THESIS

PERSISTENCE STABILITY FOR METRIC THICKENINGS

Submitted by

Michael Moy

Department of Mathematics

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2021

Master's Committee:

Advisor: Henry Adams

Emily King
Asa Ben-Hur

ABSTRACT

PERSISTENCE STABILITY FOR METRIC THICKENINGS

Persistent homology often begins with a filtered simplicial complex, such as the Vietoris–Rips complex or the Čech complex, the vertex set of which is a metric space. An important result, the stability of persistent homology, shows that for certain types of filtered simplicial complexes, two metric spaces that are close in the Gromov–Hausdorff distance result in persistence diagrams that are close in the bottleneck distance. The recent interest in persistent homology has motivated work to better understand the homotopy types and persistent homology of these commonly used simplicial complexes. This has led to the definition of metric thickenings, which agree with simplicial complexes for finite vertex sets but may have different topologies for infinite vertex sets. We prove Vietoris–Rips metric thickenings and Čech metric thickenings have the same persistence diagrams as their corresponding simplicial complexes for all totally bounded metric spaces. This immediately implies the stability of persistent homology for these metric thickenings.

ACKNOWLEDGEMENTS

I would like to thank my advisor Henry Adams for his support and guidance.

DEDICATION

I dedicate this work to my family.

TABLE OF CONTENTS

Chapter 1

Introduction

While topology studies shapes and spaces in an abstract sense, techniques from topology have recently found applications in practical areas such as data science, creating the field of applied topology. The field is based on the premise that point sets or functions arising in practical applications can have underlying geometric and topological structures. While this is natural to expect, uncovering and describing such structures can be an intricate process, requiring appropriate translations of problems into the language of topology. From the very beginning, this translation poses a challenge, in part because real world data often presents an incomplete view of the structure we might like to study.

Consider, for instance, a dataset that can be viewed as a finite subset of some Euclidean space $\mathbb{R}^n$. Regardless of whatever shape or structure the data points may have, as a subset of $\mathbb{R}^n$ the dataset simply has the discrete topology. That is, the topology of the dataset disregards the relationships between the data points. The idea that a dataset can have structure in spite of this fact suggests that we are not truly interested in the dataset alone, but in the implicit shape or pattern outlined by the dataset. This is handled in applied topology by constructing topological spaces meant to reflect the structure of the dataset, often simplicial complexes that have data points as vertices. Common simplicial complexes are Vietoris–Rips complexes and Čech complexes, which both (in slightly different ways) include all simplices up to a certain size as determined by a parameter $r \in \mathbb{R}$. We expect that because these complexes are constructed based on distances between data points, they reflect the structure of the dataset.

Given a method for constructing simplicial complexes on a dataset, the remaining task is to summarize the topology of these complexes. The main tool used is persistent homology, and in fact, simplicial complexes play such a large role in applied topology because they provide a framework for persistent homology. In this framework, we consider not just a single simplicial complex that describes a dataset, but a sequence or filtration of simplicial complexes that together give a more complete description of the dataset. Persistent homology then gives a reductive view of this entire filtration of complexes. The output of persistent homology is a persistence diagram, which is a set of points in the plane that records the evolution of certain features in the filtration. One of the main results in the study of persistent homology is the stability of persistent homology, which, roughly speaking, states that small changes to a dataset result in small changes to the persistence diagram if Vietoris–Rips or Čech complexes are used. This result establishes that persistent homology provides a reasonable summary of a dataset, as persistence diagrams will not change drastically if small amounts of noise are added to the data.

Simplicial complexes and the stability of persistent homology are the starting points for the work contained here: we expand on recent work studying simplicial complexes, often in the context of, or motivated by, applied topology. This has included the study of simplicial complexes on infinite vertex sets, which can be partially motivated by the stability of persistent homology. Since the stability of persistent homology treats finite and infinite vertex sets equally, an infinite vertex set and a close finite approximation will have similar persistent homology. In the context of applied topology, if data points are sampled from some underlying space, then as the sample becomes denser, the persistent homology of the sample will approach the persistent homology of the underlying space. An example of these ideas appears in recent work: the homotopy types of Vietoris–Rips complexes with vertex sets equal to the circle are found in [1], and the homotopy types of Vietoris–Rips complexes with finitely many vertices evenly spaced on the circle are found in [2]. Stability implies that the persistent homology must be similar between these two cases.

However, simplicial complexes with infinite vertex sets can have some unintuitive or undesirable properties. For instance, in the case of Vietoris–Rips or Čech complexes, the vertex set of the complex must be a metric space, but the simplicial complex does not always respect the topology of the metric space: specifically, the inclusion of the vertex set into the complex is not always continuous. In [3], the authors propose an alternative to simplicial complexes called metric thickenings, which agree with simplicial complexes in the case of finite vertex sets but may differ for infinite vertex sets. A metric thickening also must be built on a metric space, and in contrast to simplicial complexes, the inclusion of the underlying metric space into a metric thickening is always continuous. Another major difference is that metric thickenings are always metric spaces, whereas not all simplicial complexes are metrizable; see [3] for further discussion.

The Vietoris–Rips and Čech metric thickenings are analogs of the simplicial complexes, and in certain instances, homotopy types of these metric thickenings have been found. Theorem 4.2, Theorem 4.4, and Corollary 6.8 of [3] are analogs of previous results for simplicial complexes, establishing homotopy types in certain cases. For a more specific example, see [3] and [4] for work on the homotopy types of Vietoris–Rips metric thickenings of the circle and more generally of n-spheres; the homotopy types are established for sufficiently small parameters.

There is room for much more work on metric thickenings and their relationship with simplicial complexes. Our work here will examine their relationship from the perspective of persistent homology. Starting with the same metric space X, we can consider the persistent homology of either a filtration of simplicial complexes or a filtration of metric thickenings. Since both filtrations are meant to describe X, we should hope that they have similar persistence diagrams. We should also hope that there are stability results for the persistent homology of metric thickenings, analogous to those for simplicial complexes – this is suggested in Conjecture 6.14 of [3]. We will show that Vietoris–Rips, intrinsic Čech, and ambient Čech metric thickenings have both of these desirable properties for all totally bounded metric spaces.

We will begin in the following chapter with background information on simplicial complexes, persistent homology, and metric thickenings. Chapter 3 will cover preliminary results needed to discuss maps between metric thickenings. The main results are in Chapter 4, which is divided into three sections covering Vietoris–Rips, intrinsic Čech, and ambient Čech metric thickenings. The main results are summarized below:

- Theorem 4 states that for totally bounded metric spaces, the persistence diagrams of Vietoris–Rips complexes and Vietoris–Rips metric thickenings are identical, and Theorems 6 and 8 are the analogous statements for intrinsic and ambient Čech complexes and metric thickenings.

- Theorems 5, 7, and 9 establish the stability of persistent homology for Vietoris–Rips, intrinsic Čech, and ambient Čech metric thickenings, with bounds that match the stability theorems for the corresponding simplicial complexes.

These results help establish metric thickenings as appropriate tools for the study of persistent homology. Furthermore, they show that metric thickenings and simplicial complexes complement each other: while they agree on persistent homology, their different topologies may make one more suitable than the other depending on the context. With these goals in mind, we proceed to the background material.

Chapter 2

Background

2.1 Simplicial Complexes and Filtrations

We begin by introducing some of the common filtered simplicial complexes used in persistent homology. Each simplicial complex we consider has a vertex set that lies in a metric space, where the metric is used to specify the simplices of the complex. For any metric space, we will write d for the metric as long as there is no ambiguity. Each of the following definitions specifies an abstract simplicial complex; when discussing the topology of a simplicial complex, the geometric realization is used. In general, we will use the same name for a simplicial complex and its geometric realization. We will see certain situations in which it is helpful to use the empty simplicial complex, the geometric realization of which is the empty set.

The following definitions aim to construct simplicial complexes that reflect the shapes of their vertex sets by including simplices up to a certain size. In the context of applied topology, these complexes are used to associate a topological space to a dataset. These complexes are defined, for instance, in [5–7]. We follow [3] and define two subtly different versions of each complex.

Definition 1. Let X be a metric space. The *Vietoris–Rips complex* with the $\leq$ convention and parameter $r \in \mathbb{R}$ is defined by

$$\mathrm{VR}_{\leq}(X;r) = \Big\{\{x_1,\ldots,x_n\} \subseteq X \;\Big|\; \mathrm{diam}\big(\{x_1,\ldots,x_n\}\big) \leq r\Big\},$$

and we further define the Vietoris–Rips complex with the $<$ convention by

$$\mathrm{VR}_{<}(X;r) = \Big\{\{x_1,\ldots,x_n\} \subseteq X \;\Big|\; \mathrm{diam}\big(\{x_1,\ldots,x_n\}\big) < r\Big\}.$$

We note that $\mathrm{VR}_{\leq}(X;r)$ is the empty simplicial complex for $r < 0$ and $\mathrm{VR}_{<}(X;r)$ is the empty simplicial complex for $r \leq 0$. The next type of simplicial complex is defined in a similar manner.

Definition 2. Let X be a metric space. The *intrinsic Čech complex* with the $\leq$ convention and parameter $r \in \mathbb{R}$ is defined by

$$\check{\mathrm{C}}_{\leq}(X;r) = \Big\{ \{x_1, \ldots, x_n\} \subseteq X \;\Big|\; \text{for some } c \in X,\ d(x_i, c) \leq r \text{ for all } i \Big\},$$

and the intrinsic Čech complex with the $<$ convention is defined by

$$\check{\mathrm{C}}_{<}(X;r) = \Big\{ \{x_1, \ldots, x_n\} \subseteq X \;\Big|\; \text{for some } c \in X,\ d(x_i, c) < r \text{ for all } i \Big\}.$$

As with Vietoris–Rips complexes, we note that $\check{\mathrm{C}}_{\leq}(X;r)$ is the empty simplicial complex for $r < 0$ and $\check{\mathrm{C}}_{<}(X;r)$ is the empty simplicial complex for $r \leq 0$. If $\{x_1, \ldots, x_n\} \in \check{\mathrm{C}}_{\leq}(X;r)$, then any $c \in X$ such that $d(x_i, c) \leq r$ for all i is called an *r-center*, or simply a *center*, for the simplex, and similarly for simplices of $\check{\mathrm{C}}_{<}(X;r)$. These complexes are called *intrinsic* Čech complexes because centers are required to be in X. If instead we choose to work with an X that is a subset of some ambient space, we could consider an alternate definition that allows centers in the ambient space. For instance, X may consist of a finite set of points in a Euclidean space. In this setting, it is common to picture a ball of radius r in Euclidean space centered at each $x \in X$; we would then form a simplex $\{x_1, \ldots, x_n\}$ whenever the intersection of the balls centered at these points is nonempty. The complex we are describing is thus the *nerve* of this collection of balls, and Čech complexes are sometimes described this way. Generalizing this idea from Euclidean space leads to the following (see [5]).

Definition 3. Let L and W be subsets (containing "landmarks" and "witnesses") of some ambient metric space. The *ambient Čech complex* with the $\leq$ convention and parameter $r \in \mathbb{R}$ is defined by

$$\check{\mathrm{C}}_{\leq}(L, W;r) = \Big\{ \{x_1, \ldots, x_n\} \subseteq L \;\Big|\; \text{for some } w \in W,\ d(x_i, w) \leq r \text{ for all } i \Big\},$$

and the ambient Čech complex with the $<$ convention is defined by

$$\check{\mathrm{C}}_<(L,W;r) = \Big\{\{x_1,\ldots,x_n\} \subseteq L \;\Big|\; \text{for some } w \in W,\ d(x_i,w) < r \text{ for all } i\Big\}.$$

Here we will note that a landmark $x \in L$ is not in the complex $\check{\mathrm{C}}_\leq(L,W;r)$ if there is no witness $w \in W$ such that $d(x,w) \leq r$, and similarly for the $<$ convention. This also implies $\check{\mathrm{C}}_\leq(L,W;r)$ is the empty simplicial complex for $r < 0$ and $\check{\mathrm{C}}_<(L,W;r)$ is the empty simplicial complex for $r \leq 0$. Ambient Čech complexes in fact generalize intrinsic Čech complexes, since if X is the ambient space, $\check{\mathrm{C}}(X,X;r) = \check{\mathrm{C}}(X;r)$.

Usually the terms "Vietoris–Rips complex" and "Čech complex" refer to $\mathrm{VR}_\leq(X;r)$ and $\check{\mathrm{C}}_\leq(X;r)$ or $\check{\mathrm{C}}_\leq(L,W;r)$, but the choice of the $\leq$ or $<$ convention becomes important in certain cases involving infinite vertex sets (see [1] for instance). Following [3], we will write $\mathrm{VR}(X;r)$, $\check{\mathrm{C}}(X;r)$, and $\check{\mathrm{C}}(L,W;r)$ in cases where either the $\leq$ or $<$ convention can be used as long as the same convention is applied consistently throughout a statement or a proof.

We have noted that each of these complexes is empty for negative parameters, and in the case of the ambient Čech complexes, a landmark is not present if it is not close enough to any witnesses. In general, this means that the vertex sets of our complexes change depending on the parameter, although they are always subsets of the same metric space. While this is not always the convention and it is possible to insist that the same vertex set is present at all parameters (as in [5], for instance), allowing empty simplicial complexes and changing vertex sets provides the right setting for persistent homology, as we will describe shortly.

