Communicated by Herbert Edelsbrunner

Foundations of Computational Mathematics

SLIDING WINDOWS AND PERSISTENCE: AN APPLICATION OF TOPOLOGICAL METHODS TO SIGNAL ANALYSIS

JOSE A. PEREA (✉) AND JOHN HARER

Abstract. We develop in this paper a theoretical framework for the topological study of time series data. Broadly speaking, we describe geometrical and topological properties of sliding window embeddings, as seen through the lens of persistent homology. In particular, we show that maximum persistence at the point-cloud level can be used to quantify periodicity at the signal level, prove structural and convergence theorems for the resulting persistence diagrams, and derive estimates for their dependency on window size and embedding dimension. We apply this methodology to quantifying periodicity in synthetic data sets, and compare the results with those obtained using state-of-the-art methods in gene expression analysis. We call this new method **SW1PerS** which stands for Sliding Windows and 1-dimensional Persistence Scoring.

1. Introduction

Signal analysis is an enormous field. There are many methods to study signals and many applications of that study. Given its importance, one might conclude that there is little opportunity left for the development of totally new approaches to signals. Yet in this paper we provide a new way to find periodicity and quasi-periodicity in signals. The method is based on sliding windows (also known as time-delay reconstruction), which have been used extensively in both engineering applications and in dynamical systems. But it adds a new element not applied before, which comes from the new field of computational topology [12].

Persistent homology is a topological method for measuring the shapes of spaces and the features of functions. One of the most important applications of persistent homology is to point clouds [3], where shape is usually interpreted as the geometry of some implicit underlying object near which the point cloud is sampled. The simplest non-trivial example of this idea is a point cloud which has the shape of a circle, and this shape is captured with 1-dimensional persistence. The challenge in applying the method is that noise can reduce the persistence, and not enough points can prevent the circular shape from appearing. It's also a challenge to deal

2000 *Mathematics Subject Classification.* Primary 55U99, 37M10, 68W05; Secondary 57M99.
Key words and phrases. Persistent homology, time-delay embeddings, periodicity.
(✉) Corresponding author. Email: joperea@math.duke.edu. Phone: +1 (919) 660 - 2837.
Both authors were supported in part by DARPA under grants D12AP00001, D12AP00025-002, and by the AFOSR under grant FA9550-10-1-0436.

with the fact that features come on all scale-levels and can be nested or in more complicated relationships. But this is what persistent homology is all about.

The idea of applying 1D persistence to study time series arose in our study of gene expression data [11, 23]. The first of these papers studied a variety of existing methods for finding periodicity in gene expression patterns. The motivation of that work was the search for gene regulatory networks (more precisely, possible nodes of gene regulatory networks) that control periodic processes in cells such as the cell division cycle, circadian rhythms, metabolic cycles and periodic patterning in biological development (lateral roots and somites). The methods studied in [11] were derived from a number of fields including astronomy, geometry, biology and statistics, and all were based on a direct study of the underlying signal in either physical or frequency space. The most successful methods are based on finding cosine-like behavior, a rather limited definition of periodicity.

In this paper and in [23] we look instead at the shape of the sliding window point cloud, a totally different approach. Of course the geometry of point clouds derived from other kinds of data like images has been studied before [4, 16], but the current approach is quite different. Our method understands periodicity as repetition of patterns, whatever these may be, and quantifies this reoccurrence as the degree of circularity/roundness in the generated point-cloud. Thus, it is fundamentally agnostic.

1.1. **Previous Work.** The sliding window, or time-delay embedding, has been used mostly in the study of dynamical systems to understand the nature of their attractors. Takens' theorem [28] gives conditions under which a smooth attractor can be reconstructed from the observations of a function, with bounds related to those of the Whitney Embedding Theorem. This methodology has in turn been employed to test for non-linearity and chaotic behavior in the dynamics of ECG-EKG, EEG and MEG [25, 26].

It has been recently demonstrated by de Silva et. al. [10] that combining time-delay embeddings with topological methods provides a framework for parametrizing periodic systems.

Kantz and Schreiber provide in [16, Chapter 1] a good source of examples of time delay embeddings used in real-world data sets.

1.2. **Our Contribution.** In the above applications, little of the topology and none of the geometry of the resulting sliding window embedding has ever been used. The novelty of our approach lies in our use of this geometry and topology through persistent homology. We make this possible by showing that maximum persistence, as a measure of "roundness" of the point-cloud, occurs when the window size corresponds to the natural frequency of the signal. This means that 1D persistence is an effective quantifier of periodicity and quasi-periodicity and can be used to infer properties of the signal.

1.3. **Outline.** In section 2 we show a motivating example to illustrate our perspective. In section 3 we give a general introduction to persistent homology. More on this topic can be found in [12]. In section 4 we show that sliding windows behave well under approximations, and give explicit estimates at the point-cloud level. Section 5 is devoted to studying the geometric structure of sliding window embeddings from truncated Fourier series, as well as their dependency on embedding dimension and window size. In section 6 we prove results describing the structure of persistent

diagrams from sliding window embeddings. We present in section 7 some examples of how our method applies in the problem of quantifying periodicity in time series data.

2. Definitions and Motivation

Suppose that f is a function defined on an interval of the real numbers. Choose an integer M and a real number τ, both greater than 0. The sliding window embedding of f based at $t \in \mathbb{R}$ into $\mathbb{R}^{M+1}$ is the point

$$SW_{M,\tau}f(t) = \begin{bmatrix} f(t) \\ f(t+\tau) \\ \vdots \\ f(t+M\tau) \end{bmatrix}.$$

Choosing different values of t gives a collection of points called a **sliding window point cloud for** f. A critical parameter for this embedding is the **window-size** $M\tau$.

2.1. Motivation:

To motivate the approach we take in the paper, let us begin with the following example.

Example. Let $L \in \mathbb{N}$ and $f(t) = \cos(Lt)$. Then

$$\begin{aligned} SW_{M,\tau}f(t) &= \begin{bmatrix} \cos(Lt) \\ \cos(Lt+L\tau) \\ \vdots \\ \cos(Lt+ML\tau) \end{bmatrix} \\ &= \cos(Lt)\begin{bmatrix} 1 \\ \cos(L\tau) \\ \vdots \\ \cos(LM\tau) \end{bmatrix} - \sin(Lt)\begin{bmatrix} 0 \\ \sin(L\tau) \\ \vdots \\ \sin(LM\tau) \end{bmatrix} \\ &= \cos(Lt)\mathbf{u} - \sin(Lt)\mathbf{v} \end{aligned}$$

and therefore $t \mapsto SW_{M,\tau}f(t)$ describes a planar curve in $\mathbb{R}^{M+1}$, with winding number L, whenever $\mathbf{u}$ and $\mathbf{v}$ are linearly independent. One can in fact see how the shape of this curve changes as a function of L, M and τ. Indeed, let

$$A = \begin{bmatrix} \|\mathbf{u}\|^2 & -\langle \mathbf{u}, \mathbf{v} \rangle \\ -\langle \mathbf{u}, \mathbf{v} \rangle & \|\mathbf{v}\|^2 \end{bmatrix}$$

which can be computed using Lagrange's trigonometric formulae

$$\begin{aligned} \langle \mathbf{u}, \mathbf{v} \rangle &= \frac{1}{2}\sum_{m=0}^{M} \sin(2Lm\tau) = \frac{\sin(L(M+1)\tau)\sin(LM\tau)}{2\sin(L\tau)} \\ \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2 &= \sum_{m=0}^{M} \cos(2Lm\tau) = \frac{\sin(L(M+1)\tau)\cos(LM\tau)}{\sin(L\tau)} \\ \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 &= M+1. \end{aligned}$$

It follows that A is positive semi-definite (both its determinant and trace ar non-negative). This means the eigenvalues of A are non-negative and real: $\lambda_1 \geq \lambda_2 \geq 0$, and there is a 2×2 orthogonal matrix B so that

$$A = B^T \Lambda^2 B \quad \text{where} \quad \Lambda = \begin{bmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \end{bmatrix}.$$

Therefore, if $\mathbf{x}(t) = \begin{bmatrix} \cos(Lt) & \sin(Lt) \end{bmatrix}'$ (here $'$ denotes transpose) then

$$\begin{aligned} \|SW_{M,\tau} f(t)\|^2 &= \left\| \begin{bmatrix} | & | \\ \mathbf{u} & -\mathbf{v} \\ | & | \end{bmatrix} \mathbf{x}(t) \right\|^2 \\ &= \big\langle \mathbf{x}(t), A\, \mathbf{x}(t) \big\rangle \\ &= \big\langle \Lambda B\, \mathbf{x}(t), \Lambda B\, \mathbf{x}(t) \big\rangle. \end{aligned}$$

Since B is a rotation matrix, say by an angle α, then the map

$$SW_{M,\tau} f(t) \mapsto \begin{bmatrix} \sqrt{\lambda_1} \cos(Lt + \alpha) \\ \\ \sqrt{\lambda_2} \sin(Lt + \alpha) \end{bmatrix}$$

is an isometry.

