A COMPARISON FRAMEWORK FOR INTERLEAVED PERSISTENCE MODULES AND APPLICATIONS OF PERSISTENT HOMOLOGY TO PROBLEMS IN FLUID DYNAMICS

BY RACHEL LEVANGER

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Mathematics

Written under the direction of

Konstantin Mischaikow

and approved by

_____

_____

_____

_____

New Brunswick, New Jersey

May, 2017

ABSTRACT OF THE DISSERTATION

A comparison framework for interleaved persistence modules and applications of persistent homology to problems in fluid dynamics

by RACHEL LEVANGER

Dissertation Director: Konstantin Mischaikow

We prove an algebraic stability theorem for interleaved persistence modules that is more general than any formulations currently in the literature. We show how this generalization leads to a framework that may be used to compare persistence modules locally, enabling the computation of non-uniform error bounds for persistence diagrams. We give several examples of how to use this comparison framework, and also address an open problem on non-uniform sublevel set filtrations.

We also give two applications of persistent homology to problems in fluid dynamics. Our first application examines the structure of the dynamics of a time-evolving system on a two-dimensional domain, where we give examples for studying fixed points and periodic orbits. Our second application uses persistent homology in conjunction with techniques in computer vision to study pattern defects in the spiral defect chaos regime of Rayleigh-Bénard convection.

Preface

As with many modern scientific endeavors, much of the work contained in this dissertation is the cumulative product of many minds and could not have been accomplished by a single individual working alone. I provide here comments about the nature of the many collaborations that gave rise to the work presented in this dissertation.

Part I is work that I coauthored with Shaun Harker, Miroslav Kramár, and Konstantin Mischaikow, and is currently in preparation to be submitted. The ideas for the direction of this work were born out of many discussions with my coauthors and others, most notably Vidit Nanda, Mike Lesnick, Uli Bauer, Amit Patel, and Chuck Weibel.

Chapter 5 was previously published in *Physica-D* [1], and Chapter 6 offers a continuation of this work with the same group of coauthors in addition to Jacek Cyranka, Shaun Harker, and Logan Kageorge. The numerical simulations for both chapters were computed by my coauthors at Georgia Tech and Virginia Tech. The subsampling software used in Chapter 5 was built by Shaun Harker specifically to handle the large amount of data generated for this study. The interface for processing the persistence diagrams of images with PHAT, used extensively in the work of Chapter 6, was built by Jacek Cyranka and Shaun Harker, and some of the numerical analysis software used in this same chapter was also built by Shaun.

Acknowledgements

I sincerely thank Konstantin Mischaikow for taking me on as a graduate student and giving me informative projects as I transitioned from a naive and bright-eyed first semester graduate student to a budding mathematical researcher. Through your example, you have taught me how to see and learn and develop mathematics and its many applications, and I am forever thankful to have grown so much under your tutelage.

I must also give thanks to my unofficial academic advisor, Chuck Weibel. I still remember the first conversation we had at my visiting day at Rutgers, and I'm thankful for all of the conversations we've had since–spontaneous or otherwise. Through your patience, I have been able to get a glimpse of how to think like an algebraist, which has served me well in thinking about my chosen subject of study.

I am also indebted to my many collaborators and peers whom I have had the privilege of working with and learning from over the years. Miro Kramár, Mike Schatz, Mark Paul, Shaun Harker, Jacek Cyranka, Jeff Tithof, Takashi Ishihara, Vidit Nanda, and Mike Lesnick, it has been a pleasure and an honor to be a part of your communities.

Many thanks to other faculty in the mathematics department at Rutgers who have shaped the way I think, especially Steve Ferry, Jerrold Tunnell, and Zheng-Chao Han. I also thank other students in the department who have been there for me over the years. Justin Bush, Brian Garnett, Ed Chien, Glen Wilson, and Doug Schultz, each of you has helped me navigate the turbulent waters of graduate school.

Finally, to my immediate family members, mom, dad, Kerrie, and Tommy, you all make me proud to say I'm a Levanger, and I've been so grateful to have had all of you to lean on throughout this process. Jonathan, thank you for being my rock throughout my graduate studies. Ted, I might never have started this journey if it weren't for you. Ben, thank you for being there.

Dedication

For Kerrie.

Table of Contents

Chapter 1

Introduction

Vast amounts of high-dimensional data inundate our scientific and digital communities. To extract insight from all of this data, we must develop tools that enable us to efficiently and accurately examine the structure and nature of high-dimensional datasets. Statistical regression methods use pre-existing shapes or forms (e.g. linear, polynomial, or logistic curves) to study the structure of a dataset by testing for best fit. Many topological methods, in contrast, enable one to study the shape of a dataset without presupposing some sort of structure in advance. Techniques from these latter methods form the topic of this dissertation.

Persistent homology, a foundational tool from the field of topological data analysis, provides a method by which one can study the shape and structure of high-dimensional data on multiple spatial scales simultaneously. It provides a compact summary of the multi-scale topology of a dataset in the form of a collection of persistence diagrams. Each persistence diagram is a collection of points on the Euclidean plane, the arrangement of which provides a topological signature of the underlying dataset. Figure 1.1 shows a finite point cloud and its associated (nontrivial) persistence diagrams. The group of black persistence points, corresponding to the dimension zero homological features, encode the connectivity of the dataset. The blue persistence points, corresponding to the dimension one homological features, encode the "holes" or loops at intermediate scales of the dataset. We direct the reader to [2, 3, 4, 5] for an introduction to this rich topic. In Chapter 5, we give an intuition for how persistence diagrams encode the low-dimensional topological properties of the following types of datasets: finite point clouds and scalar fields on two-dimensional domains.

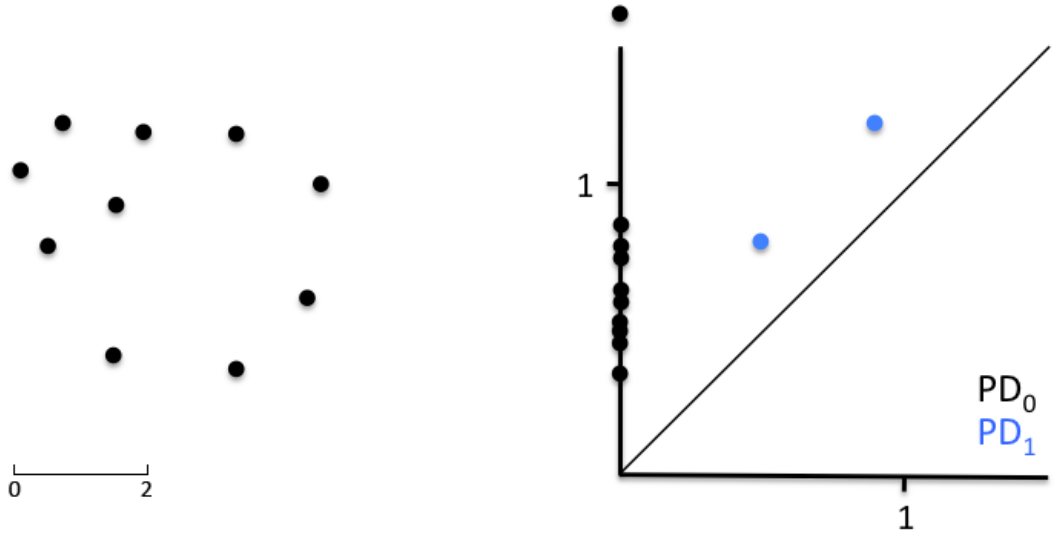


Figure 1.1: A finite point cloud (left) and its associated persistence diagrams (right). The black points in the persistence diagram encode the connectivity of the dataset in a similar way to a single-linkage hierarchical clustering algorithm. The blue points in the persistence diagram encode the two loops that are present in the dataset at intermediate scales. That is, if each of the points in the point cloud is inflated by a ball of some radius r, for large enough r, the inflated sets would contain loop structures.

1.1 Theoretical Results

In Part I of this dissertation, we give a new theoretical result which enables the statement of local error bounds for the placement of the persistence points on the Euclidean plane. That is, if an approximation must be made in order to compute the persistence diagrams of a dataset, e.g. for computational reasons, we provide a way to precisely state bounds on the errors that may have been introduced by the approximation. Until this work, statements about errors in the position of the persistence points have been written down in the form of global, uniform error bounds (in some cases after passing to a log scale) [6, 7, 8, 9, 10]. While it is possible to use uniform error bounds to make local statements by restricting the portion of the persistence plane in question, our framework gives a natural way to describe non-uniform error bounds for the entire persistence plane at once. The framework also provides a method by which one can easily and more accurately compare the errors introduced by different approximations.

These results, presented in Chapter 3, rest on the theory of persistence modules [5]

and their interleavings, which were first discussed by Chazal, et al. in [7]. We prove an algebraic stability theorem for persistence modules that extends the one established by Bauer and Lesnick in [11] by using the generalized interleavings of Bubenik, et al. in [12]. We provide enough background so that this chapter is self-contained from the perspective of persistence module theory. In Chapter 4, we give a number of examples to illustrate how our results may be used. In particular, we show how approximations arising from discretization or subsampling affect the corresponding persistence diagrams, provide a summary of error bounds for a number of approximations to Vietoris-Rips and Čech complex filtrations, and also address an open problem posed by Bendich, et al. in [13].

1.2 Applications

Our second contribution in this dissertation opens the door to a suite of new methods for studying problems in fluid dynamics via tools in computational topology. Chapters 5 and 6 showcase the versatility of persistent homology by using it as a tool for dimension reduction, symmetry reduction, shape detection, and pattern recognition. We do not assume that the reader has an extensive background in fluid dynamics, and in fact our approach highlights a collection of methods that do not require a mathematical knowledge of the underlying equations governing the dynamics of the systems studied, making them well-suited to experimentally-collected datasets. We give a brief description of each of the applications presented.

1.2.1 Dimension Reduction

Let $f : D \to \mathbb{R}$ be a Morse function on a compact subset $D \subseteq \mathbb{R}^n$. The function f is precisely described by defining its values on all of D, likely an infinite set. However, the total number of persistence points in the persistence diagrams corresponding to f is bounded above by $n \times c$, where c is the number of critical points of f [4]. For the case when $n = 2$, we approximate f by an $m_1 \times m_2$-pixel digital image g. It is possible to define the notion of a critical point in this setting, and the upper bound for the total

number of persistence points in the persistence diagrams for g is still bounded above by $n\tilde{c}$, where $\tilde{c}$ is the number of critical points of g (in general this requires that the values of each pixel or voxel in g are unique, which is always possible by performing a small perturbation of the pixel values). Thus, in the case where a digital image approximates a smooth function, as in the work contained in this dissertation, the total number of persistence points in the collection of persistence diagrams corresponding to g is significantly smaller than the number of pixels in the underlying image. We use this aspect of persistent homology extensively throughout Chapter 5, where we work in the space of persistence diagrams rather than directly with the digital images themselves.

1.2.2 Symmetry Reduction

Fluid flow simulations are often carried out under periodic boundary conditions. This is primarily due to the complications introduced when attempting to accurately model the boundary effects of the fluid. Alternatively, one may be interested in studying patterns and dynamics that arise in the bulk of the fluid flow, far away from the boundary, so the boundary effects may be assumed to be negligible and not taken into account.

The addition of periodic boundary conditions on the governing equations for the fluid flow simulation can introduce a number of symmetries into the system. That is, solutions may lie on group orbits, where each solution in a group orbit is related to many others by some symmetry, such as a translation or rotation. Often when studying dynamical structures, it is desirable to treat solutions lying in a single group orbit as equivalent, and thus to work with equivalence classes of solutions instead of the actual solutions.

Classical approaches to carry out this identification involve a variety of methods. Typically, the set of symmetries is explicitly searched in order to make the identifications. For small systems, this approach is adequate. However, when dealing with thousands of solutions sampled from a single flow trajectory, the task becomes computationally infeasible. In contrast, homology is invariant under these types of coordinate transformations, and so persistent homology is a natural tool for making these dynamical identifications. However, it is also possible that persistent homology collapses too

many symmetries, as the following remark serves to illustrate.

Remark 1.1. Let $f : D \to \mathbb{R}$ be a scalar field on a topological domain D, and let $h : D \to D$ be a homeomorphism. Then $g := f \circ h$ is also a scalar field on D, and the persistence diagrams of f are equivalent to the persistence diagrams of g.

In light of this remark, it is necessary to determine if persistence diagrams are a strong enough invariant to give different signatures for patterns that are dynamically distinct. While this is not true in general, in Chapter 5 we establish that persistent homology is comparable in practice to Fourier methods for symmetry reduction for a small dataset. We sample a chaotic trajectory for sixty-seven initial conditions for a Newton solver, which is then used to find the nearby fixed point. These solutions are then identified by both Fourier methods and using persistence diagrams. Our results show that both techniques determine the same number of equivalence classes of representative dynamical patterns.

1.2.3 Shape Detection

One of the canonical applications of persistent homology is to the study of the intrinsic structure of finite point clouds at multiple spatial scales, as in Figure 1.1 (see also [2, 14]). In Chapter 5, we first map a collection of scalar fields resulting from simulations of fluid flows on a 2D domain to the space of persistence diagrams to perform a data reduction step and also to quotient out symmetries inherent in the system as a result of the periodic boundary conditions (see previous two sections).

The space of persistence diagrams is a metric space (endowed with a variety of metrics), and the function mapping a scalar field to its corresponding persistence diagrams is Lipschitz continuous with unit constant under the sup-norm metric on the scalar fields and either the Bottleneck or Wasserstein metrics on the space of persistence diagrams. Thus, viewing a collection of fluid flow simulations as samples along a continuous trajectory in solution space, we can view their resulting persistence diagrams as a point cloud that lie along a continuous trajectory in the space of persistence diagrams. In the case of truly periodic dynamics, this trajectory is contained in the graph of a continuous function that embeds the circle S^1 into the space of persistence diagrams.

Considering that we are working with numerical approximations of the scalar fields and not true solutions, we can view this point cloud as a noisy sample of a circular structure embedded in the space of persistence diagrams.

In Chapter 5, we show how the persistence diagrams of a point cloud obtained in such a manner from Kolmogorov flow reflects this circular structure: the persistence points with finite coordinates in the dimension zero persistence diagram are close to the diagonal, indicating that the trajectory was sampled densely; and the dimension one persistence diagram consists of a single point far from the diagonal in addition to some small-scale noise, or points close to the diagonal.

We take this study one step further in the study of an almost-periodic orbit taken from simulated Rayleigh-Bénard convection. In this case, we address issues that stem from multiple time scales in the dynamics and the large number of sample points that result from sampling densely enough to resolve the dynamical structure of the trajectory. Our theoretical results in Chapter 4 (see Section 4.1.4) provide us with a precise statement on the errors introduced by down-sampling the point cloud in order to perform the persistence diagram computations. While the dynamical structure of a circle is recovered, a second prominent point in the dimension one persistence diagram indicates that the trajectory might be folded or twisted, and thus additional structure is captured by mapping the trajectory to the space of persistence diagrams.

1.2.4 Pattern Recognition

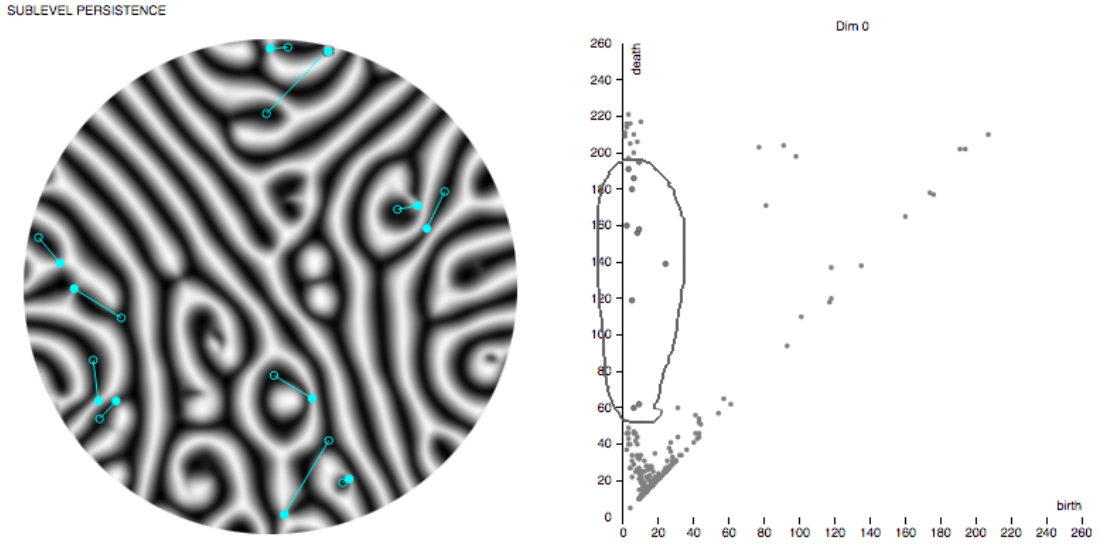
Finally, in Chapter 6 we use the Morse-theoretic underpinnings of persistent homology as a tool in two pattern recognition problems. Specifically, we use the critical point pairings that result from the computation of a persistence diagram as a way to localize certain pattern formations in the mid-plane temperature field of Rayleigh-Bènard convection at high Rayleigh number.

Figure 1.2 shows a sample temperature fields and its corresponding sublevel set persistence diagrams for dimension (a) zero and (b) one. The author of this dissertation has designed interactive software [15] that enables one to visualize the critical cell pairings corresponding to the persistence points; once encircled on the persistence plane,

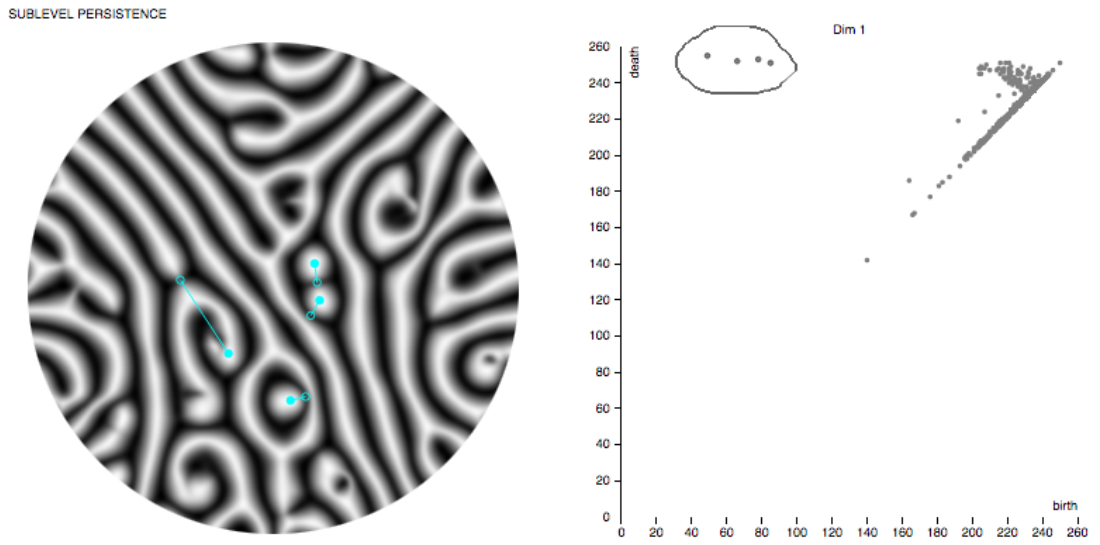
the corresponding critical cell pairings are overlaid (in cyan) on the temperature field as shown in the figure.

As a first application of persistent homology to pattern analysis, we use critical cell pairings from the dimension one persistence diagrams to differentiate targets from spirals (see Section 6.3.2). This enables an automated detection of the six canonical defect types that are most commonly studied in the existing literature.

As a second application to pattern analysis, we use persistent homology to select certain critical cells corresponding to saddle points or local extrema in the image. That is, a single image may contain many critical points, but only certain critical points will be important for distinguishing the global arrangement of the pattern; persistent homology offers a simple method by which to select the distinguishing critical cells. In this second application, we combine the critical cells selected through persistence with the more classical *topological defects* to give a more nuanced description of a local defect and for detecting recurring patterns. We then demonstrate how our methods combine with standard techniques from object recognition in computer vision through two examples.



(a)



(b)

Figure 1.2: Sample mid-plane temperature field of Rayleigh-Bènard convection with selected critical cell pairings overlaid in cyan (the birth cell is marked by an open circle, and the death cell is marked by a solid circle, joined by a straight line). (a) Sublevel dimension zero persistence diagram with persistence points near the death axis encircled. The corresponding persistence pairings indicate that the death critical cell "points" to roll instabilities. (b) Sublevel dimension one persistence diagram with persistence points far from the diagonal encircled. The corresponding persistence pairings indicate that these correspond to rolls that are disjoint from the boundary of the image, with the death critical cell "pointing" to the maximum of each roll.

PART I:

A Comparison Framework for Interleaved Persistence Modules

Persistent homology [5] is a key element in the rapidly-developing field of topological data analysis, where it is used both as a means of identifying geometric structures associated with data and as a data reduction tool. Any work with data involves approximations that arise from finite sampling, limits to measurement, and experimental or numerical errors. Part I of this dissertation focuses on obtaining rigorous bounds on the variations in persistence diagrams arising from these approximations.

To motivate this work we begin with the observation that many problems in data analysis can be rephrased as a problem concerned with the analysis of the geometry induced by a scalar function $f\colon X \to \mathbb{R}$. For example, consider a finite point cloud $\mathcal{X}$ in a metric space X with metric ρ. Single-linkage hierarchical clustering problems are naturally associated with the function $f_{\mathcal{X}}\colon X \to [0,\infty)$ given by

$$f_{\mathcal{X}}(x) := \rho(x,\mathcal{X}) = \min\left\{\rho(x,\xi) : \xi \in \mathcal{X}\right\},$$

where clusters are derived from the connected components of the sublevel set

$$\mathsf{C}(f_{\mathcal{X}},t) := \{x \in X : f_{\mathcal{X}}(x) \leq t\}$$

for a particular choice of $t \in [0,\infty)$. The collection $\{\mathsf{C}(f_{\mathcal{X}},t)\}_{t\in\mathbb{R}}$ is called the *sublevel set filtration* of X induced by f. Superlevel sets and superlevel set filtrations are defined similarly by considering the sets $\{x \in X : t \leq f_{\mathcal{X}}(x)\}$.

Alternatively, consider a scalar field

$$u\colon \Omega \to \mathbb{R}$$

arising at a particular point in time from a nonlinear physical model, e.g. the vorticity or temperature field of a fluid, chemical density in a reaction diffusion system, forces between particles in a granular system, etc. Patterns produced by these systems are often associated with sublevel or superlevel sets of u.

These examples are meant to motivate our interest in studying the geometry of the sets $\mathsf{C}(f,t)$. Homology provides a coarse but computable representation of this

geometry. For each function f, $t \in \mathbb{R}$, and $n \in \mathbb{N}$, there is an assigned vector space

$$M(f)_t = H_n(\mathsf{C}(f,t), \mathsf{k}),$$

where k is a field. Because each $t \leq s$ implies $\mathsf{C}(f,t) \subseteq \mathsf{C}(f,s)$, the inclusion maps induce the following linear maps at the level of homology:

$$\varphi_{M(f)}(t,s) \colon M(f)_t \to M(f)_s.$$

This homological information can be abstracted into a *persistence module*, which consists of a collection of vector spaces indexed by the real numbers, $\{V_t\}_{t\in\mathbb{R}}$, and linear maps $\{\varphi_V(s,t) : V_s \to V_t\}_{s\leq t\in\mathbb{R}}$ satisfying the following conditions:

(i) $\varphi_V(t,t) = \mathrm{id}_{V_t}$ for every $t \in \mathbb{R}$, and

(ii) $\varphi_V(s,t) \circ \varphi_V(r,s) = \varphi_V(r,t)$ for every $r \leq s \leq t$ in $\mathbb{R}$.

We write (V, φ_V) to denote the collection of vector spaces and compatible linear maps, and will sometimes just write V for the full persistence module when the maps are clear. We say that V is a *pointwise finite dimensional (PFD)* persistence module when every V_t is finite-dimensional.

As is described in greater detail in Chapter 4, a PFD persistence module gives rise to a *persistence diagram*, which is a multiset of points in the extended plane, i.e. a multiset with some base set consisting of a collection of points in $\overline{\mathbb{R}}^2$, where $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$. Given a PFD persistence module (V, φ_V), we denote its associated persistence diagram by $\mathsf{PD}(V)$. When an underlying object, such as a scalar field $f : X \to \mathbb{R}$, is used to induce a persistence module (through a sublevel set filtration on X, for example), we will write $\mathsf{PD}(f)$ as the persistence diagram associated to the induced persistence module; in these cases, the persistence module will be made clear by context.

Returning to our examples, observe that we have outlined a procedure by which the sublevel sets of a scalar field u produce a persistence diagram $\mathsf{PD}(u)$. If, as suggested above, u is generated by a physical system, then there is little chance of computing $\mathsf{PD}(u)$ exactly. More specifically, if u is meant to be a solution to a partial differential

equation at a given time point, then one may expect that an approximation $\tilde{u}$ of u has been obtained via a numerical method. Alternatively, for an actual physical system, $\tilde{u}$ may represent data obtained from an experiment.

Recent computational developments have led to the routine computation of $\mathsf{PD}(\tilde{u})$ in a wide variety of settings. Thus, the natural question is how are $\mathsf{PD}(u)$ and $\mathsf{PD}(\tilde{u})$ related? A fundamental result [6] in the theory of persistent homology is that a variety of metrics can be imposed on the space of persistence diagrams such that the function PD is continuous with respect to L^∞ changes in u. Recent developments by Bauer and Lesnick [11] allow for comparisons of persistence modules through a matching of the associated persistence points. In Chapter 3 we extend these recent results, and then use this extension in Chapter 4 to provide explicit bounds on the location of unmatched points and on the differences between the matched points in the persistence diagrams that are applicable to a wide variety of naturally arising problems.

To be more explicit, consider the following example. Let $(M(f), \varphi_{M(f)})$ be the persistence module associated with the sublevel sets of the function $f\colon X \to \mathbb{R}$, but assume that we can only sample the sublevel sets of f at the integers $\mathbb{Z} \subset \mathbb{R}$. As explained in Chapter 4, this sampling gives rise to a persistence module $(M(f)', \varphi_{M(f)'})$. Figure 1.3 indicates the type of result that we obtain: the location of a computed persistence point of $\mathsf{PD}(M(f)')$ is shown along with the region of the persistence diagram within which the matched persistence point of $\mathsf{PD}(M(f))$ must lie, and the possible locations of unmatched persistence points of $\mathsf{PD}(M(f))$ and $\mathsf{PD}(M(f)')$.

An outline of the chapter is as follows. In Chapter 2 we review the essential concepts associated with persistence modules required for our results. This section defines the notions of persistence modules and their morphisms, interleavings, induced matchings, and the algebraic stability theorem, which form the starting point for our work.

Our theoretical results are presented in Chapter 3. We begin by extending the Induced Matching Theorem of Bauer and Lesnick [11]. Using a combination of ideas from the theory of generalized interleavings of Bubenik, et al. [12] with the Induced Matching Theorem, we obtain the Non-uniform Induced Matching Theorem (Theorem 3.10), which is a more general statement about how the sizes of the kernel and cokernel of a

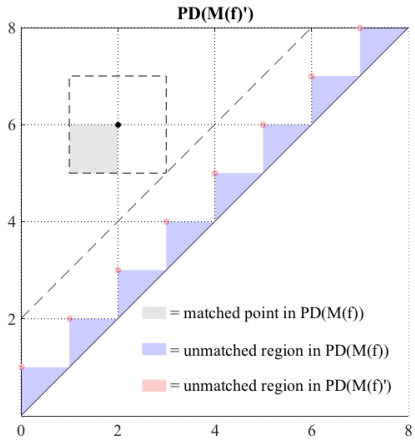


Figure 1.3: A schematic diagram illustrating the quality of the approximation given in Section 4.1.2. To illustrate the matching guaranteed by Theorem 3.13, suppose that $\mathsf{PD}(M(f)')$ contains a single off-diagonal point at $(2, 6)$. The gray region surrounding this point illustrates the possible location of the matched point in the "true" persistence diagram $\mathsf{PD}(M(f))$, and any unmatched points in $\mathsf{PD}(M(f))$ fall in the blue region. If any persistence points in $\mathsf{PD}(M(f)')$ had fallen in the pink region, then it's possible they may have been introduced as a computational artifact, and may not have corresponding matched points in $\mathsf{PD}(M(f))$. The dashed lines illustrate the error bounds given by the bottleneck distance between the approximation persistence diagram and the underlying "true" diagram.

persistence module morphism relate to the quality of the induced matching of [11].

Next we show that generalized interleavings can be combined with the Non-uniform Induced Matching Theorem to produce an Algebraic Stability Theorem for Generalized Interleavings (Theorem 3.13 and easy-to-state Corollary 3.14). Our proof closely follows Bauer and Lesnick's proof of the Algebraic Stability Theorem [11], though we use the Non-uniform Induced Matching Theorem to push the bounds given by generalized interleavings through for a more precise statement. We believe this result (and especially the statement of Corollary 4.2) clarifies how the monotone functions comprising the translation pair of a generalized interleaving relate to stability in the space of persistence diagrams.

Chapter 4 provides applications of the above-mentioned results. The most classical stability result is given using the bottleneck distance [6], which is akin to measuring the largest difference between two diagrams, regardless of where the persistence points lie on the two-dimensional persistence plane. In contrast, Theorem 3.13 gives a local

notion of distance specified by the location of the points in each of the diagrams. In Chapter 4 we show how Theorem 3.13 and Corollary 4.2 make it possible to describe the errors introduced by computing a persistence diagram from an approximation.

It can be argued that for applications, the most difficult task is the construction of the interleaving between the two persistence modules. However, as we hope the examples of Chapter 4 illustrate, once the interleavings are determined, working with our framework is straightforward. With this in mind, we include Table 4.1 in Section 4.1.5, providing an easily-referenced list of translation maps of generalized interleavings for common approximations to Vietoris-Rips and Čech filtrations.

In Section 4.2, we consider the problem of manifold identification. In particular, Proposition 4.26 provides, with high probability, information about the location of persistence points if one randomly samples sufficiently many points on a manifold and uses the distance function and Vietoris-Rips filtration to compute the persistence diagram. Finally, in Section 4.3, we show how this theory can be used to address, at least in part, a question raised by Bendich, et al. in [13] concerning variations in a persistence diagram under a non-uniform sublevel set filtration.

Chapter 2

Preliminaries

In this chapter, we summarize background material and establish notation for the work we present in Part I of this dissertation. Section 2.1 provides a short overview of persistence modules and their morphisms. Section 2.2 addresses interval decompositions of persistence modules, barcodes, and decorated endpoints. Section 2.3 introduces the the notion of an interleaving between two persistence modules and establishes some properties of monotone functions. Section 2.5 recalls the results of Bauer and Lesnick [11] concerning the matchings between barcodes induced by morphisms of persistence modules and the traditional algebraic stability theorem, both of which we generalize in Chapter 3.

2.1 Persistence Modules and their Morphisms

This section provides basic facts about persistence modules, presented with minimal reference to the language of category theory. For alternative treatments, see [5, 11, 12, 16].

Definition 2.1. A *persistence module* is an $\mathbb{R}$-indexed family of vector spaces $\{V_t\}_{t\in\mathbb{R}}$, each over a field k, together with a family of linear maps $\{\varphi_V(s,t) : V_s \to V_t\}_{s\leq t\in\mathbb{R}}$, called *transition maps*, that satisfy the following conditions:

(i) $\varphi_V(t,t) = \mathrm{id}_{V_t}$ for every $t \in \mathbb{R}$, where id_{V_t} is the identity map on V_t, and

(ii) $\varphi_V(s,t) \circ \varphi_V(r,s) = \varphi_V(r,t)$ for every $r \leq s \leq t$ in $\mathbb{R}$.

We write (V, φ_V) to denote the collection of vector spaces and compatible linear maps, and will sometimes just write V for the full persistence module when the maps are

clear. We say that V is a *pointwise finite dimensional (PFD)* persistence module when every V_t is finite-dimensional.

Definition 2.2. Let (V, φ_V) and (W, φ_W) be persistence modules. A *persistence module morphism* $\phi : V \to W$ is a collection of linear maps $\{\phi_t : V_t \to W_t\}_{t\in\mathbb{R}}$ such that the following diagram commutes for all $s, t \in \mathbb{R}$ with $s \leq t$.

$$\begin{array}{ccc} V_s & \xrightarrow{\varphi_V(s,t)} & V_t \\ \downarrow{\scriptstyle \phi_s} & & \downarrow{\scriptstyle \phi_t} \\ W_s & \xrightarrow{\varphi_W(s,t)} & W_t \end{array}$$

If ϕ_t is injective (surjective) for every $t \in \mathbb{R}$, then we say that ϕ is a monomorphism (epimorphism). A persistence module morphism that is both a monomorphism and an epimorphism is an isomorphism.

To readers familiar with category theory, persistence modules and their morphisms form an abelian category, a fact that we essentially establish in the remainder of this section. We also formulate analogs for several concepts from the category of vector spaces and linear maps, such as the notion of a persistence submodule, intersections, and direct sums. We begin with the definition of a persistence submodule.

Definition 2.3. Let (V, φ_V) be a persistence module and let $\{W_t\}_{t\in\mathbb{R}}$ be a family of vector spaces such that W_t is a subspace of V_t for every $t \in \mathbb{R}$. For $s, t \in \mathbb{R}$ with $s \leq t$, the *domain restriction* of the map $\varphi_V(s, t) : V_s \to V_t$ to the subspace W_s is expressed by the notation $\varphi_V(s, t)|_{W_s} : W_s \to V_t$.

Definition 2.4. Let (V, φ_V) be a persistence module. We say that (W, φ_W) is a *persistence submodule* of the persistence module V if W is a persistence module, $W_t \subseteq V_t$ for every $t \in \mathbb{R}$, and $\varphi_W(s, t) = \varphi_V(s, t)|_{W_s}$ for every $s, t \in \mathbb{R}$ with $s \leq t$.

It is sometimes cumbersome to establish that a persistence submodule is, indeed, a persistence module. The following lemma gives a more direct route.

Lemma 2.5. *Let (V, φ_V) be a persistence module and let $\{W_t\}_{t\in\mathbb{R}}$ be a family of vector spaces such that $W_t \subseteq V_t$ for every $t \in \mathbb{R}$. Suppose that* $\operatorname{im} \varphi_V(s, t)|_{W_s} \subseteq W_t$ *for all*

$s, t \in \mathbb{R}$ with $s \leq t$ and define $\varphi_W(s,t) := \varphi_V(s,t)|_{W_s}$. Then (W, φ_W) is a persistence submodule of (V, φ_V).

Proof. We must show that $\varphi_W(s,t) : W_s \to W_t$ for every $s, t \in \mathbb{R}$ with $s \leq t$, and that the maps φ_W satisfy properties (i) and (ii) in Definition 2.1.

Let $s \leq t$. By hypothesis, we are given that im $\varphi_V(s,t)|_{W_s} \subseteq W_t$. Thus, $\varphi_W(s,t) = \varphi_V(s,t)|_{W_s} : W_s \to W_t$, so the transition maps are defined. It remains to verify properties (i) and (ii) in Definition 2.1. We first verify property (i). Let $t \in \mathbb{R}$. Then $\varphi_W(t,t) = \varphi_V(t,t)|_{W_t} = \mathrm{id}_{W_t}$ and property (i) is verified. To verify property (ii), let $r \leq s \leq t$ in $\mathbb{R}$. Since im $\varphi_V(r,s)|_{W_r} \subseteq W_s$,

$$[\varphi_V(s,t) \circ \varphi_V(r,s)]|_{W_r} = \varphi_V(s,t)|_{W_s} \circ \varphi_V(r,s)|_{W_r}.$$

Thus, we have

$$\begin{aligned} \varphi_W(r,t) &= \varphi_V(r,t)|_{W_r} \\ &= [\varphi_V(s,t) \circ \varphi_V(r,s)]|_{W_r} \\ &= \varphi_V(s,t)|_{W_s} \circ \varphi_V(r,s)|_{W_r} \\ &= \varphi_W(s,t) \circ \varphi_W(r,s). \end{aligned}$$

Hence, the maps φ_W satisfy property (ii). □

With the notion of a persistence submodule established, it is possible to construct the notion of an intersection of two persistence modules.

Definition 2.6. Suppose that (V, φ_V) and (W, φ_W) are both persistence submodules of a persistence module (U, φ_U). The *intersection persistence submodule* $(V \cap W, \varphi_{V \cap W})$ of (U, φ_U) is defined by the vector spaces

$$(V \cap W)_t := V_t \cap W_t$$

for all $t \in \mathbb{R}$, and transition maps $\varphi_{V \cap W}(s,t) := \varphi_U(s,t)|_{(V \cap W)_s}$ defined via restriction for every $s, t \in \mathbb{R}$ with $s \leq t$.

Proposition 2.7. *Let (V, φ_V) and (W, φ_W) be persistence submodules of a persistence module (U, φ_U). The persistence module $(V \cap W, \varphi_{V\cap W})$ is a persistence submodule of (U, φ_U).*

Proof. We will use Lemma 2.5 to show that $(V \cap W, \varphi_{V\cap W})$ is a persistence submodule of (U, φ_U). Since the intersection of two vector subspaces is again a subspace, then

$$(V \cap W)_t = V_t \cap W_t \subseteq U_t$$

for every $t \in \mathbb{R}$, so the subspace hypothesis is satisfied.

The transition maps of $(V \cap W, \varphi_{V\cap W})$ are defined by $\varphi_{V\cap W}(s,t) := \varphi_U(s,t)|_{(V\cap W)_s}$ for every $s, t \in \mathbb{R}$ with $s \leq t$, so it remains to show that $\text{im } \varphi_U(s,t)|_{(V\cap W)_s} \subseteq (V \cap W)_t$ for every $s \leq t$. Since (V, φ_V) and (W, φ_W) are each persistence submodules of (U, φ_U), then $\text{im } \varphi_U(s,t)|_{V_s} = \text{im } \varphi_V(s,t) \subseteq V_t$ and $\text{im } \varphi_U(s,t)|_{W_s} = \text{im } \varphi_W(s,t) \subseteq W_t$ for every $s \leq t$. It follows that

$$\text{im } \varphi_U(s,t)|_{(V\cap W)_s} = \text{im } \varphi_U(s,t)|_{V_s\cap W_s} \subseteq V_t \cap W_t = (V \cap W)_t$$

for every $s, t \in \mathbb{R}$ with $s \leq t$. By Lemma 2.5, $(V \cap W, \varphi_{V\cap W})$ is a persistence submodule of (U, φ_U). □

A fundamental construction in the theory of persistence modules is a direct sum.

Definition 2.8. Let (V, φ_V) and (W, φ_W) be persistence modules. The direct sum $V \oplus W$ is formed by taking

$$(V \oplus W)_t := V_t \oplus W_t$$

for all $t \in \mathbb{R}$ and

$$\varphi_{V\oplus W}(s,t) := \varphi_V(s,t) \oplus \varphi_W(s,t)$$

for every $s, t \in \mathbb{R}$ with $s \leq t$.

It is straightforward to show that the direct sum of two persistence modules gives another persistence module, and we omit the proof. This notion can also be generalized

to direct sums with an arbitrary number of summands. That is, if $\mathcal{A}$ is an index set and $\{(V_a, \varphi_{V_a})\}_{a\in\mathcal{A}}$ is a family of persistence modules, the persistence module

$$\left(\bigoplus_{a\in\mathcal{A}} V_a, \varphi_{\oplus V_a}\right)$$

is defined by taking

$$\left(\bigoplus_{a\in\mathcal{A}} V_a\right)_t := \bigoplus_{a\in\mathcal{A}} (V_a)_t$$

for every $t \in \mathbb{R}$, and

$$\varphi_{\oplus V_a}(s,t) := \bigoplus_{a\in\mathcal{A}} \varphi_{V_a}(s,t)$$

whenever $s, t \in \mathbb{R}$ with $s \leq t$. As we will see in the next section, this construction is at the heart of persistence module theory and the characterization of the structure of PFD persistence modules.

The following definitions, while rather simple, give the notions of an inclusion and quotient of persistence modules.

Definition 2.9. Let (W, φ_W) be a persistence submodule of a persistence module (V, φ_V). The *inclusion morphism* $j : W \to V$ is defined via the inclusion maps $j_t : W_t \hookrightarrow V_t$ for every $t \in \mathbb{R}$.

Lemma 2.10. *Let (W, φ_W) be a persistence submodule of a persistence module (V, φ_V). The inclusion morphism $j : W \to V$ is a persistence module morphism.*

Proof. Let $s, t \in \mathbb{R}$ with $s \leq t$. Then $j_t \circ \varphi_W(s,t) = j_t \circ \varphi_V(s,t)|_{W_s} = \varphi_V(s,t)|_{W_s} = \varphi_V(s,t) \circ j_s$. By Definition 2.2, $j : W \to V$ is a persistence module morphism. □

Definition 2.11. Let (W, φ_W) be a persistence submodule of a persistence module (V, φ_V). The *quotient morphism* $q : V \to V/W$ is defined via the quotient maps $q_t : V_t \twoheadrightarrow V_t/W_t$ for every $t \in \mathbb{R}$.

Lemma 2.12. *Let (W, φ_W) be a persistence submodule of a persistence module (V, φ_V). The quotient morphism $q : V \to V/W$ is a persistence module morphism.*

Proof. Let $s, t \in \mathbb{R}$ with $s \leq t$. By the definition of the quotient persistence module, $\varphi_{V/W}(s,t)$ is defined to be the unique map such that

$$\varphi_{V/W}(s,t) \circ q_s = q_t \circ \varphi_V(s,t).$$

This is exactly what is required to make q a persistence module morphism. □

A persistence module morphism $\phi : V \to W$ gives rise to the following constructions, which are central to our work in Chapter 3.

Definition 2.13. Let (V, φ_V) and (W, φ_W) be persistence modules and let $\phi : V \to W$ be a persistence module morphism. The persistence modules $(\operatorname{im} \phi, \varphi_{\operatorname{im} \phi})$ and $(\ker \phi, \varphi_{\ker \phi})$ are defined by the vector spaces

$$(\operatorname{im} \phi)_t := \operatorname{im} \phi_t \quad \text{and} \quad (\ker \phi)_t := \ker \phi_t$$

for every $t \in \mathbb{R}$. The transition maps are defined to be the restrictions of the linear maps $\varphi_W(s,t)$ and $\varphi_V(s,t)$ to $\operatorname{im} \phi_s$ and $\ker \phi_s$, respectively, for every $s, t \in \mathbb{R}$ with $s \leq t$.

The persistence module $(\operatorname{coker} \phi, \varphi_{\operatorname{coker} \phi})$ is defined by the (quotient) vector spaces

$$(\operatorname{coker} \phi)_t := \operatorname{coker} \phi_t = W_t / \operatorname{im} \phi_t, \qquad t \in \mathbb{R}.$$

Let $q_t : W_t \to W_t / \operatorname{im} \phi_t$ be the canonical quotient map. The transition maps $\varphi_{\operatorname{coker} \phi}(s,t)$ are defined to be the unique linear maps such that for every $s, t \in \mathbb{R}$ with $s \leq t$,

$$\varphi_{\operatorname{coker} \phi}(s,t) \circ q_s = q_t \circ \varphi_W(s,t).$$

Proposition 2.14. *Let (V, φ_V) and (W, φ_W) be persistence modules and let $\phi : V \to W$ be a persistence module morphism. Then* $\operatorname{im} \phi$, $\ker \phi$, *and* $\operatorname{coker} \phi$ *are persistence modules.*

Proof. We use Lemma 2.5 to show that $(\operatorname{im} \phi, \varphi_{\operatorname{im} \phi})$ is a persistence submodule of

(W, φ_W) and that $(\ker \phi, \varphi_{\ker \phi})$ is a persistence submodule of (V, φ_V). If $t \in \mathbb{R}$, then $\operatorname{im} \phi_t \subseteq W_t$ and $\ker \phi_t \subseteq V_t$ since ϕ_t is a linear map. Since the transition maps $\varphi_{\operatorname{im} \phi}$ and $\varphi_{\ker \phi}$ are given by domain restriction of the transition maps φ_W and φ_V, respectively, it remains to show that $\operatorname{im} \varphi_W(s,t)|_{\operatorname{im} \phi_s} \subseteq \operatorname{im} \phi_t$ and $\operatorname{im} \varphi_W(s,t)|_{\ker \phi_s} \subseteq \ker \phi_t$ for every $s, t \in \mathbb{R}$ with $s \leq t$.

Let $s \leq t$. First we show that $\operatorname{im} \varphi_W(s,t)|_{\operatorname{im} \phi_s} \subseteq \operatorname{im} \phi_t$. For $\alpha \in \operatorname{im} \phi_s$, there exists some $\beta \in V_s$ such that $\phi_s(\beta) = \alpha$. Since the diagram in Definition 2.2 commutes for all $s \leq t$, this implies that

$$\varphi_W(s,t)(\alpha) = [\varphi_W(s,t) \circ \phi_s](\beta) = [\phi_t \circ \varphi_V(s,t)](\beta),$$

yielding $\varphi_W(s,t)(\alpha) \in \operatorname{im} \phi_t$. Since $\alpha \in W_s$ was arbitrary, we have succeeded in showing that $\operatorname{im} \varphi_W(s,t)|_{\operatorname{im} \phi_s} \subseteq \operatorname{im} \phi_t$. Now we show that $\operatorname{im} \varphi_W(s,t)|_{\ker \phi_s} \subseteq \ker \phi_t$. Let $\alpha \in \ker \phi_s$. Then $\phi_s(\alpha) = 0$, and hence $[\varphi_W(s,t) \circ \phi_s](\alpha) = 0$. Again, by the commutativity of the diagram in Definition 2.2, we see that

$$0 = [\varphi_W(s,t) \circ \phi_s](\alpha) = [\phi_t \circ \varphi_V(s,t)](\alpha),$$

so $\varphi_V(s,t)(\alpha) \in \ker \phi_t$. Since $\alpha \in \ker \phi_s$ was arbitrary, we have shown that $\operatorname{im} \varphi_V(s,t)|_{\ker \phi_s} \subseteq \ker \phi_t$. Thus, by Lemma 2.5, $\operatorname{im} \phi$ and $\ker \phi$ are persistence submodules of W and V, respectively, and as such they are persistence modules.

We now show that $\operatorname{coker} \phi$ is a persistence module. The vector spaces $(\operatorname{coker} \phi)_t := \operatorname{coker} \phi_t = W_t / \operatorname{im} \phi_t$ are defined for every $t \in \mathbb{R}$ since $\phi_t : V_t \to W_t$ is a linear map. Thus, it remains to show the transition maps $\varphi_{\operatorname{coker} \phi}(s,t) : (\operatorname{coker} \phi)_s \to (\operatorname{coker} \phi)_t$ are defined for every $s, t \in \mathbb{R}$ with $s \leq t$, and they satisfy properties (i) and (ii) in Definition 2.1.

First we show that for $s, t \in \mathbb{R}$ with $s \leq t$, we have $\varphi_{\operatorname{coker} \phi}(s,t) : (\operatorname{coker} \phi)_s \to (\operatorname{coker} \phi)_t$. By definition, $\varphi_{\operatorname{coker} \phi}(s,t) \circ q_s = q_t \circ \varphi_W(s,t)$, so the domain of $\varphi_{\operatorname{coker} \phi}(s,t)$ is $\operatorname{im} q_s = W_s / \operatorname{im} \phi_s$, and $\operatorname{im} \varphi_{\operatorname{coker} \phi}(s,t) \subseteq \operatorname{im} q_t = W_t / \operatorname{im} \phi_t$. Hence, $\varphi_{\operatorname{coker} \phi}(s,t) : (\operatorname{coker} \phi)_s \to (\operatorname{coker} \phi)_t$.

Now we show that the transition maps satisfy property (i) of Definition 2.1. Let $t \in \mathbb{R}$. Then $\varphi_{\operatorname{coker}\phi}(t,t) \circ q_t = q_t \circ \varphi_W(t,t) = q_t \circ \mathrm{id}_{W_t} = q_t$, so it must be the case that $\varphi_{\operatorname{coker}\phi}(t,t) = \mathrm{id}_{\operatorname{im} q_t} = \mathrm{id}_{W_t/\operatorname{im}\phi_t} = \mathrm{id}_{\operatorname{coker}\phi_t}$. Hence, property (i) is satisfied.

Finally, we show that the transition maps satisfy property (ii) of Definition 2.1. Let $r \leq s \leq t$ in $\mathbb{R}$. Then

$$\begin{aligned}
\varphi_{\operatorname{coker}\phi}(r,t) \circ q_r &= q_t \circ \varphi_W(r,t) \\
&= q_t \circ \varphi_W(s,t) \circ \varphi_W(r,s) \\
&= \varphi_{\operatorname{coker}\phi}(s,t) \circ q_s \circ \varphi_W(r,s) \\
&= \varphi_{\operatorname{coker}\phi}(s,t) \circ \varphi_{\operatorname{coker}\phi}(r,s) \circ q_r.
\end{aligned}$$

Since the map $\varphi_{\operatorname{coker}\phi}(r,t)$ is the unique map satisfying $\varphi_{\operatorname{coker}\phi}(s,t) \circ q_s = q_t \circ \varphi_W(s,t)$, it must be the case that $\varphi_{\operatorname{coker}\phi}(r,t) = \varphi_{\operatorname{coker}\phi}(s,t) \circ \varphi_{\operatorname{coker}\phi}(r,s)$. Thus, property (ii) of Definition 2.1 is also satisfied. We have shown that coker ϕ is a persistence module. □

2.2 Interval Decompositions, Barcodes, and Decorated Endpoints

As suggested in the introduction, the concept of a persistence module arose as a means of organizing families of vector spaces indexed by $\mathbb{R}$ and their associated linear maps. Thus, it makes sense to consider persistence modules that are induced, in some sense, by subsets of $\mathbb{R}$. As will be seen shortly, it is the intervals of $\mathbb{R}$ that play an essential role in the characterization of PFD persistence modules.

Definition 2.15. Let $J \subseteq \mathbb{R}$ be an interval (possibly infinite). The *interval persistence module* (I_J, φ_J) is defined by the vector spaces and transition maps

$$(I_J)_t := \begin{cases} \mathsf{k} & \text{if } t \in J, \\ 0 & \text{otherwise, and} \end{cases}$$

$$\varphi_J(s,t) := \begin{cases} \mathrm{id}_{\mathsf{k}} & \text{if } s,t \in J, \\ 0 & \text{otherwise.} \end{cases}$$

If $J = \emptyset$ is the empty interval, then $(I_J)_t = 0$ for every $t \in \mathbb{R}$; in this case, I_J is called

the *trivial persistence module.*

Lemma 2.16. *Let (I_J, φ_J) be an interval persistence module. Any persistence submodule of I_J is also an interval persistence module.*

Proof. Let (V, φ_V) be a persistence submodule of I_J. By the definition of a persistence submodule, for every $t \in \mathbb{R}$ we know that $V_t \subseteq (I_J)_t$, and so either $V_t = 0$ or $V_t = \mathsf{k}$ since I_J is an interval persistence module. Additionally, for every $s, t \in \mathbb{R}$ with $s \leq t$, we have $\operatorname{im} \varphi_V(s,t) = \operatorname{im} \varphi_J(s,t)|_{V_s} \subseteq V_t$. If $V_t = 0$ for every $t \in \mathbb{R}$, then V is the trivial persistence module, which is an interval persistence module. Suppose that there exist $r \leq t$ in $\mathbb{R}$ such that $V_r = \mathsf{k} = V_t$, and hence $r, t \in J$. Then for any $s \in \mathbb{R}$ with $r \leq s \leq t$, we have

$$\begin{aligned}
\mathrm{id}_{\mathsf{k}} &= \varphi_J(r,t) \\
&= \varphi_J(r,t)|_{V_r} \\
&= \left[(\varphi_J(s,t) \circ \varphi_J(r,s)\right]|_{V_r} \\
&= \varphi_J(s,t)|_{V_s} \circ \varphi_J(r,s)|_{V_r} \\
&= \varphi_V(s,t) \circ \varphi_V(r,s).
\end{aligned}$$

Hence, it follows that $\operatorname{im} \varphi_V(r,s) \neq 0$, yielding $V_s = \mathsf{k}$. Thus, the set $\{t \in \mathbb{R} : V_t = \mathsf{k}\} \subseteq \mathbb{R}$ is an interval, and so V must be an interval persistence module. □

The following fundamental result [16, Theorem 1.3, Theorem 1.4] [17], often referred to as the structure theorem or interval decomposition theorem of PFD persistence modules [11], is a consequence of the Krull-Remak-Schmidt theorem.

Structure Theorem 2.17. *Let V be a PFD persistence module. There exists a (possibly infinite) family of (nonempty) $\mathbb{R}$-intervals $\mathcal{J}$ such that*

$$V \cong \bigoplus_{J \in \mathcal{J}} I_J.$$

Moreover, this decomposition is unique up to a reindexing of the intervals.

The family of interval persistence modules in the Structure Theorem is referred to as the *interval decomposition of* V, and, owing to the uniqueness given by the theorem, the family of intervals itself can be viewed as the following useful invariant.

Definition 2.18. Let V be a PFD persistence module with interval decomposition given by the intervals $\mathcal{J}$. We say that $\mathcal{B}_V := \mathcal{J}$ is the *barcode* of V.

We will have occasion to refer to the intervals in the barcode of a persistence module explicitly. To do so, we find it convenient to use the conventions in [11, 16], which we summarize in the following definition and subsequent remarks.

Definition 2.19. Let $\mathcal{D} = \{+, -\}$. The set of *decorated endpoints* is defined to be the set $\mathbb{E} := \mathbb{R} \times \mathcal{D} \cup \{-\infty, +\infty\}$. We find it convenient to write $t^- := (t, -)$ and $t^+ := (t, +)$ for $t \in \mathbb{R}$. The symbol $\pm$ is used to denote both the symbols $+$ and $-$, in that order, and the symbol $\mp$ is used to denote both the symbols $-$ and $+$, in that order.

To recover the underlying extended real number from a decorated endpoint, we define the map $\pi : \mathbb{E} \to \overline{\mathbb{R}}$ as follows: $\pi(t, \pm) = t$ and $\pi(\pm\infty) = \pm\infty$. We will often write πe instead of $\pi(e)$ for $e \in \mathbb{E}$ to make notation uncluttered.

It is possible to define an addition operation $(\cdot) + (\cdot) : \mathbb{E} \times \mathbb{R} \to \mathbb{E}$ by taking $t^\pm + \delta := (t + \delta)^\pm$ and $\pm\infty + \delta := \pm\infty$ for all $t, \delta \in \mathbb{R}$. Subtraction is defined by $t^\pm - \delta := t^\pm + (-\delta)$. We define multiplication by negative one by $-(\cdot) : \mathbb{E} \to \mathbb{E}$ by taking $-(\pm\infty) := \mp\infty$ and $-(t, \pm) := (-t, \mp)$.

Endowing $\mathcal{D} = \{+, -\}$ with a total order $- < +$ induces a total order on $\mathbb{E}$ with respect to the dictionary order on $\mathbb{R} \times \mathcal{D}$, taking the minimal element to be $\{-\infty\}$ and the maximal element to be $\{+\infty\}$. In particular, for any $t \in \mathbb{R}$, it follows that $t^- < t^+$, and for any $s, t \in \mathbb{R}$ with $s < t$, we have $s^+ < t^-$.

The ordered pairs $\{(c, e) \in \mathbb{E} \times \mathbb{E} : c < e\}$ can be put into one-to-one correspondence

to the intervals in $\mathbb{R}$ as follows. Let $b \leq d \in \mathbb{R}$ and assign

$$(-\infty, d^-) \mapsto (\infty, d)$$
$$(-\infty, d^+) \mapsto (\infty, d]$$
$$(b^-, d^-) \mapsto [b, d)$$
$$(b^-, d^+) \mapsto [b, d]$$
$$(b^+, d^-) \mapsto (b, d)$$
$$(b^+, d^+) \mapsto (b, d]$$
$$(b^-, +\infty) \mapsto [b, +\infty)$$
$$(b^+, +\infty) \mapsto (b, +\infty)$$
$$(-\infty, +\infty) \mapsto (-\infty, +\infty).$$

We will often express an interval $J \subseteq \mathbb{R}$ by using its corresponding ordered pair of decorated endpoints. As in Bauer and Lesnick [11], to avoid confusion with the usual open intervals in $\mathbb{R}$, we will denote an ordered pair $(c, e) \in \mathbb{E} \times \mathbb{E}$ with $c < e$ as $\langle c, e \rangle$.

2.3 Interleavings of Persistence Modules

Persistence module morphisms give rise to the notion of persistence module interleavings, introduced by Chazal, et al. in [7] and generalized by Bubenik, et al. in [12], which are sometimes likened to the concept of *approximate isomorphisms* between persistence modules. That is, two persistence modules that are interleaved are in some way comparable to one another, if not isomorphic. Our results in Chapter 3 make this notion of comparison precise, and we illustrate its utility in the examples that follow in Chapter 4 and beyond.