An important feature of these definitions is the parameter $r \in \mathbb{R}$, which determines the size of simplices included in the complex. The need for a parameter poses a problem: without any prior knowledge about a vertex set, there is not an obvious choice of parameter that will best reflect the shape of the vertex set. One solution, the approach taken in applied topology, is to view these simplicial complexes at a range of parameters, observing the changes as the parameter varies. As such, we will generally consider a family of Vietoris–Rips complexes or a family of Čech complexes at all parameters. We write $\mathrm{VR}(X;\_)$ for the family of complexes $\mathrm{VR}(X;r)$ for all $r \in \mathbb{R}$, and similarly, we write $\check{\mathrm{C}}(X;\_)$ and $\check{\mathrm{C}}(L,W;\_)$ for families of Čech complexes. From the definitions, we can see that if $a \leq b$, then $\mathrm{VR}(X;a) \subseteq \mathrm{VR}(X;b)$, and similarly for the Čech complexes. In general, a *filtered simplicial complex*, or a *filtration* of simplicial complexes, is defined as a family $\{S_r\}_{r\in\mathbb{R}}$ of simplicial complexes such that $S_a \subseteq S_b$ whenever $a \leq b$ (see [5]). Because each complex in a filtration is a subset of those with higher parameters, a filtration is often described or visualized as growing over time.

We will refer to (filtered) Vietoris–Rips and Čech simplicial complexes throughout, and we will introduce related topological spaces, the Vietoris–Rips and Čech metric thickenings, in section 2.3. Filtered simplicial complexes, especially Vietoris–Rips and Čech complexes, play a significant role in applied topology because they assign additional topological structure to a dataset. Perhaps most importantly, filtered simplicial complexes are a common setting for persistent homology.

2.2 Persistent Homology

Persistent homology is a tool in the area of applied topology that is used to give a rough characterization of the shape of a dataset or the evolution of a space over time. In this section we will give a brief overview of persistent homology and the relevant background for discussing stability. Further details can be found in [5–10].

While persistent homology can be defined for finite sequences of spaces, and indeed practical computations will involve finitely many spaces, certain theoretical results can be stated for families of spaces indexed by the real line. Since our results are stated for such families, we will introduce persistent homology in this context. We begin with a family of topological spaces $\{X_a\}_{a\in\mathbb{R}}$ with maps $f_a^b\colon X_a \to X_b$ for any $a,b \in \mathbb{R}$ with $a \leq b$. We will require[1] that the maps satisfy $f_a^a = \mathrm{id}_{X_a}$ for all a, and $f_b^c \circ f_a^b = f_a^c$ for all $a \leq b \leq c$. The prototypical examples of such spaces are the filtered simplicial complexes described above, along with the inclusion maps.

[1] In category theoretic terms, we are considering an $\mathbb{R}$-indexed diagram of topological spaces, where $\mathbb{R}$ is viewed as a category with real numbers as objects and, for each pair $a,b \in \mathbb{R}$, a single arrow $a \to b$ if and only if $a \leq b$. Applying homology as in the following paragraphs, we obtain an $\mathbb{R}$-indexed diagram of vector spaces. For a description of persistent homology in the language of category theory, see [11].

We will let H be homology in any fixed dimension with coefficients in a field k. Thus, homology modules are vector spaces. When working with simplicial complexes, simplicial homology may be used, but later we will need to use singular homology. We will not need many facts about homology, but we list here the few that we will use. First, homology is a functor from the category of topological spaces to the category of vector spaces over k. The homology vector space $H(\varnothing)$ is the zero vector space. Finally, if $f, g\colon X \to Y$ are homotopic maps between topological spaces, then $H(f) = H(g)$.

Thus, applying homology to the family $\{X_a\}_{a\in\mathbb{R}}$ and the associated maps, we obtain a family of vector spaces over k, indexed by $\mathbb{R}$. Each $H(f_a^b)\colon H(X_a) \to H(X_b)$ is then a linear map, where $H(f_a^a)$ is the identity for each a and $H(f_b^c) \circ H(f_a^b) = H(f_a^c)$ for all $a \leq b \leq c$. We generalize this situation to the following definition (see [5]).

Definition 4. A *persistence module* $\mathbb{V}$ over the real numbers is a family $\{V_a\}_{a\in\mathbb{R}}$ of vector spaces over a field k, along with a family of linear maps $\{v_a^b\colon V_a \to V_b \mid a \leq b\}$ such that v_a^a is the identity on V_a for each a, and $v_b^c \circ v_a^b = v_a^c$ for all $a \leq b \leq c$.

Thus, $\{H(X_a)\}_{a\in\mathbb{R}}$ along with the maps $H(f_a^b)$ is an example of a persistence module, called the *persistent homology module* of $\{X_a\}_{a\in\mathbb{R}}$. If $\mathbb{V}$ is a persistence module, an element $z \in V_a$ has an image $v_a^b(z)$ in V_b, so if the index is viewed as time, the persistence module records the evolution of each element. In the case of a persistent homology module, a nonzero element of a homology vector space $H(X_a)$ is interpreted as a hole in X_a, and the evolution of this element describes the hole as the topological space evolves around it. In particular, we can record when a nonzero element first appears (i.e., is not the image of a nonzero element at a previous time), and when it is either mapped to 0 or merges with an older nonzero element. These are referred to as the birth and death times, and an element will be viewed as being born at a certain time and dying at another time. Birth and death times are allowed to be $\pm\infty$ when a nonzero element is alive at all parameters less than some value or is alive at all parameters greater than some value. Persistent homology is based on the observation that, under certain conditions, a persistence module can be summarized by recording the birth and death times of all elements with positive lifetimes. It is not immediately clear in what cases this can be done, and in particular, it is important that this process can be done consistently across the entire persistence module; for work resolving these issues, see [8, 9, 12].

A record of birth and death times of elements with positive lifetimes in $\mathbb{V}$ is kept in a *persistence diagram*[2], denoted $\mathrm{dgm}(\mathbb{V})$. This is a multiset of points in the extended half plane $\{(x, y) \mid -\infty \leq x < y \leq \infty\}$, where each point has an x-coordinate equal to an element's birth time and a y-coordinate equal to the corresponding death time (see [6, 9, 12]). A persistence diagram is the output of persistent homology, given our input of a family of topological spaces such as a filtered simplicial complex. The criterion we will use to show a persistence module has a well-defined persistence diagram is given below in Theorem 1.

Since a persistence diagram is meant to give a concise summary of the topological spaces from which a persistence module is built, we should hope that small changes to the spaces result in small changes to the persistence diagram. This turns out to be true, and we can give explicit bounds on these small changes provided we have appropriate notions of distance. The results establishing this property are generally referred to as the *stability of persistent homology*. The key result is a statement about persistence modules (Theorem 1, below), which can be used to give more explicit results applying to filtered simplicial complexes (Theorems 2 and 3). In the remainder of this section, we will develop the terminology to discuss these results. To begin, we need a method of comparing persistence diagrams.

The following definitions are from [9]; for an alternate description, see [12]. In a persistence diagram, we will view a point with multiplicity as multiple distinct copies of a point (formally, a multiset can be realized as a set), so we will treat persistence diagrams as sets of distinct points. Define a *partial matching* between two persistence diagrams $\mathrm{dgm}(\mathbb{U})$ and $\mathrm{dgm}(\mathbb{V})$ to be a set of pairs $M \subseteq \mathrm{dgm}(\mathbb{U}) \times \mathrm{dgm}(\mathbb{V})$ satisfying the following: for each $P \in \mathrm{dgm}(\mathbb{U})$, there is at most one $Q \in \mathrm{dgm}(\mathbb{V})$ such that $(P, Q) \in M$ and for each $Q \in \mathrm{dgm}(\mathbb{V})$, there is at most one $P \in \mathrm{dgm}(\mathbb{U})$ such that $(P, Q) \in M$. Let Δ denote the diagonal of the plane. We will use the l_∞ norm to measure

[2] Specifically, this is the *undecorated* persistence diagram: see [9].

distances in the plane, so let $d^\infty\big((x_1, y_1), (x_2, y_2)\big) = \max\big(|x_1 - x_2|, |y_1 - y_2|\big)$. If a point (x, y) is in a persistence diagram, then it is above the diagonal, so $d^\infty\big((x, y), \Delta\big) = \frac{1}{2}(y - x)$. Call a partial matching M between $\mathrm{dgm}(\mathbb{U})$ and $\mathrm{dgm}(\mathbb{V})$ an *ε-matching* if for any $(P, Q) \in M$, $d^\infty(P, Q) \le \varepsilon$ and if any unmatched point P in either $\mathrm{dgm}(\mathbb{U})$ or $\mathrm{dgm}(\mathbb{V})$ satisfies $d_\infty\big(P, \Delta\big) \le \varepsilon$. To understand why matchings are defined this way, notice that a point on the diagonal would be interpreted as having equal birth and death times. So in an ε-matching, points of the diagrams with short enough lifetimes (at most 2ε) can be disregarded, while the remaining points must be matched with points in the other diagram with similar birth and death times (each within a tolerance of ε). An ε-matching gives an indication of how similar two diagrams are, leading to the following definition (see [9]).

Definition 5. If persistence modules $\mathbb{U}$ and $\mathbb{V}$ have well-defined persistence diagrams, then the *bottleneck distance* between these diagrams is defined by

$$\mathrm{d_b}\big(\mathrm{dgm}(\mathbb{U}), \mathrm{dgm}(\mathbb{V})\big) = \inf\big\{\varepsilon \mid \text{there exists an } \varepsilon\text{-matching between } \mathrm{dgm}(\mathbb{U}) \text{ and } \mathrm{dgm}(\mathbb{V})\big\}.$$

See [9] for a more thorough development of persistence diagrams and the bottleneck distance. The following definition will be used to characterize when persistence diagrams are well-defined and will be important in the statement of stability.

Definition 6 (See [5,9])**.** A persistence module $\mathbb{V}$ is called *q-tame* if v_a^b has finite rank for all $a < b$.

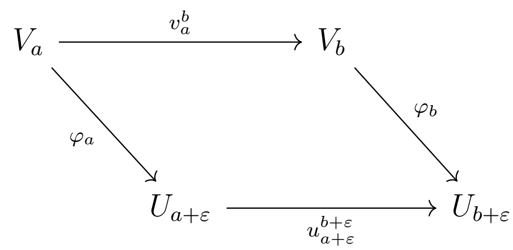
We will state the stability theorems in terms of persistence modules. A key feature of the following results is the direct comparison of persistence modules, using the definition below.

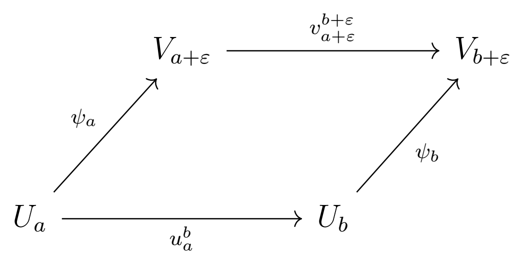
Definition 7 (See [5, 9])**.** Two persistence modules $\mathbb{U}$ and $\mathbb{V}$ are said to be *ε-interleaved* if there exist two families of linear maps $\{\varphi_a \colon V_a \to U_{a+\varepsilon}\}_{a\in\mathbb{R}}$ and $\{\psi_a \colon U_a \to V_{a+\varepsilon}\}_{a\in\mathbb{R}}$ such that for all $a < b$, the following conditions are satisfied:

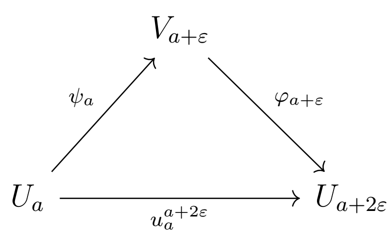
- $\varphi_b \circ v_a^b = u_{a+\varepsilon}^{b+\varepsilon} \circ \varphi_a$
- $\psi_b \circ u_a^b = v_{a+\varepsilon}^{b+\varepsilon} \circ \psi_a$

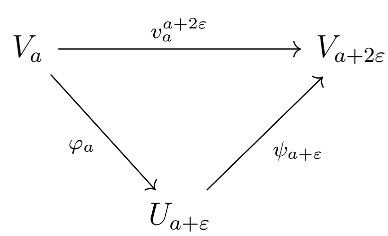
- $\varphi_{a+\varepsilon} \circ \psi_a = u_a^{a+2\varepsilon}$
- $\psi_{a+\varepsilon} \circ \varphi_a = v_a^{a+2\varepsilon}$.

These can be expressed by saying the following four diagrams commute:









In this case, we also say that the families $\{\varphi_a\}$ and $\{\psi_a\}$ form an ε-*interleaving*.

We note some basic properties of interleavings. If $\mathbb{U}$ and $\mathbb{V}$ are ε-interleaved, then they are ε'-interleaved for any $\varepsilon' \geq \varepsilon$. Given an ε_1-interleaving between $\mathbb{U}$ and $\mathbb{V}$ and an ε_2-interleaving between $\mathbb{V}$ and $\mathbb{W}$, we can compose the appropriate maps to get an $(\varepsilon_1 + \varepsilon_2)$-interleaving between $\mathbb{U}$ and $\mathbb{W}$. In the context of persistent homology modules, one of the simplest ways to obtain an interleaving is to find maps on the families of topological spaces that commute in an analogous way, then apply the homology functor. For instance, for any metric space X, any r, and any $\varepsilon > 0$, we have inclusions $\mathrm{VR}_{\leq}(X;r) \hookrightarrow \mathrm{VR}_{<}(X;r+\varepsilon)$ and $\mathrm{VR}_{<}(X;r) \hookrightarrow \mathrm{VR}_{\leq}(X;r+\varepsilon)$. These give induced maps on homology, which define an ε-interleaving between $H(\mathrm{VR}_{\leq}(X;\_))$ and $H(\mathrm{VR}_{<}(X;\_))$. Thus, $H(\mathrm{VR}_{\leq}(X;\_))$ and $H(\mathrm{VR}_{<}(X;\_))$ are ε-interleaved for any $\varepsilon > 0$, and the same is true for Čech complexes. In general, it may be too much to expect that the maps on topological spaces commute. However, it is sufficient to find maps that commute up to homotopy, since homotopic maps induce equal maps on homology – we will use this strategy later.

