots, m\}$, and the existence of the injective function α implies $m \leq n$. Dually, $\alpha(I) = \{1, \dots, n\}$ and $n \leq m$ if f is an epimorphism.

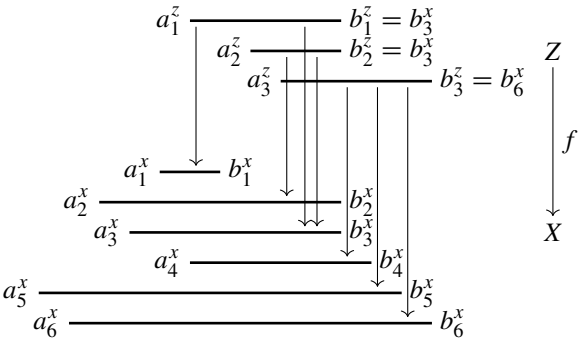
The main result of this section is the following (Theorem 3.14): given any monomorphism $f\colon Z \hookrightarrow X$, there is a bar-to-bar monomorphism $f_b\colon Z \hookrightarrow X$ such that $\| \operatorname{coker} f_b \|_p \leq \| \operatorname{coker} f \|_p$ for any $p \in [1, \infty]$. A dual statement (Theorem 3.16) holds for kernels of epimorphisms.

3.2 Finding monomorphisms with smaller cokernels

To prove our inequalities between p-norms of cokernels, we modify a strategy used in (Skraba and Turner 2020, Sect. 7.1) to obtain new inequalities between p-norms of persistence modules, based on the rearrangement inequality (Theorem 3.12) and on the comparison of pairings in certain barcode decompositions using the persistent homology algorithm. For simplicity, we fix the field with two elements $\mathbb{F}_2$ as the base field in this subsection, but our results work for any base field.

Let Z and X be persistence modules and $f\colon Z \hookrightarrow X$ a monomorphism of persistence modules. Fix $\{z_i\}_{i=1}^{m}$ and $\{x_j\}_{j=1}^{n}$ sets of generators of Z and X, respectively, and denote by $Z = \bigoplus_{i=1}^{m} K(a_i^z, b_i^z)$ and $X = \bigoplus_{j=1}^{n} K(a_j^x, b_j^x)$ the respective barcode decompositions. That is, for every z_i, a_i^z is the **degree** of $z_i \in Z_{a_i^z}$ and b_i^z is the end-point of the bar generated by z_i, and similarly for the x_j. In this section, we assume for the ease of exposition that X has no infinite bars in its decomposition. All the results we present can be adapted to the general case by setting $b_j^x = \infty$ whenever x_j generates an infinite bar. Throughout this section, we will consider an example

monomorphism f, which we represent as follows:



(4)

The persistence modules Z and X are represented in terms of their barcode decompositions. An arrow between bars indicates that the bar in the domain maps non-trivially to the bar in the codomain.

The main results of this subsection are based on matrix reduction arguments applied to a matrix M_f associated with the morphism $[f_G \; i_X] \colon G_Z \oplus R_X \to G_X$ (Sect. 3.1), which we construct as follows.

Definition 3.3 Define the sets of labels $\mathcal{G}_X := \{x_j\}_{j=1}^n$, $\mathcal{G}_Z := \{z_i\}_{i=1}^m$, and $\mathcal{R}_X := \{r_j\}_{j=1}^n$, where $\{z_i\}_{i=1}^m$ and $\{x_j\}_{j=1}^n$ are generators of Z and X respectively and r_j corresponds to the generator of R_X that is sent by i_X to the bar generated by x_j in G_X. The **degree** of r_j is b_j^x.

The **presentation matrix of** f is an $n \times (m+n)$ matrix M_f with rows labeled by $\mathcal{G}_X$ and columns labeled by $\mathcal{G}_Z \sqcup \mathcal{R}_X$, constructed as follows. For each z_i in $\mathcal{G}_Z$, we set the corresponding column of M_f to be the column vector $f_{a_i^z}(z_i) \in X_{a_i^z}$ in the basis given by the nonzero elements of $\{X_{a_j^x \le a_i^z}(x_j)\}_{j=1}^n$. Note that if $X_{a_j^x \le a_i^z}(x_j)$ is 0, then $M_f(x_j, z_i)$ is also 0. For each r_j in $\mathcal{R}_X$, we set the corresponding column of M_f to be the zero vector except with a 1 on the row x_j. Finally, we reorder the rows and columns so that the degrees of the labels are nondecreasing.

We denote by $M_f(x, c)$ the entry of M_f in row $x \in \mathcal{G}_X$ and column $c \in \mathcal{G}_Z \sqcup \mathcal{R}_X$.

As an example, one presentation matrix of the example monomorphism f from (4) is

$$M_f = \begin{array}{c} \\ x_5 \\ x_6 \\ x_2 \\ x_3 \\ x_1 \\ x_4 \end{array} \begin{array}{c} \begin{array}{ccccccccc} z_1 & r_1 & z_2 & z_3 & r_2 & r_3 & r_4 & r_5 & r_6 \end{array} \\ \left[\begin{array}{ccccccccc} 0 & & 0 & 1 & & & & 1 & \\ 0 & & 0 & 1 & & & & & 1 \\ 0 & & 1 & 0 & 1 & & & & \\ 1 & & 1 & 0 & & 1 & & & \\ 1 & 1 & 0 & 0 & & & & & \\ 0 & & 0 & 1 & & & 1 & & \end{array}\right] \end{array}, \tag{5}$$

where the columns $\mathcal{G}_Z = \{z_1, z_2, z_3\}$ are outlined, while the columns $\mathcal{R}_X = \{r_1, \ldots, r_6\}$ are represented sparsely: blank spaces are zero coefficients. Note that

the restriction of the matrix M_f to the columns $\mathcal{R}_X$ represents the morphism $i_X \colon R_X \to G_X$.

Remark 3.4 As we mentioned in Sect. 3.1, we want to determine the barcode decomposition of coker f by using the persistent homology algorithm on the matrix M_f representing the morphism $[f_G \; i_X]$. More precisely, we are interested in methods to compute barcode decompositions based on matrix reduction via left-to-right column operation, like the so-called standard algorithm for persistent homology (Edelsbrunner et al. 2000; Zomorodian and Carlsson 2005) (see Algorithm 1 in Otter et al. 2017 for a description). Even though these methods are usually presented for filtered simplicial complexes in the literature, they extend to graded free chain complexes as in our case. The barcode decomposition (of coker f in our case) can be read out from a reduced matrix, and does not depend on the way of reducing the matrix via left-to-right column operations (see Lemma 3.5).

Let $\bar{M}_f$ be a complete reduction of M_f by left-to-right column transformations, where a matrix is said to be **reduced** if no two columns have their lowest nonzero entry on the same row. Let σ_f be the function that to the k^{th} nonzero column of $\bar{M}_f$ associates the row of its lowest nonzero entry, for every $k \in \{1, \dots, n\}$. We know that σ_f is a permutation on $\{1, \dots, n\}$ since the n columns of M_f in $\mathcal{R}_X$ are linearly independent. In this section, we use square brackets for a permutation $\sigma = [\sigma(1) \cdots \sigma(n)]$ on $\{1, \dots, n\}$ expressed in one-line notation, to distinguish it from the notation for cycles, denoted by $(c_1 \, c_2 \, \cdots \, c_\ell)$. For the running example (5), we get

$$\bar{M}_f = \begin{bmatrix} 0 & & 0 & 1 & & & 1 & \\ 0 & & 0 & 1 & & 1 & & 0 \\ 0 & & 1 & 0 & 0 & & & \\ 1 & 1 & 0 & 0 & & 0 & & \\ 1 & 0 & 0 & 0 & & & & \\ 0 & & 0 & 1 & & & 0 & \end{bmatrix},$$

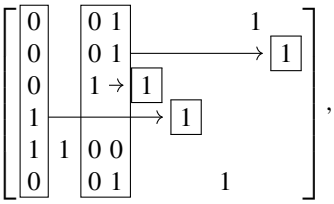
where we have outlined the lowest nonzero coefficient of each column, and so $\sigma_f = [543621]$. We do not need to specify the order of transformations in this reduction thanks to the following lemma, which is a consequence of the pairing uniqueness lemma of (Cohen-Steiner et al. 2006, Sect. 3).

Lemma 3.5 *The permutation σ_f is well-defined. In particular, it does not depend on the choice of a sequence of left-to-right column operations to obtain a reduced matrix from M_f.*

By design of the persistent homology algorithm, a barcode decomposition of coker f is completely determined by σ_f together with the degrees of the generators of Z and X. In Corollary 3.11 we will provide a precise statement.

From the matrix M_f we define the **bar-to-bar matrix** M_b by Algorithm 1. The bar-to-bar matrix M_b is the presentation matrix of a **bar-to-bar monomorphism** $f_b \colon Z \hookrightarrow X$ having the same domain and codomain as f.

Algorithm 1 also partially reduces M_f and constructs an injective function $r_{\max}\colon \mathcal{G}_Z \to \mathcal{R}_X$. Given a column z in $\mathcal{G}_Z$, we call $r_{\max}(z)$ its **rightmost matched column** in $\mathcal{R}_X$. Informally, Algorithm 1 computes the bar-to-bar matrix M_b by setting to zero each entry in the columns z of M_f in $\mathcal{G}_Z$ except for the nonzero entry on the unique row x such that $M_f(x, r_{\max}(z)) = 1$. For example, starting with the matrix M_f from (5), we get



where the arrows represent the function $r_{\max}$. The corresponding matrix M_b is

$$\begin{bmatrix} 0 & & 0\,0 & & & 1 & \\ 0 & & 0\,1 & & & & 1 \\ 0 & & 1\,0 & 1 & & & \\ 1 & & 0\,0 & & 1 & & \\ 0 & 1 & 0\,0 & & & & \\ 0 & & 0\,0 & & & 1 & \end{bmatrix}.$$

Algorithm 1 Bar-to-bar algorithm

Input: a presentation matrix M_f of a monomorphism f
Output: a partially reduced matrix M_f^*, the associated bar-to-bar matrix M_b, and a function $r_{\max}\colon \mathcal{G}_Z \to \mathcal{R}_X$

Let $M_b := M_f$
Let $M_f^* := M_f$
Set the columns $\mathcal{G}_Z$ of M_b to 0
for $r \in \mathcal{R}_X$ in decreasing order **do**
  Let x be the row associated to r (that is, $M_f^*(x, r) = 1$)
  if $\exists z \in \mathcal{G}_Z$ such that $M_f^*(x, z) = 1$ and $r_{\max}(z)$ is undefined **then**
    Let z be minimal such that $M_f^*(x, z) = 1$ and $r_{\max}(z)$ is undefined
    Set $M_b(x, z) = 1$
    Define $r_{\max}(z) := r$
    for $z' > z$ such that $M_f^*(x, z') = 1$ **do**
      Reduce column z' in M_f^* by column z to set to zero the entry in row x
      for $r' \in \mathcal{R}_X$ and x' the row associated to r', such that $r' < z'$ and $M_f^*(x', z') = 1$ **do**
        Reduce column z' in M_f^* by column r'
      end for
    end for
  end if
end for

The following two propositions prove useful facts regarding Algorithm 1.

Proposition 3.6 *In a presentation matrix M_f of a monomorphism $f: Z \hookrightarrow X$, all columns in $\mathcal{G}_Z$ are nonzero. Moreover, for every column $z \in \mathcal{G}_Z$, all columns in the set*

$$\Gamma(z) := \{r \in \mathcal{R}_X \mid r \text{ and } z \text{ have a nonzero entry on the same row}\}$$

have degree strictly larger than the degree of z, and $|\Gamma(z)|$ equals the number of nonzero entries of z.

Proof Since f is a monomorphism, it cannot send a generator of a bar of Z to zero, hence the columns in $\mathcal{G}_Z$ are nonzero. A nonzero entry in a column $z \in \mathcal{G}_Z$ indicates that the corresponding generator of a bar of Z maps non-trivially to the vector space generated by $X_{a^x \leq a^z}(x)$ for some x generating a bar in X, where a^x is the degree of x and a^z is the degree of z. This implies that the end-point of the bar of X generated by x has degree strictly larger than the degree of z. Lastly, the cardinality of $\Gamma(z)$ equals the number of nonzero entries of z because the columns in $\mathcal{R}_X$ form a permutation matrix of rank n. □

Proposition 3.7 *Let $f: Z \hookrightarrow X$ be a monomorphism and let M_f be a presentation matrix of f. The execution of Algorithm 1 on M_f returns a well-defined function $r_{\max}: \mathcal{G}_Z \to \mathcal{R}_X$ that is injective. Furthermore, for every column $z \in \mathcal{G}_Z$, the column $r_{\max}(z)$ is to the right of z.*

Proof We prove that, for every column $z \in \mathcal{G}_Z$, $r_{\max}(z)$ is well-defined and to the right of z. We proceed by induction on a natural number $m \geq 0$, proving the result for all monomorphisms $f: Z \hookrightarrow X$ with presentations such that $|\mathcal{G}_Z| = m$.

If $m = 1$ and $\mathcal{G}_Z = \{z\}$, then the algorithm sets $r_{\max}(z)$ to be the rightmost column in $\mathcal{R}_X$ having a nonzero entry on the same row as a nonzero entry of z, which exists and is to the right of z by Proposition 3.6.

Now suppose that the statement holds for every monomorphism presentation matrix with m columns in $\mathcal{G}_Z$. Let M_f be a presentation matrix such that $|\mathcal{G}_Z| = m + 1$. Algorithm 1 performs a 'for' loop (line 4) until the 'if' statement (line 6) is true, which by Proposition 3.6 must happen before the algorithm terminates. Let r_0 be the rightmost column in $\mathcal{R}_X$ such that there is a (minimal, i.e. leftmost) $z \in \mathcal{G}_Z$ with $M_f(x, z) = 1$, where x is the row associated to r_0. Again by Proposition 3.6, column r_0 is to the right of column z. The reductions in lines 11–14 of the algorithm transform M_f into a matrix M_f^* presenting a different monomorphism $f': Z \hookrightarrow X$. The morphism f' coincides with f on all generators of Z except for generator z', which is mapped to the nonzero element $f_{a^{z'}}(z') + f_{a^{z'}}(Z_{a^z \leq a^{z'}}(z))$, where a^z and $a^{z'}$ respectively denote the degrees of z and z'. We see that f' is a monomorphism via the following pointwise argument. For every degree a, the linear function $f_a: Z_a \to X_a$ has $\ker f_a = 0$, hence it maps nonzero elements in $\{Z_{a_i^z \leq a}(z_i)\}_{i=1}^m$ to linearly independent elements $\{y_j\}$ in $\operatorname{span}(X_{a_j^x \leq a}(x_j))_{j=1}^n$. We see that $f'_a: Z_a \to X_a$ satisfies the same linear independence property (which implies $\ker f'_a = 0$) because the set of image elements coincides, except for possibly an element y' replaced by $y' + y$, where y is a different element of the set.

In M_f^*, the only column in $\mathcal{G}_Z$ with nonzero entry in row x is z. By removing column z and row x, we obtain a matrix with m columns in $\mathcal{G}_Z$ which is again a presentation

matrix of a monomorphism. By induction hypothesis we know that the algorithm determines a function $r'_{\max}\colon \mathcal{G}_Z\backslash\{z\} \to \mathcal{R}_X$ whose image does not contain r_0 and the columns to its right. The function $r'_{\max}$ extends to a function $r_{\max}\colon \mathcal{G}_Z \to \mathcal{R}_X$ by defining $r_{\max}(z) := r_0$. Finally, we observe that the function $r_{\max}$ is injective by construction. □

Let us now go back to the reduction of presentation matrices. As with M_f, we can reduce M_b by left-to-right column transformations to get a reduced matrix $\bar{M}_b$. We denote by σ_b the permutation on $\{1, \ldots, n\}$ associated with the nonzero columns of $\bar{M}_b$, which is well-defined because the matrix M_b only has columns with at most one nonzero coefficient and has the same set of columns in $\mathcal{R}_X$ as M_f. In our running example, computing $\bar{M}_b$ gives us $\sigma_b = [453261]$.

After reduction via left-to-right column operations, the matrices $\bar{M}_f$ and $\bar{M}_b$ have non-zero columns with the same set of labels, as we will prove in Proposition 3.9.

Definition 3.8 Let $n \geq 1$ be an integer and σ a permutation on $\{1, \ldots, n\}$. An **inversion** of σ is a pair (i, j) of elements of $\{1, \ldots, n\}$ such that $i < j$ and $\sigma(i) > \sigma(j)$.

Given a permutation σ, we also give the name **inversion** to a transposition $(i\ j)$ such that $i < j$ and $\sigma(i) < \sigma(j)$: composing σ by $(i\ j)$ on the right creates an inversion.

Using inversions we can define a poset structure on permutations: we write $\sigma \leq \sigma'$ if there exist $k \geq 0$ and a composition of transpositions $\tau = \tau_1 \cdots \tau_k$ such that $\sigma\tau = \sigma'$ and, for all $i \leq k$, τ_i is an inversion of the permutation $\sigma\tau_1 \cdots \tau_{i-1}$. In what follows, we often call τ simply a **composition of inversions of** σ when it satisfies this property. Notice that $\leq$ is a partial order on S_n, the symmetric group on $\{1, \ldots, n\}$. With respect to this order, the identity permutation is the smallest element and the reverse permutation $[n\ n-1\ \ldots\ 2\ 1]$ is the largest element.

Proposition 3.9 *Let $f\colon Z \hookrightarrow X$ be a monomorphism, M_f be a presentation matrix of f and M_b be the bar-to-bar matrix computed via Algorithm 1. Let $\bar{M}_f$ and $\bar{M}_b$ be reduced matrices obtained from M_f and M_b respectively, and let σ_f and σ_b be the associated permutations. Then, the following facts hold:*

- *The nonzero columns of the reduced matrices $\bar{M}_f$ and $\bar{M}_b$ are in the same positions,*
- *$\sigma_f \geq \sigma_b$, that is, $\sigma_f = \sigma_b\tau$ with τ a composition of inversions of σ_b.*

Proof Since we can replace M_f with the output M_f^* of Algorithm 1, which has the same associated permutation σ_f (as it is obtained by partially reducing M_f), and following the proof of Proposition 3.7, we can assume that M_f satisfies the following property: let z_0 be the unique column of M_f in $\mathcal{G}_Z$ such that the column $r_0 := r_{\max}(z_0)$ is maximal in the total order on columns; then the only row x_0 such that $M_f(x_0, r_{\max}(z_0)) = 1$ has exactly one other nonzero entry, which is $M_f(x_0, z_0) = 1$. We prove the claims by induction on the number of columns in $\mathcal{G}_Z$.

If $\mathcal{G}_Z = \varnothing$, then there is nothing to prove: $M_f = M_b$ and they are reduced, so $\sigma_f = \sigma_b$.

Otherwise, we execute Algorithm 1 to get the bar-to-bar matrix M_b and the function $r_{\max}$. Let z_0 be the unique column of M_f in $\mathcal{G}_Z$ such that $r_0 := r_{\max}(z_0)$ is maximal in the total order on columns. By removing column z_0, we obtain a presentation matrix

M'_f of a monomorphism f' with a set of columns $\mathcal{G}'_Z$ strictly contained in $\mathcal{G}_Z$, to which we can apply our induction hypothesis: $\bar{M}'_f$ and $\bar{M}'_b$ have the same nonzero columns, and $\sigma'_f = \sigma'_b\tau$ for some composition of inversions τ of σ'_b. The matrix M'_b, computed by using Algorithm 1 on M'_f, can be equivalently obtained by removing column z_0 from M_b, since M_f satisfies the property stated at the beginning of the proof. See Example 3.10 for matrices M'_f, M'_b, $\bar{M}'_f$ and $\bar{M}'_b$ in the running example.

Let x_0 be the only row such that $M_f(x_0, r_0) = 1$. By the execution of Algorithm 1, no other column of M'_f has a nonzero coefficient on row x_0, and so we deduce that the reductions of the matrices M'_f and M'_b do not affect column r_0. Since by inductive hypothesis $\bar{M}'_f$ and $\bar{M}'_b$ have the same nonzero columns, this implies that column r_0 does not appear in the inversions of τ, meaning that $\tau = (s_1\, t_1)(s_2\, t_2)\cdots(s_k\, t_k)$ with $s_i \neq c'_{r_0}$ and $t_i \neq c'_{r_0}$ for all $i \in \{1, \ldots, k\}$, where c'_{r_0} denotes the relative position in $\{1, \ldots, n\}$ of column r_0 in the (totally ordered) set of nonzero columns of the reduced matrix $\bar{M}'_b$.