In summary, for $f(t) = \cos(Lt)$, the embedding $t \mapsto SW_{M,\tau} f(t)$ describes an ellipse on the plane $Span\{\mathbf{u}, \mathbf{v}\}$, whose shape (minor and major axes) is determined by the square roots of the eigenvalues of A. These eigenvalues can be computed explicitly as

$$\begin{aligned} \lambda_1 &= \frac{(M+1) + \left| \frac{\sin(L(M+1)\tau)}{\sin(L\tau)} \right|}{2} \\ \lambda_2 &= \frac{(M+1) - \left| \frac{\sin(L(M+1)\tau)}{\sin(L\tau)} \right|}{2}. \end{aligned}$$

It follows that the ellipse is roundest when λ_2 attains its maximum, which occurs if and only if $L(M+1)\tau \equiv 0 \mod \pi$. One such instance is

$$M\tau = \left(\frac{M}{M+1} \right) \frac{2\pi}{L}$$

which is when the window-size approximates the length of the period of $f(t)$. In other words, the roundness of the sliding window point cloud for $f(t) = \cos(Lt)$ is maximized when the window-size is close to resonating with its natural frequency.

The previous example provides the following intuition: For a generic function f, the degree to which the image of $SW_{M,\tau} f$ traces a closed curve in $\mathbb{R}^{M+1}$ reflects how periodic f is. Moreover, if f is periodic, then the roundness of $SW_{M,\tau} f$ defined as the largest radius of a ball in $\mathbb{R}^{M+1}$ so that the curve

$$t \mapsto SW_{M,\tau} f(t)$$

is tangent to at least two points of its equator, is maximized when the window-size $M\tau$ approaches the period length. The goal of this paper is to understand these relations.

The geometry of the curve $t \mapsto SW_{M,\tau}f(t)$ can be quite complicated, as shown in figure 1.

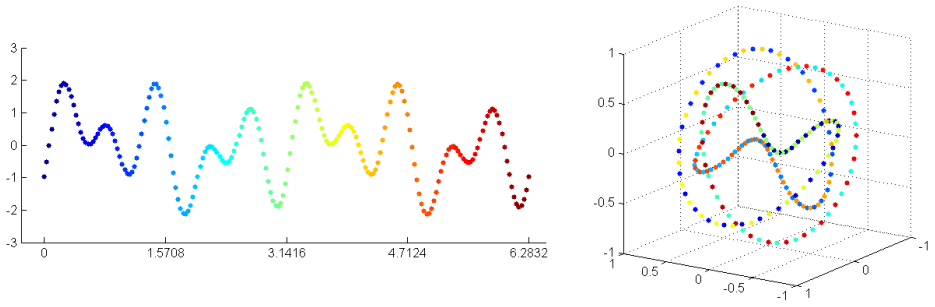

FIGURE 1. From a periodic function to its sliding window point cloud. **Left:** A periodic function f. **Right:** Multidimensional scaling into $\mathbb{R}^3$ for $SW_{20,\tau}f$. For each t, we use the same color for $f(t)$ and $SW_{20,\tau}f(t)$. Please refer to an electronic version for colors.

The 1-dimensional persistence diagram for the Vietoris-Rips filtration on a finite sample $\{SW_{M,\tau}f(t_1), \ldots, SW_{M,\tau}f(t_S)\}$, on the other hand, is readily computable [29, 30] and its maximum persistence is a measure of roundness as defined in the previous paragraph. We will review in section 3 the basic concepts behind persistent homology, and devote the rest of the paper to understanding how the geometry of $SW_{M,\tau}f$ reflects properties of f such as periodicity and period.

2.2. **Approach:** With this motivation in mind, we now describe our approach: As we have seen, understanding the algebraic properties of trigonometric functions allows one to characterize the geometry of $SW_{M,\tau}f$ when f is a trigonometric polynomial. This understanding, in turn, can be bootstrapped using Fourier analysis and stability of persistence diagrams, in an approximation step towards $SW_{M,\tau}$ of a generic periodic function. In what follows, we will establish the appropriate continuity results for approximation, as well as the necessary structural results for persistence diagrams from sliding window point clouds.

3. Background: Persistent Homology

In this section we define the key concepts that underlie the theory of persistent homology for filtered simplicial complexes. We give a terse introduction to simplicial homology, but more information can be found in [14, 21].

3.1. **Homology of Simplicial Complexes.** Let K be a simplicial complex and p a prime number. Recall that this means that K is a finite set of simplices that is closed under the face relation and that two simplices of K are either disjoint or intersect in a common face. Let $\mathbb{F}_p$ be the finite field with p elements, the $\mathbb{F}_p$ vector space generated by the k-dimensional simplices of K is denoted $C_k(K)$. It consists of all *k-chains*, which are finite formal sums $c = \sum_j \gamma_j x_j$, with $\gamma_j \in \mathbb{F}_p$ and each x_j a k-simplex in K. The boundary $\partial(x_j)$ is the alternating formal sum

of the $(k-1)$-dimensional faces of x_j and the boundary of the chain c is obtained by extending ∂ linearly

$$\partial(c) \quad = \quad \sum_j \gamma_j \partial(x_j).$$

It is not difficult to check that $\partial \circ \partial = \partial^2 = 0$. The k-chains that have boundary 0 are called *k-cycles*; they form a subspace Z_k of C_k. The k-chains that are the boundary of $(k+1)$-chains are called *k-boundaries* and form a subspace B_k of C_k. The fact that $\partial^2 = 0$ tells us that $B_k \subset Z_k$. The quotient group $H_k(K) = Z_k/B_k$ is the *k-th simplicial homology group* of K with $\mathbb{F}_p$-coefficients. The rank of $H_k(K)$ is the *k-th mod p Betti number* of K and is denoted $\beta_k(K)$. Since the prime p will be clear from the context, we do not include it in the notation.

When we have two simplicial complexes K and K', a *simplicial map* $f : K \to K'$ is a continuous map that takes simplices to simplices and is linear on each. A simplicial map induces a homomorphism on homology, $f_* : H_k(K) \to H_k(K')$, and homotopic maps induce the same homomorphism. Homotopy equivalences of spaces induce isomorphisms on homology. The simplicial approximation theorem tells us that a continuous map of simplicial complexes can be approximated by a simplicial map, so that it makes sense to talk about continuous maps inducing homomorphisms on homology.

Persistence. We next define persistence, persistent homology and the persistence diagram for a simplicial complex K. A *subcomplex* of K is a subset of its simplices that is closed under the face relation. A *filtration* of K is a nested sequence of subcomplexes that starts with the empty complex and ends with the complete complex,

$$\emptyset = K_0 \subset K_1 \subset \ldots \subset K_m = K.$$

A homology class α is *born* at K_i if it is not in the image of the map induced by the inclusion $K_{i-1} \subset K_i$. If α is born at K_i, we say that it *dies entering* K_j if the image of the map induced by $K_{i-1} \subset K_{j-1}$ does not contain the image of α but the image of the map induced by $K_{i-1} \subset K_j$ does. The *persistence* of α is $j - i$.

We code birth and death information in the persistence diagrams, one for each dimension. The diagram $\mathsf{dgm}(k)$ has a point (i, j) for every k-homology class that is born at K_i and dies entering K_j. For most of the paper the homological dimension k will be clear from the context or unimportant for the discussion. To ease notation we will simply write dgm instead of $\mathsf{dgm}(k)$, and let dgm_1, dgm_2 denote two k-persistence diagrams to be compared. Sometimes we have a function h that assigns a height or distance to each sub complex K_i, and in that case we use the pair $(h(i), h(j))$. Each diagram is now a multiset since classes can be born simultaneously and die simultaneously. We adjoin the diagonal $\Delta = \{(x, x) : x \geq 0\}$ to each diagram, and endow each point $(x, x) \in \Delta$ with countable multiplicity.

The Bottleneck distance between two persistence diagrams dgm_1 and dgm_2 is defined by

$$d_B(\mathsf{dgm}_1, \mathsf{dgm}_2) = \inf_{\phi} \sup_{x \in \mathsf{dgm}_1} ||x - \phi(x)||_\infty$$

where the infimum is taken over all bijections $\phi : \mathsf{dgm}_1 \to \mathsf{dgm}_2$. Note that such ϕ exist even if the number of points of dgm_1 and dgm_2 are different since we have included the diagonal.