We now have the terminology to state the following theorem:

Theorem 1 (Theorem 2.3 of [5])**.** *Any q-tame persistence module has a well-defined persistence diagram. If $\mathbb{U}$ and $\mathbb{V}$ are q-tame persistence modules that are ε-interleaved, then* $\mathrm{d_b}\big(\mathrm{dgm}(\mathbb{U}), \mathrm{dgm}(\mathbb{V})\big) \leq \varepsilon$.

This algebraic result can be used to show more concrete stability results. Intuitively, when we form Vietoris–Rips or Čech filtrations on a metric space X, we would like to know that small changes to the metric space result in small changes to the persistence diagram. Formalizing and proving this idea requires an appropriate notion of distance between two metric spaces. We outline the setting used by [5] in the definitions below.

Definition 8. Let X be a metric space. For an $\varepsilon > 0$, a subset $Y \subseteq X$ is a called an ε-*sample* of X if for any $x \in X$, there exists a $y \in Y$ such that $d(x,y) < \varepsilon$. Furthermore, X is called *totally bounded* if it has a finite ε-sample for all $\varepsilon > 0$.

Definition 9. A *correspondence* between two sets X and Y is a subset of $X \times Y$ that projects surjectively onto both X and Y.

Correspondences can also be discussed in the language of multivalued maps. A correspondence $C \subseteq X \times Y$ allows us to associate points in X with points in Y in a broader context than functions would; one point in X can be associated to multiple points in Y and vice versa.

Definition 10. The *Gromov–Hausdorff distance* between two nonempty[3] metric spaces (X, d_X) and (Y, d_Y) is defined as

$$d_{GH}(X,Y) = \frac{1}{2} \inf_C \sup \left\{ |d_X(x,x') - d_Y(y,y')| : (x,y),(x',y') \in C \right\},$$

where the infimum is taken over all correspondences $C \subseteq X \times Y$.

The expression $\sup \left\{ |d_X(x,x') - d_Y(y,y')| : (x,y),(x',y') \in C \right\}$ is called the *distortion* of the correspondence C, and measures how large of a change in distances can arise from the correspondence. Thus, the Gromov–Hausdorff distance measures how large the distortion must be when we compare two metric spaces by a correspondence. We will also make use of the *Hausdorff distance*, written as d_H, between subsets of a metric space.

We can now state the stability results for Vietoris–Rips and Čech complexes:

Lemma 1 (Lemmas 4.3 and 4.4 of [5])**.** *Let X and Y be metric spaces. For any $\varepsilon > 2d_{GH}(X,Y)$, the persistence modules $H(\mathrm{VR}(X;\_))$ and $H(\mathrm{VR}(Y;\_))$ are ε-interleaved, and the persistence modules $H(\check{\mathrm{C}}(X;\_))$ and $H(\check{\mathrm{C}}(Y;\_))$ are ε-interleaved.*

Lemma 2 (Corollary 4.10 of [5])**.** *Let L, L', and W be subsets of a metric space. For any $\varepsilon > d_H(L,L')$, the persistence modules $H(\check{\mathrm{C}}(L,W;\_))$ and $H(\check{\mathrm{C}}(L',W;\_))$ are ε-interleaved.*

We will briefly outline the proof in the case of Vietoris–Rips complexes. If $\varepsilon > 2d_{GH}(X,Y)$, then there exists a correspondence $C \subseteq X \times Y$ that has distortion less than ε. Then if $\sigma \in \mathrm{VR}(X;r)$, any finite subset of $C(\sigma) = \{y \in Y \mid (x,y) \in C \text{ for some } x \in \sigma\}$ is a simplex of $\mathrm{VR}(Y; r+\varepsilon)$, since by the definition of distortion, $d(y_1,y_2) < d(x_1,x_2) + \varepsilon$ whenever $(x_1,y_1),(x_2,y_2) \in C$. Since C projects surjectively onto X, there is at least one function $f\colon X \to Y$ such that $(x,f(x)) \in C$ for all x, and this function can be used to define a simplicial

[3]To extend the definition to include the empty metric space, we suggest the following (this will be necessary later if Theorems 5, 7, and 9 are to apply to the empty metric space). In the case that either X or Y is empty and the other is nonempty, let $d_{GH}(X,Y) = +\infty$; this is consistent with the convention that the infimum of the empty set is $+\infty$. Also let $d_{GH}(\varnothing,\varnothing) = 0$. This is not consistent with the convention that the supremum of the empty set is $-\infty$, but defines the distance to be nonnegative. See also [13] for details on the Gromov–Hausdorff distance.

map $\mathrm{VR}(X;r) \to \mathrm{VR}(Y;r+\varepsilon)$. The fact that any finite subset of $C(\sigma)$ is a simplex for each $\sigma \in \mathrm{VR}(X;r)$ can be used to show that any two functions defined this way are homotopic. Thus, any such functions induce the same map $H(\mathrm{VR}(X;r)) \to H(\mathrm{VR}(Y;r+\varepsilon))$. Similarly, we get a map $H(\mathrm{VR}(Y;r)) \to H(\mathrm{VR}(X;r+\varepsilon))$ for each r, and it can be checked that these families of maps define an ε-interleaving. The proofs for both types of Čech complexes are similar.

We note here that the argument relies on showing that any finite subset of $C(\sigma)$ is a simplex for each $\sigma \in \mathrm{VR}(X;r)$. This step depended on the fact that $\sigma \in \mathrm{VR}_{\leq}(X;r)$ if and only if $\mathrm{diam}(\sigma) \leq r$ and $\sigma \in \mathrm{VR}_{<}(X;r)$ if and only if $\mathrm{diam}(\sigma) < r$. In particular, it was necessary that $\mathrm{VR}(X;r)$ be empty for $r < 0$, and this is the reason for allowing the empty simplicial complexes in our definitions. More generally, in our definitions, we have allowed the vertex sets to change depending on the parameter. This is mostly a technicality for Vietoris–Rips and intrinsic Čech complexes, as negative parameters are mostly disregarded. However, in the case of ambient Čech complexes, there can be notable differences even at positive parameters if the same vertex set is kept for all parameters. In short, Lemmas 1 and 2 are the justification for allowing vertex sets to depend on the parameter in our definitions of Vietoris–Rips and Čech complexes.

The final step toward concrete stability results is to translate the interleavings above to statements about persistence diagrams. For this, we need well-defined persistence diagrams, which are guaranteed by the following propositions.

Proposition 1 (Proposition 5.1 of [5])**.** *If X is a totally bounded metric space, then the persistence modules $H(\mathrm{VR}(X;\_))$ and $H(\check{\mathrm{C}}(X;\_))$ are q-tame.*

Proposition 2 (Proposition 5.4 of [5])**.** *Let L, W be subsets of a metric space. If at least one of L, W is totally bounded, then the persistence module $H(\check{\mathrm{C}}(L,W;\_))$ is q-tame.*

Combining Theorem 1, Lemma 1, and Proposition 1, we obtain the following:

Theorem 2 (Stability of Vietoris–Rips and intrinsic Čech Complexes: Theorem 5.2 of [5])**.** *Let X and Y be totally bounded metric spaces. Then*

$$\mathrm{d_b}\Big(\mathrm{dgm}\big(H(\mathrm{VR}(X;\_))\big), \mathrm{dgm}\big(H(\mathrm{VR}(Y;\_))\big)\Big) \leq 2d_{GH}(X,Y),$$

and

$$\mathrm{d_b}\Big(\mathrm{dgm}\big(H(\check{\mathrm{C}}(X;\_))\big), \mathrm{dgm}\big(H(\check{\mathrm{C}}(Y;\_))\big)\Big) \leq 2d_{GH}(X,Y).$$

Similarly, combining Theorem 1, Lemma 2, and Proposition 2, we obtain the following:

Theorem 3 (Stability of ambient Čech Complexes: Theorem 5.6 of [5])**.** *Let L, L', and W be subsets of a metric space. Suppose that L and L' are totally bounded, or that W is totally bounded. Then*

$$\mathrm{d_b}\Big(\mathrm{dgm}\big(H(\check{\mathrm{C}}(L,W;\_))\big), \mathrm{dgm}\big(H(\check{\mathrm{C}}(L',W;\_))\big)\Big) \leq d_H(L,L').$$

In the context of Vietoris–Rips or Čech persistent homology of a dataset, these theorems show that small amounts of noise in the dataset will only lead to small changes in the persistence diagram. Moreover, in an ideal case where data is sampled from some underlying space, as the sample gets denser, its persistence diagram approaches the persistence diagram of the underlying space. This underlying space is likely to have infinite cardinality, and thus we have a connection between simplicial complexes built on finite and infinite vertex sets. This is a perspective that has guided some of the recent study of simplicial complexes, and it relies on the generality of these theorems – namely, that the theorems treat finite and infinite metric spaces equally. This leads us to metric thickenings, described in the following section, which provide a possible approach to better understand the relationship between finite and infinite simplicial complexes and their persistent homology.

2.3 Metric Thickenings

In the previous section, we saw how persistent homology motivates further study of filtered simplicial complexes such as the Vietoris–Rips and Čech complexes. In particular, the stability of persistent homology provides justification for studying complexes built on infinite vertex sets. While this may be theoretically appealing, such complexes are potentially difficult to study because they contain infinitely many simplices and may contain simplices of arbitrarily high dimension. Metric thickenings were defined in [3] as an alternative to simplicial complexes, with the primary interest being cases of infinite vertex sets. While our primary interest will be metric thickenings that are analogs of Vietoris–Rips and Čech complexes, we begin by defining them in a general sense, following [3].

Given a metric space (X, d_X), we will start by considering probability measures supported on finite subsets of X. Such measures can be written as $\sum_{i=1}^{n} \lambda_i \delta_{x_i}$, where $x_i \in X$ for each i, each δ_{x_i} is the Dirac delta measure at x_i, $\lambda_i \geq 0$ for each i, and $\sum_{i=1}^{n} \lambda_i = 1$. We will write $\operatorname{supp}(\mu)$ for the support of a measure μ, so for instance $\operatorname{supp}\big(\sum_{i=1}^{n} \lambda_i \delta_{x_i}\big) = \{x_1, \ldots, x_n\}$ if $\lambda_i > 0$ for each i. Metric thickenings of X will be defined as certain sets of finitely supported probability measures on X. We will further equip these sets with the 1-Wasserstein metric, which we describe here only for finitely supported measures; a more general definition is provided in [3]. A *matching* between two finitely supported measures $\mu = \sum_{i=1}^{n} \lambda_i \delta_{x_i}$ and $\nu = \sum_{j=1}^{n'} \lambda'_j \delta_{x'_j}$ on a metric space X is an indexed set $\pi = \{\pi_{i,j}\}_{\substack{1\leq i\leq n \\ 1\leq j\leq n'}}$ of nonnegative real numbers such that $\sum_{i=1}^{n} \pi_{i,j} = \lambda'_j$ for each j and $\sum_{j=1}^{n'} \pi_{i,j} = \lambda_i$ for each i. The 1-Wasserstein distance between these two measures is given by

$$W(\mu, \nu) = \inf_{\pi} \sum_{\substack{1\leq i\leq n \\ 1\leq j\leq n'}} \pi_{i,j} d_X(x_i, x'_j),$$

where the infimum is taken over all matchings π between μ and ν. If the probability measures are thought of as distributions of mass in X, then a matching between two measures can be interpreted as a set of instructions for redistributing the mass of one distribution to form the other: an individual $\pi_{i,j}$ represents the amount of mass moved from x_i to x'_j. The sum $\sum_{\substack{1\leq i\leq n \\ 1\leq j\leq n'}} \pi_{i,j} d_X(x_i, x'_j)$ is called the *cost* of the matching.

We can now give the definition of a metric thickening:

Definition 11. Let X be a metric space and K a simplicial complex with vertex set X. We define the *metric thickening* K^m as the following set of probability measures, equipped with the 1-Wasserstein metric:

$$K^m = \left\{ \sum_{i=1}^{n} \lambda_i \delta_{x_i} \;\middle|\; \lambda_i \geq 0 \text{ for all } i,\, \sum_{i=1}^{n} \lambda_i = 1,\, \{x_1, \ldots, x_n\} \in K \right\}.$$

A K^m as in the definition will be called a *metric thickening of* X when we need to specify the underlying metric space X. As a set, K^m is in bijection with the geometric realization of K via the map $K \to K^m$ given by $\sum_{i=1}^n \lambda_i x_i \mapsto \sum_{i=1}^n \lambda_i \delta_{x_i}$. In fact, this map is always continuous (Proposition 6.1 of [3]). However, it is not always a homeomorphism, so in general K and K^m have different topologies. Thus, a metric thickening is essentially a simplicial complex built on a metric space, endowed with an alternate topology that may not agree with the usual topology.