It is in the theory of interleavings that the rich connection between a persistence module and its indexing set ($\mathbb{R}$ in our case) becomes apparent. The following definition offers a sense of what will follow.

Definition 2.20. Let S be an ordered set. A function $\sigma : S \to S$ is called *monotone* if $x \leq y$ implies that $\sigma(x) \leq \sigma(y)$. If in addition $x \leq \sigma(x)$ for all $x \in S$, then σ is called

a *translation map*. We say that σ is *right-continuous* if

$$\lim_{\substack{x \to t, \\ t < x \in S}} \sigma(x) = \sigma(t)$$

for all $t \in S$.

Monotone functions on $\mathbb{R}$ induce the following useful operator on the space of persistence modules and their morphisms. In the language of category theory, a monotone function $\sigma : \mathbb{R} \to \mathbb{R}$ induces a functor T_σ from the category of persistence modules and their morphisms to itself.

Definition 2.21. Let $\sigma : \mathbb{R} \to \mathbb{R}$ be monotone and let (V, φ_V) be a persistence module. The *σ-shifted persistence module* $(T_\sigma(V), \varphi_{T_\sigma(V)})$ is defined by the vector spaces

$$T_\sigma(V)_t := V_{\sigma(t)}$$

for every $t \in \mathbb{R}$ and transition maps

$$\varphi_{T_\sigma(V)}(s,t) := \varphi_V(\sigma(s), \sigma(t))$$

for every $s, t \in \mathbb{R}$ with $s \leq t$.

Let (W, φ_W) be a persistence module and let $\phi : V \to W$ be a persistence module morphism. The *σ-shifted persistence module morphism* $T_\sigma(\phi) : T_\sigma(V) \to T_\sigma(W)$ is defined via the linear maps

$$T_\sigma(\phi)_t := \phi_{\sigma(t)}$$

for every $t \in \mathbb{R}$.

The following proposition shows that Definition 2.21 makes sense.

Proposition 2.22. *Let $\sigma : \mathbb{R} \to \mathbb{R}$ be monotone. Let (V, φ_V) and (W, φ_W) be persistence modules, and let $\phi : V \to W$ be a persistence module morphism. Then $(T_\sigma(V), \varphi_{T_\sigma(V)})$ is a persistence module and $T_\sigma(\phi) : T_\sigma(V) \to T_\sigma(W)$ is a persistence module morphism.*

Proof. To show that $(T_\sigma(V), \varphi_{T_\sigma(V)})$ is a persistence module, we first note that $T_\sigma(V)_t := V_{\sigma(t)}$ is defined for each $t \in \mathbb{R}$, and that

$$\varphi_{T_\sigma(V)}(t,t) = \varphi_V(\sigma(t), \sigma(t)) = \mathrm{id}_{V_{\sigma(t)}} = \mathrm{id}_{T_\sigma(V)}$$

for every $t \in \mathbb{R}$ since φ_V is a transition map. Since σ is monotone, $\sigma(s) \leq \sigma(t)$. By definition, $\varphi_{T_\sigma(V)}(s,t) = \varphi_V(\sigma(s), \sigma(t))$ is a map from $T_\sigma(V)_s = V_{\sigma(s)}$ to $T_\sigma(V)_t = V_{\sigma(t)}$.

It remains to verify property (ii) of Definition 2.1. To this end, let $r \leq s \leq t$. Since σ is monotone, $\sigma(r) \leq \sigma(s) \leq \sigma(t)$. Thus, since φ_V is a transition map, we have

$$\begin{aligned}
\varphi_{T_\sigma(V)}(r,t) &= \varphi_V(\sigma(r), \sigma(t)) \\
&= \varphi_V(\sigma(s), \sigma(t)) \circ \varphi_V(\sigma(r), \sigma(s)) \\
&= \varphi_{T_\sigma(V)}(s,t) \circ \varphi_{T_\sigma(V)}(r,s).
\end{aligned}$$

Hence, $(T_\sigma(V), \varphi_{T_\sigma(V)})$ is a persistence module.

Next we show that given a persistence module morphism $\phi : V \to W$, the map $T_\sigma(\phi) : T_\sigma(V) \to T_\sigma(W)$ is a persistence module morphism. We must show that $T_\sigma(\phi)$ satisfies Definition 2.2. By definition, $T_\sigma(\phi)_t : T_\sigma(V)_t \to T_\sigma(W)_t$ is the map $T_\sigma(\phi)_t = \phi_{\sigma(t)} : V_{\sigma(t)} \to W_{\sigma(t)}$, and so it is defined since ϕ is a persistence module morphism. Thus, it remains to show that $T_\sigma(\phi)$ satisfies the commutativity constraint. Let $s \leq t$. Then

$$\begin{aligned}
\varphi_{T_\sigma(W)}(s,t) \circ T_\sigma(\phi)_s &= \varphi_W(\sigma(s), \sigma(t)) \circ \phi_{\sigma(s)} \\
&= \phi_{\sigma(t)} \circ \varphi_V(\sigma(s), \sigma(t)) \\
&= T_\sigma(\phi)_t \circ \varphi_{T_\sigma(V)}(s,t),
\end{aligned}$$

and so $T_\sigma(\phi)$ is a persistence module morphism. □

The following propositions and corollary, while fairly elementary, are worth noting.

Lemma 2.23. *Let $\sigma : \mathbb{R} \to \mathbb{R}$ be monotone and bijective. Then its inverse σ^{-1} is also monotone.*

Proof. Let $s, t \in \mathbb{R}$ with $s \leq t$. Since σ is bijective, there exist $x \in \mathbb{R}$ with $\sigma(x) = s$ and $y \in \mathbb{R}$ with $\sigma(y) = t$. Moreover, since $\sigma(x) \leq \sigma(y)$ and σ is bijective and monotone, we must have $x \leq y$. Thus, $\sigma^{-1}(s) = \sigma^{-1}(\sigma(x)) = x \leq y \leq \sigma^{-1}(\sigma(y)) = \sigma^{-1}(t)$, and so the monotonicity of σ^{-1} is proved. □

Lemma 2.24. *Let $\sigma, \tau : \mathbb{R} \to \mathbb{R}$ be monotone functions and let (V, φ_V) be a persistence module. Then $T_{\sigma\circ\tau}(V) = T_\sigma[T_\tau(V)]$.*

Proof. Let $t \in \mathbb{R}$. Then $T_{\sigma\circ\tau}(V)_t = V_{(\sigma\circ\tau)(t)} = V_{\sigma[\tau(t)]} = T_\sigma(V_{\tau(t)}) = T_\sigma[T_\tau(V)]$. Equality at the level of transition maps follows similarly. □

Corollary 2.25. *Let $\sigma : \mathbb{R} \to \mathbb{R}$ be a bijective monotone function and let (V, φ_V) be a persistence module. Then $V = T_{\sigma^{-1}}[T_\sigma(V)]$.*

The significance of Corollary 2.25 is that when $\sigma : \mathbb{R} \to \mathbb{R}$ is both monotone and bijective, the persistence module V is completely recoverable from the persistence module $T_\sigma(V)$, and thus the barcode $\mathcal{B}_V$ is completely recoverable from the barcode $\mathcal{B}_{T_\sigma(V)}$.

The following definition is the key to defining an interleaving between two persistence modules.

Definition 2.26. Let $\tau, \sigma : \mathbb{R} \to \mathbb{R}$ be monotone. We call the pair of maps (τ, σ) a *translation pair* if $x \leq (\tau \circ \sigma)(x)$ and $x \leq (\sigma \circ \tau)(x)$ for all $x \in \mathbb{R}$.

We are now in a place to define an interleaving between two persistence modules.

Definition 2.27. Let (V, φ_V) and (W, φ_W) be persistence modules and let (τ, σ) be a translation pair, considered as an ordered pair. We say that the persistence modules V and W are (τ, σ)-*interleaved* if there exist persistence module morphisms $\phi : V \to T_\tau(W)$ and $\psi : W \to T_\sigma(V)$ such that

$$T_\tau(\psi)_t \circ \phi_t = \varphi_V[t, (\sigma \circ \tau)(t)]$$

and

$$T_\sigma(\phi)_t \circ \psi_t = \varphi_W[t, (\tau \circ \sigma)(t)]$$

for all $t \in \mathbb{R}$. We refer to these last two conditions as the *commutativity constraint* of the interleaving. The maps ϕ and ψ are called *interleaving morphisms.*

Definition 2.27 leads to the following result, which is sometimes referred to as the triangle inequality [12].

Proposition 2.28. *Let $(U, \varphi_U), (V, \varphi_V)$, and (W, φ_W) be persistence modules. If U and V are (τ, σ)-interleaved and V and W are (ρ, η)-interleaved, then U and W are $(\rho \circ \tau, \sigma \circ \eta)$-interleaved.*

Proof. Since U and V are (τ, σ)-interleaved and V and W are (ρ, η)-interleaved, there exist interleaving morphisms $\phi_1 : U \to T_\tau(V)$, $\psi_1 : V \to T_\sigma(U)$, $\phi_2 : V \to T_\rho(W)$ and $\psi_2 : W \to T_\eta(V)$ satisfying Definition 2.27. Define $\phi := T_\tau(\phi_2) \circ \phi_1$ and $\psi := T_\eta(\psi_1) \circ \psi_2$. We will show that the persistence module morphisms ϕ and ψ witness a $(\rho \circ \tau, \sigma \circ \eta)$-interleaving bewteen U and W.

First we show that $\phi : U \to T_{(\rho\circ\tau)}W$ and $\psi : W \to T_{(\sigma\circ\eta)}(U)$. Let $t \in \mathbb{R}$. Then

$$\phi_t = T_\tau(\phi_2)_t \circ (\phi_1)_t = (\phi_2)_{\tau(t)} \circ (\phi_1)(t),$$

which is indeed a map from U_t to $W_{(\rho\circ\tau)(t)} = T_{(\rho\circ\tau)}(W)_t$. Similarly,

$$\psi_t = T_\eta(\psi_1)_t \circ (\psi_2)_t = (\psi_1)_{\eta(t)} \circ (\psi_2)_t$$

is a map from W_t to $U_{(\sigma\circ\eta)(t)} = T_{\sigma\circ\eta}(U)_t$.

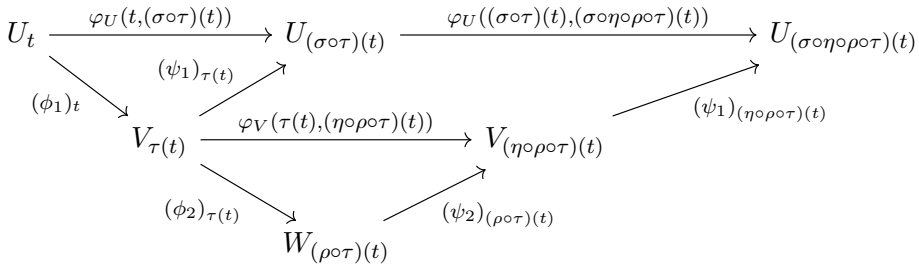
It remains to show that the commutativity constraints from Definition 2.27 are satisfied by showing that

$$T_{(\rho\circ\tau)}(\psi)_t \circ \phi_t = \varphi_U[t, (\sigma \circ \eta \circ \rho \circ \tau)(t)] \tag{2.28a}$$

and

$$T_{(\sigma\circ\eta)}(\phi)_t \circ \psi_t = \varphi_W[t, (\rho \circ \tau \circ \sigma \circ \eta)(t)] \tag{2.28b}$$

for all $t \in \mathbb{R}$. Proving (2.28a) comes down to proving the commutativity of the two triangles and the parallelogram in the following diagram.



By examining indices, one can see that the two triangles commute since U and V are (τ,σ)-interleaved and V and W are (ρ,η)-interleaved. Additionally, the parallelogram commutes since $\psi_1 : V \to T_\sigma(U)$ is a persistence module morphism. Thus, observing that $\phi_t = T_\tau(\phi_2)_t \circ (\phi_1)_t = (\phi_2)_{\tau(t)} \circ (\phi_1)_t$ and that

$$\begin{aligned} T_{(\rho\circ\tau)}(\psi)_t &= \psi_{(\rho\circ\tau)(t)} \\ &= T_\eta(\psi_1)_{(\rho\circ\tau)(t)} \circ (\psi_2)_{(\rho\circ\tau)(t)} \\ &= (\psi_1)_{(\eta\circ\rho\circ\tau)(t)} \circ (\psi_2)_{(\rho\circ\tau)(t)}, \end{aligned}$$

we achieve (2.28a). The proof of equality (2.28b) is similar. □

The following proposition shows how to stitch two persistence modules together that are interleaved with a third. The author is not aware of the appearance of the following construction in existing literature. We illustrate the use of this construction in Section 4.1.4.

Proposition 2.29. *Let $(V,\varphi_V),(V',\varphi_{V'})$, and (W,φ_W) be persistence modules. Suppose that V and W are (τ,σ)-interleaved via the morphisms $\phi : V \to T_\tau(W)$ and $\psi : W \to T_\sigma(V)$, and that V' and W are (τ',σ')-interleaved via $\phi' : V' \to T_{\tau'}(W)$ and $\psi' : W \to T_{\sigma'}(V')$.*

Let $t_0 \in \mathbb{R}$ such that $t_0 \leq (\sigma' \circ \tau)(t_0)$ be fixed and consider the vector spaces U_t for

$t \in \mathbb{R}$ and linear maps $\varphi_U(s,t)$ for $s,t \in \mathbb{R}$ with $s \le t$ defined by

$$U_t = \begin{cases} V_t & \text{if } \ t < t_0 \\ V_{t_0} & \text{if } \ t_0 \le t < (\sigma' \circ \tau)(t_0) \\ V'_t & \text{if } \ (\sigma' \circ \tau)(t_0) \le t; \end{cases}$$

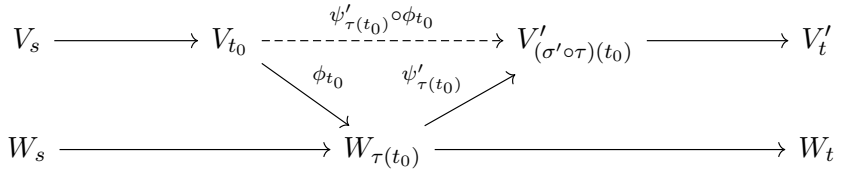
$$\varphi_U(s,t) = \begin{cases} \varphi_V(\min\{s,t_0\}, \min\{t,t_0\}) & \text{if } \ s \le t < (\sigma' \circ \tau)(t_0) \\ \varphi_{V'}(t, (\sigma' \circ \tau)(t_0)) \circ \psi'_{\tau(t_0)} \circ \phi_{t_0} \circ \varphi_V(\min\{s,t_0\}, t_0) & \text{if } \ s < (\sigma' \circ \tau)(t_0) \le t \\ \varphi_{V'}(s,t) & \text{if } \ (\sigma' \circ \tau)(t_0) \le s \le t. \end{cases}$$

Then (U, φ_U) is a persistence module and U and W are (η, ρ)-interleaved where

$$\eta(t) = \begin{cases} \tau(t) & \text{if } \ t \le t_0 \\ \tau(t_0) & \text{if } \ t_0 < t < (\sigma' \circ \tau)(t_0) \\ \tau'(t) & \text{if } \ (\sigma' \circ \tau)(t_0) \le t; \end{cases}$$

$$\rho(t) = \begin{cases} \sigma(t) & \text{if } \ t < \sigma^{-1}(t_0) \\ (\sigma' \circ \tau)(t_0) & \text{if } \ \sigma^{-1}(t_0) \le t < \tau(t_0) \\ \sigma'(t) & \text{if } \ \tau(t_0) \le t. \end{cases}$$

Proof. It is clear that U is a persistence module since $t_0 \in \mathbb{R}$ is chosen such that $t_0 \le (\sigma' \circ \tau)(t_0)$. The following diagram (with unlabeled arrows assumed to be the appropriate transition maps) shows the idea behind the vector spaces of U and the transition maps $\varphi_U(s,t)$ for $s < (\sigma' \circ \tau)(t_0) \le t$.



To show that U and W are (η, ρ)-interleaved we will show that the morphisms

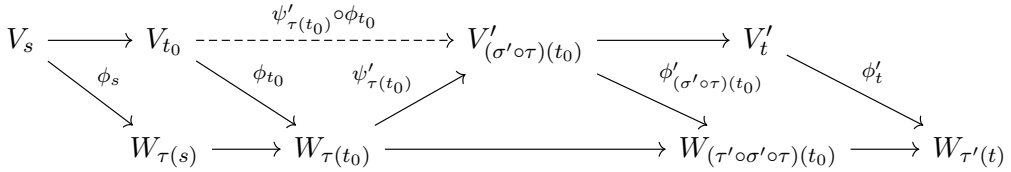
$\bar{\phi} : U \to T_\eta(W)$ and $\bar{\psi} : W \to T_\rho(U)$, where

$$\bar{\phi}_t = \begin{cases} \phi_t & \text{if} \quad t \leq t_0 \\ \phi_{t_0} & \text{if} \quad t_0 < t < (\sigma' \circ \tau)(t_0) \\ \phi'_t & \text{if} \quad (\sigma' \circ \tau)(t_0) \leq t, \end{cases}$$

$$\bar{\psi}_t = \begin{cases} \psi_t & \text{if} \quad t < \sigma^{-1}(t_0) \\ \psi'_{\tau(t_0)} \circ \varphi_W(t, \tau(t_0)) & \text{if} \quad \sigma^{-1}(t_0) \leq t < \tau(t_0) \\ \psi'_t & \text{if} \quad \tau(t_0) \leq t, \end{cases}$$

give the desired interleaving of U and W. To show that $\bar{\phi}$ and $\bar{\psi}$ are persistence module morphisms, we first note that the monotone functions η and ρ line up with the indices of the shifts of $\bar{\phi}$ and $\bar{\psi}$ by inspection. Let $s, t \in \mathbb{R}$.

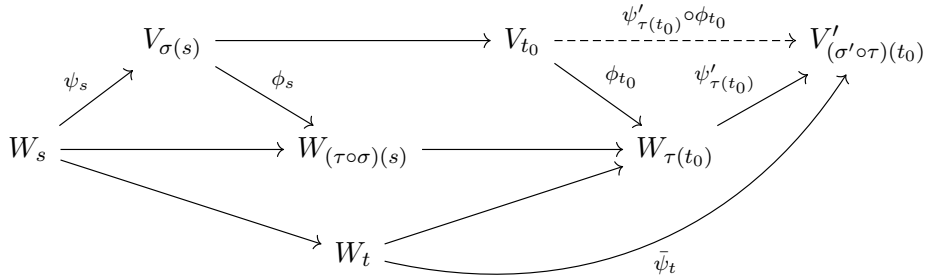
We will now show that $\bar{\phi} : U \to T_\eta(W)$ is a persistence module morphism. If $s \leq t < (\sigma' \circ \tau)(t_0)$, then $\bar{\phi}_s = \phi_{\min\{s,t_0\}}$ and $\bar{\phi}_t = \phi_{\min\{t,t_0\}}$, and if $(\sigma' \circ \tau)(t_0) \leq s \leq t$, then $\bar{\phi}_s = \phi'_s$ and $\bar{\phi}_t = \phi'_t$, and so these cases hold. Now suppose that $s < (\sigma' \circ \tau)(t_0) \leq t$. The diagram



commutes since both ϕ and ϕ' are persistence module morphisms and since V' and W are (τ', σ')-interleaved, where unlabeled arrows are transition maps. It follows that $\bar{\phi}_t \circ \varphi_U(s,t) = \varphi_{T_\eta(W)}(s,t) \circ \bar{\phi}_s$, and thus $\bar{\phi} : U \to T_\eta(W)$ is a persistence module morphism.

Now we will show that $\bar{\psi} : W \to T_\rho(U)$ is a persistence module morphism. If $s \leq t < \sigma^{-1}(t_0)$, then $\bar{\psi}_s = \psi_s$ and $\bar{\psi}_t = \psi_t$, and if $\tau(t_0) \leq s \leq t$, then $\bar{\psi}_s = \psi'_s$ and $\bar{\psi}_t = \psi'_t$, so these cases are clear. Suppose that $s < \sigma^{-1}(t_0) \leq t \leq \tau(t_0)$. This choice of s and t yield $\bar{\psi}_s = \psi_s$ and $\bar{\psi}_t = \psi'_{\tau(t_0)} \circ \varphi_W(t, \tau(t_0))$ by definition. Since $s < \sigma^{-1}(t_0)$, it follows that $s \leq (\tau \circ \sigma)(s) \leq \tau(t_0)$, and so the following diagram commutes, where

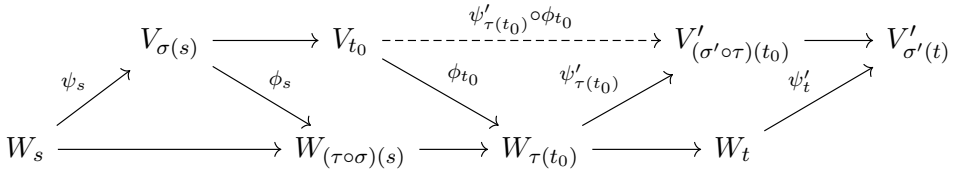
unlabeled arrows are transition maps.



Hence, $\bar{\psi}_t \circ \varphi_W(s,t) = \varphi_{T_\rho(U)}(s,t) \circ \bar{\psi}_s$ in this case as well. For $\sigma^{-1}(t_0) \leq s \leq t \leq \tau(t_0)$, then

$$\begin{aligned}
\bar{\psi}_t \circ \varphi_W(s,t) &= \psi'_{\tau(t_0)} \circ \varphi_W(t, \tau(t_0)) \circ \varphi_W(s,t) \\
&= \varphi_U((\sigma' \circ \tau)(t_0), (\sigma' \circ \tau)(t_0)) \circ \psi'_{\tau(t_0)} \circ \varphi_W(s, \tau(t_0)) \\
&= \varphi_{T_\rho(U)}(s,t) \circ \bar{\psi}_s,
\end{aligned}$$

and so this case also holds. If $s < \sigma^{-1}(t_0) \leq \tau(t_0) < t$, the commutativity of the diagram



ensures that $\bar{\psi}_t \circ \varphi_W(s,t) = \varphi_{T_\rho(U)}(s,t) \circ \bar{\psi}_s$. Finally, it remains to check the case when $\sigma^{-1}(t_0) \leq s \leq \tau(t_0) < t$. This follows from extending the case when $\sigma^{-1}(t_0) \leq s \leq t \leq \tau(t_0)$ by the right-most commutative square in the previous diagram. Thus, we have shown that $\bar{\phi}$ and $\bar{\psi}$ are persistence module morphisms that witness a (η, ρ)-interleaving between U and W.

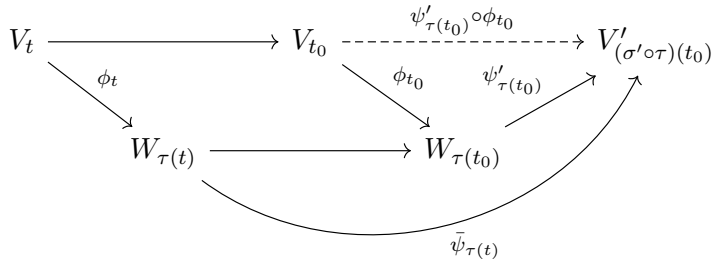
We must now show that $\bar{\phi}$ and $\bar{\psi}$ satisfy the commutativity constraints of Definition 2.27. That is, for every $t \in \mathbb{R}$, we must show that

$$T_\eta(\bar{\psi})_t \circ \bar{\phi}_t = \varphi_U(t, (\rho \circ \eta)(t)) \tag{2.29a}$$

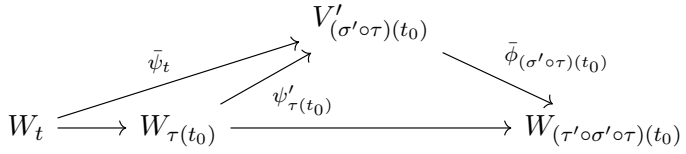
and

$$T_\rho(\bar{\phi})_t \circ \bar{\psi}_t = \varphi_W(t, (\rho \circ \eta)(t)) \tag{2.29b}$$

for all $t \in \mathbb{R}$. First we prove (2.29a) for all $t \in \mathbb{R}$. When $\tau(t) < \sigma^{-1}(t_0)$, then (2.29a) follows from the fact that V and W are (τ, σ)-interleaved. Similarly, when $(\sigma' \circ \tau)(t_0) \leq t$, (2.29a) holds since V' and W are (τ', σ')-interleaved. If $t_0 \leq t < (\sigma' \circ \tau)(t_0)$, then (2.29a) holds by the definition of (U, φ_U). Now suppose that $t \leq t_0$ such that $\sigma^{-1}(t_0) \leq \tau(t)$. Then (2.29a) holds from the commutativity of the following diagram.



Thus we have shown that (2.29a) holds for every $t \in \mathbb{R}$. We now verify that (2.29b) holds for every $t \in \mathbb{R}$. When $t < \sigma^{-1}(t_0)$, then (2.29b) follows since V and W are (τ, σ)-interleaved. Similarly, if $\tau(t_0) \leq t$, then (2.29b) follows since V' and W are (τ', σ')-interleaved. For $\sigma^{-1}(t_0) \leq t < \tau(t_0)$, then (2.29b) follows by the commutativity of the following diagram.



Thus, we have also shown that (2.29b) holds for every $t \in \mathbb{R}$. It follows that (η, ρ) is a translation pair. □

Conjecture 2.30. *The function $(\rho \circ \eta)$ from Proposition 2.29 is right-continuous if both $(\sigma \circ \tau)$ and $(\sigma' \circ \tau')$ are right-continuous.*

Recall that the trivial persistence module is one that has vector spaces $V_t = 0$ for all $t \in \mathbb{R}$, and thus all of the transition maps must be trivial maps as well. Conversely, if every transition map of a persistence module (V, φ_V) is trivial, then $\mathrm{id}_V(t,t) = 0$ for every $t \in \mathbb{R}$, and so we must have $V_t = 0$ for every $t \in \mathbb{R}$ and thus V must be the trivial persistence module. The following definition leads to the notion of an "approximately trivial" persistence module by relaxing this condition on the transition maps.

Definition 2.31. Let $\sigma : \mathbb{R} \to \mathbb{R}$ be a translation map. We say a persistence module (V, φ_V) is *σ-trivial* if $\varphi_V(t, \sigma(t)) = 0$ for all $t \in \mathbb{R}$.

The following remark establishes intuition for how Definition 2.31 relates to the sizes of the intervals in the barcode of a PFD persistence module.

Remark 2.32. Let (V, φ_V) be a PFD persistence module. By the Structure Theorem, there exists a persistence module isomorphism $\Phi : V \to \bigoplus_{J \in \mathcal{B}_V} I_J$, where each I_J is an interval persistence module. For $J \in \mathcal{B}_J$, let

$$\Pi_J : \bigoplus_{J' \in \mathcal{B}_V} I_{J'} \to I_J$$

be projection onto the persistence submodule corresponding to the interval persistence module I_J. Then the transition maps φ_{I_J} for the interval persistence module I_J satisfy

$$\varphi_{I_J}(s,t) \circ (\Pi_J)_s \circ \Phi_s = (\Pi_J)_t \circ \Phi_t \circ \varphi_V(s,t)$$

for every $s, t \in \mathbb{R}$ with $s \leq t$. Thus, if (V, φ_V) is σ-trivial, so is I_J.

Denote J as a pair of decorated endpoints, so that $J = \langle b, d \rangle$. The significance of Definition 2.31 is that if (V, φ_V) is σ-trivial, the interval J can be no longer than $\sigma(\pi b) - \pi b$, where $\pi : \mathbb{E} \to \mathbb{R}$ is the projection to the real part of the decorated endpoint

(or identity on $\pm\infty$). Thus, a statement about the σ-triviality of a persistence module may be viewed as a statement about the lengths of the intervals in its barcode.

In light of Remark 2.32, the following proposition shows how the translation pair associated to two interleaved persistence modules controls the lengths of the barcodes in the interval decompositions of the kernel and cokernel (Definition 2.13) of the interleaving morphisms.

Proposition 2.33. *Suppose that (V, φ_V) and (W, φ_W) are PFD persistence modules that are (τ, σ)-interleaved via interleaving morphisms $\phi : V \to T_\tau(W)$ and $\psi : W \to T_\sigma(V)$. Then*

(i) $\ker \phi$ *and* $\operatorname{coker} \phi$ *are $(\sigma \circ \tau)$-trivial, and*

(ii) $\ker \psi$ *and* $\operatorname{coker} \psi$ *are $(\tau \circ \sigma)$-trivial.*

Proof. To prove (i), we first show that $\ker \phi$ is $(\sigma \circ \tau)$-trivial by showing that

$$\varphi_{\ker \phi}(t, (\sigma \circ \tau)(t)) = 0$$

for every $t \in \mathbb{R}$. To this end, fix $t \in \mathbb{R}$. By the commutativity constraint of a (τ, σ)-interleaving, we know that $\varphi_V(t, (\sigma \circ \tau)(t)) = T_\tau(\psi)_t \circ \phi_t$. Thus,

$$\varphi_V(t, (\sigma \circ \tau)(t))|_{\ker \phi_t} = T_\tau(\psi)_t \circ \phi_t|_{\ker \phi_t} = 0.$$

By the definition of the persistence module $\ker \phi$, we have

$$\varphi_{\ker \phi}(t, (\sigma \circ \tau)(t)) = \varphi_V(t, (\sigma \circ \tau)(t))|_{\ker \phi_t} = 0,$$

and so we're done.

Next, we show that $\operatorname{coker} \phi$ is $(\sigma \circ \tau)$-trivial, which requires we show that

$$\varphi_{\operatorname{coker} \phi}(t, (\sigma \circ \tau)(t)) = 0$$

for every $t \in \mathbb{R}$. Recall from Definition 2.13 that the transition maps of the persistence

module coker ϕ are defined to be the unique linear maps $\varphi_{\operatorname{coker}\phi}(r,s)$ such that

$$\varphi_{\operatorname{coker}\phi}(r,s) \circ q_r = q_s \circ \varphi_{T_\tau(W)}(r,s)$$

for every $r \leq s \in \mathbb{R}$, where $q_r := (\alpha \mapsto \alpha + \operatorname{im}\phi_r)$ for every $\alpha \in T_\tau(W)_r$ is the quotient map. Thus, it suffices to show that $\operatorname{im}\varphi_{T_\tau(W)}(t,(\sigma\circ\tau)(t)) \subseteq \operatorname{im}\phi_{(\sigma\circ\tau)(t)}$ for every $t \in \mathbb{R}$. Let $t \in \mathbb{R}$. Then

$$\begin{aligned}
\varphi_{T_\tau(W)}(t,(\sigma\circ\tau)(t)) &= \varphi_W(\tau(t),(\tau\circ\sigma\circ\tau)(t)) && \text{(2.33a)}\\
&= T_\sigma(\phi)_{\tau(t)} \circ \psi_{\tau(t)} && \text{(2.33b)}\\
&= \phi_{(\sigma\circ\tau)(t)} \circ \psi_{\tau(t)} && \text{(2.33c)}
\end{aligned}$$

where (2.33a) follows from the definition of the maps $\varphi_{T_\tau(W)}$, (2.33b) follows from the commutativity constraint of the interleaving morphisms ϕ and ψ, and (2.33c) follows from the definition of $T_\sigma(\phi)$. Hence, we have shown that

$$\operatorname{im}\varphi_{T_\tau(W)}(t,(\sigma\circ\tau)(t)) \subseteq \operatorname{im}\phi_{(\sigma\circ\tau)(t)}$$

for every $t \in \mathbb{R}$, from which it follows that coker ϕ is $(\sigma\circ\tau)$-trivial.

Finally, (ii) follows from (i) by reversing the roles of ϕ and ψ, creating a (σ,τ)-interleaving of W and V; it follows directly that $\ker\psi$ and coker ψ are $(\tau\circ\sigma)$-trivial. □

Thus, the translation pair (τ,σ) controls the sizes of the intervals of the barcode decompositions of the kernel and cokernel of the corresponding interleaving morphisms. In particular, the smaller the differences $(\sigma\circ\tau)(t) - t$ and $(\tau\circ\sigma)(t) - t$ for every $t \in \mathbb{R}$, the smaller the intervals in the barcodes of the kernel and cokernel of the interleaving morphisms, effectively giving the extent to which the interleaving morphisms fail to be isomorphisms.

The following definition relates the notion of an interleaving from [11, 7, 16] to the generalized interleaving of Definition 2.27.

Definition 2.34. The persistence modules (V,φ_V) and (W,φ_W) are *δ-interleaved* if

they are (τ, σ)-interleaved with $\tau(t) = t + \delta = \sigma(t)$ for some $\delta \geq 0$.

Remark 2.35. Two persistence modules that are 0-interleaved are isomorphic as persistence modules.

Definition 2.34 and Remark 2.35 also hint at this notion of approximate isomorphism. Without being precise, if two persistence modules are δ-interleaved for $\delta \geq 0$, then they are isomorphic to within a tolerance of δ in some sense, and thus their barcodes also differ by this tolerance of δ, again in some sense. Typically, this tolerance is stated in terms of the *bottleneck distance* between two barcodes. As in Bauer and Lesnick [11], it is possible to give this definition in terms of δ-interleavings.

Definition 2.36. Let (V, φ_V) and (W, φ_W) be PFD persistence modules and let $\mathcal{B}_V$ and $\mathcal{B}_W$ be their associated barcodes. The *bottleneck distance* between $\mathcal{B}_V$ and $\mathcal{B}_W$ is defined as follows:

$$d_B(\mathcal{B}_V, \mathcal{B}_W) := \inf\{\delta \geq 0 : V \text{ and } W \text{ are } \delta\text{-interleaved}\}.$$

Note that as consequence of the Structure Theorem, any two persistence modules yielding equivalent barcodes are isomorphic, and thus an application of Remark 2.35 and Proposition 2.28 makes the bottleneck distance between two barcodes well-defined. Additionally, if two persistence modules are isomorphic, then the bottleneck distance of their corresponding barcodes is zero. The converse of this statement, referred to as the isometry theorem [11, 16], is a consequence of the Algebraic Stability Theorem (Theorem 2.67).

2.4 Remarks on Translation Pairs and Constructions for Monotone Functions

We now point out that our definition for an interleaving (Definition 2.27) differs slightly from the one given by Bubenik, et al. [12] in that it does not require both $t \leq \tau(t)$ and $t \leq \sigma(t)$ for all $t \in \mathbb{R}$, but rather that $t \leq (\tau \circ \sigma)(t)$ and $t \leq (\sigma \circ \tau)(t)$ for all $t \in \mathbb{R}$. This is due to the fact that Bubenik, et al. were concerned with constructing a

monoidal category on the collection of functors induced by translation maps, and here we are only concerned with satisfying commutativity properties, which our definition maintains. We illustrate the use of our modified definition explicitly in Sections 4.1.1 and 4.3, where it is natural to express the interleavings such that the translation maps may have either $\tau(t) < t$ or $\sigma(t) < t$ for some or for all $t \in \mathbb{R}$. Even so, it is often possible to go from our case to the situation in [12], as we show next.

Definition 2.37. Let (τ, σ) be a translation pair. If there exists a bijective increasing function $\gamma : \mathbb{R} \to \mathbb{R}$ such that $x \leq (\gamma \circ \tau)(x)$ and $x \leq (\sigma \circ \gamma^{-1})(x)$ for all $x \in \mathbb{R}$, we say that the ordered pair (τ, σ) is *alignable*, or *alignable via* γ.

Proposition 2.38. *Let (V, φ_V) and (W, φ_W) be (τ, σ)-interleaved where (τ, σ) is alignable via γ. Then V and $T_\gamma(W)$ are $(\gamma \circ \tau, \sigma \circ \gamma^{-1})$-interleaved.*

Proof. This is a straightforward application of Definition 2.37 and Prop 2.28. □

Combining Proposition 2.38 with Corollary 2.25, we see that a statement about $T_\gamma(W)$, resulting from the modified interleaving aligned via γ, will carry forward to a statement about W through an application of $T_{\gamma^{-1}}$. Additionally, the interleaving here would agree with the definition given in [12] since $t \leq (\gamma \circ \tau)(t)$ and $t \leq (\sigma \circ \gamma^{-1})(t)$ for all $t \in \mathbb{R}$. Thus, although it is unnecessary to do so, when two persistence modules V and W are (τ, σ)-interleaved by a translation pair alignable via some bijective monotone function $\gamma : \mathbb{R} \to \mathbb{R}$, one could first shift the persistence module W to $T_\gamma(W)$ to be in the setting of [12], apply the theory we give in this chapter, and then shift the result back to a statement about W.

The following proposition gives a class of alignable translation pairs (τ, σ).

Proposition 2.39. *Let (τ, σ) be a translation pair. If either of the following hold:*

(i) τ is bijective

(ii) σ is bijective

then the ordered pair (τ, σ) is alignable.

Proof. Essentially, the problem comes down to showing the existence of a bijective monotone function $\gamma : \mathbb{R} \to \mathbb{R}$ such that $x \leq (\gamma \circ \tau)(x)$ and $x \leq (\sigma \circ \gamma^{-1})(x)$ for all $x \in \mathbb{R}$. We will do this for each of the cases listed.

To prove (i), suppose that τ is bijective. Then setting $\gamma(x) = \tau^{-1}(x)$ yields $x \leq (\tau^{-1} \circ \tau)(x) = x$ and $x \leq (\sigma \circ \tau)(x) = (\sigma \circ (\tau^{-1})^{-1})(x)$ for all $x \in \mathbb{R}$ since (τ, σ) is a translation pair. Hence, (τ, σ) is alignable via $\gamma = \tau^{-1}$.

To prove (ii), suppose that σ is bijective. Then setting $\gamma(x) = \sigma(x)$ yields $x \leq (\sigma \circ \tau)$ and $x \leq (\sigma \circ \sigma^{-1})(x) = x$ for all $x \in \mathbb{R}$ since (τ, σ) is a translation pair. Hence, (τ, σ) is alignable via $\gamma = \sigma$. □

Thus, while Proposition 2.39, Proposition 2.38, and Corollary 2.25 show that the examples we give in Sections 4.1.1 and 4.3 could be put into the setting of [12] (all of the translation maps are bijective), our hope is to convince the reader that our definition for an interleaving of two persistence modules (Definition 2.27) removes a certain amount of burden when applying the Algebraic Stability Theorem for Generalized Interleavings (Theorem 3.13), as the function γ does not need to be constructed or applied in order to obtain the results given by the theorem (or reversed at the end of the computation).

To achieve the results of Chapter 3, we will need the following additional concepts. While many of the propositions we establish may at first seem unmotivated, this work is necessary to extend the work of Bauer and Lesnick [11] to more general interleavings. In particular, one must take care when treating the decorated endpoints of the intervals of barcode decompositions when the translation maps are not bijective.

Our first step is to extend a monotone function $\sigma : \mathbb{R} \to \mathbb{R}$ to the extended real numbers $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$. For the work that follows, it will make sense to define the extension of σ to $\overline{\mathbb{R}}$ as follows. The asymmetry in the definition is due to the need to establish the inequalities in Lemma 2.42.

Definition 2.40. Let $\sigma : \mathbb{R} \to \mathbb{R}$ be monotone. We extend σ to a function $\overline{\sigma} : \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ by defining

$$\overline{\sigma}(-\infty) := \lim_{x \to -\infty} \sigma(x) \quad \text{and} \quad \overline{\sigma}(+\infty) := +\infty.$$

We call $\overline{\sigma}$ the *closure* of σ.

If $\sigma : \mathbb{R} \to \mathbb{R}$ is monotone, then its closure $\overline{\sigma}$ is also monotone. If $\sigma : \mathbb{R} \to \mathbb{R}$ is monotone and right-continuous, then by definition, its closure $\overline{\sigma}$ is also right-continuous.

Definition 2.41. Let S be either $\mathbb{R}$ or $\overline{\mathbb{R}}$, and let $\sigma : S \to S$ be monotone. The *left generalized inverse* of σ, written $\sigma_L^{-1} : S \to \overline{S}$, is defined as

$$\sigma_L^{-1}(t) := \inf\{s \in S : \sigma(s) \geq t\}$$

for every $t \in S$, where by convention we set the infimum of the empty set to be $\inf \emptyset := +\infty$. The *right generalized inverse* of σ, written $\sigma_R^{-1} : \sigma(S) \to \overline{S}$ is defined as

$$\sigma_R^{-1}(t) := \sup\{s \in S : \sigma(s) = t\}$$

for every $t \in \sigma(S)$. Note the difference in the domains of σ_L^{-1} and σ_R^{-1}.

Note that if $\sigma : \mathbb{R} \to \mathbb{R}$ is bijective with inverse σ^{-1}, then $\sigma_L^{-1} = \sigma^{-1} = \sigma_R^{-1}$.

Lemma 2.42. *If $\sigma : \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ is monotone and right-continuous with $\sigma(+\infty) = +\infty$, the inequalities*

$$(\sigma_L^{-1} \circ \sigma)(t) \leq t \leq (\sigma \circ \sigma_L^{-1})(t)$$

hold for all $t \in \overline{\mathbb{R}}$.

Proof. Let $t \in \overline{\mathbb{R}}$. First we prove that $(\sigma_L^{-1} \circ \sigma)(t) \leq t$. By definition,

$$(\sigma_L^{-1} \circ \sigma)(t) = \inf\{x \in \overline{\mathbb{R}} : \sigma(x) \geq \sigma(t)\},$$

and since t satisfies $\sigma(t) \geq \sigma(t)$, then necessarily $(\sigma_L^{-1} \circ \sigma)(t) \leq t$.

Now we show that $t \leq (\sigma \circ \sigma_L^{-1})(t)$ for all $t \in \overline{\mathbb{R}}$. By the definition of σ_L^{-1}, we have $t \leq \sigma(x)$ for every $\sigma_L^{-1}(t) < x$. By right-continuity, we see that

$$t \leq \lim_{\substack{x \to \sigma_L^{-1}(t), \\ \sigma_L^{-1}(t) < x}} \sigma(x) = (\sigma \circ \sigma_L^{-1})(t),$$

and thus the inequality $t \leq (\sigma \circ \sigma_L^{-1})(t)$ holds for all $t \in \overline{\mathbb{R}}$ with $\sigma_L^{-1}(t) < +\infty$. For

$t \in \mathbb{R}$ with $\sigma_L^{-1}(t) = +\infty$, then by assumption, $t \leq +\infty = (\sigma \circ \sigma_L^{-1})(t)$, and so this case holds as well. □

We now wish to promote monotone functions $\sigma : \mathbb{R} \to \mathbb{R}$ to functions on decorated endpoints $\mathbb{E}$ in such a way that monotonicity is preserved. A naive guess would be to preserve the endpoint types, mapping $t^{\pm} \mapsto \sigma(t)^{\pm}$ for every $t \in \mathbb{R}$ with $\sigma(t) \in \mathbb{R}$, but this quickly causes issues. In particular, monotonicity is not preserved when σ is non-injective. To illustrate, suppose that $x < y$ for $x, y \in \mathbb{R}$, but that $\sigma(x) = \sigma(y) \in \mathbb{R}$. Then $x^+ < y^-$, but we would have $\sigma(y)^- < \sigma(x)^+$. This situation is the reason behind the following definition.

Definition 2.43. Let $\sigma : \mathbb{R} \to \mathbb{R}$ be monotone. We define the *promotion of* σ to the map $\tilde{\sigma} : \mathbb{E} \to \mathbb{E}$ as follows:

- for $t \in \mathbb{R}$, if there exists $x \in \mathbb{R}$ with $t < x$ and $\sigma(t) = \sigma(x)$, we define $\tilde{\sigma}(t^{\pm}) := \sigma(t)^-$, otherwise set $\tilde{\sigma}(t^{\pm}) := \sigma(t)^{\pm}$, preserving the endpoint type;
- if the closure $\overline{\sigma}$ of σ has $\overline{\sigma}(-\infty) \in \mathbb{R}$, set $\tilde{\sigma}(-\infty) := \overline{\sigma}(-\infty)^-$, otherwise set $\tilde{\sigma}(-\infty) := -\infty$; and
- set $\tilde{\sigma}(+\infty) := +\infty$.

Note that if $\sigma(t) = t + \delta$ for $\delta \geq 0$, then $\tilde{\sigma}(t^{\pm}) = (t + \delta)^{\pm}$ by the bijectivity of σ, which agrees with Bauer and Lesnick's definition in [11].

Lemma 2.44. *If $\sigma : \mathbb{R} \to \mathbb{R}$ is monotone, then the map $\tilde{\sigma} : \mathbb{E} \to \mathbb{E}$ is monotone.*

Proof. Let $c \leq e \in \mathbb{E}$. We need to show that $\tilde{\sigma}(c) \leq \tilde{\sigma}(e)$. If either $\tilde{\sigma}(c) = -\infty$ or $\tilde{\sigma}(e) = +\infty$, we're done. By the definition of $\tilde{\sigma}$, if $\tilde{\sigma}(c) = +\infty$, then necessarily $c = e = +\infty$, and so we're done in this case. The case where $\tilde{\sigma}(e) = -\infty$ similarly yields $c = -\infty$, and so this case is done as well.

Now suppose that $\tilde{\sigma}(c), \tilde{\sigma}(e) \in \mathbb{R} \times \mathcal{D}$ and $c < e$. By the monotonicity of σ, we know that $(\sigma \circ \pi)(c)^- \leq (\sigma \circ \pi)(e)^-$, and so if $\tilde{\sigma}(c) = (\sigma \circ \pi)(c)^-$, we're done. Suppose that $\tilde{\sigma}(c) = (\sigma \circ \pi)(c)^+$, so at necessarily $c = \pi(c)^+ < e$ and thus $\pi(c) < \pi(e)$. By

Definition 2.43, we know that for any $c < x$, we have $\sigma(c) < \sigma(x)$, and so we must have $\tilde{\sigma}(c) < \tilde{\sigma}(e)$. □

We will have reason to define the notion of σ_L^{-1} and σ_R^{-1} on decorated endpoints. These definitions are ultimately a consequence of Propositions 3.8 and 3.12.

Definition 2.45. Let $\sigma : \mathbb{R} \to \mathbb{R}$ be monotone. We define the maps $\tilde{\sigma}_L^{-1} : \mathbb{E} \to \mathbb{E}$ and $\tilde{\sigma}_R^{-1} : \sigma(\mathbb{R}) \times \mathcal{D} \to \mathbb{E}$ as follows. Let $e \in \mathbb{E}$. We define

$$\tilde{\sigma}_L^{-1}(e) := \begin{cases} \sigma_L^{-1}(\pi e)^+ & \text{if } (\sigma \circ \sigma_L^{-1})(\pi e) < \pi e \\ \sigma_L^{-1}(\pi e)^- & \text{if } (\sigma \circ \sigma_L^{-1})(\pi e) \geq \pi e. \end{cases}$$

Let $e \in \sigma(\mathbb{R}) \times \mathcal{D}$. We define

$$\tilde{\sigma}_R^{-1}(e) := \begin{cases} \sigma_R^{-1}(\pi e)^+ & \text{if } (\sigma \circ \sigma_R^{-1})(\pi e) = \pi e \\ \sigma_R^{-1}(\pi e)^- & \text{if } (\sigma \circ \sigma_R^{-1})(\pi e) > \pi e. \end{cases}$$

Proposition 2.46. *Let $\sigma : \mathbb{R} \to \mathbb{R}$ be monotone and right-continuous. For $s, t \in \sigma(\mathbb{R})$ and $x \in \mathbb{R}$ with $s < x \leq t$, then $\tilde{\sigma}_R^{-1}(s^+) \leq \tilde{\sigma}_L^{-1}(x^\pm) \leq \tilde{\sigma}_R^{-1}(t^+)$.*

Proof. Since σ is right-continuous, we must have $\tilde{\sigma}_L^{-1}(e) = \sigma_L^{-1}(\pi e)^-$ for every $e \in \mathbb{E}$. Since by definition $\sigma_L^{-1}(y) \leq \sigma_R^{-1}(y)$ for every $y \in \sigma(\mathbb{R})$, then by the monotonicity of σ, necessarily $\tilde{\sigma}_L^{-1}(x^\pm) \leq \tilde{\sigma}_R^{-1}(t^+)$ whenever $x \leq t$ and $t \in \sigma(\mathbb{R})$. Now suppose that $s \in \sigma(\mathbb{R})$ and $s < x$ for some $x \in \mathbb{R}$. Then by the monotonicity of σ,

$$\begin{aligned} \sigma_R^{-1}(s) &= \sup\{y \in \mathbb{R} : \sigma(y) = s\} \\ &\leq \inf\{y \in \mathbb{R} : \sigma(y) > s\} \\ &\leq \inf\{y \in \mathbb{R} : \sigma(y) \geq x\} \\ &= \sigma_L^{-1}(x). \end{aligned}$$

If $(\sigma \circ \sigma_R^{-1})(s) > s$, then $\tilde{\sigma}_R^{-1}(s^+) = \sigma_R^{-1}(s)^- \leq \sigma_L^{-1}(x)^- = \tilde{\sigma}_L^{-1}(x^\pm)$. Suppose that $(\sigma \circ \sigma_R^{-1})(s) = s$. By the right-continuity of σ and the fact that $(\sigma \circ \sigma_R^{-1})(s) < x$, it follows that $\sigma_R^{-1}(s) < \sigma_L^{-1}(x)$, whence $\tilde{\sigma}_R^{-1}(s^+) = \sigma_R^{-1}(s)^+ < \sigma_L^{-1}(x)^- = \tilde{\sigma}_L^{-1}(x^\pm)$. □

To finish this section, we give an alternative characterization of the alignability of a translation pair.

Proposition 2.47. *If (τ, σ) is a translation pair and τ is right-continuous, then (τ, σ) is alignable if and only if there exists a bijective monotone function $\gamma : \mathbb{R} \to \mathbb{R}$ such that $\tau_L^{-1}(x) \leq \gamma(x) \leq \sigma(x)$ for every $x \in \mathbb{R}$.*

Proof. Let $\gamma : \mathbb{R} \to \mathbb{R}$ be bijective and monotone. Observe that

$$(\gamma \circ \tau)(x) \geq x = (\gamma \circ \gamma^{-1})(x) \iff \tau(x) \geq \gamma^{-1}(x) \tag{2.47a}$$

for all $x \in \mathbb{R}$, and that

$$(\sigma \circ \gamma^{-1})(x) \geq x = (\gamma \circ \gamma^{-1})(x) \iff \sigma(x) \geq \gamma(x) \tag{2.47b}$$

for all $x \in \mathbb{R}$.

Now suppose that $\sigma(x) < \tau_L^{-1}(x)$ for some $x \in \mathbb{R}$. Then $\sigma(x) < \inf\{s \in \mathbb{R} : x \leq \tau(s)\}$, and so $(\tau \circ \sigma)(x) < x$, which is a contradiction since (τ, σ) is a translation pair. Hence, $\tau_L^{-1}(x) \leq \sigma(x)$ for all $x \in \mathbb{R}$.

We now want to show that $\gamma^{-1} \leq \tau$ if and only if $\tau_L^{-1} \leq \gamma$. For the forward implication, suppose that $\gamma^{-1}(y) \leq \tau(y)$ for all $y \in \mathbb{R}$. Choose $x \in \mathbb{R}$ and let $y \in \mathbb{R}$ such that $x = \gamma^{-1}(y)$. Then

$$\begin{aligned} \tau_L^{-1}(x) &= \inf\{s \in \mathbb{R} : x \leq \tau(s)\} \\ &= \inf\{s \in \mathbb{R} : \gamma^{-1}(y) \leq \tau(s)\} \\ &\leq y \\ &= \gamma(x). \end{aligned}$$

For the reverse implication, suppose that $\tau_L^{-1}(x) \leq \gamma(x)$ for all $x \in \mathbb{R}$. Then

$$\tau_L^{-1}(\gamma^{-1}(x)) \leq \gamma(\gamma^{-1}(x)) = x.$$

Thus, if τ is right-continuous, Lemma 2.42 yields

$$\gamma^{-1}(x) \leq (\tau \circ \tau_L^{-1})[\gamma^{-1}(x)] \leq \tau(x).$$

Hence, we have shown that if τ is right-continuous, then $\gamma^{-1} \leq \tau$ if and only if $\tau_L^{-1} \leq \gamma$. Thus, to satisfy (2.47a) and (2.47b), it suffices to find a bijective, monotone function γ such that $\tau_L^{-1}(x) \leq \gamma(x) \leq \sigma(x)$ for every $x \in \mathbb{R}$. □

The author of this dissertation thanks Chuck Weibel for providing the following example of a translation pair that is not alignable.

Example 2.48. Let $\tau(t) = t + \lfloor t \rfloor$ for all $t \in \mathbb{R}$, where $\lfloor \cdot \rfloor : \mathbb{R} \to \mathbb{R}$ rounds every real number down to the nearest integer (commonly the notation used is $[\cdot] : \mathbb{R} \to \mathbb{R}$, but we will have reason to use the function $\lceil \cdot \rceil : \mathbb{R} \to \mathbb{R}$ as well). Then τ is right-continuous and τ_L^{-1} is constant from $2n$ to $2n+1$ for $n \in \mathbb{Z}$ and has unit slope otherwise. If we set $\sigma(t) := \tau_L^{-1}(t)$, then it is not hard to show that (τ, σ) is a translation pair. However, Proposition 2.47 forces $\gamma = \sigma$, and thus γ cannot be bijective since σ is not.

2.5 Induced Matchings on Barcodes and the Algebraic Stability Theorem

We now summarize the results from Bauer and Lesnick's work on induced matchings in [11] that show how a PFD persistence module morphism $\phi : V \to W$ guarantees the existence of a matching between the barcodes $\mathcal{B}_V$ and $\mathcal{B}_W$ satisfying certain properties. First we make clear what we mean by a multiset and its representation.

Definition 2.49. A *multiset* (S, n) is a set S together with a function $n : S \to \mathbb{Z}_+$ with values in the (strictly) positive integers, called a *counting function*. For any $x \in S$, the number $n(x)$ denotes the *multiplicity* of the element x in the multiset. We will often write S for the multiset, dropping explicit reference to the counting function.

Definition 2.50. Let (S, n) be a multiset. We define a *representation of a multiset* as the set

$$\mathrm{Rep}(S, n) = \{(x, k) \in S \times \mathbb{Z}_+ : k \leq n(x)\}.$$

For $(x, k) \in \mathrm{Rep}(S, n)$, the number k is called the *copy number* of x. We will sometimes write $\mathrm{Rep}(S)$ for $\mathrm{Rep}(S, n)$, dropping the reference to the counting function when it is clear.

Remark 2.51. Given a PFD persistence module V, we may view $\mathcal{B}_V$ as a multiset. For each $J \in \mathcal{B}_V$, we define $\mathbb{R}(J) := \{x \in \mathbb{R} : x \in J\}$. We then take

$$S = \{\mathbb{R}(J) \subseteq \mathbb{R} : J \in \mathcal{B}_V\},$$

the set of the unique intervals in the barcode of V, and define the counting function as

$$n(\mathbb{R}(J)) := |\{K \in \mathcal{B}_V : \mathbb{R}(K) = \mathbb{R}(J)\}|,$$

which counts the number of times the interval $\mathbb{R}(J)$ appears in the family of intervals $\mathcal{B}_V$. The counting function n satisfies Definition 2.49 since V is a PFD persistence module, and hence $n(I)$ must be finite for each interval I in S.

Representations of multisets give us a way to go from statements about multisets to statements about sets, and so we are now in a position to give the following definitions.

Definition 2.52. Let S and T be two sets. A *matching* $\mathcal{X} : S \nrightarrow T$ is given by a bijection $\mathcal{X} : S' \to T'$, where $S' \subseteq S$ and $T' \subseteq T$. We define $\mathrm{im}\, \mathcal{X} := T'$ and $\mathrm{coim}\, \mathcal{X} := S'$. We say that elements in $S \setminus S'$ and $T \setminus T'$ are *unmatched by* $\mathcal{X}$.

We say a matching $\mathcal{X} : S \nrightarrow T$ is *left-perfect* (a monomorphism) if $\mathrm{coim}\, \mathcal{X} = S$, *right-perfect* (an epimorphism) if $\mathrm{im}\, \mathcal{X} = T$, and *perfect* (an isomorphism) if it is both left-perfect and right-perfect. If $\mathcal{X} : S \nrightarrow T$, we say that $\mathcal{X}^{-1} : T \nrightarrow S$ is the *reverse matching of* $\mathcal{X}$ and define it via the inverse of the bijection $\mathcal{X} : S' \to T'$.

We say that a finite, totally ordered set S with N elements is an *enumerated set* if it is equipped with an order-preserving bijection from $\{1, ..., N\}$ to S. If S is an enumerated set via the bijection $\{1, ..., N\} \ni i \mapsto s_i \in S$, we say that s_i is the i^{th} element of S. If $S = \{s_i\}_{i=1}^N$ and $T = \{t_j\}_{j=1}^M$ are two enumerated sets, the *minimal matching* $\mathcal{X}$ from S to T is the unique matching such that $\mathcal{X}(s_i) = t_i$ for every $i = 1, ..., \min\{N, M\}$.