Rips Complex. Let $X \subset \mathbb{R}^n$ be a compact set, for example a finite point cloud. We define $d_X(y)$ to be the distance from the point $y \in \mathbb{R}^n$ to X. We are interested in how the homology of the sub-level sets $X_r = d_X^{-1}([0, r])$ changes as we increase r. To make this computationally feasible, we replace the continuous family of spaces X_r with a discrete family of approximations called the Rips complexes defined as follows. Fix $r \geq 0$, $R_r(X)$ is the simplicial complex whose vertices are the points of X and whose k-simplices are the $k+1$ tuples $[x_0, \cdots, x_k]$ such that the pairwise distances $||x_i - x_j||$ are less than or equal to r for all $0 \leq i < j \leq k$. Note that the edges determine the simplices of $R_r(X)$, a higher dimensional simplex is added if and only if all its edges have been added, and that the Rips construction makes sense for any metric space.

Since $R_r(X) \subset R_s(X)$ whenever $r < s$, the Rips complexes form a filtration of R_∞, which denotes the largest simplicial complex having X as its vertex set. Changes occur at the finite set of r values that are pairwise distances between points, so we can work with just these r_j to get the filtration

$$X = R_0 \subset R_1 \subset \cdots \subset R_m,$$

where $R_j(X) = R_{r_j}(X)$ and $R_m = R_\infty$. We will use this filtered complex to study the persistence and the persistence diagrams of the point cloud X. We thus denote by $\mathsf{dgm}(X)$ the persistence diagram of the homology filtration induced from the Rips filtration on X, where we use homology with coefficents in $\mathbb{F}_p$.

A key property of persistence is that it is *stable* [6]. In our context this means that if X, Y are two point clouds and d_H, d_{GH} are the Hausdorff and Gromov-Hausdorff distances, then

$$d_B(\mathsf{dgm}(X), \mathsf{dgm}(Y)) \leq 2d_{GH}(X, Y) \leq 2d_H(X, Y). \tag{1}$$

4. The Approximation Theorem

In this section we show that one can study $SW_{M,\tau}f$ and the persistence of the point cloud it generates for a generic function $f \in L^2(\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z})$, by using its Fourier Series approximation. While it seems quite difficult to study $SW_{M,\tau}f$ directly, it is not hard to understand $SW_{M,\tau}\cos(nt)$ and $SW_{M,\tau}\sin(nt)$, so we will build our understanding of the geometry of a general $SW_{M,\tau}f$ from these special cases using the Fourier series of f. To do this we will need to show that $SW_{M,\tau}$ behaves well under approximations and that these approximations work in the context of stability for persistence diagrams.

Let $C(X, Y)$ denote the set of continuous functions from X to Y equipped with the sup norm. The sliding window embedding induces a mapping

$$SW_{M,\tau} : C(\mathbb{T}, \mathbb{R}) \longrightarrow C(\mathbb{T}, \mathbb{R}^{M+1}).$$

The first fact about this map that we need is the following:

Proposition 4.1. *Let $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$. Then for all $M \in \mathbb{N}$ and $\tau > 0$, the mapping $SW_{M,\tau} : C(\mathbb{T}, \mathbb{R}) \longrightarrow C(\mathbb{T}, \mathbb{R}^{M+1})$ is a bounded linear operator with norm $\|SW_{M,\tau}\| \leq \sqrt{M+1}$.*

Proof. Linearity of $SW_{M,\tau}$ follows directly from its definition. To see that it is bounded, notice that for every $f \in C(\mathbb{T}, \mathbb{R})$ and $t \in \mathbb{T}$ we have

$$\begin{aligned}\|SW_{M,\tau}f(t)\|^2_{\mathbb{R}^{M+1}} &= |f(t)|^2 + |f(t+\tau)|^2 + \cdots + |f(t+M\tau)|^2 \\ &\leq (M+1)\|f\|^2_\infty\end{aligned}$$

□

We now consider approximating a function f by its Fourier polynomials and study how the sliding windows behave in this context. In particular, let

$$f(t) = S_N f(t) + R_N f(t)$$

where

$$S_N f(t) = \sum_{n=0}^{N} a_n \cos(nt) + b_n \sin(nt) = \sum_{n=-N}^{N} \widehat{f}(n) e^{int}$$

is the N-truncated Fourier series expansion of f, $R_N f$ is the remainder, and

$$\widehat{f}(n) = \begin{cases} \frac{1}{2}a_n - \frac{i}{2}b_n & \text{if } n > 0, \\ \frac{1}{2}a_{-n} + \frac{i}{2}b_{-n} & \text{if } n < 0, \\ a_0 & \text{if } n = 0. \end{cases} \tag{2}$$

We can easily compute that

$$SW_{M,\tau}f(t) = \sum_{n=0}^{N} \cos(nt)\big(a_n \mathbf{u}_n + b_n \mathbf{v}_n\big) + \sin(nt)\big(b_n \mathbf{u}_n - a_n \mathbf{v}_n\big) + SW_{M,\tau}R_N f(t)$$

where

$$\mathbf{u}_n = SW_{M,\tau}\cos(nt)\big|_{t=0} \quad \text{and} \quad \mathbf{v}_n = SW_{M,\tau}\sin(nt)\big|_{t=0}.$$

The vectors $\mathbf{u}_n$ and $\mathbf{v}_n$ form a fundamental basis out of which we can build our understanding of the structure of the point clouds that sliding windows create. We introduce the notation:

$$\phi_\tau(t) = \sum_{n=0}^{N} \cos(nt)\big(a_n \mathbf{u}_n + b_n \mathbf{v}_n\big) + \sin(nt)\big(b_n \mathbf{u}_n - a_n \mathbf{v}_n\big),$$

for the sliding window embedding for $S_N f(t)$. Also, when f, M and N are clear from the context we will simply write $\phi_\tau = SW_{M,\tau}S_N f$.

The next step is to find a bound on the term $SW_{M,\tau}R_N f(t)$. We will actually find a series of bounds, one for each of the derivatives $f^{(k)} = \frac{d^k f}{dt^k}$, whenever they exist and are continuous.

Proposition 4.2. *Let $k \in \mathbb{N}$. If $f \in C^k(\mathbb{T}, \mathbb{R})$ then for all $t \in \mathbb{T}$*

$$\|SW_{M,\tau}f(t) - \phi_\tau(t)\|_{\mathbb{R}^{M+1}} \leq \sqrt{4k-2}\left\|R_N f^{(k)}\right\|_2 \cdot \frac{\sqrt{M+1}}{(N+1)^{k-\frac{1}{2}}}$$

Proof. If $k \in \mathbb{N}$ and $f \in C^k(\mathbb{T}, \mathbb{R})$, then integration by parts yields the well known identity

$$\left|\widehat{f^{(k)}}(n)\right| = |n|^k \left|\widehat{f}(n)\right|$$

for the length of $\widehat{f^{(k)}}(n)$, the n-th complex Fourier coefficient of $f^{(k)}$, $n \in \mathbb{Z}$. Thus for all $t \in \mathbb{T}$, the Cauchy-Schwartz inequality, Young's inequality and Parseval's theorem together imply that

$$\begin{aligned}
|R_N f(t)| &\leq \sum_{n=N+1}^{\infty} \frac{\left|\widehat{f^{(k)}}(n)\right| + \left|\widehat{f^{(k)}}(-n)\right|}{n^k} \\
&\leq \left(\sum_{n=N+1}^{\infty} \left(\left|\widehat{f^(k)}(n)\right| + \left|\widehat{f^{(k)}}(-n)\right| \right)^2 \right)^{1/2} \cdot \left(\sum_{n=N+1}^{\infty} \frac{1}{n^{2k}} \right)^{1/2} \\
&\leq \left(2 \sum_{|n| \geq N+1} \left|\widehat{f^{(k)}}(n)\right|^2 \right)^{1/2} \cdot \left(\int_{N+1}^{\infty} \frac{1}{x^{2k}} dx \right)^{1/2} \\
&= \sqrt{2} \left\| R_N f^{(k)} \right\|_2 \cdot \frac{\sqrt{2k-1}}{(N+1)^{k-\frac{1}{2}}}
\end{aligned}$$

and hence, by proposition 4.1

$$\begin{aligned}
\|SW_{M,\tau} f(t) - \phi_\tau(t)\|_{\mathbb{R}^{M+1}} &\leq \sqrt{M+1} \|R_N f\|_\infty \\
&\leq \sqrt{4k-2} \left\| R_N f^{(k)} \right\|_2 \cdot \frac{\sqrt{M+1}}{(N+1)^{k-\frac{1}{2}}}
\end{aligned}$$

□

These bounds readily imply estimates for the Hausdorff distance between the sliding window point clouds of f and $S_N f$. Indeed, let X and Y be the images of $T \subset \mathbb{T}$ through $SW_{M,\tau} f$ and ϕ_τ respectively. It follows that if $f \in C^k(\mathbb{T}, \mathbb{R})$ and

$$\epsilon > \sqrt{4k-2} \left\| R_N f^{(k)} \right\|_2 \frac{\sqrt{M+1}}{(N+1)^{k-\frac{1}{2}}}$$

then $X \subset Y^\epsilon$, $Y \subset X^\epsilon$ and therefore $d_H(X, Y) \leq \epsilon$. Letting ϵ approach its lower bound and using the stability of d_B with respect to d_H (equation 1), we obtain the relation

$$d_B\big(\mathsf{dgm}(X), \mathsf{dgm}(Y)\big) \leq 2\sqrt{4k-2} \left\| R_N f^{(k)} \right\|_2 \frac{\sqrt{M+1}}{(N+1)^{k-\frac{1}{2}}}$$

As described in the introduction, the maximum persistence of $\mathsf{dgm}(X)$ will serve to quantify the periodicity of f when measured with sliding windows of length $M\tau$. By the maximum persistence of a diagram dgm we mean the following

Definition 4.3. Let $(x, y) \in \mathsf{dgm}$ and define $\mathsf{pers}(x, y) = y - x$ for $(x, y) \in \mathbb{R}^2$, and as ∞ otherwise. We let

$$mp(\mathsf{dgm}) = \max_{\mathbf{x} \in \mathsf{dgm}} \mathsf{pers}(\mathbf{x})$$

denote the maximum persistence of dgm.