Now, let M_g be the matrix M_f where we modify the column z_0 by setting to zero all its entries except the one on row x_0. We reduce the matrix M_f first as for M'_f, and then we reduce the column z_0 by columns to its left, which does not affect the nonzero coefficient on row x_0: we denote the resulting matrix by M''_f. $\bar{M}_f$ is then obtained by completing the reduction using column z_0. We reduce M_g and M_b in similar fashion, following M'_f and M'_b, respectively. We observe the following facts.

- The nonzero columns of $\bar{M}_f$, $\bar{M}_g$, and $\bar{M}_b$ are the nonzero columns of $\bar{M}'_f$ and $\bar{M}'_b$, except we replace r_0 with z_0. This is clear by construction for the matrices $\bar{M}_g$ and $\bar{M}_b$, as the column z_0 coincides with r_0. For the matrix $\bar{M}_f$, observe that for every nonzero entry $M_f(x, z_0)$ on column z_0, there is a nonzero entry $M_f(x, r)$ in a column r to the left of r_0, which implies that r_0 gets zeroed out after the reduction as it is linearly dependent with a number of columns to its left.
- $\sigma_f = \sigma_g\tau'$ where $\tau' := (c_{z_0}\, c_1)(c_1\, c_2)\cdots(c_{k-1}\, c_k)$ and $c_1, \ldots, c_k, c_{r_0}$ are the relative positions in $\{1, \ldots, n\}$ of the nonzero columns of M''_f whose lowest nonzero entry is modified (is moved to a different row) when reducing to $\bar{M}_f$, with c_{z_0} and c_{r_0} respectively denoting the relative positions of column z_0 and r_0 in the set of nonzero columns of M''_f.
- $\sigma_g = \sigma'_f\gamma^{-1}$ and $\sigma_b = \sigma'_b\gamma^{-1}$ where $\gamma := (c_{z_0}\, c_{z_0}+1\, \cdots\, c_{r_0})$ represents a cyclic permutation of the nonzero columns between z_0 and r_0.

See Example 3.10 for concrete examples of these relationships. We deduce that

$$\begin{aligned}\sigma_f &= \sigma_g\tau' \\ &= \sigma'_f\gamma^{-1}\tau' \\ &= \sigma'_b\tau\gamma^{-1}\tau' \\ &= \sigma_b\gamma\tau\gamma^{-1}\tau'.\end{aligned}$$

By the definition of τ', it is a composition of inversions of σ_g. We conclude the induction step by showing that $\gamma\tau\gamma^{-1}$ is a composition of inversions of σ_b.

More precisely, we know that $\tau = (s_1\, t_1)\cdots(s_k\, t_k)$ is a composition of inversions of σ'_b, meaning that $(s_i\ t_i)$ is an inversion of the permutation $\sigma'_b(s_1\, t_1)\cdots(s_{i-1}\, t_{i-1})$, for every $i \in \{1,\ldots,k\}$, and we want to prove that $\gamma\tau\gamma^{-1} = (\gamma(s_1)\,\gamma(t_1))\cdots(\gamma(s_k)\,\gamma(t_k))$ is a composition of inversions of σ_b, meaning that $(\gamma(s_i)\,\gamma(t_i))$ is an inversion of the permutation $\sigma_b(\gamma(s_1)\,\gamma(t_1))\cdots(\gamma(s_{i-1})\,\gamma(t_{i-1}))$, for every $i \in \{1,\ldots,k\}$. First, we observe that $s_i < t_i$ implies $\gamma(s_i) < \gamma(t_i)$, since as observed earlier the relative position c'_{r_0} of column r_0 in the set of nonzero columns of $\bar{M}'_b$ does not appear in τ. Let us now denote

$$\begin{aligned}\sigma'_{i-1} &:= \sigma'_b(s_1\, t_1)\cdots(s_{i-1}\, t_{i-1}),\\ \sigma_{i-1} &:= \sigma_b(\gamma(s_1)\,\gamma(t_1))\cdots(\gamma(s_{i-1})\,\gamma(t_{i-1})).\end{aligned}$$

We have to prove that $\sigma'_{i-1}(s_i) < \sigma'_{i-1}(t_i)$ implies $\sigma_{i-1}(\gamma(s_i)) < \sigma_{i-1}(\gamma(t_i))$. This is a consequence of the equalities

$$\sigma_{i-1}(\gamma(s_i)) = \sigma_b\gamma(s_1\, t_1)\gamma^{-1}\gamma\cdots\gamma^{-1}\gamma(s_{i-1}\, t_{i-1})\gamma^{-1}\gamma(s_i) = \sigma'_{i-1}(s_i)$$

and of similar equalities for t_i. □

Example 3.10 From the example matrix M_f in (5), the induction hypothesis of Proposition 3.9 with $\mathcal{G}'_Z = \mathcal{G}_Z\backslash\{z_3\}$ gives the matrices

$$M'_f = \begin{bmatrix}0&0&&1\\0&0&&1\\0&1&1&\\1&0&&\\1&0&&\\0&0&1&\end{bmatrix},\quad M'_b = \begin{bmatrix}0&0&&1\\0&0&&1\\0&1&1&\\1&0&&\\0&0&1&\\0&0&&\end{bmatrix},$$

where the column z_3 is omitted, and the reduced matrices

$$\bar{M}'_f = \begin{bmatrix}0&0&&1\\0&0&&1\\0&1&0&\\1&0&&\\1&0&&\\0&0&1&\end{bmatrix},\quad \bar{M}'_b = \begin{bmatrix}0&0&&1\\0&0&&1\\0&1&0&\\1&0&&\\0&1&&\\0&0&1&\end{bmatrix}.$$

We find that $\sigma'_f = [543612]$ and $\sigma'_b = [453612]$, and so $\sigma'_f = \sigma'_b(1\,2)$, where $(1, 2)$ is indeed an inversion.

The induction step of Proposition 3.9 reduces the matrices

$$M_f = \begin{bmatrix}0&0&1&&1\\0&0&1&&1\\0&1&0&1&\\1&0&0&&\\1&0&0&&\\0&0&1&&\end{bmatrix},\quad M_g = \begin{bmatrix}0&0&0&&1\\0&0&1&&1\\0&1&0&1&\\1&0&0&&\\1&0&0&&\\0&0&1&&\end{bmatrix},\quad M_b = \begin{bmatrix}0&0&0&&1\\0&0&1&&1\\0&1&0&1&\\1&0&0&&\\0&0&0&&\\0&0&1&&\end{bmatrix}.$$

to

$$\bar{M}_f = \begin{bmatrix}0&&0&1&&1\\0&&0&1&1&0\\1&1&0&0&&\\1&0&0&0&&\\0&&0&1&&\end{bmatrix},\quad \bar{M}_g = \begin{bmatrix}0&&0&0&&1\\0&&0&1&&0\\1&1&0&0&0&\\1&0&0&0&&\\0&&0&0&1&\end{bmatrix},\quad \bar{M}_b = \begin{bmatrix}0&&0&0&&1\\0&&0&1&&0\\1&0&0&0&0&\\0&1&0&0&&\\0&&0&0&1&\end{bmatrix}.$$

We find that $\sigma_f = [543621]$, $\sigma_g = [543261]$, and $\sigma_b = [453261]$. Thus $\sigma_f = \sigma_g(4\,5)$, $\sigma_g = \sigma'_f(4\,5\,6)$, and $\sigma_b = \sigma'_b(6\,5\,4)$.

Corollary 3.11 *Let $f\colon Z \hookrightarrow X$ be a monomorphism, and let $f_b\colon Z \hookrightarrow X$ be the associated bar-to-bar monomorphism. Let $a_1 \leq a_2 \leq \ldots \leq a_n$ be the start-points of the bars of X, and let $b_1 \leq b_2 \leq \ldots \leq b_n$ be the degrees of the non-zero columns of $\bar{M}_f$. Then*

$$\operatorname{coker} f = \bigoplus_{j=1}^{n} K(a_j, b_{\sigma_f(j)}) \qquad \textit{and} \qquad \operatorname{coker} f_b = \bigoplus_{j=1}^{n} K(a_j, b_{\sigma_b(j)}).$$

Proof By Proposition 3.9, the real numbers $b_1 \leq b_2 \leq \ldots \leq b_n$ are also the degrees of the non-zero columns of $\bar{M}_b$. By design of the persistent homology algorithm, the barcode decomposition of coker f and coker f_b is then determined by pairing start-points $\{a_j\}$ with end-points $\{b_j\}$ following the permutations σ_f and σ_b respectively, and the claim follows. □

We state below the rearrangement inequality following (Vince 1990). Since the statement we need is slightly different from those we found in the literature, we include a short proof, which is a slight modification of the argument in (Vince 1990) and can be found also in (Steele 2004, p. 82).

Theorem 3.12 (Rearrangement inequality) *Let $g_1, g_2, \ldots, g_n$ be real valued functions defined on an interval $I \subseteq \mathbb{R}$ such that $g_{k+1} - g_k$ is a non-decreasing function, for all $k \in \{1, \ldots, n-1\}$, and let $b_1 \leq b_2 \leq \ldots \leq b_n$ be a sequence of elements of I. If $\rho \leq \sigma$ in S_n, then*

$$\sum_{k=1}^{n} g_k(b_{\rho(k)}) \geq \sum_{k=1}^{n} g_k(b_{\sigma(k)}).$$

Proof Since the argument we present can be iterated, it is enough to prove the statement for $\sigma = \rho\tau$ where $\tau = (i\ j)$ is an inversion: $i < j$ and $\rho(i) < \rho(j)$. We have

$$\begin{aligned}\sum_{k=1}^{n} g_k(b_{\rho(k)}) - \sum_{k=1}^{n} g_k(b_{\sigma(k)}) &= g_i(b_{\rho(i)}) + g_j(b_{\rho(j)}) - g_i(b_{\sigma(i)}) - g_j(b_{\sigma(j)}) \\ &= g_i(b_{\rho(i)}) + g_j(b_{\rho(j)}) - g_i(b_{\rho(j)}) - g_j(b_{\rho(i)}) \\ &= \big(g_j(b_{\rho(j)}) - g_i(b_{\rho(j)})\big) - \big(g_j(b_{\rho(i)}) - g_i(b_{\rho(i)})\big) \geq 0,\end{aligned}$$

where the last inequality follows from $b_{\rho(i)} \leq b_{\rho(j)}$ and from the fact that $g_j - g_i$ is non-decreasing. □

Corollary 3.13 *Let $a_1 \leq a_2 \leq \ldots \leq a_n$ and $b_1 \leq b_2 \leq \ldots \leq b_n$ be sequences of real numbers, and let $p \in [1, \infty)$. If $\rho \leq \sigma$ in S_n, then*

$$\sum_{k=1}^{n} |a_k - b_{\rho(k)}|^p \leq \sum_{k=1}^{n} |a_k - b_{\sigma(k)}|^p.$$

Proof Let $h_k(x) = |a_k - x|^p$. It is easy to check that the function $h_{k+1} - h_k$ is non-increasing for all $k \in \{1, \dots, n-1\}$, so we can apply Theorem 3.12 to the sequence of functions $g_k := -h_k$. □

Theorem 3.14 *For any monomorphism $f\colon Z \hookrightarrow X$ it is possible to determine (via Algorithm 1) a bar-to-bar monomorphism $f_b\colon Z \hookrightarrow X$ such that $\| \operatorname{coker} f_b \|_p \leq \| \operatorname{coker} f \|_p$, for all $p \in [1, \infty]$.*

Proof First, assume $p \in [1, \infty)$. The persistence modules coker f and coker f_b have barcode decompositions as in Corollary 3.11. Then, the claim follows from Corollary 3.13 applied to the permutations $\sigma_b \leq \sigma_f$ (Proposition 3.9). The claim for $p = \infty$ follows from taking the limit for $p \to \infty$ of both sides of the inequality $\|\operatorname{coker} b\|_p \leq \|\operatorname{coker} f\|_p$, recalling that $\lim_{p\to\infty} \|u\|_p = \|u\|_\infty$ for any vector $u \in \mathbb{R}^n$ (Sect. 2.6). □

3.3 Bar-to-bar epimorphisms

A similar result to Theorem 3.14 exists for epimorphisms and their kernels. To show this, we use a duality argument. The dualization of persistence modules has been studied extensively, see e.g. (Bauer and Lesnick 2015; Miller 2020; Bauer and Schmahl 2023). Here we dualize tame persistence modules indexed by $[0, \infty)$, which requires some special handling since in the setting of this article bars are only supported on left-closed right-open intervals. In this subsection, we abuse the terminology introduced in Sect. 2.1 by calling bars more general persistence modules with interval support and pointwise dimension at most 1, including persistence modules supported on intervals of the form $(a, b]$ instead of $[a, b)$. To simplify the exposition, we explicitly work with finite direct sums of bars instead of general tame persistence modules, which are only equal to direct sums of bars up to isomorphism.

Definition 3.15 Let $\mathsf{Bar}_{[0,\infty)}$ be the full subcategory of tame persistence modules whose objects are finite direct sums of bars. Similarly, let $\mathsf{Bar}_{(-\infty,0]}$ be the category of finite direct sums of bars indexed by the poset $(-\infty, 0]$ with the usual order. These categories are abelian and we can represent morphisms as matrices of morphisms between summands.

We consider the contravariant functor $(-)^\vee\colon \mathsf{Bar}_{[0,\infty)} \to \mathsf{Bar}_{(-\infty,0]}$ sending an object X in $\mathsf{Bar}_{[0,\infty)}$ to the functor $X^\vee\colon (-\infty, 0] \to \mathrm{vect}_K$ defined by $X^\vee_t := \hom(X_{-t}, K)$, for all t in $(-\infty, 0]$, with transition functions $X^\vee_{s\leq t} := \hom(X_{-t\leq -s}, K)$ given by precomposition by $X_{-t\leq -s}$, for all $s \leq t$ in $(-\infty, 0]$. Similarly, a contravariant functor $\mathsf{Bar}_{(-\infty,0]} \to \mathsf{Bar}_{[0,\infty)}$ is defined, which we also denote by $(-)^\vee$ with an abuse of notation. Both functors $(-)^\vee$ are exact.

If $X = K(a, b)$, we observe that $X^\vee$ is also a bar, but its support (i.e., the set of poset elements for which the functor $X^\vee$ takes a nonzero value) is the left-open, right-closed interval $(-b, -a]$. In this article, we are considering bars whose support is a left-closed, right-open interval in $\mathbb{R}$. To fix this, we can consider the **pointwise direct limit** functor $\varinjlim_{[0,-)}(-)\colon \mathsf{Bar}_{[0,\infty)} \to \mathsf{Bar}_{[0,\infty)}$ sending X to the persistence module

whose value at a is $\varinjlim_{t<a} X_t$ and whose transition functions are naturally defined. If $X = K(a,b)$, applying the pointwise direct limit functor yields the bar supported on $(a,b] \cap [0,\infty)$, whose dual $K(a,b)^\vee$ is the bar $K(-b,-a)$ in $\mathsf{Bar}_{(-\infty,0]}$, which is supported on $[-b,-a) \cap (-\infty,0]$. Similarly, one defines the pointwise direct limit functor $\varinjlim_{(-\infty,-)}(-)\colon \mathsf{Bar}_{(-\infty,0]} \to \mathsf{Bar}_{(-\infty,0]}$. The pointwise direct limit functors are exact. Applying the functor $(-)^\vee$ after the pointwise direct limit functor is therefore an exact functor, which we denote by $(-)^\dagger$.

To summarize, we have contravariant exact functors

$$(-)^\dagger\colon \mathsf{Bar}_{[0,\infty)} \to \mathsf{Bar}_{(-\infty,0]} \qquad \text{and} \qquad (-)^\dagger\colon \mathsf{Bar}_{(-\infty,0]} \to \mathsf{Bar}_{[0,\infty)}$$

sending the bar $K(a,b)$ supported on $[a,b)$ to the bar $K(-b,-a)$ supported on $[-b,-a) \cap (-\infty,0]$, and extended to the rest of the objects by additivity. These functors send morphisms to their transpose (when viewed as matrices). More precisely, given a morphism $f\colon \bigoplus_i K(a_i,b_i) \to \bigoplus_j K(c_j,d_j)$ written as the matrix $[f_{i,j}]_{i,j}$ with $f_{i,j}\colon K(a_i,b_i) \to K(c_j,d_j)$ for all i and j, the morphism $f^\dagger\colon \bigoplus_j K(-d_j,-c_j) \to \bigoplus_i K(-b_i,-a_i)$ can be written as the matrix $[f^\dagger_{j,i}]_{i,j}$.

As a consequence of the exactness of $(-)^\dagger$, for any morphism $f\colon X \to Y$ in $\mathsf{Bar}_{[0,\infty)}$ (resp. in $\mathsf{Bar}_{(-\infty,0]}$) we have

$$(\ker f)^\dagger \cong \operatorname{coker} f^\dagger \qquad \text{and} \qquad (\operatorname{coker} f)^\dagger \cong \ker f^\dagger.$$

Since the functors $(-)^\dagger$ send morphisms to their transpose, they send bar-to-bar morphisms to bar-to-bar morphisms. In particular, they send bar-to-bar monomorphisms to bar-to-bar epimorphisms, and vice-versa. It is also clear that the functors $(-)^\dagger$ preserve p-norms: $\|X^\dagger\|_p = \|X\|_p$.

Moreover, we can apply the theory of Sect. 3.2 to the category $\mathsf{Bar}_{(-\infty,0]}$, and in particular apply Theorem 3.14. We conclude with the following result:

Theorem 3.16 *For any epimorphism $f\colon Z \twoheadrightarrow X$ of persistence modules, it is possible to determine a bar-to-bar epimorphism $f_b\colon Z \twoheadrightarrow X$ such that $\|\ker f_b\|_p \leq \|\ker f\|_p$, for all $p \in [1,\infty]$.*

Proof As for the case of monomorphisms, we assume that X and Z are finite direct sums of finite bars. We apply Theorem 3.14 to $f^\dagger\colon X^\dagger \hookrightarrow Z^\dagger$ to get a bar-to-bar monomorphism g. We set $f_b := g^\dagger\colon Z \twoheadrightarrow X$ a bar-to-bar epimorphism, and we observe that

$$\|\ker f_b\|_p = \|\operatorname{coker} g\|_p \leq \|\operatorname{coker} f^\dagger\|_p = \|\ker f\|_p.$$

□

In the case $p = \infty$, a result similar to Theorems 3.14 and 3.16 based on induced matchings (see Sect. 3.4) instead of bar-to-bar monomorphisms and epimorphisms is proved in (Bauer and Lesnick 2020, Theorem 2).

3.4 Comparing bar-to-bar morphisms with induced matchings

In Bauer and Lesnick (2015) the authors introduce a construction similar to bar-to-bar morphisms, namely **induced matchings**. In particular, given persistence modules X and Z such that there exists a monomorphism $Z \hookrightarrow X$, the authors define a **canonical injection** from the multiset of bars of Z to the multiset of bars of X, where for all real numbers $b \in [0, \infty]$ the i^{th} longest bar of Z with end-point b is sent to the i^{th} longest bar of X with the same end-point. This injection induces a bar-to-bar monomorphism, which we define here:

Definition 3.17 Let Z and X be persistence modules such that there exists a monomorphism $Z \hookrightarrow X$, and fix barcode decompositions $Z = \bigoplus_{i=1}^{m} K(a_i^z, b_i^z)$ and $X = \bigoplus_{j=1}^{n} K(a_j^x, b_j^x)$. The induced matching of Bauer and Lesnick (2015) then corresponds to a canonical injection $\varphi\colon \{1, \dots, m\} \hookrightarrow \{1, \dots, n\}$ such that $a_i^z \geq a_{\varphi(i)}^x$ and $b_i^z = b_{\varphi(i)}^x$, for all $i \in \{1, \dots, m\}$.

We define the **monomorphism induced by the canonical injection** φ as the monomorphism $f_\varphi\colon Z \hookrightarrow X$ given by

$$f_\varphi = \bigoplus_{i=1}^{m} (K(a_i^z, b_i^z) \hookrightarrow K(a_{\varphi(i)}^x, b_{\varphi(i)}^x)) \oplus \bigoplus_{j \in \{1,\dots,n\} \setminus \operatorname{im} \varphi} (0 \hookrightarrow K(a_j^x, b_j^x)).$$

Note that this is a bar-to-bar monomorphism.