A matching $\mathcal{X} : S \nrightarrow T$ can be seen as a relation on $S \times T$ given by

$$\mathrm{Gr}(\mathcal{X}) := \left\{(s, \mathcal{X}(s)) \in S \times T | s \in S'\right\},$$

where $\mathrm{Gr}(\mathcal{X})$ is called the *graph of* $\mathcal{X}$. We define the *composition of two matchings* $\mathcal{X} : S \nrightarrow T$ and $\mathcal{X}' : T \nrightarrow U$ as the relation

$$\mathcal{X}' \circ \mathcal{X} = \{(s, u) | (s, t) \in \mathrm{Gr}(\mathcal{X}), (t, u) \in \mathrm{Gr}(\mathcal{X}') \text{ for some } t \in T\}.$$

The following definitions, inspired by Bauer and Lesnick [11], will enable us to describe a particular set of matchings on representations of barcodes.

Definition 2.53. Let $\mathcal{B}$ be a barcode. Given $b, d \in \mathbb{E}$, we define

$$\langle b, \cdot \rangle_{\mathcal{B}} := \{(\langle b', d' \rangle, j) \in \mathrm{Rep}(\mathcal{B}) : b' = b\} \subseteq \mathrm{Rep}(\mathcal{B});$$

$$\langle \cdot, d \rangle_{\mathcal{B}} := \{(\langle b', d' \rangle, j) \in \mathrm{Rep}(\mathcal{B}) : d' = d\} \subseteq \mathrm{Rep}(\mathcal{B}).$$

We also define the sets

$$\ell(\mathcal{B}) := \{b \in \mathbb{E} : (\langle b, d \rangle, k) \in \mathrm{Rep}(\mathcal{B})\}$$

and

$$r(\mathcal{B}) := \{d \in \mathbb{E} : (\langle b, d \rangle, k) \in \mathrm{Rep}(\mathcal{B})\},$$

i.e. the sets of unique left and right endpoints of intervals in $\mathcal{B}$.

Remark 2.54. Notice that if $\mathcal{B}$ is a barcode, then

$$\mathrm{Rep}(\mathcal{B}) = \bigcup_{b \in \ell(\mathcal{B})} \langle b, \cdot \rangle_{\mathcal{B}} = \bigcup_{d \in r(\mathcal{B})} \langle \cdot, d \rangle_{\mathcal{B}},$$

where the unions are disjoint.

Definition 2.55. Let $\mathcal{B}$ be a barcode. The *left-handed ordering* on $\mathrm{Rep}(\mathcal{B})$ is the one

inherited from the dictionary ordering on $\mathbb{E} \times \mathbb{E} \times \mathbb{Z}$ under the bijective correspondence

$$\mathrm{Rep}(\mathcal{B}) \ni (\langle b, d\rangle, k) \mapsto (b, -d, k) \in \mathbb{E} \times \mathbb{E} \times \mathbb{Z},$$

where multiplication of a decorated endpoint is defined as in Definition 2.19. Similarly, the *right-handed ordering* on $\mathrm{Rep}(\mathcal{B})$ is the one inherited from the dictionary ordering on $\mathbb{E} \times \mathbb{E} \times \mathbb{Z}$ under the bijective correspondence

$$\mathrm{Rep}(\mathcal{B}) \ni (\langle b, d\rangle, k) \mapsto (d, b, k) \in \mathbb{E} \times \mathbb{E} \times \mathbb{Z}.$$

If $\mathrm{Rep}(\mathcal{B})$ is endowed with the left-handed ordering, then we write $\mathrm{Rep}^{\ell}(\mathcal{B})$, and if it is endowed with the right-handed ordering, we write $\mathrm{Rep}^{r}(\mathcal{B})$. Thus, for the left-handed (right-handed) ordering on $\mathrm{Rep}(\mathcal{B})$ and a fixed left-hand endpoint $b \in \mathbb{E}$ (right-hand endpoint $d \in \mathbb{E}$), the longest intervals with the lowest copy numbers are listed first.

The sets given in Definition 2.53 with the left-handed and right-handed orderings defined in Definition 2.55 serve to set us up to establish an analog of what Bauer and Lesnick refer to as the *canonical injections between barcodes* [11]. In their construction, intervals in barcode representations beginning (ending) with the same decorated endpoint are first sorted longest to shortest, with increasing copy number. Then, the injection on representations matches the barcodes in the two barcode representations greedily, preserving this ordering. The following definitions make this precise.

Definition 2.56. Let V and W be PFD persistence modules. The matching

$$\lfloor \mathcal{X}_V^W \rfloor : \mathrm{Rep}^{\ell}(\mathcal{B}_V) \twoheadrightarrow \mathrm{Rep}^{\ell}(\mathcal{B}_W)$$

is defined as follows. For each $b \in \ell(\mathcal{B}_V)$, enumerate the sets $\langle b, \cdot\rangle_{\mathcal{B}_V} \subseteq \mathrm{Rep}^{\ell}(\mathcal{B}_V)$ and $\langle b, \cdot\rangle_{\mathcal{B}_W} \subseteq \mathrm{Rep}^{\ell}(\mathcal{B}_W)$ according to their left-handed orderings, and define the restriction of the matching $\lfloor \mathcal{X}_V^W \rfloor$ to $\langle b, \cdot\rangle_{\mathcal{B}_V}$ to be the minimal matching between $\langle b, \cdot\rangle_{\mathcal{B}_V}$ and $\langle b, \cdot\rangle_{\mathcal{B}_W}$ that respects the enumerations. By Remark 2.54, this defines a matching from $\mathrm{Rep}(\mathcal{B}_V)$ to $\mathrm{Rep}(\mathcal{B}_W)$.

The matching

$$\lceil \mathcal{X}_V^W \rceil : \mathrm{Rep}^r(\mathcal{B}_V) \nrightarrow \mathrm{Rep}^r(\mathcal{B}_W)$$

is defined similarly. For each $d \in r(\mathcal{B}_V)$, enumerate the sets $\langle \cdot, d\rangle_{\mathcal{B}_V} \subseteq \mathrm{Rep}^r(\mathcal{B}_V)$ and $\langle \cdot, d\rangle_{\mathcal{B}_W} \subseteq \mathrm{Rep}^r(\mathcal{B}_W)$ according to their right-handed orderings, and define the restriction of the matching $\lceil \mathcal{X}_V^W \rceil$ to $\langle \cdot, d\rangle_{\mathcal{B}_V}$ to be the minimal matching between $\langle \cdot, d\rangle_{\mathcal{B}_V}$ and $\langle \cdot, d\rangle_{\mathcal{B}_W}$ that respects the enumerations. By Remark 2.54, this defines a matching from $\mathrm{Rep}(\mathcal{B}_V)$ to $\mathrm{Rep}(\mathcal{B}_W)$.

Since minimal matchings on enumerated sets are unique, so are $\lfloor \mathcal{X}_V^W \rfloor$ and $\lceil \mathcal{X}_V^W \rceil$. Thus, we say that $\lfloor \mathcal{X}_V^W \rfloor$ is the *left minimal matching* from $\mathrm{Rep}(\mathcal{B}_V)$ to $\mathrm{Rep}(\mathcal{B}_W)$ and that $\lceil \mathcal{X}_V^W \rceil$ is the *right minimal matching* from $\mathrm{Rep}(\mathcal{B}_V)$ to $\mathrm{Rep}(\mathcal{B}_W)$.

Definition 2.57. Let V and W be PFD persistence modules with barcode representations $\mathrm{Rep}(\mathcal{B}_V)$ and $\mathrm{Rep}(\mathcal{B}_W)$. We say that a matching $\mathcal{X} : \mathrm{Rep}(\mathcal{B}_V) \nrightarrow \mathrm{Rep}(\mathcal{B}_W)$ is

- *left-admissible* if it is right-perfect and $\mathcal{X}(\langle b, d\rangle, k) = (\langle b', d'\rangle, j)$ implies $b = b'$ and $d' \leq d$ for every $(\langle b', d'\rangle, j) \in \mathrm{Rep}(\mathcal{B}_W)$.
- *right-admissible* if it is left-perfect and $\mathcal{X}(\langle b, d\rangle, k) = (\langle b', d'\rangle, j)$ implies $d = d'$ and $b' \leq b$ for every $(\langle b, d\rangle, k) \in \mathrm{Rep}(\mathcal{B}_V)$.

Let U be a PFD persistence module with barcode representation $\mathrm{Rep}(\mathcal{B}_U)$. We say that a matching $\mathcal{X} : \mathrm{Rep}(\mathcal{B}_V) \nrightarrow \mathrm{Rep}(\mathcal{B}_W)$ is *admissible via Rep*$(\mathcal{B}_U)$ if it can be expressed as the composition $\mathcal{X} = \mathcal{X}_r \circ \mathcal{X}_\ell$ where $\mathcal{X}_\ell : \mathrm{Rep}(\mathcal{B}_V) \nrightarrow \mathrm{Rep}(\mathcal{B}_U)$ is a left-admissible matching and $\mathcal{X}_r : \mathrm{Rep}(\mathcal{B}_U) \nrightarrow \mathrm{Rep}(\mathcal{B}_W)$ is a right-admissible matching. (Note that this matching is defined since $\mathcal{X}_\ell$ is surjective and $\mathcal{X}_r$ is injective.)

Definition 2.58. Let $\phi : V \to W$ be a persistence module morphism. A matching $\mathcal{X} : \mathrm{Rep}(\mathcal{B}_V) \nrightarrow \mathrm{Rep}(\mathcal{B}_W)$ is called an *admissible matching for* ϕ if it is an admissible matching via $\mathrm{Rep}(\mathcal{B}_{\mathrm{im}\,\phi})$.

The following theorem is a rephrasing of [11, Theorem 4.2], which we state without proof.

Theorem 2.59. *Let V and W be PFD persistence modules.*

(i) If there exists an epimorphism $\phi : V \to W$, then $\lfloor \mathcal{X}_V^W \rfloor : Rep(\mathcal{B}_V) \twoheadrightarrow Rep(\mathcal{B}_W)$ is left-admissible.

(ii) If there exists a monomorphism $\psi : V \to W$, then $\lceil \mathcal{X}_V^W \rceil : Rep(\mathcal{B}_V) \twoheadrightarrow Rep(\mathcal{B}_W)$ is right-admissible.

Corollary 2.60. *Let $\phi : V \to W$ be a PFD persistence module morphism. The matching $\lceil \mathcal{X}_{\text{im }\phi}^W \rceil \circ \lfloor \mathcal{X}_V^{\text{im }\phi} \rfloor : Rep(\mathcal{B}_V) \twoheadrightarrow Rep(\mathcal{B}_W)$ is an admissible matching for ϕ.*

Proof. Every persistence module morphism $\phi\colon V \to W$ can be factored as a composition of injective and surjective persistence module morphisms:

$$V \xrightarrow{q_\phi} \text{im } \phi \xhookrightarrow{j_\phi} W.$$

By Theorem 2.59, the matching $\lfloor \mathcal{X}_V^{\text{im }\phi} \rfloor : \text{Rep}(\mathcal{B}_V) \twoheadrightarrow \text{Rep}(\mathcal{B}_{\text{im }\phi})$ is left-admissible and the matching $\lceil \mathcal{X}_{\text{im }\phi}^W \rceil : \text{Rep}(\mathcal{B}_{\text{im }\phi}) \twoheadrightarrow \text{Rep}(\mathcal{B}_W)$ is right-admissible. By definition, $\lceil \mathcal{X}_{\text{im }\phi}^W \rceil \circ \lfloor \mathcal{X}_V^{\text{im }\phi} \rfloor$ is an admissible matching for ϕ. □

Definition 2.61. Let $\phi : V \to W$ be a PFD persistence module morphism. The *matching witnessed by* ϕ is defined to be the minimal admissible matching given by the composition $\mathcal{X}_\phi := \lceil \mathcal{X}_{\text{im }\phi}^W \rceil \circ \lfloor \mathcal{X}_V^{\text{im }\phi} \rfloor$.

Remark 2.62. Let $\phi : V \to W$ be a PFD persistence module morphism. If ϕ is an epimorphism, then $\mathcal{X}_\phi = \lceil \mathcal{X}_W^W \rceil \circ \lfloor \mathcal{X}_V^W \rfloor = \lfloor \mathcal{X}_V^W \rfloor$ is left-admissible. If ϕ is a monomorphism, then $\mathcal{X}_\phi = \lceil \mathcal{X}_V^W \rceil \circ \lfloor \mathcal{X}_V^V \rfloor = \lceil \mathcal{X}_V^W \rceil$ is right-admissible.

Proposition 2.63. *Let U, V, and W be PFD persistence modules and suppose that $\mathcal{X} : Rep(\mathcal{B}_V) \twoheadrightarrow Rep(\mathcal{B}_W)$ is an admissible matching via $Rep(\mathcal{B}_U)$. If $\mathcal{X}(\langle b, d\rangle, k) = (\langle b', d'\rangle, j)$, then $b' \leq b < d' \leq d$.*

Proof. Since $\mathcal{X} : \text{Rep}(\mathcal{B}_V) \twoheadrightarrow \text{Rep}(\mathcal{B}_W)$ is an admissible matching via $\text{Rep}(\mathcal{B}_U)$, there exists a left-admissible matching $\mathcal{X}_\ell : \text{Rep}(\mathcal{B}_V) \twoheadrightarrow \text{Rep}(\mathcal{B}_U)$ and a right-admissible matching $\mathcal{X}_r : \text{Rep}(\mathcal{B}_U) \twoheadrightarrow \text{Rep}(\mathcal{B}_W)$ such that $\mathcal{X} = \mathcal{X}_r \circ \mathcal{X}_\ell$. Thus, there exists some

$(\langle b, d'\rangle, k) \in \mathrm{Rep}(\mathcal{B}_U)$ such that $\mathcal{X}_\ell(\langle b, d\rangle, j) = (\langle b, d'\rangle, k)$ and $\mathcal{X}_r(\langle b, d'\rangle, k) = (\langle b', d'\rangle, l)$. Since $\langle b, d'\rangle$ is an interval, we know that $b < d'$. Hence, by the definition of left-admissible and right-admissible matchings, we see that $b' \leq b < d' \leq d$. □

Corollary 2.64. *Let $\phi : V \to W$ be a PFD persistence module morphism. If $\mathcal{X}_\phi\langle b, d\rangle = \langle b', d'\rangle$, then $b' \leq b < d' \leq d$.*

Proof. By Definition 2.61, we see that $\mathcal{X}_\phi = \lceil \mathcal{X}^W_{\mathrm{im}\,\phi} \rceil \circ \lfloor \mathcal{X}_V^{\mathrm{im}\,\phi} \rfloor$. By Corollary 2.60, it follows that $\mathcal{X}_\phi$ is an admissible matching for ϕ, and hence it is an admissible matching via $\mathrm{Rep}(\mathcal{B}_{\mathrm{im}\,\phi})$. Thus, Proposition 2.63 guarantees that $\mathcal{X}_\phi\langle b, d\rangle = \langle b', d'\rangle$ yields $b' \leq b < d' \leq d$. □

The following is proved by Bauer and Lesnick in [11] and gives a condition of functoriality for barcode matchings witnessed by morphisms of PFD persistence modules.

Proposition 2.65. *Let U, V, and W be PFD persistence modules and let $\phi : U \to V$ and $\psi : V \to W$ be morphisms. If ϕ and ψ are either both injective or both surjective, then $\mathcal{X}_{\psi\circ\phi} = \mathcal{X}_\psi \circ \mathcal{X}_\phi$.*

We remark that the construction of the minimal matchings guarantees the partial functoriality property exhibited in Proposition 2.65. While this proposition is used in the proof of the Non-uniform Induced Matching Theorem, there are perhaps other ways to achieve the commutativity in the matchings as required by the proof. In particular, the task is to move from a diagram of commutative persistence module morphisms to a commutative diagram of barcode representations, but it is not explicitly necessary that the minimal matchings are used.

Throughout the remainder of this chapter and the next, since we will be working with the uniquely determined minimal matchings witnessed by persistence module morphisms, we will often suppress the reference to both the representation and the copy number in what follows, expressing $(\langle b, d\rangle, k) \in \mathrm{Rep}(\mathcal{B}_V)$ as $\langle b, d\rangle \in \mathcal{B}_V$. Additionally, we will often write $\mathcal{X}\langle b, d\rangle$ instead of $\mathcal{X}(\langle b, d\rangle, k)$ to denote a matching applied to an element of a barcode representation. However, when it is required, we will use the more explicit notation.

To complete this section, we consider intervals in the barcode of a PFD persistence module exceeding a certain minimal length. Let $\mathcal{B}$ be a barcode and let $\delta \geq 0$. We define the set

$$\mathrm{Rep}(\mathcal{B})^{\delta} := \{(\langle b, d\rangle, k) \in \mathrm{Rep}(\mathcal{B}) : b + \delta < d\}.$$

Bauer and Lesnick refer to the following theorem as the *induced matching theorem* [11]. We state it here for the convenience of the reader as a comparison to the results we present in Chapter 3.

Induced Matching Theorem 2.66. *Let $\phi : V \to W$ be a PFD persistence module morphism, let $\delta \geq 0$. Suppose that $\mathcal{X}_{\phi}\langle b, d\rangle = \langle b', d'\rangle$.*

(i) If coker ϕ *is δ-trivial, then $Rep(\mathcal{B}_W)^{\delta} \subseteq \operatorname{im} \mathcal{X}_{\phi}$ and $b' \leq b \leq b' + \delta$.*

(ii) If ker ϕ *is δ-trivial, then $Rep(\mathcal{B}_V)^{\delta} \subseteq \operatorname{coim} \mathcal{X}_{\phi}$ and $d - \delta \leq d' \leq d$.*

This theorem can be thought of as the *uniform* induced matching theorem, since a uniform bound of length δ is placed on the sizes of the intervals of the barcodes in $\ker \phi$ and coker ϕ. In [11], Bauer and Lesnick use the previous theorem to prove the following version of the *algebraic stability theorem*, establishing the following important relationship between the barcodes of two δ-interleaved PFD persistence modules.

Algebraic Stability Theorem 2.67. *Let (V, φ_V) and (W, φ_W) be δ-interleaved PFD persistence modules. Then there exists a matching $\mathcal{X} : Rep(\mathcal{B}_V) \nrightarrow Rep(\mathcal{B}_W)$ that satisfies the following. If $\mathcal{X}\langle b, d\rangle = \langle b', d'\rangle$, then*

$$b' - \delta \leq b \leq b' + \delta,$$

$$d - \delta \leq d' \leq d + \delta,$$

and unmatched intervals in either $Rep(\mathcal{B}_V)$ or $Rep(\mathcal{B}_W)$ have length at most 2δ.

In the next chapter, we show how the notion of and interleaving given in Definition 2.27 generalizes Theorem 2.67.

Chapter 3

The Algebraic Stability Theorem for Generalized Interleavings

Using the ideas from Chapter 2, we now state and prove the main results of this chapter: generalizations of the Induced Matching Theorem and the Algebraic Stability Theorem (Theorems 2.66 and 2.67) in the context of the more general interleaving framework of Bubenik, et al. [12] via Definition 2.27. For the remainder of this chapter, we consider (V, φ_V) and (W, φ_W) to be PFD persistence modules.

We first establish an analog to the set $\text{Rep}(\mathcal{B})^\delta$ for more general translation maps. Recall from Definition 2.19 that $\pi : \mathbb{E} \to \overline{\mathbb{R}}$ is the map that forgets the decoration of a decorated endpoint. The reader may want to recall the following additional maps related to a monotone function $\sigma : \mathbb{R} \to \mathbb{R}$:

- $\overline{\sigma} : \overline{\mathbb{R}} \to \overline{\mathbb{R}}$, the extension of σ to the extended real numbers (Definition 2.40).
- $\tilde{\sigma} : \mathbb{E} \to \mathbb{E}$, the promotion of σ to decorated endpoints (Definition 2.43).
- $\sigma_L^{-1} : \mathbb{R} \to \overline{\mathbb{R}}$ and $\sigma_R^{-1} : \sigma(\mathbb{R}) \to \overline{\mathbb{R}}$, the left and right generalized inverses of σ (Definition 2.41).
- $\tilde{\sigma}_L^{-1} : \mathbb{E} \to \mathbb{E}$ and $\tilde{\sigma}_R^{-1} : \sigma(\mathbb{R}) \times \mathcal{D} \to \mathbb{E}$, the analagous left and right generalized inverses of $\tilde{\sigma}$ (Definition 2.45).

Let $\mathcal{B}$ be a barcode and let $\sigma : \mathbb{R} \to \mathbb{R}$ be a translation map (Definition 2.20). We define the following barcode representation subset:

$$\text{Rep}(\mathcal{B})^\sigma := \{(\langle b, d\rangle, k) \in \text{Rep}(\mathcal{B}) : \tilde{\sigma}(b) < d\}.$$

Thus, intervals in $\text{Rep}(\mathcal{B})^\sigma$ are those that are non-empty after moving the left-hand

endpoint forward by $\tilde{\sigma}$, taking endpoint decorations into consideration. The following partition of decorated endpoints will be useful.

Definition 3.1. Let $\sigma : \mathbb{R} \to \mathbb{R}$ be monotone. We say that a decorated endpoint $e \in \mathbb{E}$ has property $B(\sigma)$ if $e = \pi(e)^+$ and $\pi(e) \in \sigma(\mathbb{R})$. Otherwise, we will say that e has property $A(\sigma)$.

In particular, if a decorated endpoint $e \in \mathbb{R} \times \mathcal{D}$ has property $B(\sigma)$ and σ is right-continuous, then $e = \pi(e)^+$ and $\pi(e) = (\sigma \circ \sigma_L^{-1})(\pi e)$. We will sometimes have reason to refer to the set of decorated endpoints with property $A(\sigma)$ by simply writing $A(\sigma) \subseteq \mathbb{E}$, and similarly with $B(\sigma) \subseteq \mathbb{E}$.

Proposition 3.8, below, shows how the two types of endpoints in Definition 3.1 determine the bounds on the matching given in the Non-uniform Induced Matching Theorem. Lemma 3.9 gives the motivation behind Definitions 3.3 and 3.6. The following remarks will be used in the proofs of both Proposition 3.5 and 3.8.

Remark 3.2. Recall from the Structure Theorem, for every PFD persistence module morphism (V, φ_V) there exists an isomorphism to a direct sum of interval persistence modules

$$\Phi : V \to \bigoplus_{J \in \mathcal{B}_V} I_J.$$

For an interval $(\langle c, e\rangle, k) \in \mathrm{Rep}(\mathcal{B}_V)$, we write $I_{(\langle c,e\rangle,k)}$ for the interval persistence module in the direct sum corresponding to interval $(\langle c, e\rangle, k)$. Define

$$I^V_{(\langle c,e\rangle,k)} := \Phi^{-1}[I_{(\langle c,e\rangle,k)}] \subseteq V,$$

the inverse image of the interval summand $I_{(\langle c,e\rangle,k)}$ under Φ. Then

$$V = \bigoplus_{(\langle c,e\rangle,k)\in \mathrm{Rep}(\mathcal{B}_V)} I^V_{(\langle c,e\rangle,k)}.$$

From this point forward we will drop the copy number so notation is not cluttered, writing just $I^V_{\langle c,e\rangle}$, understanding that it designates a particular summand. Let $s, t \in \mathbb{R}$

with $s \leq t$. Since Φ is a persistence module isomorphism, we have

$$\Phi_t^{-1} \circ \varphi_{\Phi(V)}(s,t) = \varphi_V(s,t) \circ \Phi_s^{-1},$$

and thus we see

$$\Phi_t^{-1} \circ \varphi_{\Phi(V)}(s,t)|_{(I_{\langle c,e\rangle})_s} = \varphi_V(s,t) \circ \Phi_s^{-1}|_{(I_{\langle c,e\rangle})_s} = \varphi_V(s,t)|_{\left(I^V_{\langle c,e\rangle}\right)_s}.$$

We will need to define transition maps indexed by decorated real numbers. Let (V, φ_V) be a persistence module. For any finite interval corresponding to $\langle s,t\rangle \in \mathbb{E} \times \mathbb{E}$, we can consider the associated transition map

$$\varphi_V(\pi(s), \pi(t)).$$

However, to keep the notation uncluttered, we will write $\varphi_V(s,t)$, where the projection via π is understood.

Here are two constructions which will be used in the proof of Theorem 3.10.

Definition 3.3. Let (W, φ_W) be a persistence module and let $\sigma : \mathbb{R} \to \mathbb{R}$ be a translation map. To each $t \in \mathbb{R}$ we associate the set

$$S(t) := \{x \in \mathbb{R} : \sigma(x) \leq t\}.$$

We define a persistence submodule $(W^\sigma, \varphi_{W^\sigma})$ of W via

$$W_t^\sigma := \bigcup_{x \in S(t)} \text{im } \varphi_W(x,t) \quad \text{for every} \quad t \in \mathbb{R}.$$

Clearly W_t^σ is a subspace of W_t for every t since if $x_1 \leq x_2 \in S(t)$, then im $\varphi_W(x_1,t) \subseteq$ im $\varphi_W(x_2,t)$.

Proposition 3.4. *Let (W, φ_W) be a persistence module and let $\sigma : \mathbb{R} \to \mathbb{R}$ be a translation map. Then $(W^\sigma, \varphi_{W^\sigma})$ is a persistence submodule of W.*

Proof. To show that W^σ is a persistence submodule of W, by Lemma 2.5, we must

show that $\varphi_W(s,t)|_{W_s^\sigma} \subseteq W_t^\sigma$ for all $s \leq t$. To this end, let $s \leq t$ and let $\alpha \in W_s^\sigma$. By the definition of W_s^σ, there exists some $x \in S(s)$ such that $\alpha \in \operatorname{im} \varphi_W(x,s)$. Thus, there exists some $\beta \in W_x$ such that $\varphi_W(x,s)(\beta) = \alpha$, and so

$$\varphi_W(s,t)(\alpha) = \varphi_W(s,t)[\varphi_W(x,s)(\beta)] = \varphi_W(x,t)(\beta),$$

yielding that $\varphi_W(s,t)(\alpha) \in \operatorname{im} \varphi_W(x,t)$. Since σ is a translation map and $x \in S(s)$, we have $x \leq \sigma(x) \leq s \leq t$, so $x \in S(t)$ as well. It follows that $\operatorname{im} \varphi_W(x,t) \subseteq W_t^\sigma$, so we have proved that $\varphi_W(s,t)(\alpha) \in W_t^\sigma$. Since α was chosen arbitrarily, we have shown that $\operatorname{im} \varphi_W(s,t)|_{W_s^\sigma} \subseteq W_t^\sigma$. Hence, W^σ is a persistence submodule of W. □

Proposition 3.5. *Let (W, φ_W) be a persistence module and let $\sigma : \mathbb{R} \to \mathbb{R}$ be a translation map. Then the minimal matching $\mathcal{X}_{j_\sigma} : Rep(\mathcal{B}_{W^\sigma}) \nrightarrow Rep(\mathcal{B}_W)$ witnessed by $j_\sigma : W^\sigma \hookrightarrow W$ has $\operatorname{im} \mathcal{X}_{j_\sigma} = Rep(\mathcal{B}_W)^\sigma$. Moreover, if $\mathcal{X}_{j_\sigma}\langle b, d\rangle = \langle b', d'\rangle$, then $b = \tilde{\sigma}(b')$.*

Proof. We will first prove a preliminary result by examining the intersection of W^σ with each of the interval persistence submodules of W corresponding to the intervals in $\mathcal{B}_W$ (see Remark 3.2). In particular, we will show that

$$W^\sigma = \bigoplus_{J \in \mathcal{B}_W} \left(W^\sigma \cap I_J^W\right). \tag{3.5a}$$

First, it is clear that the right side of (3.5a) is a persistence submodule of W^σ. We will show that W^σ is a persistence submodule of the right-hand side by showing that

$$\operatorname{im} \varphi_W(s,t) \subseteq \bigoplus_{J \in \mathcal{B}_W} \operatorname{im} \varphi_W(s,t) \cap (I_J^W)_t$$

for every $s, t \in \mathbb{R}$ with $s \leq t$. To this end, let $s \leq t$. The following diagram commutes, where Φ is the isomorphism giving the direct sum decomposition of W into interval persistence modules, and the $\varphi_{\Phi(W)}$ are the linear maps in the interval decomposition.

$$\begin{array}{ccc} \bigoplus_{J\in\mathcal{B}_W}(I_J)_s & \xrightarrow{\varphi_{\Phi(W)}(s,t)} & \bigoplus_{J\in\mathcal{B}_W}(I_J)_t \\ \downarrow{\scriptstyle \Phi_s^{-1}} & & \downarrow{\scriptstyle \Phi_t^{-1}} \\ W_s = \bigoplus_{J\in\mathcal{B}_W}\Phi_s^{-1}[(I_J)_s] & \xrightarrow{\varphi_W(s,t)} & \bigoplus_{J\in\mathcal{B}_W}\Phi_t^{-1}[(I_J)_t] = W_t \end{array} \tag{3.5b}$$

In particular, we see that $\operatorname{im}\varphi_W(s,t)|_{(I_J^W)_s} \subseteq (I_J^W)_t$. Thus, we have

$$\begin{aligned} \operatorname{im}\varphi_W(s,t) &= \bigoplus_{J\in\mathcal{B}_W} \operatorname{im}\varphi_W(s,t)|_{(I_J^W)_s} \\ &= \bigoplus_{J\in\mathcal{B}_W} \operatorname{im}\varphi_W(s,t)|_{(I_J^W)_s} \cap (I_J^W)_t \\ &\subseteq \bigoplus_{J\in\mathcal{B}_W} \operatorname{im}\varphi_W(s,t) \cap (I_J^W)_t, \end{aligned}$$

which is what we wanted to show. Thus, taking the union over all $x \in S(t)$, we have succeeded in proving equation (3.5a).

Now we will show that for every $\langle c,e\rangle \in \operatorname{Rep}(\mathcal{B}_W)$, if $\tilde{\sigma}(c) < e$, then $(W^\sigma \cap I^W_{\langle c,e\rangle}) \cong I_{\langle\tilde{\sigma}(c),e\rangle}$, otherwise $(W^\sigma \cap I^W_{\langle c,e\rangle})$ is the trivial persistence module. By Proposition 2.16, a persistence submodule of an interval persistence module is isomorphic to an interval persistence module, and so this makes sense to try to show. Let $(\langle c,e\rangle, k) \in \operatorname{Rep}(\mathcal{B}_W)$ and denote by $I^W_{\langle c,e\rangle}$ the corresponding interval persistence submodule of W. For $t \in \mathbb{R}$, we will show that $(I^W_{\langle c,e\rangle} \cap W^\sigma)_t \neq 0$ if and only if $\tilde{\sigma}(c) \leq t^- < e$. By the definitions of W^σ and $I^W_{\langle c,e\rangle}$, we know that $(I^W_{\langle c,e\rangle} \cap W^\sigma)_t \neq 0$ if and only if $t^- < e$ and there exists some $x \in S(t)$ such that $c < x^+$. Let $t \in \mathbb{R}$ such that $t^- < e$. We will show that there exists some $x \in S(t)$ with $c < x^+$ if and only if $\tilde{\sigma}(c) \leq t^-$.

For the forwards implication, suppose there exists some $x \in S(t)$ with $c < x^+$. Then $c < x^+$ implies that $\tilde{\sigma}(c) \leq \tilde{\sigma}(x^+)$ by the monotonicity of $\tilde{\sigma}$ (Lemma 2.44). But $x \in S(t)$ implies that $\sigma(x) \leq t$ which yields $\sigma(x)^- \leq t^-$. Thus, if we can show that $\tilde{\sigma}(c) \leq \sigma(x)^-$, it will follow that $\tilde{\sigma}(c) \leq t^-$. Suppose first that $\tilde{\sigma}(c) = (\pi\circ\tilde{\sigma})(c)^-$. Then $(\pi\circ\tilde{\sigma})(c)^- \leq \tilde{\sigma}(x^+)$. By the definition of $\tilde{\sigma}$, this implies that $(\pi\circ\tilde{\sigma})(c)^- \leq \sigma(x)^-$, and so this case is done. Now suppose that $\tilde{\sigma}(c) = (\pi\circ\tilde{\sigma})(c)^+$, so that it is necessarily

the case that $c = c^+$ by Definition 2.43, and by the same definition, we know there does not exist a $y > \pi c$ such that $\sigma(\pi c) = \sigma(y)$. Recall that since $\pi(c)^+ < x^+$, by the definition of $\mathbb{E}$ we have $\pi c < \pi x$. By Definition 2.43 again, it must be the case that $(\pi \circ \tilde{\sigma})(c) < (\pi \circ \tilde{\sigma})(x^+)$, and thus we have $\tilde{\sigma}(c) \leq \sigma(x)^-$ in this case as well.

For the reverse implication, suppose that $\tilde{\sigma}(c) \leq t^-$. We will show that there exists some $x \in S(t)$ with $c < x^+$. By our assumption and the definition of $\tilde{\sigma}$, we know that $\sigma(\pi c) \leq t$, and so $\pi c \in S(t)$. If $c = \pi(c)^-$, we can put $x = \pi c \in S(t)$ and we achieve $c = \pi(c)^- < \pi(c)^+ = x^+$ Now suppose that $c = \pi(c)^+$. First we examine the case when $\tilde{\sigma}(c) = (\pi \circ \tilde{\sigma})(c)^-$. By Definition 2.43, there must exist a $y > \pi c$ such that $\sigma(\pi c) = \sigma(y) \leq t$, and so $y \in S(t)$ and $c < y^+$, so we're done in this case. If, on the other hand, $c = c^+$ and $\tilde{\sigma}(c) = (\pi \circ \tilde{\sigma})(c)^+$, then again by Definition 2.43, there cannot exist a $y > \pi c$ such that $\sigma(\pi c) = \sigma(y)$. Also, $\tilde{\sigma}(c) = (\pi \circ \tilde{\sigma})(c)^+ \leq t^-$ by assumed, and so $(\pi \circ \tilde{\sigma})(c)^+ < t$. By the right-continuity of σ, we thus have $\lim_{y \to \pi c, y > \pi c} \sigma(\pi c) < t$, and so there exists some $y \in \mathbb{R}$ with $\sigma(y) < t$, whereby $y \in S(t)$ and $c < y^+$, which is what we desired to show.

We will now show that $\mathcal{X}_{j_\sigma} : \mathrm{Rep}(\mathcal{B}_{W^\sigma}) \nrightarrow \mathrm{Rep}(\mathcal{B}_W)$ is left-perfect. By Definition 2.61 and the fact that j_σ is a monomorphism, $\mathcal{X}_{j_\sigma} = \lceil \mathcal{X}^W_{W^\sigma} \rceil \circ \lfloor \mathcal{X}^{W^\sigma}_{W^\sigma} \rfloor = \lceil \mathcal{X}^W_{W^\sigma} \rceil$, and so it is right-admissible. By the definition of right-admissibility, $\mathcal{X}_{j_\sigma}$ is left-perfect. Next we will show that $\mathrm{im}\, \mathcal{X}_{j_\sigma} = \mathrm{Rep}(\mathcal{B}_W)^\sigma$.

Recall we have already shown that for every $\langle c, e \rangle \in \mathrm{Rep}(\mathcal{B}_W)^\sigma \subseteq \mathrm{Rep}(\mathcal{B}_W)$, we have

$$(I_{\langle c,e \rangle} \cap W^\sigma) \cong I_{\langle \tilde{\sigma}(c), e \rangle} \tag{3.5c}$$

for some $\langle \tilde{\sigma}(c), e \rangle \in \mathrm{Rep}(\mathcal{B}_{W^\sigma})$, otherwise if $\langle c, e \rangle \in \mathrm{Rep}(\mathcal{B}_W) \setminus \mathrm{Rep}(\mathcal{B}_W)^\sigma$ then $(I_{\langle c,e \rangle} \cap W^\sigma)$ is the trivial persistence module. It follows that

$$W^\sigma \cong \bigoplus_{\langle c,e \rangle \in \mathrm{Rep}(\mathcal{B}_W)^\sigma} (I_{\langle c,e \rangle} \cap W^\sigma).$$

Let $d \in r(\mathcal{B}_{W^\sigma})$ and consider the enumerated set $\langle \cdot, d \rangle_{\mathcal{B}_W}$ with the right-handed ordering. Then $\langle \cdot, d \rangle_{\mathcal{B}_W} \cap \mathrm{Rep}^r(\mathcal{B}_W)^\sigma$ is also an enumerated set with the same indices

since the enumeration on $\langle \cdot, d\rangle_{\mathcal{B}_W}$ lists longest intervals first by definition. It follows that (3.5c) informs a one-to-one correspondence between $\langle \cdot, d\rangle_{B_W} \cap \mathrm{Rep}^r(\mathcal{B}_W)^\sigma$ and $\langle \cdot, d\rangle_{B_{W^\sigma}}$, and so they have the same number of elements, say N_d elements. Again since $\mathcal{X}_{j_\sigma} = \lceil \mathcal{X}^W_{W^\sigma} \rceil$ is left-perfect, and since $\lceil \mathcal{X}^W_{W^\sigma} \rceil$ maps each interval in $\langle \cdot, d\rangle_{\mathcal{B}_{W^\sigma}}$ to the first N_d intervals in the enumeration $\langle \cdot, d\rangle_{\mathcal{B}_W}$, the matching $\mathcal{X}_{j_\sigma}$ maps $\langle \cdot, d\rangle_{\mathcal{B}_{W^\sigma}}$ to precisely the set $\langle \cdot, d\rangle_{\mathcal{B}_W} \cap \mathrm{Rep}^r(\mathcal{B}_W)^\sigma$. By the construction of $\mathcal{X}_{j_\sigma} = \lceil \mathcal{X}^W_{W^\sigma} \rceil$ in Definition 2.56, and since $d \in r(\mathcal{B}_{W^\sigma})$ was arbitrary, we have proved that im $\mathcal{X}_{j_\sigma} = \mathrm{Rep}(\mathcal{B}_W)^\sigma$.

Finally, we will show that if $\mathcal{X}_{j_\sigma}\langle b, d\rangle = \langle b', d'\rangle$, then $b = \tilde{\sigma}(b')$ and $d' = d$. Again let $d \in r(\mathcal{B}_{W^\sigma})$ and consider $(\langle b, d\rangle, k) \in \mathrm{Rep}(\mathcal{B}_{W^\sigma})$ such that $\mathcal{X}_{j_\sigma}(\langle b, d\rangle, k) = (\langle b', d'\rangle, j) \in \mathrm{Rep}(\mathcal{B}_W)^\sigma$. Since $\mathcal{X}_{j_\sigma} = \lceil \mathcal{X}^W_{W^\sigma} \rceil$ then $d' = d$ by the definition of right-admissibility. By the right-handed orderings associated to both $\langle \cdot, d\rangle_{\mathcal{B}_{W^\sigma}}$ and $\langle \cdot, d\rangle_{\mathcal{B}_W} \cap \mathrm{Rep}^r(\mathcal{B}_W)^\sigma$, intervals are sorted ascending by their left-hand endpoint, and the copy numbers are used as secondary orderings. Thus, by the correspondence given by (3.5c) and the monotonicity of $\tilde{\sigma}$, it must also follow that $b = \tilde{\sigma}(b')$. □

Definition 3.6. Let (V, φ_V) be a persistence module and let $\sigma : \mathbb{R} \to \mathbb{R}$ be a translation map. For each $t \in \mathbb{R}$, define

$$K^\sigma_t := \ker \varphi_V(t, \sigma(t))$$

and form the quotient vector spaces

$$(V/K^\sigma)_t := V_t/K^\sigma_t$$

for every $t \in \mathbb{R}$. To every $s, t \in \mathbb{R}$ with $s \leq t$, associate the linear maps given by the unique maps $\varphi_{V/K^\sigma}(s, t)$ such that

$$\varphi_{V/K^\sigma}(s, t) \circ q^\sigma_s = q^\sigma_t \circ \varphi_V(s, t),$$

where $q^\sigma_t : V_t \to V_t/K^\sigma_t$ is the quotient map taking $\alpha \mapsto \alpha + K^\sigma_t$ (see Definition 2.11). Denote this collection of vector spaces and linear maps by $(V/K^\sigma, \varphi_{V/K^\sigma})$.

Lemma 3.7. *Let (V, φ_V) be a persistence module and let $\sigma : \mathbb{R} \to \mathbb{R}$ be a translation map. Then $(V/K^\sigma, \varphi_{V/K^\sigma})$ is a persistence module.*

Proof. Define a persistence module morphism $\phi : V \to T_\sigma(V)$ via $\phi_t := \varphi_V(t, \sigma(t))$. Then $K^\sigma = \operatorname{im} \phi$ as defined, and V/K^σ is the corresponding quotient persistence module (see Definition 2.11). □

Proposition 3.8. *Let (V, φ_V) be a persistence module, let $\sigma : \mathbb{R} \to \mathbb{R}$ be a translation map. The minimal matching $\mathcal{X}_{q^\sigma} : Rep(\mathcal{B}_V) \nrightarrow Rep(\mathcal{B}_{V/K^\sigma})$ witnessed by q^σ has* $\operatorname{coim} \mathcal{X}_{q^\sigma} = Rep(\mathcal{B}_V)^\sigma$, *and if $\mathcal{X}_{q^\sigma}\langle b, d\rangle = \langle b', d'\rangle$ and*

- *if $d \in A(\sigma)$, then $d' = \tilde{\sigma}_L^{-1}(d)$;*
- *if $d \in B(\sigma)$, then $d' = \tilde{\sigma}_R^{-1}(d)$.*

Proof. As in the beginning of proof of Proposiiton 3.5, we will first establish a preliminary result. Pursuant to Remark 3.2, denote by I_J the interval persistence submodule of V corresponding to the interval $J \in \mathcal{B}_V$ in the barcode decomposition of V, dropping the reference to V. We now show that

$$V/K^\sigma \cong \bigoplus_{J \in \mathcal{B}_V} \frac{(I_J)}{K^\sigma \cap (I_J)}. \tag{3.8a}$$

Consider the natural surjection

$$\bigoplus_{J \in \mathcal{B}_V} (I_J)_t \to \bigoplus_{J \in \mathcal{B}_V} \frac{(I_J)_t}{K_t^\sigma \cap (I_J)_t},$$

which has kernel $\bigoplus_{J \in \mathcal{B}_V} K_t^\sigma \cap (I_J)_t$. Since $\varphi_V(t, \sigma(t)) = \bigoplus_{J \in \mathcal{B}_V} \varphi_V(t, \sigma(t))|_{(I_J)_t}$ for all $t \in \mathbb{R}$, then $K_t^\sigma = \bigoplus_{J \in \mathcal{B}_V} K_t^\sigma \cap (I_J)_t$. Thus,

$$V_t/K_t^\sigma \cong \bigoplus_{J \in \mathcal{B}_V} \frac{(I_J)_t}{K_t^\sigma \cap (I_J)_t}$$

and so (3.8a) holds.

In what follows, we establish necessary conditions on $t \in \mathbb{R}$ such that

$$\frac{(I_{\langle c,e\rangle})_t}{K_t^\sigma \cap (I_{\langle c,e\rangle})_t} \neq 0. \tag{3.8b}$$

By the definition of a quotient of vector spaces, (3.8b) is true if and only if there exists some $\alpha \in (I_{\langle c,e\rangle})_t$ such that $\alpha \notin \ker \varphi_V(t,\sigma(t))$, which is true if and only if $\alpha \neq 0$ implies that $\varphi_V(t,\sigma(t))(\alpha) \neq 0$ since $(I_{\langle c,e\rangle})_t \cong \mathsf{k}$. But this only occurs when $c < t^+$ and $\sigma(t)^- < e$. Thus, it remains to determine for which $t \in \mathbb{R}$ we have $\sigma(t)^- < e$.

Let $t \in \mathbb{R}$. If $t < \sigma_L^{-1}(\pi e)$, then by the definition of σ_L^{-1}, we must have $\sigma(t) < \pi e$. Hence, $\sigma(t)^- < e$, and so this case is done. Suppose now that $\sigma_L^{-1}(\pi e) \leq t$, which implies $\pi e \leq (\sigma \circ \sigma_L^{-1})(\pi e) \leq \sigma(t)$ by the right-continuity and monotonicity of σ. If $\pi e < \sigma(t)$, then it cannot be the case that $\sigma(t)^- < e$, and so this case is done. Now consider the case that $\pi e = \sigma(t)$ and suppose that $e = \pi(e)^-$. Then $\sigma(t)^- = \pi(e)^- = e$ and so it cannot be the case that $\sigma(t)^- < e$. Finally, if $\pi e = \sigma(t)$ and $e = \pi(e)^+$, then $\sigma(t)^- < \pi(e)^+ = e$, and so this case holds.

Thus, for any $t \in \mathbb{R}$ with $c < t^+$, we have shown that (3.8b) holds for two cases and these two cases alone:

(i) $t < \sigma_L^{-1}(\pi e)$

(ii) the following two conditions hold:

(a.) $\pi e = \sigma(t)$, and

(b.) $e = \pi(e)^+$.

We now determine conditions on the interval $\langle c, e\rangle$ for which finding a $t \in \mathbb{R}$ with $c < t^+$ and satisfying either condition (i) or condition (ii) is possible. Condition (i) is possible only when $\pi c \leq t < \sigma_L^{-1}(\pi e)$, and thus when $\sigma(\pi c) < \pi e$. Hence, we must have $\tilde{\sigma}(c) < e$. If $e \in A(\sigma)$, then these are the only such t satisfying (i). Thus, in this case, (3.8b) gives us:

$$\text{if } \langle c, e\rangle \in \mathrm{Rep}(\mathcal{B}_V)^\sigma \quad \text{and} \quad e \in A(\sigma), \quad \text{then} \quad \frac{I_{\langle c,e\rangle}}{K^\sigma \cap I_{\langle c,e\rangle}} \cong I_{\langle c, \tilde{\sigma}_L^{-1}(e)\rangle}. \tag{3.8c}$$

Condition (ii) can only occur when $e \in B(\sigma)$. To find the largest $c \in \mathbb{E}$ such that (a.) can be satisfied, suppose that $\sigma_L^{-1}(\pi e) \leq \pi c$ so that we are not in case (i). Since $c < t^+$ implies that $\pi(c) \leq t$, the monotonicity and right-continuity of σ, together with

Lemma 2.42, yield

$$\pi e \leq (\sigma \circ \sigma_L^{-1})(\pi e) \leq \sigma(\pi c) \leq \sigma(t).$$

Thus, the largest $c \in \mathbb{E}$ for which we might achieve (a.) is when $\pi e = \sigma(\pi c)$. Additionally, we must also find a condition on c for which it is possible to have $c < t^+$ and $\sigma(\pi c) = \sigma(t)$, otherwise (a.) could not be satisfied for any $t \in \mathbb{R}$. Note that $c < t^+$ and $\sigma(\pi c) = \sigma(t)$ occurs either when $c = c^-$, since in this case we can choose $t = \pi c$, or when $c = c^+$ and there exists some $t > \pi c \in \mathbb{R}$ with $\sigma(t) = \sigma(\pi c)$. Both of these scenarios occur precisely when $\tilde{\sigma}(c) = \tilde{\sigma}(c)^-$. Hence, in this case we have $\tilde{\sigma}(c) < e$, and so $\langle c, e\rangle \in \mathrm{Rep}(\mathcal{B}_V)^\sigma$. Thus, we have established that for there to exist a $t \in \mathbb{R}$ with $c < t^+$ and satisfying property (ii), we must have $e \in B(\sigma)$ and $\langle c, e\rangle \in \mathrm{Rep}(\mathcal{B}_V)^\sigma$.

In order to establish a result analogous to (3.8c) for intervals $\langle c, e\rangle$ with $e \in B(\sigma)$, we must find the largest $t \in \mathbb{R}$ such that condition (ii) holds. By (a.) and the monotonicity of σ, we must characterize the largest $t \in \mathbb{R}$ with $c < t^+$ such that $\pi e = \sigma(t)$. Consider the number

$$\zeta := \sup\{t \in \mathbb{R} : \sigma(t) = \pi(e)\}.$$

If σ is continuous at ζ, then $\sigma(\zeta) = \pi e$. If σ is discontinuous at ζ, then $\sigma(\zeta) > \pi e$ by the monotonicity and right-continuity of σ. Thus, we see that we achieve the following characterization:

$$\text{if } \langle c, e\rangle \in \mathrm{Rep}(\mathcal{B}_V)^\sigma \quad \text{and} \quad e \in B(\sigma), \quad \text{then} \quad \frac{I_{\langle c,e\rangle}}{K^\sigma \cap I_{\langle c,e\rangle}} \cong I_{\langle c, \tilde{\sigma}_R^{-1}(e)\rangle}. \tag{3.8d}$$

We will now show that $\mathcal{X}_{q^\sigma} : \mathrm{Rep}(\mathcal{B}_V)^\sigma \nrightarrow \mathrm{Rep}(\mathcal{B}_{V/K^\sigma})$ is right-perfect. Since q^σ is an epimorphism, Definition 2.61 yields $\mathcal{X}_{q^\sigma} = \lceil \mathcal{X}_{V/K^\sigma}^{V/K^\sigma} \rceil \circ \lfloor \mathcal{X}_V^{V/K^\sigma} \rfloor = \lfloor \mathcal{X}_V^{V/K^\sigma} \rfloor$, and so it is left-admissible. By the definition of left-admissibility, $\mathcal{X}_{q^\sigma}$ is right-perfect.

Next we show that $\mathrm{coim}\, \mathcal{X}_{q^\sigma} = \mathrm{Rep}(\mathcal{B}_V)^\sigma$. By the correspondences given by equations (3.8c) and (3.8d) and the fact that the persistence module for V/K^σ is isomorphic to the direct sum in (3.8a), we have shown that

$$V/K^\sigma \cong \bigoplus_{(\langle c,e\rangle,k) \in \mathrm{Rep}(\mathcal{B}_V)^\sigma} \frac{(I_{(\langle c,e\rangle,k)})}{K^\sigma \cap (I_{(\langle c,e\rangle,k)})}.$$

Fix some $b \in \ell(\mathcal{B}_{V/K^\sigma})$. Considering only the intervals with left-hand endpoint b, it follows that (3.8c) and (3.8d) determine a one-to-one correspondence between $\langle b, \cdot \rangle_{B_V} \cap \mathrm{Rep}^\ell(\mathcal{B}_V)^\sigma$ and $\langle b, \cdot \rangle_{B_{V/K^\sigma}}$, and so they have the same number of elements, say N_b elements. Again since $\mathcal{X}_{q^\sigma} = \lfloor \mathcal{X}_V^{V/K^\sigma} \rfloor$ is right-perfect, and since $\lfloor \mathcal{X}_V^{V/K^\sigma} \rfloor$ maps each interval in the enumeration $\langle \cdot, d \rangle_{\mathcal{B}_W}$ from the first N_b intervals in $\langle b, \cdot \rangle_{\mathcal{B}_{V/K^\sigma}}$, the matching $\mathcal{X}_{q^\sigma}$ maps $\langle b, \cdot \rangle_{V/K^\sigma}$ to precisely the set $\langle b, \cdot \rangle_{\mathcal{B}_V} \cap \mathrm{Rep}^\ell(\mathcal{B}_V)^\sigma$. By the construction of $\mathcal{X}_{q^\sigma} = \lfloor \mathcal{X}_V^{V/K^\sigma} \rfloor$ in Definition 2.56, and the fact that $b \in \ell(\mathcal{B}_{V/K^\sigma})$ was arbitrary, we conclude that $\mathrm{coim}\, \mathcal{X}_{q^\sigma} = \mathrm{Rep}(\mathcal{B}_V)^\sigma$.

Finally, we will show that if $\mathcal{X}_{q^\sigma} \langle b, d \rangle = \langle b', d' \rangle$, then $b = b'$. Additionally, if $d \in A(\sigma)$, then $d' = \tilde{\sigma}_L^{-1}(d)^-$, and if $d \in B(\sigma)$, then $d' = \tilde{\sigma}_R^{-1}(d)$. Let $b \in \ell(\mathcal{B}_{V/K^\sigma})$ and consider $(\langle b', d' \rangle, j) \in \mathrm{Rep}(\mathcal{B}_{V/K^\sigma})$ such that $\mathcal{X}_{q^\sigma}(\langle b, d \rangle, k) = (\langle b', d' \rangle, j)$. Since $\mathcal{X}_{q^\sigma} = \lfloor \mathcal{X}_V^{V/K^\sigma} \rfloor$ then $b = b'$ by the definition of left-admissibility. By the left-handed orderings associated to both $\langle b, \cdot \rangle_{\mathcal{B}_V} \cap \mathrm{Rep}^\ell(\mathcal{B}_V)^\sigma$ and $\langle b, \cdot \rangle_{\mathcal{B}_{V/K^\sigma}}$, intervals are sorted descending by their right-hand endpoint, and the copy numbers are used as secondary orderings. The result follows from the correspondences given in (3.8c) and (3.8d) and Proposition 2.46. □

We have to establish one more useful lemma. It also serves to justify the formulations of Definitions 3.3 and 3.6.

Lemma 3.9. *Let (V, φ_V) and (W, φ_W) be persistence modules and let $\sigma : \mathbb{R} \to \mathbb{R}$ be a translation map. Define K_t^σ and W_t^σ as in Definitions 3.3 and 3.6 for every $t \in \mathbb{R}$. For any morphism $\phi : V \to W$,*

(i) if $\mathrm{im}\, \phi$ *is σ-trivial then $W_t^\sigma \subseteq \mathrm{im}\, \phi_t \subseteq W_t$ for every $t \in \mathbb{R}$, and*

(ii) if $\ker \phi$ *is σ-trivial then* $\ker \phi_t \subseteq K_t^\sigma \subseteq V_t$ *for every $t \in \mathbb{R}$.*

Proof. We first prove (i). By definition, given a morphism $\phi : V \to W$, the cokernel $\mathrm{coker}\, \phi$ is σ-trivial if and only if

$$\varphi_{\mathrm{coker}\, \phi}(x, \sigma(x)) = 0 \quad \text{for all} \quad x \in \mathbb{R},$$

which is true if and only if

$$\text{im } \varphi_W(x, \sigma(x)) \subseteq \text{im } \phi_{\sigma(x)} \quad \text{for all} \quad x \in \mathbb{R},$$

which again is true if and only if for each $x \in \mathbb{R}$ and each $\alpha \in W_x$, there exists some $\beta \in V_{\sigma(x)}$ such that

$$\varphi_W(x, \sigma(x))(\alpha) = \phi_{\sigma(x)}(\beta).$$

For $x \in S(t)$, the commutativity of diagram (3.9a) then implies that for such $\alpha \in W_x$ and $\beta \in V_{\sigma(x)}$, we have

$$\varphi_W(x,t)(\alpha) = \varphi_W(\sigma(x),t)[\varphi_W(x,\sigma(x))(\alpha)] = \phi_t[\varphi_V(\sigma(x),t)(\beta)],$$

and so $\text{im } \varphi_W(x,t) \subseteq \text{im } \phi_t$. Since $x \in S(t)$ was arbitrary, we have succeeded in showing that $W_t^\sigma = \bigcup_{x \in S(t)} \text{im } \varphi_W(x,t) \subseteq \text{im } \phi_t$.

$$\begin{array}{ccccc} & & V_{\sigma(x)} & \xrightarrow{\varphi_V(\sigma(x),t)} & V_t \\ & & \downarrow{\scriptstyle \phi_{\sigma(x)}} & & \downarrow{\scriptstyle \phi_t} \\ W_x & \xrightarrow{\varphi_W(x,\sigma(x))} & W_{\sigma(x)} & \xrightarrow{\varphi_W(\sigma(x),t)} & W_t \end{array} \tag{3.9a}$$

To prove (ii), we will prove both that for all morphisms ϕ with a σ-trivial kernel, it follows that $\ker \phi_t \subseteq K_t^\sigma$ for all $t \in \mathbb{R}$. Let $\phi : V \to W$ be a morphism such that $\ker \phi$ is σ-trivial. By definition, $\ker \phi$ is σ-trivial if and only if

$$\varphi_V(t, \sigma(t))|_{\ker \phi_t} = \phi_{\ker \phi}(t, \sigma(t)) = 0$$

for all $t \in \mathbb{R}$. Hence, $\ker \phi_t \subseteq \ker \varphi_V(t, \sigma(t)) = K_t^\sigma$ for all $t \in \mathbb{R}$ if and only if $\ker \phi$ is σ-trivial. □

We are now ready to prove our first main result. We label this generalization of the Induced Matching Theorem with the word *non-uniform* since the bounds placed on

the sizes of the barcodes in $\ker\phi$ and $\operatorname{coker}\phi$ is based on an arbitrary translation map $\sigma : \mathbb{R} \to \mathbb{R}$, whereby the quantity $\sigma(t) - t$ may be of variable length on $t \in \mathbb{R}$ (compare to the uniform induced matching theorem which is the special case of what follows by taking $\sigma(t) = t + \delta$ for some $\delta \geq 0$).