Some discussion of the differences between simplicial complexes and metric thickenings can be found in [3]. For now, we will note one of the most important differences: if X is a metric space and K a simplicial complex with vertex set X, then the inclusion $X \hookrightarrow K$ is continuous only if X has the discrete topology. On the other hand, if K^m is a metric thickening of X, then the map $X \to K^m$ given by $x \mapsto \delta_x$ embeds X isometrically in K^m and is thus continuous. Note that this is enough to show that the map $K \to K^m$ above is not always a homeomorphism. Because of this embedding, we can view the Wasserstein metric as extending the metric of X: for this reason, we will generally write both the metric on X and the Wasserstein metric on a metric thickening of X as d. If we need to distinguish between metrics, we will use d_{K^m} to indicate the Wasserstein metric on the metric thickening K^m.

While the topologies of K^m and K can be different in general, there is an important case in which they agree:

Proposition 3 (Proposition 6.2 of [3])**.** *If K^m is a metric thickening of a finite metric space X, then the function $K \to K^m$ given by $\sum_{i=1}^n \lambda_i x_i \mapsto \sum_{i=1}^n \lambda_i \delta_{x_i}$ is a homeomorphism.*

While the general definition of a metric thickening of X allows for any simplicial complex, we will be most interested in the cases where the simplicial complex is a Vietoris–Rips complex or a Čech complex. If the simplicial complex used to construct the metric thickening is $\mathrm{VR}(X;r)$, the metric thickening will be written as $\mathrm{VR}^m(X;r)$, and similarly for $\check{\mathrm{C}}^m(X;r)$ and $\check{\mathrm{C}}^m(L,W;r)$. As with the simplicial complexes, we may consider a family of such metric thickenings for all real r. The family $\{\mathrm{VR}^m(X;r)\}_{r\in\mathbb{R}}$ will be written as $\mathrm{VR}^m(X;\_)$ and comes with the inclusions $\mathrm{VR}^m(X;a) \hookrightarrow \mathrm{VR}^m(X;b)$ for all $a \le b$. This forms a filtration of metric thickenings, analogous to our filtrations of simplicial complexes. We will similarly consider the filtrations $\check{\mathrm{C}}^m(X;\_)$ and $\check{\mathrm{C}}^m(L,W;\_)$. We also note that $\mathrm{VR}^m_{\leq}(X;r)$ and $\check{\mathrm{C}}^m_{\leq}(X;r)$ are empty when $r<0$ and $\mathrm{VR}^m_{<}(X;r)$ and $\check{\mathrm{C}}^m_{<}(X;r)$ are empty when $r\leq 0$, since in these cases the underlying simplicial complexes are empty. For ambient Čech metric thickenings, $\check{\mathrm{C}}^m_{\leq}(L,W;r)$ does not contain any measure with x in its support if there is no $w\in W$ such that $d(x,w)\leq r$; this also implies $\check{\mathrm{C}}^m_{\leq}(L,W;r)$ is empty for $r<0$. Similar statements hold for the $<$ convention.

The basic facts we have seen about metric thickenings establish close connections to simplicial complexes. Using Vietoris–Rips metric thickenings as an example, we have seen that $\mathrm{VR}(X;r)$ and $\mathrm{VR}^m(X;r)$ agree for finite X, while they may disagree for infinite X. Thus, Vietoris–Rips metric thickenings are an alternate way of extending the concept of Vietoris–Rips complexes on finite vertex sets to infinite vertex sets, and this opens up questions about homotopy types and persistent homology of Vietoris–Rips metric thickenings.

It is natural to wonder if the stability results of Theorems 2 and 3 have analogs for metric thickenings. As a first approach, one could try applying the method used in [5] to prove Lemma 1. An issue quickly arises with this approach. A map between simplicial complexes can be specified by sending vertices to vertices in a way that preserves simplices and extending linearly: this defines a simplicial map. An analogous map on metric thickenings is not necessarily continuous because in a metric thickening, the vertex set (the underlying metric space) retains its topology as a metric space. Thus, mapping vertices to vertices does not always give a continuous map, as the topology of a metric thickening imposes additional restrictions on how a continuous function can behave on the vertex set. The proof of Lemma 1, which we briefly outlined, relies crucially on simplicial maps, where the functions on vertices are defined using a correspondence and are not necessarily continuous. Thus, attempting to construct the analogs of these simplicial maps between metric thickenings does not necessarily produce continuous functions. The maps considered in Lemma 3.7 of [3] offer one possible solution to this problem: requiring the maps between vertex sets be Lipschitz implies that the maps on metric thickenings analogous to simplicial maps are also Lipschitz. However, these are not general enough to allow for a proof of stability, since correspondences between spaces with low distortion may not always give rise to Lipschitz maps.

Nevertheless, it is still reasonable to expect that stability holds for metric thickenings. We might expect this to be true simply because of the close relationship of between simplicial complexes and metric thickenings. In particular, it is certainly true if we restrict to finite metric spaces, since simplicial complexes and metric thickenings agree in this case: this has already been noted in Corollary 6.13 of [3]. It is also this relationship between simplicial complexes and metric thickenings that suggest metric thickenings may be a useful tool for future work. In general, by studying the properties of a filtration of metric thickenings of X, we may hope to gain a better understanding of the corresponding filtration of simplicial complexes or a filtration of simplicial complexes built on a finite approximation of X. Furthermore, there remains plenty of work to be done to understand the relationship between simplicial complexes and metric thickenings. These are the themes of the following chapters. We will improve the understanding of metric thickenings and simplicial complexes by proving the stability of persistent homology for Vietoris–Rips and Čech metric thickenings, also proving along the way that the metric thickenings have the same persistence diagrams as the corresponding simplicial complexes. We begin with some simple results on metric thickenings in the next chapter.

Chapter 3

Preliminary Results

In this chapter, we prove several lemmas concerning metric thickenings that will be used in the proof of stability. While they will be used in the specific cases of Vietoris–Rips and Čech metric thickenings, we will prove them here in the more general setting of arbitrary metric thickenings. The first lemma will be used in the proofs of two of the following lemmas. It bounds the distance between convex combinations of measures in a metric thickening in a way that is reminiscent of norms of linear combinations in a vector space.

Lemma 3. *Let K^m be a metric thickening of a metric space X. If $\mu_1, \dots, \mu_n, \mu'_1, \dots, \mu'_n \in K^m$ and $c_1, \dots, c_n$ are nonnegative real numbers satisfying $\sum_{k=1}^n c_k = 1$ such that $\sum_{k=1}^n c_k\mu_k \in K^m$ and $\sum_{k=1}^n c_k\mu'_k \in K^m$, then*

$$d\Big(\sum_{k=1}^n c_k\mu_k, \sum_{k=1}^n c_k\mu'_k\Big) \leq \sum_{k=1}^n c_k d(\mu_k, \mu'_k).$$

Proof. Let $\{x_1, \dots, x_m\} \subseteq X$ be the union of the supports of $\mu_1, \dots, \mu_n$ and similarly let $\{x'_1, \dots, x'_{m'}\} \subseteq X$ be the union of the supports of $\mu'_1, \dots, \mu'_n$. Then for each k, we can write

$$\mu_k = \sum_{i=1}^m \lambda_{k,i}\delta_{x_i}$$

and

$$\mu'_k = \sum_{j=1}^{m'} \lambda'_{k,j}\delta_{x'_j}.$$

Let $\varepsilon > 0$. For each k, there is a matching $\{\pi_{k,i,j}\}$ such that

$$\sum_j \pi_{k,i,j} = \lambda_{k,i},$$

$$\sum_i \pi_{k,i,j} = \lambda'_{k,j},$$

and

$$\sum_{i,j} \pi_{k,i,j} d(x_i, x'_j) < d(\mu_k, \mu'_k) + \varepsilon.$$

Then $\{\sum_k c_k \pi_{k,i,j}\}$ is a matching between $\sum_k c_k \mu_k$ and $\sum_k c_k \mu'_k$, since

$$\sum_j \Big(\sum_k c_k \pi_{k,i,j} \Big) = \sum_k c_k \sum_j \pi_{k,i,j} = \sum_k c_k \lambda_{k,i}$$

and

$$\sum_i \Big(\sum_k c_k \pi_{k,i,j} \Big) = \sum_k c_k \sum_i \pi_{k,i,j} = \sum_k c_k \lambda'_{k,j}.$$

So we have

$$\begin{aligned} d\Big(\sum_k c_k \mu_k, \sum_k c_k \mu'_k \Big) &\le \sum_{i,j} \Big(\sum_k c_k \pi_{k,i,j} \Big) d(x_i, x'_j) \\ &= \sum_k c_k \sum_{i,j} \pi_{k,i,j} d(x_i, x'_j) \\ &< \sum_k c_k \big(d(\mu_k, \mu'_k) + \varepsilon \big) \\ &= \varepsilon + \sum_k c_k d(\mu_k, \mu'_k). \end{aligned}$$

Since this holds for all $\varepsilon > 0$, the claimed inequality holds. □

The following is a simple generalization of Lemma 3.9 of [3]. In subsets of Euclidean spaces, straight line homotopies are some of the simplest to construct. The following lemma shows that analogous straight line homotopies can be used in metric thickenings as well. Straight line homotopies are, in fact, the only homotopies we will use in the following sections. The proof from [3] applies with only a small modification. We also provide a proof here that requires less knowledge about the Wasserstein metric.

Lemma 4 (Straight line homotopies in metric thickenings)**.** *Suppose $f, g\colon Z \to K^m$ are continuous functions from any topological space Z to a metric thickening K^m such that $H\colon Z \times I \to K^m$*

given by $H(z,t) = (1-t)f(z) + tg(z)$ is well-defined. Then H is continuous, and thus f and g are homotopic.

Proof. Suppose $\varepsilon > 0$; we show continuity of H at (z_0, t_0). Since f and g are continuous, there is some open set $U \subseteq Z$ containing z_0 such that $d(f(z_0), f(z)) < \frac{\varepsilon}{2}$ and $d(g(z_0), g(z)) < \frac{\varepsilon}{2}$ for all $z \in U$. We will suppose $z \in U$ and $|t - t_0| < \frac{\varepsilon}{2d(f(z_0), g(z_0))}$, and show $d(H(z_0, t_0), H(z, t)) < \varepsilon$. First note by Lemma 3 that

$$d\big(H(z_0, t), H(z, t)\big) \leq (1-t)d\big(f(z_0), f(z)\big) + t\,d\big((g(z_0), g(z)\big) < \frac{\varepsilon}{2}.$$

We next bound $d(H(z_0, t_0), H(z_0, t))$. Let $\{x_1, \ldots, x_n\}$ be the union of the supports of $f(z_0)$ and $g(z_0)$, so we can write these measures as $f(z_0) = \sum_{i=1}^{n} \lambda_i \delta_{x_i}$ and $g(z_0) = \sum_{i=1}^{n} \lambda'_i \delta_{x_i}$. Then

$$H(z_0, t_0) = (1-t_0)f(z_0) + t_0 g(z_0) = \sum_{i=1}^{n} \big((1-t_0)\lambda_i + t_0\lambda'_i\big)\delta_{x_i}$$

and

$$H(z_0, t) = (1-t)f(z_0) + tg(z_0) = \sum_{j=1}^{n} \big((1-t)\lambda_j + t\lambda'_j\big)\delta_{x_j}.$$

Without loss of generality, suppose $t \geq t_0$. If $\{\pi_{i,j}\}$ is a matching from $f(z_0)$ to $g(z_0)$, we may define a matching from $H(z_0, t_0)$ to $H(z_0, t)$ by leaving mass $(1-t)\lambda_i + t_0\lambda'_i$ stationary at each x_i, and distributing the remaining mass according to $\{\pi_{i,j}\}$. This matching $\{\rho_{i,j}\}$ is given by

$$\rho_{i,j} = \begin{cases} (1-t)\lambda_i + t_0\lambda'_i + (t-t_0)\pi_{i,i} & \text{if } i = j \\ (t-t_0)\pi_{i,j} & \text{if } i \neq j. \end{cases}$$

We verify this is a matching. For any j,

$$\begin{aligned}\sum_{i=1}^{n}\rho_{i,j} &= (1-t)\lambda_j + t_0\lambda'_j + (t-t_0)\pi_{j,j} + \sum_{\substack{1\le i\le n\\ i\neq j}}(t-t_0)\pi_{i,j}\\ &= (1-t)\lambda_j + t_0\lambda'_j + (t-t_0)\sum_{i=1}^{n}\pi_{i,j}\\ &= (1-t)\lambda_j + t_0\lambda'_j + (t-t_0)\lambda'_j\\ &= (1-t)\lambda_j + t\lambda'_j.\end{aligned}$$

Similarly, for any i, we find $\sum_{j=1}^{n}\rho_{i,j} = (1-t_0)\lambda_i + t_0\lambda'_i$, so $\{\rho_{i,j}\}$ is a matching from $H(z_0,t_0)$ to $H(z_0,t)$. The cost of this matching is given by $\sum_{i,j}\rho_{i,j}d(x_i,x_j) = \sum_{i,j}(t-t_0)\pi_{i,j}d(x_i,x_j)$, which is $(t-t_0)$ times the cost of $\{\pi_{i,j}\}$. Since this holds for any matching $\{\pi_{i,j}\}$ between $f(z_0)$ and $g(z_0)$, we find

$$d\big(H(z_0,t),H(z_0,t_0)\big) \le (t-t_0)d\big(f(z_0),g(z_0)\big) < \frac{\varepsilon}{2}.$$

Combining with the other bound, we have

$$d\big(H(z_0,t_0),H(z,t)\big) \le d\big(H(z_0,t_0),H(z_0,t)\big) + d\big(H(z_0,t),H(z,t)\big) < \frac{\varepsilon}{2}+\frac{\varepsilon}{2} = \varepsilon.$$

This shows H is continuous at (z_0,t_0).