Remark 3.18 Let $f\colon Z \hookrightarrow X$ be a monomorphism. If the bars of X all have distinct end-points, then the monomorphism induced by the canonical injection (Definition 3.17) coincides, up to isomorphism, with the bar-to-bar monomorphism f_b as determined in Sect. 3.2. This is because, in this case, there is only one bar-to-bar monomorphism from Z to X (up to isomorphism).

Remark 3.19 In general, this is not necessarily the case. For example, starting from the monomorphism $f\colon K(2,3) \hookrightarrow K(1,3) \oplus K(2,3)$ defined by

$$f = (K(2,3) \hookrightarrow K(2,3)) \oplus (0 \hookrightarrow K(1,3)),$$

then we have $f_\varphi = (K(2,3) \hookrightarrow K(1,3)) \oplus (0 \hookrightarrow K(2,3))$, while $f_b = f$.

More generally, monomorphisms induced by canonical injections do not commute with direct sums of monomorphisms (Bauer and Lesnick 2015, Example 5.8), while the bar-to-bar monomorphisms we introduced in Sect. 3.2 do, as an immediate consequence of their definition via Algorithm 1. That is, given $f\colon Z \hookrightarrow X$ and $f'\colon Z' \hookrightarrow X'$, it is not true in general that $(f \oplus f')_\varphi = f_\varphi \oplus f'_\varphi$, while $(f \oplus f')_b = f_b \oplus f'_b$ holds.

Maintaining a close relation between a given monomorphism $f\colon Z \hookrightarrow X$ and the bar-to-bar monomorphism f_b is instrumental to proving Proposition 3.9, the key technical result of Sect. 3.2, and in turn Theorem 3.14 and the dual Theorem 3.16. Since the canonical injection (Bauer and Lesnick 2015) does not depend on the specific given

monomorphism between Z and X, but just on the existence of such a monomorphisms, it is not possible to prove Proposition 3.9 using the same strategy with the canonical injection instead of the bar-to-bar monomorphism we introduced.

We conclude this section with two corollaries to Theorems 3.14 and 3.16. Using Theorem 3.14, the problem of minimizing the p-norm of cokernels of all possible monomorphisms between two given persistence modules can be solved by restricting to bar-to-bar monomorphisms. This simplification allows for an easy combinatorial proof of the fact that, among all monomorphisms, the one induced by the canonical injection minimizes the p-norm of the cokernel. A dual result holds for epimorphisms.

Corollary 3.20 *Let Z and X be two persistence modules such that there exists a monomorphism $Z \hookrightarrow X$. Then $\min_{f\colon Z\hookrightarrow X} \| \operatorname{coker} f\|_p = \| \operatorname{coker} f_\varphi\|_p$ for all $p \in [1,\infty]$, where f_φ is the bar-to-bar monomorphism induced by the canonical injection (Definition 3.17).*

Proof By Theorem 3.14, it suffices to show the inequality $\| \operatorname{coker} f_\varphi\|_p \leq \| \operatorname{coker} f\|_p$ for bar-to-bar monomorphisms $f\colon Z \hookrightarrow X$. Let $\{b_k\}_{k=1}^{\ell}$ be the set of distinct end-points of X. Then, by hypothesis that there exists a monomorphism $Z \hookrightarrow X$, we can decompose $Z = \bigoplus_{k=1}^{\ell} Z^{(k)}$ and $X = \bigoplus_{k=1}^{\ell} X^{(k)}$ where $Z^{(k)}$ (resp. $X^{(k)}$) is the direct sum of the bars in Z (resp. X) with end-point b_k. Given a bar-to-bar monomorphism $f\colon Z \hookrightarrow X$, this induces monomorphisms $f^{(k)}\colon Z^{(k)} \hookrightarrow X^{(k)}$. We then observe that $f = \bigoplus_{k=1}^{\ell} f^{(k)}$, and so

$$\| \operatorname{coker} f\|_p = \left\| \left(\| \operatorname{coker} f^{(k)}\|_p \right)_{k=1,\dots,\ell} \right\|_p .$$

Since p-norms are nondecreasing with respect to the coordinate-wise order on $[0,\infty)^\ell$ (Sect. 2.6), proving that $\| \operatorname{coker} f_\varphi^{(k)}\|_p \leq \| \operatorname{coker} f^{(k)}\|_p$ for each $k \in \{1,\dots,\ell\}$ implies that $\| \operatorname{coker} f_\varphi\|_p \leq \| \operatorname{coker} f\|_p$.

Thus it suffices to prove the result in the case where the bars of Z and X all have the same end-point, which is what we assume for the rest of the proof. Denote by b_0 this common end-point and write $Z = \bigoplus_{i=1}^{m} K(a_i^z, b_0)$ and $X = \bigoplus_{j=1}^{n} K(a_j^x, b_0)$, where the a_i^z and a_j^x are in nondecreasing order. Define $a_i^z = b_0$ for $i \in \{m+1,\dots,n\}$. Then every bar-to-bar monomorphism $f\colon Z \hookrightarrow X$ has the form

$$f = \bigoplus_{i=1}^{n} (K(a_i^z, b_0) \hookrightarrow K(a_{\alpha(i)}^x, b_0)),$$

where $\alpha\colon \{1,\dots,n\} \to \{1,\dots,n\}$ is a permutation and $K(b_0, b_0)$ denotes the zero module. By Remark 3.2,

$$\operatorname{coker} f = \bigoplus_{i=1}^{n} K\left(a_{\alpha(i)}^x, a_i^z\right).$$

In particular, the bar-to-bar monomorphism induced by the canonical injection corresponds to the permutation $\alpha = \mathrm{id}$. For all $p \in [1,\infty)$, we then apply the rearrangement

inequality of Corollary 3.13 to deduce that $\| \operatorname{coker} f_\varphi \|_p \leq \| \operatorname{coker} f \|_p$. The case $p = \infty$ follows by taking the limit (as in the proof of Theorem 3.14). □

We now dualize the definitions and results for the case of epimorphisms. An epimorphism $f : Z \twoheadrightarrow X$ induces a canonical injection (Bauer and Lesnick 2015) from the multiset of bars of X to the multiset of bars of Z, where for all $a \in [0, \infty]$, the i^{th} longest bar of X with start-point a is sent to the i^{th} longest bar of Z with the same start-point.

Definition 3.21 Let Z and X be persistence modules such that there exists an epimorphism $Z \twoheadrightarrow X$, and fix barcode decompositions $Z = \bigoplus_{i=1}^{m} K(a_i^z, b_i^z)$ and $X = \bigoplus_{j=1}^{n} K(a_j^x, b_j^x)$. The induced matching of Bauer and Lesnick (2015) then corresponds to a canonical injection $\psi : \{1, \ldots, n\} \hookrightarrow \{1, \ldots, m\}$ such that $a_j^x = a_{\psi(j)}^z$ and $b_j^x \leq b_{\psi(j)}^z$, for all $j \in \{1, \ldots, n\}$.

We define the **epimorphism induced by the canonical injection** ψ as the epimorphism $f_\psi : Z \twoheadrightarrow X$ given by

$$f_\psi = \bigoplus_{j=1}^{n} \left(K\left(a_{\psi(j)}^z, b_{\psi(j)}^z\right) \twoheadrightarrow K\left(a_j^x, b_j^x\right) \right) \oplus \bigoplus_{i \in \{1,\ldots,m\} \setminus \operatorname{im} \psi} \left(K\left(a_i^x, b_i^x\right) \twoheadrightarrow 0 \right).$$

As in the case of monomorphisms, given an epimorphism $f : Z \twoheadrightarrow X$, the associated bar-to-bar epimorphisms f_ψ and f_b (see Sect. 3.3) are not necessarily the same.

Corollary 3.22 *Let Z and X be two persistence modules such that there exists an epimorphism $Z \twoheadrightarrow X$. Then $\min_{f : Z \twoheadrightarrow X} \| \ker f \|_p = \| \ker f_\psi \|_p$ for all $p \in [1, \infty]$, where f_ψ is a bar-to-bar epimorphism induced by the canonical injection (Definition 3.21).*

We omit the proof of Corollary 3.22, which is based on Theorem 3.16 and analogous to the proof of Corollary 3.20.

4 Noise systems and Wasserstein pseudometrics

In this section we study algebraic Wasserstein pseudometrics between persistence modules. After introducing in Sect. 4.1 a generalization of the pseudometrics associated with a noise system, we study in Sect. 4.2 noise systems determined by p-norms of persistence modules and regular contours. Section 4.3 is devoted to the associated algebraic Wasserstein pseudometrics. For some choices of parameters, these pseudometrics have a combinatorial interpretation, as we show in Sect. 4.4. Finally, in Sect. 4.5 we present formulas to compute the algebraic Wasserstein pseudometric between persistence modules in some specific cases.

4.1 Pseudometrics associated to noise systems

Given a noise system $\mathcal{S}$ and $p \in [1, \infty]$, in this section we will introduce pseudometrics $d_{\mathcal{S}}^p$ between persistence modules. These pseudometrics are a simple generalization to

$p > 1$ of the pseudometric associated to a noise system in Scolamiero et al. (2017) (see Sect. 2.4), where $p = 1$. Although the statements in this section hold true for tame functors indexed by $[0, \infty)^r$ for every positive natural number r, as in Scolamiero et al. (2017), we will limit the presentation to $r = 1$, since this is the setting of the following sections.

Definition 4.1 Let X and Y be persistence modules. A **span** of X, Y is a triplet (Z, f, g) with Z a persistence module and $f\colon Z \to X$ and $g\colon Z \to Y$ morphisms between persistence modules. A span of X, Y is therefore a diagram in Tame of the form

$$X \xleftarrow{f} Z \xrightarrow{g} Y.$$

Definition 4.2 Let X and Y be persistence modules, and let $\mathcal{S}$ be a noise system. A span $X \xleftarrow{f} Z \xrightarrow{g} Y$ is called a $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$**-span** if

$$\ker f \in \mathcal{S}_{\varepsilon_1}, \quad \operatorname{coker} f \in \mathcal{S}_{\varepsilon_2}, \quad \ker g \in \mathcal{S}_{\varepsilon_3} \quad \text{and} \quad \operatorname{coker} g \in \mathcal{S}_{\varepsilon_4}.$$

Definition 4.3 Let X and Y be persistence modules, and let $\mathcal{S}$ be a noise system. For $p \in [1, \infty]$ and $\varepsilon \in [0, \infty)$, we say that X and Y are ε**-close in** p**-norm** $\|\cdot\|_p$ if there exists a $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$-span $X \xleftarrow{f} Z \xrightarrow{g} Y$ for some $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4 \in [0, \infty)$ such that $\|(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)\|_p \leq \varepsilon$. We define

$$d_{\mathcal{S}}^p(X, Y) := \inf \{\varepsilon \in [0, \infty) \mid X \text{ and } Y \text{ are } \varepsilon\text{-close in } p\text{-norm}\},$$

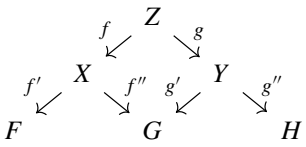
adopting the convention $\inf \varnothing = \infty$.

Our next aim is to prove that $d_{\mathcal{S}}^p$ is a pseudometric on Tame. We start by generalizing Proposition 8.5 in Scolamiero et al. (2017) to our current framework. Even if the generalization is not difficult, we include the proof to highlight how the properties of p-norms on $\mathbb{R}^4$ are used. We note that a similar result can be obtained for a larger family of subadditive functions on $\mathbb{R}^4$ which include p-norms (see Giunti et al. 2024, Sect. 2.1).

Proposition 4.4 *Let F, G, H be persistence modules. Assume that F and G are ε-close in p-norm, and that G and H are τ-close in p-norm. Then F and H are $(\varepsilon + \tau)$-close in p-norm.*

Proof By assumption there exists a $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$-span $F \xleftarrow{f'} X \xrightarrow{f''} G$ with $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4 \in [0, \infty)$ such that $\|(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)\|_p \leq \varepsilon$ and a $(\tau_1, \tau_2, \tau_3, \tau_4)$-span $G \xleftarrow{g'} Y \xrightarrow{g''} H$ with $\tau_1, \tau_2, \tau_3, \tau_4 \in [0, \infty)$ such that $\|(\tau_1, \tau_2, \tau_3, \tau_4)\|_p \leq \tau$. Con-

sider the following diagram, where the square is a pullback:



By Scolamiero et al. (2017, Proposition 8.1), $\ker f \in \mathcal{S}_{\tau_1}$ and $\operatorname{coker} f \in \mathcal{S}_{\tau_2}$, hence by Scolamiero et al. (2017, Proposition 8.2) $\ker f' f \in \mathcal{S}_{\varepsilon_1+\tau_1}$ and $\operatorname{coker} f' f \in \mathcal{S}_{\varepsilon_2+\tau_2}$. By a similar argument, $\ker g'' g \in \mathcal{S}_{\varepsilon_3+\tau_3}$ and $\operatorname{coker} g'' g \in \mathcal{S}_{\varepsilon_4+\tau_4}$. This proves that F and H are η-close in p-norm, where $\eta := \|(\varepsilon_1+\tau_1, \varepsilon_2+\tau_2, \varepsilon_3+\tau_3, \varepsilon_4+\tau_4)\|_p$. Our claim follows from the inequality

$$\begin{aligned}\|(\varepsilon_1+\tau_1, \varepsilon_2+\tau_2, \varepsilon_3+\tau_3, \varepsilon_4+\tau_4)\|_p &\leq \|(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)\|_p + \|(\tau_1, \tau_2, \tau_3, \tau_4)\|_p \\ &\leq \varepsilon + \tau,\end{aligned}$$

which expresses the subadditivity of $\|\cdot\|_p$ and the hypotheses. □

We are now ready to prove that $d_{\mathcal{S}}^p$ is a pseudometric on Tame.

Proposition 4.5 *Given $p \in [1, \infty]$ and a noise system $\mathcal{S}$, the function $d_{\mathcal{S}}^p$ in Definition 4.3 is a pseudometric on* Tame *(see Sect. 2.4).*

Proof If $g : X \to Y$ is an isomorphism of persistence modules, the span $X \xleftarrow{\mathrm{id}} X \xrightarrow{g} Y$ shows that $d_{\mathcal{S}}^p(X, Y) = 0$. For all persistence modules X and Y, the bijection between spans $X \xleftarrow{f} Z \xrightarrow{g} Y$ between X and Y and spans $Y \xleftarrow{g} Z \xrightarrow{f} X$ between Y and X implies that $d_{\mathcal{S}}^p(X, Y) = d_{\mathcal{S}}^p(Y, X)$. Proposition 4.4 shows that the triangle inequality holds true. □

Remark 4.6 Given a noise system $\mathcal{S}$, the pseudometrics $d_{\mathcal{S}}^p$ for all $p \in [1, \infty]$ are strongly equivalent. Assuming $p \leq q$, for any pair of persistence modules X, Y we have

$$d_{\mathcal{S}}^q(X, Y) \leq d_{\mathcal{S}}^p(X, Y) \leq 4^{\left(\frac{1}{p}-\frac{1}{q}\right)} d_{\mathcal{S}}^q(X, Y),$$

as can be easily concluded from the properties on p-norms on $\mathbb{R}^4$ stated in Sect. 2.6.

4.2 *p*-norms of persistence modules and contours

The aim of this section is to introduce and study a generalization of the notion of p-norm of a persistence module (see Sect. 2.6) first introduced in Skraba and Turner (2020), that coincides with the original definition if C is the standard contour (see Sect. 2.2).

Definition 4.7 Let C be a regular contour. For $p \in [1, \infty]$, define the (p, C)**-norm** of a persistence module $X \cong \bigoplus_{i=1}^{k} K(a_i, b_i)$ by

$$\|X\|_{p,C} := \begin{cases} \left(\sum_{i=1}^{k} \ell(a_i, b_i)^p\right)^{\frac{1}{p}} & \text{for } p \in [1, \infty), \\ \max\{\ell(a_i, b_i)\}_{i=1}^{k} & \text{for } p = \infty, \end{cases}$$

where $\ell(a_i, b_i)$ denotes the lifetime of the bar $K(a_i, b_i)$ with respect to C (see Sect. 2.2).

We see that $\|X\|_{p,C}$ does not depend on the choice of barcode decomposition for X. For $p \in [1, \infty]$ and $\varepsilon \in [0, \infty)$, consider the class of tame persistence modules

$$\mathcal{S}_{\varepsilon}^{p,C} := \{X \in \mathsf{Tame} \mid \|X\|_{p,C} \leq \varepsilon\},$$

and denote $\mathcal{S}^{p,C} := \{\mathcal{S}_{\varepsilon}^{p,C}\}_{\varepsilon \in [0,\infty)}$. If D is the standard contour (see Sect. 2.2), then $\ell(a_i, b_i) = b_i - a_i$ and we have $\|X\|_{p,D} = \|X\|_p$ and $\mathcal{S}^{p,D} = \mathcal{S}^p$. The main result in this subsection is showing that $\mathcal{S}^{p,C}$ is a noise system (see Sect. 2.3) whenever C is an action, for any $p \in [1, \infty]$. For the standard contour, this result together with Proposition 4.5 provide an alternative proof to the one in Skraba and Turner (2020) that the algebraic p-Wasserstein distance is a pseudometric, as will be later highlighted in Remark 4.20.

Given a contour C, the function $C(0, -) \colon [0, \infty) \to [0, \infty)$ is nondecreasing (and it is an increasing bijection if the contour is regular). Hence it can be viewed as a functor from the poset category $[0, \infty)$ to itself, that can be effectively used to re-parameterize $[0, \infty)$. For any persistence module X, the composition of functors $T_C(X) := XC(0, -) \colon [0, \infty) \to \mathrm{vect}_K$ is a persistence module. As we will show, $T_C(X)$ is in Tame whenever X is in Tame and C is a regular contour (Corollary 4.9). The assignment $X \mapsto T_C(X)$ can be extended to a functor $T_C \colon \mathsf{Tame} \to \mathsf{Tame}$ sending a morphism $f \colon X \to Y$ of persistence modules to the morphism $T_C(f) \colon T_C(X) \to T_C(Y)$ defined as the natural transformation between $T_C(X)$ and $T_C(Y)$ whose component at $a \in [0, \infty)$ is $T_C(f)_a = f_{C(0,a)} \colon X_{C(0,a)} \to Y_{C(0,a)}$.

We now explain the relationship between the barcode decompositions of X and $T_C(X)$ when C is a regular contour.

Proposition 4.8 *Let C be a regular contour, and let ℓ be the associated lifetime function. Consider a bar $K(a, b)$. Then*

$$T_C(K(a, b)) \cong K(\ell(0, a), \ell(0, b)).$$

Proof The functor $T_C(K(a, b)) \colon [0, \infty) \to \mathrm{vect}_K$ sends $c \leq d$ in $[0, \infty)$ to the linear function

$$K(a, b)_{C(0,c) \leq C(0,d)} \colon K(a, b)_{C(0,c)} \to K(a, b)_{C(0,d)},$$

which is the identity on K if $a \leq C(0,c) \leq C(0,d) < b$ and the zero function otherwise. Since C is regular, $\ell(0,-)$ is a strictly increasing function, hence the condition $a \leq C(0,c) \leq C(0,d) < b$ is equivalent to $\ell(0,a) \leq c \leq d < \ell(0,b)$. □

Corollary 4.9 *Let X be a tame persistence module with barcode decomposition $\bigoplus_{i=1}^{k} K(a_i,b_i)$, and let C be a regular contour. Then $T_C(X) \cong \bigoplus_{i=1}^{k} K(\ell(0,a_i), \ell(0,b_i))$. In particular $T_C(X)$ is also in* Tame.

Proof Since direct sums in Tame are defined pointwise (Sect. 2.1), if $\{X_i\}_{i\in I}$ is a finite collection of persistence modules and C is a regular contour, then $T_C(\bigoplus_{i\in I} X_i) \cong \bigoplus_{i\in I} T_C(X_i)$, this together with Proposition 4.8, gives the following.

$$\begin{aligned} T_C(X) &\cong T_C\left(\bigoplus_{i=1}^{k} K(a_i,b_i)\right) && \text{(by Proposition 4.8)} \\ &\cong \bigoplus_{i=1}^{k} T_C\left(K(a_i,b_i)\right) \\ &\cong \bigoplus_{i=1}^{k} K\left(\ell(0,a_i), \ell(0,b_i)\right). \end{aligned}$$

Given that bars $K(\ell(0,a_i), \ell(0,b_i))$ are in Tame and tameness is preserved by finite direct sums, $T_C(X)$ is also in Tame. □

We now show that the functor T_C is exact.