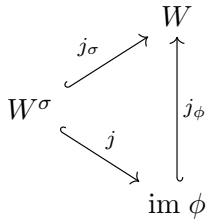
Non-uniform Induced Matching Theorem 3.10. *Let σ be a right-continuous translation map on $\mathbb{R}$ and let $\phi : V \to W$ be a persistence module morphism. Suppose that $\mathcal{X}_\phi : \mathcal{B}_V \nrightarrow \mathcal{B}_W$ is the admissible matching witnessed by ϕ, and that $\mathcal{X}_\phi\langle b, d\rangle = \langle b', d'\rangle$.*

(i) If $\operatorname{coker}\phi$ *is σ-trivial, then $Rep(\mathcal{B}_W)^\sigma \subseteq \operatorname{im}\mathcal{X}_\phi$ and $b' \leq b \leq \tilde{\sigma}(b')$.*

(ii) If $\ker\phi$ *is σ-trivial, then $Rep(\mathcal{B}_V)^\sigma \subseteq \operatorname{coim}\mathcal{X}_\phi$. If $\langle b, d\rangle \in Rep(\mathcal{B}_V)^\sigma$ and $d \in B(\sigma)$, then $\tilde{\sigma}_R^{-1}(d) \leq d' \leq d$; otherwise, $\tilde{\sigma}_L^{-1}(d) \leq d' \leq d$.*

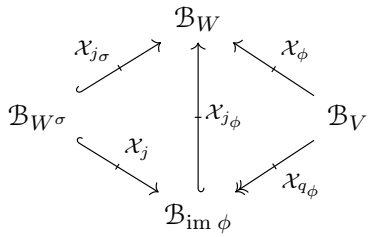
Proof. The proofs of (i) and (ii) proceed very similarly. The general technique is to construct a commutative diagram which sandwiches $\operatorname{im}\phi$ between W and W^σ for (i) and in between V and V/K^σ for (ii). Both of these triangles contain only monomorphisms or epimorphisms, and so by Proposition 2.65, we get a corresponding commutative diagram on barcode matchings. After this, it is a short argument in either case to conclude the bounds given in the statement of the theorem.

To prove (i), we note that by Lemma 3.9(i), the inclusion map $j : W^\sigma \hookrightarrow \operatorname{im}\phi$ is such that diagram (3.10a) commutes, where $j_\phi : \operatorname{im}\phi \hookrightarrow W$ is the canonical injection associated to the map ϕ and let $j_\sigma : W^\sigma \hookrightarrow W$ be the inclusion morphism (Proposition 3.4).



(3.10a)

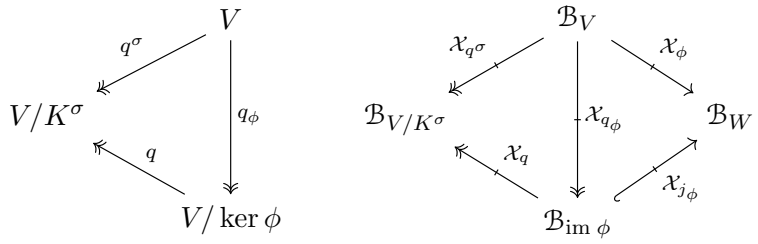
Since $\mathcal{X}_\phi$ is the admissible matching witnessed by ϕ, the definition of $\mathcal{X}_\phi$ and Remark 2.62 imply that $\mathcal{X}_\phi = \mathcal{X}_{j_\phi} \circ \mathcal{X}_{q_\phi}$. By Proposition 2.65 and the commutativity of diagram (3.10a), which consists only of injective morphisms, diagram (3.10b) commutes.



(3.10b)

By commutativity of the left triangle in (3.10b), im $\mathcal{X}_{j_\sigma} \subseteq$ im $\mathcal{X}_{j_\phi}$. Hence, we see that $\text{Rep}(\mathcal{B}_W)^\sigma = \text{im } \mathcal{X}_{j_\sigma} \subseteq \text{im } \mathcal{X}_{j_\phi} = \text{im } \mathcal{X}_\phi$. We will now prove that if $\mathcal{X}_\phi\langle b, d\rangle = \langle b', d'\rangle$, then $b' \leq b \leq \tilde{\sigma}(b')$. Suppose that $\mathcal{X}_{j_\phi}\langle b, d'\rangle = \langle b', d'\rangle$. Since $\mathcal{X}_{j_\phi}$ is right-admissible, then $b' \leq b$. If $\langle b', d'\rangle \in \text{Rep}(\mathcal{B}_W)^\sigma$, then there exists some $\langle \tilde{\sigma}(b'), d'\rangle \in \text{Rep}(\mathcal{B}_{W^\sigma})$ with $\mathcal{X}_{j_\sigma}\langle \tilde{\sigma}(b'), d'\rangle = \langle b', d'\rangle$. Thus, by commutativity of the the left-hand triangle in (3.10b), we must have $\mathcal{X}_j\langle \tilde{\sigma}(b'), d'\rangle = \langle b, d'\rangle$. By Theorem 2.59 and Remark 2.62, $\mathcal{X}_j$ is right-admissible, and so $b \leq \tilde{\sigma}(b')$. If $\langle b', d'\rangle \notin \text{Rep}(\mathcal{B}_W)^\sigma$, then $d' \leq \tilde{\sigma}(b')$. By Corollary 2.64, it follows that $b < d' \leq \tilde{\sigma}(b')$.

The proof of (ii) is similar to the proof of (i). By Lemma 3.9(ii), the quotient map $q : V/(\ker \phi) \to V/K^\sigma$ is surjective, well-defined, and makes the triangle on the left of diagram (3.10c) commute, where q^σ is the quotient morphism from V to V/K^σ, and q_ϕ is the canonical surjective morphism from the factorization of ϕ. Hence every map on the triangle on the left of diagram (3.10c) is surjective. By Proposition 2.65, the diagram at the level of barcodes, right, also commutes because $V/\ker \phi \cong \text{im } \phi$, and $\mathcal{X}_\phi$ is the admissible matching witnessed by ϕ.



(3.10c)

By Proposition 3.8, we have coim $\mathcal{X}_{q^\sigma} = \mathrm{Rep}(\mathcal{B}_V)^\sigma$. Thus, $\mathrm{Rep}(\mathcal{B}_V)^\sigma = \text{coim}\ \mathcal{X}_{q^\sigma} \subseteq \text{coim}\ \mathcal{X}_{q_\phi} = \text{coim}\ \mathcal{X}_\phi$. Now suppose that $\mathcal{X}_\phi\langle b, d\rangle = \langle b', d'\rangle$. Since $\mathcal{X}_{q_\phi}$ is left-admissible, we have that $d' \leq d$.

Suppose that $\langle b, d\rangle \in \mathrm{Rep}(B_V)^\sigma$ has $d \in B(\sigma)$. Since $\langle b, d\rangle \in \mathrm{Rep}(B_V)^\sigma$ there exists some $\langle b, \tilde{\sigma}_R^{-1}(d)\rangle \in \mathrm{Rep}(\mathcal{B}_{V/K^\sigma})$ such that $\mathcal{X}_q\langle b, d'\rangle = \langle b, \tilde{\sigma}_R^{-1}(d)\rangle$. By Theorem 2.59 and Remark 2.62, $\mathcal{X}_q$ is left-admissible, so $\tilde{\sigma}_R^{-1}(d) \leq d'$.

Suppose that $\langle b, d\rangle \in \mathrm{Rep}(B_V)^\sigma$ has $d \in A(\sigma)$. Then there exists some $\langle b, \tilde{\sigma}_L^{-1}(d)\rangle \in \mathrm{Rep}(\mathcal{B}_{V/K^\sigma})$ such that $\mathcal{X}_q\langle b, d'\rangle = \langle b, \tilde{\sigma}_L^{-1}(d)\rangle$. By Theorem 2.59 and Remark 2.62, $\mathcal{X}_q$ is left-admissible, and so $\tilde{\sigma}_L^{-1}(d) \leq d'$.

Finally, suppose that $\langle b, d\rangle \notin \mathrm{Rep}(\mathcal{B}_V)^\sigma$. Then $(\sigma_L^{-1} \circ \pi)(d) \leq \pi(b)$. Hence, since Corollary 2.64 yields $b' \leq b < d \leq d'$, we have $\tilde{\sigma}_L^{-1}(d) \leq (\sigma_L^{-1} \circ \pi)(d)^+ \leq \pi(b)^+ \leq d'$. □

Remark 3.11. By Proposition 2.46, item (ii) of the Non-uniform Induced Matching Theorem may be simplified by relaxing the bounds somewhat on intervals with right-hand endpoint in $B(\sigma)$: if $\mathcal{X}_\phi\langle b, d\rangle = \langle b', d'\rangle$ and $\ker\phi$ is σ-trivial, then $\tilde{\sigma}_L^{-1}(d) \leq d' \leq d$.

We use this simplified version in the Algebraic Stability Theorem for Generalized Interleavings, which we present shortly. Let $\sigma : \mathbb{R} \to \mathbb{R}$ be monotone. For what follows, we will define

$$\tilde{\sigma}_\star^{-1} := \begin{cases} \tilde{\sigma}_L^{-1}(e) & \text{if } e \in A(\sigma), \text{ and} \\ \tilde{\sigma}_R^{-1}(e) & \text{if } e \in B(\sigma) \end{cases}$$

for every decorated endpoint $e \in \mathbb{E}$.

Proposition 3.12. *Let V be a persistence module and let $\tau : \mathbb{R} \to \mathbb{R}$ be monotone. There exists a right-perfect matching $\mathcal{X} : \mathit{Rep}(\mathcal{B}_V) \twoheadrightarrow \mathit{Rep}(\mathcal{B}_{T_\tau(V)})$ such that* $\operatorname{coim} \mathcal{X} = \{(\langle c, e\rangle, k) \in \mathit{Rep}(\mathcal{B}_V) : \tilde{\tau}_\star^{-1}(c) < \tilde{\tau}_\star^{-1}(e)\}$ *and* $\langle c, e\rangle \mapsto \langle \tilde{\tau}_\star^{-1}(c), \tilde{\tau}_\star^{-1}(e)\rangle$.

Proof. Suppose that $\langle c, e\rangle \in \mathcal{B}_V$ and consider the corresponding interval persistence submodule $I_{\langle c,e\rangle}$. Then $I_{\langle c,e\rangle} \cap B_{T_\tau(V)}$ is an interval persistence submodule of $B_{T_\tau(V)}$ with interval

$$\{t \in \mathbb{R} : c \le \tau(t)^- < \tau(t)^+ \le e\}, \tag{3.12a}$$

which is nonempty if and only if $\langle c, e\rangle \cap \tau(\mathbb{R}) \neq \emptyset$, where we regard $\langle c, e\rangle$ as an interval in $\mathbb{R}$. We will now find the $t \in \mathbb{R}$ such that $c \le \tau(t)^-$.

We first consider the case when $(\tau \circ \tau_L^{-1})(\pi c) < \pi c$, whereby we must have $\tau_L^{-1}(\pi c) < t$ to get $c \le \tau(t)^-$. If $c = \pi(c)^-$, then it suffices to choose $t \in \mathbb{R}$ such that $\tau_L^{-1}(\pi c) < t$ since then $c = \pi(c)^- \le \tau(t)^-$. Suppose that $c = \pi(c)^-$ and $\pi c \notin \tau(\mathbb{R})$. Then for all $\tau_L^{-1}(\pi c) < t$ we also have $c < \tau(t)^-$. Now suppose that $c = \pi(c)^+$ and $\pi c \in \tau(\mathbb{R})$ (i.e. $c \in B(\tau)$). If $(\tau \circ \tau_R^{-1})(\pi c) = \pi c$, then we must have $\tau_R^{-1}(\pi c) < t$ for $c \le \tau(t)^-$. However, if $(\tau \circ \tau_R^{-1})(\pi c) > \pi c$, then we can take $\tau_R^{-1}(\pi c) = t$ for $c \le \tau(t)^-$.

Next we consider the case when $(\tau \circ \tau_L^{-1})(\pi c) \ge \pi c$. If $(\tau \circ \tau_L^{-1})(\pi c) > \pi c$ then setting $t = \tau_L^{-1}(\pi c)$ yields $c \le \tau(t)^-$. Suppose that $(\tau \circ \tau_L^{-1})(\pi c) = \pi c$. If $c = \pi(c)^-$ then $t = \tau_L^{-1}(\pi c)$ yields $c = \pi(c)^- \le \tau(t)^-$ and so this case is done. Finally, suppose that $c = \pi(c)^+$ (i.e. $c \in B(\tau)$). If $(\tau \circ \tau_R^{-1})(\pi c) = \pi c$, we must have $\tau_R^{-1}(\pi c) < t$ to get $c \le \tau(t)^-$. If $(\tau \circ \tau_R^{-1})(\pi c) > \pi c$, then we can set $\tau_R^{-1}(\pi c) = t$ and have $c \le \tau(t)^-$.

Thus, we have succeeded in showing that if $c \in A(\tau)$, the left-hand endpoint of interval (3.12a) is $\tilde{\tau}_L^{-1}(c)$, and if $c \in B(\tau)$, the left-hand endpoint of interval (3.12a) is $\tilde{\tau}_R^{-1}(c)$. It remains to find $t \in \mathbb{R}$ such that $\tau(t)^+ \le e$ and establish the right-hand endpoint of interval (3.12a).

For the first case we consider $e = \pi(e)^-$. If $(\tau \circ \tau_L^{-1})(\pi e) < \pi e$, then setting $t = \tau_L^{-1}(\pi e)$ yields $\tau(t)^+ \le e$. If $(\tau \circ \tau_L^{-1})(\pi e) \ge \pi e$, then we must have $t < \tau_L^{-1}(\pi e)$ for $\tau(t)^+ \le e$. Now suppose that $e = \pi(e)^+$. If $\pi e \notin \tau(\mathbb{R})$, then if $(\tau \circ \tau_L^{-1})(\pi e) < \pi e$ we can set $t = \sigma_L^{-1}(\pi e)$ and get $\tau(t)^+ \le e$, and if $(\tau \circ \tau_L^{-1})(\pi e) > \pi e$ we must have $t < \sigma_L^{-1}(\pi e)$ for $\tau(t)^+ \le e$. Finally, suppose that $\pi e \in \tau(\mathbb{R})$ (i.e. $e \in B(\tau)$). Then if $(\tau \circ \tau_R^{-1})(\pi e) =$

πe, we can set $t = \tau_R^{-1}(\pi e)$ to achieve $\tau(t)^+ \leq e$, and if $(\tau \circ \tau_R^{-1})(\pi e) > \pi e$, we must have $t < \tau_R^{-1}(\pi e)$ for $\tau(t)^+ \leq e$. Hence, if $e \in A(\tau)$, the right-hand endpoint of interval (3.12a) is $\tilde{\tau}_L^{-1}(e)$, and if $e \in B(\tau)$, then the right-hand endpoint is $\tilde{\tau}_R^{-1}(e)$.

To conclude the proof, note that $\langle c, e\rangle \cap \tau(\mathbb{R}) \neq \emptyset$ if and only if $\tilde{\tau}_\star^{-1}(c) < \tilde{\tau}_\star^{-1}(e)$. Thus, by the work above, we get a one-to-one correspondence between the intervals $\langle c, e\rangle \in \mathcal{B}_V$ with $\tilde{\tau}_\star^{-1}(c) < \tilde{\tau}_\star^{-1}(e)$ and the intervals in $\mathcal{B}_{T_\tau(V)}$. Define $\mathcal{X} : \mathrm{Rep}(\mathcal{B}_V) \nrightarrow \mathrm{Rep}(\mathcal{B}_{T_\tau(V)})$ to be this matching, minimizing copy numbers. □

Algebraic Stability Theorem for Generalized Interleavings 3.13. *Let V and W be (τ, σ)-interleaved persistence modules. If the composition $(\sigma \circ \tau)$ is right-continuous, then there exists a matching $\mathcal{X} : Rep(\mathcal{B}_V) \nrightarrow Rep(\mathcal{B}_W)$ on the barcodes of V and W such that for each $\mathcal{X}\langle b, d\rangle = \langle b', d'\rangle$, the inequalities*

$$\tilde{\tau}_\star^{-1}(b') \leq b \leq [\widetilde{(\sigma \circ \tau)} \circ \tilde{\tau}_\star^{-1}](b') \tag{3.13a}$$

and

$$\widetilde{(\sigma \circ \tau)}_L^{-1}(d) \leq \tilde{\tau}_\star^{-1}(d') \leq d \tag{3.13b}$$

are satisfied, and any unmatched intervals are contained in the sets

$$\{\langle b, d\rangle \in \mathcal{B}_V : \widetilde{(\sigma \circ \tau)}_L^{-1}(d) \leq b\}$$

and

$$\{\langle b', d'\rangle \in \mathcal{B}_W : \tilde{\tau}_\star^{-1}(d') \leq [\widetilde{(\sigma \circ \tau)} \circ \tilde{\tau}_\star^{-1}](b')\}.$$

Proof. By Proposition 3.12, consider the matching on barcodes $\mathcal{X}' : \mathrm{Rep}(\mathcal{B}_W) \to \mathrm{Rep}(\mathcal{B}_{T_\tau(W)})$ given by $\langle c, e\rangle \mapsto \langle \tilde{\tau}_\star^{-1}(c), \tilde{\tau}_\star^{-1}(e)\rangle$. The restriction of this matching to the subset

$$S = \{(\langle c, e\rangle, k) \in \mathrm{Rep}(\mathcal{B}_W) : \tilde{\tau}_\star^{-1}(c) < \tilde{\tau}_\star^{-1}(e)\}$$

is a bijection. That is for each $(\langle \bar{c}, \bar{e}\rangle, j) \in \mathcal{B}_{T_\tau(W)}$, there exists a $(\langle c, e\rangle, k) \in S$ such that $\bar{c} = \tilde{\tau}_\star^{-1}(c)$ and $\bar{e} = \tilde{\tau}_\star^{-1}(e)$.

By Proposition 2.33, the Non-uniform Induced Matching Theorem, and Remark 3.11,

the interleaving morphism $\phi : V \to T_\tau(W)$ witnesses a minimal admissible matching $\mathcal{X}_\phi : \mathrm{Rep}(\mathcal{B}_V) \nrightarrow \mathrm{Rep}(\mathcal{B}_{T_\tau(W)})$ for ϕ such that for matched intervals $\mathcal{X}_\phi\langle b, d\rangle = \langle \bar{b}, \bar{d}\rangle$,

$$\bar{b} \leq b \leq \widetilde{(\sigma \circ \tau)}(\bar{b}) \quad \text{and} \quad \widetilde{(\sigma \circ \tau)}_L^{-1}(d) \leq \bar{d} \leq d, \tag{3.13c}$$

and unmatched intervals are contained in the sets

$$\{\langle b, d\rangle \in \mathcal{B}_V : \widetilde{(\sigma \circ \tau)}_L^{-1}(d) \leq b\} \quad \text{and} \quad \{\langle \bar{b}, \bar{d}\rangle \in \mathcal{B}_{T_\tau(W)} : \bar{d} \leq \widetilde{(\sigma \circ \tau)}(\bar{b})\}. \tag{3.13d}$$

Applying the reverse matching $\mathcal{X}'^{-1}$, we obtain the relations $\bar{b} = \tilde{\tau}_\star^{-1}(b')$ and $\bar{d} = \tilde{\tau}_\star^{-1}(d')$ for some $b', d' \in \mathbb{E}$. Substituting these into (3.13c) and (3.13d), we obtain a matching $\mathcal{X} : \mathrm{Rep}(\mathcal{B}_V) \nrightarrow \mathrm{Rep}(\mathcal{B}_W)$ such that inequalities (3.13a) and (3.13b) are satisfied for matched intervals, and the unmatched intervals are contained in the sets in the statement of the theorem. To show that every unmatched interval in $\mathrm{Rep}(\mathcal{B}_W)$ is contained in the set $\{\langle b', d'\rangle \in \mathcal{B}_W : \tilde{\tau}_\star^{-1}(d') \leq [\widetilde{(\sigma \circ \tau)} \circ \tilde{\tau}^{-1}](b')\}$, it remains to show that any interval $\langle b', d'\rangle \in \mathrm{Rep}(\mathcal{B}_W) \setminus S$ must also fall in this set. By the definition of the set S, we have $\tilde{\tau}_\star^{-1}(d') \leq \tilde{\tau}_\star^{-1}(b')$, so by the monotonicity of $\widetilde{(\sigma \circ \tau)}$ and the fact that $(\sigma \circ \tau)$ is a translation map, we have $\tilde{\tau}_\star^{-1}(d') \leq [\widetilde{(\sigma \circ \tau)} \circ \tilde{\tau}_\star^{-1}](b')$, which concludes the proof. □

While the inequalities in the Algebraic Stability Theorem for Generalized Interleavings may seem impractical at first glance, when the translation maps τ and σ are bijective, we have the following intuitive relationships between the matched intervals $\langle b, d\rangle \in \mathcal{B}_V$ and $\langle b', d'\rangle \in \mathcal{B}_W$. Since we ultimately want to estimate $\langle b', d'\rangle$ from $\langle b, d\rangle$, we rewrite the inequalities given by the theorem so that this relationship is more clear.

Corollary 3.14. *Let V and W be PFD persistence modules that are (τ, σ)-interleaved as in Theorem 3.13. If the maps τ and σ are invertible, then inequalities* (3.13a) *and* (3.13b) *become $\tilde{\sigma}^{-1}(b) \leq b' \leq \tilde{\tau}(b)$ and $\tilde{\sigma}^{-1}(d) \leq d' \leq \tilde{\tau}(d)$, respectively. The unmatched intervals $\langle c, e\rangle$ in either barcode have $e \leq \widetilde{(\sigma \circ \tau)}(c)$.*

Note that if $\tau(t) = t + \delta = \sigma(t)$, then we are in the case of a δ-interleaving and recover the Algebraic Stability Theorem.

Chapter 4

Applications of the Algebraic Stability Theorem for Generalized Interleavings

We illustrate the use of the Algebraic Stability Theorem for Generalized Interleavings (Theorem 3.13) as part of a comparison framework for interleaved persistence modules through a collection of applications. We find it convenient to visualize the barcode of a persistence module using an associated multiset.

Definition 4.1. Let $\mathcal{B}_V$ be the barcode associated to a persistence module (V, φ_V). The *(undecorated) persistence diagram* of V is the multiset

$$\mathsf{PD}(V) := (\{(\pi b, \pi d) \in \overline{\mathbb{R}}^2 : (\langle b, d\rangle, k) \in \mathrm{Rep}(\mathcal{B}_V)\}, n),$$

where $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$, and the counting function $n : \overline{\mathbb{R}}^2 \to \mathbb{Z}_+$ is defined as

$$n(b, d) := |\{(\langle b', d'\rangle, k) \in \mathrm{Rep}(\mathcal{B}_V) : \pi b' = b, \pi d' = d\}|.$$

By projecting decorated endpoints to extended real numbers, we obtain the following corollary of the Algebraic Stability Theorem for Generalized Interleavings in the context of persistence diagrams. As in Corollary 3.14, we rewrite the first inequality so that we can estimate the location of b' in terms of b.

Corollary 4.2. *Let V and W be (τ, σ)-interleaved persistence modules and that the maps τ and σ are invertible. If the composition $(\sigma \circ \tau)$ is right-continuous, then there exists a matching*

$$\mathsf{PD}_{\mathcal{X}} : \mathit{Rep}(\mathsf{PD}(V)) \nrightarrow \mathit{Rep}(\mathsf{PD}(W))$$

on the persistence diagrams of V and W such that for $\mathsf{PD}_{\mathcal{X}}(b,d) = (b',d')$, the inequalities $\sigma^{-1}(b) \leq b' \leq \tau(b)$ and $\sigma^{-1}(d) \leq d' \leq \tau(d)$ hold, and any unmatched points (c,e) in either persistence diagram have $e \leq (\sigma \circ \tau)(c)$.

Unfortunately, without knowledge of the endpoint decorations, there is not a direct corollary of the Algebraic Stability Theorem for Generalized Interleavings in the context of persistence diagrams for arbitrary translation pairs. This is due to the nature of inequality (3.13b), whereby the middle term could be either the left or right generalized inverse depending on the endpoint decoration. In this case, one must apply the Algebraic Stability Theorem for Generalized Interleavings to the barcodes and then take the projection to extended real numbers.

4.1 Applications for Image Analysis and Large Point Clouds

Our first example shows how the persistence diagram of a real-valued function changes after composing with a monotone perturbation of the range. Our second example examines the relationship between $\mathbb{Z}$-indexed and $\mathbb{R}$-indexed persistence modules, and the third concerns errors introduced when approximating a smooth scalar field by a digital image. Finally, we close out this section with two examples that apply to computing persistence diagrams of large point clouds.

In the sections that follow, it will be useful to define persistence modules induced by the sublevel sets of a scalar field.

Definition 4.3. Let D be a topological space and let $f : D \to \mathbb{R}$ be a scalar field. Fix a field k and choose $n \in \mathbb{N}$. The *persistence module induced by the sublevel set filtration of D by f*, denoted by $(M_n(f), \varphi_{M_n(f)})$, is defined by

$$M_n(f)_t := H_n(\mathsf{C}(f,t), \mathsf{k}), \qquad t \in \mathbb{R},$$

the vector spaces given by taking n-dimensional homology of the sublevel sets $\mathsf{C}(f,t) := \{x \in D : f(x) \leq t\}$ for every $t \in \mathbb{R}$, and the associated linear maps induced by the inclusion maps. That is, for $s,t \in \mathbb{R}$ with $s \leq t$, we define $\varphi_{M_n(f)}(s,t)$ to be the map at

the level of homology induced by the inclusion $\mathsf{C}(f,s) \subseteq \mathsf{C}(f,t)$. We will often consider some $n \in \mathbb{N}$ to be fixed in advance and write $(M(f), \varphi_{M(f)})$ instead of $(M_n(f), \varphi_{M_n(f)})$.

We will also find it useful to define persistence modules induced by finite point clouds.

Definition 4.4. Let X be a finite metric space with metric d. The *Vietoris-Rips complex of X at scale t*, denoted by $\mathsf{R}(X,t)$, is the simplicial complex with vertices given by X and containing the N-simplex $[x_{i_0}, \ldots, x_{i_N}]$ if and only if $d(x_{i_j}, x_{i_k}) \leq 2t$ for all $j,k = 0, \ldots, N$. The collection $\{\mathsf{R}(X,t)\}_{t\in\mathbb{R}}$ is called the *Vietoris-Rips filtration associated to X*.

Definition 4.5. Let X be a finite metric space with metric d. Fix a field k and choose $n \in \mathbb{N}$. We form the *persistence module induced by the Vietoris-Rips filtration associated to X*, denoted by $(M_n^{\mathsf{R}}(X), \varphi_{M_n^{\mathsf{R}}(X)})$, by defining

$$M_n^{\mathsf{R}}(f)_t := H_n(\mathsf{R}(X,t), \mathsf{k}), \qquad t \in \mathbb{R},$$

the vector spaces given by taking n-dimensional homology of the Vietoris-Rips complex $\mathsf{R}(X,t)$ at scale $t \in \mathbb{R}$, and the associated linear maps induced by the inclusion maps. That is, for $s,t \in \mathbb{R}$ with $s \leq t$, we define $\varphi_{M_n^{\mathsf{R}}(X)}(s,t)$ to be the map at the level of homology induced by the inclusion $\mathsf{R}(X,s) \subseteq \mathsf{R}(X,t)$. We will often consider some $n \in \mathbb{N}$ to be fixed in advance and write $(M^{\mathsf{R}}(X), \varphi_X)$ instead of $(M_n^{\mathsf{R}}(X), \varphi_{M_n^{\mathsf{R}}(X)})$, with an additional simplification in the notation of the transition maps.

Definition 4.6. Let X be a finite point cloud in a metric space $\mathcal{M}$ with metric d. The *Čech complex of X at scale t*, denoted by $\check{\mathsf{C}}(X,t)$, is the simplicial complex with vertices given by X and containing the N-simplex $[x_{i_0}, \ldots, x_{i_N}]$ if and only if $\bigcap_{j=0}^{N} B_t(x_{i_j}) \neq \emptyset$, where $B_t(x) \subseteq \mathcal{M}$ is the open ball of radius t about the point x. The collection $\{\check{\mathsf{C}}(X,t)\}_{t\in\mathbb{R}}$ is called the *Čech filtration* associated to X.

Definition 4.7. Let X be a point cloud in a finite metric space $\mathcal{M}$ with metric d. Fix a field k and choose $n \in \mathbb{N}$. We form the *persistence module induced by the Čech filtration*

associated to X, denoted by $(M_n^{\check{\mathsf{C}}}(X), \varphi_{M_n^{\check{\mathsf{C}}}(X)})$ by defining

$$M_n^{\check{\mathsf{C}}}(f)_t := H_n(\check{\mathsf{C}}(X,t), \mathsf{k}), \qquad t \in \mathbb{R},$$

the vector spaces given by taking n-dimensional homology of the Čech complex $\check{\mathsf{C}}(X,t)$ at scale $t \in \mathbb{R}$, and the associated linear maps induced by the inclusion maps. That is, for $s, t \in \mathbb{R}$ with $s \leq t$, we define $\varphi_{M_n^{\check{\mathsf{C}}}(X)}(s,t)$ to be the map at the level of homology induced by the inclusion $\check{\mathsf{C}}(X,s) \subseteq \check{\mathsf{C}}(X,t)$. We will often consider some $n \in \mathbb{N}$ to be fixed in advance and write $(M^{\check{\mathsf{C}}}(X), \varphi_X)$ instead of $(M_n^{\check{\mathsf{C}}}(X), \varphi^{\check{\mathsf{C}}}_{M_n(X)})$, with an additional simplification in the notation of the transition maps.

The following lemma gives a useful relationship between Vietoris-Rips and Čech complexes.

Lemma 4.8. *Let X be a point cloud in a finite metric space $\mathcal{M}$ with metric d and let $t \in \mathbb{R}$. Then $\check{\mathsf{C}}(X,t) \subseteq \mathsf{R}(X,t) \subseteq \check{\mathsf{C}}(X,2t)$.*

Proof. The first inclusion is an immediate consequence of Definition 4.4 and 4.6. The second inclusion is a consequence of the triangle inequality and Definition 4.6. □

In the sections that follow, we will assume that a field k and a dimension $n \in \mathbb{N}$ are fixed in advance, and so we drop reference to the dimension when referencing the induced persistence modules defined above.

4.1.1 Perturbations of a Real-Valued Function

Our first example, while rather simple, serves to illustrate the convenience of Definition 2.27 (versus the analogous definition in [12]) in applying Theorem 3.13 and its corollaries. While the relationship between the two persistence diagrams could be computed via different means, it is a demonstration of how to apply the results from the previous chapter.

Proposition 4.9. *Let $f : D \to \mathbb{R}$ and let $g = h(f)$ where $h : \mathbb{R} \to \mathbb{R}$ is a monotone increasing bijection. Let $M(f), M(g)$ be the persistence modules induced by the sublevel*

set filtrations of D by f and g. Then $M(g)$ and $M(f)$ are (τ, σ)-interleaved where

$$\tau(t) = h^{-1}(t) \quad \textit{and} \quad \sigma(t) = h(t),$$

and there exists a perfect matching $\mathsf{PD}_{\mathcal{X}} : \mathsf{PD}(M(g)) \nrightarrow \mathsf{PD}(M(f))$ such that if $\mathsf{PD}_{\mathcal{X}}(b, d) = (b', d')$,

$$b' = h^{-1}(b) \quad \textit{and} \quad d' = h^{-1}(d).$$

In particular, every point in $\mathsf{PD}(M(f))$ and $\mathsf{PD}(M(g))$ is matched.

Proof. We begin by establishing the interleaving at the level of the sublevel set filtrations. Let $t \in \mathbb{R}$. Then

$$\mathsf{C}(g, t) = \mathsf{C}(h \circ f, t) = \mathsf{C}(f, h^{-1}(t))$$

and

$$\mathsf{C}(f, t) = \mathsf{C}(h \circ f, h(t))) = \mathsf{C}(g, h(t)).$$

Thus, for $s \leq t$, after applying homology, we have the following commutative diagram, where horizontal maps are induced by inclusion and vertical maps are isomorphisms.

$$\begin{array}{ccc} M(g)_s & \xrightarrow{\varphi_{M(g)}(s,t)} & M(g)_t \\ \downarrow & & \downarrow \\ M(f)_{h^{-1}(s)} & \xrightarrow{\varphi_{M(f)}(h^{-1}(s), h^{-1}(t))} & M(f)_{h^{-1}(t)} \end{array}$$

Thus, we see that $M(g)$ and $M(f)$ are (τ, σ)-interleaved with $\tau(t) = h^{-1}(t)$ and $\sigma(t) = h(t)$. Since h is bijective, Corollary 4.2 guarantees the existence of the matching given in the statement of the proposition. □

Thus, even if $h(t) < t$ or $h^{-1}(t) < t$ for some $t \leq \mathbb{R}$, the result still shows that we can determine the location of the persistence points of $\mathsf{PD}(f)$ by computing the persistence points of $\mathsf{PD}(g)$, without first applying the inverse function h^{-1} to the function g.

4.1.2 Discretizing a Persistence Module

As indicated in the introduction to Part I, in many applications we wish to use the $\mathbb{R}$-index persistence module $M(f)$ to characterize the geometry of a function $f\colon X \to \mathbb{R}$ by examining the sublevel set filtration it induces on X. However, in practice we can only perform a finite number of calculations. An idealized approximation is to assume that we can compute the *discretized $\mathbb{Z}$-indexed persistence module* $M(f)_t^{\mathbb{Z}}$ defined as follows. Set

$$M(f)_t^{\mathbb{Z}} := M(f)_{\lfloor t \rfloor} \quad \text{and} \quad \varphi_{M(f)^{\mathbb{Z}}}(s,t) := \varphi_{M(f)}(\lfloor s \rfloor, \lfloor t \rfloor),$$

where $\lfloor \cdot \rfloor$ is the floor function. Proposition 4.11 provides an answer to the following question: given the discretized $\mathbb{Z}$-indexed persistence module $M(f)^{\mathbb{Z}}$, what are the constraints on the persistence diagram associated to the persistence module $M(f)$? To begin answering this question, we find the following corollary of the Algebraic Stability Theorem for Generalized Interleavings useful. (Note as in Corollaries 3.14 and 4.2, we rewrite the inequalities so that we can estimate the position of the persistence points from the approximated persistence module via the position of the persistence points from the approximation persistence module.)

Corollary 4.10. *Let V and W be (τ,σ)-interleaved with $\tau(t) = t$ and σ right-continuous. Then there exists a matching $\mathsf{PD}_{\mathcal{X}} : \mathit{Rep}(\mathsf{PD}(V)) \nrightarrow \mathit{Rep}(\mathsf{PD}(W))$ on the persistence diagrams of V and W such that for $\mathsf{PD}_{\mathcal{X}}(b,d) = (b',d')$, the inequalities $\sigma_L^{-1}(b) \leq b' \leq b$ and $\sigma_L^{-1}(d) \leq d' \leq d$ hold, and any unmatched points (c,e) in either persistence diagram have $e \leq \sigma(c)$.*

Proposition 4.11. *Let (V,φ_V) be an $\mathbb{R}$-indexed PFD persistence module and let $(V^{\mathbb{Z}},\varphi_{V^{\mathbb{Z}}})$ be the associated discretized $\mathbb{Z}$-indexed PFD persistence module. Then $V^{\mathbb{Z}}$ and V are (τ,σ)-interleaved, where $\tau(t) = t$ and $\sigma(t) = \lfloor t+1 \rfloor$.*

Proof. Define a map $\phi : V^{\mathbb{Z}} \to V$ by $\phi_t := \varphi_V(\lfloor t \rfloor, t)$ and a map $\psi : V \to T_\sigma(V^{\mathbb{Z}})$ by $\psi_t := \varphi_V(t, \lfloor t+1 \rfloor)$. It is left to the reader to check that ϕ and ψ are persistence module morphisms which yield a (τ,σ)-interleaving of $V^{\mathbb{Z}}$ with V. □

Corollary 4.12. *Let $(V^{\mathbb{Z}}, \varphi_{V^{\mathbb{Z}}})$ be the associated discretized $\mathbb{Z}$-indexed persistence module of an $\mathbb{R}$-indexed PFD persistence module (V, φ_V). There exists a matching $\mathsf{PD}_{\mathcal{X}} : \mathsf{PD}(V^{\mathbb{Z}}) \nrightarrow \mathsf{PD}(V)$ such that if $\mathsf{PD}_{\mathcal{X}}(b,d) = (b',d')$, then*

$$b - 1 \leq b' \leq b \quad \textit{and} \quad d - 1 \leq d' \leq d.$$

Additionally, any unmatched points $(c,e) \in \mathsf{PD}(V^{\mathbb{Z}})$ satisfy $e = c + 1$, and unmatched points in $(c',e') \in \mathsf{PD}(V)$ satisfy $e' \leq \lfloor c' + 1 \rfloor$.

Proof. First observe that $\sigma_L^{-1}(t) = \lceil t - 1 \rceil$ and any persistence point $(b,d) \in \mathsf{PD}(V^{\mathbb{Z}})$ has $b, d \in \overline{\mathbb{Z}}$. Thus, a direct application of Corollary 4.10 to Proposition 4.11 yields the result. In particular, any unmatched points $(c,e) \in \mathsf{PD}(V^{\mathbb{Z}})$ actually satisfy $e = c + 1$ (versus $e \leq c + 1$, as is given by the theorem). □

See Figure 1.3 for an illustration of an estimate of $\mathsf{PD}(V)$ from $\mathsf{PD}(V^{\mathbb{Z}})$.

4.1.3 An Example from Image Analysis

We now consider an example that illustrates the effect of using a pixelated image as an approximation for a continuous function. To be more precise, consider a function $f\colon D \to \mathbb{R}$ where D is a rectangular region tiled by a cubical complex K [18]. An *ε-digital approximation of f over K* is a piecewise constant function $g\colon D \to \mathbb{Q}$ that is constant on the interior of each cell in K and satisfies

$$\|f - g\|_\infty \leq \varepsilon.$$

Assume that we can bound values of g, for example by a function $h\colon D \to \mathbb{Z}$ defined by

$$h(x) = \lceil g(x) \rceil,$$

where $\lceil \cdot \rceil : \mathbb{R} \to \mathbb{Z}$ is the function that rounds up to the next integer ($\lceil \cdot \rceil$ is the identity on $\mathbb{Z}$). We are interested in bounds on the persistence module $M(f)$ in terms of the persistence module that we can compute, $M(h)$.

Proposition 4.13. *Let $g : K \to \mathbb{Q}$ be an ε-digital approximation of $f\colon D \to \mathbb{R}$. Define $h\colon D \to \mathbb{Z}$ by $h(x) = \lceil g(x) \rceil$. Consider the sublevel set PDF persistence modules $(M(f), \varphi_f)$ and $(M(h), \varphi_h)$. Then $M(h)$ and $M(f)$ are (τ, σ)-interleaved, where*

$$\tau(t) = t + \varepsilon \quad \textit{and} \quad \sigma(t) = \lfloor t + 1 \rfloor + \varepsilon.$$

Proof. Since $||g - f||_\infty \leq \varepsilon$, $M(g)$ and $M(f)$ are ε-interleaved with translation maps $\varepsilon_g(t) = t + \varepsilon = \varepsilon_f(t)$. This can be seen by noting that

$$\mathsf{C}(f, t) \subseteq \mathsf{C}(g, t + \varepsilon) \qquad \text{and} \qquad \mathsf{C}(g, t) \subseteq \mathsf{C}(f, t + \varepsilon)$$

for all $t \in \mathbb{R}$, and taking the interleaving maps to be the linear maps induced by the corresponding inclusion maps at the level of homology. We leave it to the reader to check that a similar idea as in Proposition 4.11 leads to showing that $M(h)$ and $M(g)$ are (ρ, η)-interleaved with $\rho(t) = t$ and $\eta(t) = \lfloor t+1 \rfloor$. Thus, by Proposition 2.28, $M(h)$ and $M(f)$ are (τ, σ)-interleaved. □

Corollary 4.14. *Let $(M(f), \varphi_f)$ and $(M(h), \varphi_h)$ be the PDF persistence modules defined in Proposition 4.13. Then there exists a matching $\mathsf{PD}_{\mathcal{X}} : \mathsf{PD}(M(h)) \nrightarrow \mathsf{PD}(M(f))$ such that if $\mathsf{PD}_{\mathcal{X}}(b, d) = (b', d')$, then*

$$\lceil b - \varepsilon \rceil - 1 \leq b' \leq b + \varepsilon \quad \textit{and} \quad \lceil d - \varepsilon \rceil - 1 \leq d' \leq d + \varepsilon,$$

and unmatched points in $\mathsf{PD}(M(h))$ have $d \leq \lfloor b + 1 + \varepsilon \rfloor + \lfloor \varepsilon \rfloor$, while unmatched points in $\mathsf{PD}(M(f))$ have $d' \leq \lfloor b' + 1 \rfloor + 2\varepsilon$.

Proof. First note that $(\sigma \circ \tau)$ is a right-continuous translation map and that

$$\tau^{-1}(t) = t - \varepsilon \quad \text{and} \quad \sigma_L^{-1}(t) = \lceil t - \varepsilon \rceil - 1.$$

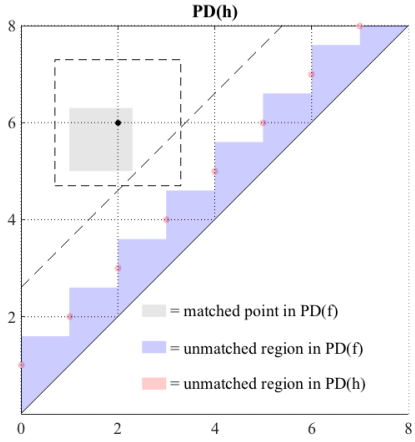


Figure 4.1: A schematic diagram illustrating the quality of the approximation given in Section 4.1.3 where h is the discretization of an ε-digital approximation of f with $\varepsilon = .3$. For clarity, we suppose that a persistence diagram with a single off-diagonal point at $(2, 6)$ has been computed from the approximation persistence module. Using the classical bottleneck distance, the square and diagonal strip defined by the dashed lines indicate the possible locations of the persistence points in the persistence diagrams of $M(f)$. In contrast, the matchings guaranteed by Theorem 3.13 limit the possible location of the associated matched point in the persistence diagrams of interest to the grey region. If any persistence points in $\mathsf{PD}(M(h))$ had fallen in the pink region, then it's possible they may have been introduced as a computational artifact, and may not have corresponding matched points in the persistence diagrams of interest.

In addition, we have

$$(\sigma \circ \tau)(t) = \lfloor t + 1 + \varepsilon \rfloor + \varepsilon,$$
$$(\sigma \circ \tau)_L^{-1}(t) = \lceil t - \varepsilon \rceil - 1 - \varepsilon,$$
$$(\sigma \circ \tau \circ \tau^{-1})(t) = \lfloor t + 1 \rfloor + \varepsilon.$$

Furthermore, for any $(b, d) \in \mathsf{PD}(M(h))$, we know that $b, d \in \mathbb{Z}$. Thus, with a bit of work one can verify the stated matching directly. □

For an illustration of an estimate of $\mathsf{PD}(M(f))$ from $\mathsf{PD}(M(h))$, see Figure 4.1. To conclude this section, we point out a possible implication of the quality of the error bounds computed in these two examples. One common problem in topological data analysis and persistent homology is the determination of a level of noise inherent in

the system; that is, a level beyond which lifespans of points in the persistence diagram are considered to be significant. This noise level is sometimes set by the known measurement error inherent in the system. If an approximation persistence module is (τ, σ)-interleaved with a 'true' persistence module, one possibility is to set the noise level of the system to be the maximum of the functions defining the unmatched regions for both the approximation and the 'true' persistence diagrams. In the context of Propositions 4.11 and 4.13, Figure 4.1 illustrates that the gray region around point $(2, 6)$ in the computed diagram is bounded away from the noise level, while the gray region in Figure 4.1 intersects it. Thus, it cannot be guaranteed that the matched point in $\mathsf{PD}(M(f))$ associated to point $(2, 6)$ is distinguishable from noise, while this is guaranteed in Proposition 4.11.

4.1.4 Two Examples for Large Point Clouds

Consider a finite a metric space (X, d). The persistence modules induced by the Vietoris-Rips filtrations associated to X (one for each homological dimension $n \in \mathbb{N}$) provide geometric information that can be associated to X. The following two sections both show how the Algebraic Stability Theorem for Generalized Interleavings (and its corollaries) can be used to keep track of errors introduced when computing the persistence diagrams induced by approximations of Vietoris-Rips complex filtrations.

Subsampling a large point cloud

Observe that for sufficiently large $t \in \mathbb{R}$, the complex $\mathsf{R}(X, t)$ is a simplex of dimension one less than the number of points in X. Thus, for large X and t, calculating the homology of $\mathsf{R}(X, s)$ for every $0 \leq s \leq t$ is computationally infeasible. However, given a dense point cloud, it is reasonable to assume that if one chooses an appropriate subset $Y \subset X$ and constructs the associated Vietoris-Rips filtration $\{\mathsf{R}(Y, t)\}_{t \in \mathbb{R}}$, the resulting persistence module $(M^{\mathsf{R}}(Y), \varphi_Y)$ will provide a reasonable approximation of $(M^{\mathsf{R}}(X), \varphi_X)$. Since it is understood that we are only concerned with the persistence modules induced by Vietoris-Rips filtrations associated to X and Y, in what follows we will drop the reference to the complex type and write $(M(Y), \varphi_Y)$ and $(M(X), \varphi_X)$.

The following proposition, whose proof is provided at the end of this section, quantifies this assumption.

Proposition 4.15. *Let (X, d) be a finite metric space. If $Y \subset X$ is such that for every $x \in X$ there exists a $y \in Y$ such that $d(x, y) \leq \delta$, then the persistence modules $M(Y)$ and $M(X)$ are (τ, σ)-interleaved, where $\tau(t) = t$ and $\sigma(t) = t + \delta$.*

Since both τ and σ are invertible, we can apply Corollary 4.2 to estimate $\mathsf{PD}(M(X))$ from $\mathsf{PD}(M(Y))$ as follows.

Corollary 4.16. *There exists a matching $\mathsf{PD}_{\mathcal{X}} : \mathsf{PD}(M(Y)) \nrightarrow \mathsf{PD}(M(X))$ of the persistence diagrams corresponding to the persistence modules $M(Y)$ and $M(X)$ with the following properties. If $\mathsf{PD}_{\mathcal{X}}(b, d) = (b', d')$ then $b - \delta \leq b' \leq b$ and $d - \delta \leq d' \leq d$. Moreover, all unmatched points in $\mathsf{PD}(M(Y))$ and $\mathsf{PD}(M(X))$ are at most δ above the diagonal.*

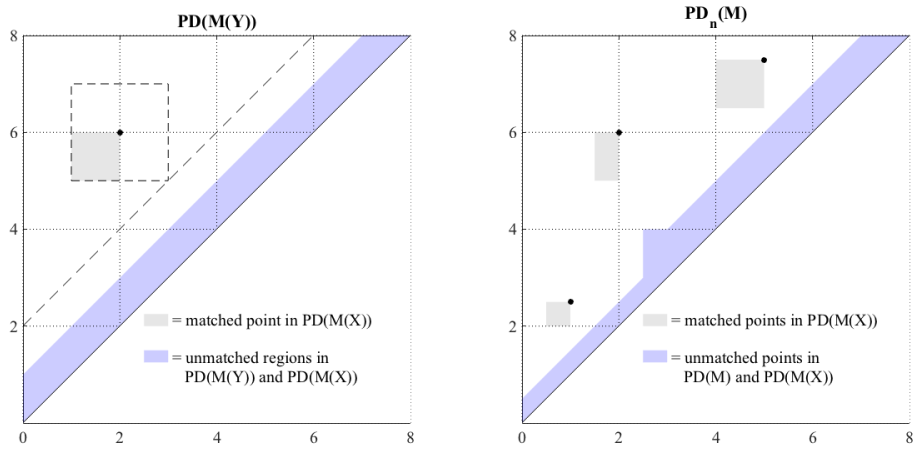
Figure 4.2(a) illustrates this approximation. To prove Proposition 4.15, we need a series of lemmas and definitions.

Lemma 4.17. *Let S and S' be subsets of a point cloud X and $t, \delta \geq 0$. Suppose that $\gamma : S \to S'$ is a map which satisfies $d(x, \gamma(x)) \leq \delta$ for all $x \in S$. Then the map*

$$\tilde{\gamma} : \mathsf{R}(S, t) \to \mathsf{R}(S', t + \delta)$$

defined by $\tilde{\gamma}[x_0, \cdots, x_k] = [\gamma(x_0), \cdots, \gamma(x_k)]$ is a simplicial map.

Proof. To prove that $\tilde{\gamma}$ is a simplicial map, we need to show that for every k-simplex $[x_0, ..., x_k] \in \mathsf{R}(S, t)$, the k-simplex $[\gamma(x_0), ..., \gamma(x_k)]$ is a simplex in $\mathsf{R}(S', t+\delta)$. Since the simplices in a Vietoris-Rips complex are fully determined by its 1-skeleton, we only need to show that the 1-skeleton of $\mathsf{R}(S, t)$ is mapped to the 1-skeleton of $\mathsf{R}(S', t+\delta)$. If $[x, y]$ is an edge in $\mathsf{R}(S, t)$, then by the definition of the Vietoris-Rips complex, $d(x, y) \leq 2t$.



(a) (b)

Figure 4.2: Schematic diagrams illustrating the quality of the matching from Corollaries 4.2 and 4.10 for (a) $M(Y)$ the persistence module given by a δ-dense subsample $Y \subseteq X$ with $\delta = 1$, (b) M the adaptive persistence module for a δ-cover tree of length one with $\delta = 0.5$ and $t_1 = 3$. For (b), we show a persistence diagram with three off-diagonal points at $(1, 2.5), (2, 6)$, and $(5, 7.5)$. The gray regions surrounding these points illustrate the possible locations of the associated matched points in the persistence diagram of $M(X)$. The blue region indicates the possible locations of unmatched points in the approximated persistence diagrams as well as unmatched points in the persistence diagram of $M(X)$.

Thus, we have

$$\begin{aligned} d(\gamma(x), \gamma(y)) &\leq d(\gamma(x), x) + d(x, y) + d(y, \gamma(y)) \\ &\leq \delta + 2t + \delta \\ &= 2(t + \delta), \end{aligned}$$

and so $[\gamma(x), \gamma(y)]$ is either a 1-simplex or a 0-simplex in $\mathsf{R}(S', t + \delta)$. □

We aim to construct the (τ, σ)-interleaving of $M(Y)$ and $M(X)$ in Proposition 4.15 explicitly, where $\tau(t) = t$ and $\sigma(t) = t + \delta$. This amounts to showing the existence of a persistence module morphism $\psi : M(X) \to T_\sigma(M(Y))$ such that the collection of maps

$\psi_t \colon M(X)_t \to M(Y)_{t+\delta}$ makes the diagram

$$\begin{array}{ccc} M(Y)_t & \xrightarrow{\varphi_{M(Y)}(t,t+\delta)} & M(Y)_{t+\delta} \\ \downarrow & \overset{\psi_t}{\nearrow} & \downarrow \\ M(X)_t & \xrightarrow{\varphi_{M(X)}(t,t+\delta)} & M(X)_{t+\delta} \end{array} \tag{4.18}$$

commute for all $t \in \mathbb{R}$, where the vertical arrows are induced by inclusion maps $i_t \colon \mathsf{R}(Y,t) \hookrightarrow \mathsf{R}(X,t)$. To achieve this goal we will use a simplicial homotopy. For a treatment of this material, see [19, Chapter 8].

Definition 4.19. Let K and L be simplicial complexes and $f, g : K \to L$ simplicial maps. Let $[v_0, \cdots, v_k]$ be a k-simplex in K, and define the maps

- Remove vertex v_i: $\partial_i[v_0, \cdots, v_k] = [v_0, \cdots, v_{i-1}, v_{i+1}, \cdots, v_k]$
- Repeat vertex v_i: $\rho_i[v_0, \cdots, v_k] = [v_0, \cdots, v_i, v_i, \cdots, v_k]$

The maps f and g are said to be *simplicially homotopic* if there exist morphisms $h_i : K_k \to L_{k+1}$ $(i = 0, ..., k)$ from the k-skeleton of K to the $(k+1)$-skeleton of L such that $\partial_0 h_0 = f$ and $\partial_{k+1} h_k = g$. Additionally, the maps $\{h_i\}$ must satisfy

$$\partial_i h_j = \begin{cases} h_{j-1}\partial_i & \text{if } i < j \\ \partial_i h_{i-1} & \text{if } i = j \neq 0 \\ h_j \partial_{i-1} & \text{if } i > j+1 \end{cases}$$

and

$$\rho_i h_j = \begin{cases} h_{j+1}\rho_i & \text{if } i \leq j \\ h_j \rho_{i-1} & \text{if } i > j. \end{cases}$$

The collection of maps $\{h_i\}$ is called a *simplicial homotopy* from f to g, written $f \simeq g$.

It is known [19, Lemma 8.3.13, Theorem 8.3.8] that if $f \simeq g$, then the induced maps on homology have $f_* = g_*$. Let S, S' be subsets of some point cloud X. By Lemma 4.17, any map $\gamma \colon S \to S'$ satisfying $d(x, \gamma(x)) \leq \delta$ can be extended to a

simplicial map $\tilde{\gamma}\colon \mathsf{R}(S,t) \to \mathsf{R}(S', t+\delta)$ by defining $\tilde{\gamma}[x_0, \cdots, x_k] = [\gamma(x_0), \cdots, \gamma(x_k)]$. In the following lemma we show another useful result concerning the map $\tilde{\gamma}$.

Lemma 4.20. *Let S, S' be subsets of a point cloud X and $t, \delta \geq 0$. Suppose that $\gamma\colon S \to S'$ is a map satisfying $d(x, \gamma(x)) \leq \delta$ for all $x \in S$ and let $\tilde{\gamma}\colon \mathsf{R}(S,t) \to \mathsf{R}(S', t+\delta) \subseteq \mathsf{R}(X, t+\delta)$ be its simplicial extension. If the maps $i'\colon \mathsf{R}(S',t) \hookrightarrow \mathsf{R}(X, t+\delta)$ and $i\colon \mathsf{R}(S, t+\delta) \to \mathsf{R}(X, t+\delta)$ are inclusions, then $i \circ \tilde{\gamma} \simeq i'$ and so $(i \circ \tilde{\gamma})_* = i'_*$ at the level of homology.*

Proof. We will define the simplicial homotopy $\{h_i\}$ from the map $i \circ \tilde{\gamma}$ to the inclusion map i' by $h_i([x_0, \cdots, x_n]) := [x_0, \cdots, x_i, \gamma(x_i), \cdots, \gamma(x_n)]$. To prove that the maps h_i define a simplicial homotopy, we start by showing that

$$\begin{aligned} \partial_0 h_0[x_0, \cdots, x_n] &= \partial_0[x_0, \gamma(x_0), \cdots, \gamma(x_n)] \\ &= [\gamma(x_0), \cdots, \gamma(x_n)] \\ &= (i \circ \tilde{\gamma})([x_0, \cdots, x_n]) \end{aligned}$$

and

$$\begin{aligned} \partial_{n+1} h_n[x_0, \cdots, x_n] &= \partial_{n+1}[x_0, \cdots, x_n, \gamma(x_n)] \\ &= [x_0, \cdots, x_n] \\ &= i'[x_0, \cdots, x_n]. \end{aligned}$$

The remaining properties of the maps h_i can be verified by a simple calculation as well. Thus, the map $i \circ \tilde{\gamma}$ is simplicially homotopic to ι and so $(i \circ \tilde{\gamma})_* = \iota_*$ at the level of homology. □

Proof of Proposition 4.15. Recall that we need to construct a family of maps ψ_t such that diagram (4.18) commutes for all $t \in \mathbb{R}$. By assumption, there is a map $\gamma\colon X \to Y$ such that $d(x, \gamma(x)) \leq \delta$. For each $t \in \mathbb{R}$, define a map $\tilde{\gamma}_t\colon \mathsf{R}(X,t) \to \mathsf{R}(Y, t+\delta)$ by $\tilde{\gamma}_t[x_0, \cdots, x_k] := [\gamma(x_0), \cdots, \gamma(x_k)]$. By Lemma 4.17, $\tilde{\gamma}_t$ is a simplicial map. Set $\psi_t := (\tilde{\gamma}_t)_*$. It is easy to check that $\psi : M(X) \to T_\sigma(M(Y))$ is a persistence module

morphism since each map ψ_t is a simplicial extension of the same vertex map γ.

The upper triangle of diagram (4.18) commutes by the following argument. If $i_t \colon \mathsf{R}(Y,t) \to \mathsf{R}(X,t)$ is the inclusion map, then the map $\tilde{\gamma}_t \circ i_t$ restricts to the identity on $\mathsf{R}(Y,t)$ and so at the level of homology it is equal to $\varphi_M(t,t+\delta)$. The commutativity of the lower triangle is a direct consequence of Lemma 4.20. □

An Application for Cover Trees

Suppose that $Y \subset X$ such that for any $y \in Y$, there exists an $x \in X$ with $d(y,x) \leq \delta$. The computational cost of constructing the interval decomposition of the persistence module $M(Y)$ decreases when we increase the value of δ. Unfortunately, the approximation error also increases with δ. Therefore, choosing a large value of δ might lead to a poor approximation of $\mathsf{PD}(M(X))$ by $\mathsf{PD}(M(Y))$. Note that the number of simplices in the Vietoris-Rips complex increases with the filtration parameter t. This implies that the barcodes of the persistence module $M(Y)$ might be computable only for a restriction of $M(Y)$ to a short interval $[0,t]$. By a restriction of the persistence module $M(Y)$ we mean a persistence module with the vector spaces $M(Y)_s$ for $s \leq t$ and $M(Y)_t$ for $s > t$.