□

The following two lemmas will allow us to construct continuous functions into metric thickenings. The first, Lemma 5, allows us to define maps by continuously varying the distribution of mass among a finite set of points. Specifically, the f_i in the lemma form a *partition of unity*: the fact that they sum to 1 allows them to represent masses in probability measures.

Lemma 5. *Let K^m be a metric thickening of a metric space X, let $x_1,\ldots,x_n \in X$, and let Z be a topological space. Suppose $f_1,\ldots,f_n\colon Z \to \mathbb{R}^{\geq 0}$ are continuous functions such that*

$\sum_{i=1}^n f_i(z) = 1$ *for all* $z \in Z$ *(that is, the* f_i *form a partition of unity). Then if the function* $f\colon Z \to K^m$ *given by* $f(z) = \sum_{i=1}^n f_i(z)\delta_{x_i}$ *is well-defined, then it is continuous.*

This lemma can be proven using Proposition 3. We also provide a direct proof here.

Proof. Let $\varepsilon > 0$. If $n = 1$, then the function f is constant and the result holds; so suppose $n > 1$ and set $C = \max_{i,j} d(x_i, x_j)$, which implies $C > 0$. We show continuity at $z_0 \in Z$. By continuity of the functions $f_1, \ldots, f_n$, there exists an open neighborhood $U \subseteq Z$ containing z_0 such that $|f_k(z) - f_k(z_0)| < \frac{\varepsilon}{Cn}$ for all $z \in U$ and all k. So letting $z \in U$, we define a matching between $f(z) = \sum_{i=1}^n f_i(z)\delta_{x_i}$ and $f(z_0) = \sum_{j=1}^n f_j(z_0)\delta_{x_j}$ by keeping as much of the mass as possible fixed, and distributing the remaining mass evenly. Let $m_k = \min(f_k(z), f_k(z_0))$. If $f(z) = f(z_0)$, then $d(f(z), f(z_0)) = 0 < \varepsilon$ as required; if not, then $\sum_{k=1}^n m_k < 1$, and we may define a matching $\{\pi_{i,j}\}$ by

$$\pi_{i,j} = \begin{cases} m_i & \text{if } i = j \\ \frac{1}{1-\sum_{k=1}^n m_k}(f_i(z) - m_i)(f_j(z_0) - m_j) & \text{if } i \neq j. \end{cases}$$

We can verify that this is a matching: noting that $(f_k(z) - m_k)(f_k(z_0) - m_k) = 0$ for each k, we have, for any j,

$$\begin{aligned} \sum_{i=1}^n \pi_{i,j} &= m_j + \sum_{\substack{1\le i\le n \\ i\neq j}} \frac{1}{1-\sum_{k=1}^n m_k}(f_i(z) - m_i)(f_j(z_0) - m_j) \\ &= m_j + \sum_{i=1}^n \frac{1}{1-\sum_{k=1}^n m_k}(f_i(z) - m_i)(f_j(z_0) - m_j) \\ &= m_j + \frac{(f_j(z_0) - m_j)}{1-\sum_{k=1}^n m_k}\sum_{i=1}^n (f_i(z) - m_i) \\ &= m_j + \frac{(f_j(z_0) - m_j)}{1-\sum_{k=1}^n m_k}\Big(1 - \sum_{i=1}^n m_i\Big) \\ &= f_j(z_0). \end{aligned}$$

A similar method shows $\sum_{j=1}^n \pi_{i,j} = f_i(z)$ for each i, so $\{\pi_{i,j}\}$ is in fact a matching. We can bound the cost of the matching as follows:

$$\begin{aligned}
\sum_{i,j} \pi_{i,j} d(x_i, x_j) &= \sum_{i \neq j} \frac{1}{1 - \sum_{k=1}^n m_k} (f_i(z) - m_i)(f_j(z_0) - m_j) d(x_i, x_j) \\
&\leq \frac{C}{1 - \sum_{k=1}^n m_k} \sum_{i,j} (f_i(z) - m_i)(f_j(z_0) - m_j) \\
&= \frac{C}{1 - \sum_{k=1}^n m_k} \Big(\sum_{i=1}^n (f_i(z) - m_i) \Big) \Big(\sum_{j=1}^n (f_j(z_0) - m_j) \Big) \\
&= \frac{C}{1 - \sum_{k=1}^n m_k} \Big(\sum_{i=1}^n (f_i(z) - m_i) \Big) \Big(1 - \sum_{j=1}^n m_j \Big) \\
&= C \sum_{i=1}^n (f_i(z) - m_i) \leq C \sum_{i=1}^n |f_i(z) - f_i(z_0)| < C \sum_{i=1}^n \frac{\varepsilon}{Cn} = \varepsilon.
\end{aligned}$$

Therefore, $d(f(z), f(z_0)) < \varepsilon$ whenever $z \in U$, so f is continuous at z_0.

□

Our final lemma of this section provides a method for constructing a map from one metric thickening to another, which will be a key ingredient for proving stability. In some sense, this provides a replacement for the notion of simplicial maps between simplicial complexes, which, as we have seen, do not have an exact analog for metric thickenings. Instead of beginning with a function between the underlying metric spaces, which would be the exact analog of a simplicial map, we begin with a map f from one underlying metric space X into a metric thickening L^m. We then check that we can extend f to be defined on a metric thickening of X by mapping measures to the corresponding convex combinations in L^m.

Lemma 6 (Induced maps on metric thickenings)**.** *Let K^m be a metric thickening of X, let L^m be any metric thickening, and suppose $f \colon X \to L^m$ is continuous and bounded. If the map $\tilde{f} \colon K^m \to L^m$ given by $\tilde{f}(\sum_i \lambda_i \delta_{x_i}) = \sum_i \lambda_i f(x_i)$ is well-defined, then it is also continuous.*

The induced maps $\tilde{f}$ described here are a generalization of the induced maps described in section 3.1 of [3]. The proof is almost identical to Lemma 5.2 of [3], but since Euclidean space is replaced by another metric thickening, we will need to use Lemma 3.

Proof. We will write the metrics of X, K^m, and L^m as d_X, d_{K^m}, and d_{L^m}. Since f is bounded, let $C > 0$ be such that $d_{L^m}(f(x), f(y)) < C$ for all $x, y \in X$. Let $\varepsilon > 0$: we show continuity of $\tilde{f}$ at a fixed $\sum_{i=1}^n \lambda_i \delta_{x_i} \in K^m$. By continuity of f at the points $x_1, \ldots, x_n$, there is a $\delta > 0$ such that for each i and any $x \in X$, $d_X(x_i, x) < \delta$ implies $d_{L^m}(f(x_i), f(x)) < \frac{\varepsilon}{2}$. We will further assume $0 < \delta < \frac{\varepsilon}{2C}$ and show that $d_{K^m}\big(\sum_{i=1}^n \lambda_i \delta_{x_i}, \sum_{j=1}^{n'} \lambda'_j \delta_{x'_j}\big) < \delta^2$ implies $d_{L^m}\big(\tilde{f}(\sum_{i=1}^n \lambda_i \delta_{x_i}), \tilde{f}(\sum_{j=1}^{n'} \lambda'_j \delta_{x'_j})\big) < \varepsilon$.

Suppose $\sum_{j=1}^{n'} \lambda'_j \delta_{x'_j} \in K^m$ and that $\{\pi_{i,j}\}$ is a matching between $\sum_{i=1}^n \lambda_i \delta_{x_i}$ and $\sum_{j=1}^{n'} \lambda'_j \delta_{x'_j}$ with $\sum_{i,j} \pi_{i,j} d_X(x_i, x'_j) < \delta^2$. Let $A = \{(i,j) \mid d_X(x_i, x'_j) \geq \delta\}$ and $B = \{(i,j) \mid d_X(x_i, x'_j) < \delta\}$. First, we have

$$\delta \sum_{(i,j)\in A} \pi_{i,j} \leq \sum_{(i,j)\in A} \pi_{i,j} d_X(x_i, x'_j) \leq \sum_{i,j} \pi_{i,j} d_X(x_i, x'_j) < \delta^2.$$

So $\sum_{(i,j)\in A} \pi_{i,j} < \delta$. Thus, applying Lemma 3 and the choice of δ, we have

$$\begin{aligned}
d_{L^m}\Big(\tilde{f}\big(\sum_i \lambda_i \delta_{x_i}\big), \tilde{f}\big(\sum_j \lambda'_j \delta_{x'_j}\big)\Big) &= d_{L^m}\Big(\sum_i \lambda_i f(x_i), \sum_j \lambda'_j f(x'_j)\Big) \\
&= d_{L^m}\Big(\sum_{i,j} \pi_{i,j} f(x_i), \sum_{i,j} \pi_{i,j} f(x'_j)\Big) \\
&\leq \sum_{i,j} \pi_{i,j} d_{L^m}(f(x_i), f(x'_j)) \\
&= \sum_{(i,j)\in A} \pi_{i,j} d_{L^m}(f(x_i), f(x'_j)) + \sum_{(i,j)\in B} \pi_{i,j} d_{L^m}(f(x_i), f(x'_j)) \\
&< \sum_{(i,j)\in A} \pi_{i,j} C + \sum_{(i,j)\in B} \pi_{i,j} \frac{\varepsilon}{2} \leq C\delta + \frac{\varepsilon}{2} < \varepsilon.
\end{aligned}$$

This shows $\tilde{f}$ is continuous at $\sum_{i=1}^n \lambda_i \delta_{x_i}$.

□

Chapter 4

Stability

Given the preparation from the previous chapter, we can now prove that for all totally bounded metric spaces, Vietoris–Rips metric thickenings and both intrinsic and ambient Čech metric thickenings have the same persistence diagrams as their corresponding simplicial complexes. Combining with the stability results for simplicial complexes, this will imply the stability of persistent homology for these metric thickenings. The strategy is to compare metric thickenings of X to metric thickenings of a suitable finite subset F. Maps from the metric thickenings of X to the metric thickenings of F will be constructed using Lemma 6 as induced maps, beginning with maps from X to the metric thickenings of F. The key will be to define maps that distort distances by only a controlled amount, since both Vietoris–Rips and Čech metric thickenings are defined in terms of distances. Since metric thickenings and simplicial complexes agree for finite vertex sets, this comparison to F will let us apply the known results for simplicial complexes.

4.1 Vietoris–Rips Metric Thickenings

Suppose $F = \{f_1, \dots, f_n\}$ is a finite $\frac{\varepsilon}{2}$-sample of a metric space X for some $\varepsilon > 0$. We will assume X is nonempty; the results can be checked separately in the case that X is empty. We will aim to define a map $\varphi\colon X \to \mathrm{VR}^m(F;\varepsilon)$ by $\varphi(x) = \sum_{l=1}^{n} m_l(x)\delta_{f_l}$, where the real-valued functions m_l will be defined below. Since the total mass of a measure in $\mathrm{VR}^m(F;\varepsilon)$ must be 1, we will need the functions m_l to form a partition of unity. We will further define the m_l so that $m_l(x) = 0$ if $d(x, f_l) \geq \frac{\varepsilon}{2}$, which will ensure that each $\varphi(x)$ has a support of diameter less than ε. The m_l will also have the property that $m_l(f_k) = 1$ if $k = l$ and $m_l(f_k) = 0$ if $k \neq l$, which will imply $\varphi(f_l) = \delta_{f_l}$.

To begin, for each $1 \leq l \leq n$, define $w_l\colon X \setminus \{f_l\} \to \mathbb{R}^{\geq 0}$ by

$$w_l(x) = \begin{cases} \frac{\frac{\varepsilon}{2} - d(x,f_l)}{d(x,f_l)} & \text{if } d(x,f_l) \leq \frac{\varepsilon}{2} \\ 0 & \text{if } d(x,f_l) \geq \frac{\varepsilon}{2}. \end{cases}$$

Since each $x \mapsto d(x, f_l)$ is continuous and $d(x, f_l) \neq 0$ for $x \neq f_l$, each w_l is continuous. We note that $\lim_{x \to f_l} w_l(x) = +\infty$.

Next, for each l, define $m_l\colon X \to \mathbb{R}^{\geq 0}$ by

$$m_l(x) = \begin{cases} \frac{w_l(x)}{\sum_{k=1}^n w_k(x)} & \text{if } x \in X \setminus F \\ 1 & \text{if } x = f_l \\ 0 & \text{if } x = f_k \text{ for some } k \neq l. \end{cases}$$

Since F is an $\frac{\varepsilon}{2}$-sample, for any $x \in X \setminus F$, there is a k such that $d(x, f_k) < \frac{\varepsilon}{2}$, and hence $w_k(x) > 0$. Thus, $\sum_{k=1}^n w_k(x) > 0$ for all $x \in X \setminus F$, so m_l is well-defined. Continuity of each w_l shows that each m_l is continuous on $X \setminus F$. Continuity of each m_l at each f_k can be checked by finding $\lim_{x \to f_k} m_l(x)$, using the fact that $\lim_{x \to f_l} w_l(x) = +\infty$. We can also verify $\sum_{l=1}^n m_l(x) = 1$ for all $x \in X$, so the m_l form a partition of unity.