Proposition 4.10 *Let $0 \to X \to Y \to Z \to 0$ be an exact sequence in* Tame*, and let C be a regular contour. Then the sequence $0 \to T_C(X) \to T_C(Y) \to T_C(Z) \to 0$ is also exact in* Tame.

Proof Exactness in Tame is defined pointwise: $0 \to X \to Y \to Z \to 0$ is exact if and only if $0 \to X_a \to Y_a \to Z_a \to 0$ is exact in vect_K, for every $a \in [0,\infty)$. As a consequence, $0 \to X_{C(0,b)} \to Y_{C(0,b)} \to Z_{C(0,b)} \to 0$ is exact in vect_K, for every $b \in [0,\infty)$, hence by definition the sequence $0 \to T_C(X) \to T_C(Y) \to T_C(Z) \to 0$ is exact. The fact that the exact sequence $0 \to T_C(X) \to T_C(Y) \to T_C(Z) \to 0$ is in Tame follows from Corollary 4.9 and how T_C is defined on morphisms. □

Remark 4.11 As is clear from its proof, Proposition 4.10 in fact holds for the precomposition of persistence modules by any increasing bijection of $[0,\infty)$.

In the rest of the article, we will focus on contours that are regular and actions (see Sect. 2.2), called regular actions for brevity. We prove here a simple but important property of regular actions, and the associated lifetime function ℓ, which is used to prove the subsequent results.

Lemma 4.12 *If C is a regular action, then $\ell(a,c) = \ell(a,b) + \ell(b,c)$ for any $a \leq b \leq c$ in $[0,\infty)$.*

Proof Let $a \leq b \leq c$. Using the definitions and the assumption that C is an action, we have $C(C(a, \ell(a, b)), \ell(b, c)) = C(a, \ell(a, b) + \ell(b, c))$. Again by definition, we observe that the left-hand side equals c, and that $c = C(a, \ell(a, b) + \ell(b, c))$ implies $\ell(a, c) = \ell(a, b) + \ell(b, c)$. □

Proposition 4.13 *Let X be a persistence module, let $p \in [1, \infty]$, and let C be a regular action. Then $\|X\|_{p,C} = \|T_C(X)\|_p$.*

Proof Let $X \cong \bigoplus_{i=1}^{k} K(a_i, b_i)$. For any fixed $p \in [1, \infty)$, we have

$$\begin{aligned} \|T_C(X)\|_p &= \left(\sum_{i=1}^{k} (\ell(0, b_i) - \ell(0, a_i))^p \right)^{\frac{1}{p}} \\ &= \left(\sum_{i=1}^{k} \ell(a_i, b_i)^p \right)^{\frac{1}{p}} \\ &= \|X\|_{p,C} , \end{aligned}$$

where the first equality is by Corollary 4.9, the second one is by Lemma 4.12, and the third one is by definition of $\|\cdot\|_{p,C}$. The case $p = \infty$ is similar. □

We are now ready to prove that $\mathcal{S}^{p,C}$, with C a regular action, satisfies the axioms in the definition of noise system (see Sect. 2.3).

Lemma 4.14 *Let $0 \to X \to Y \to Z \to 0$ be an exact sequence in* Tame*, and let C be a regular contour. Then $\|X\|_{p,C} \leq \|Y\|_{p,C}$ and $\|Z\|_{p,C} \leq \|Y\|_{p,C}$.*

For the standard contour, our statement coincides with Lemma 7.8 in Skraba and Turner (2020), which is easily proven using the induced matchings (Bauer and Lesnick 2015) for monomorphisms and epimorphisms of persistence modules. For the sake of completeness, we include the proof for $\|\cdot\|_{p,C}$, which does not present any additional difficulty.

Proof The existence of a monomorphism from X to Y implies the existence of the bar-to-bar monomorphism $f_\varphi \colon X \hookrightarrow Y$ of Definition 3.17, induced by the canonical injection (Bauer and Lesnick 2015). The monomorphism f_φ decomposes as a finite direct sum of monomorphisms of the form $K(a', b) \hookrightarrow K(a, b)$, with $a \leq a'$, and of the form $0 \hookrightarrow K(a, b)$. By monotonicity of contours, $a \leq a'$ implies $\ell(a', b) \leq \ell(a, b)$. The inequality $\|X\|_{p,C} \leq \|Y\|_{p,C}$ then follows from the definition of $\|\cdot\|_{p,C}$, since every term in the expression for $\|X\|_{p,C}$ is upper bounded by a term in the expression for $\|Y\|_{p,C}$.

The proof of the inequality $\|Z\|_{p,C} \leq \|Y\|_{p,C}$ is obtained similarly, using the epimorphism induced by the canonical injection (see Definition 3.21). □

Lemma 4.15 *Let $0 \to X \to Y \to Z \to 0$ be an exact sequence in* Tame*, and let C be a regular action. Then $\|Y\|_{p,C} \leq \|X\|_{p,C} + \|Z\|_{p,C}$.*

Proof First, we prove the statement assuming that C is the standard contour. Let $0 \to X \xrightarrow{f} Y \xrightarrow{g} Z \to 0$ be a short exact sequence of persistence modules, and let us show that $\|Y\|_p \leq \|X\|_p + \|Z\|_p$. We consider the monomorphism f and observe that $Z \cong \operatorname{coker} f$ implies that Z and $\operatorname{coker} f$ have the same barcode decomposition, hence $\|Z\|_p = \|\operatorname{coker} f\|_p$. Theorem 3.14 tells us that, among all monomorphisms between two fixed persistence modules, the norm $\|\cdot\|_p$ of the cokernel is minimized by a bar-to-bar monomorphism. We therefore just need to prove that $\|Y\|_p \leq \|X\|_p + \|\operatorname{coker} f\|_p$, for any bar-to-bar monomorphism f between X and Y.

By Remark 3.2, if $f\colon X \to Y$ is a bar-to-bar monomorphism, then there exist barcode decompositions $\bigoplus_{i=1}^m X_i$ and $\bigoplus_{j=1}^n Y_j$ of X and Y, respectively, such that $m \leq n$ and, up to permutation of the Y_j, there are monomorphisms $f_i\colon X_i \to Y_i$ between bars such that $\operatorname{coker} f = \bigoplus_{i=1}^m \operatorname{coker} f_i \oplus \bigoplus_{j=m+1}^n Y_j$. We observe that, for each bar $Y_i = K(a_i, b_i)$ of Y with $i \in \{1, \ldots, m\}$, there is a bar $X_i = K(a_i', b_i)$ of X and a corresponding summand $\operatorname{coker} f_i$ of $\operatorname{coker} f$, which is a bar $K(a_i, a_i')$ if $a_i < a_i'$, and it is the zero module if $a_i = a_i'$. Similarly, we observe that each bar $Y_j = K(a_j, b_j)$ of Y with $j \in \{m+1, \ldots, n\}$ is also a bar of $\operatorname{coker} f$. By definition, $\|Y\|_p$ is the p-norm of the following element of $\mathbb{R}^n$:

$$(b_j - a_j)_{j\in\{1,\ldots,n\}} = (((b_i - a_i') + (a_i' - a_i))_{i\in\{1,\ldots,m\}}, (b_j - a_j)_{j\in\{m+1,\ldots,n\}}).$$

Then, by the triangular inequality of p-norms in $\mathbb{R}^n$, we have $\|Y\|_p \leq \|X\|_p + \|\operatorname{coker} f\|_p$, which completes the proof when C is the standard contour.

Let now C be any regular action. By Proposition 4.10, exactness of $0 \to X \to Y \to Z \to 0$ implies exactness of $0 \to T_C(X) \to T_C(Y) \to T_C(Z) \to 0$. Applying the previous part of the proof to the latter exact sequence yields $\|T_C(Y)\|_p \leq \|T_C(X)\|_p + \|T_C(Z)\|_p$, which by Proposition 4.13 coincides with our claim. □

For the standard contour, the statement of Lemma 4.15 is given in Remark 7.32 of Skraba and Turner (2020). However, to our knowledge, we provide the first proof of this inequality that does not assume the fact that the p-norm of persistence modules induces a pseudometric. Indeed in Skraba and Turner (2020) the fact that the algebraic Wasserstein distance satisfies the triangular inequality is used as an hypothesis.

We can now prove the main result of this subsection.

Theorem 4.16 *For any $p \in [1, \infty]$ and any regular action C, $\mathcal{S}^{p,C}$ is a noise system.*

Proof We show that $\mathcal{S}^{p,C}$ satisfies all axioms of the definition of noise system (see Sect. 2.3). Since the norm $\|\cdot\|_{p,C}$ of the zero module 0 is zero, we have $0 \in \mathcal{S}_\varepsilon^{p,C}$, for all $\varepsilon \in [0, \infty)$. By definition of $\mathcal{S}^{p,C}$, it is clear that $\mathcal{S}_\tau^{p,C} \subseteq \mathcal{S}_\varepsilon^{p,C}$ whenever $\tau \leq \varepsilon$. Lemmas 4.14 and 4.15 complete the proof, showing that $\mathcal{S}^{p,C}$ satisfies both conditions on short exact sequences of persistence modules. □

Remark 4.17 For $p < \infty$, the noise system $\mathcal{S}^{p,C}$ in not closed under direct sums (Sect. 2.3), since $\|X \oplus Y\|_{p,C} = \left\|\left(\|X\|_{p,C}, \|Y\|_{p,C}\right)\right\|_p$ by Eq. (2).

Remark 4.18 Let us briefly highlight the role of our hypotheses on contours, which are required to be regular actions in Theorem 4.16. The regularity assumption ensures

for instance that the associated lifetime function ℓ is well-defined, and that the functor T_C is an endofunctor on Tame. The assumption that $C(0,-)\colon [0,\infty) \to [0,\infty)$ is an increasing bijection is sufficient to prove many results of this subsection (see Remark 4.11), but we choose to assume the stronger condition of regularity on C to facilitate a comparison with the results of Chachólski and Riihimäki (2020), observing in addition that many examples of regular contours can be found, for example the contours of distance type (Sect. 2.2) that are used in our experiments (see Sect. 5). The hypothesis that the considered contours are actions is motivated by the use of Proposition 4.13 in the proof of Lemma 4.15. If C is not an action, the equalities of Lemma 4.12 and Proposition 4.13 are replaced by the inequalities $\ell(a,c) \leq \ell(a,b) + \ell(b,c)$, for any $a \leq b \leq c$, and $\|T_C(X)\|_p \leq \|X\|_{p,C}$. A proof of Lemma 4.15 removing the action hypothesis on C eludes us.

4.3 Contours and algebraic Wasserstein distances

We now turn to considering the pseudometrics $d^q_{\mathcal{S}^{p,C}}$ associated (as in Sect. 4.1) with the noise systems $\mathcal{S}^{p,C}$ introduced in Sect. 4.2, for fixed $p, q \in [1,\infty]$ and a regular action C. We also refer to these pseudometrics as **algebraic Wasserstein distances**. First, we show that the functor T_C introduced in Sect. 4.2 allows us to switch between a pseudometric $d^q_{\mathcal{S}^{p,C}}$ and the pseudometric $d^q_{\mathcal{S}^p}$ associated with the standard contour. More precisely, we show that T_C can be viewed as an isometry

$$T_C\colon \left(\mathsf{Tame}, d^q_{\mathcal{S}^{p,C}}\right) \to \left(\mathsf{Tame}, d^q_{\mathcal{S}^p}\right).$$

Let us recall that, if C is a regular contour, the function $C(0,-) : [0,\infty) \to [0,\infty)$ is an increasing bijection. Its inverse $\ell(0,-) := C(0,-)^{-1}$ is therefore an increasing bijection as well. Mimicking the definition of T_C given in Sect. 4.2, we can define a functor $T_\ell\colon \mathsf{Tame} \to \mathsf{Tame}$ given by precomposition by the increasing function $\ell(0,-)$. By Proposition 4.10, the functor $T_C\colon \mathsf{Tame} \to \mathsf{Tame}$ preserves kernels and cokernels, and T_ℓ has the same property by Remark 4.11. Furthermore, since $C(0,-)$ and $\ell(0,-)$ are inverse to each other, the compositions $T_C T_\ell$ and $T_\ell T_C$ are the identity functor 1_{Tame} on Tame.

To prove the following result, it is convenient to define the (p,q,C)**-cost** of a span $X \xleftarrow{f} Z \xrightarrow{g} Y$ of persistence modules as the element $c \in [0,\infty]$ defined by

$$c := \left\| \left(\|\ker f\|_{p,C}, \|\mathrm{coker}\, f\|_{p,C}, \|\ker g\|_{p,C}, \|\mathrm{coker}\, g\|_{p,C} \right) \right\|_q .$$

Proposition 4.19 *Let C be a regular action, and let X, Y be persistence modules. Then*

$$d^q_{\mathcal{S}^{p,C}}(X,Y) = d^q_{\mathcal{S}^p}(T_C(X), T_C(Y)).$$

Proof Let D denote the standard contour, and let us recall that the (p,D)-norm of a persistence module coincides with its p-norm (Sect. 4.2). We describe a correspondence between spans having the same cost, calculated with respect to (p,q,C) and (p,q,D) respectively.

Let $X \xleftarrow{f} Z \xrightarrow{g} Y$ be a span and let c be its (p, q, C)-cost. Applying the functor T_C, we obtain the span $T_C(X) \xleftarrow{T_C(f)} T_C(Z) \xrightarrow{T_C(g)} T_C(Y)$, whose (p, q, D)-cost is

$$\begin{aligned} c' &= \left\| \left(\|\ker T_C(f)\|_p , \|\operatorname{coker} T_C(f)\|_p , \|\ker T_C(g)\|_p , \|\operatorname{coker} T_C(g)\|_p \right) \right\|_q \\ &= \left\| \left(\|T_C(\ker f)\|_p , \|T_C(\operatorname{coker} f)\|_p , \|T_C(\ker g)\|_p , \|T_C(\operatorname{coker} g)\|_p \right) \right\|_q \\ &= c, \end{aligned}$$

where the second equality holds because the functor T_C preserves kernels and cokernels, and the last equality holds by Proposition 4.13.

To prove the other direction of the correspondence, we start from a span $T_C(X) \xleftarrow{\varphi} T_C(Z) \xrightarrow{\psi} T_C(Y)$ whose (p, q, D)-cost is

$$k := \left\| \left(\|\ker \varphi\|_p , \|\operatorname{coker} \varphi\|_p , \|\ker \psi\|_p , \|\operatorname{coker} \psi\|_p \right) \right\|_q ,$$

and we exhibit a span between X and Y whose (p, q, C)-cost equals k. Applying the functor T_ℓ, we obtain the span $X \xleftarrow{T_\ell(\varphi)} Z \xrightarrow{T_\ell(\psi)} Y$. To determine the (p, q, C)-cost of this span we observe that

$$\|\ker T_\ell(\varphi)\|_{p,C} = \|T_\ell(\ker \varphi)\|_{p,C} = \|T_C T_\ell(\ker \varphi)\|_p = \|\ker \varphi\|_p ,$$

where the first equality holds because T_ℓ preserves kernels, the second equality is by Proposition 4.13, and the third equality holds because $T_C T_\ell = 1_{\mathsf{Tame}}$. Since similar equalities hold for $\operatorname{coker} T_\ell(\varphi)$, $\ker T_\ell(\psi)$, and $\operatorname{coker} T_\ell(\psi)$, the (p, q, C)-cost of the span $X \xleftarrow{T_\ell(\varphi)} Z \xrightarrow{T_\ell(\psi)} Y$ equals k. □

Remark 4.20 Some of the pseudometrics between persistence modules that have been studied by other authors fall within the framework we have presented in this subsection and in Sect. 4.1. If C is a regular contour, the pseudometric denoted by d_C in (Chachólski and Riihimäki 2020, Sect. 6) coincide with our pseudometrics of the type $d^1_{\mathcal{S}^{\infty},C}$. In particular, for the standard contour (Sect. 2.2) the pseudometric $d^1_{\mathcal{S}^\infty}$ coincides with the standard pseudometric already introduced in Scolamiero et al. (2017). As we already mentioned, the algebraic pseudometrics introduced in Skraba and Turner (2020, Sect. 7) are of the form $d^p_{\mathcal{S}^p}$, thus coinciding with our pseudometrics with the choice $p = q$ and for the standard contour. In Giunti et al. (2024), the authors propose a framework to study distances on abelian categories which is equivalent to noise systems on abelian categories. The authors of Bubenik et al. (2023) also study distances on abelian categories, introducing the notion of exact weight, which is more general than noise systems as the first axiom on short exact sequences is relaxed. The so-called path metric associated with an exact weight is defined for zigzags of morphisms of arbitrary finite length, but for the particular case of path metrics on noise systems considering spans is sufficient. In this case, the path metric coincides with a pseudometric of the form $d^1_{\mathcal{S}}$. In particular, the path metric $d_{\mu\circ\dim}$ between persistence modules studied

in (Bubenik et al. 2023, Sect. 4) coincides with $d^1_{\mathcal{S}^1}$ in our notations, while the p-Wasserstein distances introduced by the authors are different from our pseudometrics $d^q_{\mathcal{S}^{p,C}}$.

4.4 Algebraic and combinatorial (p, C)-Wasserstein distances

In this subsection we consider Wasserstein distances between persistence diagrams. Here, we call these pseudometrics *combinatorial* Wasserstein distances, to distinguish them from the *algebraic* pseudometrics $d^q_{\mathcal{S}^{p,C}}$ defined on the class of persistence modules. We introduce a new family of combinatorial Wasserstein distances, parametrized by $p, q \in [1, \infty]$ and a regular action C, which generalize the Wasserstein distances commonly used in persistence theory. Finally, we prove isometry results involving the combinatorial Wasserstein distances and the algebraic Wasserstein distances $d^q_{\mathcal{S}^{p,C}}$ introduced in Sect. 4.2.

Let $U := \{(a, b) \in [0, \infty) \times [0, \infty] \mid a \leq b\}$ be a subset of the extended plane. A **persistence diagram** is a finite multiset $D = \{x_i\}_{i \in S}$ of elements of U. Since D is a multiset, it may happen that $x_i = x_k$ for some $i \neq k$. The **diagonal** Δ of $[0, \infty)$ is the set $\Delta := \{(a, a) \mid a \in [0, \infty)\} \subset U$. For all $p \in [1, \infty]$, we denote by d_p the metric on U induced by the p-norm, defined by $d_p(x, y) := \|x - y\|_p$ for all $x, y \in U$, and we denote $d_p(x, \Delta) := \inf_{z \in \Delta} d_p(x, z)$. As is easy to show, if $x = (a, b)$, then $d_p(x, \Delta) = d_p(x, \overline{z})$ with $\overline{z} := (\frac{a+b}{2}, \frac{a+b}{2})$.

Let $D = \{x_i\}_{i \in \{1,\dots,m\}}$ and $D' = \{x'_j\}_{j \in \{1,\dots,n\}}$ be persistence diagrams. For any $p, q \in [1, \infty]$, the (p, q)-**Wasserstein distance** between D and D' is defined by

$$W^q_p(D, D') := \inf_{\alpha} \left\| \left(\left\| (d_p(x_i, x'_{\alpha(i)}))_{i \in I} \right\|_q, \left\| (d_p(x_i, \Delta))_{i \in \{1,\dots,m\} \setminus I} \right\|_q, \left\| (d_p(\Delta, x'_j))_{j \in \{1,\dots,n\} \setminus \alpha(I)} \right\|_q \right) \right\|_q,$$

where the infimum is over all injective functions $\alpha \colon I \to \{1, \dots, n\}$, with $I \subseteq \{1, \dots, m\}$.

Remark 4.21 We note that in the literature, the letters p and q are sometimes interchanged with respect to our notation of the parameters of Wasserstein distances between persistence diagrams. This is the case for instance in (Skraba and Turner 2020, Def. 2.7). Our choice of notation is motivated by symmetry with the definition of algebraic Wasserstein distances, where a norm $\|\cdot\|_q$ is used to "aggregate" costs expressed with respect to a norm $\|\cdot\|_p$.