Remark 4.4. If dgm_Δ denotes the diagram with the diagonal as underlying set, each point endowed with countable multiplicity, then

$$mp(\mathsf{dgm}) = 2 d_B(\mathsf{dgm}, \mathsf{dgm}_\Delta).$$

Indeed, for any bijection $\phi : \mathsf{dgm} \longrightarrow \mathsf{dgm}_\Delta$ and every $\mathbf{x} \in \mathsf{dgm}$

$$\|\mathbf{x} - \phi(\mathbf{x})\|_\infty \geq \frac{1}{2}\mathsf{pers}(\mathbf{x})$$

with equality if and only if $\phi(x,y) = \left(\frac{x+y}{2}, \frac{x+y}{2}\right)$. Thus

$$\max_{\mathbf{x}\in\mathsf{dgm}} \|\mathbf{x} - \phi(\mathbf{x})\| \geq \frac{1}{2}mp(\mathsf{dgm})$$

and therefore $d_B(\mathsf{dgm}, \mathsf{dgm}_\Delta) = \min_\phi \max_{\mathbf{x}\in\mathsf{dgm}} \|\mathbf{x}-\phi(\mathbf{x})\| \geq \frac{1}{2}mp(\mathsf{dgm})$. For the reverse inequality, notice that the map

$$(x,y) \mapsto \left(\frac{x+y}{2}, \frac{x+y}{2}\right)$$

extends to a bijection $\phi_0 : \mathsf{dgm} \longrightarrow \mathsf{dgm}_\Delta$ of multisets, such that for all $\mathbf{x} \in \mathsf{dgm}$ one has $\|\mathbf{x} - \phi_0(\mathbf{x})\|_\infty = \frac{1}{2}\mathsf{pers}(\mathbf{x})$.

We summarize the results of this section in the following theorem:

Theorem 4.5 (**Approximation**). *Let $T \subset \mathbb{T}$, $f \in C^k(\mathbb{T}, \mathbb{R})$, $X = SW_{M,\tau}f(T)$ and $Y = SW_{M,\tau}S_N f(T)$. Then*

(1)
$$d_H(X,Y) \leq \sqrt{4k-2}\left\|R_N f^{(k)}\right\|_2 \frac{\sqrt{M+1}}{(N+1)^{k-\frac{1}{2}}}$$

(2)
$$\left|mp\big(\mathit{dgm}(X)\big) - mp\big(\mathit{dgm}(Y)\big)\right| \leq 2d_B\big(\mathit{dgm}(X), \mathit{dgm}(Y)\big)$$

(3)
$$d_B\Big(\mathit{dgm}(X), \mathit{dgm}(Y)\Big) \leq 2\sqrt{4k-2}\left\|R_N f^{(k)}\right\|_2 \frac{\sqrt{M+1}}{(N+1)^{k-\frac{1}{2}}}$$

It follows that the persistent homology of the sliding window point cloud of a function $f \in C^k(\mathbb{T}, \mathbb{R})$ can, in the limit, be understood in terms of that of its truncated Fourier series.

Remark 4.6. Regarding the hypothesis of f being at least C^1, Proposition 4.2 (which is the basis of the Approximation Theorem, 4.5) only uses that $f' \in L^2(\mathbb{T})$, thus everything up to this point (and in fact, for the rest of the paper) holds true for functions in the Sobolev space $W^{1,2}(\mathbb{T})$. The reason why we have phrased the results in terms of the spaces $C^k(\mathbb{T})$ is because it provides the following interpretation: If the function f has certain degree of niceness, then one should expect the approximation of the persistence diagrams from $SW_{M,\tau}f$ by those of $SW_{M,\tau}S_N f$ to improve at an explicit rate. Moreover, the nicer the function the better the rate.

Another function space for which our arguments apply is the set of Hölder continuous functions with exponent $\alpha \in \left(\frac{1}{2}, 1\right)$. Indeed, if for such an f one considers the Fejér approximation

$$\sigma_N f(t) = \sum_{|n|\leq N}\left(1 - \frac{|n|}{N+1}\right)\widehat{f}(n)e^{int}$$

then (see [22, Theorem 1.5.3])

$$\|\sigma_N f - f\|_\infty \leq \frac{C_\alpha K_\alpha}{N^\alpha}$$

where K_α is the Hölder constant of f and C_α is a constant depending solely on α. Hence one gets the following version of Proposition 4.2: For every $t \in \mathbb{T}$

$$\left\| SW_{M,\tau} f(t) - SW_{M,\tau}\sigma_N f(t) \right\|_{\mathbb{R}^{M+1}} \leq C_\alpha K_\alpha \frac{\sqrt{M+1}}{N^\alpha}$$

and the corresponding version of the Approximation Theorem follows. Later on (e.g. in Theorem 6.8) we will use some bounds in terms of $\|f'\|_2$ and $\|S_N f'\|_2$. Adapting results involving said bounds to the Holder-continuous setting requires some work: One can use the fact that every Holder function has a Lipschitz approximation, and then invoke Rademacher's theorem. We leave the details to the interested reader.

5. The Geometric Structure of $SW_{M,\tau}S_N f$

We now turn our attention to the sliding window construction when applied to the truncated Fourier series of a periodic function. More specifically, we study the geometric structure of the sliding window point cloud, and its dependency on τ, N and M.

Our focus on geometry contrasts with methods used by others to determine τ and M. Traditionally $M\tau$, the window size, is estimated using the autocorrelation function [17]. While M is sometimes estimated directly using the method of false nearest neighbors [16].

5.1. Dimension of the Embedding.

One way of interpreting the dimension of the embedding, $M+1$, is as the level of detail (from the function) one hopes to capture with the sliding window representation. Given the advantages of a description which is as detailed as possible, it can be argued that large dimensions are desirable. From a computational perspective, however, this is a delicate point as our ultimate goal is to compute the persistent homology of the associated sliding window point cloud. Indeed, as the dimension of the embedding grows, it follows that the point cloud needs to be (potentially) more densely populated. This causes the size of the Rips complex to outweigh the computational resources, making the persistent homology calculation unfeasible.

While there has been considerable progress on dealing with the size problem of the Rips complex [20], it is important to have a sense of the amount of retained information given the computational constraints on the embedding dimension. Fortunately, when dealing with trigonometric polynomials the answer is clear: One loses no information if and only if the embedding dimension is greater than twice the maximum frequency. Indeed, recall the linear decomposition

$$SW_{M,\tau}S_N f(t) = \sum_{n=0}^{N} \cos(nt)\big(a_n \mathbf{u}_n + b_n \mathbf{v}_n\big) + \sin(nt)\big(b_n \mathbf{u}_n - a_n \mathbf{v}_n\big)$$

where

$$\mathbf{u}_n = SW_{M,\tau}\cos(nt)\big|_{t=0} \quad , \quad \mathbf{v}_n = SW_{M,\tau}\sin(nt)\big|_{t=0}$$

and a_n, b_n are as defined in equation 2. Since the angles between the $\mathbf{u}_n$'s and the $\mathbf{v}_m$'s, as well as their norms can be determined from M and τ (see Example in section 2.1), then $S_N f$ can be recovered from $SW_{M,\tau}S_N f$ if the $\mathbf{u}_n$'s and the $\mathbf{v}_m$'s are linearly independent. This is the sense in which we say that there is no loss of information.