To overcome this problem, we construct an approximation of $M(X)$ that is based on the concept of a cover tree [20].

Definition 4.21. We say that the sequence of subsets $\{Y_i \subseteq X\}_{i=0}^{m}$ is a *δ-cover tree of length m* for X, denoted $\mathcal{Y}_\delta$, if

- (Nesting) The sequence satisfies $Y_i \supset Y_{i+1}$ for all $i < m$.
- (Covering) For every $p \in Y_{i-1}$, $0 < i \leq m$, there exists a $q \in Y_i$ such that $d(p,q) \leq \delta 2^i$.
- (Sparsity) For distinct $p, q \in Y_i$, $d(p,q) > \delta 2^i$.

We call the subset $Y_i \subset X$ the cover at level i.

Suppose that each $M(Y_i)$ is computable on some interval $[0, t_{i+1}]$, and $0 = t_0 \leq t_1 \leq \cdots \leq t_m$. To obtain an approximation of $M(X)$, we will stitch the restricted

persistence modules $M(Y_i)$ together into a single persistence module M (see Definition 4.22, below). We call this persistence module *adaptive* in that the intervals on which each persistence module $M(Y_i)$ is computable may be system-dependent, and not based on some theoretical term computed in advanced. Ideally, we would like the vector spaces M_t to be equal to $M(Y_i)_t$ for $t \in [t_i, t_{i+1}]$. However, to ensure that M is a persistence module, we will need to modify this condition slightly. First, we will explain how to build the adaptive approximation persistence module M for $m = 1$, where $\mathcal{Y}_\delta$ is a δ-cover tree of length one for X. In this case, we will stitch together the two persistence modules $(M(Y_0), \varphi_0)$ and $(M(Y_1), \varphi_1)$ at some index $t_1 \in \mathbb{R}$. To do this, we need to define the transition maps between them. By the covering condition on $\mathcal{Y}_\delta$, this guarantees the existence of a map $\phi_0 \colon Y_0 \to Y_1$ such that $d(x, \phi_0(x)) < 2\delta$ for all $x \in Y_0$. Thus, by Lemma 4.17, the map $\tilde{\phi_0}$ induces a map $\Phi_0 : M(Y_0)_{t_1} \to M(Y_1)_{t_1+2\delta}$ after passing to the level of homology. The gap of length 2δ in the indices of the map motivates the following choice of vector spaces for the persistence module M.

$$M_t := \begin{cases} M(Y_0)_{\min(t,t_1)} & \text{for } \ 0 \leq t < t_1 + 2\delta \\ M(Y_1)_t & \text{for } \ t_1 + 2\delta \leq t. \end{cases}$$

The transition maps for the persistence module M are defined by

$$\varphi(s,t) := \begin{cases} \varphi_0(\min(s,t_1), \min(t,t_1)) & \text{for } \ 0 \leq s \leq t < t_1 + 2\delta \\ \varphi_1(t_1 + 2\delta, t) \circ \Phi_0 \circ \varphi_0((\min(s,t_1), t_1) & \text{for } \ 0 \leq s < t_1 + 2\delta \leq t \\ \varphi_1(s,t) & \text{for } \ t_1 + 2\delta \leq s \leq t. \end{cases}$$

To construct the full adaptive persistence module M for a δ-cover tree $\mathcal{Y}_c$ with $m > 1$, we repeat the above merging process for each set Y_i. Again by the covering condition, there exist maps $\phi_i \colon Y_i \to Y_{i+1}$ such that $d(x, \phi_i(x)) \leq 2^{i+1}\delta$ for all $x \in Y_i$. As before, these maps induce a collection of maps $\Phi_i : M(Y_i)_{t_{i+1}} \to M(Y_{i+1})_{t_{i+1}+2^{i+1}\delta}$. Thus, we can define the adaptive persistence module M as follows.

Definition 4.22. Let $\mathcal{Y}_\delta$ be a δ-cover tree for X of length m and $0 = t_0 < t_1 < \ldots < t_m < \infty$ be an increasing sequence such that $t_{i+1} > t_i + 2^i\delta$ for $i = 0, ..., m-1$ and for

which $M(Y_i)$ is computable on $[0, t_i)$. We define the *adaptive PDF persistence module* M as follows. The vector spaces of the persistence module M are given by

$$M_t := M(Y_i)_{\min(t,t_{i+1})} \text{ for } t \in I_i,\ i = 0, ..., m,$$

where

$$I_i := \begin{cases} [0, t_1 + 2\delta) & \text{for } i = 0 \\ [t_i + 2^i\delta, t_{i+1} + 2^{i+1}\delta) & \text{for } i = 1, ..., m-1 \\ [t_m + 2^m\delta, \infty) & \text{for } i = m. \end{cases}$$

The transition maps are defined by

$$\varphi_M(s,t) := \begin{cases} \varphi_i(\min(s, t_{i+1}), \min(t, t_{i+1})) & \text{for } s,t \in I_i, \\ & i = 0, ..., m \\ \varphi_j(t_j + 2^j\delta, \min(t, t_{j+1})) \circ \Phi_{i,j} \circ \varphi_i(\min(s, t_{i+1}), t_{i+1}) & \text{for } s \in I_i, t \in I_j, \\ & 0 \le i < j \le m, \end{cases}$$

where $\Phi_{i,j} := \Phi_{j-1} \circ \varphi_{j-1}(t_{j-1} + 2^{j-1}\delta, t_j) \circ \cdots \circ \Phi_{i+1} \circ \varphi_{i+1}(t_{i+1} + 2^{i+1}\delta, t_{i+2}) \circ \Phi_i$.

We remark that the adaptive persistence module M depends on the choice of $\delta > 0$, the values $\{t_i\}$, and the subsamples $\{Y_i\}$. Indeed, different choices of Y_i could result in different adaptive persistence modules, even if δ and $\{t_i\}$ remain fixed. However, regardless of the exact choice of each of the subsets Y_i, the persistence modules $M(X)$ and M are interleaved in such a way that the upper bound of the error of the approximation depends only on δ and $\{t_i\}$.

Theorem 4.23. *Let X be a point cloud and M the adaptive PFD persistence module given by Definition 4.22. Then M and $M(X)$ are (τ, σ)-interleaved, where $\tau(t) = t$ and*

$$\sigma(t) := \begin{cases} t + 2^i\delta & \textit{for } t \in [t_i, t_{i+1} - 2^i\delta) \textit{ and } i = 0, ..., m \\ t_i + 2^i\delta & \textit{for } t \in [t_i - 2^{i-1}\delta, t_i) \textit{ and } i = 1, ..., m-1. \end{cases}$$

Proof. We recall that the adaptive persistence module M is determined by $\delta > 0$

and the sequences $\{t_i\}_{i=0}^m$ and $\{Y_i\}_{i=0}^m$. The proof proceeds by induction on m. For the case $m = 0$, the result is immediate by Proposition 4.15. Now suppose that M is an adaptive persistence module defined by $\delta > 0$, and the sequences $\{t_i\}_{i=0}^{m+1}$ and $\{Y_i\}_{i=0}^{m+1}$, while M' is an adaptive persistence module defined by the truncated sequences $\{t_i\}_{i=0}^m$ and $\{Y_i\}_{i=0}^m$. By our inductive hypothesis, M' and $M(X)$ are (τ, σ)-interleaved per the definitions in Theorem 4.23. By Proposition 4.15, $M(Y_{m+1})$ and $M(X)$ are $2^{m+1}\delta$ interleaved. Thus, Proposition 2.29 enables us to stitch together the persistence modules M' and $M(Y_{m+1})$ at t_{m+1}. The resulting persistence module M is as described since the interleaving maps are induced by nearest neighbor maps. Moreover, the translation maps specified by the interleaving between M and $M(X)$ also follows from Proposition 2.29. □

Thus, an application of Corollary 4.10 yields the following.

Corollary 4.24. *There exists a matching* $\mathsf{PD}_{\mathcal{X}} : \mathsf{PD}(M) \nrightarrow \mathsf{PD}(M(X))$ *of the persistence diagrams corresponding to the persistence modules* M *and* $M(X)$ *with the following properties. If* $\mathsf{PD}_{\mathcal{X}}(b, d) = (b', d')$*, then* $\sigma_L^{-1}(b) \leq b' \leq b$ *and* $\sigma_L^{-1}(d) \leq d' \leq d$*, where*

$$\sigma_L^{-1}(t) = \min(t - 2^i\delta, t_{i+1}) \text{ for } t \in I_i, i = 0, .., m.$$

Furthermore, all unmatched points $(c, e) \in \mathsf{PD}(M)$ *satisfy* $\sigma_L^{-1}(e) \leq c$*, and all unmatched points* $(c', e') \in \mathcal{B}_{M(X)}$ *satisfy* $e' \leq \sigma(c')$*.*

See Figure 4.2 for a visualization of the relationship between the persistence diagram of a persistence module $M(X)$ and the persistence diagram associated to an adaptive persistence module as in Definition 4.22. As a final remark, note that the persistence module above contains the familiar transition maps induced by inclusion of the underlying Vietoris-Rips complexes as well as transition maps induced by simplicial maps. Computing the interval decomposition of persistence modules stemming from simplicial maps is addressed in [21].

4.1.5 A Summary of Generalized Interleavings for Filtrations of Point Clouds

It is often not possible to compute the persistence diagram of a given filtration associated to a dataset directly, and instead an approximate persistence diagram is computed from an interleaved persistence module. For example, often a Vietoris-Rips filtration (Definition 4.4) is used in place of a Čech filtration (Definition 4.6), since the latter is very expensive to compute [4]. Persistence module interleavings are commonly stated in terms of an additive δ-interleaving with translation maps $\tau(t) = t + \delta = \sigma(t)$, $\delta \geq 0$, or a multiplicative interleaving (also called a c-approximation [10]) with translation maps $\tau(t) = ct = \sigma(t)$, $c \geq 1$. Accordingly, the numerical error introduced by these approximations is typically stated in terms of a bottleneck error in the first case, or a log-bottleneck error in the latter case, in many instances considerably overstating the amount of error introduced by the approximation.

In Table 1, we provide a list of persistence modules that approximate persistence modules induced by common filtrations of a finite point cloud in $\mathbb{R}^n$ or an arbitrary metric space. To use the table, suppose that a persistence module (V, φ_V) is induced by taking homology of the filtration in the first column, and a persistence module (W, φ_W) is induced by taking homology of the filtration in the second column. The persistence modules V and W are (τ, σ)-interleaved, where the maps τ and σ are given by the third and fourth columns in the table. In all instances, Corollary 3.14 may then be used to make precise statements about the errors introduced by the approximation.

The references in the first column give the location for where the interleaving is constructed. Most often, as in the case of Lemma 4.8, this is stated in terms of inclusions of the approximation and the original filtration, and the interleaving at the level of persistence modules is achieved by computing the linear maps induced by inclusion at the level of homology. However, as in Lemma 4.20 and Proposition 4.15, sometimes a simplicial homotopy (or similar notion) is required to obtain commutativity at the level of persistence modules.

Point clouds in $\mathbb{R}^n$			
Approximate filtration	*Original filtration*	$\tau(t)$	$\sigma(t)$
Vietoris-Rips [4, Proof of Vietoris-Rips Lemma.]	Čech	$t\sqrt{\frac{2n}{n+1}}$	t
Net-tree [8, Proof of Proposition 20]	Čech	t	$(1+\varepsilon)^2 t$
Graph induced complex [9, Proposition 2.8]	Vietoris-Rips	$t+2\varepsilon$	t
Sparsified Vietoris-Rips [21, Claim 6.1]	Vietoris-Rips	t	$(1+\varepsilon)t$
Point clouds in an arbitrary metric space			
Approximate filtration	*Original filtration*	$\tau(t)$	$\sigma(t)$
Vietoris-Rips (Lemma 4.8)	Čech filtration	$2t$	t
Relaxed Vietoris-Rips [10, Lemma 4]	Vietoris-Rips	t	$\left(\frac{1}{1-2\varepsilon}\right)t$
Sparse weighted Rips [22, Lemma 6.13]	Vietoris-Rips	t	$\left(\frac{1+\sqrt{1+\delta^2}\varepsilon}{1-\varepsilon}\right)t$

Table 4.1: A table of approximations for Vietoris-Rips and Čech filtrations for point clouds. The first column gives the approximation and a reference to the construction of the approximation (including the explicit construction of the interleavings), and the second column gives the filtration that is being approximated. The third and fourth columns list the translation maps for (τ, σ)-interleavings of the associated persistence modules induced by taking homology of the filtrations. The values $\delta, \varepsilon \geq 0$ are parameters specified by the approximations where applicable.

4.2 Manifold Estimation

An important application for homological methods is in manifold learning. Consider a compact k-dimensional Riemannian submanifold $\mathcal{M} \subset \mathbb{R}^N$. Let $\tau \in \mathbb{R}$ be the largest real number such that the open normal disk bundle of radius r of M is embedded in $\mathbb{R}^N$ for all $r < \tau$. Let $B_\varepsilon(x) \subset \mathbb{R}^N$ denote the open ball of radius ε about x. Define $\theta_1 := \arcsin(\varepsilon/8\tau)$ and $\theta_2 := \arcsin(\varepsilon/16\tau)$ and let $\varepsilon > 0$. Set

$$\beta_1 := \frac{vol(\mathcal{M})}{\cos^k(\theta_1) vol(B^k_{\varepsilon/4})} \quad \text{and} \quad \beta_2 := \frac{vol(\mathcal{M})}{\cos^k(\theta_2) vol(B^k_{\varepsilon/8})}$$

where $vol(B^k_\varepsilon)$ denotes the k-dimensional volume of the standard k-dimensional ball of radius ε. The following theorem is proven by Niyogi, Smale, and Weinberger in [23].

Theorem 4.25. *Let $\mathcal{M} \subset \mathbb{R}^N$ be a compact k-dimensional Riemannian manifold and let $\delta > 0$. Let $X = \{x_i \mid i = 0, \ldots, n\} \subset \mathcal{M}$ be a finite set of points drawn i.i.d. using the uniform probability measure on $\mathcal{M}$. Let $0 < \varepsilon < \tau(\mathcal{M})/2$. Set $U := \bigcup_{x \in X} B_\varepsilon(x)$. If*

$$n > \beta_1 \left(\log(\beta_2) + \log\left(\delta^{-1}\right) \right)$$

then $H_*(U) \cong H_*(\mathcal{M})$ with probability greater than $1 - \delta$.

The importance of this theorem is that it provides an abstract guarantee that with high probability it is possible to compute the homology of a manifold from a finite sampling of points. There are at least two challenges to applying this theorem in practice. First, the proof is based on determining the Čech complex associated with the balls used to define U, which is computationally expensive. Second, the manifold $\mathcal{M}$ is unknown and thus τ is unknown, and hence it is not clear a priori how ε and n are to be chosen. The advantage of persistent homology is that there is no particular scale that needs to be chosen. Thus, it is natural to ask if the persistence module of the Vietoris-Rips complex associated to X can provide information about $H_*(\mathcal{M})$. To simplify notation in what follows, we define

$$\nu_N := \sqrt{\frac{2N}{N+1}}.$$

The following proposition provides an alternative interpretation of the constants and probabilities of Theorem 4.25.

Proposition 4.26. *Under the assumptions of Theorem 4.25, let* $(b, d) \in \mathsf{PD}(M^{\mathsf{R}}(X))$ *with either* b *or* d *in the interval* $(\varepsilon, \tau\nu_N^{-1}/2)$*. Then with probability greater than* $1 - \delta$*,* $d \leq b\nu_N$ *and the point* (b, d) *is a numerical artifact that does not correspond to a feature of the underlying manifold.*

Proof. By Theorem 4.25, with probability greater than $1 - \delta$, the complex $\check{\mathsf{C}}(X, t)$ has the homology of $\mathcal{M}$ for $\varepsilon \leq t < \tau/2$. Hence, $\mathsf{PD}(M^{\check{\mathsf{C}}}(X))$ has no persistence points (c, e) with c or e in the interval $(\varepsilon, \tau/2)$. From Table 1, the persistence modules associated to the Vietoris-Rips and Čech filtrations are (ρ, σ)-interleaved with $\rho(t) = t\nu_N$ and $\sigma(t) = t$.

By Corollary 4.2, if the point (b, d) is matched to a point (c, e) in the persistence diagram corresponding to the Čech filtration of X, we must have $b \leq c \leq b\nu_N$ and $d \leq e \leq d\nu_N$. If $c, e \leq \varepsilon$ then $b, d \leq \varepsilon$, and if $\tau \leq c, e$ then $\tau\nu_N^{-1} \leq b, d$. Hence, if (b, d) has either b or d in the interval $(\varepsilon, \tau\nu_N^{-1}/2)$, then (b, d) must be unmatched and so $d \leq b\nu_N$. □

4.3 Addressing an Open Problem for Non-Uniform Sublevel Set Filtrations

Non-uniform sublevel set filtrations are examined by Bendich, et al. in [13]. Let D a compact topological space, $f : D \to \mathbb{R}$, and let $e : D \to (0, \infty)$. Consider the sets

$$\mathsf{C}(f, re) := \{x \in D : f(x) \leq re(x)\}$$

for each $r \in \mathbb{R}$, which can be viewed as a non-uniform sublevel set filtration of X. Taking n-dimensional homology with respect to a field k, we form a persistence module $M(f, e)$ with vector spaces $M(f, e)_r := H_n(\mathsf{C}(f, re), \mathsf{k})$ and transition maps induced by inclusion. Let $M(f)$ be the persistence module induced by the usual sublevel set filtration of D under f (with respect to the same dimension n and field k).

In the concluding remarks of [13], the authors note that the relationship between the difference of $\mathsf{PD}(M(f))$ and $\mathsf{PD}(M(f,e))$ and the function e is unknown. We show how Corollary 4.2 gives insight into this relationship by examining an interleaving of the persistence modules $M(f,e)$ and $M(f)$.

Take $\alpha = \inf_{x\in D} e(x)$, $\beta = \sup_{x\in D} e(x)$, and suppose that $\alpha > 0$, $\beta < \infty$. Set $r = a/\alpha$. Then $a = r\alpha \leq re(x)$, and thus $\mathsf{C}(f,a) \subseteq \mathsf{C}(f,re)$. Similarly, setting $r = b\beta$, we have $be(x) \leq b\beta = r$, and so $\mathsf{C}(f,be) \subseteq \mathsf{C}(f,r)$. Taking $\tau(t) = t/\alpha$ and $\sigma(t) = t\beta$, we see that $(\tau\circ\sigma)(t) = (\sigma\circ\tau)(t) = (\beta/\alpha)t \geq t$. Taking the interleaving morphisms induced by these inclusions, we see that $M(f,e)$ and $M(f)$ are (τ,σ)-interleaved and the translation maps are invertible. Hence there is a matching $\mathsf{PD}_{\mathcal{X}} : \mathsf{PD}(M(f,e)) \nrightarrow \mathsf{PD}(M(f))$ with any $\mathsf{PD}_{\mathcal{X}}(b,d) = (b',d')$ satisfying

$$\frac{b}{\beta} \leq b' \leq \frac{b}{\alpha} \quad \text{and} \quad \frac{d}{\beta} \leq d' \leq \frac{d}{\alpha},$$

and unmatched points (c,e) satisfy $e \leq c(\beta/\alpha)$ in both diagrams. Thus, in the case that $\alpha > 0$ and $\beta < \infty$, the difference between $\mathsf{PD}(M(f))$ and $\mathsf{PD}(M(f,e))$ is controlled by the minimum and maximum values that $e(x)$ takes on X.

4.4 Comments on Future Directions

Our hope is that this new perspective will yield statements of more precise error bounds for the computations of other objects used in TDA, for example posets [24], Reeb Graphs [25], merge trees [26], and multi-dimensional persistence modules [27], all of which have established interleaving theories. Recent developments give a notion of homotopy interleavings [28], various flavors of which are exemplified in the construction of the generalized interleavings in this chapter, so this work hints at a possible extension of this theory as well. Additionally, this approach might shed some light on problems concerning the Wasserstein metric for persistence diagrams, which involve statements about non-uniform distances between points in two persistence diagrams, as opposed to the Bottleneck metric, which only considers the single largest difference between the collections of points [6].

PART II:

Applications of Persistent Homology to Problems in Fluid Dynamics

Chapter 5

Dynamical Structures in Kolmogorov and Rayleigh-Bénard Convection Flow

We present here a very lightly edited version of [1], with many portions reproduced verbatim. In this work, we introduce new mathematical techniques for analyzing complex spatiotemporal nonlinear dynamics and demonstrate their efficacy in problems from two different paradigms in hydrodynamics. Our approach employs methods from persistent homology [4, 3]; earlier efforts have shown that computing the homology of topological spaces associated to scalar or vector fields generated by complex systems can provide new insights into dynamics [29, 30, 31, 32, 33, 34].

Complex spatiotemporal systems often exhibit complicated pattern evolution. The patterns are given by scalar or vector fields representing the state of the system under study. As discussed in the introduction to the previous chapter, persistent homology can be viewed as a map PD that assigns to every scalar field f a multiset of points in $\mathbb{R}^2$ called a *persistence diagram* $\mathsf{PD}(f)$. Essentially, we can view the persistence diagram as a highly simplified representation of the pattern encoded by the scalar field f.

The space of all persistence diagrams, Per, can be endowed with a variety of metrics under which PD is a Lipschitz function. This has several important implications that we exploit in this chapter. First, the Lipschitz property implies that small changes in the field pattern, e.g. bounded errors associated with measurements or numerical approximations, lead to small changes in the persistence diagrams. Second, by using different metrics, we can vary our focus of interest between larger and smaller changes in the persistence diagrams. Moreover, by comparing different metrics, we can infer if the changes in a pattern affect geometric features with longer or shorter lifespans. Finally, since applying the map PD to a time series of patterns produces a time series

in Per, the distance between the consecutive data points in Per can be used to quantify the rate at which the geometry of the patterns is changing.

As mentioned above, the dynamics of spatiotemporal systems are characterized by the time-evolution of the patterns corresponding to the scalar fields generated by the system. However, capturing these scalar fields, either experimentally or numerically, results in multi-scale high dimensional data sets. In order to efficiently analyze these data sets, a dimension reduction must be performed. We use persistent homology to perform nonlinear dimension reduction from a time series of patterns to a time series of persistence diagrams. We show that this reduction can cope with redundancies introduced by symmetries (both discrete and continuous) present in the system. In particular, this approach directly quotients out symmetries and, thereby, permits easy identification of solutions that lie on a group orbit. Alternative approaches to nonlinear dimension and symmetry reduction include both the method of slices [35] and recent advances in identifying unstable exact solutions of nonlinear partial differential equations [36]. While a detailed comparison of these methods is beyond the scope of this work, it is worth pointing out that the application of persistent homology does not rely on knowledge of the underlying governing equations.

Separately, we also apply persistent homology to extract information about dynamical structures in the reduced data. Characterizing dynamics in the space of persistence diagrams cannot currently be done using conventional methods (e.g., linear analysis of time delay embeddings), since choosing a coordinate system in Per is currently an open problem [37]. However, since Per is a metric space, the geometry of the point cloud X, generated by the time series of the reduced data, is encoded by a scalar field which assigns to each point in Per its distance to X. We show how persistent homology may be applied to describe dynamics by characterizing the geometry of X.

An outline of this chapter is as follows. In Section 5.1 we present an abreviated overview of the two fluid flows examined: (1) Kolmogorov flow and (2) Rayleigh-Bénard convection. We note here, for emphasis, that while persistent homology can be applied to vector fields, it will be sufficient for this chapter to focus on scalar fields drawn from these systems (specifically, one component of the vorticity field for Kolmogorov flow,

and the temperature field for Rayleigh-Bénard convection).

In Section 5.2 we discuss key issues related to the application of persistent homology. By now, the mathematical theory of persistent homology is well developed. Therefore, our main emphasis is on the computational aspect of passing from the data to the persistence diagrams. Section 5.3 describes the correspondence between the geometric features of a scalar field and the points in its corresponding persistence diagram. Section 5.4 discusses the structure of the space Per and the properties of the associated metrics.

In Sections 5.5 and 5.6 we discuss how these metrics can be used to analyze dynamics. First, we interpret distance between the persistence diagrams representing consecutive data points in the time series as a rate at which the geometry of the corresponding scalar fields is changing. Second, we motivate and explain the procedure for extracting the geometric structure of the point cloud in Per.

We close the chapter by applying the techniques developed to the following problems. In Section 5.7, we identify distinct classes of symmetry-related equilibria for Kolomogorov flow. In Section 5.8, we show that a relative periodic orbit for Kolmogorov flow collapses to a closed loop in Per. Finally, in Section 5.9, we deal with identifying recurrent dynamics that occur on different time scales in our study of Rayleigh-Bénard convection flow.

5.1 The Systems to be Studied

5.1.1 Kolmogorov Flow

For the study of turbulence in two dimensions, Kolmogorov proposed a model flow where the evolution of a two-dimensional (2D) velocity field $\mathbf{u}(x, y, t)$ is given by

$$\begin{aligned} \frac{\partial \mathbf{u}}{\partial t} + \beta \mathbf{u} \cdot \nabla \mathbf{u} &= -\frac{1}{\rho} \nabla p + \nu \nabla^2 \mathbf{u} - \alpha \mathbf{u} + \mathbf{f} \\ \nabla \cdot \mathbf{u} &= 0 \end{aligned} \tag{5.1}$$

(with $\beta = 1$ and $\alpha = 0$). In the above equation, $p(x, y)$ is the pressure field, ν is the kinematic viscosity, ρ is fluid density, and $\mathbf{f} = \chi \sin(\kappa y)\hat{\mathbf{x}}$ is the forcing that drives the flow [38]. Laboratory experiments in electromagnetically-driven shallow layers of electrolyte can exhibit flow dynamics that are well-described by Equations (5.1) with appropriate choices of β and α to capture three-dimensional effects, which are commonly present in experiments [39]. In this chapter, we refer to all models described by Equations (5.1) (including experimentally-realistic versions) as Kolmogorov flows.

It is convenient to use the vorticity-stream function formulation [40] to study Kolmogorov flow analytically and numerically. Equations (5.1), written in terms of the z-component of the vorticity field $\omega = (\nabla \times \mathbf{u}) \cdot \hat{\mathbf{k}}$, a scalar field, take the form

$$\frac{\partial \omega}{\partial t} + \beta \mathbf{u} \cdot \nabla \omega = \nu \nabla^2 \omega - \alpha \omega + \chi \kappa \cos(\kappa y). \tag{5.2}$$

For the current study, we choose $\beta = 0.83$, $\nu = 3.26 \times 10^{-6}$ m^2/s, $\alpha = 0.063$ s^{-1}, $\rho = 959$ kg/m^3, and $\lambda = 2\pi/\kappa = 0.0254$ m. We express the strength of the forcing in terms of a non-dimensional parameter, the Reynolds number $Re = \sqrt{\frac{\lambda^3 \chi}{8\nu^2}}$.

Equation (5.2) is solved numerically by using a semi-discrete, pseudo-spectral method [41], assuming periodic boundary conditions in both x and y directions, i.e., $\omega(x, y) = \omega(x + L_x, y) = \omega(x, y + L_y)$, where $L_x = 0.085$ m and $L_y = 4\lambda = 0.1016$ m are the dimensions of the domain in the x and y directions, respectively. The vorticity field is discretized in the Fourier space using 128×128 modes, which corresponds to spatially resolving the domain on a 2D mesh with spacing $\Delta x = L_x/128$ and $\Delta y = L_y/128$ in the x and y directions, respectively. A time step of $dt = 1/32$ is chosen for the temporal discretization.

It is important to note that Equation (5.2), with periodic boundary conditions, is invariant under any combination of three distinct coordinate transformations: (1) a translation along x: $\mathcal{T}_{\delta x}(x, y) = (x + \delta x, y)$, $\delta x \in [0, L_x]$; (2) a rotation by π: $\mathcal{R}(x, y) = (-x, -y)$; and (3) a reflection and a shift: $\mathcal{D}(x, y) = (-x, y + \lambda/2)$. Because of these symmetries, each particular solution to Equation (5.2) generates a set of solutions which are dynamically equivalent. Physically, invariance under continuous translation leads to

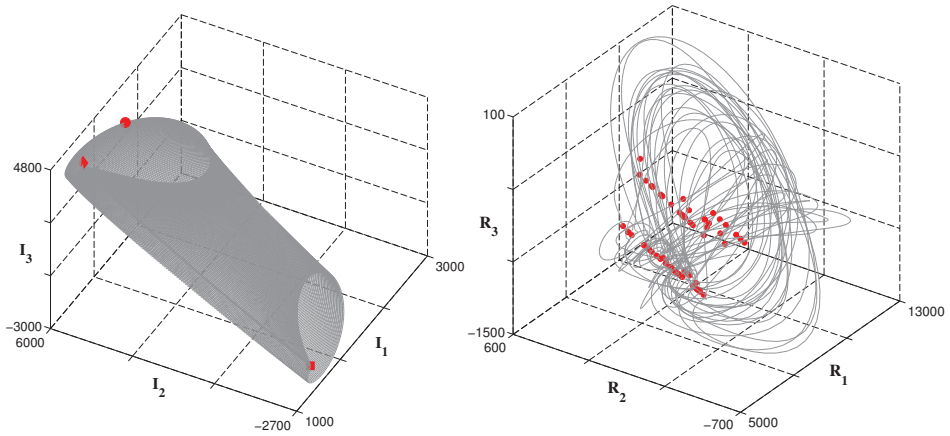
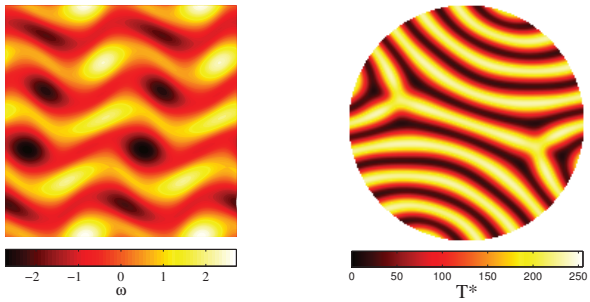


Figure 5.1: (a) Three-dimensional projections of a stable RPO at $Re = 25.43$ from the Kolmogorov flow using the imaginary part of the three dominant Fourier modes, I_1, I_2, and I_3. The gray line indicates the evolution of a RPO; three snapshots sampled from that orbit are indicated by a red diamond, a red circle, and a red square, which are analyzed below. (b) Three-dimensional projections of a turbulent trajectory, at $Re = 26.43$, using the real parts of the three dominant Fourier modes, R_1, R_2, and R_3. The gray line indicates the chaotic evolution of the flow, which is influenced by the presence of unstable fixed points, indicated by red circles, which are also analyzed below.

the existence of relative equilibria (REQ) and relative periodic orbit (RPO) solutions, in addition to equilibria (EQ) and periodic orbit (PO) solutions.

For $Re = 25.43$, the flow is characterized by a steady RPO; Figure 5.1(a) shows a projection, plotted using the three dominant Fourier modes of this RPO. The RPO has a period 34.78 seconds and a drift speed 1.354×10^{-6} m/s. The tunnel-like structure is a result of the periodic motion superposed over the slow drift along the x-direction. For larger forcing ($Re = 26.43$), the flow becomes weakly turbulent, as can be seen from the Fourier projections in Figure 5.1(b). The turbulent dynamics in this regime are of great interest as the flow explores a region of the state space which contains "weakly" unstable EQ, PO, REQ, and RPO solutions. Recent theoretical advances have shown that the identification of these solutions could aid the understanding of weakly turbulent dynamics [42]. For instance, if the turbulent trajectory is close to an EQ solution (ω_0), which is characterized by $\partial\omega_0/\partial t = 0$, we would expect the instantaneous rate of change of ω to be relatively small, i.e., $\partial\omega/\partial t \approx 0$. Similarly, a close pass to a PO solution would



(a) (b)

Figure 5.2: (a) A snapshot of the z-component of the vorticity field ω for Kolmogorov flow from the stable relative periodic orbit found at $Re = 25.43$. (b) A snapshot of the renormalized 8-bit mid-plane temperature field T^* for Rayleigh-Bénard convection from the stable almost-periodic orbit found at $Ra = 3000$ and $Pr = 1$.

mean $\omega(t+T) \approx \omega(t)$, where T is the period of the PO that is guiding the dynamics of the turbulent trajectory. The turbulent trajectory depicted in Figure 5.1(b) passes close to unstable EQ and REQ solutions which are indicated by the red dots.

5.1.2 Rayleigh-Bénard Convection

Rayleigh-Bénard convection is a canonical pattern forming system that has been used to gain many new fundamental insights into the spatiotemporal dynamics of systems that are driven far-from-equilibrium [43, 44]. Rayleigh-Bénard convection is the buoyancy driven fluid flow that occurs when a shallow layer of fluid is heated uniformly from below in a gravitational field. The dynamics are governed by the Boussinesq equations,

$$Pr^{-1}\left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u}\cdot\nabla\mathbf{u}\right) = -\nabla p + \nabla^2\mathbf{u} + RaT\hat{\mathbf{z}}, \tag{5.3}$$

$$\frac{\partial T}{\partial t} + \mathbf{u}\cdot\nabla T = \nabla^2 T, \tag{5.4}$$

$$\nabla\cdot\mathbf{u} = 0, \tag{5.5}$$

where $\mathbf{u}(x,y,z,t)$ is a vector field of the fluid velocity, $p(x,y,z,t)$ is the pressure field, and $T(x,y,z,t)$ is the temperature field. In our notation, the origin of the Cartesian coordinates (x,y,z) at the center of the domain are at the lower heated plate where $\hat{\mathbf{z}}$ is

a unit vector opposing gravity. Equations (5.3)-(5.5) represent the conservation of momentum, energy, and mass, respectively. The equations have been nondimensionalized using the vertical diffusion time of heat as the time scale, the layer depth as the length scale, and the constant temperature difference between the lower and upper plates as the temperature scale.

In our work, we consider Rayleigh-Bénard convection in a shallow domain with a cylindrical cross-section. The no-slip fluid boundary condition $\mathbf{u} = 0$ is applied to all material surfaces. The lower and upper plates are held at a constant temperature where $T(z = 0) = 1$ and $T(z = 1) = 0$, respectively. The lateral sidewalls of the cylindrical container are assumed to be perfectly conducting, which yields $T(z) = 1 - z$.

The dynamics can be described using three non-dimensional parameters. The Rayleigh number Ra represents the ratio of buoyancy to viscous forces. At the critical value $Ra_c = 1708$, an infinite layer of fluid undergoes a bifurcation to straight and parallel convection rolls. For increasing values of the Rayleigh number $Ra > Ra_c$, the dynamics become periodic, chaotic, and eventually turbulent. The Prandtl number Pr is the ratio of the momentum and thermal diffusivities. For typical experiments using compressed gasses, $Pr \approx 1$. Lastly, the aspect ratio of the cylindrical domain Γ is the ratio of the domain's radius to its depth.

We numerically integrate Eqs. (5.3)-(5.5) using a highly parallel spectral element algorithm that has been tailored for the study of convection (c.f. [45]). Figure 5.2(b) shows a typical pattern from a numerical simulation of Rayleigh-Bénard convection. In this simulation, $Ra = 3000$, $Pr = 1$, and the aspect ratio of the domain is $\Gamma = 10$. The numerical simulation is initiated from a field of small random perturbations to the temperature field and is integrated for long times. Figure 5.2(b) illustrates the fluid temperature field at the horizontal mid-plane ($z = 1/2$), where light is warm rising fluid and dark is cool falling fluid. This image is a snap shot in time of a time-dependent pattern where the dynamics are nearly periodic in time. The pattern shown does not include the region near the sidewall. Specifically, a distance of one-layer depth from the lateral sidewall is not shown (this distance is approximately the width of a convection roll). This is done to remove the complex fluid flow that occurs in the small region near

the sidewalls to allow our diagnostics to focus upon the bulk patterns and dynamics (c.f. [44]).

5.2 Persistent Homology

The aim of this work is to introduce an approach for analyzing the dynamics of the pattern evolution in spatiotemporal systems. This is done in two steps. First, we perform nonlinear data reduction, and then we extract information about the dynamical structures from this reduced data. We formulate both of these tasks in terms of analyzing the structure of the sub-level sets of a scalar function $f\colon D \to \mathbb{R}$, where D is a topological space. Tools from algebraic topology, homology in particular, are used to capture and quantify the geometry of the sub-level sets.

Recall that given any topological space Z, homology assigns to Z a sequence of groups $H_k(Z)$, $k = 0, 1, \ldots$. In the work that follows, we compute the homology groups using $\mathbb{Z}_2$ coefficients, and thus each $H_k(Z)$ is a vector space. The dimension of $H_k(Z)$ is called the k-th *Betti number* and is denoted by $\beta_k(Z)$. Betti numbers provide geometric information about X: $\beta_0(Z)$ is the number of connected components, or pieces, of Z; $\beta_1(Z)$ indicates the number of loops or tunnels in Z; and $\beta_2(Z)$ is the number of cavities.

Our goal is to understand structure of the sublevel sets

$$\mathsf{C}(f, \theta) = \{x \in D : f(x) \leq \theta\}, \tag{5.6}$$

for all values of $\theta \in \mathbb{R}$. As we vary θ, the number of components, loops, and cavities in $\mathsf{C}(f, \theta)$ changes, implying that $\beta_k(\mathsf{C}(f, \theta))$, $k = 0, 1, 2$, also changes. (See Section 5.3 for examples.) What is remarkable is that, under very weak conditions, we can choose bases for the vector spaces $H_k(\mathsf{C}(f, \theta))$ over all values of θ such that, given a basis element of $H_k(\mathsf{C}(f, \theta))$, we can identify a unique value θ_b at which this basis element appears and a unique value θ_d at which this basis element disappears. We refer to θ_b as the *birth value*, θ_d as the *death value*, and the pair $(\theta_b, \theta_d) \in \mathbb{R}^2$ as a *persistence point* corresponding to the chosen basis element of $H_k(\mathsf{C}(f, \theta))$. The difference $\theta_d - \theta_b$ is

called the *lifespan* of the persistence point. Longer lifespans are associated with geometric features that persist through larger variations of θ, and persistence diagrams are a codification of this information. Given a scalar field f, the set of associated *persistence diagrams* are denoted by $\mathsf{PD}(f) = \{\mathsf{PD}_k(f)\}$, where $\mathsf{PD}_k(f)$ consists of all persistence points corresponding to the k-th level of homology. Let us stress that different features might be represented by persistence points with the same coordinates. Therefore, the persistence diagram is a multiset, where the multiplicity of each persistence point corresponds to the number of features represented by the persistence point. Moreover, every persistence diagram contains infinitely many points at each point along the diagonal $\theta_b = \theta_d$. The reason for the inclusion of the diagonal is made clear in Definition 5.8, when we define metrics on the space of persistence diagrams.

For the systems introduced in Section 5.1, we first use persistent homology as a nonlinear data reduction method. For Kolmogorov flow we study the scalar field $\omega\colon D \to \mathbb{R}$, the z-component of the vorticity field, while for Rayleigh-Bénard convection we study the scalar field $T\colon D \to \mathbb{R}$, the temperature field at the mid-plane. It is important to note that the domains for these two scalar fields are different. For Kolmogorov flow, the domain D is a torus since we are using periodic boundary conditions, while for Rayleigh-Bénard convection, D is a disk. For the disk, we need only to concern ourselves with the vector spaces $H_k(\mathsf{C}(T,\theta))$ for $k = 0, 1$. However, for the torus, the vector spaces $H_2(\mathsf{C}(\omega,\theta))$ also need to be considered, since the torus encloses a three-dimensional cavity. In section 5.3, we explain how the persistence diagrams $\mathsf{PD}(f)$ capture important information about the patterns given by the scalar fields ω and T.

The set of all persistence diagrams PD is a metric space, denoted by Per (see Section 5.4). Since we are studying the evolution of Kolmogorov flow and Rayleigh-Bénard convection, we have time series of the vorticity $\{\omega_i\}$ and temperature $\{T_i\}$ fields, and, therefore, we have time series of persistence diagrams $\{\mathsf{PD}(\omega_i)\}$ and $\{\mathsf{PD}(T_i)\}$. We view each of these time series as a point cloud $X \subset \mathsf{Per}$. To extract information about the dynamical structures present in the time series, we use persistent homology a second time to quantify the geometry associated with this point cloud. This is achieved by introducing a scalar function $g\colon \mathsf{Per} \to [0,\infty)$ that gives the distance from any point in

Per to the point cloud X and is defined by

$$g(x) := d(x, X) := \min_{x_i \in X} d(x, x_i), \tag{5.7}$$

where d is an appropriate metric on the space of persistence diagrams. The associated sublevel sets $\mathsf{C}(g, \theta)$ are once again given by (5.6).

To carry out the steps mentioned above requires the ability to compute the persistence diagrams $\mathsf{PD}(f)$. To do this, we need to calculate $H_k(\mathsf{C}(f, \theta))$, which requires a discrete representation of $\mathsf{C}(f, \theta)$ called a *complex*. In the context of nonlinear data reduction, we make use of a *cubical* complex. When analyzing the geometry of the point cloud, we approximate $\mathsf{C}(f, \theta)$ using a *Vietoris-Rips* complex, which is a special form of a simplicial complex (Definition 4.4). This is a classical subject and thus there are a variety of references providing precise definitions of complexes, e.g. [4] for Vietoris-Rips complexes and [18] for cubical complexes, discussions of issues related to approximations [31], and how one proceeds from a complex to computing persistent homology [46, 4]. The homological computations in this chapter were performed using the Perseus software [47].

The numerical data for the vorticity and the temperature fields is presented in the form of piecewise-constant functions defined on a rectangular lattice. For Kolmogorov flow, values of ω are reported in double precision. Recall that the vector spaces $H_k(\mathsf{C}(\omega, \theta))$ can only change for $\theta \in \Theta$, where Θ is the finite set of values that ω attains on the given lattice. Each of the sets $\mathsf{C}(\omega, \theta)$ is a cubical complex, and we use the Perseus software [47] to compute the corresponding persistence diagrams using only the values $\theta \in \Theta$. Numerical simulations for Rayleigh-Bénard convection are carried out with high precision as well. However, keeping in mind our goal to compare the numerical simulations with experimental data, we convert the temperature field to an 8-bit temperature field T^* (an integer-valued function with values between 0 and 255), which can be obtained experimentally. Consequences of this rescaling are examined in Section 5.5.

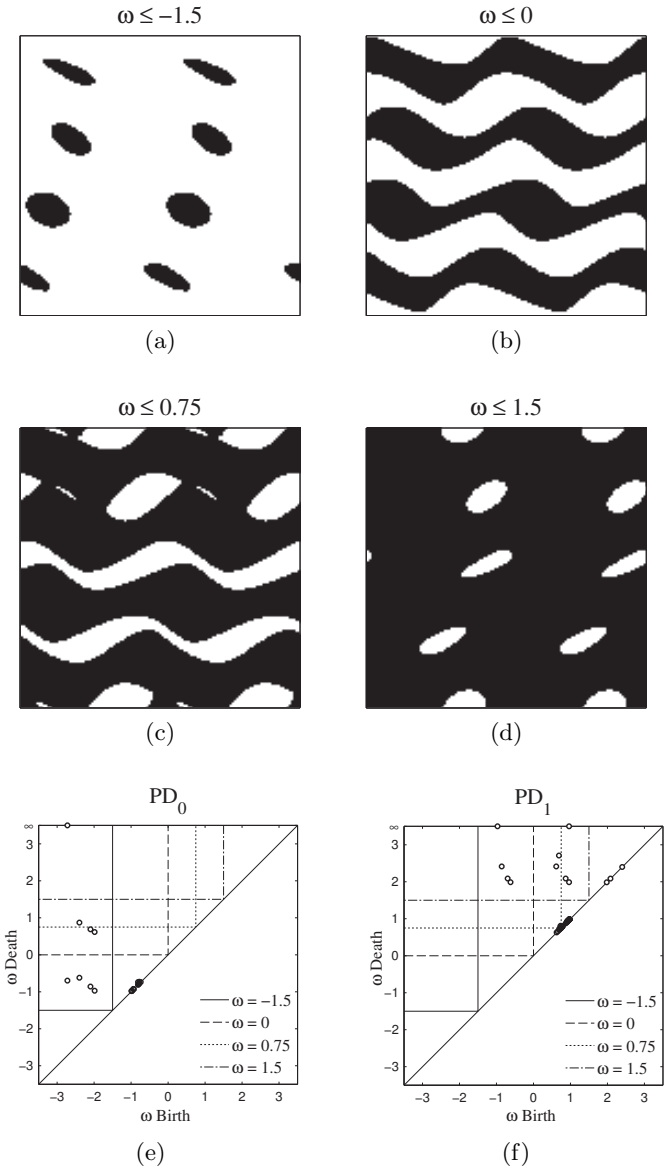


Figure 5.3: (a-d) Sublevel sets $\mathsf{C}(\omega, \theta) = \{x \in D : \omega(x) \leq \theta\}$ of the vorticity field, shown in Figure 5.2(a), for different values of θ, depicted in black. (e) $\mathsf{PD}_0(\omega)$ and (f) $\mathsf{PD}_1(\omega)$ persistence diagrams of the vorticity field indicate the values of θ at which the connected components and loops appear and disappear (merge together). Video 1 of the supplementary materials of [1] provides an animation.

5.3 Interpreting Persistence Diagrams

The purpose of this section is to provide intuition and interpretation of the information that persistence diagrams present. As indicated in the previous section, we are interested in the diagrams $\mathsf{PD}_k(\omega)$, $k = 0, 1, 2$, of the vorticity field for Kolmogorov

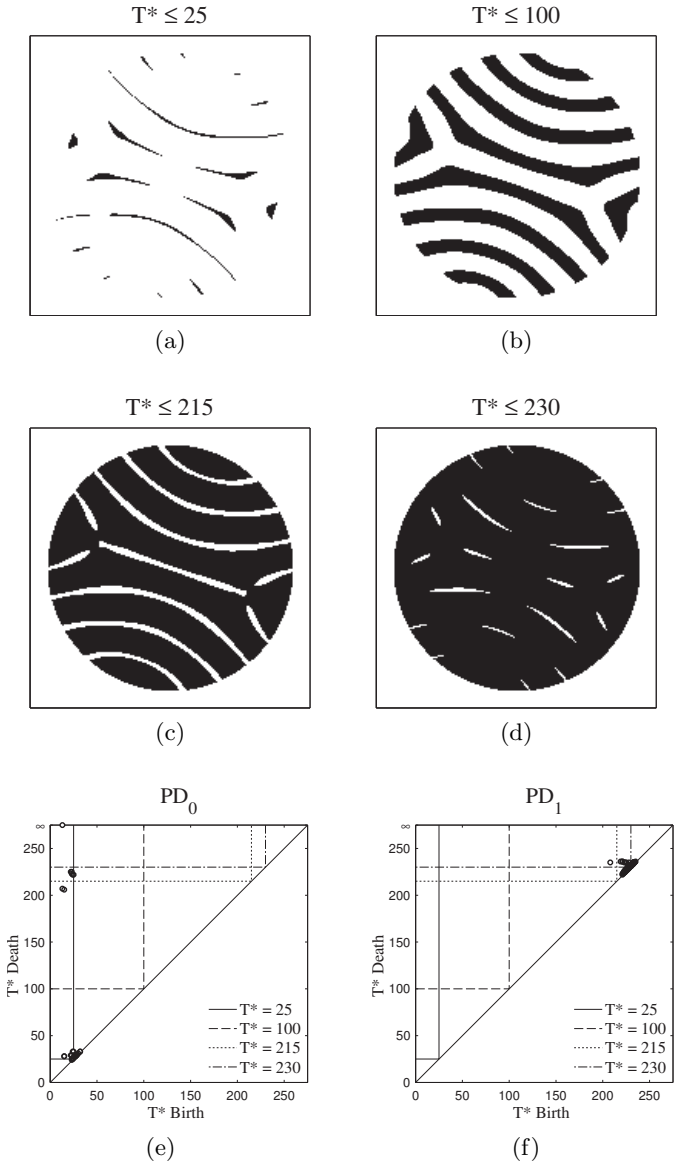


Figure 5.4: (a-d) Sublevel sets $\mathsf{C}(T^*, \theta) = \{x \in D : \mathbb{T}^*(x) \leq \theta\}$ of the the renormalized 8-bit temperature field T^*, shown in Figure 5.2(b), for different values of θ, depicted in black. As in Figure 5.3, the persistence diagrams (e) $\mathsf{PD}_0(\omega)$ and (f) $\mathsf{PD}_1(\omega)$ indicate the values of θ at which the connected components and loops appear and disappear (merge together). Video 2 of the supplementary materials of [1] provides an animation.

flow, and the diagrams $\mathsf{PD}_k(T^*)$, $k = 0, 1$, of the temperature field for Rayleigh-Bénard convection, shown in Figure 5.2.

We begin by discussing $\mathsf{PD}_0(\omega)$, shown in Figure 5.3(e), computed from a single time snapshot of the vorticity field for the Kolmogorov flow. The minimum value of the vorticity field is -2.7206, and therefore, $\mathsf{C}(\omega, \theta) = \emptyset$ for all $\theta < -2.7206$. At

$\theta = -2.7206$, two components appear, indicated by the two persistence points with birth value $\theta_b = -2.7206$. The death value of one of these two persistence points is $\theta_d = -0.697$, and so the two components merge at this value, resulting in a single component. This explains the persistence point $(-2.7206, -0.697)$. The reason the other persistence point is denoted by $(-2.7206, \infty)$, with $\theta_d = \infty$, is because when features merge, a choice must be made about which topological feature (in this case, a connected component) dies. Having a consistent choice of basis over all values of θ requires that the homology generator associated with the geometric feature that has the larger birth value die first. If the birth values are the same, then it does not matter which topological feature with this birth value is chosen to be the one that persists. In particular, this implies that the generator associated with one of these two initial components can never die.

Figure 5.3(a) indicates the subset of D corresponding to $\mathsf{C}(\omega, -1.5)$. We remind the reader that the domain D for Kolmogorov flow is a torus, since the left (top) and right (bottom) boundaries are identified. Therefore, $\mathsf{C}(\omega, -1.5)$ consists of eight distinct connected components instead of nine.

The existence of these eight connected components can also be extracted from $\mathsf{PD}_0(\omega)$, shown in Figure 5.3(e). Observe that these connected components correspond to connected regions with birth value $\theta_b \leq -1.5$ and death value $\theta_d > -1.5$. In Figure 5.3(e), this corresponds to the eight points in the rectangular region $R_{-1.5} := \big\{(\theta_b, \theta_d) \in \mathbb{R}^2 : \theta_b \leq -1.5 \text{ and } \theta_d > -1.5\big\}$.

Figure 5.3(b) indicates that $\mathsf{C}(\omega, 0)$ consists of four connected horizontal bands, which agrees with the number of persistence points in the rectangular region $R_0 = \big\{(\theta_b, \theta_d) \in \mathbb{R}^2 : \theta_b \leq 0 \text{ and } \theta_d > 0\big\}$ of $\mathsf{PD}_0(\omega)$. Each stripe is created as two distinct components present in Figure 5.3(a) grow and merge, causing one component to die each time. The deaths of these components are captured by the points in the rectangle $R_{-1.5}$ which are not in the rectangle R_0, since these are components that are born before $\theta = -1.5$ but die before $\theta = 0$.

Three horizontal stripes merge together before $\theta = 0.75$, as indicated by two points inside the rectangle $R_{-1.5}$ that are not in the rectangle $R_{0.75}$. The two remaining

connected components merge together soon thereafter, and for all greater threshold values, there is only one connected component.

To finish our analysis of $\mathsf{PD}_0(\omega)$, we turn our attention to the persistence points close to the diagonal. These have very short lifespans, which suggests that these features may be numerical artifacts. In our example, these points represent the narrow horizontal bands formed in between two connected components before they merge into a single band (see video 1 available in the supplementary materials of [1]). These narrow bands are formed by small oscillations of the vorticity field at the places where the field is almost constant.

We now turn our attention to the $\mathsf{PD}_1(\omega)$ persistence diagram, which characterizes loops in $\mathsf{C}(\omega, \theta)$. From $\mathsf{PD}_1(\omega)$, we see that the first loop appears at threshold $\theta = -0.963$. It corresponds to one of the four horizontal bands shown in Figure 5.3(b). Each horizontal band generates a single independent loop, corroborated by the existence of four persistence points in the rectangle R_0 of $\mathsf{PD}_1(\omega)$.

We note that the full torus has two loops captured by homology. This is expressed in $\mathsf{PD}_1(\omega)$ by the two persistence points with $\theta_d = \infty$. Observe that the first loop that appears at $\theta = -0.963$ is equivalent to one of the toral loops, thus it cannot be killed by any other loop, and hence is captured by the persistence point $(-0.963, \infty)$. The other three loops present at $\theta = 0$ correspond to the same toral loop and thus must die. In fact, they do so by $\theta = 2.5$. Note that the birth values θ_b of these persistence points are close to the death values θ_d of the persistence points in $R_{-1.5} \setminus R_0$ of $\mathsf{PD}_0(\omega)$. This implies that shortly after the components merge, they form horizontal bands across the entire domain.

New loops are also created as the bands start merging. If two horizontal bands are connected by n links, then the number of loops generated by this object (two bands plus the links) is $(n + 1)$. Thus, the first additional loop appears when a second link is created. In our example, this happens near the threshold 0.75.

In Figure 5.3(c), there are four distinct links between the two horizontal bands at the top of the figure. The small punctures visible in Figure 5.3(c) are filled in almost immediately, and the four links merge into two distinct links. The points in

$\mathsf{PD}_1(\omega)$ that are close to the diagonal capture this behavior. The other two links are present for a wider range of thresholds, and the loop they generate is represented by one of the persistence points in $\mathsf{PD}_1(\omega)$ with birth coordinate slightly smaller than 0.75. The horizontal band at the top and the horizontal band at the bottom are linked in a similar fashion. This explains the presence of another point with birth coordinate slightly smaller than 0.75.

At $\theta = 0.932$, a connection from the top to the bottom boundary is created. This loop is homologically equivalent to the second of the two independent loops of the torus, and hence is identified by the persistence point $(0.932, \infty)$. As the threshold passes the value 1.988, the punctures shown in Figure 5.3(d) start disappearing and the corresponding loops start dying. Again, there are $1 + n$ independent loops for $n > 0$ punctures. Since the maximum value of ω is 2.7092, the sub-level set is the whole torus for any threshold above this, i.e. $\mathsf{C}(\omega, \theta) = D$ for all $\theta \geq 2.7092$. In this case, there are no more punctures, and the rectangle $R_{2.7092}$ contains only two persistence points.

Finally, we address the $\mathsf{PD}_2(\omega)$ persistence diagram, not shown for brevity. This diagram contains a single persistence point at $(\theta_b, \infty) = (3.01, \infty)$. The birth coordinate, θ_b, corresponds to the minimum value of θ for which $\mathsf{C}(\omega, \theta) = D$, the whole torus. Since $\mathsf{C}(\omega, \theta) = D$ for all $\theta \geq \theta_b$, this point never dies.

We now discuss the persistence diagrams for the temperature field T^* shown in Figure 5.4 for Rayleigh-Bénard convection. Again, beginning with $\mathsf{PD}_0(T^*)$, the points with short lifespans correspond the large number of small connected components that make up $\mathsf{C}(T^*, 25)$, as shown in Figure 5.4(a). Points with long lifespans represent the well-defined connected components shown in Figure 5.4(b). From the persistence diagram, we can see that these components merge almost simultaneously at two threshold values, $\theta \approx 210$ and $\theta \approx 225$.

Turning to $\mathsf{PD}_1(T^*)$, we note that the domain of the temperature field is a disk, so the independent loops correspond to punctures inside of the disk. The diagram $\mathsf{PD}_1(T^*)$ indicates that there are no loops with long lifespans, and the loops that do appear do so roughly at the same threshold values at which the dominant components merge. These features are due to the small fluctuations of the temperature field close

to the critical values at which different rolls merge together. This is consistent with their short lifespans.

5.4 The Space of Persistence Diagrams

As explained in the previous section, a persistence diagram codifies, in a reasonably compact form, considerable information about the geometry of a scalar function. As suggested by the examples, we use persistence diagrams to provide a reduced description of the state of the dynamical system of interest at any given point in time. Therefore, to analyze the dynamics, we need to be able to compare one collection of persistence diagrams PD (corresponding to a snapshot of the flow pattern at an instant of time) to another collection of diagrams PD' (from another flow snapshot). There are a variety of metrics that can be imposed on persistence diagrams [6, 16, 48]. The metrics used herein rely on pairing the points $p \in \mathsf{PD}_k$ in a one-to-one correspondence (bijection) with the points in PD'_k. According to the definition used here, every persistence diagram contains an infinite number of copies of the diagonal. Hence, there are many different bijections γ between PD_k and PD'_k. Roughly speaking, the distance between PD and PD' is defined using the bijections that minimize the maximum distance observed in the mapping of the points p from PD_k to $\gamma(p)$ in PD'_k. This notion is made more precise in the following definition.