Let $\mathcal{D}$ denote the set of all persistence diagrams. We define the class function $\mathrm{Dgm} \colon \mathsf{Tame} \to \mathcal{D}$ sending any persistence module X to the persistence diagram $\mathrm{Dgm}(X)$ such that $X \cong \bigoplus_{(a,b) \in \mathrm{Dgm}(X)} K(a, b)$, where we note that in the right-hand term each bar $K(a, b)$ appears the same number of times as the multiplicity of (a, b) in the multiset $\mathrm{Dgm}(X)$. By virtue of the barcode decomposition theorem (Theorem 2.2), the function $\mathrm{Dgm} \colon \mathsf{Tame} \to \mathcal{D}$ induces a bijection between the set $\mathsf{Tame}/\sim$ of isomorphism classes of persistence modules and $\mathcal{D}$.

As proven in Skraba and Turner (2020), if $p = q$ then the algebraic distance $d^q_{\mathcal{S}^p}$ between persistence modules coincides with the combinatorial distance W^q_p between the associated persistence diagrams.

Theorem 4.22 (Skraba and Turner 2020) *For any $p \in [1, \infty]$ and for any persistence modules X and Y we have*

$$d^p_{\mathcal{S}^p}(X, Y) = W^p_p(\mathrm{Dgm}(X), \mathrm{Dgm}(Y)).$$

It is worth observing that the equality of Theorem 4.22 does not hold when $p \neq q$. For example, we can consider the persistence modules

$$X = K(a_1, a_1 + \ell_1) \oplus K(a_2, a_2 + \ell_2) \oplus K(a_3, a_3 + \ell_3)$$

with ℓ_1, ℓ_2, ℓ_3 positive real numbers, and 0, the zero module. Then, assuming $q < \infty$,

$$d^q_{\mathcal{S}^p}(X, 0) = \left(\left\| \left(\frac{\ell_1}{2}, \frac{\ell_2}{2}, \frac{\ell_3}{2} \right) \right\|^q_p + \left\| \left(\frac{\ell_1}{2}, \frac{\ell_2}{2}, \frac{\ell_3}{2} \right) \right\|^q_p \right)^{\frac{1}{q}}$$

(as we will prove in Lemma 4.23), while

$$W^q_p(\mathrm{Dgm}(X), \mathrm{Dgm}(0)) = \left(\left\| \left(\frac{\ell_1}{2}, \frac{\ell_1}{2} \right) \right\|^q_p + \left\| \left(\frac{\ell_2}{2}, \frac{\ell_2}{2} \right) \right\|^q_p + \left\| \left(\frac{\ell_3}{2}, \frac{\ell_3}{2} \right) \right\|^q_p \right)^{\frac{1}{q}}.$$

Given a regular contour C, we now define a function $\tau_C : U \to U$ as follows: for $x = (a, b) \in U$, we set $\tau_C(x) = (\ell(0, a), \ell(0, b))$, where $\ell(0, -)$ is the lifetime function associated with C (Sect. 2.2). If D is a persistence diagram, then by applying τ_C to each element of D we obtain a persistence diagram that we denote by $\tau_C(D)$. Hence, we have a function $\mathcal{D} \to \mathcal{D}$ which we denote again by τ_C, with a slight abuse of notation. If C is the standard contour, then τ_C is the identity function and in particular $\tau_C(D) = D$. Figure 1 illustrates a persistence diagram transformed by applying τ_C for a contour C of distance type.

Given a regular contour C, we define the **combinatorial (p, C)-Wasserstein distance** $W^p_{p,C}$ pulling back the pseudometric W^p_p via $\tau_C : \mathcal{D} \to \mathcal{D}$. Explicitly, for all persistence diagrams D and D', we define $W^p_{p,C}(D, D') := W^p_p(\tau_C(D), \tau_C(D'))$. If C is a regular action, then as a consequence of Corollary 4.9 we have $\mathrm{Dgm}(T_C(X)) = \tau_C(\mathrm{Dgm}(X))$, for every persistence module X. This implies, by virtue of Proposition 4.19 and Theorem 4.22, that

$$d^p_{\mathcal{S}^{p,C}}(X, Y) = W^p_{p,C}(\mathrm{Dgm}(X), \mathrm{Dgm}(Y)),$$

for all persistence modules X and Y.

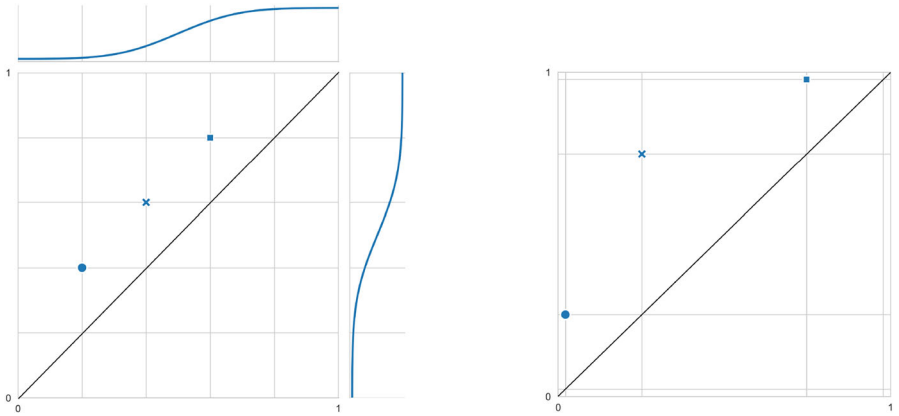


Fig. 1 **Left:** A persistence diagram $D = \{(0.2, 0.4), (0.4, 0.6), (0.6, 0.8)\}$. A contour C of distance type parametrized by a Gaussian density ($\mu = 0.5, \sigma = 0.15$) is chosen and the corresponding function $f(x) = \ell(0, x)$ (i.e. the Gaussian cumulative distribution function) is shown above and to the right of the persistence diagram. **Right:** The transformed persistence diagram $\tau_C(D) = \{(\ell(0, 0.2), \ell(0, 0.4)), (\ell(0, 0.4), \ell(0, 0.6)), (\ell(0, 0.6), \ell(0, 0.8))\}$. The regular grid from the left diagram has also been transformed to illustrate how τ_C stretches the plane

To summarize, for any $p \in [1, \infty]$ and any regular action C, we have a commutative diagram of isometries

$$\begin{array}{ccc} (\mathsf{Tame}, d^p_{\mathcal{S}^{p,C}}) & \xrightarrow{\mathrm{Dgm}} & (\mathcal{D}, W^p_{p,C}) \\ {\scriptstyle T_C}\downarrow & & \downarrow{\scriptstyle \tau_C} \\ (\mathsf{Tame}, d^p_{\mathcal{S}^p}) & \xrightarrow{\mathrm{Dgm}} & (\mathcal{D}, W^p_p) \end{array}$$

4.5 Algebraic parametrized Wasserstein distances

The equivalence between algebraic and combinatorial Wasserstein distances for the case $p = q$, described in Sect. 4.4 or in Skraba and Turner (2020) for the standard contour, implies that in general Wasserstein distances cannot be expressed by an explicit (i.e., not involving an optimization problem) formula depending on the barcode decompositions of the persistence modules we are comparing. However for some special classes of persistence modules this is the case. The focus of this section is to present such formulas for the exact computation of algebraic Wasserstein distances. To avoid distinguishing the cases $q < \infty$ and $q = \infty$ in stating the results of this subsection, for $q = \infty$ we set by convention $\frac{1}{q} = 0$ and $2^{\frac{1-q}{q}} = 2^{-1}$.

Lemma 4.23 *For all persistence modules X and all $p, q \in [1, \infty]$ we have*

$$d^q_{\mathcal{S}^p}(X, 0) = 2^{\frac{1-q}{q}} \|X\|_p .$$

Proof Let $X = \bigoplus_{i=1}^{k} K(a_i, b_i)$ be a barcode decomposition of X, consider a persistence module of the form $Z = \bigoplus_{i=1}^{k} K(\frac{a_i+b_i}{2}, b_i)$ and a bar-to-bar morphism $f = \bigoplus_{i=1}^{k} f_i \colon Z \to X$, with each $f_i \colon K(\frac{a_i+b_i}{2}, b_i) \to K(a_i, b_i)$ a monomorphism between bars. Denote by $g = \bigoplus_{i=1}^{k} g_i \colon Z \to 0$ the zero map. The existence of the span $X \xleftarrow{f} Z \xrightarrow{g} 0$ implies that X and 0 are $2^{\frac{1-q}{q}} \|X\|_p$ close in q-norm (Definition 4.3), proving that $d^q_{\mathcal{S}^p}(X, 0) \leq 2^{\frac{1-q}{q}} \|X\|_p$. Indeed $\ker f = \operatorname{coker} g = 0$ and $\|\operatorname{coker} f\|_p = \|\ker g\|_p = \left\| (\frac{b_i-a_i}{2})_{i\in\{1,\dots,k\}} \right\|_p$. The bound is obtained by computing $\left\| \left(\left\| (\frac{b_i-a_i}{2})_{i\in\{1,\dots,k\}} \right\|_p, \left\| (\frac{b_i-a_i}{2})_{i\in\{1,\dots,k\}} \right\|_p \right) \right\|_q = 2^{\frac{1-q}{q}} \|X\|_p$.

To prove the converse inequality, let us show that if $d^q_{\mathcal{S}^p}(X, 0) < \varepsilon$ then $2^{\frac{1-q}{q}} \|X\|_p < \varepsilon$. If $d^q_{\mathcal{S}^p}(X, 0) < \varepsilon$, then there exists a $(\varepsilon_1, \varepsilon_2, \varepsilon_3, 0)$-span $X \xleftarrow{\varphi} Z \twoheadrightarrow 0$ for some $\varepsilon_1, \varepsilon_2, \varepsilon_3$ in $[0, \infty)$ such that $\|(\varepsilon_1, \varepsilon_2, \varepsilon_3)\|_q < \varepsilon$. Note that $X \hookleftarrow \operatorname{im}\varphi \twoheadrightarrow 0$ is then a $(0, \varepsilon_2, \varepsilon_3, 0)$-span. Consider the short exact sequence $\operatorname{im}\varphi \hookrightarrow X \twoheadrightarrow \operatorname{coker}\varphi$. Since $\operatorname{coker}\varphi \in \mathcal{S}^p_{\varepsilon_2}$ and $\operatorname{im}\varphi \in \mathcal{S}^p_{\varepsilon_3}$, by the third axiom of noise systems we get $X \in \mathcal{S}^p_{\varepsilon_2+\varepsilon_3}$, and so we get $\|X\|_p \leq \varepsilon_2 + \varepsilon_3$ by definition of $\mathcal{S}^p$. Furthermore, by inequalities (1) between p-norms on $\mathbb{R}^2$, $\varepsilon_2 + \varepsilon_3 = \|(\varepsilon_2, \varepsilon_3)\|_1 \leq 2^{1-\frac{1}{q}} \|(\varepsilon_2, \varepsilon_3)\|_q < 2^{1-\frac{1}{q}} \varepsilon$. Therefore we have $\|X\|_p < 2^{1-\frac{1}{q}} \varepsilon$ or equivalently $2^{\frac{1-q}{q}} \|X\|_p < \varepsilon$. We conclude that $d^q_{\mathcal{S}^p}(X, 0) \geq 2^{\frac{1-q}{q}} \|X\|_p$, and therefore $d^q_{\mathcal{S}^p}(X, 0) = 2^{\frac{1-q}{q}} \|X\|_p$. □

Remark 4.24 The formula $d^q_{\mathcal{S}^p}(X, 0) = 2^{\frac{1-q}{q}} \|X\|_p$ of Lemma 4.23 was already shown for the case $p = q$ in Skraba and Turner (2020) by using the correspondence between combinatorial and algebraic Wasserstein distances.

The proof of Lemma 4.23 can be easily extended to the case of a regular action C. In this case, we have

$$d^q_{\mathcal{S}^p,C}(X, 0) = d^q_{\mathcal{S}^p}(T_C(X), 0) = 2^{\frac{1-q}{q}} \|T_C(X)\|_p = 2^{\frac{1-q}{q}} \|X\|_{p,C}\,, \tag{6}$$

where the first equality holds by Proposition 4.19, the second by Lemma 4.23 and the third by Proposition 4.13. Similar arguments can be applied to all the results of this subsection. For exposition purposes we consider the case of the standard contour throughout the section and collect generalizations of the main results at the end of the subsection in Proposition 4.32.

Proposition 4.25 *Let X, Y, V be persistence modules. Then, for every $p, q \in [1, \infty]$,*

$$d^q_{\mathcal{S}^p}(X \oplus V, Y \oplus V) \leq d^q_{\mathcal{S}^p}(X, Y).$$

Proof It suffices to observe that for any span $X \xleftarrow{f} Z \xrightarrow{g} Y$, the span $X \oplus V \xleftarrow{f\oplus 1} Z \oplus V \xrightarrow{g\oplus 1} Y \oplus V$ has the same cost. □

Remark 4.26 Note that by considering $Y = 0$, Proposition 4.25 gives

$$d_{\mathcal{S}^p}^q(X \oplus V, V) \leq d_{\mathcal{S}^p}^q(X, 0) = 2^{\frac{1-q}{q}} \|X\|_p$$

The converse inequality $d_{\mathcal{S}^p}^q(X \oplus V, V) \geq d_{\mathcal{S}^p}^q(X, 0) = 2^{\frac{1-q}{q}} \|X\|_p$ does not hold in general, as illustrated in the following example. Consider $p = q = 2$, $X = K(0, 6)$ and $V = K(1, 5) \oplus K(2, 4)$. By Lemma 4.23 we have that $d_{\mathcal{S}^p}^q(X, 0) = \frac{1}{\sqrt{2}} \cdot 6 = \sqrt{18}$. However, $X \oplus V$ and $Y \oplus V$ are $\sqrt{6}$-close via the following $(0, \sqrt{3}, \sqrt{3}, 0)$-span

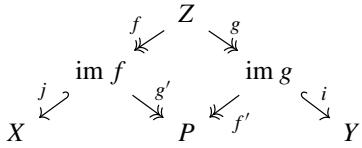
$$\begin{aligned} K(0,6) \oplus K(1,5) \oplus K(2,4) &\xleftarrow{f_1 \oplus f_2 \oplus f_3} K(1,6) \oplus K(2,5) \\ \oplus\, K(3,4) &\xrightarrow{g_1 \oplus g_2 \oplus g_3} K(1,5) \oplus K(2,4) \oplus 0 \end{aligned}$$

implying that $d_{\mathcal{S}^p}^q(X \oplus V, Y \oplus V) \leq \sqrt{6} < \sqrt{18} = d_{\mathcal{S}^p}^q(X, Y)$. This example is based on the fact that given a span $X \xleftarrow{f} Z \xrightarrow{g} Y$ realizing the distance between X and Y, the span $X \oplus V \xleftarrow{f \oplus 1} Z \oplus V \xrightarrow{g \oplus 1} Y \oplus V$ not always is the one achieving the distance between $X \oplus V$ and $Y \oplus V$.

Let $\{K(a_i, b_i)\}_{i \in \{1,\ldots,k\}}$ be a sequence of bars ordered non-decreasingly by length, that is, $b_1 - a_1 \leq b_2 - a_2 \leq \cdots \leq b_k - a_k$. For $j \in \{1, \ldots, k\}$, consider $Z := \bigoplus_{i=1}^{j} K(a_i, b_i)$ and $Y := \bigoplus_{i=j+1}^{k} K(a_i, b_i)$. The remainder of this section is devoted to proving that, in this case, $d_{\mathcal{S}^p}^q(Y \oplus Z, Y) = d_{\mathcal{S}^p}^q(Z, 0) = 2^{\frac{1-q}{q}} \|Z\|_p$. In Sect. 5, this result will be used for the computation of the stable rank of a persistence module with respect to $d_{\mathcal{S}^p}^q$.

Proposition 4.27 *Let $\mathcal{S}$ be a noise system. For any $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$-span $X \leftarrow Z \rightarrow Y$ of persistence modules there is a mono-epi $(0, \varepsilon_2, \varepsilon_3, 0)$-span $X \hookleftarrow \operatorname{im} f \twoheadrightarrow P$ such that* $\operatorname{rank}(P) \leq \operatorname{rank}(Y)$.

Proof By Theorem 3.14 and Remark 3.2, if $U \hookrightarrow V$ is a monomorphism between persistence modules, then $\operatorname{rank}(U) \leq \operatorname{rank}(V)$, and similarly if $V \twoheadrightarrow U$ is an epimorphism, then $\operatorname{rank}(U) \leq \operatorname{rank}(V)$. Let $X \xleftarrow{f} Z \xrightarrow{g} Y$ be a $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$-span of persistence modules, and consider the following diagram in Tame, where the square is a push-out:



Since f' is an epimorphism and i is a monomorphism, $\operatorname{rank}(P) \leq \operatorname{rank}(\operatorname{im} g) \leq \operatorname{rank}(Y)$. We consider the span $X \xleftarrow{j} \operatorname{im} f \xrightarrow{g'} P$. Clearly, the kernel of the corestriction $g\colon Z \twoheadrightarrow \operatorname{im} g$ still belongs to $\mathcal{S}_{\varepsilon_3}$, and its cokernel is zero. Then, by Proposition 8.1

in Scolamiero et al. (2017), $\ker g' \in \mathcal{S}_{\varepsilon_3}$ and $\operatorname{coker} g' = 0$. The kernel of j is 0, while its cokernel belongs to $\mathcal{S}_{\varepsilon_2}$, as it coincides with the cokernel of $f\colon Z \to X$. □

Lemma 4.28 *Let $p, q \in [1, \infty]$, and let $[a_i, b_i]$ be nonempty intervals in $[0, \infty)$, for $i \in \{1, \dots, j\}$. The function $\gamma\colon \prod_{i=1}^{j} [a_i, b_i] \to [0, \infty)$ defined by*

$$\gamma(x_1, \dots, x_j) := \left\| \left(\left\| (x_1 - a_1, \dots, x_j - a_j) \right\|_p, \left\| (b_1 - x_1, \dots, b_j - x_j) \right\|_p \right) \right\|_q$$

has a global minimum at $(\frac{a_1+b_1}{2}, \dots, \frac{a_j+b_j}{2})$.

Proof The function γ is continuous with a compact domain, so it admits a global minimum by the extreme value theorem. Moreover, it is convex because norms are convex functions.

Write $a = (a_1, \dots, a_j)$, $b = (b_1, \dots, b_j)$ and $x = (x_1, \dots, x_j)$ in $\mathbb{R}^j$. Since $\gamma(x) = \gamma(a + b - x)$ for every x, the function γ is invariant under point reflection through $\frac{a+b}{2}$. By convexity, we conclude that $\frac{a+b}{2}$ is a global minimum of γ. □

Proposition 4.29 *Let $X = \bigoplus_{i=1}^{k} K(a_i, b_i)$, with the bars ordered non-decreasingly by length. Let $j \in \{1, \dots, k\}$, and let $p, q \in [1, \infty]$. Then, any persistence module Y with $\operatorname{rank}(Y) \leq \operatorname{rank}(X) - j$ is such that*

$$d_{\mathcal{S}^p}^q(X, Y) \geq 2^{\frac{1-q}{q}} \left\| \bigoplus_{i=1}^{j} K(a_i, b_i) \right\|_p.$$

Proof We prove the claim by contradiction. Suppose that there exists a persistence module Y such that $\operatorname{rank}(Y) \leq \operatorname{rank}(X) - j$ and

$$d_{\mathcal{S}^p}^q(X, Y) < 2^{\frac{1-q}{q}} \left\| \bigoplus_{i=1}^{j} K(a_i, b_i) \right\|_p.$$

By definition, there exists a span $X \xleftarrow{f} Z \xrightarrow{g} Y$ such that

$$\left\| \left(\|\ker f\|_p, \|\operatorname{coker} f\|_p, \|\ker g\|_p, \|\operatorname{coker} g\|_p \right) \right\|_q < 2^{\frac{1-q}{q}} \left\| \bigoplus_{i=1}^{j} K(a_i, b_i) \right\|_p. \tag{7}$$

By Proposition 4.27 we can assume (possibly after replacing Y with a persistence module of smaller or equal rank) that the span above is mono-epi, that is, of the form $X \xleftarrow{f} Z \xrightarrow{g} Y$. By Theorems 3.14 and 3.16, we can moreover assume that f and g are bar-to-bar morphisms.