Proposition 5.1. *Let $M\tau < 2\pi$. Then $\mathbf{u}_0, \mathbf{u}_1, \mathbf{v}_1, \ldots, \mathbf{u}_N, \mathbf{v}_N$ are linearly independent if and only if $M \geq 2N$.*

Proof. If $2N+1$ vectors in $\mathbb{R}^{M+1}$ are linearly independent, it readily follows that $2N \leq M$. Let us assume now that $\mathbf{u}_0, \mathbf{u}_1, \mathbf{v}_1, \ldots, \mathbf{u}_N, \mathbf{v}_N$ are linearly dependent and let us show that $2N > M$, or equivalently that $2N \geq M+1$. Indeed, let $\gamma_0, \beta_0, \ldots, \gamma_N, \beta_N \in \mathbb{R}$ be scalars not all zero (set $\beta_0 = 0$) so that

$$\gamma_0 \mathbf{u}_0 + \beta_0 \mathbf{v}_0 + \cdots + \gamma_N \mathbf{u}_N + \beta_N \mathbf{v}_N = \mathbf{0}.$$

That is, for all $m = 0, \ldots, M$ we have

$$0 = \sum_{n=0}^{N} \gamma_n \cos(nm\tau) + \beta_n \sin(nm\tau) = Re\Big(\sum_{n=0}^{N} (\gamma_n - i\beta_n) e^{inm\tau} \Big).$$

Let $\xi_m = e^{im\tau}$, $\quad p(z) = \sum\limits_{n=0}^{N} (\gamma_n + i\beta_n) z^n$, $\quad \overline{p}(z) = \sum\limits_{n=0}^{N} (\gamma_n - i\beta_n) z^n$ $\quad$ and

$$q(z) = z^N \cdot \left(\overline{p}(z) + p\left(\frac{1}{z} \right) \right).$$

It follows that $q(z)$ is a non-constant complex polynomial of degree at most $2N$, and that for $m = 0, \ldots, M$ we have $0 = Re(\overline{p}(\xi_m))$. This implies that

$$\begin{aligned} q(\xi_m) &= (\xi_m)^N \left(\overline{p}(\xi_m) + p\left(\frac{1}{\xi_m} \right) \right) \\ &= (\xi_m)^N \left(\overline{p}(\xi_m) + p\Big(\overline{\xi_m} \Big) \right) \\ &= 2(\xi_m)^N \; Re\Big(\overline{p}(\xi_m) \Big) \\ &= 0 \end{aligned}$$

and therefore $\xi_0, \ldots, \xi_M$ are roots of $q(z)$. Since $M\tau < 2\pi$ then $\xi_0, \xi_1, \ldots, \xi_M$ are distinct, and we have that

$$M + 1 \leq degree\big(q(z)\big) \leq 2N.$$

□

It is useful to contrast Proposition 5.1 with two important results in signal analysis: Takens' theorem from dynamical systems [28], and the Nyquist-Shannon sampling theorem from information theory [27]. Takens' theorem gives sufficient conditions on the length of a sequence of observation, so that the resulting embedding recovers the topology of a smooth attractor in a chaotic dynamical system. The aforementioned condition is that the dimension of the embedding should be greater than twice (an appropriate notion of) that of the attractor. The Nyquist-Shannon sampling theorem, on the other hand, contends that a band-limited signal can be recovered (exactly) from a sequence of observations whenever the sampling frequency is greater than twice the position, in the frequency domain, of the limiting band. The conclusion: In the case of trigonometric polynomials and the sliding window construction, the usual sufficient condition on dimension of the embedding and maximum frequency is also necessary.

Important Assumption: Unless otherwise stated, given $N \in \mathbb{N}$ we will always set $M = 2N$, and require $\tau > 0$ to be so that $M\tau < 2\pi$.

5.2. **Window Size and Underlying Frequency.** As we saw in the Motivation Section 2, the sliding window point cloud for $\cos(Lt)$ describes a planar ellipse which is roundest when $\|\mathbf{u}\| - \|\mathbf{v}\| = \langle \mathbf{u}, \mathbf{v} \rangle = 0$, or equivalently, when

$$L(M+1)\tau \equiv 0 \pmod{\pi}.$$

This uncovers a fundamental relation between window size, 1-dimensional persistence and underlying frequency: The maximum persistence of the sliding window point cloud from $\cos(Lt)$ is largest when the window size $M\tau$ is proportional to the underlying frequency $\frac{2\pi}{L}$, with proportionality constant $\frac{M}{M+1}$.

For the case of the truncated Fourier series $S_N f$ from a periodic function f, we will see shortly that if the same proportionality relation between window size and underlying frequency holds then

$$SW_{M,\tau} S_N f(t) = \sum_{n=0}^{N} \cos(nt)\big(a_n \mathbf{u}_n + b_n \mathbf{v}_n\big) + \sin(nt)\big(b_n \mathbf{u}_n - a_n \mathbf{v}_n\big) \tag{3}$$

is a linear decomposition into mutually orthogonal vectors. We begin with the now familiar case of the restriction to $Span\{\mathbf{u}_n, \mathbf{v}_n\}$.

Proposition 5.2. *For $n \geq 1$, $\langle \mathbf{u}_n, \mathbf{v}_n \rangle = \|\mathbf{u}_n\|^2 - \|\mathbf{v}_n\|^2 = 0$ if and only if*

$$n(M+1)\tau \equiv 0 \ \ (mod\ \pi).$$

Proof.

$$\begin{aligned}
\langle \mathbf{u}_n, \mathbf{v}_n \rangle &= \sum_{m=1}^{M} \cos(nm\tau)\sin(nm\tau) = \frac{1}{2}\sum_{m=1}^{M} \sin(2nm\tau) \\
&= \frac{1}{2} Im\left(\sum_{m=1}^{M} z_{2n\tau}^{m}\right), \ \text{ where } z_\theta = e^{i\theta} \\
&= \frac{1}{2} Im\left(\frac{1 - z_{2n(M+1)\tau}}{1 - z_{2n\tau}} - 1\right) \\
&= \frac{1}{2} Im\left(\frac{1 - z_{2n(M+1)\tau}}{1 - z_{2n\tau}}\right) \\
\|\mathbf{u}_n\|^2 - \|\mathbf{v}_n\|^2 &= \sum_{m=0}^{M} \cos^2(nm\tau) - \sin^2(nm\tau) \\
&= Re\left(\frac{1 - z_{2n(M+1)\tau}}{1 - z_{2n\tau}}\right)
\end{aligned}$$

and therefore

$$4\langle \mathbf{u}_n, \mathbf{v}_n \rangle^2 + (\|\mathbf{u}_n\|^2 - \|\mathbf{v}_n\|^2)^2 = \left\| \frac{1 - z_{2n(M+1)\tau}}{1 - z_{2n\tau}} \right\|^2.$$

It follows that $\langle \mathbf{u}_n, \mathbf{v}_n \rangle = \|\mathbf{u}_n\|^2 - \|\mathbf{v}_n\|^2 = 0$ if and only if $z_{2n(M+1)\tau} = 1$, which holds true if and only if $n(M+1)\tau \equiv 0 \pmod{\pi}$. □

It can be checked that $nM\tau \equiv 0 \pmod{\pi}$ also yields $\langle \mathbf{u}_n, \mathbf{v}_n \rangle = 0$, but letting $n(M+1)\tau \equiv 0 \pmod{\pi}$ implies that $a_n \mathbf{u}_n + b_n \mathbf{v}_n$ is perpendicular to $b_n \mathbf{u}_n - a_n \mathbf{v}_n$ for all $a_n, b_n \in \mathbb{R}$. Now, in order to extend the perpendicularity results to components from different harmonics, we will use the following:

Definition 5.3. We say that a function f is L-periodic on $[0, 2\pi]$, $L \in \mathbb{N}$, if

$$f\left(t + \frac{2\pi}{L}\right) = f(t)$$

for all t.

Remark 5.4. If f is an L-periodic function, a_n and b_n are its n-th real Fourier coefficients (see equation 2), and we let $a_n + ib_n = r_n e^{i\alpha_n}$, with $\alpha_n = 0$ whenever $r_n = 0$; then $r_n \neq 0$ implies $n \equiv 0 \pmod{L}$. Indeed, $g(t) = f(t/L)$ is a 1-periodic function and therefore has a Fourier series expansion

$$g(t) = \sum_{r=0}^{\infty} a'_r \cos(rt) + b'_r \sin(rt)$$

with equality almost everywhere. Thus

$$f(t) = g(tL) = \sum_{r=0}^{\infty} a'_r \cos(rLt) + b'_r \sin(rLt) = \sum_{n=0}^{\infty} a_n \cos(nt) + b_n \sin(nt)$$

for almost every t, and the result follows from the uniqueness of the Fourier expansion in $L^2(\mathbb{T})$.

We are now ready to see that the potentially non-zero terms in the linear decomposition of $SW_{M,\tau}S_N f$ (equation 3), can be made mutually orthogonal by choosing the window size proportional to the underlying frequency, with proportionality constant $\frac{M}{M+1}$.