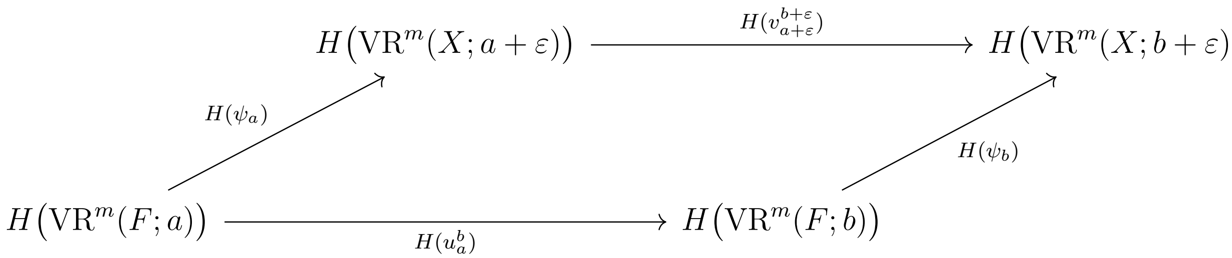
We can now define the map $\varphi\colon X \to \mathrm{VR}^m(F;\varepsilon)$ by setting $\varphi(x) = \sum_{l=1}^n m_l(x)\delta_{f_l}$. To see that φ is well-defined, note that $m_l(x) = 0$ if $d(x, f_l) \geq \frac{\varepsilon}{2}$. So for any $x \in X$, the support of $\varphi(x)$ is contained in an open ball of radius $\frac{\varepsilon}{2}$ centered at x and thus has diameter less than ε. Finally, since the m_l form a partition of unity, the total mass of each $\varphi(x)$ is 1, so φ is well-defined. Continuity of φ follows from Lemma 5 because the m_l form a partition of unity.

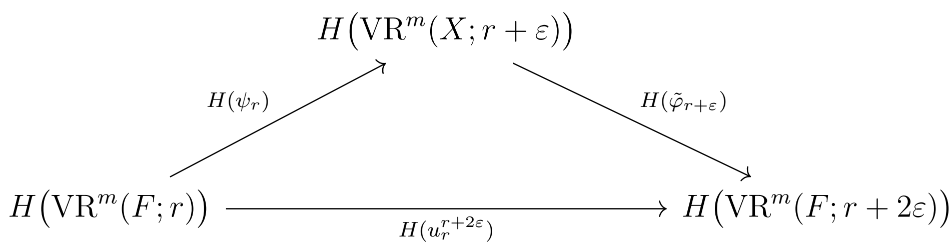
For each $r \geq 0$, let $\varphi_r\colon X \to \mathrm{VR}^m(F; r+\varepsilon)$ be the composition of φ with the inclusion map $\mathrm{VR}^m(F;\varepsilon) \hookrightarrow \mathrm{VR}^m(F; r+\varepsilon)$. We will use Lemma 6 to define an induced map $\tilde{\varphi}_r\colon \mathrm{VR}^m(X;r) \to \mathrm{VR}^m(F; r+\varepsilon)$ for each r. Since each φ_r is continuous and bounded (because F is finite and thus $\mathrm{VR}^m(F; r+\varepsilon)$ is bounded), we only need to check that the induced maps are well-defined. Letting $\sum_i \lambda_i \delta_{x_i} \in \mathrm{VR}^m(X;r)$, the induced map is defined by $\tilde{\varphi}_r(\sum_i \lambda_i \delta_{x_i}) = \sum_i \lambda_i \varphi_r(x_i) = \sum_i \lambda_i \varphi(x_i)$. We know $d(x_j, x_k) \leq r$ for all $x_j, x_k \in \operatorname{supp}(\sum_i \lambda_i \delta_{x_i})$, and for each x_j, the support of $\varphi(x_j)$ is contained in an open ball of radius $\frac{\varepsilon}{2}$ around x_j. Therefore, if y_1 is in the support of $\varphi(x_j)$ and y_2 is in the support of $\varphi(x_k)$, we have $d(y_1, y_2) \leq d(y_1, x_j) + d(x_j, x_k) + d(x_k, y_2) < \frac{\varepsilon}{2} + r + \frac{\varepsilon}{2} = r + \varepsilon$. This shows that the support of $\tilde{\varphi}_r(\sum_i \lambda_i \delta_{x_i})$ has diameter less than $r + \varepsilon$, so $\tilde{\varphi}_r$ is well-defined and is thus continuous by Lemma 6. We will further let $\tilde{\varphi}_r$ be the empty function for all $r < 0$ in the case of the $\leq$ convention and for all $r \leq 0$ in the case of the $<$ convention[4].

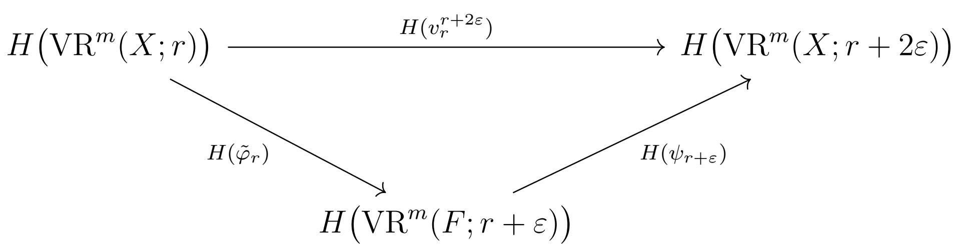
To define an interleaving, let $\psi_r \colon \mathrm{VR}^m(F; r) \hookrightarrow \mathrm{VR}^m(X; r + \varepsilon)$ be the inclusion (this is also empty if $\mathrm{VR}^m(F; r)$ is empty). We also let $v_a^b \colon \mathrm{VR}^m(X; a) \hookrightarrow \mathrm{VR}^m(X; b)$ and $u_a^b \colon \mathrm{VR}^m(F; a) \hookrightarrow \mathrm{VR}^m(F; b)$ be the inclusions for any $a \leq b$. The families of maps $\{\tilde{\varphi}_r\}_{r \in \mathbb{R}}$ and $\{\psi_r\}_{r \in \mathbb{R}}$ do not necessarily commute with the inclusions, but we will show that they commute up to homotopy. This will be enough to show that the induced maps on homology commute, since homotopic maps on spaces give equal maps on homology vector spaces. We will thus obtain the ε-interleaving shown in the diagrams (for the general requirements for maps to form an interleaving, refer to Definition 7).

$$
\begin{array}{ccccc}
H\big(\mathrm{VR}^m(X;a)\big) & & \xrightarrow{H(v_a^b)} & & H\big(\mathrm{VR}^m(X;b)\big) \\
& \searrow^{H(\tilde{\varphi}_a)} & & & & \searrow^{H(\tilde{\varphi}_b)} \\
& & H\big(\mathrm{VR}^m(F;a+\varepsilon)\big) & \xrightarrow[H(u_{a+\varepsilon}^{b+\varepsilon})]{} & & & H\big(\mathrm{VR}^m(F;b+\varepsilon)\big)
\end{array}
$$

[4] This is where it is important to have $\mathrm{VR}^m(X; r)$ be empty for negative parameters. If instead we had let Vietoris–Rips complexes consist of the vertex set for negative r, we would have $\mathrm{VR}^m(X; r) \cong X$ and $\mathrm{VR}^m(F; r) \cong F$ for negative r. To obtain maps that commute, as described shortly, this would then require us to find a continuous map $X \to F$ that is the identity on F, which in general does not exist.







Any maps originating from the empty set are empty, in which case there is nothing to check, so we will assume in each case that the composition of maps originates from a nonempty space. We can first check that for any $a \leq b$, we have

$$\tilde{\varphi}_b\Big(v_a^b\Big(\sum_i \lambda_i \delta_{x_i}\Big)\Big) = \tilde{\varphi}_b\Big(\sum_i \lambda_i \delta_{x_i}\Big) = \sum_i \lambda_i \varphi(x_i) = u_{a+\varepsilon}^{b+\varepsilon}\Big(\tilde{\varphi}_a\Big(\sum_i \lambda_i \delta_{x_i}\Big)\Big).$$

So $\tilde{\varphi}_b \circ v_a^b = u_{a+\varepsilon}^{b+\varepsilon} \circ \tilde{\varphi}_a$. We also have $\psi_b \circ u_a^b = v_{a+\varepsilon}^{b+\varepsilon} \circ \psi_a$ since each map involved is an inclusion.

Next, if $\sum_{k=1}^n \lambda_k \delta_{f_k} \in \mathrm{VR}^m(F;r)$, we have

$$\tilde{\varphi}_{r+\varepsilon}\Big(\psi_r\Big(\sum_k \lambda_k \delta_{f_k}\Big)\Big) = \tilde{\varphi}_{r+\varepsilon}\Big(\sum_k \lambda_k \delta_{f_k}\Big) = \sum_k \lambda_k \varphi(f_k) = \sum_k \lambda_k \sum_{l=1}^n m_l(f_k)\delta_{f_l} = \sum_k \lambda_k \delta_{f_k},$$

so $\tilde{\varphi}_{r+\varepsilon} \circ \psi_r = u_r^{r+2\varepsilon}$. Finally, we will show that $\psi_{r+\varepsilon} \circ \tilde{\varphi}_r \simeq v_r^{r+2\varepsilon}$. For any $\sum_i \lambda_i \delta_{x_i} \in \mathrm{VR}^m(X;r)$, we have $\psi_{r+\varepsilon} \circ \tilde{\varphi}_r(\sum_i \lambda_i \delta_{x_i}) = \sum_i \lambda_i \varphi(x_i)$. We check that $\mathrm{supp}(\sum_i \lambda_i \delta_{x_i}) \cup \mathrm{supp}(\sum_i \lambda_i \varphi(x_i))$ has diameter less than $r + 2\varepsilon$. We know the distance between two points in $\mathrm{supp}(\sum_i \lambda_i \delta_{x_i})$ is at most r, since $\sum_i \lambda_i \delta_{x_i} \in \mathrm{VR}^m(X;r)$, and similarly the distance between any two points in

$\mathrm{supp}(\sum_i \lambda_i \varphi(x_i))$ is at most $r + \varepsilon$. And if $x_k \in \mathrm{supp}(\sum_i \lambda_i \delta_{x_i})$ and $y \in \mathrm{supp}(\sum_i \lambda_i \varphi(x_i))$, then $y \in \mathrm{supp}(\varphi(x_j))$ for some j, and thus $d(x_k, y) \leq d(x_k, x_j) + d(x_j, y) < r + \frac{\varepsilon}{2}$. Therefore $\mathrm{supp}(\sum_i \lambda_i \delta_{x_i}) \cup \mathrm{supp}(\sum_i \lambda_i \varphi(x_i))$ has diameter less than $r + 2\varepsilon$, which means that a straight line homotopy between $\psi_{r+\varepsilon} \circ \tilde{\varphi}_r$ and $v_r^{r+2\varepsilon}$ is well-defined. So by Lemma 4, $\psi_{r+\varepsilon} \circ \tilde{\varphi}_r \simeq v_r^{r+2\varepsilon}$.

Letting H be homology in any dimension over a fixed field, the facts above show that the induced maps $H(\tilde{\varphi}_r)$ and $H(\psi_r)$ define an ε-interleaving of the persistence modules $H(\mathrm{VR}^m(X; \_))$ and $H(\mathrm{VR}^m(F; \_))$. We have used the fact that $H(\psi_{r+\varepsilon} \circ \tilde{\varphi}_r) = H(v_r^{r+2\varepsilon})$, since $\psi_{r+\varepsilon} \circ \tilde{\varphi}_r \simeq v_r^{r+2\varepsilon}$. We state this as the following lemma:

Lemma 7. *If F is a finite $\frac{\varepsilon}{2}$-sample of a metric space X for some $\varepsilon > 0$, then $H(\mathrm{VR}^m(X; \_))$ and $H(\mathrm{VR}^m(F; \_))$ are ε-interleaved.*

This quickly leads to the following results:

Proposition 4. *If X is a totally bounded metric space, then $H(\mathrm{VR}^m(X; \_))$ is q-tame.*

Proof. We follow the method used in [5] to prove Proposition 1 (this is Proposition 5.1 of [5]). As above, we let $v_a^b \colon \mathrm{VR}^m(X; a) \to \mathrm{VR}^m(X; b)$ be the inclusion for any $a \leq b$. We must show that for any $a < b$, the map $H(v_a^b) \colon H(\mathrm{VR}^m(X; a)) \to H(\mathrm{VR}^m(X; b))$ has finite rank. Let $\varepsilon = \frac{b-a}{2}$. Since X is totally bounded, there exists a finite $\frac{\varepsilon}{2}$-sample F, and by the theorem, $H(\mathrm{VR}^m(X; \_))$ and $H(\mathrm{VR}^m(F; \_))$ are ε-interleaved. Using the interleaving maps, $H(v_a^b)$ factors as

$$H(\mathrm{VR}^m(X; a)) \longrightarrow H(\mathrm{VR}^m(F; a + \varepsilon)) \longrightarrow H(\mathrm{VR}^m(X; b)).$$

Since F is finite and $\mathrm{VR}^m(F; a + \varepsilon)$ is homeomorphic to $\mathrm{VR}(F; a + \varepsilon)$ by Proposition 3, $H(\mathrm{VR}^m(F; a + \varepsilon))$ has finite dimension, so $H(v_a^b)$ must have finite rank. □

This proposition implies that if X is totally bounded, then $\mathrm{dgm}\big(H(\mathrm{VR}^m(X; \_))\big)$ is well-defined, by Theorem 1.