Thus, we can consider a barcode decomposition $Z = \bigoplus_{i=1}^{k} Z_i$, with some of the Z_i possibly zero, and a barcode decomposition $Y = \bigoplus_{i=1}^{k} Y_i$, with at least j of the Y_i equal to zero by assumption, together with morphisms between bars $K(a_i, b_i) \xleftarrow{f_i} Z_i \xrightarrow{g_i} Y_i$ such that $f = \bigoplus_{i=1}^{k} f_i$ and $g = \bigoplus_{i=1}^{k} g_i$. Let $I \subseteq \{1, \dots, k\}$, with $|I| \geq j$, be the subset of the indices i such that $Y_i = 0$. For every $i \in I$, we have

$K(a_i, b_i) \xleftarrow{f_i} Z_i \xrightarrow{g_i} 0$, with $Z_i = K(x_i, b_i)$ for some $a_i \leq x_i \leq b_i$, where $K(b_i, b_i)$ denotes the zero module. Since $\ker f = \bigoplus_{i=1}^k \ker f_i$ and $\operatorname{coker} f = \bigoplus_{i=1}^k \operatorname{coker} f_i$, by Remark 3.2 we observe that $\bigoplus_{i\in I} K(a_i, x_i)$ is a direct summand of $\operatorname{coker} f$, and similarly that $\bigoplus_{i\in I} K(x_i, b_i)$ is a direct summand of $\ker g$, which gives

$$\begin{aligned} \|\operatorname{coker} f\|_p &\geq \|\oplus_{i\in I} K(a_i, x_i)\|_p = \|(x_i - a_i)_{i\in I}\|_p, \\ \|\ker g\|_p &\geq \|\oplus_{i\in I} K(x_i, b_i)\|_p = \|(b_i - x_i)_{i\in I}\|_p. \end{aligned}$$

If $b_i < \infty$ for all $i \in I$, it is easy to show using Lemma 4.28 that the cost of the span is

$$\left\|\left(\|\operatorname{coker} f\|_p, \|\ker g\|_p\right)\right\|_q \geq 2^{\frac{1-q}{q}} \|(b_i - a_i)_{i\in I}\|_p = 2^{\frac{1-q}{q}} \left\|\bigoplus_{i\in I} K(a_i, b_i)\right\|_p,$$

and the same inequality clearly holds if $b_i = \infty$ for some $i \in I$. However, since $|I| \geq j$, the right-hand side of the inequality cannot be smaller than

$$2^{\frac{1-q}{q}} \left\|(b_i - a_i)_{i\in\{1,\ldots,j\}}\right\|_p = 2^{\frac{1-q}{q}} \left\|\bigoplus_{i=1}^j K(a_i, b_i)\right\|_p,$$

and this contradicts (7). □

Proposition 4.30 *Let $X = \bigoplus_{i=1}^k K(a_i, b_i)$, with the bars ordered non-decreasingly by length. Let $j \in \{1, \ldots, k\}$, and let $Y = \bigoplus_{i=j+1}^k K(a_i, b_i)$ (with $Y = 0$ when $j = k$). Then, for all $p, q \in [1, \infty]$,*

$$d_{\mathcal{S}^p}^q(X, Y) = 2^{\frac{1-q}{q}} \left\|\bigoplus_{i=1}^j K(a_i, b_i)\right\|_p. \tag{8}$$

Proof Since $\operatorname{rank}(Y) = \operatorname{rank}(X) - j$, Proposition 4.29 gives us the inequality

$$d_{\mathcal{S}^p}^q(X, Y) \geq 2^{\frac{1-q}{q}} \left\|\bigoplus_{i=1}^j K(a_i, b_i)\right\|_p.$$

The other inequality, as already noticed in Remark 4.26, follows from Proposition 4.25 and Lemma 4.23 showing that $d_{\mathcal{S}^p}^q(Z \oplus Y, Y) \leq d_{\mathcal{S}^p}^q(Z, 0) = 2^{\frac{1-q}{q}} \|Z\|_p$ with $Z = \bigoplus_{i=1}^j K(a_i, b_i)$. □

In the final part of this subsection we generalize some results from the case of the standard contour to the case of any regular action C.

Definition 4.31 Let C be a regular contour, and let $X = \bigoplus_{i=1}^k K(a_i, b_i)$. We say that (the barcode decomposition of) X has **bars ordered non-decreasingly by lifetime** if $\ell(a_1, b_1) \leq \ell(a_2, b_2) \leq \cdots \leq \ell(a_k, b_k)$, where ℓ denotes the lifetime function associated with C (see Sect. 2.2).

Proposition 4.32 *Let C be a regular action, and let $p, q \in [1, \infty]$. Let $X = \bigoplus_{i=1}^{k} K(a_i, b_i)$, with bars ordered non-decreasingly by lifetime, and let $j \in \{1, \dots, k\}$. Then, for all persistence modules Y,*

1. *if $\mathrm{rank}(Y) \leq \mathrm{rank}(X) - j$, then*

$$d^q_{\mathcal{S}^{p,C}}(X, Y) \geq 2^{\frac{1-q}{q}} \left\| \bigoplus_{i=1}^{j} K(a_i, b_i) \right\|_{p,C};$$

2. *if $Y = \bigoplus_{i=j+1}^{k} K(a_i, b_i)$ (with the convention $Y = 0$ when $j = k$), then*

$$d^q_{\mathcal{S}^{p,C}}(X, Y) = 2^{\frac{1-q}{q}} \left\| \bigoplus_{i=1}^{j} K(a_i, b_i) \right\|_{p,C}.$$

Proof The first statement follows from

$$\begin{aligned} d^q_{\mathcal{S}^{p,C}}(X, Y) &= d^q_{\mathcal{S}^p}(T_C(X), T_C(Y)) \\ &\geq 2^{\frac{1-q}{q}} \left\| \bigoplus_{i=1}^{j} T_C(K(a_i, b_i)) \right\|_p \\ &= 2^{\frac{1-q}{q}} \left\| \bigoplus_{i=1}^{j} K(a_i, b_i) \right\|_{p,C} \end{aligned}$$

where we are using in sequence Proposition 4.19, Proposition 4.29 (observing that the length of a bar $T_C(K(a, b))$ coincides with the lifetime $\ell(a, b)$ of $K(a, b)$, see Proposition 4.8), and Proposition 4.13. The second statement is proven similarly, using Proposition 4.30. □

5 Wasserstein stable ranks: computations and stability

In Sect. 4 it was shown that the Wasserstein distances $d^q_{\mathcal{S}^{p,C}}$ are pseudometrics on Tame. They can therefore be used in the framework of hierarchical stabilization (see Sect. 2.5) to produce stable invariants of persistence modules. The focus of this section is on one type of such invariants, the **Wasserstein stable ranks**, which are the hierarchical stabilization of the rank function with respect to Wasserstein distances $d^q_{\mathcal{S}^{p,C}}$. Denoting $d^q_{\mathcal{S}^{p,C}}$ by d, the stability result for stable ranks (Proposition 2.5) states that for every pair of persistence modules X and Y

$$d(X, Y) \geq d_{\bowtie}\left(\widehat{\mathrm{rank}}_d(X), \widehat{\mathrm{rank}}_d(Y)\right).$$

In the case where $p = q$ and C is the standard contour, combining the above inequality with the stability results of Skraba and Turner (2020) gives several stability results of Wasserstein stable ranks with respect to perturbation of the original data. In particular, (Skraba and Turner 2020, Theorem 4.8) expresses stability with respect to sublevel set filtrations of monotone functions on cellular complexes, (Skraba and Turner 2020, Theorem 5.1) expresses stability with respect to the construction of

cubical complexes from grey scale images, and (Skraba and Turner 2020, Theorem 5.9), expresses stability with respect to Wasserstein distance between point clouds when using the Vietoris-Rips construction.

In order to use the Wasserstein stable ranks in applications, it is important to be able to efficiently compute them as well as distances between them. In this section we use computations of Wasserstein distances from Sect. 4 to derive a formula for the Wasserstein stable rank and propose a convenient formulation of the interleaving distance between stable ranks.

Having defined a rich family of Wasserstein distances $d^q_{\mathcal{S}^{p,C}}$, it is natural to ask whether we can in a supervised learning context search for an optimal distance for a problem at hand. Choosing a suitable parametrization of a contour and leveraging the simple expression of the interleaving distance between Wasserstein stable ranks, in Sect. 5.3 we set up a simple metric learning problem with the aim of observing the interaction between the parameter p and the parameters related to the contour C within the learning. Preliminary results on the optimization of only a contour in a metric learning framework are presented in Gäfvert (2018).

5.1 Computation of the stable rank with Wasserstein distances

The results of this subsection provide explicit formulas to compute the stable rank with respect to the Wasserstein distances $d^q_{\mathcal{S}^{p,C}}$ introduced in Sect. 4. We begin by considering the case $p < \infty$. As in the previous section, if $q = \infty$ we set by convention $\frac{1}{q} = 0$ and $2^{\frac{1-q}{q}} = 2^{-1}$.

Proposition 5.1 *Let $p \in [1, \infty)$ and $q \in [1, \infty]$, let C be a regular action, and let d denote the pseudometric $d^q_{\mathcal{S}^{p,C}}$. Let $X = \bigoplus_{i=1}^k K(a_i, b_i)$, with bars ordered non-decreasingly by lifetime (Definition 4.31), and let $n := |\{i \in \{1, \dots, k\} \mid b_i < \infty\}|$ denote the number of finite bars of X. Then, there exist real numbers $0 = t_0 < t_1 < t_2 < \cdots < t_n$ such that the stable rank function $\widehat{\operatorname{rank}}_d(X) \colon [0, \infty) \to [0, \infty)$ is constant on the intervals $[t_0, t_1)$, $[t_1, t_2)$,..., $[t_{n-1}, t_n)$, $[t_n, \infty)$, and*

$$\widehat{\operatorname{rank}}_d(X)(t_j) = \operatorname{rank}(X) - j,$$

for every $j \in \{0, 1, \dots, n\}$. Furthermore,

$$t_j = 2^{\frac{1-q}{q}} \left\| \bigoplus_{i=1}^j K(a_i, b_i) \right\|_{p,C} = 2^{\frac{1-q}{q}} \left\| (\ell(a_1, b_1), \dots, \ell(a_j, b_j)) \right\|_p$$

for every $j \in \{1, \dots, n\}$, where ℓ is the lifetime function associated with C.

Proof For every $j \in \{1, \dots, k\}$, by Proposition 4.32 $Y_j := \bigoplus_{i=j+1}^k K(a_i, b_i)$ is the closest persistence module to X (in the pseudometric $d^q_{\mathcal{S}^{p,C}}$) such that $\operatorname{rank}(Y_j) = \operatorname{rank}(X) - j$. We have

$$d^q_{\mathcal{S}^{p,C}}(X, Y_j) = 2^{\frac{1-q}{q}} \left\| \bigoplus_{i=1}^j K(a_i, b_i) \right\|_{p,C} =: t_j,$$

with $t_j < \infty$ if, and only if, $j \in \{1, \dots, n\}$. Lastly, we observe that $0 = t_0 < t_1 < t_2 < \cdots < t_n$ as a consequence of the assumption $p < \infty$. □

In particular, when $p < \infty$, the value of the piecewise constant function $\widehat{\text{rank}}_d(X)$ can only decrease by 1 at every discontinuity point t_j. For $p = \infty$, the stable rank has a slightly different behavior. Even though we can define the sequence of real numbers $(t_j)_j$ as in Proposition 5.1, we only have $0 = t_0 \leq t_1 \leq t_2 \leq \cdots \leq t_n$ instead of strict inequalities. Letting s_m denote the m^{th} smallest value in $\{t_j\}_j$ we obtain a sequence $0 = s_0 < s_1 < s_2 < \cdots < s_{n'}$ such that the stable rank with respect to the pseudometric $d := d^q_{\mathcal{S}^{\infty,C}}$ is constant on the intervals $[s_0, s_1), \dots, [s_{n'}, \infty)$, taking the values

$$\widehat{\text{rank}}_d(X)(s_m) = \text{rank}(X) - \max\{j \mid t_j = s_m\}.$$

An explicit formula for the stable rank in the case $p = \infty$ and $q = 1$ was first given in Chachólski and Riihimäki (2020).

Remark 5.2 We observe that for a persistence module X of rank k, once the k bars in the barcode decomposition of X have been ordered non-decreasingly by lifetime, the complexity of computing the discontinuity points of the the Wasserstein stable rank using Proposition 5.1 is linear in k. Therefore the computational complexity of the Wasserstein stable rank is $O(k \log k)$, determined by the complexity of the sorting algorithm to order the bars non-decreasingly by lifetime.

5.2 Interleaving distance between stable ranks

The aim of this subsection is to propose a convenient expression for the interleaving distance (Sect. 2.5) between two non-increasing piecewise constant functions. We assume functions to take only finitely many values, that is the case of stable ranks which will be the object of our study. Let $f, g \colon [0, \infty) \to [0, \infty)$ be non-increasing piecewise constant functions. If $\lim_{t\to\infty} f(t) \neq \lim_{t\to\infty} g(t)$, then $d_{\bowtie}(f, g) = \infty$. For the computation of the interleaving distance we can therefore assume that the functions f and g have the same limit value and denote it by L. Given a non-increasing piecewise constant function $f \colon [0, \infty) \to [0, \infty)$ with limit value L, we define the non-increasing piecewise constant function $f^{-1} \colon [L, \infty) \to [0, \infty)$ with values $f^{-1}(y) := \inf\{t \mid f(t) \leq y\}$. If in addition the function f is right-continuous, then $f^{-1}(y) = \min\{t \mid f(t) \leq y\}$. We observe that for every right-continuous non-increasing piecewise constant function f we have $f^{-1}(f(t)) \leq t$ for all t, and equality holds if t is a discontinuity point of f. Moreover, $f(f^{-1}(y)) \leq y$ for all $y \geq L$, and equality holds if $y \in \text{im } f$. Our focus in this subsection will be on the discontinuity points $\{t_i\}$ of f and on the values in $\text{im } f$, rather than on the full domain and codomain of f, thus justifying our use of the notation f^{-1}.

Proposition 5.3 *Consider two right-continuous non-increasing piecewise constant functions $f, g \colon [0, \infty) \to [0, \infty)$ having the same limit value L. Using the nota-*

tion introduced above, we have:

$$d_{\bowtie}(f,g) = \|f^{-1} - g^{-1}\|_{\infty}.$$

Proof Let us define the following subset of $[0,\infty)$,

$$A(f,g) := \{\varepsilon \in [0,\infty) \mid f(t) \geq g(t+\varepsilon) \text{ and } g(t) \geq f(t+\varepsilon), \text{ for all } t \in [0,\infty)\}.$$

Remember that, by definition, $d_{\bowtie}(f,g) = \inf A(f,g)$.

We first prove that $d_{\bowtie}(f,g) \geq \|f^{-1} - g^{-1}\|_{\infty}$. Let $\varepsilon \in A(f,g)$. Then, for all $y \geq L$, we have $y \geq f(f^{-1}(y)) \geq g(f^{-1}(y)+\varepsilon)$. Composing by the non-increasing function g^{-1} and recalling that $g^{-1}(g(t)) \leq t$ for all t, we obtain $f^{-1}(y)+\varepsilon \geq g^{-1}(y)$. We have thus shown that $g^{-1}(y) - f^{-1}(y) \leq \varepsilon$, for all $y \geq L$ and $\varepsilon \in A(f,g)$, which implies $g^{-1}(y) - f^{-1}(y) \leq d_{\bowtie}(f,g)$, for all $y \geq L$. By symmetry in the roles of f and g, we conclude that $|g^{-1}(y) - f^{-1}(y)| \leq d_{\bowtie}(f,g)$, for all $y \geq L$.

We now prove that $d_{\bowtie}(f,g) \leq \|f^{-1} - g^{-1}\|_{\infty}$ by showing that $\varepsilon := \|f^{-1} - g^{-1}\|_{\infty}$ is in $A(f,g)$. By the definition of ε, $g^{-1}(y) \leq f^{-1}(y) + \varepsilon$ for every $y \in [L,\infty)$. Moreover if $y = f(t)$ for $t \in [0,\infty)$, as discussed above, we have $f^{-1}(y) = f^{-1}(f(t)) \leq t$, which together with the fact that g is a non-increasing function proves the following inequalities:

$$f(t) = y \geq g(g^{-1}(y)) \geq g(f^{-1}(y)+\varepsilon) \geq g(t+\varepsilon).$$

By symmetry, we also get $g(t) \geq f(t+\varepsilon)$, and we conclude that $\varepsilon \in A(f,g)$. □

Let $X = \bigoplus_{i=1}^{k} K(a_i^x, b_i^x)$ be a persistence module with bars ordered non-decreasingly by lifetime, and with n finite bars. Let $f := \widehat{\text{rank}}_d(X)$ denote the corresponding Wasserstein stable rank with respect to the distance $d = d^q_{\mathcal{S}^{p,C}}$, for some $p, q \in [1,\infty]$ and a regular action C. The sequence $0 = t_0 \leq t_1 \leq t_2 \leq \cdots \leq t_n$ such that f is constant on the intervals $[t_0,t_1)$, $[t_1,t_2)$,..., $[t_{n-1},t_n)$, $[t_n,\infty)$ defined in Sect. 5.1 is enough to encode f^{-1} as a finite vector $\hat{f}^{-1} := (\hat{f}_i^{-1})_{i\in\{0,\ldots,n\}}$, where

$$\hat{f}_i^{-1} := t_{n-i} = 2^{\frac{1-q}{q}} \left\| \left(\ell\left(a_1^x, b_1^x\right), \ldots, \ell\left(a_{n-i}^x, b_{n-i}^x\right)\right)\right\|_p$$

for $i \in \{0,\ldots,n-1\}$ and $\hat{f}_n^{-1} = 0$. Indeed, the limit value of f is the number $L = k - n$ of infinite bars of X; for all $i \in \{0,\ldots,n\}$ we have $f^{-1}(L+i) = \hat{f}_i^{-1}$, and for any $y \in [L,\infty)$ the value $f^{-1}(y)$ equals $f^{-1}(L+i)$, where $i \in \{0,\ldots,n\}$ is the largest integer such that $L + i \leq y$.

Let $Y = \bigoplus_{i=1}^{l} K(a_i^y, b_i^y)$ be another persistence module with bars ordered non-decreasingly by lifetime, and with m finite bars. Suppose that X and Y have the same number of infinite bars, $L = k - n = l - m$. Let $g := \widehat{\text{rank}}_d(Y)$ denote the Wasserstein stable rank of Y with respect to the distance $d = d^q_{\mathcal{S}^{p,C}}$. The interleaving distance between f and g can then be written as the L^∞ norm of the vector $(\hat{f}_i^{-1} -$

$\hat{g}_i^{-1})_{i\in\{0,\dots,\min(n,m)\}}$ with components

$$\hat{f}_i^{-1} - \hat{g}_i^{-1} = 2^{\frac{1-q}{q}} \Big(\big\| \big(\ell - (a_1^x, b_1^x), \dots, \ell\big(a_{n-i}^x, b_{n-i}^x\big)\big) \big\|_p - \big\| \big(\ell(a_1^y, b_1^y), \dots, \ell\big(a_{m-i}^y, b_{m-i}^y\big)\big) \big\|_p \Big) \tag{9}$$

for $i \in \{0, \dots, \min(n, m) - 1\}$, and last component

$$\hat{f}^{-1}_{\min\{n,m\}} - \hat{g}^{-1}_{\min\{n,m\}} = \begin{cases} -\|(\ell(a_1^y, b_1^y), \dots, \ell(a_{m-n}^y, b_{m-n}^y))\|_p & \text{if } \min(n,m) = n < m \\ \|(\ell(a_1^x, b_1^x), \dots, \ell(a_{n-m}^x, b_{n-m}^x))\|_p & \text{if } \min(n,m) = m < n \\ 0 & \text{if } n = m. \end{cases}$$

We observe that the considered vector encodes the function $(f^{-1} - g^{-1})$ on the intervals $[L, L+1), \dots, [\min(k, l), \min(k, l) + 1)$ of length one where both f^{-1} and g^{-1} are constant. Since f^{-1} and g^{-1} are nonincreasing and for $x \geq \min(k, l)$ either $f^{-1}(x) = 0$ or $g^{-1}(x) = 0$, it is enough to consider those intervals.

Remark 5.4 For two persistence modules X and Y both of rank k, the complexity of computing the interleaving distance is dominated by the sorting of the bars in the respective barcode decompositions of X and Y, since forming the vector as in (9) and computing its L^∞ norm can be done linearly in k. The computational complexity of the interleaving distance between Wasserstein stable ranks is thus $O(k \log k)$.