Definition 5.8. *Let* $\mathsf{PD} = \{\mathsf{PD}_k\}$ *and* $\mathsf{PD}' = \{\mathsf{PD}'_k\}$ *be two collections of persistence diagrams. The* bottleneck distance *between* PD *and* PD' *is defined to be*

$$d_B(\mathsf{PD}, \mathsf{PD}') = \max_k \inf_{\gamma\colon \mathsf{PD}_k \to \mathsf{PD}'_k} \sup_{p \in \mathsf{PD}_k} \|p - \gamma(p)\|_\infty, \tag{5.9}$$

where $\|(a_0, b_0) - (a_1, b_1)\|_\infty := \max\{|a_0 - a_1|, |b_0 - b_1|\}$ *and* γ *ranges over all bijections between persistence points. Similarly, the* degree-p Wasserstein distance *is defined as*

$$d_{W^p}(\mathsf{PD}, \mathsf{PD}') = \left[\sum_k \inf_{\gamma\colon \mathsf{PD}_k \to \mathsf{PD}'_k} \sum_{p \in \mathsf{PD}_k} \|p - \gamma(p)\|_\infty^p \right]^{1/p}. \tag{5.10}$$

Roughly speaking, a function $f\colon D \to \mathbb{R}$ is *tame* if, for every $\theta \in \mathbb{R}$, the vector space

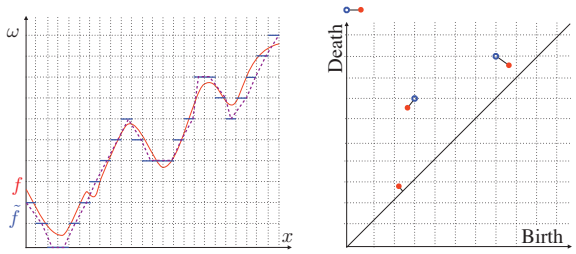
$H_k(f^{-1}((-\infty, \theta]))$ is finite dimensional for every k, and there are only finitely-many thresholds at which the vector spaces change (for a precise definition see [4]). For our purposes, it suffices to remark that if f is a piecewise-constant function on a finite complex, then f is tame. In particular, the numerically-computed vorticity field ω and 8-bit temperature field T^* are tame functions.

For the remainder of this chapter, we use Per_k to denote the set of persistence diagrams corresponding to H_k and Per to denote the set of all persistence diagrams. Let $T(D, \mathbb{R})$ denote the set of tame functions $f\colon D \to \mathbb{R}$ equipped with the L^∞ norm. A fundamental result [4] is that, using the Wasserstein or bottleneck metrics, $\mathsf{PD}\colon T(D, \mathbb{R}) \to \mathsf{Per}$ is a Lipschitz-continuous function. In particular, if $f, g \in T(D, \mathbb{R})$, then

$$d_B(\mathsf{PD}(f), \mathsf{PD}(g)) \leq \sup_{x \in D} |f(x) - g(x)|. \tag{5.11}$$

These results on Lipschitz continuity have two important implications for this work, both stemming from the fact that our analysis is based on numerical simulations. Assume for the moment that $f\colon D \to \mathbb{R}$ denotes the exact solution at a given time to either Kolmogorov flow or the Boussinesq equations. Ideally, we want to understand $\mathsf{PD}(f)$. Our computations of persistent homology are based on $\mathsf{C}(\tilde{f}, \theta_i)$, a cubical complex defined in terms of the numerically-reported values θ_i, where $\tilde{f}$ represents the associated piecewise-constant function. If the numerical approximation $\tilde{f}$ satisfies $\sup_{x \in D} |f(x) - \tilde{f}(x)| \leq \varepsilon$, then by (5.11) we have a bound on the bottleneck distance between the actual persistence diagram $\mathsf{PD}(f)$ and the computed persistence diagram $\mathsf{PD}(\tilde{f})$, so that $d_B(\mathsf{PD}(f), \mathsf{PD}(\tilde{f})) \leq \varepsilon$. Figure 5.5 provides a schematic justification of this claim.

As indicated in the introduction, persistent homology is invariant under certain continuous deformations of the domain. For example, if $g\colon D \to D$ is a homeomorphism and $f\colon D \to \mathbb{R}$, then $\mathsf{PD}(f \circ g) = \mathsf{PD}(f)$. Of particular relevance to this work is a function g which arises as a symmetric action on the domain. In this chapter, we work with piecewise-constant numerical approximations of the actual functions of interest, and we cannot assume that this equality holds. However, if f is given and $f' = f \circ g$, where g



(a) (b)

Figure 5.5: (a) A one-dimensional scalar field $f\colon D \to \mathbb{R}$ and its piecewise constant approximation $\tilde{f}$. The dashed line shows a tame approximation of f whose persistence diagram is identical to $\mathsf{PD}(\tilde{f})$. (b) Points in the persistence diagram $\mathsf{PD}(f)$ are given by closed symbols and the points in $\mathsf{PD}(\tilde{f})$ are represented by open symbols. Two points on the top have infinite death coordinate. Lines connecting the points represent matching of the persistence points for which the bottleneck distance $d_B(\mathsf{PD}(f), \mathsf{PD}(\tilde{f}))$ is realized. It follows from (5.11) that $d_B(\mathsf{PD}(f), \mathsf{PD}(\tilde{f})) \leq \sup_{x\in D} |f(x) - \tilde{f}(x)|$.

is as above, and we have an L^∞ bound ε on the difference between the approximation and the true function, then by (5.11),

$$d_B(\mathsf{PD}(f), \mathsf{PD}(f')) \leq \varepsilon. \tag{5.12}$$

In summary, under the assumption of bounded noise or errors from numerical simulations (or experimental data), we have explicit control of the errors of the distances in Per.

5.5 Using Metrics in the Space of Persistence Diagrams

The goal of this section is twofold: one, to provide intuition about the information contained in the different metrics, and two, to suggest how viewing a time series in Per can provide insight into the underlying dynamics.

We begin by remarking that the bottleneck distance d_B measures only the single largest difference between the persistence diagrams and ignores the rest. The Wasserstein distance d_{W^p} includes all differences between the diagrams. Thus, it is always true that

$$d_B \leq d_{W^p}. \tag{5.13}$$

	d_B	d_{W^2}	d_{W^1}
$(\mathsf{PD}^a, \mathsf{PD}^b)$	0.01	0.049	0.497
$(\mathsf{PD}^a, \mathsf{PD}^c)$	0.864	2.648	12.35

Table 5.1: Distances between selected persistence diagrams (rounded to 3 decimal places) shown in Figure 5.6.

The sensitivity of the Wasserstein metric to small differences (possibly due to noise) can be modulated by the choice of the value of p, i.e. if $p > q$, then one expects d_{W^p} to be less sensitive to small changes than d_{W^q}. In the work that follows, we restrict ourselves to the bottleneck distance d_B and the Wasserstein distances d_{W^p} for $p = 1, 2$.

The most obvious use of these metrics is to identify or distinguish patterns. As an example, we consider patterns along an orbit from the Kolmogorov flow. As indicated in Section 5.1.1, this particular trajectory arises from a periodic orbit with a slow drift along an orbit of continuous symmetry. In particular, we consider the three time points indicated in Figure 5.1(a): two that appear to differ by the continuous symmetry, and a third that lies on the 'opposite' side of the periodic orbit. Plots of the associated vorticity fields at these points (see Figure 5.6(a)-(c)) agree with this characterization of the time points. We want to identify this information through the associated persistence diagrams PD^a, PD^b, and PD^c, shown in Figure 5.6(d)-(e). Indeed, the plots of PD^a_k and PD^b_k are difficult to distinguish, but PD^c_k is clearly distinct. To quantify this difference, the values of the distances between these persistence diagrams under the metrics $d_\star$ for $\star \in \left\{B, W^2, W^1\right\}$ are recorded in Table 5.1. Not surprisingly, the distances between PD^a and PD^b are much smaller than the distances between PD^a and PD^c. We consider the distances between PD^a and PD^b to give a rough estimate of the noise level of the system with respect to that metric.

We now turn to the question of understanding dynamics from the time series in Per. Let f_i denote the scalar field of the system at time t_i. If $\Delta t = t_{i+1} - t_i$ is small and the evolution of the system is continuous, then because $d_\star$ (for $\star \in \left\{B, W^2, W^1\right\}$) is a metric,

$$s_\star(t_i) = \frac{d_\star(\mathsf{PD}(f_i), \mathsf{PD}(f_{i+1}))}{\Delta t} \tag{5.14}$$

can be interpreted as an average speed in the space of persistence diagrams over the

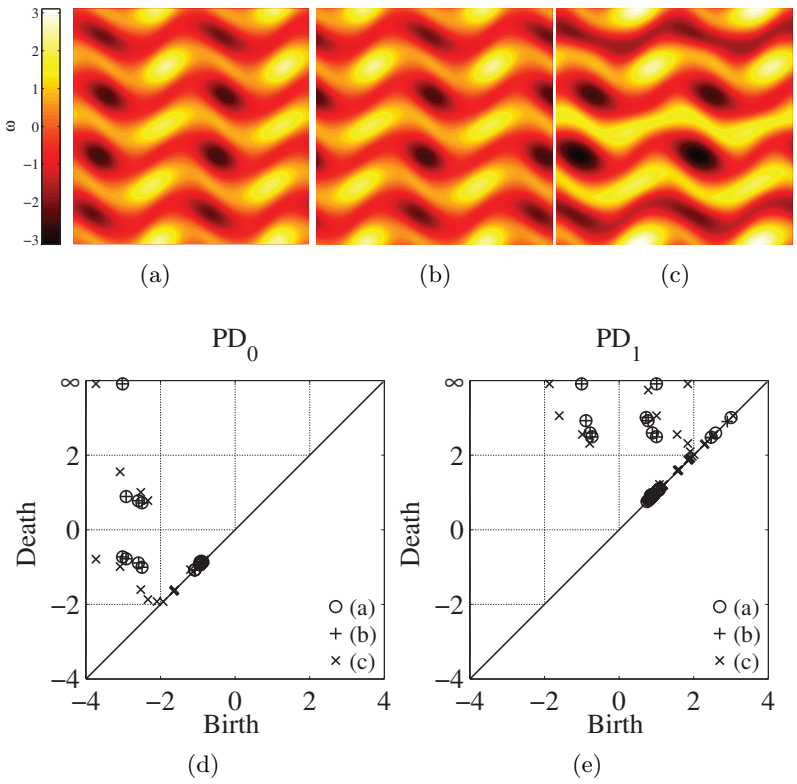
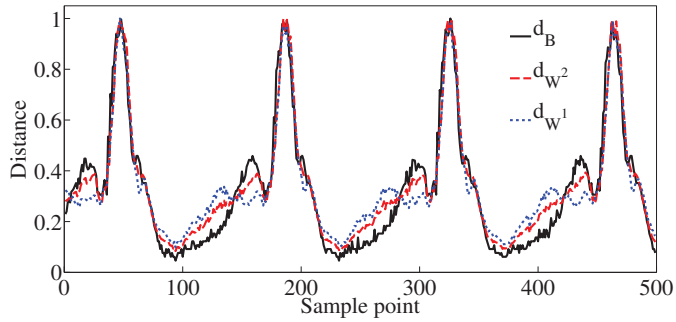


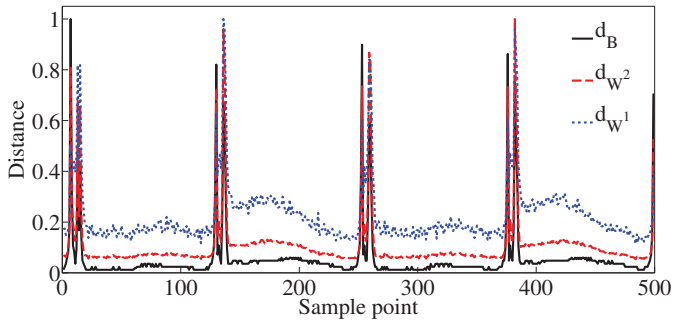
Figure 5.6: Three snapshots of the vorticity fields ω from the stable relative periodic orbit of the Kolmogorov flow, found at $Re = 25.43$. The vorticity fields correspond to the (a) diamond, (b) square, and (c) circle in Figure 5.1(a). The persistence diagrams for these three snapshots are compared in (d) and (e). The points in PD^a and PD^b are almost identical because the corresponding vorticity fields are similar. The points in PD^c are more spread out and do not shadow the points in PD^a so well. The same is true for the PD_1 persistence diagrams which are not shown. Thus, $d_\star(\mathsf{PD}^a, \mathsf{PD}^b) < d_\star(\mathsf{PD}^a, \mathsf{PD}^c)$ for $\star \in \left\{B, W^2, W^1\right\}$, as indicated by Table 5.1.

time interval $[t_i, t_{i+1}]$. The value of $s_\star$ depends on the choice of metric. For example, s_{d_B} is the rate at which the largest change between the geometric features of the scalar fields occurs. The speeds measured by d_{W^p}, $p = 1, 2$, keep track of the rate of change between all geometric features, though to some extent, d_{W^2} suppresses the effect of noise.

Figure 5.7(a) shows distances d_* between consecutive sample points, normalized by the maximum distance, of samples taken along approximately three periods of the stable relative periodic orbit of the Kolmogorov flow described in Section 5.1.1. Normalizing s_* by the maximum speed along the orbit furnishes the same curves. Each of the graphs of $s_\star$ indicate that speed is not uniform along the orbit; there are parts of the



(a)

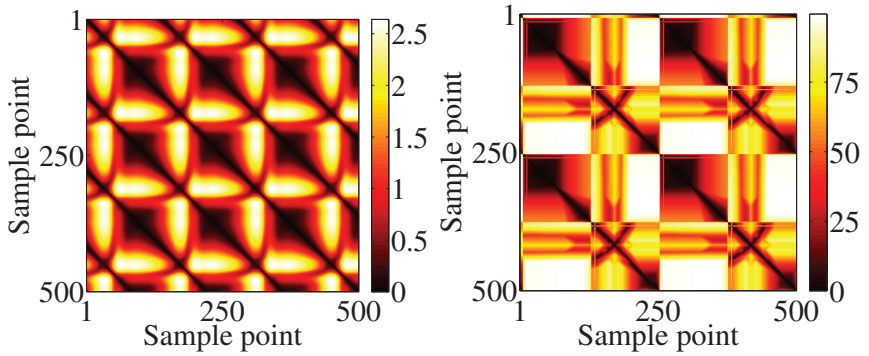


(b)

Figure 5.7: (a) Distances $d_\star$ between the consecutive sample points along the stable relative periodic orbit of the Kolmogorov flow are shown for just over three periods. Distances are normalized by their maximum value which is 0.0654 for d_B, 0.2266 for d_{W^2}, and 1.9143 for d_{W^1}. Distance between the consecutive sample points can be interpreted as the speed at which the orbit is traversed in the space of persistence diagrams. Speed is not uniform along the orbit. Instead, there are parts of the orbit where the dynamics are slow, separated by relatively fast evolution. (b) Distances $d_\star$ between the consecutive sample points along an almost-periodic orbit from Rayleigh-Bénard convection are shown for approximately 2 periods. Distances are again normalized by their maximum value which is 83.5 for d_B, 113.66 for d_{W^2}, and 383 for d_{W^1}.

orbit where the geometry is changing slowly, separated by intervals of relatively fast evolution. The evolution is extremely slow around the states $100, 240$, and 380. The values of the speed (before normalizing) are below the noise (fluctuation) levels given by the first row of Table 5.1. This suggests that the orbit may be passing close to a fixed point.

While the general shapes of the speed profiles for different distances are similar, there are places where the signs of their derivatives differ. As the system starts accelerating around $t = 100$, all three speeds are increasing. Around $t = 130$, the speed s_{W^1} starts decreasing while the other two speeds are still increasing. Note that around $t = 130$,



(a) (b)

Figure 5.8: (a) Distance matrix D, generated by the d_{W^2} metric, for approximately three periods of the stable relative periodic orbit of the Kolmogorov flow. The large black patches correspond to the parts of the orbit with slow dynamics. Equally spaced black lines parallel to the diagonal suggest periodicity of the orbit with period equal to the distance between these lines. (b) Distance matrix D, generated by the d_B metric, for 2 periods of the almost-periodic orbit of Rayleigh-Bénard convection. The checkerboard pattern indicates that sampling is too sparse, and fast dynamics are not resolved with the level of sampling.

the speeds rise above the noise level (fluctuations). The fact that s_B and s_{W^2} are both increasing means that the changes between the prominent geometric features are growing in this region. The speed s_{W^1} is decreasing in this region and so the noise (error) fluctuations are decreasing. At $t = 170$, the dominant geometric features start to evolve considerably. Changes of the dominant features are the most important contributions to all three metrics. Therefore, the derivatives of the speeds $s_\star$ have the same sign again (see video 3, 4, or 5 in the supplementary materials of [1]).

Figure 5.7(b) shows the normalized speed profiles for the Rayleigh-Bénard convection simulations. As in the case of the Kolmogorov flow, all three metrics indicate that there are two distinct speed scales along the orbit. However, the speed profile for d_B differs significantly from those of d_{W^1} and d_{W^2}. In particular, it suggests that for significant time periods, the major geometric features of the temperature field vary only slightly, followed by two rapid bursts of change. This can be verified by viewing video 6, 7, or 8 of the supplementary materials of [1]. Away from the rapid bursts, s_B is close to 1. The temperature field has integer values, so the changes cannot be smaller than 0.5. This implies that both s_{W^1} and s_{W^2} are dominated by the small fluctuations

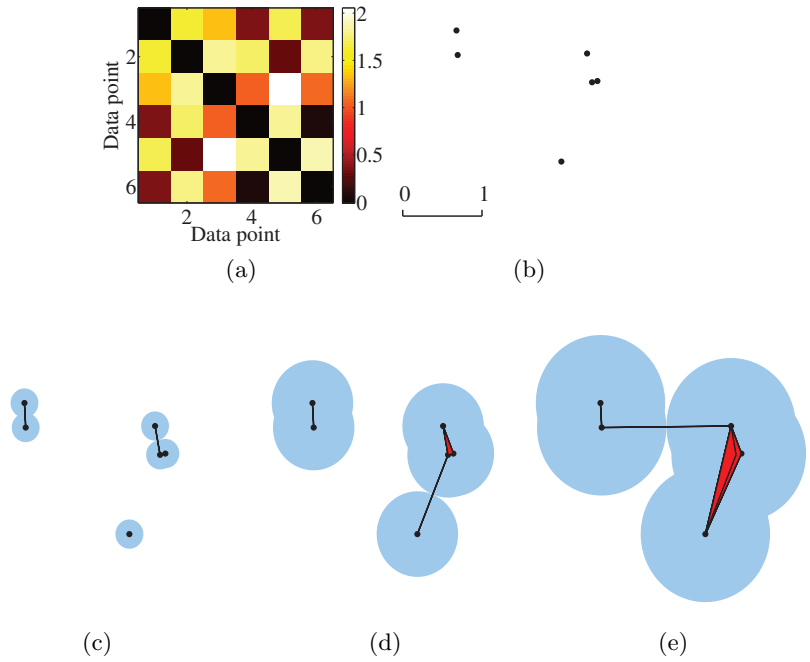


Figure 5.9: (a) Distance matrix representing pairwise Euclidian distances d_E between the points in (b) a point cloud X. (c-e) The blue shaded regions indicate the sub-level sets $\mathsf{C}(X, \theta)$ for $\theta = 0, 0.1755, 0.5135$, and 0.816. The points, edges and triangles indicate the Vietoris-Rips complexes $\mathsf{R}(X, \theta)$. (c) For $\theta = 0.1775$ the set $\mathsf{C}(f, \theta)$ consist of three distinct connected clusters. The same is true for $\mathsf{R}(X, \theta)$. The points in each connected component of $\mathsf{C}(f, \theta)$ are connected by edges in $\mathsf{R}(X, \theta)$. (c) The three components remain distinct until $\theta = 0.5135$, at which point two components of $\mathsf{C}(f, \theta)$ merge and an edge connecting the points in the merged components appears in $\mathsf{R}(X, \theta)$.

which are roughly of order 1. Hence, the relative speeds s_{W^1} and s_{W^2} have essentially the same shape.

The plots of $s_\star$ hint at the underlying dynamics being that of a periodic orbit. However, it is important to keep in mind that Per is an infinite-dimensional space, and that periodicity in the speed of a trajectory does not imply that the trajectory lies on a closed curve. As an example, Figure 5.7(b) suggests that there are just over four periods of Rayleigh-Bénard convection shown, and that a single period is roughly 125 frames long. However, looking at video 6, 7, or 8 in the supplementary materials of [1], it is clear that a full period is closer to 250 frames. Similarly, it is not obvious that extended periods of high speed imply that the pattern changes significantly over that time period (a periodic orbit of small diameter can exhibit high speed). This requires

a more global geometric analysis of the time series, which we discuss shortly.

With the same data set used to generate Figure 5.7(a) and letting ω_j denote the vorticity field at time t_j, Figure 5.8(a) exhibits the d_{W^2} distance matrix D for Kolmogorov flow, with color-coded entries $D(i,j) := d_{W^2}(\mathsf{PD}(\omega_i), \mathsf{PD}(\omega_j))$. (The d_B and d_{W^1} distance matrices look very similar and are not shown.) Observe that $D(i,i) = 0$ and D is symmetric since $D(i,j) = D(j,i)$. Furthermore, Figure 5.7(a) is a plot of the immediate off-diagonal entries. A striking feature of the distance matrix in Figure 5.8(a) is the existence of dark lines parallel to the diagonal, spaced at intervals of roughly 110 samples. This indicates that, in the space of persistence diagrams, the trajectory periodically repeats the same, or nearly the same, state. Since the diagonals are spaced at roughly 110 samples, we can indeed say that the orbit revisits very similar states at intervals of roughly 110 samples. Similarly, the light regions close to the diagonal in Figure 5.8(a) correspond to the times in Figure 5.7(a) at which the speed is large, indicating significant changes in the pattern at these times.

To obtain a more global analysis we turn to Figure 5.8(b) that shows the distance matrix $D(i,j) := d_{W^2}(\mathsf{PD}(T_i^*), \mathsf{PD}(T_j^*))$ for the temperature fields T_i^* corresponding to the trajectory from Rayleigh-Bénard convection. Distances between the consecutive temperature fields are shown in Figure 5.7(b). The dark diagonal lines are spaced at intervals of roughly 250 samples. Thus, even though the Figure 5.7(b) suggests a period of approximately 125, the orbit does not revisit the same state in the space of persistence diagrams every 125 samples, but instead every 250 samples.

5.6 Analyzing a Point Cloud using Persistent Homology

The discussion in the previous section suggests that interesting information concerning the dynamics of the geometry of time-evolving scalar fields can be obtained by studying the time series in the space of persistence diagrams. Note that each scalar field is represented by a persistence diagram $\mathsf{PD}(f)$ and thus corresponds to a point in Per. We argue that viewing the time series as a point cloud in the space of persistence diagrams and studying its geometry provides useful information about the dynamics.

For a point cloud $X \subset \mathsf{Per}$ and the scalar function $f\colon X \to [0,\infty)$ given by (5.7) (for any of the metrics d_B, d_{W^2} or d_{W^1}), the sub-level set $\mathsf{C}(f,\theta)$ defined by (5.6) is a union of balls

$$\mathsf{C}(f,\theta) = \bigcup_{\mathsf{PD}\in X} B(\mathsf{PD},\theta), \tag{5.15}$$

where $B(\mathsf{PD},\theta) = \left\{\mathsf{PD}' \in \mathsf{Per} : d(\mathsf{PD}',\mathsf{PD}) \leq \theta\right\}$, and d is the appropriate metric. In general, one should expect that the sets $\mathsf{C}(f,\theta)$ are complicated. Therefore, computing $H_*(\mathsf{C}(f,\theta))$ directly is not practical. Instead, we make use of the Vietoris-Rips complex, given by Definition 4.4.

Observe that the Vietoris-Rips complex is determined by the distance matrix associated with X, and hence, there is a finite set of threshold values $\Theta = \{\theta_i\}$ at which the complex changes. Thus, given a point cloud X in a metric space with metric d, the associated persistence diagrams $\mathsf{PD}(X,d)$ are determined by the Vietoris-Rips complexes $\mathsf{R}(X,\theta)$ for $\theta \in \Theta$.

We emphasize that the only data used to analyze a point cloud based on the persistent homology of Vietoris-Rips complexes are the pairwise distances between the points given by the distance matrix associated with X.

5.6.1 Detecting Clusters

Since β_0 counts components, it is reasonable to use persistent homology as a clustering tool. We demonstrate this idea on a point cloud with pairwise distances given by the distance matrix shown in Figure 5.9(a). A possible configuration of the six points in $\mathbb{R}^2$ is depicted in Figure 5.9(b). Using the length scale presented in Figure 5.9(b) as an indicator of the order of magnitude at which we want to declare a separation length for the clusters, there are three clusters. We now focus on the geometric information conveyed by $\mathsf{PD}_0(X,d_E)$, shown in Figure 5.10.

Observe that $\mathsf{C}(f,0) = \mathsf{R}(X,0)$ consists of 6 vertices. As θ increases, the distinct connected components of $\mathsf{C}(f,\theta)$ (as defined in (5.15)) start merging together. In fact, when the balls $B(x_i,\theta)$ and $B(x_j,\theta)$ merge together, an edge $\langle x_i, x_j\rangle$ appears in $\mathsf{R}(X,\theta)$. Therefore, $H_0(\mathsf{C}(f,\theta)) = H_0(\mathsf{R}(X,\theta))$ for all $\theta \in \mathbb{R}$, and $\mathsf{PD}_0(f) = \mathsf{PD}_0(X,d_E)$. Note

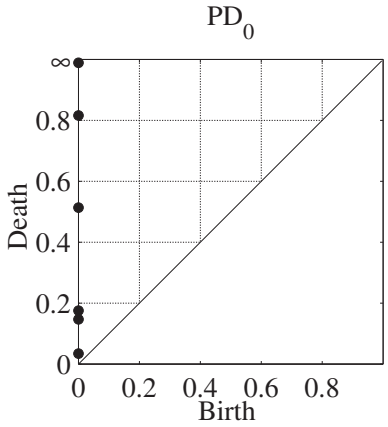


Figure 5.10: Persistence diagram $\mathsf{PD}_0(X, d_E)$ corresponding to the distance matrix in Figure 5.9(a).

that it is impossible for a new connected component to appear for $\theta > 0$. Hence, all persistence points in $\mathsf{PD}_0(X, d_E)$ have a birth value equal to zero. The death coordinates represent the spatial scales at which distinct connected components (clusters) merge together. Say that we are interested in clusters where the minimal separation is on the order of 1 length scale. These clusters correspond to the points in $\mathsf{PD}_0(X, d_E)$ with the death coordinate greater than approximately 0.5, and there are three persistence points that satisfy this criterion. Thus, we declare that there are three clusters. If the relevant scale for separation is of an order of magnitude smaller, then there are five clusters, since, in addition to the three points with death value greater than 0.5, two points have death values slightly larger than 0.05.

Alternatively, if we are interested in dividing the data into two clusters, then $\mathsf{PD}_0(X, d_E)$ can be used to determine the magnitude of the separation between the clusters. Observe that the persistence point $(0, \infty)$ corresponds to the final connected component. The persistence point $(0, 0.816)$, with the largest finite death coordinate, indicates that the components merged at a distance 0.816. Hence, the minimal distance between points from the point cloud X that belong to two distinct clusters is 1.632.

5.6.2 Detecting Loops

Since β_1 counts loops, it is reasonable to use persistent homology as a tool for identifying cycles that arise from dynamics. Consider any point cloud that generates a distance

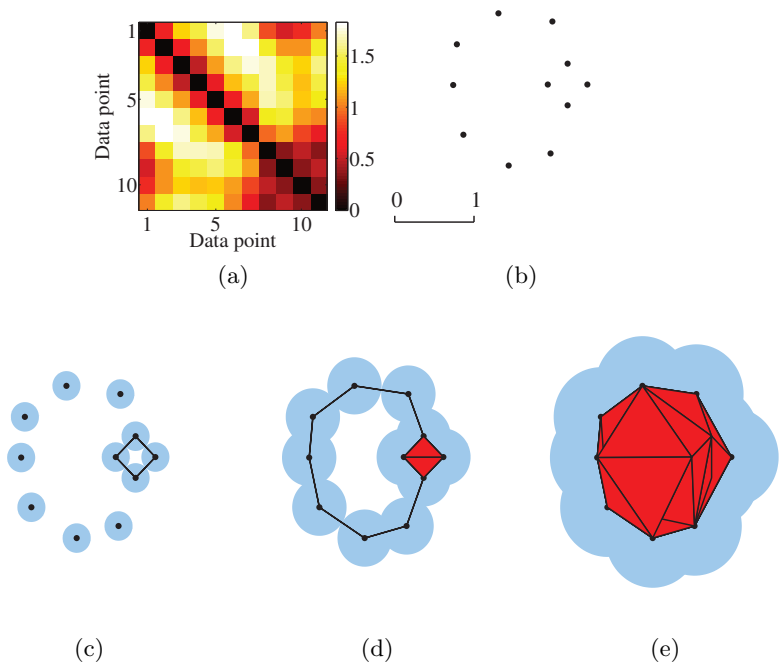


Figure 5.11: (a) Distance matrix representing pairwise Euclidian distances d_E between the points in (b) a point cloud X. (c-e) Sets $\mathsf{C}(f, \theta)$ for $\theta = 0, 0.177, 0.343$ and 0.596. The homology of $\mathsf{C}(f, \theta)$ can be approximated by a Vietoris-Rips complex $\mathsf{R}(X, \theta)$ given by the vertices, edges, and triangles shown in (b-e). The first loop in $\mathsf{C}(f, \theta)$ is created at $\theta = 0.177$. This loop is due to the noisy sampling and is filled in almost immediately. The dominant loop shown in (c) is formed at $\theta = 0.343$ and persists until $\theta = 0.596$.

matrix as in Figure 5.11(a). Again, for the sake of intuition, we present in Figure 5.11(b) an example of a point cloud $X \subset \mathbb{R}^2$ with pairwise distances given by the distance matrix shown. The persistence diagrams for the associated Vietoris-Rips complex filtrations are shown in Figure 5.12.

Applying the reasoning from the previous section, we can ask whether there is a natural or interesting clustering of the data. If, as before, we insist that we are interested in clusters where the minimal separation is on the order of length scale 1, shown in Figure 5.11(b), then $(0, \infty)$ is the only persistence point with death value greater than 0.5, i.e. at this scale there is only one component. Thus, we conclude that from a geometric perspective we may treat the point cloud as arising from a single dynamical structure.

We now look for cyclic structures. Observe that $\mathsf{PD}_1(X, d_E)$ contains two persistence

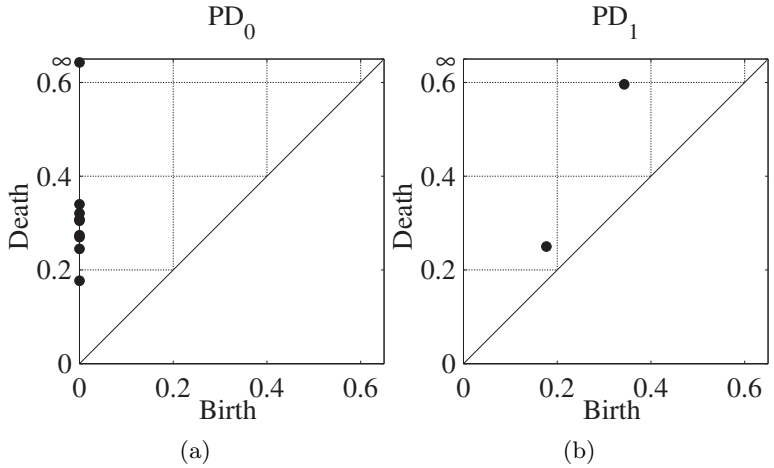


Figure 5.12: Persistence diagrams (a) $\mathsf{PD}_0(X, d_E)$ and (b) $\mathsf{PD}_1(X, d_E)$ corresponding to the distance matrix in Figure 5.11(a). The persistence diagram $\mathsf{PD}_1(X, d_E)$ contains a dominant point $(0.343, 0.596)$ corresponding to the robust loop shown in Figure 5.11(d) while the point $(0.177, 0.25)$ represents the small loop visible in Figure 5.11(c).

points. The lifespan of point $(0.177, 0.250)$ is 0.06, which is short compared to the order 1 length scale, and thus it is reasonable to think of this as a result of noise in the data. This is substantiated by Figure 5.11(c), in which the loop in the Vietoris-Rips complex $\mathsf{R}(X, 0.177)$ consists of four edges. An additional edge and two triangles (two 2-simplicies) appear in $\mathsf{R}(X, 0.250)$, see Figure 5.11(d). The triangles fill in the loop formed by the edges of $\mathsf{R}(X, 0.177)$. Two of the four data points that are involved in the construction of this loop can be viewed as arising from noise or errors associated with sampling points from a smooth cycle.

The lifespan of persistence point $(0.343, 0.596)$ is 0.253 and suggests that the point cloud is generated by a loop with a minimal radius of 0.596, which is on the order of the scale of the data. This suggests that the associated cycle, indicated in Figure 5.11(d), represents an observable, robust dynamical feature.

5.6.3 Application to Systems with Multiple Time Scales and Large Data Sets

Characterizing the geometry of a continuous orbit via an approximation by a discrete time series depends on the frequency of sampling, and thus becomes a challenge in

the setting of dynamics with multiple time scales, i.e. when the rate of change of the patterns is far from constant. If the sampling rate is too slow, then parts of the orbit will be poorly (or not at all) sampled. Note that the geometry of the continuous trajectory may be more complicated than that of a circle; secondary structures might occur if the orbit is twisted, pinched, or bent in Per. Thus, the missing parts of the orbit could distort (or entirely miss) significant features in the geometry of the sampled trajectory as compared to the geometry of the underlying (continuous) dynamics. Thus, in order to obtain a description of the geometry on all relevant spatial scales, including information about secondary structures, the sampling rate needs to be fast enough.

To determine if a trajectory has been sampled densely enough to resolve the geometry of the underlying dynamics, it is useful to compare the following three values related to the point cloud in Per: the noise threshold of the system, the maximum consecutive distance in the sampled trajectory, and the diameter of the point cloud. Ideally, once a noise threshold has been computed, one would like distances between consecutive points from the sampled trajectory to be on the length scale of the noise. If sampling faster than this, the features detected from the sample that are on the scale of the noise would be indistinguishable from artifacts generated from the noise in the sample. Thus, ideally, the speed profiles (e.g. Figure 5.7 for Kolmogorov flow and Rayleigh-Bènard convection) should have maximums no larger than the noise. Unfortunately, this is not practical for reasons that will be explained next, and fortunately it is often not necessary. For example, the length scale of the computational noise could be much smaller than the relevant length scale of interest for studying the geometry of the dynamics. In this case, a comparison of the maximum consecutive distance in the sample to the diameter of the point cloud in Per is often useful. For instance, if a point cloud has diameter 100 and the smallest relevant length scale for the geometry to be studied is 10, then a maximum consecutive distance of 10 is sufficient for the sampling of the time series, even if the noise threshold is on length scale 1. Thus, it is the interplay of these three numbers that determine if one has sampled a continuous time series densely enough.

Evaluating these three quantities from an initial time sample may indicate that an

increase in the sample rate is required to resolve the dynamics at the relevant spatial scale. In the context of a large-scale computation such as that required for the 3D simulation of Rayleigh-Bénard convection, it is easier to save the data at a higher sampling frequency than to develop numerical methods that save data based on an adaptive time step. In Section 5.9, we demonstrate the approach introduced here using approximately 7×10^5 equally-spaced snapshots of the temperature field of Rayleigh-Bénard convection. It should be immediately apparent that the set X is too large to compute the associated persistence diagrams $\mathsf{PD}(X, d_\star)$, for $\star \in B, W^2, W^1$, directly. The first step would require computing the distance matrix for X, which would involve 49×10^{10} distance computations. Note, however, that using a fast sampling rate leads to collecting unnecessarily many samples at the places where the dynamics are slow. This suggests that an appropriate choice of down-sampling will allow us to capture the global geometry of the point cloud.

Definition 5.16. *Let X be a point cloud in a metric space (M, d). Fix $\delta > 0$. A set $Y \subseteq X$ is a* δ-dense *subsample of X if for every $x \in X$, there exits a $y \in Y$ such that $d(x, y) < \delta$.*

Remark 5.17. By Corollary 4.16, there is a bijection $\gamma : \mathsf{PD}(Y, d) \to \mathsf{PD}(X, d)$ with the following property: if $\gamma(\theta_b, \theta_d) = (\theta'_b, \theta'_d)$ and $\theta'_b \neq \theta'_d$, then $0 \leq \theta'_b - \theta_b < \delta$ and $0 \leq \theta'_d - \theta_d < \delta$.

To optimize the computational cost, we wish to choose a subsample of the point cloud Y as small as possible. A point cloud Y is *δ-sparse* if, for every pair of distinct points $y_1, y_2 \in Y$, the distance $d(y_1, y_2) \geq \delta$. Given a point cloud X and a value $\delta \geq 0$, a δ-dense and δ-sparse subsample Y may always be constructed [9]. Due to the size of the point cloud X and the complexity of computing $d_\star$ for $\star \in B, W^2, W^1$, we use an alternate algorithm [49], which takes advantages of parallel computing structures and metric trees to construct a subsampling which is both δ-dense and δ-sparse .

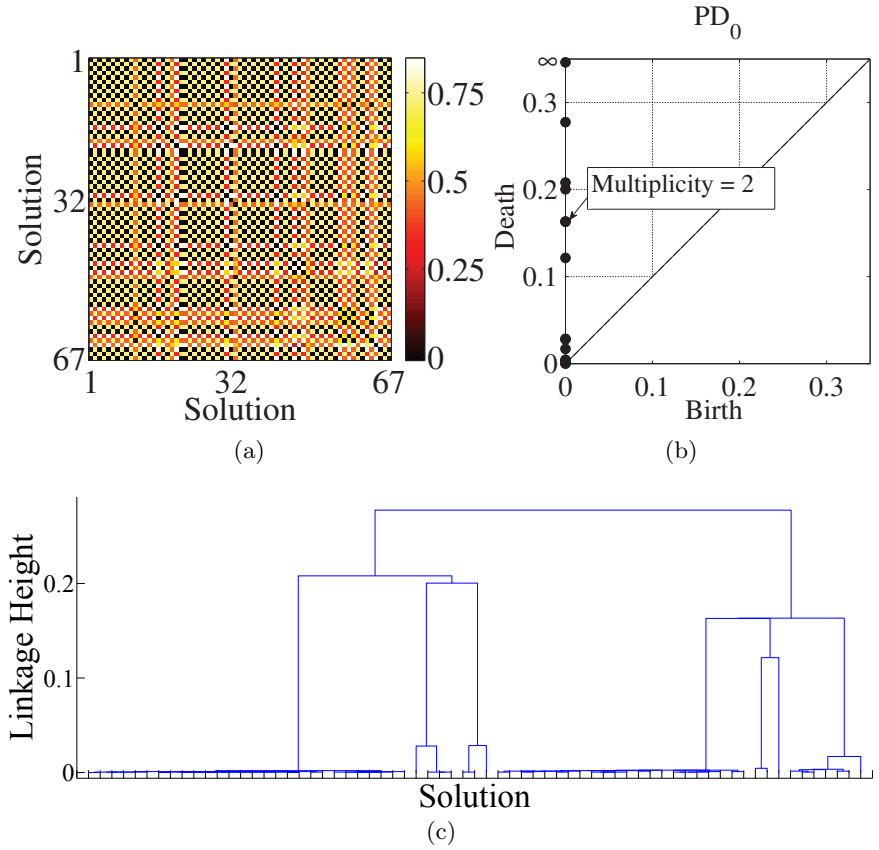


Figure 5.13: (a) Pairwise d_B distances between the EQ and REQ points in $X = \{\mathsf{PD}(\omega_n) \mid n = 1, \ldots, 67\}$ of the Komogorov flow found at $Re = 26.43$ using Newton's method. (b) Corresponding persistence diagram $\mathsf{PD}_0(X, d_B)$ shows a clear gap between the points with death value $\theta_d = 0.1215$ and $\theta_d = 0.0285$. We interpret this gap as the separation between the signal and noise. (c) Dendrogram for the EQ and REQ points with the pairwise distances given by the matrix shown in part (a), with values halved to show the correspondence to the persistence diagram $\mathsf{PD}_0(X, d_B)$ in (b). The dendrogram indicates the presence of seven clusters with separation the same size as the above-mentioned gap.

5.7 Distinguishing Equilibria

We now apply the ideas presented in Section 5.6.1 to the problem of clustering symmetry-related equilibria of the Kolmogorov flow at $Re = 26.43$. As discussed in Section 5.1.1, we sample a turbulent trajectory, shown in Figure 5.1(b). To obtain the EQ and REQ solutions, we use a Newton method. The initial guesses for the Newton method are the vorticity fields ω that are local minima of the L^2 norm of $\partial\omega/\partial t$. In this way, we find a collection $X = \{\mathsf{PD}(\omega_n) : n = 1, \ldots, 67\}$ of persistence diagrams of vorticity

fields corresponding to EQ and REQ of the Kolmogorov flow. These 67 solutions may be related to one another through any composition of the coordinate transformations listed in Section 5.1.1. Hence, it is non-trivial to determine how many unique classes of solutions there are and which solutions belong to which class. To perform this analysis, we use persistent homology.

We start by analyzing $\mathsf{PD}_0(X, d_B)$. The pairwise distances between the points in X are shown in Figure 5.13(a). As is discussed in Section 5.4, the distance between persistence diagrams of vorticity fields related by symmetry is small, while persistence diagrams corresponding to the vorticity fields that are not symmetry related differ by a larger amount. This implies that we can reformulate the question of identifying symmetry classes of equilibria as a clustering problem.

The persistence diagram $\mathsf{PD}_0(X, d_B)$, depicted in Figure 5.13(b), shows a clear gap between the persistence point with death value $\theta_d = 0.0285$ and the persistence point with death value $\theta_d = 0.1215$. We interpret this gap as separation between the signal and noise (numerical errors). Indeed, 0.0285 is just twice the estimate of the lower bound on numerical errors for the Kolmogorov flow obtained in Section 5.5. There are 7 points in $\mathsf{PD}_0(X, d_B)$ with death coordinate greater than 0.12, and so we conclude that there are seven distinct symmetry classes of solutions.

The same conclusions can be achieved using a dendrogram based on single-linkage clustering. Figure 5.13 (c) shows the dendrogram for the different EQ a REQ points with the pairwise distances given by the distance matrix shown in Figure 5.13 (a), with values halved to show the correspondence with $\mathsf{PD}_0(X, d_B)$. Clearly, there are again seven different clusters with distances larger that 0.12.

We validate the results of the persistence homology analysis by performing clustering using the Fourier amplitudes as follows. If $\hat{\omega}(k_x, k_y)$ is the Fourier amplitude of a mode (k_x, k_y), then a translation of the pattern in the x or y directions in real space merely adds to the phase of $\hat{\omega}(k_x, k_y)$, leaving the magnitude unchanged. Hence, by comparing the amplitudes of the Fourier modes we could group vorticity fields which are related by translations. Since the conjugate modes $\hat{\omega}(\pm k_x, \pm k_y)$ relate fields which are related by inversion, to group the vorticity fields which are related by a combination of inversion

and translation, we sum the amplitudes of the conjugate modes. Adding the amplitudes of conjugate modes yields a "reduced matrix," which is unique for all the vorticity fields related by the coordinate transformations that leave the governing equations of the flow invariant. This approach also yields 7 distinct classes.

An analysis of $\mathsf{PD}_0(X, d_{W^p})$, $p = 1, 2$, yields the same results. There are several gaps between the death values of the points in the persistence diagrams. Again, one of the gaps starts at roughly twice the value of the estimated lower bound of the noise. However, the separation is less pronounced. As discussed in Section 5.5, the d_{W^p} metrics capture all the differences between the persistence diagrams, and the local numerical errors are summed together. Thus, a large number of small errors can obscure the distinction between the signal and noise.

5.8 Stable Periodic Orbit of the Kolmogorov Flow

In the previous section, we demonstrated the practicality of using persistent homology to cluster equilibria that are symmetry-related. In this section, we extend these ideas to the setting of recurrent orbits in the context of the Kolmogorov flow.

As is indicated in Figure 5.1(a), the projection of the orbit onto the real parts of the three dominant eigenvectors suggests a periodic orbit that is undergoing a slow drift in the direction of the continuous symmetry. The nature of this drift is reinforced by tracking this orbit in the space of persistence diagrams; since persistent homology is invariant under the continuous symmetry, this type of drift is not present in Per. As a result, we expect the time series to lie on a closed loop in Per. This is consistent with the information provided by the distance matrix of Figure 5.7, in which the dark lines parallel to the diagonal indicate that the distance between persistence diagrams becomes very small at regular time intervals.

For the remainder of this section, we use the ideas of Section 5.6.2 to verify that a circle provides a good description of geometry of the point cloud $X \subset \mathsf{Per}$ generated by the time series sampled from the Kolmogorov flow. More precisely, we show that there is a single dominant feature in $\mathsf{PD}_0(X, d_B)$ and a single dominant feature in $\mathsf{PD}_1(X, d_B)$,

which agrees with the persistent diagrams for a circle.

There are two issues that need to be considered: the first is the size of the data set, and the second is the spacing between the data points. As is indicated in Section 5.6, we use the Vietoris-Rips complex to compute persistent homology of point clouds. We remark that given N data points, the full Vietoris-Rips complex has 2^N cells. Considering this, we complete our analysis with the distance matrices corresponding to d_B, d_{W^1}, and d_{W^2} for 500 points, or roughly three periods of the Kolmogorov flow. In the next section, we introduce techniques for computing persistence on larger point clouds, which could arise due to increased sampling rates, sampling more periods, or both.

Since we are sampling from a single continuous trajectory, the fact that $\mathsf{PD}_0(X, d_B)$, as shown in Figure 5.14(a), suggests the existence of a single component does not come as a surprise. The persistence diagrams for $\mathsf{PD}_0(X, d_{W^p})$, $p = 1, 2$, yield similar results and are not shown. However, it is worth noting that this is not a foregone conclusion as the location of and spacing between the points of the time series are dependent upon the speed along the periodic orbit. As is clear from Figure 5.7(a), the speed of the trajectory is not constant. However, it is fairly smooth, thus we do not expect extreme differences in the spacings between points.

As discussed in Section 5.6.3, we compare the noise threshold, $\theta = 0.01$ (Table 5.1), to the maximum consecutive sample distance, $d_B = 0.0654$ (Figure 5.7(a) caption), and the diameter of the point cloud, 2.64 (Figure 5.8(a)). The maximum consecutive sample distance is more than six times larger than the length scale of the noise for this system. However, the diameter of the point cloud is more than forty times larger than the consecutive sample distance. Thus, features on the length scale of one fortieth of the diameter of the entire point cloud will be resolved, which is sufficiently small to consider this an adequate sampling. We will return to this issue in the next section.

As indicated in Figure 5.14(b), the persistence diagram $\mathsf{PD}_1(X, d_B)$ clearly detects a single dominant loop along which the data is organized. Thus, we conclude that in Per, equipped with the metric d_B, the point cloud X generated by the time series forms a loop with a minimal radius of 0.1344. Table 5.2 shows the coordinates of the persistence point

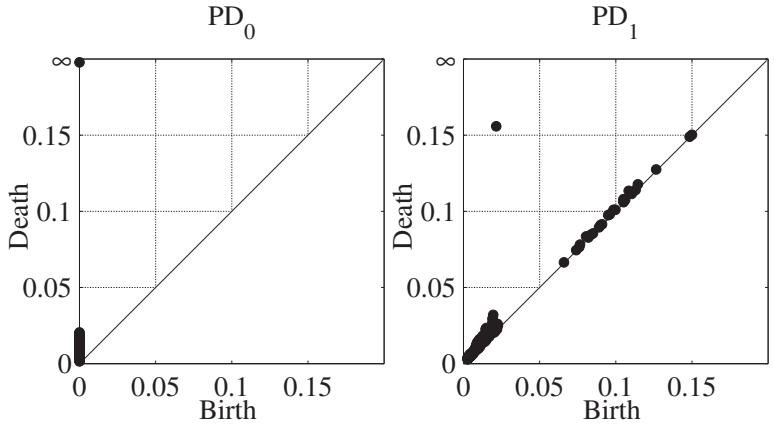


Figure 5.14: (a) The persistence diagram $\mathsf{PD}_0(X, d_B)$ for Kolmogorov flow at $Re = 26.43$. Since all points with finite death coordinates die before 0.025, there is only a single dominant point. (b) The persistence diagram $\mathsf{PD}_1(X, d_B)$, showing the single dominant generator at $(0.0215, 0.1559)$.

PD_1	Dominant coordinate	Max lifespan	2^{nd} largest lifespan
d_B	$(0.022, 0.156)$	0.134	0.013
d_{W^2}	$(0.075, 0.405)$	1.366	0.105
d_{W^1}	$(0.703, 2.069)$	0.330	0.016

Table 5.2: The coordinate of the dominant point in the persistence diagram $\mathsf{PD}_1(X, d_\star)$ for $\star = B, W^2, W^1$, its lifespan, and the second largest lifespan.

with the longest lifespan, its lifespan, and the second longest lifespan for each of the persistence diagrams $\mathsf{PD}_1(X, d_\star)$, $\star \in B, W^2, W^1$. As the table indicates, the lifespan of the dominant point is an order of magnitude larger than the next longest lifespan in each case, and so there is a single dominant feature in $\mathsf{PD}_1(X, d_\star)$. Additionally, note that the second longest lifespans are as small or smaller than the lower bounds on numerical errors indicated by the first row of Table 5.1.

5.9 Almost-Periodic Orbit of Rayleigh-Bénard Convection

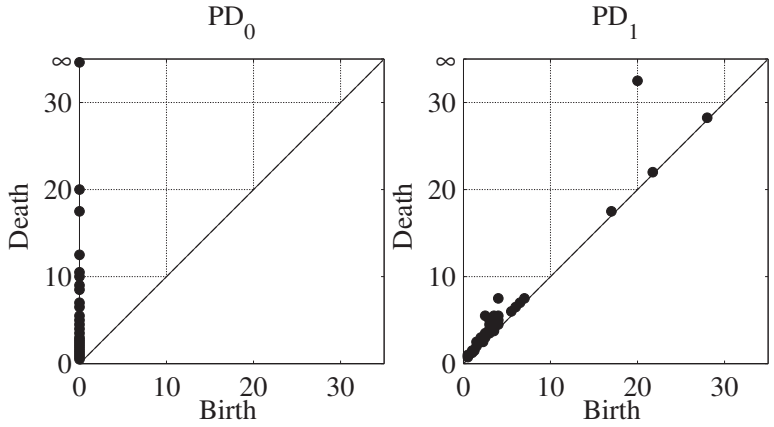
As mentioned in Section 5.6.3, characterizing the geometry of a continuous trajectory becomes a challenge in the setting of dynamics with multiple time scales. To demonstrate this, we consider the numerical simulation of Rayleigh-Bénard convection, where from multiple perspectives it appears that the trajectory is close to a periodic orbit and

that the rate of change in the patterns of the temperature field is far from constant. This can be clearly seen visually (see video 6, 7, or 8 in the supplementary materials of [1]). Moreover, both the speed plot, Figure 5.7(b), and the distance matrix, Figure 5.8(b), suggest recurrent dynamics. However, we note that the rate of change, especially using the bottleneck distance, is typically small except for short periods of time at which the speed spikes. The distance matrix has a distinct checkerboard pattern, with the edges corresponding to the spikes, again indicating a rapid and large change in location in the space of persistence diagrams.

The maximum bottleneck distance between the consecutive sampling points is 83.5 (Figure 5.7(b) caption), while the diameter of the point cloud is only $d_B = 99.5$ (Figure 5.8(b)). Therefore, we expect that significant portions of the trajectory are missing. Indeed, Figure 5.15(a) shows that there are several persistence points in $\mathsf{PD}_0(X, d_b)$ with a (finite) death coordinate larger than ten. Thus, at a length scale of 20 (which is forty times larger than the noise threshold), the sample of the trajectory is broken into several pieces. The largest gap between different pieces of the trajectory is 40, as indicated by the persistence point with coordinates $(0, 20)$. This means that the sampling rate is far from adequate.

The diagram $\mathsf{PD}_1(X, d_B)$ in Figure 5.15(b) contains a single dominant point at $(20, 32.5)$ with lifespan 12.5. However, unlike in our analysis of the Kolmogorov flow in the previous section, we cannot argue that this point corresponds to a single dominant loop along which the data is organized because of the gaps in the sampling of the orbit. As mentioned in Section 5.6.3, the missing parts of the orbit could introduce loops of similar size corresponding to secondary structures. These structures might occur due to the fact that the loop corresponding to the underlying almost-periodic dynamics might be twisted, pinched, or bent in Per. In order to obtain information about secondary structures, we require a faster sampling rate.

We increased the sampling rate considerably and collected approximately 7×10^5 equally-spaced snapshots of the temperature field over four-and-a-half periods and compute the associated persistence diagrams, producing a point cloud $Y \subset \mathsf{Per}$. The maximal distances between the consecutive frames for the increased sampling rate drop to



(a) (b)

Figure 5.15: Persistence diagrams for 500 points of Rayleigh-Bénard convection, generated from the distance matrix shown in Figure 5.8(b). (a) The diagram $\mathsf{PD}_0(X, d_B)$ shows the appearance of persistence points with death values significantly greater than the noise threshold, indicating that the sampled trajectory is broken into pieces and sampling is not fast enough to resolve the periodic orbit. (b) The diagram $\mathsf{PD}_1(X, d_B)$ shows the presence of a loop that is born when the pieces of the orbit merge together.

$d_B = 4$, $d_{W^1} = 28$, and $d_{W^2} = 6.52$. The new value of d_B is much closer to our estimate of the numerical error ($e = 0.5$) and it is more than 24 times smaller than the diameter of the point cloud generated from the slower sampling. Since the point cloud could only increase in diameter through increasing the sample rate, we consider this sampling rate to be satisfactory.

Our next step is to use the ideas introduced in Section 5.6.3 to reduce the size of the sample and to complete our analysis. First we construct a δ-dense, δ-sparse subsample Y' of the point cloud Y. The smallest value of δ for which we were able to compute the persistence diagrams $\mathsf{PD}(Y', d_B)$, using 256 GB of memory, is $\delta = 4.5$. This value is only slightly larger than the largest distance between the consecutive states and, since the diameter of the subsampled point cloud is 99.5, the relationship between the length scale of the smallest detectable feature and the length scale of the diameter of the point cloud is still sufficient to resolve the geometry of the dynamics. The resulting persistence diagrams $\mathsf{PD}(Y', d_B)$ are shown in Figure 5.16.

As shown by $\mathsf{PD}_0(Y', d_B)$, Figure 5.16(a), the point cloud merges to a single connected component at $\theta = 4.5$. This indicates that the sample of the trajectory is not

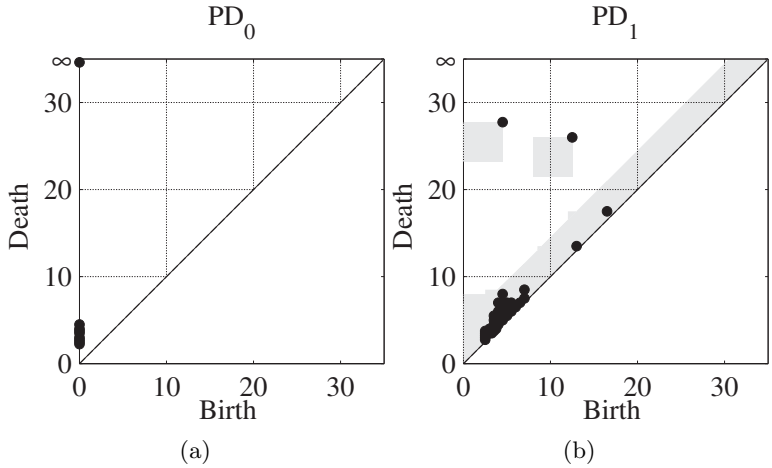


Figure 5.16: Persistence diagrams for 70,000 points of Rayleigh-Bénard convection subsampled with $\delta = 4.5$, resulting in a point cloud $Y' \subset \mathsf{Per}$ with 523 points. (a) Persistence diagram $\mathsf{PD}_0(Y', d_B)$ indicating a single dominant component above the noise threshold. (b) Persistence diagram $\mathsf{PD}_1(Y', d_B)$ with subsampling error bounds shaded in gray.

broken into different pieces separated from each other. Since the maximum consecutive distance between any two points in Y is 4, the loop along which the data is organized should be present for $\theta = 2$. However, after subsampling, it is possible that the loop will not be born until $\theta = 2+\delta$. Looking at the diagram $\mathsf{PD}_1(Y', d_B)$ in Figure 5.16(b), we see that it contains a dominant point at $(4.5, 27.75)$, and so the loop was indeed born before $\theta = 2 + \delta$. This is the loop along which the point cloud is organized. Now, there is another point, $(12.5, 26) \in \mathsf{PD}_1(Y', d_B)$, with lifespan 13.5. This point corresponds to a secondary structure of the orbit. Indeed, it can be seen from the distance matrix for the δ-sparse, δ-dense subsample (not shown for brevity) that the part of the orbit corresponding to the fast dynamics (missing for the slow sampling rate) revisits very similar states before continuing along the main loop. However, the development of more sensitive tools is required to fully understand these secondary features.