Proposition 5.5. *Let f be L-periodic, and let $\tau = \frac{2\pi}{L(M+1)}$. Then the vectors in*

$$\{\mathbf{u}_n, \mathbf{v}_n \mid 0 \leq n \leq N,\ n \equiv 0 \pmod{L}\}$$

are mutually orthogonal, and we have $\|\mathbf{u}_n\| = \|\mathbf{v}_n\| = \sqrt{\frac{M+1}{2}}$ for $n \equiv 0 \pmod{L}$.

Proof. Let $k = pL$ and $n = qL$. If $k = n$, it follows from Proposition 5.2 that $\langle \mathbf{u}_n, \mathbf{v}_n \rangle = 0$ and

$$\begin{aligned} \|\mathbf{u}_n\|^2 = \|\mathbf{v}_n\|^2 &= \frac{\|\mathbf{u}_n\|^2 + \|\mathbf{v}_n\|^2}{2} = \frac{1}{2}\sum_{m=0}^{M} \cos(nm\tau)^2 + \sin(nm\tau)^2 \\ &= \frac{M+1}{2}. \end{aligned}$$

Let us assume now that $p \neq q$. If we let $z_\theta = e^{i\theta}$, $\theta \in \mathbb{R}$, then

$$\begin{aligned} \langle \mathbf{u}_n, \mathbf{u}_k \rangle &= \sum_{m=0}^{M} \cos(nm\tau)\cos(km\tau) \\ &= \frac{1}{2}\sum_{m=0}^{M} \cos((n-k)m\tau) + \cos((n+k)m\tau) \\ &= \frac{1}{2} Re\left(\frac{1 - z_{(n-k)(M+1)\tau}}{1 - z_{(n-k)\tau}} + \frac{1 - z_{(n+k)(M+1)\tau}}{1 - z_{(n+k)\tau}}\right) \\ &= \frac{1}{2} Re\left(\frac{1 - z_{(q-p)2\pi}}{1 - z_{(n-k)\tau}} + \frac{1 - z_{(q+p)2\pi}}{1 - z_{(n+k)\tau}}\right) = 0. \end{aligned}$$

Notice that

$$0 < \min\{|n-k|, |n+k|\} \leq \max\{|n-k|, |n+k|\} \leq 2N \leq M < \frac{2\pi}{\tau}$$

implies that the denominators are never zero. Similarly

$$\begin{aligned}
\langle \mathbf{u}_n, \mathbf{v}_k \rangle &= \sum_{m=1}^{M} \cos(nm\tau)\sin(km\tau) \\
&= \frac{1}{2}\sum_{m=1}^{M} \sin((n+k)m\tau) - \sin((n-k)m\tau) \\
&= \frac{1}{2} Im\left(\frac{1-z_{(q+p)2\pi}}{1-z_{(n+k)\tau}} - \frac{1-z_{(q-p)2\pi}}{1-z_{(n-k)\tau}}\right) = 0 \\
\langle \mathbf{v}_n, \mathbf{u}_k \rangle &= \frac{1}{2} Im\left(\frac{1-z_{(p+q)2\pi}}{1-z_{(k+n)\tau}} - \frac{1-z_{(p-q)2\pi}}{1-z_{(k-n)\tau}}\right) = 0 \\
\langle \mathbf{v}_n, \mathbf{v}_k \rangle &= \frac{1}{2} Re\left(\frac{1-z_{(q-p)2\pi}}{1-z_{(n-k)\tau}} - \frac{1-z_{(q+p)2\pi}}{1-z_{(n+k)\tau}}\right) = 0.
\end{aligned}$$

□

When computing persistent homology it is sometimes advantageous to pointwise center and normalize the set of interest. The next theorem describes the result of such operations on the sliding window point cloud for $SW_{M,\tau}S_N f$, when f is L-periodic and $L(M+1)\tau = 2\pi$.

Theorem 5.6 (**Structure**). *Let $C : \mathbb{R}^{M+1} \longrightarrow \mathbb{R}^{M+1}$ be the centering map*

$$C(\mathbf{x}) = \mathbf{x} - \frac{\langle \mathbf{x}, \mathbf{1}\rangle}{\|\mathbf{1}\|^2}\mathbf{1} \quad \textit{where} \quad \mathbf{1} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{R}^{M+1}.$$

If f is L-periodic, $L(M+1)\tau = 2\pi$ and $\phi_\tau = SW_{M,\tau}S_N f$, then

(1)
$$\phi_\tau(t) = \widehat{f}(0)\cdot\mathbf{1} + C(\phi_\tau(t))$$

(2)
$$\big\|C\big(\phi_\tau(t)\big)\big\| = \sqrt{M+1}\left(\|S_N f\|_2^2 - \widehat{f}(0)^2\right)^{1/2}$$

(3) *There exists an orthonormal set*

$$\left\{\widetilde{\mathbf{x}}_n, \widetilde{\mathbf{y}}_n \in \mathbb{R}^{M+1} \;\middle|\; 1 \leq n \leq N,\ n \equiv 0 \ (mod\ L)\right\}$$

such that

$$\varphi_\tau(t) = \frac{C\left(\phi_\tau(t)\right)}{\|C\left(\phi_\tau(t)\right)\|} = \sum_{\substack{n\,=\,1 \\ n \,\equiv\, 0 \ (mod\ L)}}^{N} \widetilde{r}_n\big(\cos(nt)\widetilde{\mathbf{x}}_n + \sin(nt)\widetilde{\mathbf{y}}_n\big) \tag{4}$$

where

$$\widetilde{r}_n = \frac{2\left|\widehat{f}(n)\right|}{\sqrt{\|S_N f\|_2^2 - \widehat{f}(0)^2}}$$

Proof. If f is an L-periodic function on $[0, 2\pi]$ and $L(M+1)\tau = 2\pi$, then Remark 5.4 and Proposition 5.5 imply that for all $t \in \mathbb{R}$

$$\begin{aligned}\phi_\tau(t) &= \sum_{\substack{n=0 \\ n \equiv 0 \pmod{L}}}^{N} \cos(nt)\big(a_n \mathbf{u}_n + b_n \mathbf{v}_n\big) + \sin(nt)\big(b_n \mathbf{u}_n - a_n \mathbf{v}_n\big) \\ &= \sum_{\substack{n=0 \\ n \equiv 0 \pmod{L}}}^{N} r_n\big(\cos(nt)\mathbf{x}_n + \sin(nt)\mathbf{y}_n\big)\end{aligned}$$

is a linear combination of the mutually orthogonal vectors $\mathbf{x}_n = \cos(\alpha_n)\mathbf{u}_n + \sin(\alpha_n)\mathbf{v}_n$ and $\mathbf{y}_n = \sin(\alpha_n)\mathbf{u}_n - \cos(\alpha_n)\mathbf{v}_n$.

Moreover, from Proposition 5.5 we have that if $n \geq 1$ is so that $n \equiv 0 \pmod{L}$ then $\|\mathbf{x}_n\| = \|\mathbf{y}_n\| = \sqrt{\frac{M+1}{2}}$. It follows that if

$$\widetilde{\mathbf{x}}_n = \frac{\mathbf{x}_n}{\|\mathbf{x}_n\|}, \;\; \widetilde{\mathbf{y}}_n = \frac{\mathbf{y}_n}{\|\mathbf{y}_n\|}$$

then

$$\phi_\tau(t) = \big(a_0\sqrt{M+1}\big)\frac{\mathbf{1}}{\|\mathbf{1}\|} \quad + \sum_{\substack{n=1 \\ n \equiv 0 \pmod{L}}}^{N} \sqrt{\frac{M+1}{2}} r_n\big(\cos(nt)\widetilde{\mathbf{x}}_n + \sin(nt)\widetilde{\mathbf{y}}_n\big)$$

is a linear decomposition of $\phi_\tau(t)$ in terms of the orthonormal set

$$\left\{ \frac{\mathbf{1}}{\|\mathbf{1}\|}, \, \widetilde{\mathbf{x}}_n, \widetilde{\mathbf{y}}_n \;\middle|\; 1 \leq n \leq N, \; n \equiv 0 \pmod{L} \right\}.$$

Hence $C\big(\phi_\tau(t)\big) = \sum_{\substack{n=1 \\ n \equiv 0 \pmod{L}}}^{N} \sqrt{\frac{M+1}{2}} r_n\big(\cos(nt)\widetilde{\mathbf{x}}_n + \sin(nt)\widetilde{\mathbf{y}}_n\big)$ and therefore

$$\begin{aligned}\varphi_\tau(t) &= \frac{C\big(\phi_\tau(t)\big)}{\big\|C\big(\phi_\tau(t)\big)\big\|} \\ &= \sum_{\substack{n=1 \\ n \equiv 0 \pmod{L}}}^{N} \frac{r_n}{\sqrt{r_1^2 + \cdots + r_N^2}}\big(\cos(nt)\widetilde{\mathbf{x}}_n + \sin(nt)\widetilde{\mathbf{y}}_n\big)\end{aligned}$$

which we write as

$$\varphi_\tau(t) = \sum_{\substack{n=1 \\ n \equiv 0 \pmod{L}}}^{N} \widetilde{r}_n\big(\cos(nt)\widetilde{\mathbf{x}}_n + \sin(nt)\widetilde{\mathbf{y}}_n\big) \quad , \quad \sum_{n=1}^{N} \widetilde{r}_n^2 = 1.$$