Example 5.5 Consider a persistence module $Y = \bigoplus_{i=1}^3 K(a_i, b_i)$ with bars ordered non-decreasingly by lifetime and $X = K(a_0, b_0) \oplus Y$ such that $\varepsilon := \ell(a_0, b_0) \leq \ell(a_1, b_1)$. By using the formula (9) and observing that

$$\|(\ell(a_0, b_0), \dots, \ell(a_i, b_i))\|_p - \|(\ell(a_1, b_1), \dots, \ell(a_i, b_i))\|_p \leq \ell(a_0, b_0)$$

for $i \in \{1, 2, 3\}$ by properties (1) and (2) of p-norms, we see that the interleaving distance between $\widehat{\text{rank}}_d(X)$ and $\widehat{\text{rank}}_d(Y)$ with $d = d^q_{\mathcal{S}^{p,C}}$ is given by $2^{\frac{1-q}{q}}\varepsilon$. Note that by Proposition 4.32 we know $d^q_{\mathcal{S}^{p,C}}(X, Y) = 2^{\frac{1-q}{q}} \|K(a_0, b_0)\|_{p,C} = 2^{\frac{1-q}{q}}\varepsilon$. Therefore in this case the interleaving distance between stable ranks with respect to Wasserstein distance coincides with the Wasserstein distance between X and Y. Note however that this is not always the case. The Wasserstein stable ranks of X and Y with respect to $d^q_{\mathcal{S}^{p,C}}$, with parameters $q = 1$, $p = 2$ and C the standard contour, are shown in Fig. 2, together with their "inverse" functions which are used for the computation of the interleaving distance.

Let us keep denoting $d^q_{\mathcal{S}^{p,C}}$ by d. It follows from triangle inequality and Lemma 4.23 that:

$$d(X, Y) \geq 2^{\frac{1-q}{q}} |\|X\|_p - \|Y\|_p|.$$

This inequality can be refined by

$$d(X, Y) \geq d_{\bowtie}\left(\widehat{\text{rank}}_d(X), \widehat{\text{rank}}_d(Y)\right) \geq 2^{\frac{1-q}{q}} |\|X\|_p - \|Y\|_p|,$$

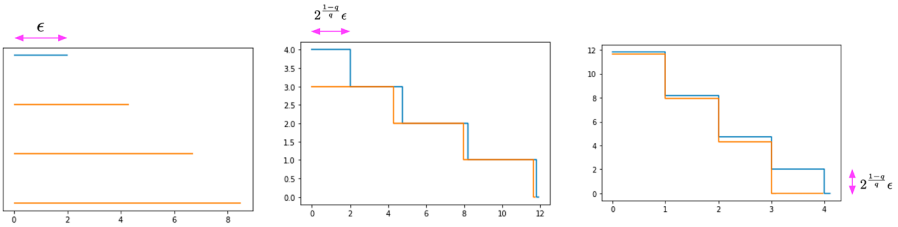


Fig. 2 Schematic representation of the computation of the interleaving distance in Example 5.5. **Left:** Barcode decomposition of Y in orange and bar $K(a_0, b_0)$ in blue. **Middle:** Stable ranks computed with standard contour, $q = 1$ and $p = 2$. The functions $\widehat{\text{rank}}_d(X)$ and $\widehat{\text{rank}}_d(Y)$ are represented in blue and orange, respectively. **Right:** Inverse stable ranks for the computation of interleaving distance, with $\widehat{\text{rank}}_d^{-1}(X)$ in blue and $\widehat{\text{rank}}_d^{-1}(Y)$ in orange. The interleaving distance between stable ranks can be computed as $\|\widehat{\text{rank}}_d^{-1}(X) - \widehat{\text{rank}}_d^{-1}(Y)\|_\infty = 2^{\frac{1-q}{q}}\varepsilon$, illustrated with the pink arrow (colour figure online)

where the first inequality is given by the stability theorem of hierarchical stabilization (Proposition 2.5) and the second inequality is provided by the characterization of interleaving distances between stable ranks in Proposition 5.3 and Eq. (9) stating that the interleaving distance between the Wasserstein stable ranks of persistence modules X and Y is the L^∞ norm of a vector with 0^{th} component $2^{\frac{1-q}{q}}|\|X\|_p - \|Y\|_p|$.

An example where the second inequality is strict is provided by Example 5.5 for $p > 1$, while an example where this is an equality is provided in the case $Y = 0$ by Lemma 4.23. A simple example in which the first inequality is strict is provided instead by $X = K(0, 1)$, $Y = K(0, 2)$ and $q = 2$.

Remark 5.6 Since stable ranks are measurable functions $[0, \infty) \to [0, \infty)$, there are many pseudometrics to compare them other than the interleaving distance $d_{\bowtie}$. In particular, one can consider the standard L^p-pseudometrics, here denoted by $d_p(f, g) := \left(\int_0^\infty |f(t) - g(t)|^p \, dt\right)^{\frac{1}{p}}$. As shown in Chachólski and Riihimaki (2020, Prop. 2.1), the stability theorem of hierarchical stabilization implies the following bounds for d_p:

$$c\, d(X, Y)^{\frac{1}{p}} \geq d_p\left(\widehat{\text{rank}}_d(X), \widehat{\text{rank}}_d(Y)\right),$$

for any persistence modules X and Y, where $c := \max\{\text{rank}(X), \text{rank}(Y)\}$ and d denotes any pseudometric between persistence modules. In this article we have chosen to work with the interleaving distance between Wasserstein stable ranks because of the strong stability result, expressed as a 1-Lipschitz condition. Lipschitz stability for Wasserstein distances other than W_1 can not be obtained for example by considering linear representations of persistence diagrams (Hofer et al. 2017; Adams et al. 2017; Chen et al. 2015; Kusano et al. 2017; Reininghaus et al. 2015) as proved in Theorem 6.3 in Skraba and Turner (2020). The trade-off between stability and the possibility of exploiting a Banach or Hilbert space structure is still to be explored.

5.3 Metric learning

We have defined distances $d^q_{\mathcal{S}^{p,C}}$ between persistence modules, parametrized by q, p and by a contour C, and computable stable rank invariants with corresponding interleaving distances. These distances can be pulled back to compare persistence modules in Tame via the function $\widehat{\text{rank}}_d$, with $d = d^q_{\mathcal{S}^{p,C}}$. Recalling that the stable ranks depend on the pseudometric $d^q_{\mathcal{S}^{p,C}}$, we now turn to the question of how to choose p and C. The optimization of the parameter q is not relevant, since it determines a constant multiplicative factor to the distance of each pair of persistence modules. We thus fix $q = 1$ for a direct comparison with the original framework of noise systems.

For brevity, we write $d := d^1_{\mathcal{S}^{p,C}}$ and $d_{\bowtie,p,C}(X, Y) := d_{\bowtie}(\widehat{\text{rank}}_d(X), \widehat{\text{rank}}_d(Y))$. The field of metric learning provides a variety of loss functions suited for different machine learning problems. For example, if we consider a simple binary classification problem we have a dataset of persistence modules $\{X_i\}_{i\in I}$ and the index set I is partitioned into two sets A and B, to represent the labeling. For this problem, a loss function (from Zhao and Wang 2019), designed to yield small intra-class distances and large inter-class distances can be formulated as:

$$\mathcal{L} = \frac{\sum_{i,j\in A}(d_{\bowtie,p,C}(X_i, X_j))^2}{\sum_{i\in A, j\in I}(d_{\bowtie,p,C}(X_i, X_j))^2} + \frac{\sum_{i,j\in B}(d_{\bowtie,p,C}(X_i, X_j))^2}{\sum_{i\in B, j\in I}(d_{\bowtie,p,C}(X_i, X_j))^2}. \tag{10}$$

In order to proceed we need to choose a family of contours that is practically searchable when minimizing the loss function above. We work with contours of distance type which are parametrized by densities (see Sect. 2.2). In turn, in order to use gradient optimization methods, we want the densities to be parametrized by a finite real-valued parameter vector. To this aim we choose as densities unnormalized Gaussian mixtures $f(x) = \sum_{i=1}^k \lambda_i \mathcal{N}(x \mid \mu_i, \sigma_i)$ for some chosen k, where $\mathcal{N}$ is Gaussian with mean μ_i and standard deviation σ_i, and $\lambda_1 = 1$.

In summary, the metric learning problem amounts to minimizing the loss function with respect to a parameter vector $\theta \in \mathbb{R}^{3k}$, i.e. $\theta = (\mu_1, \ldots, \mu_k, \sigma_1, \ldots, \sigma_k, \lambda_2, \ldots, \lambda_k, p)$, designed to learn conjointly the parameter p and the parameters of the contour of the algebraic Wasserstein distance. The loss function is a simple function of the pairwise interleaving distances between Wasserstein stable ranks of persistence modules in the dataset. As can be seen in Proposition 5.3 and the expression (9), the interleaving distance between stable ranks is the L^∞ norm of differentiable functions with respect to θ and is therefore differentiable almost everywhere with respect to θ, implying the same behavior for the loss function. Hence the metric learning problem is amenable to gradient-based optimization methods such as gradient descent.

6 Examples of analyses with Wasserstein stable ranks

In a first experiment, we show how varying the parameter p affects the distance space of the Wasserstein stable ranks and can serve as a way to weight the importance of long bars versus short bars, for a set of synthetic persistence modules. In a second

experiment, we illustrate on a real-world dataset how learning the parameter p together with the parameters of a contour can lead to more discriminative Wasserstein stable ranks in a classification problem.

6.1 Synthetic data

A straightforward way to apply persistent homology in the context of computer vision is to construct a complex (e.g. cubical complex) from the grid of pixels constituting an image. The complex is then filtered based on the grayscale intensity of the pixels (or based on the color channels for color images).

It is easy to see that, in this context, what should be considered as signal versus noise in a barcode representation of the data is highly dependent on the application. For example, for classification of handwritten digits from the MNIST dataset (Garin and Tauzin 2019; Turkeš et al. 2021) the dominant topological features are often the most discriminative (for instance the existence of a 1-dimensional cycle may be enough to distinguish between digits 0 and 1). On the other hand, in biomedical imaging (Chung et al. 2018; Qaiser et al. 2019) pathological states can translate into images with irregularities or lack of homogeneity, associated with high numbers of short-lived components as observed in Garside et al. (2019).

Inspired by these applications, we construct two much simpler synthetic datasets of images and associated persistence modules, with the goal of illustrating the effect of choosing the parameter p when using Wasserstein stable ranks. The parameter q is set to 1 and the contour is fixed to be the standard contour. In other words, we study the effect of the parameter p on how the function $\widehat{\mathrm{rank}}_d$, with $d = d_{\mathcal{S}^p}^1$, maps persistence modules onto the space of stable ranks, endowed with the interleaving distance. Each dataset is composed of 100 images together with their class label, A or B. Each image is composed of one block of high-intensity pixels and a number of blocks of low-intensity pixels (while the size of the pixel blocks does not have a direct impact on the following persistent homology analysis, the high-intensity block is made larger for visual clarity, see Figs. 3, 4). The images are represented as cubical complexes on which super-level set filtration is performed and we analyze the H_0 barcodes obtained from this process. Since we use pixel intensity [0, 255] and super-level sets are used, the resulting filtration scale is $[255, -\infty)$. This is capped to the minimum pixel value, 0, and transformed as $255 - x$ to obtain a filtration scale [0, 255] as can be seen in the barcodes in Figs. 3, 4.

- In Dataset 1 the pixels in the high-intensity block have slightly higher intensity in images from class A (uniformly distributed between 245 and 255) compared to images of class B (between 200 and 210). The low-intensity blocks however follow the same distribution for images of both classes (the number of blocks is uniformly distributed between 50 and 100 and the intensity is between 1 and 10). Sample images and barcodes are shown in Fig. 3.
- In Dataset 2 on the other hand, the intensity of the high-intensity blocks follows the same distribution for both classes (uniformly distributed between 100 and 255). The number of low-intensity blocks however follows a different distribution for Class A (between 20 and 30) and Class B (between 120 and 130). Their intensity

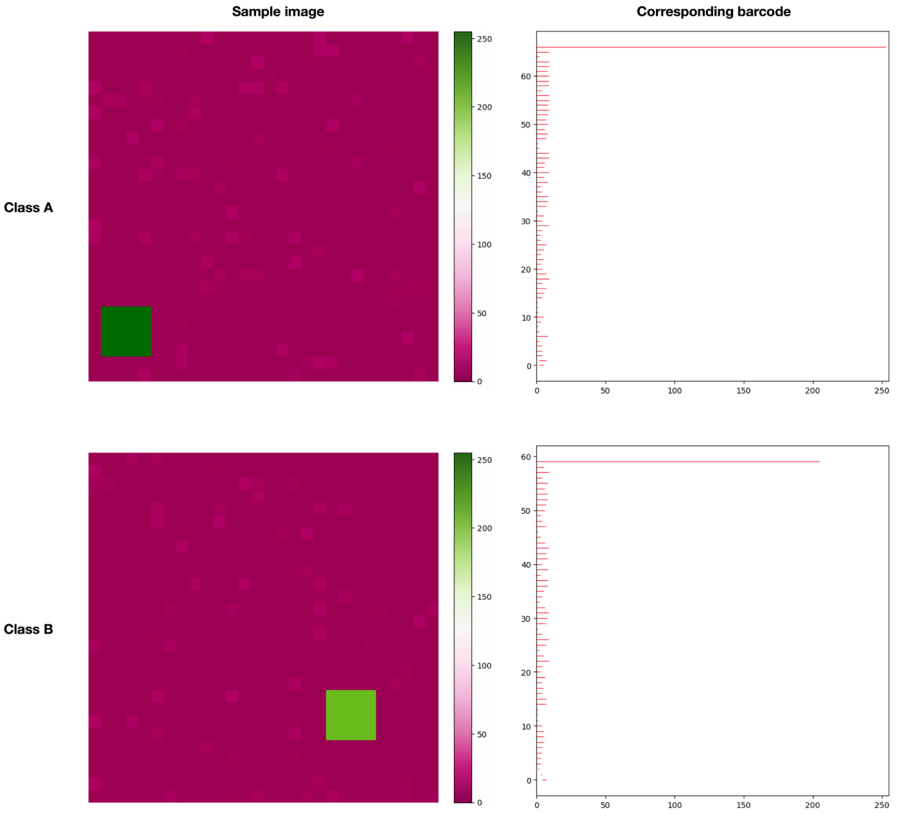


Fig. 3 Dataset 1. **Left:** Sample images from classes A and B. **Right**: H_0 barcodes corresponding to the sample images

is the same for both classes (between 1 and 10). Sample images and barcodes are shown in Fig. 4.

This construction induces distributions of barcodes where barcode features (i.e. length of longest bar or number of bars in a given length range) are expected to be statistically distinct or indistinguishable. In terms of the barcodes, for Dataset 1 the signal is by construction the single dominant topological feature (the long bar, whose length follows statistically different distributions between classes) while the noise is composed of the numerous short bars (corresponding to low intensity blocks, for which the number and intensity follows the same distribution in both classes).

In accordance with the intuition, for Dataset 1, choosing a value of $p = \infty$ when generating the stable ranks effectively "denoises" the barcodes and organizes the space of Wasserstein stable ranks in a way where stable ranks of the same class are close to each other in interleaving distance but far from elements of the other class. Stable ranks corresponding to $p = 1$ however fail to organize the corresponding distance space in this clear-cut way, being too sensitive to the noisy short bars in the barcodes. To illustrate this effect, in Fig. 5 we show the hierarchical clustering (with average

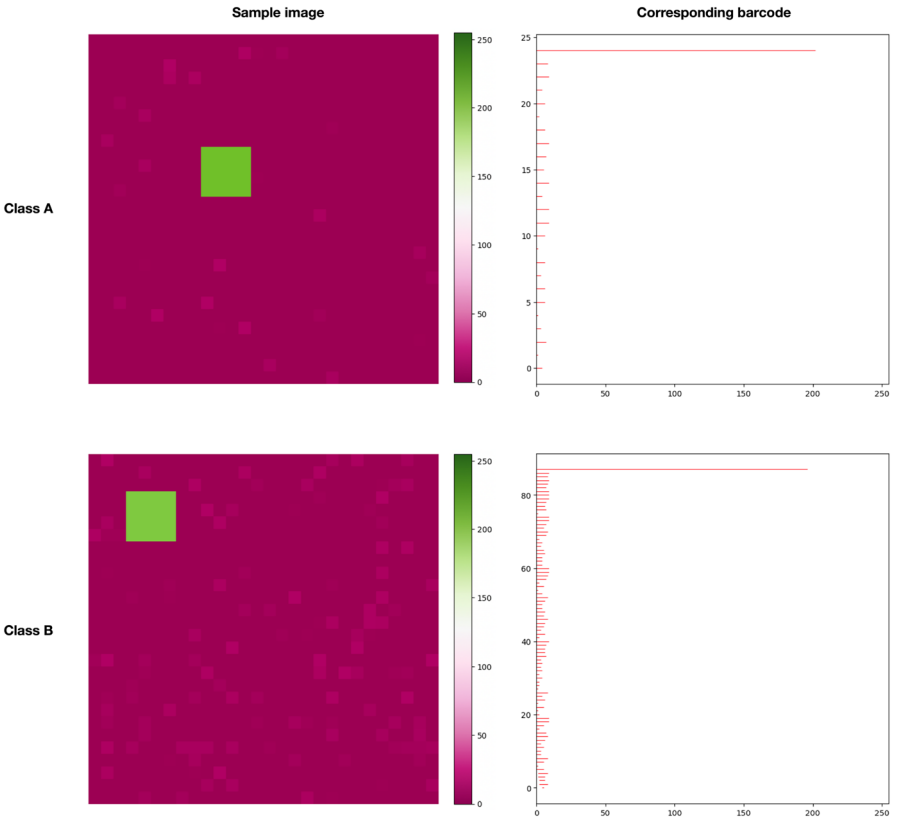


Fig. 4 Dataset 2. **Left:** Sample images from classes A and B. **Right**: H_0 barcodes corresponding to the sample images

linkage, similar results were observed for complete and single linkage) corresponding to the distance spaces of Wasserstein stable ranks for $p = 1$ and $p = \infty$.

On the contrary, for Dataset 2 the signal is by construction the number of short bars (numbers which follow statistically different distributions) while the noise is the single long bar (generated by blocks following the same distribution for both classes). In this case a choice of $p = 1$ organizes the space of stable ranks such that elements of the same class cluster together, while $p = \infty$, being too sensitive to the (for this dataset) noisy long bar, fails to do so. This is illustrated in Fig. 6.

While the effect of changing p on the structure of the distance space is clear for the parameters used to generate our artificial datasets, some class-based structure remains on a small scale, also when choosing $p = \infty$. By increasing the amount of noise it is however possible to induce e.g. a nearest neighbor classifier to perform arbitrary poorly for the $p = \infty$ while still distinguishing the classes for $p = 1$ (and vice versa for Dataset 1).

The choice of the parameter value p, which we have demonstrated can have a large impact, is essentially related to the underlying distance between persistence modules.

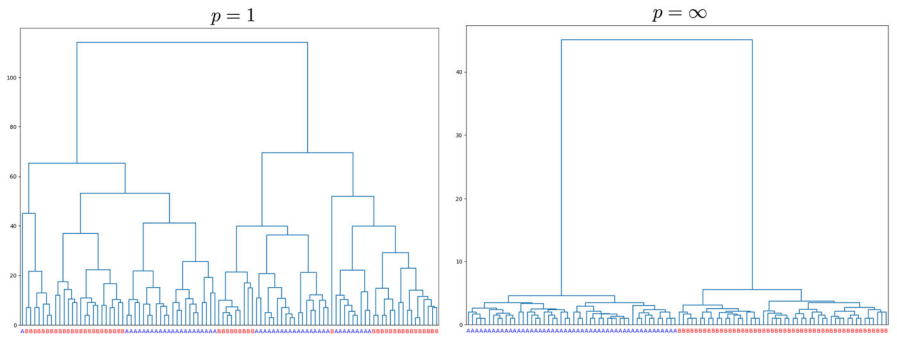


Fig. 5 Dataset 1. Hierarchical clustering on the Wasserstein stable ranks for $p = 1$ (left) and $p = \infty$ (right) with respect to the interleaving distance. The leaves (stable ranks in the dataset) are labeled and colored according to their class

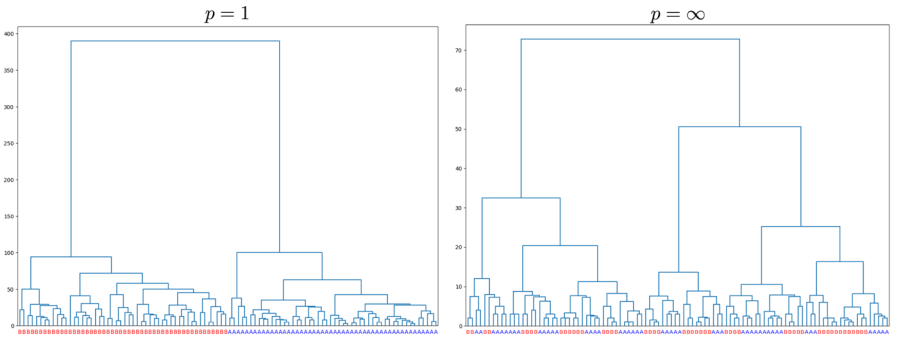


Fig. 6 Dataset 2. Hierarchical clustering on the Wasserstein stable ranks for $p = 1$ (left) and $p = \infty$ (right) with respect to the interleaving distance. The leaves (stable ranks in the dataset) are labeled and colored according to their class

Using Wasserstein-stable invariants however has computational advantages, facilitates learning the right parameters for a particular problem and allows for a richer use of machine learning methods as we illustrate in the next subsection on a real-world dataset.