Theorem 4. *If X is a totally bounded metric space, then $H(\mathrm{VR}^m(X; \_))$ and $H(\mathrm{VR}(X; \_))$ have identical persistence diagrams.*

Proof. For any $\varepsilon_1 > 0$, X has a finite $\frac{\varepsilon_1}{2}$-sample F, and Lemma 7 shows that $H(\mathrm{VR}^m(X;\_))$ and $H(\mathrm{VR}^m(F;\_))$ are ε_1-interleaved. Furthermore, $H(\mathrm{VR}^m(F;\_))$ and $H(\mathrm{VR}(F;\_))$ are 0-interleaved (in this case they are called isomorphic) by Proposition 3. The set $\{(x,f) \in X \times F \mid d(x,f) < \frac{\varepsilon_1}{2}\}$ is a correspondence[5] with distortion at most ε_1, so $d_{GH}(X,F) \leq \frac{\varepsilon_1}{2}$. By Lemma 1, $H(\mathrm{VR}(F;\_))$ and $H(\mathrm{VR}(X;\_))$ are ε_2-interleaved for any $\varepsilon_2 > \varepsilon_1$. So $H(\mathrm{VR}^m(X;\_))$ and $H(\mathrm{VR}(X;\_))$ are $(\varepsilon_1+\varepsilon_2)$-interleaved, and since $\varepsilon_1 > 0$ was arbitrary, $H(\mathrm{VR}^m(X;\_))$ and $H(\mathrm{VR}(X;\_))$ are ε-interleaved for any $\varepsilon > 0$. Thus, the bottleneck distance between their persistence diagrams is 0. The fact that the persistence diagrams are truly identical follows from Theorem 4.20 of [9], where we observe that $H(\mathrm{VR}^m(X;\_))$ is q-tame by Proposition 4 and that $H(\mathrm{VR}(X;\_))$ is q-tame by Proposition 1. □

Theorem 5 (Persistence Stability for Vietoris–Rips Metric Thickenings)**.** *If X and Y are totally bounded metric spaces, then*

$$\mathrm{d_b}\Big(\mathrm{dgm}\big(H(\mathrm{VR}^m(X;\_))\big), \mathrm{dgm}\big(H(\mathrm{VR}^m(Y;\_))\big)\Big) \leq 2d_{GH}(X,Y).$$

Proof. The persistence diagrams of these metric thickenings are equal to the persistence diagrams of their corresponding simplicial complexes by Theorem 4. Therefore this bound follows from Theorem 2. □

While our proof of stability (Theorem 5) relied on Theorem 4, this is not strictly necessary. We alternately could have started with totally bounded X and Y and finite $\frac{\varepsilon}{2}$-samples F_X and F_Y, obtaining the interleavings given by Lemma 7. Then an interleaving between $H(\mathrm{VR}^m(F_X;\_))$ and $H(\mathrm{VR}^m(F_Y;\_))$ could be found using Proposition 3 and Lemma 1. This would give an interleaving between $H(\mathrm{VR}^m(X;\_))$ and $H(\mathrm{VR}^m(Y;\_))$, then applying Theorem 1 and letting ε approach 0 would give the same result.

However, Theorem 4 is interesting in its own right, as it establishes a strong relationship between Vietoris–Rips complexes and metric thickenings. From the viewpoint of persistent homol-

[5]This is used in the proof of Proposition 5.1 of [5] and establishes the expected relationship between the persistent homology of a space and that of a finite sample.

ogy, this further justifies the idea that Vietoris–Rips complexes can be better understood by studying the corresponding metric thickenings.

4.2 Intrinsic Čech Metric Thickenings

We will prove stability for intrinsic Čech metric thickenings using the same overall technique as we used for Vietoris–Rips metric thickenings; we only need to verify that the analogous maps used in the proof are well-defined.

We begin again with $F = \{f_1, \ldots, f_n\}$, a finite $\frac{\varepsilon}{2}$-sample of a metric space X for some $\varepsilon > 0$. We use the functions m_l, defined in the same way as before, and recall the important properties here. Each m_l is continuous, and we have $\sum_{l=1}^{n} m_l(x) = 1$ for all $x \in X$. For each l, $m_l(f_k) = 1$ if $k = l$ and $m_l(f_k) = 0$ if $k \neq l$. For any $x \in X$ and any l, $m_l(x) = 0$ if $d(x, f_l) \geq \frac{\varepsilon}{2}$.

We define $\varphi\colon X \to \check{\mathrm{C}}^m(F;\varepsilon)$ by setting $\varphi(x) = \sum_{l=1}^{n} m_l(x)\delta_{f_l}$. To show φ is well-defined, we only need to check that the support of $\varphi(x)$ is a simplex in the Čech complex. Let $x \in X$, and since F is an $\frac{\varepsilon}{2}$-sample of X, let $c \in F$ satisfy $d(x, c) < \frac{\varepsilon}{2}$. Then if $y \in \mathrm{supp}(\varphi(x))$, we have $d(y, c) \leq d(y, x) + d(x, c) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$, so c is an ε-center for the simplex $\mathrm{supp}(\varphi(x))$. Therefore φ is well-defined. Continuity of φ follows from Lemma 5 because the m_l form a partition of unity.

For each $r \geq 0$, let $\varphi_r\colon X \to \check{\mathrm{C}}^m(F; r+\varepsilon)$ be the composition of φ with the inclusion $\check{\mathrm{C}}^m(F;\varepsilon) \hookrightarrow \check{\mathrm{C}}^m(F; r+\varepsilon)$. As before we let $\tilde{\varphi}_r\colon \check{\mathrm{C}}^m(X;r) \to \check{\mathrm{C}}^m(F; r+\varepsilon)$ be the induced maps, and we must check that these are well-defined. If $\sum_i \lambda_i\delta_{x_i} \in \check{\mathrm{C}}^m(X;r)$, then there exists a center $c \in X$ such that $d(x_i, c) \leq r$ for all i. Since F is an $\frac{\varepsilon}{2}$-sample of X, choose $z \in F$ such that $d(c, z) < \frac{\varepsilon}{2}$. Then for any y in the support of $\tilde{\varphi}_r(\sum_i \lambda_i\delta_{x_i}) = \sum_i \lambda_i\varphi(x_i)$, y must be in the support of $\varphi(x_i)$ for some i. Thus, $d(y, z) \leq d(y, x_i) + d(x_i, c) + d(c, z) < \frac{\varepsilon}{2} + r + \frac{\varepsilon}{2} = r + \varepsilon$. This shows that z is a center for $\mathrm{supp}(\tilde{\varphi}_r(\sum_i \lambda_i\delta_{x_i}))$, so $\tilde{\varphi}_r$ is well-defined, and by Lemma 6 it is continuous.

To define an interleaving, let $\psi_r\colon \check{\mathrm{C}}^m(F;r) \hookrightarrow \check{\mathrm{C}}^m(X; r+\varepsilon)$ be the inclusion. Also let $v_a^b\colon \check{\mathrm{C}}^m(X;a) \hookrightarrow \check{\mathrm{C}}^m(X;b)$ and $u_a^b\colon \check{\mathrm{C}}^m(F;a) \hookrightarrow \check{\mathrm{C}}^m(F;b)$ be the inclusions for any $a \leq b$. As before, the families of maps $\{\tilde{\varphi}_r\}_{r\in\mathbb{R}}$ and $\{\psi_r\}_{r\in\mathbb{R}}$ will commute with the inclusions up to homotopy, resulting in an ε-interleaving on the persistence modules. The diagrams for the interleaving are the same as those in section 4.1, but with Čech metric thickenings instead of Vietoris–Rips metric thickenings. We verify $\psi_{r+\varepsilon} \circ \tilde{\varphi}_r \simeq v_r^{r+2\varepsilon}$; the other three conditions can be checked by the same arguments used in section 4.1 (and are equalities instead of homotopy equivalences, as before). If $\sum_i \lambda_i \delta_{x_i} \in \check{\mathrm{C}}^m(X; r)$, then there exists a $c \in X$ such that $d(x_i, c) \le r$ for all i. For any y in the support of $\psi_{r+\varepsilon}(\tilde{\varphi}_r(\sum_i \lambda_i \delta_{x_i})) = \sum_i \lambda_i \varphi(x_i)$, we know y is in the support of $\varphi(x_i)$ for some i. Then $d(y, c) \le d(y, x_i) + d(x_i, c) < r + \frac{\varepsilon}{2}$. Thus, every point in $\mathrm{supp}\big(v_r^{r+2\varepsilon}(\sum_i \lambda_i \delta_{x_i})\big) \cup \mathrm{supp}\big(\psi_{r+\varepsilon}(\tilde{\varphi}_r(\sum_i \lambda_i \delta_{x_i}))\big)$ is within a distance of $r + 2\varepsilon$ of c, so this forms a simplex in the Čech complex with parameter $r + 2\varepsilon$. Therefore a straight line homotopy from $\psi_{r+\varepsilon} \circ \tilde{\varphi}_r$ to $v_r^{r+2\varepsilon}$ is well-defined, so by Lemma 4, $\psi_{r+\varepsilon} \circ \tilde{\varphi}_r \simeq v_r^{r+2\varepsilon}$.

As before, applying H gives an ε-interleaving between the persistence modules, so we have the following lemma:

Lemma 8. *If F is a finite $\frac{\varepsilon}{2}$-sample of a metric space X for some $\varepsilon > 0$, then $H(\check{\mathrm{C}}^m(X; \_))$ and $H(\check{\mathrm{C}}^m(F; \_))$ are ε-interleaved.*

The proofs of the following results follow the same methods as the proofs of Proposition 4, Theorem 4, and Theorem 5.

Proposition 5. *If X is a totally bounded metric space, then $H(\check{\mathrm{C}}^m(X; \_))$ is q-tame.*

Theorem 6. *If X is a totally bounded metric space, then $H(\check{\mathrm{C}}^m(X; \_))$ and $H(\check{\mathrm{C}}(X; \_))$ have identical persistence diagrams.*

Theorem 7 (Persistence Stability for Intrinsic Čech Metric Thickenings)**.** *If X and Y are totally bounded metric spaces, then*

$$\mathrm{d_b}\Big(\mathrm{dgm}\big(H(\check{\mathrm{C}}^m(X; \_))\big), \mathrm{dgm}\big(H(\check{\mathrm{C}}^m(Y; \_))\big)\Big) \le 2 d_{GH}(X, Y).$$

4.3 Ambient Čech Metric Thickenings

The proof of stability for ambient Čech metric thickenings will follow the same outline, with only minor adjustments. Suppose L and W are subsets of some ambient metric space, and suppose

$F = \{f_1, \ldots, f_n\}$ is a finite ε-sample of L. For ambient Čech metric thickenings, the technique used before will not be able to define a map $L \to \check{\mathrm{C}}^m(F, W; \varepsilon)$ because there may be an $f \in F$ such that δ_f is not in $\check{\mathrm{C}}(F, W; \varepsilon)$. We will instead start by letting L_r be the vertex set of $\check{\mathrm{C}}(L, W; r)$ for each r, recalling that the vertex set may be a strict subset of L. Then we will define a map $L_r \to \check{\mathrm{C}}^m(F, W; r + \varepsilon)$ for each r using the previous approach.

For each $1 \le l \le n$, we define $w_l \colon L \setminus \{f_l\} \to \mathbb{R}^{\ge 0}$ by

$$w_l(x) = \begin{cases} \frac{\varepsilon - d(x, f_l)}{d(x, f_l)} & \text{if } d(x, f_l) \le \varepsilon \\ 0 & \text{if } d(x, f_l) \ge \varepsilon, \end{cases}$$

and define $m_l \colon L \to \mathbb{R}^{\ge 0}$ by

$$m_l(x) = \begin{cases} \frac{w_l(x)}{\sum_{k=1}^n w_k(x)} & \text{if } x \in L \setminus F \\ 1 & \text{if } x = f_l \\ 0 & \text{if } x = f_k \text{ for some } k \ne l. \end{cases}$$

Note we have replaced the $\frac{\varepsilon}{2}$ in the definitions of the previous sections with ε. We recall the important properties here. Each m_l is continuous, and we have $\sum_{l=1}^n m_l(x) = 1$ for all $x \in L$. For each l, $m_l(f_k) = 1$ if $k = l$ and $m_l(f_k) = 0$ if $k \ne l$. For any $x \in L$ and any l, $m_l(x) = 0$ if $d(x, f_l) \ge \varepsilon$.

For each r, define $\varphi_r \colon L_r \to \check{\mathrm{C}}^m(F, W; r+\varepsilon)$ by setting $\varphi_r(x) = \sum_{l=1}^n m_l(x)\delta_{f_l}$. We show φ_r is well-defined: we only need to check that the support of each $\varphi_r(x)$ is a simplex in $\check{\mathrm{C}}(F, W; r+\varepsilon)$. Let $x \in L_r$, so that there exists a $w \in W$ such that $d(x, w) \le r$. Then if y is in the support of $\varphi_r(x)$, we have $d(y, w) \le d(y, x) + d(x, w) < r + \varepsilon$. Therefore w forms a center for $\varphi_r(x)$, so φ_r is well-defined. Continuity of φ_r follows from Lemma 5. Note that for each r and each $x \in L_r$, the support of $\varphi_r(x)$ is contained in an open ball of radius ε centered at x. Also note that if $r_1 < r_2$, then for any $x \in L_{r_1} \subseteq L_{r_2}$, we have $\varphi_{r_1}(x) = \varphi_{r_2}(x)$ in $\check{\mathrm{C}}^m(F, W; r_2 + \varepsilon)$.