The result follows from the identities $r_n = 2\left|\widehat{f}(n)\right| = \left|\widehat{f}(n)\right| + \left|\widehat{f}(-n)\right|$, $n \geq 1$. □

Theorem 5.6 allows us to paint a very clear geometric picture of the centered and normalized sliding window point cloud for $S_N f$ (see equation 4). Indeed, if $S^1(r) \subset \mathbb{C}$ denotes the circle of radius r centered at zero, then $t \mapsto \varphi_\tau(t)$ can be regarded as the curve in the N-torus

$$\mathcal{T} = S^1(\widetilde{r}_1) \times \cdots \times S^1(\widetilde{r}_N)$$

which when projected onto $S^1(\widetilde{r}_n)$, $\widetilde{r}_n > 0$, goes around n times at a constant speed. Another interpretation, in terms of flat (polar) coordinates, is as the image through the quotient map

$$\mathbb{R}^N = \mathbb{R} \times \cdots \times \mathbb{R} \longrightarrow (\mathbb{R}/\widetilde{r}_1\mathbb{Z}) \times \cdots \times (\mathbb{R}/\widetilde{r}_N\mathbb{Z})$$

of the line segment in $\mathbb{R}^N$ joining $(0, 0, \ldots, 0)$ and $(\widetilde{r}_1, 2\widetilde{r}_2, \ldots, N\widetilde{r}_N)$. Figure 2 depicts $\varphi_\tau(t)$ inside $\mathcal{T}$ for $N = 3$.

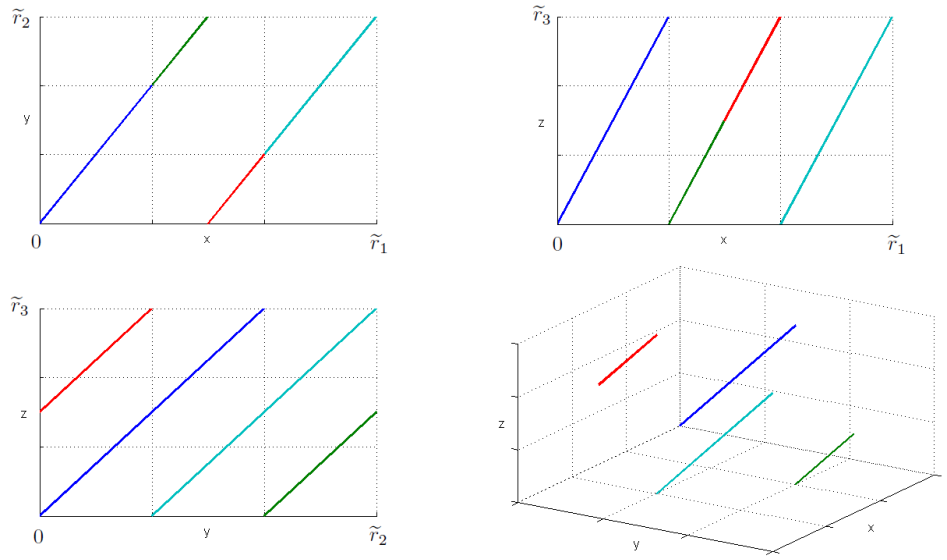

FIGURE 2. The curve $\varphi_\tau(t)$, in colors, with respect to its flat coordinates $(t, 2t, 3t) \in (\mathbb{R}/\widetilde{r}_1\mathbb{Z}) \times (\mathbb{R}/\widetilde{r}_2\mathbb{Z}) \times (\mathbb{R}/\widetilde{r}_3\mathbb{Z})$. Please refer to an electronic version for colors. **Bottom Right:** $\varphi_\tau(t)$ in the fundamental domain $[0, \widetilde{r}_1) \times [0, \widetilde{r}_2) \times [0, \widetilde{r}_3)$. **Top Left:** Projection onto the xy-plane. **Top Right:** Projection onto the xz-plane. **Bottom left:** Projection onto the yz-plane.

6. The Persistent Homology of φ_τ and $SW_{M,\tau}f$

The structural observations from the previous section, as well as the approximation results from Section 4, set the stage for understanding the persistent homology of the image of ϕ_τ (or rather of φ_τ) and how it relates to that of $SW_{M,\tau}f$.

6.1. **Some convergence results.** Let $T \subset \mathbb{T}$, and let $SW_{M,\tau}f(T)$ and $\phi_\tau(T)$ be the images of T through $SW_{M,\tau}f$ and ϕ_τ respectively. An immediate consequence of Proposition 4.2 is that as N (and thus $M = 2N$) gets larger, $\phi_\tau(T)$ gets closer to $SW_{M,\tau}f(T)$ with respect to the Hausdorff metric on subspaces of $\mathbb{R}^\infty$. Here $\mathbb{R}^\infty$ denotes the set of sequences $x = (x_k)_{k\in\mathbb{N}}$, $x_k \in \mathbb{R}$, so that $x_n = 0$ for all $n \geq N_0$, and some $N_0 = N_0(x) \in \mathbb{N}$. We endow $\mathbb{R}^\infty$ with the L^2 metric, and regard $SW_{M,\tau}f(t)$, $t \in T$, as an element of $\mathbb{R}^\infty$ by identifying it with

$$\big(f(t), f(t+\tau), \ldots, f(t+M\tau), 0, 0, \ldots\big) \in \mathbb{R}^\infty.$$

Notice, however, that while increasing the dimension $M+1$ of the sliding window embedding yields better approximations

$$SW_{M,\tau}S_N f(T) \approx SW_{M,\tau} f(T),$$

the object being approximated, $SW_{M,\tau}f(T)$, is changing. Since $(\mathbb{R}^\infty, \|\cdot\|_2)$ is not complete there is no reason to believe this process converges or stabilizes, even with a sensible way of comparing, say, $SW_{M,\tau}f(t)$ and $SW_{2M,\frac{\tau}{2}}f(t)$. This is the case, since they are samplings at different rates from the same window. Perhaps considering the Gromov-Hausdorff distance instead of the Hausdorff distance would yield such comparison, but at least at the moment we do not have a natural embedding to make this case. In addition, even when the metric completion $\overline{\mathbb{R}^\infty} = \ell^2(\mathbb{R})$, the space of square-summable sequences, is well understood, it is also big enough so that tracking global geometric features requires some work. It is in situations like this that a succinct and informative summary, such as persistence diagrams, is critical.

It is known that the space of persistence diagrams is not complete with respect to the Bottleneck distance, but that it can be completed by allowing diagrams with countably many points with at most countable multiplicity, satisfying a natural finiteness condition. See [2, Theorem 3.4] and [19, Theorem 6]. Moreover, features such as maximum persistence can be easily tracked, and there is no ambiguity on how to compare the diagrams from, say, $SW_{M,\tau}f(T)$ and $SW_{2M,\frac{\tau}{2}}f(T)$.

Proposition 6.1. *Let f be L-periodic, $N < N'$, $M = 2N$, $M' = 2N'$ and*

$$\tau = \frac{2\pi}{L(M+1)} \quad , \quad \tau' = \frac{2\pi}{L(M'+1)}.$$

If $T \subset \mathbb{T}$ is finite, $Y = SW_{M,\tau}S_N f(T)$ and $Y' = SW_{M',\tau'}S_{N'}f(T)$, then

$$d_B\left(\frac{\mathit{dgm}(Y)}{\sqrt{M+1}}, \frac{\mathit{dgm}(Y')}{\sqrt{M'+1}}\right) \leq 2\left\|S_N f - S_{N'} f\right\|_2$$

where $\lambda \cdot \mathit{dgm}(Z)$ is defined as $\{(\lambda x, \lambda y) \mid (x,y) \in \mathit{dgm}(Z)\}$, for $\lambda \geq 0$.