6.2 Brain artery data

In Bullitt et al. (2010) a dataset of brain artery trees corresponding to 97 subjects aged 18 to 72 is introduced. Each data point is modeled as a tree embedded in $\mathbb{R}^3$. In Bendich et al. (2016) the dataset is further analyzed with persistent homology. To be able to apply a sublevel set filtration on the tree, a real-valued function is defined on the vertices as the height of the vertex in the 3D-embedding. This is extended to a function on the edges by taking the maximum value of the weights of the vertices connected by the edge. After applying persistent homology, each tree is represented by a vector containing the sorted lengths of the 100 longest bars in a barcode decomposition of the corresponding persistence module. This feature is further used to demonstrate, among

Table 1 Converged values for the parameters of the metric learning problem (mean and standard deviation over the LOOCV folds)

p	μ_1	μ_2	σ_1	σ_2	λ
1.10 ± 0.04	16.45 ± 0.75	81.88 ± 0.27	15.51 ± 0.45	4.55 ± 0.24	1.15 ± 0.05

other things, an age effect of brain artery structure, by showing that the projection of the vectors on the first principal component of the dataset is correlated with age.

The authors note that using vectors of sorted length was computationally more feasible than computing Wasserstein distances between the persistence diagrams and they are more amenable to statistical analysis. In addition, the authors observed that it was not necessary to use the whole vector of lengths to establish the correlation and in fact the topological features of medium length, rather than the longest ones, were the most discriminatory.

Analyzing the dataset with stable ranks offers computational and statistical advantages. Moreover, for this problem where the discriminative information is not contained in the most persistent feature, considering other distances than the bottleneck ($p = \infty$) and more generally tuning the parameter p might be beneficial. Finally, combining the tuning of the parameter p with a contour might increase the power of the method. Indeed the parameter p and the contour, intuitively are related to different features of a persistence barcode: while the parameter p globally weights the importance of long versus short bars as illustrated in Sect. 6.1, the contour allows to focus on the most informative filtration scales.

While we also study age effects of brain artery structure, we choose to binarize the problem by creating two classes of similar size: *young* ($age < 45$, 50 subjects) and *old* ($age \geq 45$, 47 subjects) and treat the problem as a classification.

We start by studying the effect of varying p alone. We compute the distances between Wasserstein stable ranks with standard contour. We classify the samples in the distance space thus obtained by using the k-nearest neighbors algorithm (Pedregosa et al. 2011) (the parameter k is chosen in a cross-validation procedure). Repeating this for various values of p we observe a difference in the resulting accuracy, plotted in Fig. 7, with the highest values obtained for p in the medium range (2–3). This is in line with the conclusion in Bendich et al. (2016) that the highest persistent features alone have a small distinguishing power, while medium sized bars reflect variations in brain artery trees within subject of different ages. In Bendich et al. (2016) only the length of bars in a barcode is used to compare barcodes of different classes. With our features parametrized by p and a contour C, we can take into consideration both the length of bars and their position in the parameter space.

We therefore next turn to the problem of learning the contour of the stable ranks as well. We use the metric learning method described in Sect. 5.3. Using leave-one-out cross-validation (LOOCV), for each training fold we learn the metric that optimally separates training samples from the two classes by minimizing the loss defined in (10). We then classify using KNN in the obtained distance space.

For metric learning, the contours are parametrized by densities which are unnormalized Gaussian mixtures with two components. The loss function is implemented

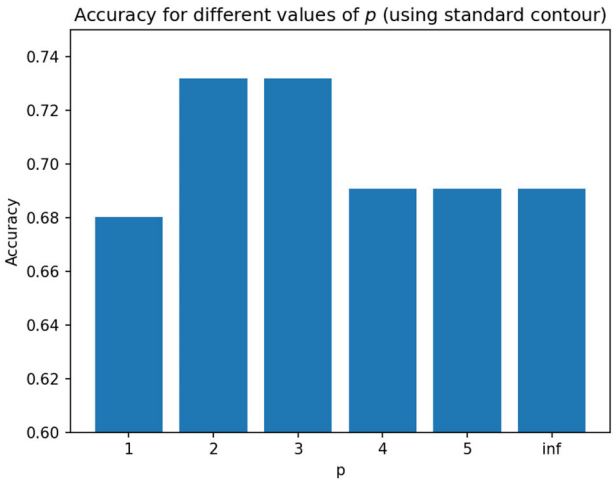


Fig. 7 Accuracy on the brain artery problem using distances between stable ranks and KNN, for standard contour and different values of p

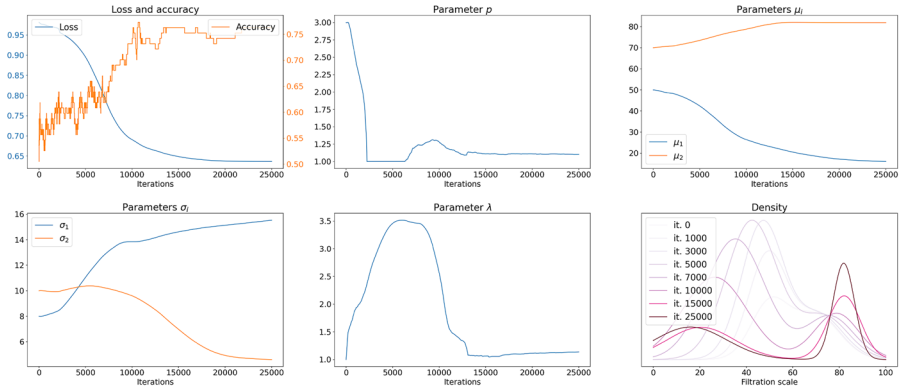


Fig. 8 Results for one example run of the metric learning optimization for Wasserstein stable ranks (see Sect. 5.3) over 25000 iterations. **Top Left:** Progression of the loss and the KNN training fold accuracy over the iterations. **Top Middle, Top Right, Bottom Left, Bottom Middle**: Progression of the parameters in $\theta = (\mu_1, \mu_2, \sigma_1, \sigma_2, \lambda_2, p)$ parametrizing Wasserstein stable ranks: p, mean μ_i, standard deviation σ_i and λ_2 respectively over the iterations. **Bottom Right**: Density at different iterations

in PyTorch (Paszke et al. 2019). After a random initialization of the parameters, projected gradient descent (to respect the constraints $p \geq 1, \lambda_i, \sigma_i > 0$) with momentum is used to achieve a lower loss. An example of an optimization on a training fold over 25000 iterations is shown in Fig. 8. The average and standard deviation of the optimized parameters, over the LOOCV folds can instead be found in Table 1.

The metric learning is effective in finding distances that improve the classification performance: running the optimization problem not only decreases the loss but also increases the corresponding classification accuracy (as is seen in Fig. 8 in the top left plot), reaching 76.3% with the parameter values summarized in Table 1.

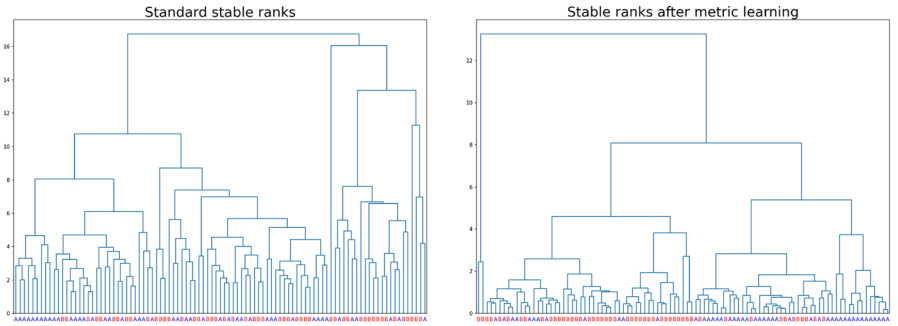


Fig. 9 Hierarchical clustering on the standard stable ranks (left) and the optimized stable ranks resulting from the metric learning problem (right) with respect to the interleaving distance. The leaves (stable ranks in the dataset) are labeled and colored according to their class (A is $age \geq 45$, B is $age < 45$)

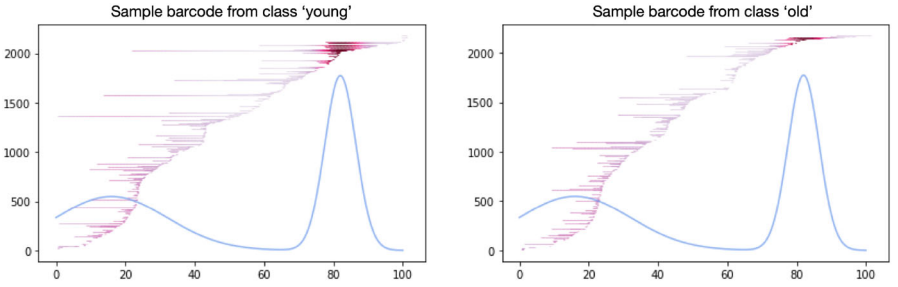


Fig. 10 Sample barcodes from the two classes with superposed learned density. Bars are colored according to the density

This is an improvement compared to the accuracies obtained at random initialization (between 44.3 and 71.1% for 10 random initializations of the parameters in Table 1), showing the benefit of learning, but also compared to the results obtained when only varying p and considering the standard contour in Fig. 7. It is thus when we learn both p and the contour that the best loss and corresponding classification accuracy is achieved.

In Fig. 9 we illustrate the effect of the metric learning by plotting the hierarchical clustering (with average linkage) corresponding to the standard stable ranks (i.e., with $p = \infty$ and standard contour) and to the optimized stable ranks. We see that the optimized stable ranks (with the exception of two outliers) group into two clusters: one with a majority of class A and the other with a majority of class B, while the pattern for standard stable ranks is less clear.

The optimal parameters found with the metric learning method are of interest because they allow to construct a distance space in which machine learning methods can be carried out, but they are also interpretable: they contain information about which features of the dataset are important to distinguish the two classes.

This is illustrated in Fig. 10 where two sample barcodes - one from each class - are displayed with the optimal density superposed and the bars colored according to the density. From the insight that some parts of the filtration scale are more important in

distinguishing younger from older subjects, one may pursue the analysis by looking for characteristics of bars in that region of the barcode. One can also take the analysis a step further by looking at the object from which the filtered simplicial complex was created. In our case, since the filtration scale corresponds to the height (z-coordinate) in the 3D-embedding of the brain artery tree, one may for example investigate whether differences in brain artery between subjects of different ages in this particular region carries a biological meaning.

Acknowledgements This work was partially supported by the Swedish Research Council (Vetenskapsrådet) and the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation.

Funding Open access funding provided by Royal Institute of Technology.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no Conflict of interest.

References

Adams, H., Moy, M.: Topology applied to machine learning: from global to local. Front. Artif. Intell. **4**, 668302 (2021)

Adams, H., Emerson, T., Kirby, M., Neville, R., Peterson, C., Shipman, P., Chepushtanova, S., Hanson, E., Motta, F., Ziegelmeier, L.: Persistence images: a stable vector representation of persistent homology. J. Mach. Learn. Res. **18**, 1–35 (2017)

Agerberg, J., Ramanujam, R., Scolamiero, M., Chachólski, W.: Supervised learning using homology stable rank kernels. Front. Appl. Math. Stat. **7**, 668046 (2021)

Bauer, U., Lesnick, M.: Induced matchings and the algebraic stability of persistence barcodes. J. Comput. Geom. **6**(2), 162–191 (2015)

Bauer, U., Lesnick, M.: Persistence diagrams as diagrams: a categorification of the stability theorem. In: Baas, N.A., Carlsson, G.E., Quick, G., Szymik, M., Thaule, M. (eds.) Topological Data Analysis, pp. 67–96. Springer, Cham (2020)

Bauer, U., Schmahl, M.: Lifespan functors and natural dualities in persistent homology. Homol. Homotopy Appl. **25**(2), 297–327 (2023)

Bendich, P., Marron, J.S., Miller, E., Pieloch, A., Skwerer, S.: Persistent homology analysis of brain artery trees. Ann. Appl. Stat. **10**(1), 198 (2016)

Bubenik, P., Milićević, N.: Homological algebra for persistence modules. Found. Comput. Math. **21**(5), 1233–1278 (2021)

Bubenik, P., De Silva, V., Scott, J.: Metrics for generalized persistence modules. Found. Comput. Math. **15**(6), 1501–1531 (2015)

Bubenik, P., Scott, J., Stanley, D.: Exact weights, path metrics, and algebraic Wasserstein distances. J. Appl. Comput. Topol. **7**(2), 185–219 (2023)

Bullitt, E., Zeng, D., Mortamet, B., Ghosh, A., Aylward, S.R., Lin, W., Marks, B.L., Smith, K.: The effects of healthy aging on intracerebral blood vessels visualized by magnetic resonance angiography. Neurobiol. Aging **31**(2), 290–300 (2010)

Carrière, M., Chazal, F., Ike, Y., Lacombe, T., Royer, M., Umeda, Y.: Perslay: A neural network layer for persistence diagrams and new graph topological signatures. In: International Conference on Artificial Intelligence and Statistics, pp. 2786–2796. PMLR (2020)

Chachólski, W., Riihimäki, H.: Metrics and stabilization in one parameter persistence. SIAM J. Appl. Algebra Geom. **4**(1), 69–98 (2020)

Chazal, F., De Silva, V., Glisse, M., Oudot, S.: The Structure and Stability of Persistence Modules. Springer Briefs in Mathematics, vol. 10. Springer, New York (2016)

Chen, C., Edelsbrunner, H.: Diffusion runs low on persistence fast. In: 2011 International Conference on Computer Vision, pp. 423–430. IEEE (2011)

Chen, Y.-C., Wang, D., Rinaldo, A., Wasserman, L.: Statistical analysis of persistence intensity functions. arXiv preprint arXiv:1510.02502v1 (2015)

Chubet, O.A., Gardner, K.P., Sheehy, D.R.: A theory of sub-barcodes. arXiv preprint arXiv:2206.10504v1 (2022)

Chung, Y.-M., Hu, C.-S., Lawson, A., Smyth, C.: Topological approaches to skin disease image analysis. In: 2018 IEEE International Conference on Big Data (Big Data), pp. 100–105. IEEE (2018)

Cohen-Steiner, D., Edelsbrunner, H., Morozov, D.: Vines and vineyards by updating persistence in linear time. In: Proceedings of the Twenty-second Annual Symposium on Computational Geometry, pp. 119–126 (2006)

Cohen-Steiner, D., Edelsbrunner, H., Harer, J., Mileyko, Y.: Lipschitz functions have L_p-stable persistence. Found. Comput. Math. **10**(2), 127–139 (2010)

de Silva, V., Munch, E., Stefanou, A.: Theory of interleavings on categories with a flow. Theory Appl. Categ. **33**(21), 583–607 (2018)

Edelsbrunner, H., Letscher, D., Zomorodian, A.: Topological persistence and simplification. In: Proceedings 41st Annual Symposium on Foundations of Computer Science, pp. 454–463. IEEE (2000)

Gäfvert, O., Chachólski, W.: Stable invariants for multidimensional persistence. arXiv preprint arXiv:1703.03632v3 (2017)

Gäfvert, O.: Topology-based metric learning. Conference Poster. Available: https://people.kth.se/~oliverg/TAGS_poster.pdf (2018)

Garin, A., Tauzin, G.: A topological "reading" lesson: classification of MNIST using TDA. In: 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA), pp. 1551–1556. IEEE (2019)

Garside, K., Henderson, R., Makarenko, I., Masoller, C.: Topological data analysis of high resolution diabetic retinopathy images. PLoS ONE **14**(5), 0217413 (2019)

Giunti, B., Nolan, J.S., Otter, N., Waas, L.: Amplitudes in persistence theory. J. Pure Appl. Algebra **228**, 107770 (2024)

Golub, G.H., Hoffman, A., Stewart, G.W.: A generalization of the Eckart-Young-Mirsky matrix approximation theorem. Linear Algebra Appl. **88**, 317–327 (1987)

Hiraoka, Y., Nakamura, T., Hirata, A., Escolar, E.G., Matsue, K., Nishiura, Y.: Hierarchical structures of amorphous solids characterized by persistent homology. Proc. Natl. Acad. Sci. **113**(26), 7035–7040 (2016)

Hofer, C., Kwitt, R., Niethammer, M., Uhl, A.: Deep learning with topological signatures. Advances Neural Inf. Process. Syst. **30** (2017)

Kerber, M., Morozov, D., Nigmetov, A.: Geometry helps to compare persistence diagrams. ACM J. Exp. Algorithmics **22**(1.4), 1–20 (2017)

Kusano, G., Fukumizu, K., Hiraoka, Y.: Kernel method for persistence diagrams via kernel embedding and weight factor. J. Mach. Learn. Res. **18**(1), 6947–6987 (2017)

Lesnick, M.: The theory of the interleaving distance on multidimensional persistence modules. Found. Comput. Math. **15**(3), 613–650 (2015)

Miller, E.: Essential graded algebra over polynomial rings with real exponents. arXiv preprint arXiv:2008.03819v1 (2020)

Otter, N., Porter, M.A., Tillmann, U., Grindrod, P., Harrington, H.A.: A roadmap for the computation of persistent homology. EPJ Data Sci. **6**, 1–38 (2017)

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., et al.: Pytorch: An imperative style, high-performance deep learning library. Adv. Neural Inf. Process. Syst. **32** (2019)

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al.: Scikit-learn: Machine learning in python. J. Mach. Learn. Res. **12**, 2825–2830 (2011)

Qaiser, T., Tsang, Y.-W., Taniyama, D., Sakamoto, N., Nakane, K., Epstein, D., Rajpoot, N.: Fast and accurate tumor segmentation of histology images using persistent homology and deep convolutional features. Med. Image Anal. **55**, 1–14 (2019)

Reinauer, R., Caorsi, M., Berkouk, N.: Persformer: A transformer architecture for topological machine learning. arXiv preprint arXiv:2112.15210v2 (2021)

Reininghaus, J., Huber, S., Bauer, U., Kwitt, R.: A stable multi-scale kernel for topological machine learning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4741–4748 (2015)

Rotman, J.: An Introduction to Homological Algebra. Universitext, 2nd edn. Springer, New York (2009)

Scolamiero, M., Chachólski, W., Lundman, A., Ramanujam, R., Öberg, S.: Multidimensional persistence and noise. Found. Comput. Math. **17**(6), 1367–1406 (2017)

Skraba, P., Turner, K.: Wasserstein stability for persistence diagrams. arXiv preprint arXiv:2006.16824v5 (2020)

Steele, J.M.: The Cauchy-Schwarz Master Class: An Introduction to the Art of Mathematical Inequalities. Cambridge University Press, Cambridge (2004)

Stolz, B.J., Harrington, H.A., Porter, M.A.: Persistent homology of time-dependent functional networks constructed from coupled time series. Chaos Interdiscip. J. Nonlinear Sci. **27**(4), 047410 (2017)

Turkeš, R., Nys, J., Verdonck, T., Latré, S.: Noise robustness of persistent homology on greyscale images, across filtrations and signatures. PLoS ONE **16**(9), 0257215 (2021)

Vince, A.: A rearrangement inequality and the permutahedron. Am. Math. Mon. **97**(4), 319–323 (1990)

Zhao, Q., Wang, Y.: Learning metrics for persistence-based summaries and applications for graph classification. Adv. Neural Inf. Process. Syst. **32** (2019)

Zomorodian, A., Carlsson, G.: Computing persistent homology. Discrete Comput. Geom. **33**, 249–274 (2005)