We now turn our attention to the differences between the persistence diagrams of the original point cloud Y and its subsample Y'. The traditional stability theorem implies that $d_B(\mathsf{PD}(Y, d_B), \mathsf{PD}(Y', d_B)) \leq \delta$, and so there exists a bijection between the points in $\mathsf{PD}(Y, d_B)$ and $\mathsf{PD}(Y', d_B)$ such that the distance between matched points is less than 4.5. According to Remark 5.17, for the dominant point $(4.5, 27.75) \in \mathsf{PD}_1(Y', d_B)$,

there is exactly one corresponding point in $\mathsf{PD}_1(Y, d_B)$. This point is the unique point in $\mathsf{PD}_1(Y, d_B)$ that lies inside of the shaded box touching the point $(4.5, 27.75)$, see Figure 5.16(b). The same is true for the other dominant point. Moreover, there are no points in $\mathsf{PD}_1(Y, d_B)$ outside of the shaded regions. Points in $\mathsf{PD}_1(Y, d_B)$ that do not correspond to the off-diagonal points in $\mathsf{PD}_1(Y', d_B)$ can appear only $\delta/\sqrt{2} \approx 3.18$ far away from the diagonal.

Chapter 6

Spatiotemporal Pattern Analysis for Spiral Defect Chaos

In this final chapter, we return to another problem in Rayleigh-Bénard convection (RBC) flow. As in Chapter 5, we study the mid-plane temperature field of simulated RBC, although at a much higher Rayleigh number (Ra $\approx$ 4000). As a result, the patterns generated in the temperature field are much more complex (compare Figure 6.1 to Figure 5.2(b)). The dynamics at this value of Rayleigh number exhibit what is called *spiral defect chaos*, named for the spirals that are commonly seen in such flows (there are at least three spirals in the temperature field shown in Figure 6.1). In this chapter, instead of focusing on the structural dynamics of the system as we did in Chapter 5, we will focus on the evolution of the patterns themselves.

Figure 6.1: A sample mid-plane temperature field from simulated Rayeigh-Bénard convection at Ra $\approx$ 4000.

The patterns simulated from RBC flows are called *locally-striped* due to the appearance of parallel black and white bands, and are plentiful in nature in the form of stripes on animals such as zebra and tropical fish, ripples on sand dunes, and cirrocumulus cloud formations, to give just a few (of many) examples. In the laboratory, locally-striped patterns are not only seen in RBC, but also in condensed matter physics in the

nematic and smectic phases of liquid crystals [50].

Our work lays the foundation for a set of techniques that can be used to study the complicated patterns of SDC, and is designed to be transferrable from numerical simulations to experimentally-collected data. The approaches we present build on classical topological methods for studying locally-striped patterns, augmented by persistent homology and established techniques in computer vision that fall under the general heading of object recognition.

An outline of this chapter is as follows. Section 6.1 describes classical methods for locating point defects in a locally-striped pattern, and Section 6.2 shows how persistent homology can be used to augment this set of defects. In Section 6.3, we describe six canonical defect types (Figure 6.2) that are well-studied for locally-striped patterns, and present an algorithm that locates and classifies each of these canonical defects. Due to the complexity of the patterns in the SDC regime (Figure 6.1), not all defects can be classified as one of these six types. We use this algorithm to compute descriptive statistics showing the frequency of each of the canonical defects across a time series of simulated RBC flow, as well as statistics showing the percentage of the area of the temperature field that is considered to be defect-free, classified, or unclassified. In Section 6.4, we use a technique from the field of computer vision to give a more nuanced description of a defect pattern, using both topological point defects and defects located using persistent homology to localize the pattern. Finally, in Section 6.5, we build on the work of Section 6.4 to show how methods from computer vision typically used for object recognition might be used to locate and analyze coherent and recurrent structures in the SDC regime of RBC flow.

6.1 Classical Methods for Studying Locally-Striped Patterns

The simplest type of pattern seen in the mid-plane temperature field of RBC is a striped pattern, which is due to the shape of the convection rolls that characterize the flow. This phenomenon occurs when ΔT is large enough so that the system passes from a conductive state to a convective state [43]. Each stripe is called a *roll* in the pattern,

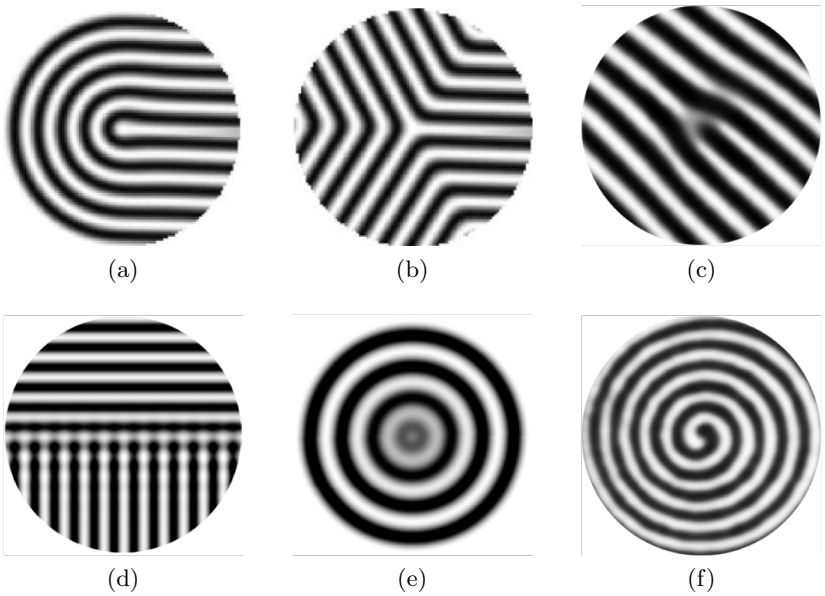


Figure 6.2: Canonical defects in convection patterns: (a) convex disclination, (b) concave disclination, (c) dislocation, (d) grain boundary, (e) target, and (f) spiral. Images from [52, 53, 54].

since each convection roll contributes one stripe in the pattern. Any deviation from the ideal state is considered a *defect* in the pattern. Typically, defects seen in RBC fall into one of six canonical types, pictured in Figure 6.2, although there are also studies of more complicated configurations [43, 51]. We call regions of the pattern that are locally-striped and free of defects *defect-free*.

Classical methods for studying locally-striped patterns use the *phase equation* [43] as a simplistic model for the pattern. We briefly describe this approach here. Suppose that the mid-plane temperature field is given by $u : D \to \mathbb{R}$, where $D \subseteq \mathbb{R}^2$ is a rectangular or circular domain. Following [55], in an ideal state the temperature field may be described completely by the equation

$$u(\vec{x}) = A(\vec{x}) \cos[\phi(\vec{x})],$$

where $A : D \to \mathbb{R}$ gives the amplitude (constant and nonzero in the ideal case) and $\phi(x)$ is the phase. In the case of an ideal striped pattern, the phase exhibits a continuous translational symmetry. This description also applies when the phase ϕ and amplitude

A change continuously over the domain, allowing for distortions from the ideal state.

Another useful scalar field in the study of locally-striped patterns is the *orientation field*, $\theta : D \to [-\pi/2, \pi/2)$, which gives the orientation of the rolls in the pattern. The orientation field is theoretically computed by expressing the *local wave vector*

$$\vec{k}(\vec{x}) = \vec{\nabla}\phi(\vec{x})$$

as a director field (so that positive and negative directions are equivalent), and then expressing the director field in polar coordinates with angles in $[0, \pi)$ as

$$\overleftrightarrow{k}(\vec{x}) = \left(k(\vec{x}), \theta(\vec{x}) + \frac{\pi}{2}\right).$$

The phase shift of the director field is a matter of convention, so that the orientation field points along the rolls instead of normal to them.

6.1.1 Topological Defects

Using the orientation field θ defined in the previous section, we compute a quantity called the *Poincaré index* for each point in the domain, which essentially computes a winding number with respect to the orientation field around each point. Suppose that γ is a closed curve in D and that a function $\tilde{\theta}$ is chosen such that $\tilde{\theta} \equiv \theta \mod \pi$ is continuous on all but, perhaps, a single point on γ. That is, $\tilde{\theta}$ gives a continuous branch of the orientation field θ on the curve γ, with the exception of at most a single point. In general, the Poincaré index associated to the closed curve γ, oriented counterclockwise, may be computed by taking

$$I(\gamma) = \frac{1}{\pi} \oint_{\gamma} d\tilde{\theta}(\vec{x}),$$

which will take on integer values. Letting γ_r be a closed, circular curve centered at $\vec{x}$ with radius r, we can define the Poincaré index at a point by taking

$$I(\vec{x}) = \lim_{r \to 0} \frac{1}{\pi} \oint_{\gamma_r} d\tilde{\theta}(\vec{y}),$$

where $\tilde{\theta}$ is chosen (for each r) so that, if possible, it is continuous on all but a single point on γ_r. Points $\vec{x}$ for which $I(\vec{x}) \neq 0$ are called *topological defects*, and occur where $\tilde{\theta}$ exhibits a local discontinuity regardless of the way $\tilde{\theta}$ is chosen. For two-dimensional domains D, these are the only types of topological defects [43], and are often called *point defects* since they occur at a point. That is, $I(\vec{y}) = 0$ for $\vec{y} \neq \vec{x}$ in a small neighborhood of $\vec{x}$. (In contrast, the region of local discontinuities of $\tilde{\theta}$ for orientation fields of smooth scalar fields defined on three-dimensional domains may form a line.)

The *charge* of the topological defect $\vec{x}$ is the Poincarè index $I(\vec{x})$. The charges of the defects exhibited in Figure 6.2 (a) and (b) are +1 and -1, respectively, and are called convex and concave *disclinations*, respectively. Intuitively, a convex disclination occurs at a focal point, and a concave disclination occurs where a roll branches. Notice that a dislocation pattern (c) is composed of a pair of convex and concave disclinations. The grain boundary exhibited in (d) is a chain of topological defects with alternating charges +1 and -1. A perfect target (e) exhibits a single topological defect with charge +2, but this only occurs when the rolls have perfect rotational symmetry in a neighborhood of the point defect. In real patterns, there will be a slight asymmetry which will cause the orientation field to have a pair of topological defects with charges both +1. Finally, the spiral pattern (f) has a pair of topological defects with charges both +1, the same as the target. To differentiate targets and spirals, the actual value of u may be used at the location of the topological defect, although this is not so reliable for real patterns; we give an alternative method for distinguishing spirals from targets in Section 6.3.2.

6.1.2 Computing the orientation field and its topological defects

While there are now many algorithms commonly used to compute the orientation fields for locally-striped patterns [56, 55, 57], these algorithms are formulated for systems that either have large aspect ratios or have patterns that exhibit large defect-free regions, which is not the case in our setting (see Figure 6.3). We find that the algorithm by Bazen and Gerez [58], originally formulated to locate singular points in fingerprint patterns (cores and deltas), is both accurate and performant in our setting and has the advantage of a rather simplistic description.

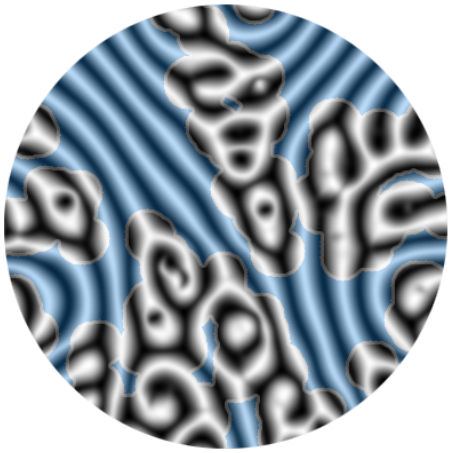


Figure 6.3: A sample mid-plane temperature field from simulated Rayeigh-Bénard convection with defect-free regions highlighted in blue.

While in theory the scalar field u is continuously defined on some domain $D \subset \mathbb{R}^2$, in our setting we are working with digital image approximations. Thus, D can be given a cubical complex structure, and the values of u are integers in $[0, 255]$. We shall detail the algorithm we use to compute the orientation fields here, which is based on the one by Bazen and Gerez [58]. Let $[G_x, G_y]^T$ be the gradient of u and define

$$G_{xy} = G_x G_y$$
$$G_{xx} = G_x^2$$
$$G_{yy} = G_y^2$$

pointwise. Let

$$\angle(x,y) = \begin{cases} \tan^{-1}(y/x) & \text{if } x \geq 0 \\ \tan^{-1}(y/x) + \pi & \text{if } x < 0 \text{ and } y \geq 0 \\ \tan^{-1}(y/x) - \pi & \text{if } x < 0 \text{ and } y < 0. \end{cases}$$

Then the orientation field is given by

$$\theta_\sigma = \frac{1}{2}\angle\Big(\mathcal{N}_\sigma(G_{xx} - G_{yy}), \mathcal{N}_\sigma(2G_{xy})\Big),$$

where $\mathcal{N}_\sigma$ is the application of a Gaussian blur of radius σ.

We then compute the topological defects by summing the differences in the orientations counter-clockwise around each 2×2 grid of pixels in the orientation field, with phase shifts added as necessary to define what would be locally-continuous values for $\tilde{\theta}$ (see Section 6.1.1). Any computations that result in a non-zero charge (after rounding to the nearest integer) are considered to be topological defects. Figure 6.4 shows a portion of the temperature field and two associated orientation fields using a Gaussian smoothing of $\sigma = 0.5$ and $\sigma = 2$ and their corresponding topological defects. As is readily seen, the amount of smoothing has a significant effect on the number of topological defects detected in the orientation field.

One drawback of the algorithm we use is that, unlike the algorithms in [56, 57], it does not provide the local amplitude A of the pattern. Changes in the local amplitude of a locally-striped pattern are studied in [59], where they are shown to be important in locating pattern defects. We present a method in Section 6.2.1 for detecting variations in the local amplitude of the pattern that uses homological methods.

6.2 Persistent Homology and Morse Theory

In general terms, persistent homology is a tool used to encode the topological features of a scalar field as points in the Euclidean plane, as explained in Chapter 5. Discrete Morse theory, described first by Forman in [60], gives an alternative description of this encoding, whereby the persistence points correspond to pairings of critical points of the underlying scalar field. During the computation of the persistence diagrams, it is possible to keep track of the underlying critical point pairings that give rise to the persistent homology generators, such as with the algorithms in the software packages Diamorse [61] and PHAT [62]. A summary of the algorithms behind these persistence computations is beyond the scope of this work, and we encourage the curious reader to refer to the sources provided for details. The computations contained in this analysis were performed using the PHAT software [62].

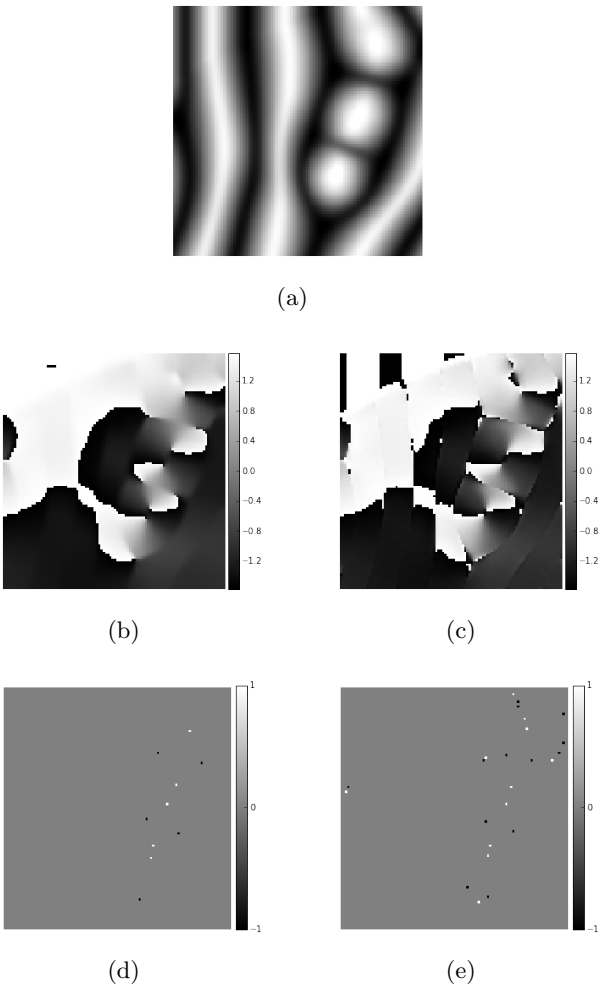


Figure 6.4: (a) A region of temperature field. The associated orientation field with smoothing at (b) $\sigma = 2$ and (c) $\sigma = 0.5$. Poincaré index computations at each point for smoothing at (d) $\sigma = 2$ and (e) $\sigma = 0.5$.

6.2.1 Persistent Homology and Local Amplitude Defects

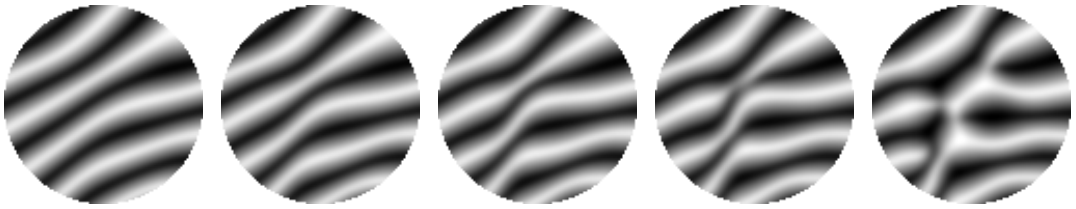
While topological defects are useful for detecting deviations from an ideal striped pattern that cannot be described by the phase equation due to singularities in the orientation field, defects that do result from smooth perturbations of the phase equation cannot be detected using these methods. One particularly compelling example of a pattern defect that initially involves a smooth deformation to the phase equation description is given by the *skew-varicose* instability, pictured in Figure 6.5. In this pattern formation, a region consisting of straight parallel rolls deforms continuously until the rolls pinch off and merge together. Figures 6.5 (a-d) exhibit changes that can be described by

continuous perturbations of the phase equation description, none of which result in the appearance of a topological defect with nonzero charge. It is not until frame (e) (more precisely at frame number 508) that the change in topology in the pattern is detected using these classical methods.

Egolf, et. al [55] examine the importance of the local wave number (the magnitude of the director field $\vec{k}$ in the previous section) to such types of instabilities. Approaches based on the local wave number are able to detect, for example, a decrease in spacing between the rolls as compared to the ideal state, or an increase in the local wave number. However, due to the nature of the algorithm, this computation is not very precise for complex patterns that have a low ratio of defect-free regions to regions near defects.

If the phase equation of the ideal state has constant amplitude, a local perturbation of A at an extreme value of u (e.g. on the top or bottom of a roll), while holding the phase ϕ fixed, will result in the formation of at least one saddle point in the temperature field. Thus, it is reasonable to assume that locating saddle points in the mid-plane temperature field will imply the detection of local amplitude fluctuations in a pattern that is locally-striped. Examining Figure 6.5 (d) visually, one can see that a variation in the local amplitude has taken place, evident by the appearance of saddle points on, or variations in the intensity of, the stripes.

Essentially, then, we have reframed the question of detecting local amplitude fluctuations in terms of a problem of locating saddle points in u. While there are existing methods for locating saddle points in digital images [63, 64], these methods determine all saddle points in the image. In an ideal locally-striped pattern, the values along ridges or valleys will have a constant value. However, real data will exhibit small fluctuations in these values, causing the appearance of many saddle points along the ridges and in the valleys of the striped pattern, which may not be considered true defects in the pattern. Thus, even after locating saddle points in the scalar field u, one needs a method by which interesting saddle points may be selected. We give here a method that utilizes persistent homology to locate saddle points in the convection patterns and which also gives a method by which to filter the saddle points considered in the analysis.



(a) (b) (c) (d) (e)

Figure 6.5: Example of a skew-varicose instability. Successive frames are from samples $t = 150, 350, 450, 500, 550$.

Using persistent homology to locate saddle points in a digital image

Recall that dimension zero persistent homology using a sublevel set filtration encodes how local minima in u merge together at saddle points. Each generator will be born at a local minimum, and will die at the saddle point to which the local minimum is paired. Thus, the critical cells corresponding to the death of a dimension zero generator will encode saddle points of the image. For dimension one persistent homology (also with a sublevel set filtration), a loop is first born at a saddle point and dies at the corresponding local maximum to which the saddle point is paired. Thus, critical cells corresponding to the birth of a dimension one generator also encode saddle points of the image. Figure 6.6 shows a sample mid-plane temperature field together with its sublevel filtration persistence diagrams PD_0 and PD_1. Panels (a) and (b) show the saddle points captured by PD_0, while panels (c) and (d) show the saddle points captured by PD_1. The saddle points overlaid on the temperature field have been colored according to the lifespan of the persistence points on the persistence plane. Saddle points detected by performing superlevel set filtrations are similar and are not displayed.

As is readily seen in Figure 6.6(a) and (c), a considerable number of saddle points that are not truly of interest have been captured by the persistence computations, the majority of which correspond to persistence points that are close to the diagonal. By filtering the saddle points via the location of their corresponding persistence points on the persistence plane, it is possible to get a cleaner signature. Figure 6.7 shows saddle points selected by regions on the persistence plane for both sublevel (red) and superlevel (blue) set filtrations. The saddle points selected from the sublevel set filtrations, in red, correspond to saddle points which have a value at or below the median pixel value

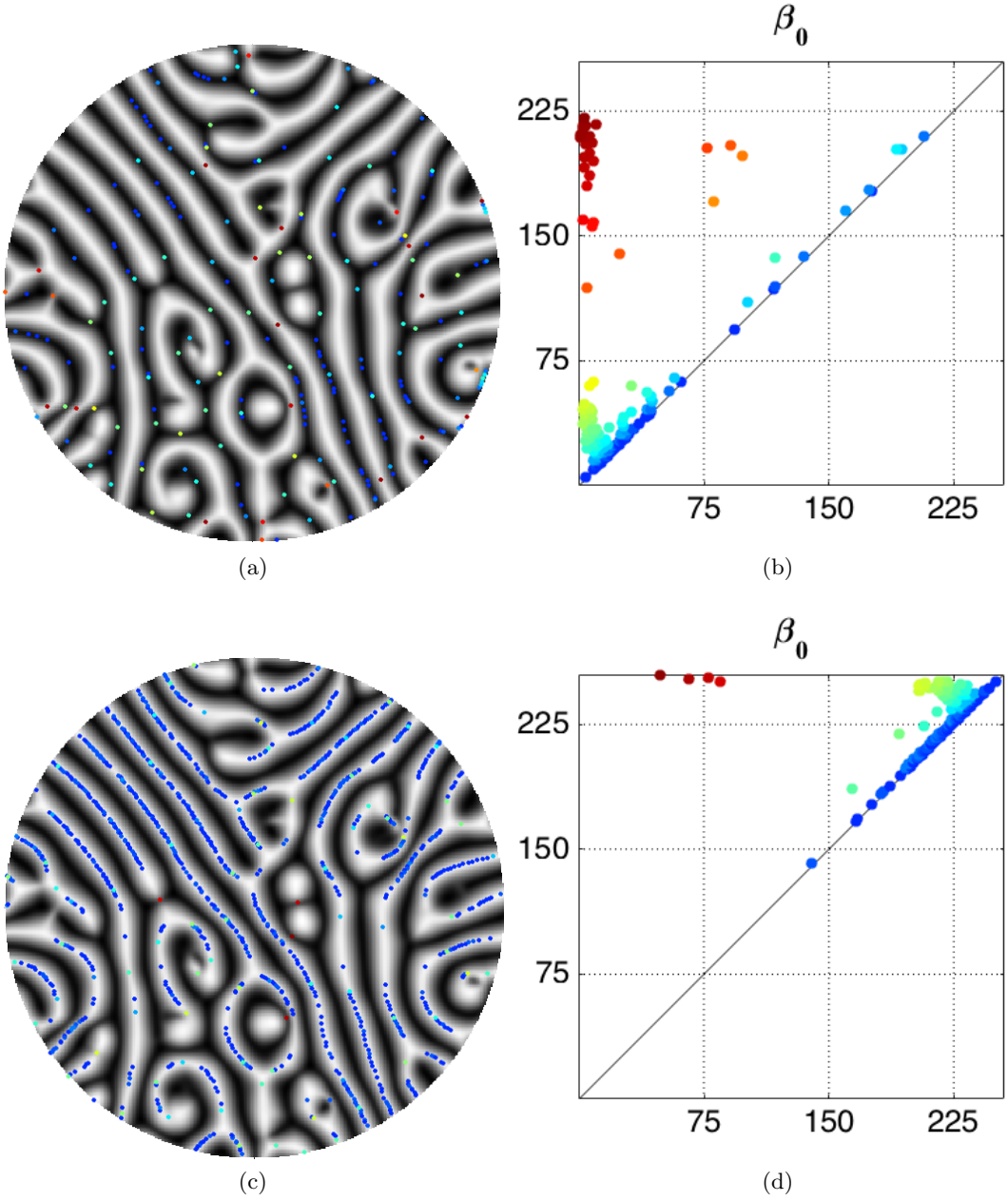


Figure 6.6: (a) Temperature field with saddle points detected by (b) PD_0 (sublevel filtration). Points are colored according to the lifespan of the persistence points. (c) Temperature field with saddle points detected by (d) PD_1 (sublevel filtration).

(black line in the persistence diagram), while those selected from the superlevel set filtrations, in blue, correspond to saddle points that have values at or above the median pixel value.

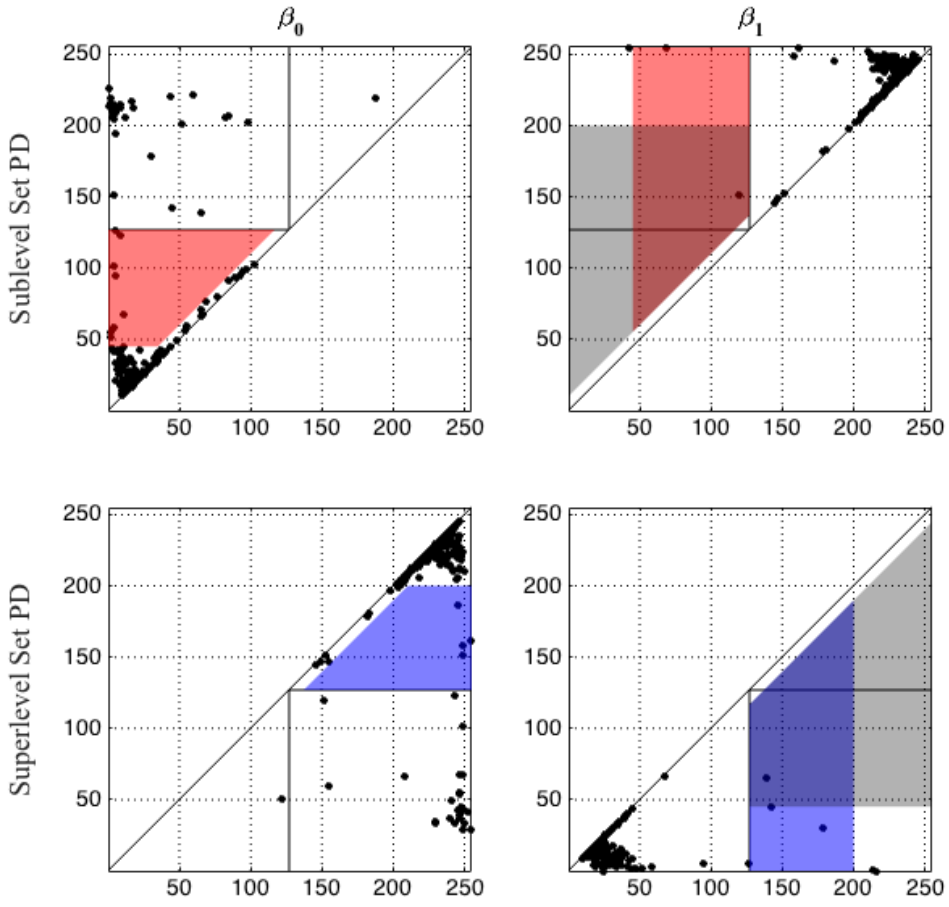
Using persistent homology to locate the emergence of new circular rolls

Another example of a local amplitude defect that can be detected by persistent homology is the emergence of a new roll in the target pattern (Figure 6.2 (e)). In this type of pattern, there is a radial symmetry about some central point x_0 in the phase ϕ of the ideal equation $u(x) = A(x)\cos(\lambda(x)\phi(x))$, where λ is the wavelength of the pattern. Assuming that $u(x)$ is smooth, the appearance of a new ring or radial roll in the pattern at a point x_0 is modulated instead by a modification of wavelength and local amplitude. Indeed, modeling this phenomenon by a shift in phase alone (keeping amplitude fixed) may result in a sharp corner if $\lim_{x\to x_0}\phi(x) \neq k\pi$.

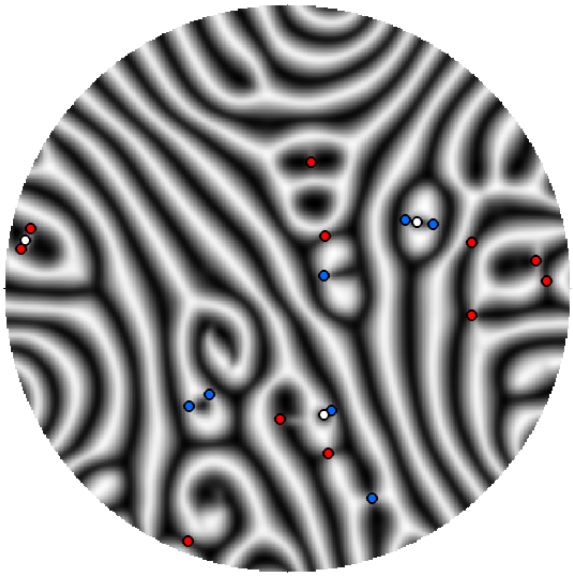
These local minima and maxima which mark the emergence of a new ring-shaped roll in the pattern may be characterized by the emergence of a persistent homology generator in $\mathsf{PD}_1(u)$. For the sublevel set filtration, these local maxima will be marked by the critical cell corresponding to the death of the dimension one homology generators that have a value at or below some cutoff value, and were born in lower temperature values. Similarly, for the superlevel set filtration, these local minima will be marked by the critical cell corresponding to the death of the dimension one homology generators that have a value at or above some cutoff value, and were born in higher temperature values. The choice of the value cutoffs are parameters that can be changed to capture either more or fewer of these types of defects.

6.3 Analysis of Canonical Defects in Numerically-Simulated SDC

In this section, we will perform an analysis of the six canonical defects exhibited in Figure 6.2 across $3,500$ frames of simulated RBC in the SDC regime. In Section 6.3.1 we describe the algorithm we use to classify the six canonical defects, which relies on persistent homological methods described in Section 6.3.2 for differentiating spirals from targets. We then give our analysis in Section 6.3.3.



(a)



(b)

Figure 6.7: (a) Partitions of the persistence plane for both sublevel and superlevel set filtrations used to find saddle points (red and blue regions) and the emergence of new circular rolls (gray). (b) Temperature field with corresponding saddle points (red and blue) and local minima/maxima (white) corresponding to circular rolls overlaid.

6.3.1 Algorithm for Identifying Canonical Defects

Our method for identifying the canonical defects in the temperature field is similar to existing methods [65], where the singular points are first clustered into groups, and then those groups are further analyzed to assign a defect type. The relevant spatial length scale for this data is $\lambda = 30$ pixels, which corresponds to a single wavelength. We find that single-linkage clustering at length-scale $\lambda/3$ gives reasonable results, and we will see in Section 6.4.1 that this factor of the wavelength is a natural length scale in the pattern. Topological defects are clustered using single-linkage hierarchical clustering at scale $\lambda/3 = 10$ pixels, and then each cluster is analyzed according to the following algorithm.

Monopoles

Clusters consisting of a single topological defect are called monopoles. Necessarily, these correspond to a topological defect with charge either $+1$ or -1, and so this type of defect is classified according to its topological charge, either a *convex disclination* $(+1)$ or a *concave disclination* (-1).

Dipoles

Clusters containing two topological defects are called dipoles. There are two likely combinations of topological defects for a dipole, which we describe below. The combination of two -1 charged topological defects is rare and is not included, since this would not result in a canonical defect pattern; instead, this combination would be considered unclassified using this schema.

Two opposite charge topological defects

When the dipole consists of a +1 and -1 pair of topological defects, then the defect is classified as a *dislocation pair*.

Two +1 charge topological defects

A dipole consisting of two +1 charged topological defects is either a *spiral* or a *target*. As discussed in Section 6.3.2, the presence of a persistent homology dimension one generator is used to differentiate the two types of defects. We also compute the center and radius of the spiral or target as follows. The center is computed as the

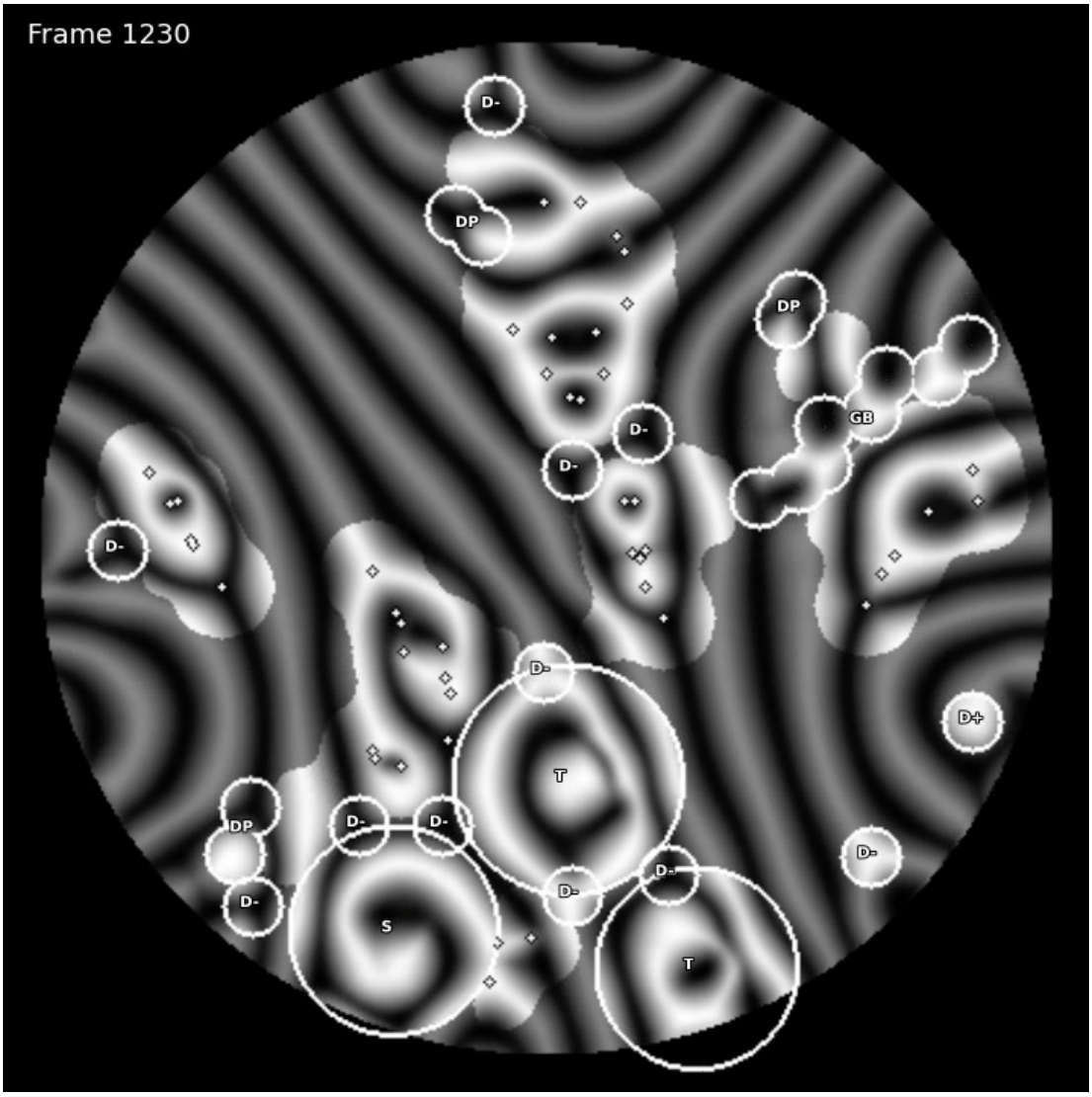


Figure 6.8: Classification of the six canonical defects using the algorithm described in Section 6.3.1. Labels are as follows: D+ = convex disclination; D- = concave disclination; DP = dislocation pair; GB = grain boundary; S = spiral; T = target. Defect-free regions are shaded 50% gray, and unclassified regions and their corresponding topological defects are left unshaded and unlabeled.

midpoint between the two +1 topological defects. The radius is given by the distance from this center point to the nearest topological defect (excluding the pair that define the spiral or target).

Tripoles and larger clusters

Clusters consisting of three or more topological defects are checked to be either *grain boundaries*, *spirals* or *targets*. While additional classifications are certainly possible [65], we are only concerned with identifying the six canonical types at this stage of the investigation.

Grain boundaries

We identify a grain boundary as a linear chain of alternating +1 and -1 topological defects that are offset from each other by some amount. To determine if the cluster of topological defects fits this description, we compute the line of best fit through the defects using a singular value decomposition of the defect points, and then take the coefficients of each of the data points with respect to this primary axis of the singular value decomposition. This ordering is then used to check to see if the signs of the defects are alternating.

If so, then we use the second-largest eigenvalue, ξ, in the singular value decomposition to get a measure of spread of the datapoints along the primary axis. If the eigenvalue ξ satisfies

$$\frac{2}{15}\lambda \leq |\xi| \leq \frac{\lambda}{2},$$

then we classify the cluster as a grain boundary defect, where the values appearing in the inequalities were chosen by the author as a result of direct observation of many frames. We say that the length of the grain boundary is the number of topological defects in the cluster.

Spirals and targets

It is often the case that a spiral or a target will be small and surrounded by concave disclinations. We identify these types of patterns by taking the centroid of the two positively-charged topological defects in the cluster, and then compute the distance from this centroid to every negatively-charged topological defect in the cluster. We then

compute the minimum distance of the two positively-charged defects to the centroid, m^+, and the minimum distance of all negatively-charged defects to the centroid, m^-. If m^- is greater than the max of either $1.25m^+$ or $(2/3)\lambda$ (to eliminate defect clusters that are clustered too tightly), then the negatively-charged topological defects are marked as concave disclinations. The two positively-charged topological defects are then processed as either spirals or targets in the same way that dipoles are classified.

Unidentified clusters

Due to the complicated nature of the patterns seen in SDC, many clusters of topological defects will not result in a classification using the above algorithm. These clusters are marked as *unclassified*.

Defect-free regions

The portion of the temperature field in the complement of the region that falls within $(2/3)\lambda$ of the topological defects are considered to be defect-free. These are considered to be portions of the temperature field that consist of parallel rolls.

Figure 6.8 shows a labeling of the temperature field at frame 1230 with the six canonical defects, labeled using the algorithm described above. The unclassified regions visually contain both a spiral and many targets, though the arrangement of the topological defects in these regions was too complex to yield a positive identification. The grain boundary on the right-hand side of the image seems to have a mis-identification on the left of the chain. Further refinement to the algorithm will both add to the number of correctly-identified canonical defects and also make the classifications more accurate. However, we believe the algorithm described here shows that this method is feasible to use with this dataset, and accurate enough for a preliminary analysis.

6.3.2 Using Persistent Homology to Differentiate Spirals from Targets

Ideal spirals and targets are easily differentiated from one another by the using the value of the temperature field in conjunction with the appearance of a pair of topological defects of positive unit charge: a spiral is classified when the temperature field at the topological defects has one high and one low value, and a target is classified when the temperature field at the topological defects are either both high or both low. However,

in the complicated patterns exhibited in SDC, the location of the topological defects relative to the visible pattern may cause misclassifications if this method is used.

To address this issue, we will use persistent homology to obtain a more robust signature for differentiating spirals from targets. In an ideal spiral, one of the arms can be contracted along the spiral to the boundary, and so the first dimension persistent homology should be relatively trivial for this pattern. In contrast, an ideal target must be accompanied by the presence of a dominant first dimension persistence point: for a target with a low temperature value in the center, there will be a dominant generator in the superlevel set dimension one persistence diagram, and for a target with a high temperature value in the center, there will be a dominant generator in the sublevel set dimension one persistence diagram. Thus, we can use the presence or absence of a critical cell from a dominant first homology generator to differentiate a spiral from a target.

6.3.3 Analysis

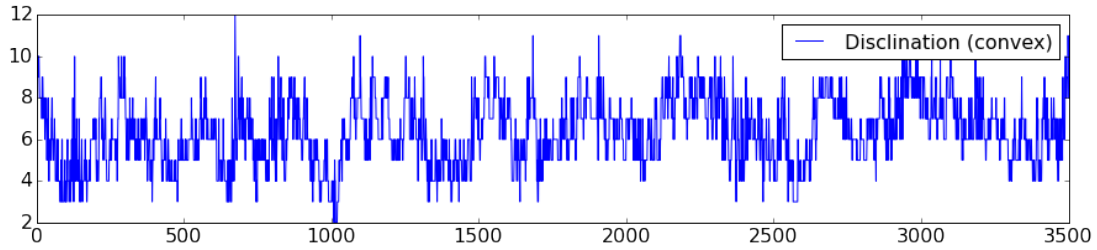
We use the algorithm in Section 6.3.1 to study $3,500$ frames of numerically-simulated data taken 0.1 time units apart. For each frame in the time series we compute:

- The number of each of the six canonical defects: disclinations (concave or convex), dislocation pairs, grain boundaries, spirals, and targets.
- The histogram distributions for the counts of each of the six canonical defects.
- The total area of each of the classified cluster regions, unidentified cluster regions, and defect-free regions as a percentage of the total area of the temperature field.

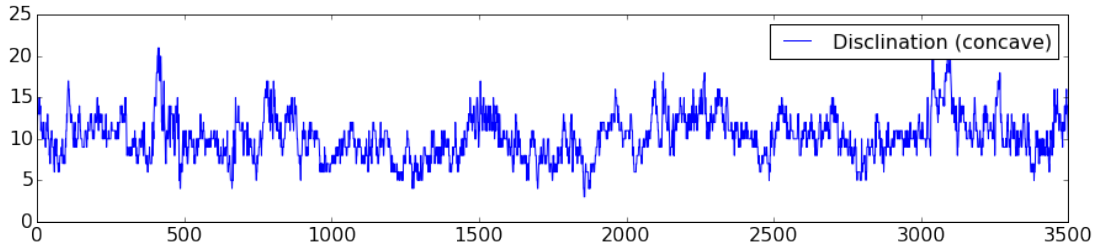
We then plot these totals as a function of the frame number (Figure 6.9).

6.4 Moving Beyond the Canonical Defects

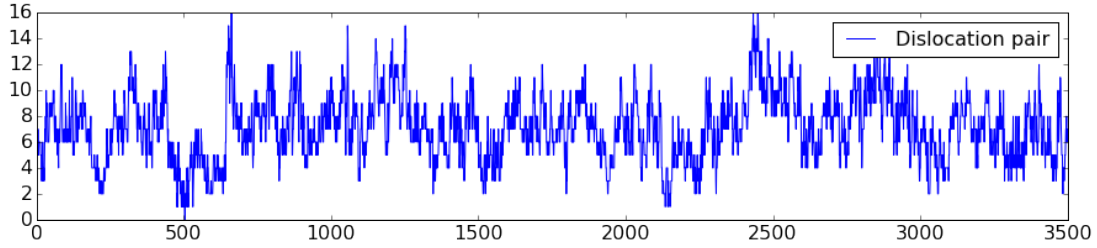
To extend defect classification beyond the six canonical defects, we turn to ideas from the field of computer vision. While the topological defects assign a global property to a local point (by encoding the behavior of the orientation field around a defect point),



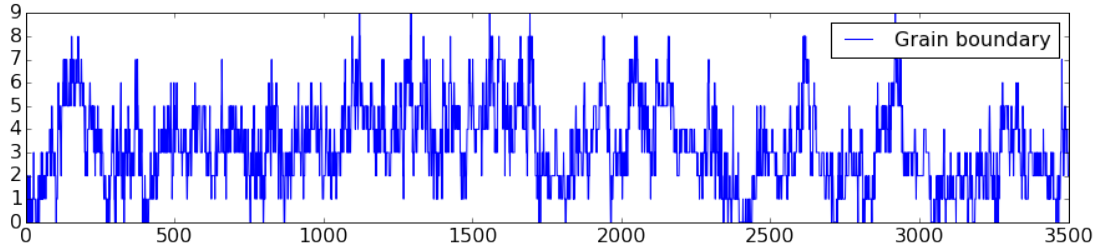
(a)



(b)

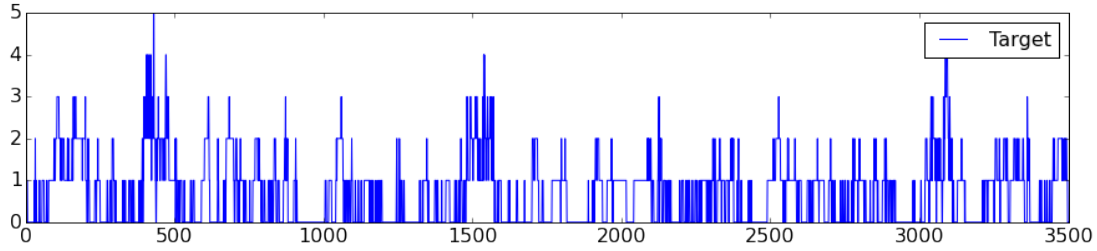


(c)

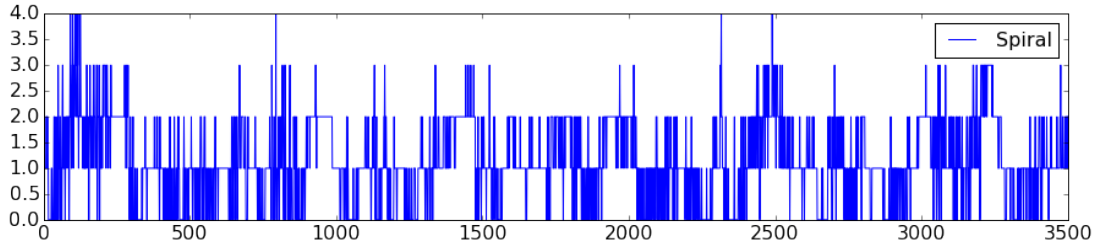


(d)

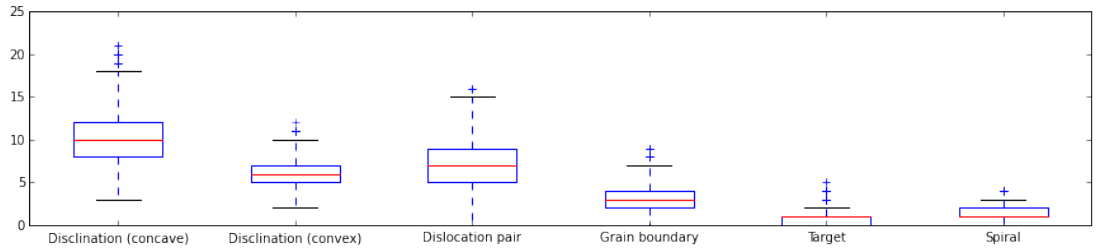
Figure 6.9: Counts of the number of occurrences of each of the six types of canonical defect patterns over 3,500 frames sampled 0.1 time units apart: (a) convex disclination, (b) concave disclination, (c) dislocation pairs, (d) grain boundaries, (e) target, and (f) spiral. The distributions of each of the counts in (a) to (f) are shown as box and whisker plots in (g). The percentage of the temperature field that is unclassified, classified, or considered to be defect-free, is shown in (h). (Continued in next figure.)



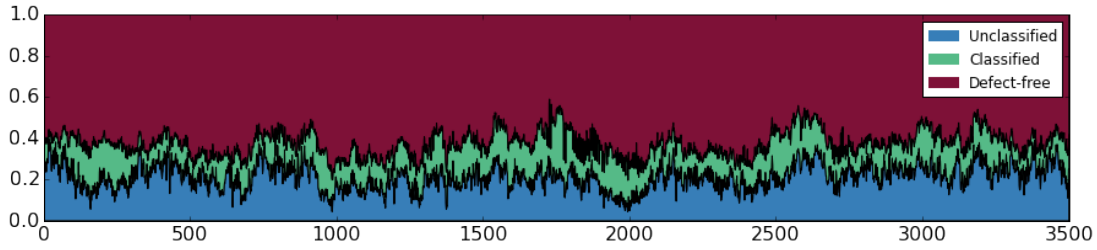
(e)



(f)



(g)



(h)

Figure 6.9 (cont.): Counts of the number of occurrences of each of the six types of canonical defect patterns over 3,500 frames sampled 0.1 time units apart: (a) convex disclination, (b) concave disclination, (c) dislocation pairs, (d) grain boundaries, (e) target, and (f) spiral. The distributions of each of the counts in (a) to (f) are shown as box and whisker plots in (g). The percentage of the temperature field that is unclassified, classified, or considered to be defect-free, is shown in (h).

many different configurations of the temperature field could give rise to a topological defect of the same charge. That is, the invariant is too coarse to use as a more nuanced descriptor of the defect. In computer vision, there are many algorithms that will generate a high-dimensional feature vector to encode the shape of the scalar field about a point in a way that it is invariant under rotations. In general, these types of algorithms have the same basic structure:

- *Keypoint generation.* Keypoints are points of interest in the image. There are many ways to generate keypoints for an image, typically involving gradient methods for identifying corners (e.g. for architectural images). However, the temperature fields we study are smooth, so we use an alternative method for generating keypoints for RBC flows.

- *Orientation assignment.* In order to build rotation invariance into the matching computation, an orientation must be assigned to each keypoint. Occasionally, there may be more than one likely orientation, so the keypoint will be duplicated in this step, one for each dominant orientation.

- *Local feature descriptors.* Instead of using local image patches directly, a high-dimensional local feature descriptor that is rotation invariant is generated for each keypoint. To compare the similarity of the image at two different keypoints, these local feature descriptors are compared rather than the local image patch.

We now describe each of the components above in more detail.

Keypoint Generation

To generate keypoints on the simulated RBC temperature fields, we use topological defects (see Section 6.1.1) and saddle points and roll emergence points as identified by persistent homology (see Section 6.2.1). The amount of smoothing applied to the orientation field and the regions selected in the partition of the persistence plane controls the number of keypoints. One common algorithm from computer vision that is used to identify keypoints in an image is the SIFT algorithm [66], a detailed description of which is beyond the scope of this dissertation. Figure 6.10 shows a comparison of keypoints

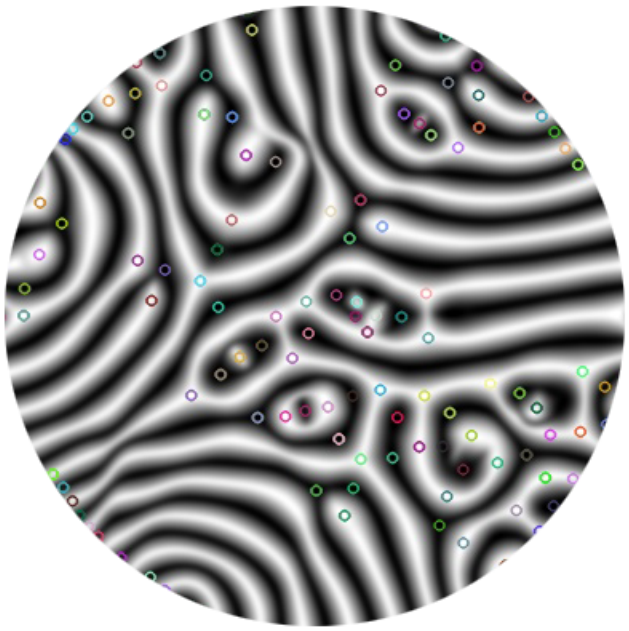
generated from the SIFT algorithm as implemented in the Python library OpenCV versus keypoints given by topological and persistent homology defects. Notice that at approximately the 8 o'clock region on the image, there is a group of saddle points detected by persistence that are not detected by the SIFT algorithm. In addition, the keypoints detected through topological methods seem to have a much more precise placement with respect to the visible defects in the temperature field.

Orientation Assignment

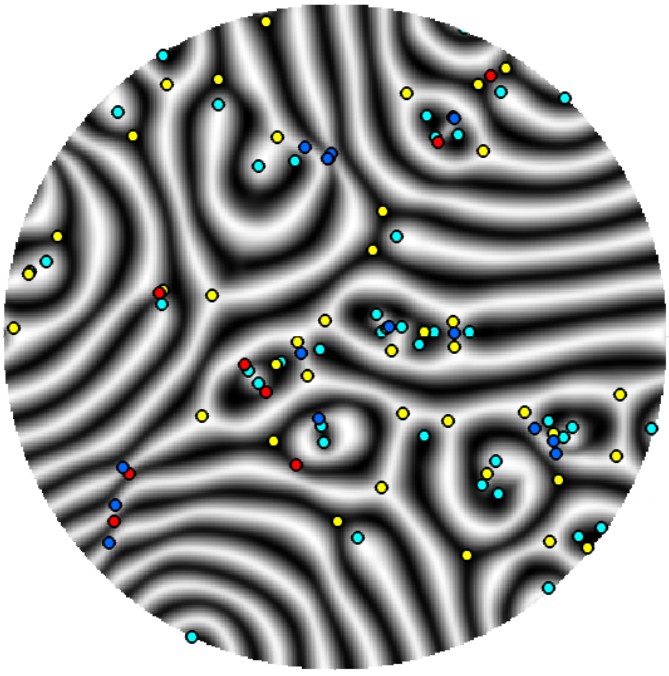
To assign an orientation to each keypoint, we first crop a circular region from the orientation field centered at the keypoint. From this circular region, a histogram of orientations is generated and each dominant mode in the histogram is used to assign an orientation to the keypoint. Notice that if an orientation field has a concentration of orientations near $\pi/2$, the corresponding histogram is likely to have values near $-\pi/2$ as well. Thus, we view the data as a histogram on a circle, where two modes at $\pi/2$ and $-\pi/2$ would be considered a single mode. We then compute the maximum value of the histogram, and then any peaks that are within some percentage of this value are used to generate additional keypoints, each designated by its corresponding maximum orientation. By adding additional keypoints for each peak in the orientation histogram, the local image feature will be more robust with respect to rotation. For example, consider a local image patch that is bimodal in orientation, with the peak orientations nearly the same. A small perturbation in the orientation field may switch the peak which represents the maximum, so if each keypoint is only assigned a single orientation, small perturbations may significantly affect the rotation-invariance. By adding a keypoint for each peak, even a small perturbation such as this will result in a positive match. Thus, orientation assignment involves the choice of two parameters: the radius of the circular image patch to use to build the histogram of orientations, and the choice of percentage to use to select secondary peaks from the histogram.

Local Feature Descriptors

For each keypoint resulting from the previous computation, a local rotation-invariant



(a)



(b)

Figure 6.10: Keypoints generated by (a) Python OpenCV SIFT implementation (colorings random), and (b) topological defects (yellow and cyan) with $\sigma = 2$ as the orientation smoothing parameter and persistent homology defects (red and dark blue) given by the partition in Figure 6.7.

feature is produced. This is again accomplished by first cropping a circular region of the orientation field centered at the keypoint location with a given radius. Then the cropped orientation field is rotated by the orientation value assigned to the keypoint orientation, so that the orientation field is centered consistently across the dataset (e.g., at zero). Then the circular region is divided into eight regions according to an inner radius width. Histograms of orientations with a common number of bins are generated for each of eight regions from this local orientation patch. So that orientations farthest away from the keypoint are not weighted as heavily as orientations close to the keypoint, a Gaussian weight is applied to the local orientation field when computing the histograms. Thus, the local feature descriptor will be $8 \times b$, where b is the number of possible bins for the orientation histograms in each of eight local regions. There are a number of parameters which must be set for this step: outer radius, inner radius, Gaussian weight parameters, and number of binning values for the histograms. See Figure 6.11 for an illustration of the regions used and the local histograms produced for a sample keypoint.