We now check that the induced maps $\tilde{\varphi}_r\colon \check{C}^m(L,W;r) \to \check{C}^m(F,W;r+\varepsilon)$ are well-defined. If $\sum_i \lambda_i\delta_{x_i} \in \check{C}^m(L,W;r)$, then there exists a $w \in W$ such that $d(x_i,w) \le r$ for all i. Then for any y in the support of $\tilde{\varphi}_r(\sum_i \lambda_i\delta_{x_i}) = \sum_i \lambda_i\varphi_r(x_i)$, y must be in the support of $\varphi_r(x_i)$ for some i. Thus, $d(y,w) \le d(y,x_i) + d(x_i,w) < r+\varepsilon$, so w is an $(r+\varepsilon)$-center for $\tilde{\varphi}_r(\sum_i \lambda_i\delta_{x_i})$. This shows $\tilde{\varphi}_r$ is well-defined, and by Lemma 6 it is continuous. We will note that φ_r is the empty function when L_r is empty, and in this case the induced map $\tilde{\varphi}_r$ is empty as well.

To define an interleaving, let $\psi_r\colon \check{C}^m(F,W;r) \hookrightarrow \check{C}^m(L,W;r+\varepsilon)$ be the inclusion. Also let $v_a^b\colon \check{C}^m(L,W;a) \hookrightarrow \check{C}^m(L,W;b)$ and $u_a^b\colon \check{C}^m(F,W;a) \hookrightarrow \check{C}^m(F,W;b)$ be the inclusions for any $a \le b$. We show that the families of maps $\{\tilde{\varphi}_r\}_{r\in\mathbb{R}}$ and $\{\psi_r\}_{r\in\mathbb{R}}$ commute with the inclusions up to homotopy. If $a \le b$ and $\sum_i \lambda_i\delta_{x_i} \in \check{C}^m(L,W;a)$, then for each i, $x_i \in L_a \subseteq L_b$, so $\varphi_a(x) = \varphi_b(x)$. Thus,

$$\tilde{\varphi}_b\Big(v_a^b\Big(\sum_i \lambda_i\delta_{x_i}\Big)\Big) = \sum_i \lambda_i\varphi_b(x_i) = \sum_i \lambda_i\varphi_a(x_i) = u_{a+\varepsilon}^{b+\varepsilon}\Big(\tilde{\varphi}_a\Big(\sum_i \lambda_i\delta_{x_i}\Big)\Big).$$

So $\tilde{\varphi}_b \circ v_a^b = u_{a+\varepsilon}^{b+\varepsilon} \circ \tilde{\varphi}_a$. We also have $\psi_b \circ u_a^b = v_{a+\varepsilon}^{b+\varepsilon} \circ \psi_a$ and $\tilde{\varphi}_{r+\varepsilon} \circ \psi_r = u_r^{r+2\varepsilon}$ following the methods from the previous sections. Finally, we will show that $\psi_{r+\varepsilon} \circ \tilde{\varphi}_r \simeq v_r^{r+2\varepsilon}$. If $\sum_i \lambda_i\delta_{x_i} \in \check{C}^m(L,W;r)$, then there exists a $w \in W$ such that $d(x_i,w) \le r$ for all i. For any y in the support of $\psi_{r+\varepsilon}(\tilde{\varphi}_r(\sum_i \lambda_i\delta_{x_i})) = \sum_i \lambda_i\varphi_r(x_i)$, we know y is in the support of $\varphi_r(x_i)$ for some i. Then $d(y,w) \le d(y,x_i) + d(x_i,w) < r+\varepsilon$. Therefore every point in $\operatorname{supp}\big(v_r^{r+2\varepsilon}(\sum_i \lambda_i\delta_{x_i})\big) \cup \operatorname{supp}\big(\psi_{r+\varepsilon}(\tilde{\varphi}_r(\sum_i \lambda_i\delta_{x_i}))\big)$ is within a distance of $r+\varepsilon$ of w, so this forms a simplex in the Čech complex with parameter $r+2\varepsilon$. Therefore a straight line homotopy from $\psi_{r+\varepsilon} \circ \tilde{\varphi}_r$ to $v_r^{r+2\varepsilon}$ is well-defined, so by Lemma 4, $\psi_{r+\varepsilon} \circ \tilde{\varphi}_r \simeq v_r^{r+2\varepsilon}$.

As before, applying H gives an ε-interleaving between the persistence modules, so we have the following lemma:

Lemma 9. *Suppose L and W are subsets of some ambient metric space and F is a finite ε-sample of L for some $\varepsilon > 0$. Then $H(\check{C}^m(L,W;\_))$ and $H(\check{C}^m(F,W;\_))$ are ε-interleaved.*

For the following, the proof of Proposition 6 follows the method for Proposition 4 and is omitted, and the proofs given for Theorems 8 and 9 are similar to those for Theorems 4 and 5.

Proposition 6. *Suppose L and W are subsets of some ambient metric space. If L is totally bounded, then $H(\check{\mathrm{C}}^m(L, W; \_))$ is q-tame.*

Theorem 8. *Suppose L and W are subsets of some ambient metric space. If L is totally bounded, then the persistence diagrams for $H(\check{\mathrm{C}}^m(L, W; \_))$ and $H(\check{\mathrm{C}}(L, W; \_))$ are identical.*

Proof. For any $\varepsilon_1 > 0$, L has a finite ε_1-sample F, and Lemma 9 shows that $H(\check{\mathrm{C}}^m(L, W; \_))$ and $H(\check{\mathrm{C}}^m(F, W; \_))$ are ε_1-interleaved. Furthermore, $H(\check{\mathrm{C}}^m(F, W; \_))$ and $H(\check{\mathrm{C}}(F, W; \_))$ are 0-interleaved by Proposition 3. Since $d_H(L, F) \leq \varepsilon_1$, by Lemma 2, $H(\check{\mathrm{C}}(F, W; \_))$ and $H(\check{\mathrm{C}}(L, W; \_))$ are ε_2-interleaved for any $\varepsilon_2 > \varepsilon_1$. So $H(\check{\mathrm{C}}^m(L, W; \_))$ and $H(\check{\mathrm{C}}(L, W; \_))$ are $(\varepsilon_1 + \varepsilon_2)$-interleaved, and since $\varepsilon_1 > 0$ was arbitrary, $H(\check{\mathrm{C}}^m(L, W; \_))$ and $H(\check{\mathrm{C}}(L, W; \_))$ are ε-interleaved for any $\varepsilon > 0$. The result now follows from Theorem 4.20 of [9], where we observe that $H(\check{\mathrm{C}}^m(L, W; \_))$ is q-tame by Proposition 6 and that $H(\check{\mathrm{C}}(L, W; \_))$ is q-tame by Proposition 2. □

Theorem 9 (Persistence Stability for Ambient Čech Metric Thickenings)**.** *Suppose L, L' and W are subsets of some ambient metric space. If L and L' are totally bounded, then*

$$\mathrm{d_b}\Big(\mathrm{dgm}\big(H(\check{\mathrm{C}}^m(L, W; \_))\big), \mathrm{dgm}\big(H(\check{\mathrm{C}}^m(L', W; \_))\big)\Big) \leq d_H(L, L').$$

Proof. This follows from Theorem 8 and Theorem 3. □

For each type of metric thickening we have considered, there is nothing particularly special about the construction of the maps φ_r. They just have the important property that they only distort points in a controlled fashion, which is what allows the induced maps to only increase the parameter by ε. This is analogous to the use of correspondences with bounded distortion in the proof of Lemma 1. An interesting feature of these proofs of stability for metric thickenings is the reliance on finite samples in the construction of interleavings. This is in contrast with the case of simplicial complexes, in which Lemmas 1 and 2 were obtained without assuming the spaces had finite samples, and the requirement of totally bounded spaces came from Propositions 1 and 2.

It should be noted that Theorem 9 does not imply Theorem 7, even though intrinsic Čech metric thickenings are special cases of ambient Čech metric thickenings. This is because in Theorem 9, the two metric thickenings must have the same witness set and the two landmark sets must be in subsets of the same space. Thus, Theorem 9 does not give a way to compare the persistent homology of $\check{\mathrm{C}}^m(X, X; \_)$ and $\check{\mathrm{C}}^m(Y, Y; \_)$ if X and Y are unrelated metric spaces.

Chapter 5

Conclusion

The theorems proved here give further evidence that metric thickenings could be helpful tools for understanding simplicial complexes and persistent homology. In particular, Vietoris–Rips and Čech metric thickenings of totally bounded metric spaces are essentially equivalent to their corresponding simplicial complexes from the point of view of persistent homology, as shown by Theorems 4, 6, and 8. These theorems imply the stability results stated in Theorems 5, 7, and 9. Simplicial complexes and metric thickenings built on infinite metric spaces can be seen as tools used to improve our understanding of those built on finite metric spaces, such as those that arise in applications. From this point of view, since metric thickenings of finite metric spaces agree with their corresponding simplicial complexes, metric thickenings of infinite metric spaces provide an alternate method for examining finite simplicial complexes. Furthermore, the stability theorems proved above validate this method in the context of persistent homology.

We will finish by outlining some potential future work on metric thickenings and their relationship with simplicial complexes. There has been recent work to determine homotopy types of both simplicial complexes and metric thickenings. As mentioned, some of this work has been on the Vietoris–Rips complexes and metric thickenings of spheres. This work is likely to continue, and current results on metric thickenings are promising because metric thickenings allow for relatively simple homotopies to be used (see again section 5 of [3]). If we ask only for the persistent homology, then either Vietoris–Rips complexes or metric thickenings will give the same answer by Theorem 4. As an example, we note that the persistent homology of the Vietoris–Rips metric thickenings of the circle is now known, because it is known for the corresponding simplicial complexes by the work in [1].

More directly related to our work, there is a slight difference between Theorem 9 and Theorem 3, which we have cited from [5]. Theorem 3 supposes that either L and L' are totally bounded, or that W is totally bounded, whereas in Theorem 9, we only consider the case where L and L' are totally bounded. In [5], the result in the case that W is totally bounded is obtained from the other case using what the authors call *Dowker duality*, which is an extension of Dowker's theorem on simplicial complexes to filtered simplicial complexes. Future work could examine whether there are versions of Dowker's theorem and Dowker duality for metric thickenings.

Finally, work on these questions and others may require more elaborate maps on metric thickenings. In our work, we have made use of the induced maps defined in Lemma 6, which allowed us to construct the maps used in the proofs of the main theorems. New techniques could lead to a better ability to construct maps between metric thickenings and filtrations of metric thickenings. We should also note that the only homotopies we used here were straight line homotopies. These are certainly not the only possible homotopies in metric thickenings: for instance, it is possible to define homotopies that continuously move the points in the support of a measure. More flexibility in the construction of homotopies may be key to new results on homotopy types of metric thickenings.

There remains plenty of work to be done in these areas. Here we have demonstrated further connections between metric thickenings and simplicial complexes, suggesting that the two complement each other. Hopefully, these results will inspire further work exploring the connections between simplicial complexes and metric thickenings and leading to a better understanding of their persistent homology.

Bibliography

[1] Michal Adamaszek and Henry Adams. The Vietoris–Rips complexes of a circle. *Pacific Journal of Mathematics*, 290:1–40, July 2017.

[2] Michał Adamaszek. Clique complexes and graph powers. *Israel Journal of Mathematics*, 196(1):295–319, Aug 2013.

[3] Michał Adamaszek, Henry Adams, and Florian Frick. Metric reconstruction via optimal transport. *SIAM Journal on Applied Algebra and Geometry*, 2(4):597–619, 2018.

[4] Henry Adams, Johnathan Bush, and Florian Frick. Metric thickenings, Borsuk–Ulam theorems, and orbitopes. *Mathematika*, 66(1):79–102, January 2020.

[5] Frédéric Chazal, Vin de Silva, and Steve Oudot. Persistence stability for geometric complexes. *Geometriae Dedicata*, 173(1):193–214, 2014.

[6] Herbert Edelsbrunner and John Harer. *Computational Topology - an Introduction.* American Mathematical Society, 2010.

[7] Robert Ghrist. Barcodes: The persistent topology of data. *Bulletin of the American Mathematical Society*, 45:61–75, January 2008.

[8] Afra Zomorodian and Gunnar Carlsson. Computing persistent homology. *Discrete & Computational Geometry*, 33:249–274, 2005.

[9] Frederic Chazal, Vin de Silva, Marc Glisse, and Steve Oudot. The structure and stability of persistence modules, 2013.

[10] Frédéric Chazal, David Cohen-Steiner, Leonidas J. Guibas, Facundo Mémoli, and Steve Y. Oudot. Gromov–Hausdorff stable signatures for shapes using persistence. In *Computer Graphics Forum*, volume 28, pages 1393–1403, 2009.

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

[12] Frédéric Chazal, David Cohen-Steiner, Marc Glisse, Leonidas Guibas, and Steve Oudot. Proximity of persistence modules and their diagrams. *Proceedings of the twenty-fifth annual symposium on Computational geometry*, pages 237–246, June 2009.

[13] Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A Course in Metric Geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001.