6.4.1 Parameter Selection for Local Feature Descriptors

As discussed above, there are six parameter values which must be set to generate the local feature descriptors generated at each keypoint (including the keypoint orientation parameters). Ideally, we would construct a cost function and find the parameter settings which resulted in the lowest cost, testing each possible parameter setting. However, choosing only 10 possible values for each parameter results in 10^6 computations, and more if additional parameter values are searched, making this approach infeasible. Thus, instead of doing an exhaustive search through parameter space, we use a coordinate descent optimization method [67]. To test the quality of the local feature descriptor and construct the cost function for choosing the parameters, we work with the following hypothesis:

The temperature field for RBC varies slowly and continuously in time. For close enough time points, the local feature descriptors at a given keypoint should be matched to the "same" keypoint in the next frame, where the spatial location of that keypoint may have moved somewhat due to the change in pattern. By computing the percentage

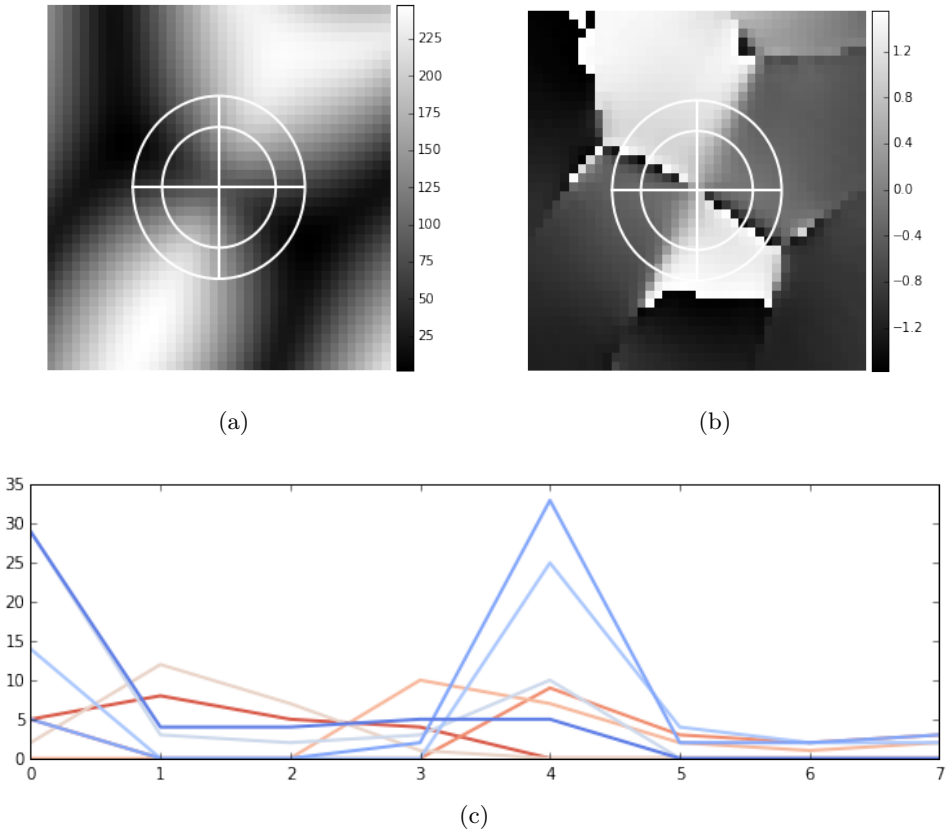


Figure 6.11: (a) The circular region used to generate the keypoint descriptor. (b) The associated orientation field. (c) The histogram of orientations for each of the eight local regions. Histograms go from dark orange to dark blue, inner circle to outer circle, top to bottom, left to right.

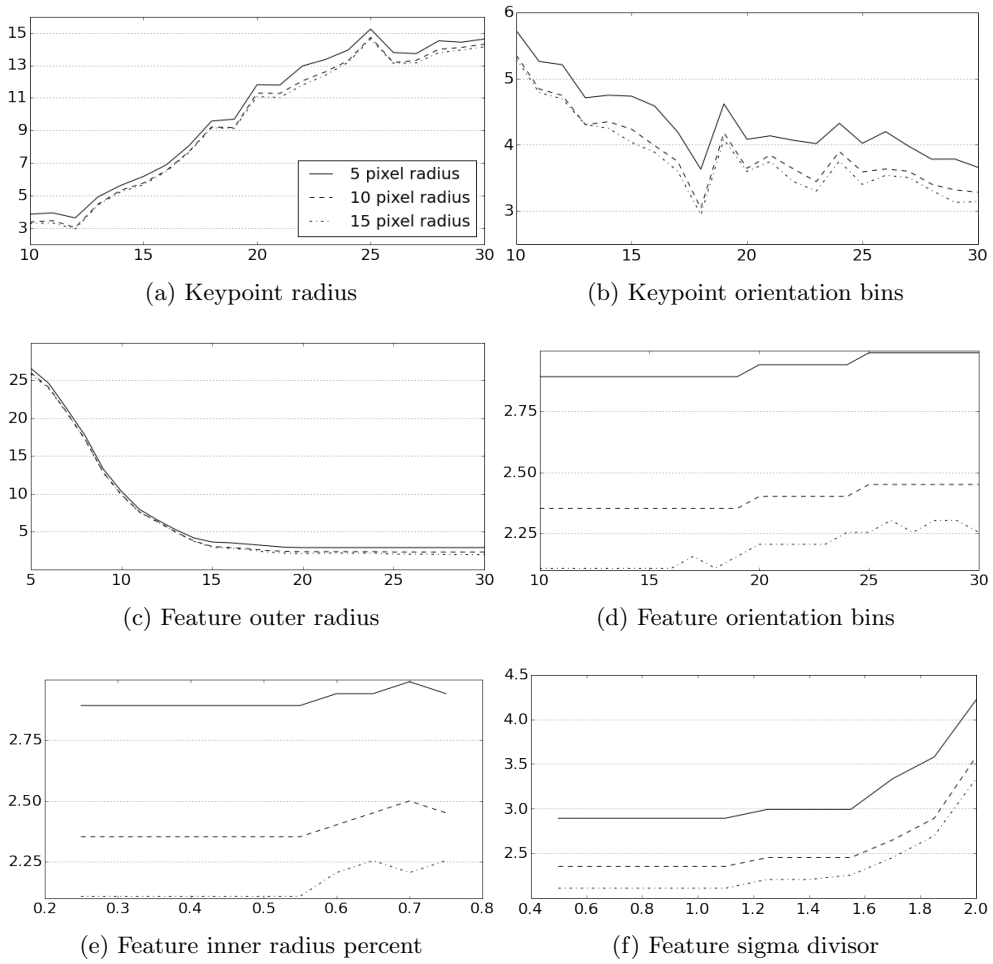
of keypoints that were not matched to nearby keypoints in the next sample frame, we obtain a function whose local minima will give local optimizations of the parameter selections used to generate the local keypoint descriptors.

The one parameter which does not fit this general idea is the secondary peak selector. By lowering the percentage parameter, more and more peaks will be selected, adding additional keypoints to the dataset. If that percentage is at 100%, then only one peak will be used, and, theoretically, lowering the percentage will decrease the percentage that aren't matched correctly. However, lowering the percentage too far will cause the number of added keypoints that have a good match to overwhelm the cost function computation, and so doing a blind optimization for this value is not such a good idea. Thus, we use a value of 80% as in [66].

Coordinate Descent Optimization of Local Feature Parameters

We construct the optimization function for the coordinate descent algorithm as follows. Our full dataset consists of $5,000$ frames sampled 0.01 time units apart (note that this is ten times faster than the dataset analyzed in Section 6.3.3). For frames separated by five sample points, we compute the nearest neighbor keypoint based on a local feature match and then compute the Euclidean distance between the corresponding matched keypoints on the temperature field. We do this for frames 1000 to 1950 in steps of 50. Thus, we are analyzing the keypoints from 20 different temperature fields, and each analysis generates roughly 1700 to 2100 keypoints (the initial settings on the radius and binning levels for the keypoints will affect this number). We compute the percentage of matched keypoints which fall outside of a circular region of radius of 5, 10, and 15 pixels from the original keypoint. The idea being that a more robust parameter setting will result in fewer percentage of keypoints that fall outside of each of these regions.

We initialize the coordinate descent algorithm with the set of values given in the first column of Table 6.1. The second and third columns give the values used in successive iterations. Figure 6.12 gives the plots of the loss function for each coordinate in the first iteration of the algorithm for the first set of initial conditions. Each test iteration



(a) Keypoint radius

(b) Keypoint orientation bins

(c) Feature outer radius

(d) Feature orientation bins

(e) Feature inner radius percent

(f) Feature sigma divisor

Figure 6.12: Plots of the loss function as a function of the parameter setting, keeping all other parameters fixed. The y-axis gives the percentage of keypoint matches that fall outside of a circle of radius r about the originating keypoint.

Parameter	Initialization	Iteration 1	Iteration 2
Keypoint radius	*	12	12
Keypoint orientation bins	18	18	23
Feature radius	15	20	20
Feature orientation bins	18	10	10
Feature inner radius percent	50%	50%	30%
Feature sigma divisor	1	1	0.8

Table 6.1: Three iterations of the coordinate descent optimization.

Parameter	Initialization	Iteration 1	Iteration 2
Keypoint radius	*	5	11
Keypoint orientation bins	10	19	19
Feature radius	5	20	20
Feature orientation bins	10	12	12
Feature inner radius percent	20%	50%	50%
Feature sigma divisor	0.5	1	1

Table 6.2: Three iterations of the coordinate descent optimization for a different set of parameter initializations and a different underlying dataset.

takes on average $75 - 400$ seconds, depending on the parameter settings. After just a few iterations, the parameters stabilize such that approximately 97% of keypoints are matched to within five pixels of their original location when compared to the frame five sample points later.

Note that we are locating a particular local minimum of the loss function, so we begin the process again with a different set of initial parameters and a different input dataset to test for mode sensitivity. This time we choose frames 1025 to 1975, again in steps of 50, yielding again 20 frames of images for the analysis. Again, the coordinate descent algorithm stabilizes after just a few iterations to similar values as we found in the first round, as shown in Table 6.2. By beginning the process with a different set of initial parameters and by varying the input dataset, we can be relatively confident that the parameters signify a robust local minimum of the system with respect to the chosen loss function.

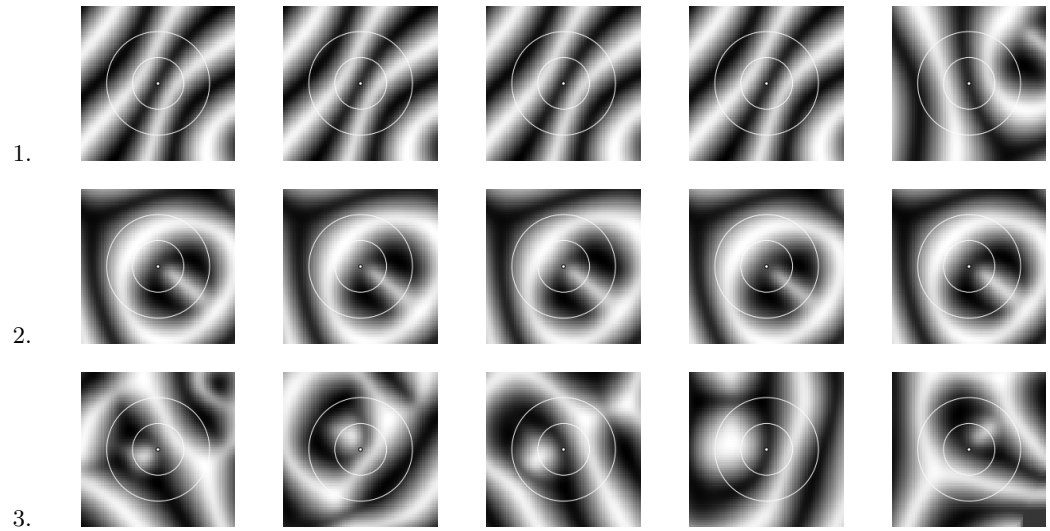


Figure 6.13: Sample temperature fields centered at keypoints for three selected clusters for the sublevel pd_0 defect types.

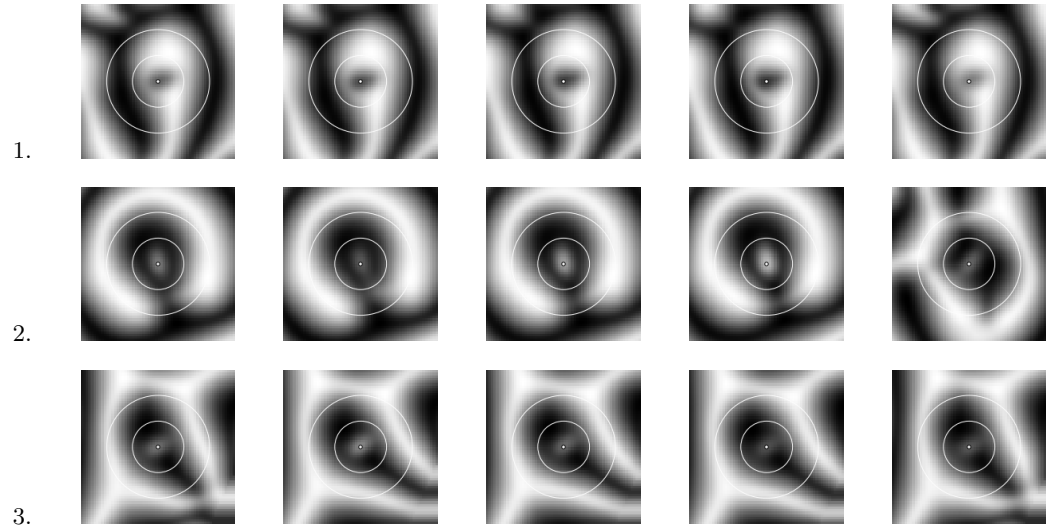


Figure 6.14: Sample temperature fields centered at keypoints for three selected clusters for the sublevel pd_1 defect types.

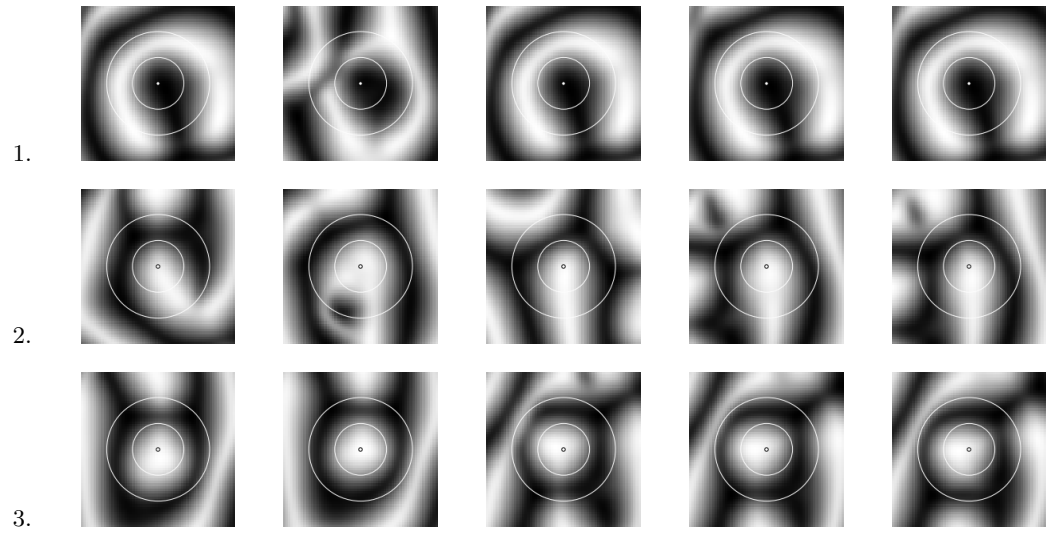


Figure 6.15: Sample temperature fields centered at keypoints for three selected clusters for the td_{+1} defect types.

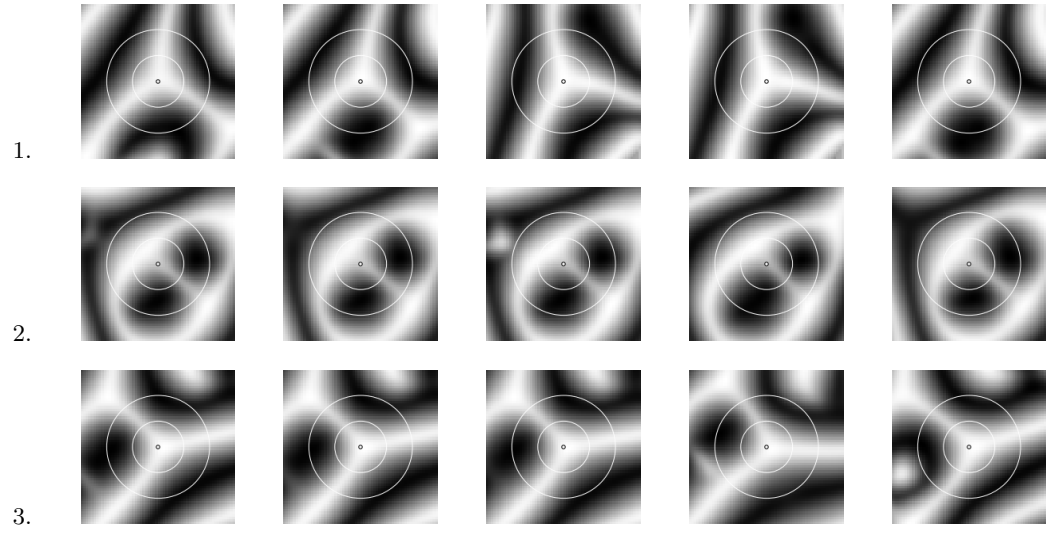


Figure 6.16: Sample temperature fields centered at keypoints for three selected clusters for the td_{-1} defect types.

6.4.2 Visualizing Defects in Feature Space

To visualize the defects in the constructed local feature space, we use an implementation of a Gaussian mixture model as described in [68] to cluster the data into groups, and then view defects from each group that are close to the means of the clusters found by the output of the algorithm. Due to the covariant nature of time series data (data points are sampled at 0.01 time units apart), rather than use every solution from the 1000 time point series, we subsample the the time series and then compute keypoints and local feature vectors for the reduced dataset. Since they are more numerous, we subsample every 25^{th} time point for persistent homology dimension 0 and topological defect keypoints, while we sample every 5^{th} time point for persistent homology dimension 1 keypoints. The number of keypoints we use to for visualization are $pd_0 = 651$, $pd_1 = 199$, $td_{+1} = 1,190$, and $td_{-1} = 1,661$.

Using only the keypoint descriptors, we perform clustering using mean-field variational inference for a mixture of Gaussians, using coordinate ascent (CAVI) as the optimization algorithm, as in [68]. We place a Gaussian prior on the cluster means with mean given by the center of the point cloud of keypoint features, and unit variance (i.e. we assume that the cluster means originate from a Gaussian distribution with mean given by the center of the point cloud and unit variance). We also assume a uniform categorical prior on the cluster assignments.

We first separate each keypoint by type (e.g., topological defect with charge ± 1, persistent homology defect dimension 0 or 1) and then cluster the data into 20 subgroups within each of these four defect types. Figures 6.13 through 6.16 show the cropped temperature fields for five patterns from three of the 20 clusters, where the keypoint features of the local temperature fields shown are closest to the means of the respective clusters. We also mark the keypoint in white and give circles for the inner and outer radii used to build the local feature vectors.

6.5 Towards an Analysis of Coherent Structures in SDC

In general terms, a *coherent structure* is a localized dynamical structure that persists beyond the time-scale of the pattern evolution of the system, and is repeatedly observed in the system given enough time [69]. To make sense of a data-driven analysis of coherent structures, we must establish a method by which groups of defects can be treated together as a single pattern, and to be able to search the temperature fields for the recurrence of such a pattern. This matching must be rotation-invariant and robust to small perturbations, since an exact match in a system driven far from equilibrium is highly unlikely.

In this section we leverage established techniques for object recognition for searching an image for the presence of a training pattern (e.g., the core of a spiral or other localized defect structure). These algorithms often give a matching that is invariant under rotations and certain affine transformations of the image, which is a desirable property to preserve for pattern recognition in this setting. In the SIFT and SURF algorithms [66, 70], keypoints and their local feature vectors are used for object recognition using a method called Hough transform voting. We illustrate its use here in the context of identifying mid-scale coherent structures in RBC. That is, structures that repeat on a scale larger than the defects identified above, but only up to the size of, say, two or three wavelengths in the pattern.

6.5.1 The Hough Transform

We summarize the method of Hough transform voting from [70], modified so that it works with the framework we have already established (in particular, we disregard scale). The general idea is to define a procedure by which a given arrangement of keypoints (and their local feature vectors) may be located in a target image and its keypoints (and their local feature vectors). This is done by accumulating votes in what is commonly referred to as *Hough space*, in this case a three-dimensional lattice space consisting of a vector (x, y, θ), where (x, y) is a vote for the location of the localized pattern, discretized by bins, and θ is a vote for the orientation, also discretized by bins.

The coordinate corresponding to the highest value in the Hough space accumulator gives the position and orientation considered to be those corresponding to the candidate match. Finally, the collection of keypoints that voted for this coordinate are considered the keypoints of the matched pattern. We now describe this procedure more precisely.

Let p_0 be a location in the training image marking the centroid of the pattern being searched for, and let K_0 be the keypoints in an ε neighborhood of p_0 along with their local feature vectors (see Figure 6.17(a) for the set K_0). Let $c_0(x, y)$ be the location of a keypoint $c_0 \in K_0$ being searched (each image in the first column of Figure 6.18 represents a keypoint c_0), and let $c_0(\theta)$ be its orientation. Compute

$$v' = (c_0(x, y) - p_0),$$

the displacement of the keypoint relative to the pattern's centroid. Take the set $K(c_0)$, which we will define below (see Choosing keypoints for Hough transform voting), to be the keypoints of the target image that are candidate matches to the keypoint c_0 (each image in columns 2+ of Figure 6.18, with distance *not* in parenthesis, represents a keypoint $c \in K(c_0)$).

The following procedure is used to estimate the relative pattern centroid in the target image corresponding to a candidate matched keypoint $c \in K(c_0)$. Let $c(x, y)$ be the keypoint's location in the training image and let $c(\theta)$ be its orientation. Compute the difference in orientations between c_0 and c by taking $\alpha = |c_0(\theta) - c(\theta)|$ in radians, and let v be the vector v' rotated by α radians. We estimate the corresponding centroid of the pattern in the target image by computing

$$p = v + c(x, y).$$

Thus, we have computed the rotation α and center p of the target pattern if the correspondence between the keypoint c_0 and c were actually correct. We then compute which integer-valued bin α and p belong to, where the bins correspond to a discretization of the possible locations and orientations, and increment the Hough space accumulator at this coordinate. As in [70], we also increment neighboring coordinates in Hough

space to avoid inaccuracies stemming from discretization of the space and improving the number of matches (for example, two keypoints could be mapped to Hough space coordinates that are off by a single bin due to rounding).

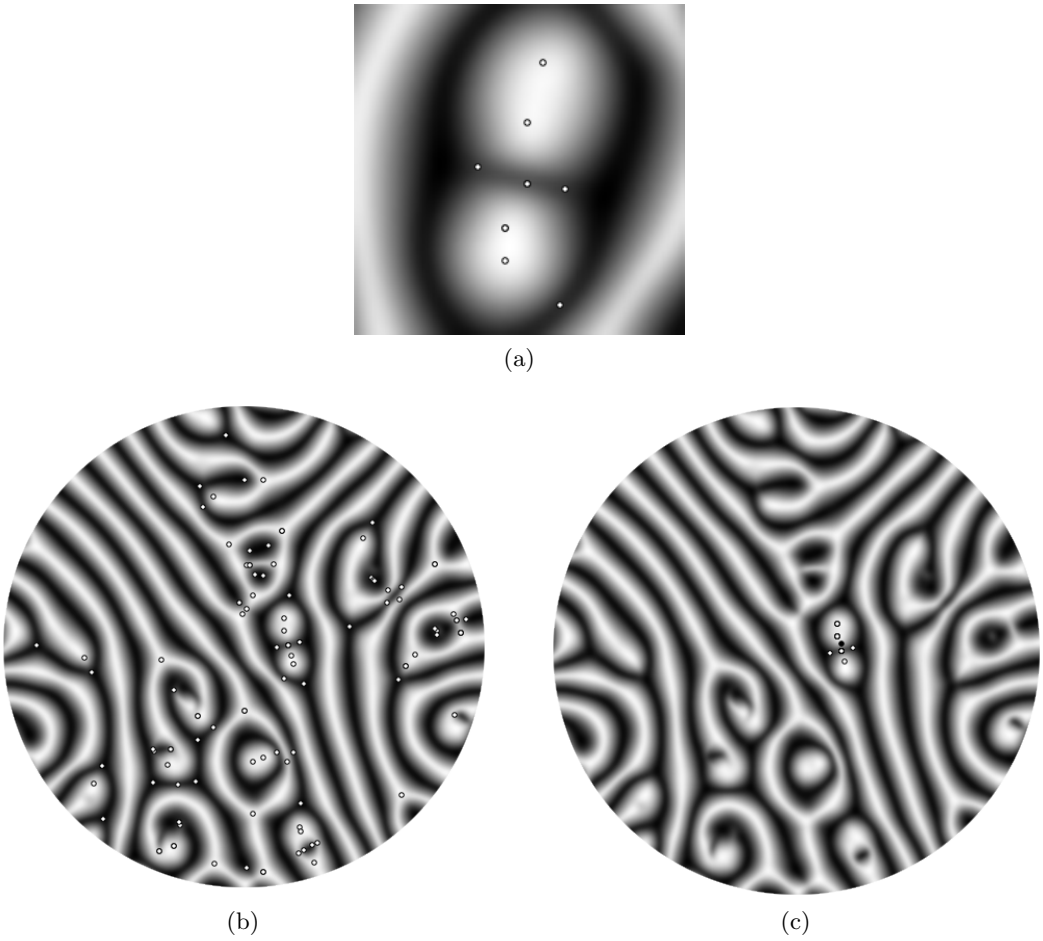
Once all votes have been cast by keypoints $c \in K(c_0)$, the coordinate in the accumulator with the maximum value is considered the winning configuration. We then include any configurations of unit sup norm distance to the winning configuration to further reduce eliminating good keypoint matches due to discretization and rounding. Any keypoints in $K(c_0)$ that cast a vote for one of these lattice points is considered a matched keypoint for the pattern. We also check to make sure more than $r|K_0|$ keypoints are matched from the training image, where r is a rate or percentage. If enough keypoints from the training pattern are matched, then we consider the pattern to have a positive match. Finally, we give a fitness measurement for the pattern by computing the median keypoint distance in local feature vector space for the corresponding matched keypoints.

Choosing keypoints for Hough transform voting

Let $c_0 \in K_0$ be a keypoint in the training image. To select the candidate keypoint matches $K(c_0)$, the entire collection of keypoints K of the target image is compared against the keypoint c_0 using the Euclidean distance of the corresponding local feature vectors. In the case of a perfect match, the idea would be that the nearest neighbor would be considered the candidate that gets to add a vote to the Hough space accumulator. However, there could be many local patterns that are fairly good matches. To resolve this problem, we choose a threshold cutoff C and consider any keypoint $c \in K$ such that if $d_0 = \min_{x \in K} d(c_0, x)$, then $d(c_0, c) \leq d_0 + C$. Thus, we define

$$K(c_0) := \{c \in K : d(c_0, c) \leq d_0 + C\}$$

for some cutoff distance C, where d_0 is the distance from c_0 to the set K as measured in local feature space. While there are certainly more sophisticated methods for choosing candidate keypoints, we use this simplified method merely as an illustration for



(a) (b) (c)

Figure 6.17: (a) A cropped portion of the temperature field at frame $1,530$ with keypoints overlaid. (b) The temperature field at frame $1,100$ with keypoints overlaid. (c) The matched keypoints after performing the Hough transform voting.

how these methods may be effective for the problem at hand conceptually. A more rigorous analysis of each of the aspects of this method should be employed for performing a statistical analysis concerning the study of coherent structures (for example, see Section 6.4.1 where optimization was used to choose parameter values).

Sample pattern match using Hough transform voting

We now illustrate the above ideas with an example. Figure 6.17(a) shows a cropped portion of a temperature field at frame number $1,530$, which we will refer to as the *training pattern*, while Figure 6.17(b) shows the temperature field at frame number $1,100$, referred to as the *target pattern*. Each image has its keypoints overlaid. The training pattern has ten keypoints including multiple keypoints for multi-modal orientations.

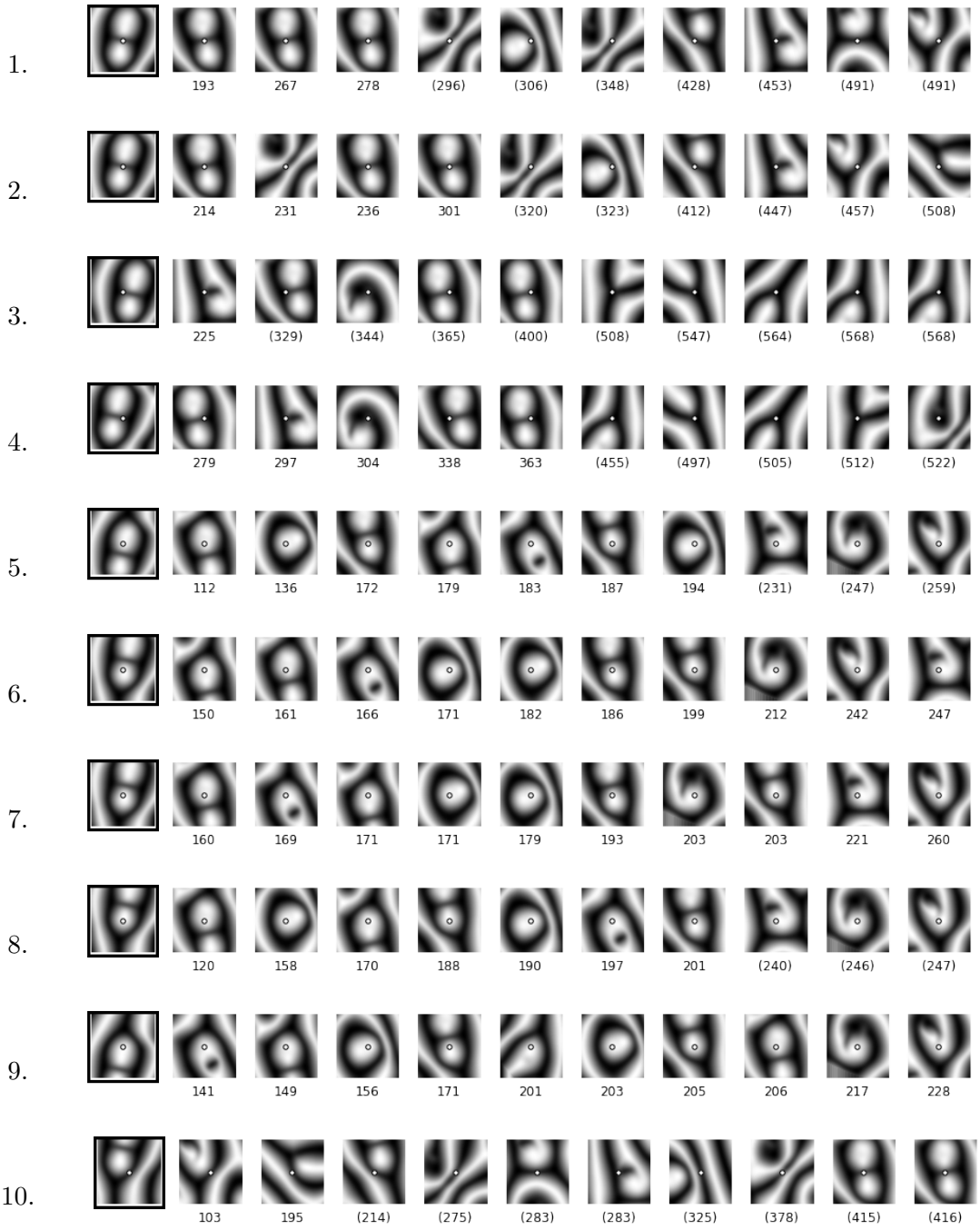


Figure 6.18: Nearest neighbor search results for each of the keypoints in the training pattern. Each row is a different keypoint in the training pattern. The first frame in each list of images is the local feature being searched (i.e. $c_0 \in K_0$). The frames that follow are the ten nearest neighbor matches in local feature space, where each keypoint has the same type (e.g. topological defect ± 1, or persistent homology dimension 0 or 1). The distance in local feature space is printed beneath each image, listed in parenthesis if the distance is beyond the cutoff. Thus, matches with distance not in parenthesis are in the set $K(c_0)$.

We set the parameters of the Hough transform voting algorithm to use 40 bins for the location coordinates and 8 bins for the rotation coordinate. The cutoff distance C for qualifying matches is set to 100. The rate r is set to 0.75, so that 75% of keypoints in each localized region must be matched in order to consider the match successful. Figure 6.18 shows, for each keypoint in the training image, the ten nearest neighbor matches in local feature space together with their distances. If the distance for a match is greater than the minimum neighbor match plus $C = 100$, it is listed in parenthesis and was not used for Hough transform voting.

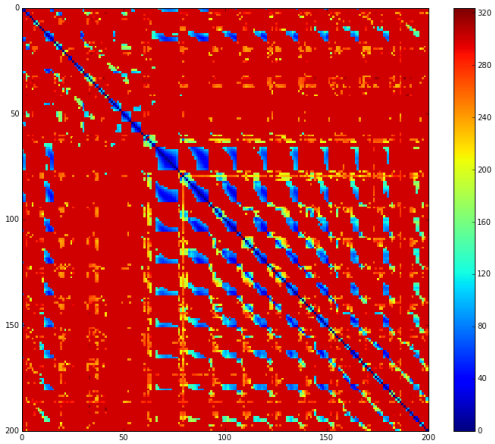
Figure 6.17(c) shows the pattern with the highest number of votes from the Hough transform. Matched keypoints are indicated in white, and the black circle is the estimated centroid. As is evident in the figure, the training pattern was correctly located in the target pattern, and the computed centroid of the matched pattern falls (roughly) near the corresponding center of the training pattern.

6.5.2 Tracking a Spiral Pattern Using Hough Transform Matching

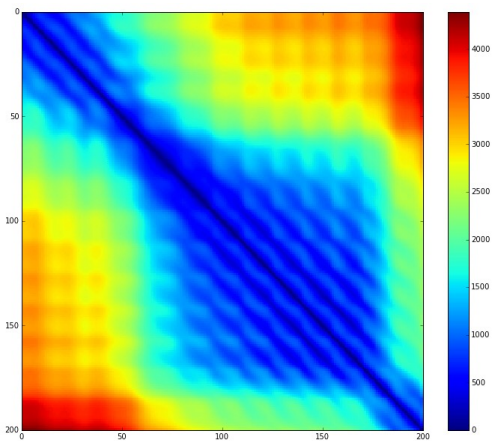
We now demonstrate the efficacy of this type of approach for searching a time series for recurrent patterns. The time series consists of $5,000$ frames sampled 0.01 time units apart (this is ten times faster than the dataset analyzed in Section 6.3.3). We perform Hough transform voting, again using 40 bins for the location coordinates and 8 bins for the rotation coordinate. We use a radius of 30 for both the training pattern and for the matched patterns. The cutoff distance C for qualifying matches is again set to 100. To achieve a more accurate matching, the rate r is set to 0.9, so that 90% of keypoints in each localized region must be matched in order to consider the match successful.

We first localize to a region of the temperature field a depicted in Figure 6.19(a) and track the evolution of center of a spiral pattern by using the fitness measurement as a distance between patterns. Figure 6.19(b) shows the distance matrix computed from the Hough transform fitness measurement for this region of the temperature field, compared to Figure 6.19(c), which shows the usual L^2 distance between cropped portions of the temperature field. The recurrent dynamics are evident in the Hough transform fitness matrix, as seen in the checkerboard type pattern.

(a)



(b)



(c)

Figure 6.19: (a) Temperature field showing cropped region (white square) used to localize a time-evolving spiral pattern. (b-c) Distance matrices of a time-evolving spiral from frames 1000 to 1999, selected every fifth frame to generate 200 patterns. (b) Hough transform voting fitness measurement. (c) L^2 distance between the cropped temperature fields.

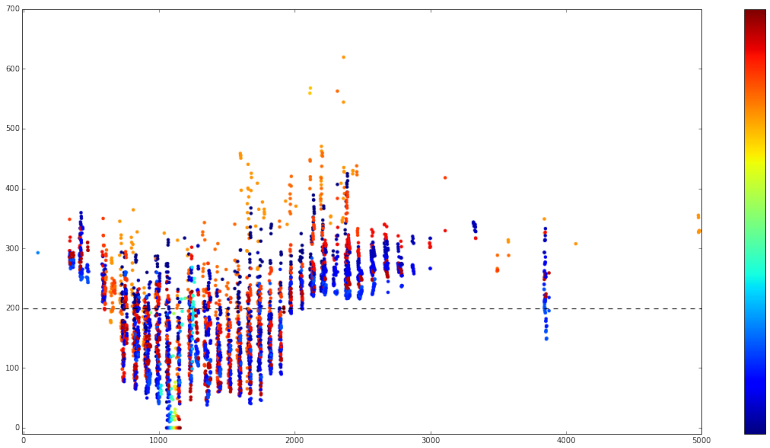
Next, we perform an ensemble pattern match by localizing the spiral pattern depicted in Figure 6.19(a) from frame 1055 to frame 1155, sampling at intervals of 5 frames, resulting in 21 sample patterns. Each of these 21 sample patterns are matched against frames 1 to 4999 in the complete time series. Any positive matches have their Hough transform fitness measurement recorded. Figure 6.20(a) shows the results of the matching. As the picture shows, this sample ensemble pattern is representative of the spiral from frames 750 to 2000, approximately, and a signature is somewhat detected from frame 2000 to 3000.

We repeat the experiment, localizing over the "same" spiral pattern as seen by the author in direct observation of the temperature field, but from different frames. Figure 6.20(b) shows the result of the matching from sampling frames 2850 to 2950 at intervals of 5 frames, and Figure 6.20(c) shows the result of the matching from sampling frames 1970 to 2070 at intervals of 5 frames. We will call the spiral pattern states A, B, and C for the patterns sampled in Figure 6.20(a), (b), and (c), respectively. The matching signatures seem to indicate that the pattern begins in state B from roughly frames 350 to 600, then transitions to state A from frames 750 to 2000, at which point it enters a transient state passing through C. At frame roughly 2500, it again passes through state B before there are no more positive matches at frame 3500.

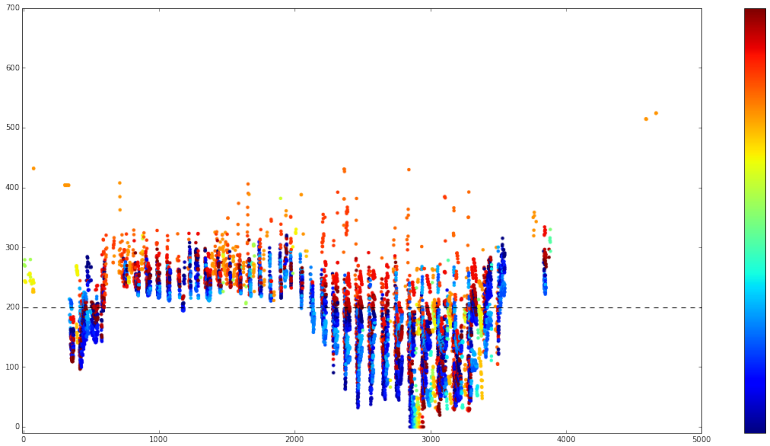
Figure 6.21 shows cropped regions of the temperature field, where each cropping window is the same absolute size, centered at the spiral structure being studied. Sample (a) was taken before the spiral formed, and sample (b) marks the emergence of the spiral (state B). Samples (c-e) show the transitions from state A to C to B, in that order. Note that state C is significantly pinched and elongated compared to states A and B, and the spiral in state A is slightly smaller than the spiral structure in state B. Sample (f) gives match at the tail-end of the spiral's lifetime, and sample (g) is taken after the spiral signature is no longer visible in the Hough transform matching diagrams.

6.5.3 Conclusions and Future Work

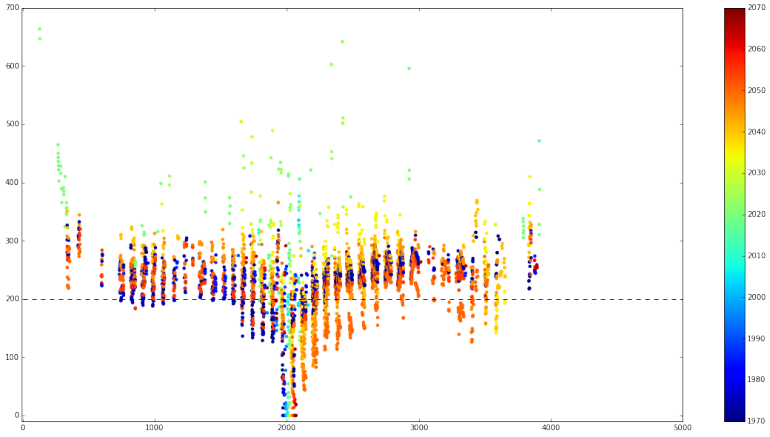
The study carried out in this section only scratches the surface at what might be possible given these techniques, as well as points to current limitations in the methodology.



(a)

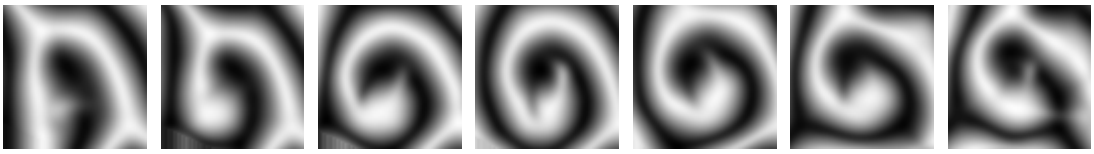


(b)



(c)

Figure 6.20: Ensemble pattern match fitness measurement plots for the spiral pattern from Figure 6.19(a). (a) Samples from frame 1055 to frame 1155, (b) samples from frame 2850 to 2950, (c) samples from frame 1970 to 2070, each sampled at regular intervals of five frames, resulting in 21 patterns. Colorbars are colored from 1 to 21, showing the sample index being matched. The horizontal dashed line at 200 gives a heuristic for "good" matches versus spurious matches as observed by the author.



(a) (b) (c) (d) (e) (f) (g)

Figure 6.21: Spiral pattern sampled from frames (a) 300 (b) 450 (c) 1200 (d) 2000 (e) 2800 (f) 3700 (g) 3900.

Figures 6.20 and 6.21, taken together, indicate that the current pattern matching setup is too fine. That is, the spiral shape must match the pattern being searched almost exactly for it to register. Our analysis only has rotation-invariance built into the pattern matching algorithms. However, the SIFT algorithm [66] also includes a notion of scale-invariance that could be leveraged. By adding a certain amount of scale-invariance to the computations, we suspect that an ensemble pattern match such as the one performed will result in more positive matches.

Once scale-invariance has been addressed, a more systematic analysis of different recurrent structures could be undertaken. For example, portions of the temperature field for a given portion of the time series could be either systematically or randomly sampled, and the remainder of the time series could be searched for the recurrence of the sampled patterns. Recurrent patterns that are statistically significant as measured by the Hough transform fitness measurement could then be classified, perhaps through clustering methods that use the Hough transform fitness measurement itself (in the spirit of the distance matrix in Figure 6.19(a)).

These methods have only been thus far applied to numerically-simulated data. A comparison to experimental data could also be beneficial. Finally, it would be prudent to study the connection between patterns in the temperature field and the leading-order Lyapunov vectors for the numerically-simulated data, which may pave the way to establishing predictive analytics for these types of fluid flows.

References

[1] M. Kramár, R. Levanger, J. Tithof, B. Suri, M. Xu, M. Paul, M. F. Schatz, and K. Mischaikow, "Analysis of Kolmogorov flow and Rayleigh-Bénard convection using persistent homology," *Physica D: Nonlinear Phenomena*, vol. 334, pp. 82 – 98, 2016. Topology in Dynamics, Differential Equations, and Data.

[2] R. Ghrist, "Barcodes: The persistent topology of data," *Bulletin of the American Mathematical Society*, vol. 45, pp. 61–76, oct 2007.

[3] G. Carlsson, "Topology and data," *Bulletin of the American Mathematical Society*, vol. 46, pp. 255–308, Jan. 2009.

[4] H. Edelsbrunner and J. L. Harer, *Computational topology : an introduction*. Providence (R.I.): American Mathematical Society, 2010.

[5] A. Zomorodian and G. Carlsson, "Computing persistent homology," *Discrete & Computational Geometry*, vol. 33, no. 2, pp. 249–274, 2004.

[6] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer, "Stability of persistence diagrams," *Discrete Comput. Geom.*, vol. 37, pp. 103–120, Jan. 2007.

[7] F. Chazal, D. Cohen-Steiner, M. Glisse, L. J. Guibas, and S. Y. Oudot, "Proximity of persistence modules and their diagrams," in *Proceedings of the Twenty-fifth Annual Symposium on Computational Geometry*, SCG '09, (New York, NY, USA), pp. 237–246, ACM, 2009.

[8] M. Botnan and G. Spreemann, "Approximating persistent homology in euclidean space through collapses," *Applicable Algebra in Engineering, Communication and Computing*, vol. 26, no. 1-2, pp. 73–101, 2015.

[9] T. K. Dey, F. Fan, and Y. Wang, "Graph induced complex on point data," in *Proceedings of the Twenty-ninth Annual Symposium on Computational Geometry*, SoCG '13, (New York, NY, USA), pp. 107–116, ACM, 2013.

[10] D. Sheehy, "Linear-size approximations to the vietorisrips filtration," *Discrete and Computational Geometry*, vol. 49, no. 4, pp. 778–796, 2013.

[11] U. Bauer and M. Lesnick, "Induced matchings of barcodes and the algebraic stability of persistence.," *Proceedings of the Thirtieth Annual Symposium Computational Geometry*, p. 355, 2014.

[12] P. Bubenik, J. Scott, and V. De Silva, "Metrics for generalized persistence modules.," *Foundations of Computational Mathematics*, p. 31p., 2014.

[13] P. Bendich, H. Edelsbrunner, M. Kerber, and A. Patel, *Mathematical Foundations of Computer Science 2010: 35th International Symposium, MFCS 2010, Brno, Czech Republic, August 23-27, 2010. Proceedings*, ch. Persistent Homology under Non-uniform Error, pp. 12–23. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010.

[14] G. Carlsson, "Topological pattern recognition for point cloud data," *Acta Numerica*, vol. 23, pp. 289–368, 005 2014.

[15] "tda-persistence-explorer." `http://www.github.com/rachellevanger/tda-persistence-explorer`, April 2017.

[16] F. Chazal, V. de Silva, M. Glisse, and S. Oudot, "The structure and stability of persistence modules," 2012.

[17] W. Crawley-Boevey, "Decomposition of pointwise finite-dimensional persistence modules," *Journal of Algebra and Its Applications*, vol. 14, no. 05, p. 1550066, 2015.

[18] T. Kaczynski, K. Mischaikow, and M. Mrozek, *Computational Homology*, vol. 157 of *Applied Mathematical Sciences*. New York: Springer-Verlag, 2004.

[19] C. Weibel, *An Introduction to Homological Algebra*. Cambridge Studies in Advanced Mathematics, Cambridge University Press, 1995.

[20] A. Beygelzimer, S. Kakade, and J. Langford, "Cover trees for nearest neighbor," in *Proceedings of the 23rd International Conference on Machine Learning*, ICML '06, (New York, NY, USA), pp. 97–104, ACM, 2006.

[21] T. K. Dey, F. Fan, and Y. Wang, "Computing topological persistence for simplicial maps," in *Proceedings of the Thirtieth Annual Symposium on Computational Geometry*, SOCG'14, (New York, NY, USA), pp. 345:345–345:354, ACM, 2014.

[22] M. Buchet, F. Chazal, S. Y. Oudot, and D. R. Sheehy, "Efficient and robust persistent homology for measures," in *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '15, pp. 168–180, SIAM, 2015.

[23] P. Niyogi, S. Smale, and S. Weinberger, "Finding the homology of submanifolds with high confidence fromrandomsamples," *Discrete & Computational Geometry*, vol. 39, no. 1, pp. 419–441, 2008.

[24] M. Botnan, J. Curry, and E. Munch, "The poset interleaving distance." Preprint, 2016.

[25] V. de Silva, E. Munch, and A. Patel, "Categorified reeb graphs," *Discrete & Computational Geometry*, vol. 55, no. 4, pp. 854–906, 2016.

[26] D. Morozov, K. Beketayev, and G. Weber, "Interleaving distance between merge trees." Manuscript, 2013.

[27] M. Lesnick, "The theory of the interleaving distance on multidimensional persistence modules," *Foundations of Computational Mathematics*, vol. 15, no. 3, pp. 613–650, 2015.

[28] A. Blumberg and M. Lesnick, "Universality of the homotopy interleaving distance." Joint Mathematics Meetings 2017, Abstract 1125-55-1597, sep 2016.

[29] M. Gameiro, K. Mischaikow, and W. Kalies, "Topological characterization of spatial-temporal chaos," *Phys. Rev. E*, vol. 70, p. 035203, Sep 2004.

[30] M. Gameiro, K. Mischaikow, and T. Wanner, "Evolution of pattern complexity in the Cahn-Hilliard theory of phase separation," *Acta Materialia*, vol. 53, no. 3, pp. 693 – 704, 2005.

[31] M. Kramr, A. Goullet, L. Kondic, and K. Mischaikow, "Quantifying force networks in particulate systems," *Physica D: Nonlinear Phenomena*, vol. 283, pp. 37 – 55, 2014.

[32] K. Krishan, H. Kurtuldu, M. F. Schatz, M. Gameiro, K. Mischaikow, and S. Madruga, "Homology and symmetry breaking in Rayleigh-Bénard convection: Experiments and simulations," *Physics of Fluids*, vol. 19, pp. 117105–117105, Nov. 2007.

[33] H. Kurtuldu, K. Mischaikow, and M. F. Schatz, "Measuring the departures from the Boussinesq approximation in Rayleigh-Bénard convection experiments," *Journal of Fluid Mechanics*, vol. 682, pp. 543–557, 009 2011.

[34] H. Kurtuldu, K. Mischaikow, and M. F. Schatz, "Extensive scaling from computational homology and Karhunen-Loève decomposition analysis of Rayleigh-Bènard convection experiments," *Phys. Rev. Lett.*, vol. 107, p. 034503, Jul 2011.

[35] P. Cvitanovic, D. Borrero-Echeverry, K. M. Carroll, B. Robbins, and E. Siminos, "Cartography of high-dimensional flows: A visual guide to sections and slices," *Chaos*, vol. 22, p. 047506, Dec. 2012.

[36] G. Kawahara, M. Uhlmann, and L. van Veen, "The significance of simple invariant solutions in turbulent flows," *Annual Review of Fluid Mechanics*, vol. 44, no. 1, pp. 203–225, 2012.

[37] A. Adcock, E. Carlsson, and G. Carlsson, "The ring of algebraic functions on persistence bar codes," 2013.

[38] V. I. Arnold and L. D. Meshalkin, "Seminar led by A. N. Kolmogorov on selected problems of analysis (1958-1959)," *Usp. Mat. Nauk*, vol. 15, no. 247, pp. 20–24, 1960.

[39] B. Suri, J. Tithof, R. Mitchell, R. O. Grigoriev, and M. F. Schatz, "Velocity profile in a two-layer Kolmogorov-like flow," *Phys. Fluids*, vol. 26, 2014.

[40] R. L. Panton, *Incompressible flow*. John Wiley & Sons, 2006.

[41] R. Mitchell, *Transition to turbulence and mixing in a quasi-two-dimensional Lorentz force-driven Kolmogorov flow*. PhD thesis, Georgia Institute of Technology, 2013.

[42] G. J. Chandler and R. R. Kerswell, "Invariant recurrent solutions embedded in a turbulent two-dimensional Kolmogorov flow," *Journal of Fluid Mechanics*, vol. 722, pp. 554–595, 005 2013.

[43] M. C. Cross and P. C. Hohenberg, "Pattern formation outside of equilibrium," *Rev. Mod. Phys.*, vol. 65, pp. 851–1112, Jul 1993.

[44] E. Bodenschatz, W. Pesch, and G. Ahlers, "Recent developments in Rayleigh-Bénard convection," *Annual Review of Fluid Mechanics*, vol. 32, no. 1, pp. 709–778, 2000.

[45] M. Paul, K.-H. Chiam, M. Cross, P. Fischer, and H. Greenside, "Pattern formation and dynamics in Rayleigh–Bénard convection: numerical simulations of experimentally realistic geometries," *Physica D*, vol. 184, no. 1, pp. 114–126, 2003.

[46] K. Mischaikow and V. Nanda, "Morse theory for filtrations and efficient computation of persistent homology," *Discrete & Computational Geometry*, vol. 50, no. 2, pp. 330–353, 2013.

[47] "Perseus." `http://people.maths.ox.ac.uk/nanda/perseus/index.html`, January 2017.

[48] P. Bubenik, "Statistical topological data analysis using persistence landscapes," *J. Mach. Learn. Res.*, vol. 16, pp. 77–102, Jan. 2015.

[49] S. Harker, R. L. amd M. Kramár, and K. Mischaikow, "An adaptive subsample approximation for vietoris-rips filtrations of large point clouds." In preparation.

[50] P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics*. Cambridge University Press, 2000.

[51] C. Harrison, Z. Cheng, S. Sethuraman, D. A. Huse, P. M. Chaikin, D. A. Vega, J. M. Sebastian, R. A. Register, and D. H. Adamson, "Dynamics of pattern coarsening in a two-dimensional smectic system," *Phys. Rev. E*, vol. 66, p. 011706, Jul 2002.

[52] N. Ercolani, R. Indik, A. Newell, and T. Passot, "Global description of patterns far from onset: A case study," *Physica D: Nonlinear Phenomena*, vol. 184, pp. 127–140, 10 2003.

[53] W. Meevasana and G. Ahlers, "Rayleigh-bénard convection in elliptic and stadium-shaped containers," *Physical Review E*, vol. 66, no. 4, p. 046308, 2002.

[54] M. Cross and H. Greenside, *Pattern Formation and Dynamics in Nonequilibrium Systems*. Cambridge University Press, 2009.

[55] D. A. Egolf, I. V. Melnikov, and E. Bodenschatz, "Importance of local pattern properties in spiral defect chaos," *Phys. Rev. Lett.*, vol. 80, pp. 3228–3231, Apr 1998.

[56] M. C. Cross, D. Meiron, and Y. Tu, "Chaotic domains: A numerical investigation," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 4, pp. 607–619, dec 1994.

[57] C. Bowman and A. C. Newell, "Natural patterns and wavelets," *Rev. Mod. Phys.*, vol. 70, pp. 289–301, Jan 1998.

[58] A. M. Bazen and S. H. Gerez, "Systematic methods for the computation of the directional fields and singular points of fingerprints," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, pp. 905–919, July 2002.

[59] N. Becker and G. Ahlers, "Local wave director analysis of domain chaos in Rayleigh-Bénard convection," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2006, no. 12, p. P12002, 2006.

[60] R. Forman, "Morse theory for cell complexes," *Advances in Mathematics*, vol. 134, no. 1, pp. 90 – 145, 1998.

[61] V. Robins, P. J. Wood, and A. P. Sheppard, "Theory and algorithms for constructing discrete morse complexes from grayscale digital images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, pp. 1646–1658, Aug. 2011.

[62] U. Bauer, M. Kerber, J. Reininghaus, and H. Wagner, "Phat persistent homology algorithms toolbox," *Journal of Symbolic Computation*, vol. 78, pp. 76 – 90, 2017. Algorithms and Software for Computational Topology.

[63] J. Staal, S. Kalitzin, B. T. H. Romeny, and M. Viergever, "M.: Detection of critical structures in scale space," *LNCS*, p. 105, 1999.

[64] A. Kuijper, "On detecting all saddle points in 2d images," *Pattern Recognition Letters*, vol. 25, no. 15, pp. 1665 – 1672, 2004.

[65] N. M. Abukhdeir and A. D. Rey, "Defect kinetics and dynamics of pattern coarsening in a two-dimensional smectic-a system," *New Journal of Physics*, vol. 10, no. 6, p. 063025, 2008.

[66] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision*, vol. 60, pp. 91–110, Nov. 2004.

[67] J. C. Bezdek, R. J. Hathaway, R. E. Howard, C. A. Wilson, and M. P. Windham, "Local convergence analysis of a grouped variable version of coordinate descent," *Journal of Optimization Theory and Applications*, vol. 54, no. 3, pp. 471–477, 1987.

[68] D. M. Blei, A. Kucukelbir, and J. D. McAuliffe, "Variational inference: A review for statisticians," 2016.

[69] H. Fiedler, "Coherent structures in turbulent flows," *Progress in Aerospace Sciences*, vol. 25, no. 3, pp. 231 – 269, 1988.

[70] V. Seib, M. Kusenbach, S. Thierfelder, and D. Paulus, "Object recognition using hough-transform clustering of surf features," in *Workshops on Electronical and Computer Engineering Subfields*, pp. 169–176, 2012.