Proof. Let us fix the notation $\mathbf{u}_n = \mathbf{u}_n(M,\tau)$, $\mathbf{v}_n = \mathbf{v}_n(M,\tau)$, $\mathbf{u}'_n = \mathbf{u}_n(M',\tau')$ and $\mathbf{v}'_n = \mathbf{v}_n(M',\tau')$, in order to specify the dependencies of $\mathbf{u}_n$ and $\mathbf{v}_k$ on M and τ. Then we have linear maps

$$\begin{array}{rccc} P: & \mathbb{R}^{M'+1} & \longrightarrow & \mathbb{R}^{M'+1} \\ & \sum_{n=0}^{N'} x_n \mathbf{u}'_n + y_n \mathbf{v}'_n & \mapsto & \sum_{n=0}^{N} x_n \mathbf{u}'_n + y_n \mathbf{v}'_n \end{array}$$

$$\begin{array}{rccc} Q: & Img(P) & \longrightarrow & \mathbb{R}^{M+1} \\ & \mathbf{u}'_n & \mapsto & \sqrt{\frac{M'+1}{M+1}}\mathbf{u}_n \\ & \mathbf{v}'_n & \mapsto & \sqrt{\frac{M'+1}{M+1}}\mathbf{v}_n \end{array}$$

which are well defined by Proposition 5.1. Moreover, Proposition 5.5 implies that P can be interpreted as an orthogonal projection when restricted to Y', and that

Q is an isometry on $P(Y')$. Notice that for every $y' \in Y'$

$$\|y' - P(y')\| = \sqrt{\frac{M'+1}{2} \sum_{n=N+1}^{N'} r_n^2}$$

where r_n is as defined in Remark 5.4, and therefore

$$d_H(Y', P(Y')) \leq \sqrt{\frac{M'+1}{2} \sum_{n=N+1}^{N'} r_n^2}.$$

Finally, since $Q \circ P(Y') = \sqrt{\frac{M'+1}{M+1}} Y$ and $\mathsf{dgm}(\,\cdot\,)$ is invariant under isometries then

$$\begin{aligned} \sqrt{M'+1} \cdot d_B \left(\frac{\mathsf{dgm}(Y')}{\sqrt{M'+1}}, \frac{\mathsf{dgm}(Y)}{\sqrt{M+1}} \right) &= d_B\big(\mathsf{dgm}(Y'), \mathsf{dgm}(Q \circ P(Y'))\big) \\ &= d_B\big(\mathsf{dgm}(Y'), \mathsf{dgm}(P(Y'))\big) \\ &\leq 2 d_H(Y', P(Y')) \\ &\leq \sqrt{2(M'+1) \sum_{n=N+1}^{N'} r_n^2} \end{aligned}$$

and the result follows from the identity $r_n = 2\left|\widehat{f}(n)\right| = \left|\widehat{f}(n)\right| + \left|\widehat{f}(-n)\right|$, $n \geq 1$. □

This result, paired with the fact that $\|f - S_N f\|_2 \to 0$ as $N \to \infty$, and the Structure Theorem 5.6(1,2), imply the following:

Corollary 6.2. *Let $f \in L^2(\mathbb{T})$ be L-periodic, $N \in \mathbb{N}$, $\tau_N = \frac{2\pi}{L(2N+1)}$, $T \subset \mathbb{T}$ finite, and let $\overline{Y}_N$ be the set resulting from pointwise centering and normalizing the point cloud*

$$SW_{2N,\tau_N} S_N f(T) \subset \mathbb{R}^{2N+1}.$$

Then for any field of coefficients, the sequence $\mathsf{dgm}(\overline{Y}_N)$ of persistence diagrams is Cauchy with respect to d_B.

Completeness of the set of (generalized) diagrams [2, Theorem 3.4], allows one to make the following definition:

Definition 6.3. Let $w = \frac{2\pi}{L}$ and denote by $\mathsf{dgm}_\infty(f, T, w)$ the limit in the Bottleneck distance of the sequence $\mathsf{dgm}(\overline{Y}_N)$.

We hope the notation $\mathsf{dgm}_\infty(f, T, w)$ is suggestive enough to evoke the idea that there exists a limiting diagram from the sequence of pointwise-centered and normalized versions of $SW_{M,\tau} f(T)$, as $M \to \infty$, and while keeping the window size $M\tau = \frac{M}{M+1} w \approx w$. The First Convergence Theorem (6.6) bellow, asserts the validity of this notation. Before presenting the proof, we start with a technical result:

Proposition 6.4. *Let $f \in C(\mathbb{T})$ be L-periodic, $N \in \mathbb{N}$ and $\tau_N = \frac{2\pi}{L(2N+1)}$. Then*

$$\lim_{N\to\infty} \frac{\left\|C\big(SW_{2N,\tau_N} f(t)\big)\right\|}{\sqrt{2N+1}} = \left\| f - \widehat{f}(0) \right\|_2 \tag{5}$$

uniformly in $t \in \mathbb{T}$.

Proof. Since the result is trivially true if f is constant, let us assume $f \neq \widehat{f}(0)$ and let

$$g(t) = \frac{f(t) - \widehat{f}(0)}{\left\| f - \widehat{f}(0) \right\|_2}$$

It follows that $g \in C(\mathbb{T})$ is L-periodic,

$$\widehat{g}(0) = \frac{1}{2\pi} \int_0^{2\pi} g(t) dt = 0 \quad \text{and} \quad \|g\|_2 = \frac{1}{\sqrt{2\pi}} \left(\int_0^{2\pi} |g(t)|^2 dt \right)^{1/2} = 1.$$

Using the identity $L(2N+1)\tau_N = 2\pi$, a Riemann sums argument, and the fact that g is L-periodic, it follows that if

$$c_N(t) = \frac{g(t) + g(t+\tau_N) + \cdots + g(t+2N\tau_N)}{2N+1}$$

then for all $t \in \mathbb{T}$

$$\begin{aligned} \lim_{N\to\infty} c_N(t) &= \lim_{\tau_N \to 0} \frac{L}{2\pi} \Big(\tau_N g(t) + \tau_N g(t+\tau_N) + \cdots + \tau_N g(t+2N\tau_N) \Big) \\ &= \frac{L}{2\pi} \int_t^{t+\frac{2\pi}{L}} g(r) dr \\ &= \frac{1}{2\pi} \int_0^{2\pi} g(r) dr \\ &= 0. \end{aligned}$$

We contend that the convergence $c_N(t) \to 0$ is uniform in $t \in \mathbb{T}$. Indeed, the fact that g is uniformly continuous implies that the sequence $c_N(t)$ is uniformly equicontinuous. This means that for every $\epsilon > 0$ there exists $\delta > 0$ independent of N such that for every $t, t' \in \mathbb{T}$ and all $N \in \mathbb{N}$

$$|t - t'| < \delta \text{ implies } |c_N(t) - c_N(t')| < \frac{\epsilon}{2}.$$

Let $N_t \in \mathbb{N}$, for $t \in \mathbb{T}$, be such that $N \geq N_t$ implies $|c_N(t)| < \frac{\epsilon}{2}$. These two inequalities together imply that if $N \geq N_t$ and $|t - t'| < \delta$ then $|c_N(t')| < \epsilon$.

By choosing a finite open cover of $[0, 2\pi]$ with intervals of length δ and letting N_0 be the maximum of the N_t's corresponding to their centers, we get that $N \geq N_0$ implies $|c_N(t)| < \epsilon$ for all $t \in \mathbb{T}$. Thus the convergence $c_N(t) \to 0$ is uniform. A similar argument shows that

$$\begin{aligned} \lim_{N\to\infty} \frac{\left\| C\big(SW_{2N,\tau_N} g(t)\big) \right\|^2}{2N+1} &= \lim_{\tau_N \to 0} \frac{L}{2\pi} \sum_{n=0}^{2N} \tau_N \big(g(t+n\tau_N) - c_N(t) \big)^2 \\ &= \frac{1}{2\pi} \int_0^{2\pi} g(r)^2 dr = 1 \end{aligned}$$

uniformly in $t \in \mathbb{T}$, and replacing g by $\frac{f - \widehat{f}(0)}{\left\| f - \widehat{f}(0) \right\|_2}$ yields the result. □

Remark 6.5. Notice that an alternative proof of Proposition 6.4 follows from combining the Structure Theorem 5.6(2), Parseval's Theorem, and the fact that

$$\|S_N f\|_2^2 - \widehat{f}(0)^2 = \left\| S_N \left(f - \widehat{f}(0) \right) \right\|_2^2$$

Theorem 6.6 (**Convergence I**). *Let $f \in C^1(\mathbb{T})$ be an L-periodic function, $N \in \mathbb{N}$, $\tau_N = \frac{2\pi}{L(2N+1)}$, $T \subset \mathbb{T}$ finite, and let $\overline{Y}_N$ be as in Corollary 6.2. Let $\overline{X}_N$ be the set resulting from pointwise centering and normalizing the point cloud*

$$SW_{2N,\tau_N} f(T) \subset \mathbb{R}^{2N+1}.$$

Then for any field of coefficients, the sequence $\mathit{dgm}(\overline{X}_N)$ of persistence diagrams is Cauchy with respect to d_B, and

$$\lim_{N\to\infty} \mathit{dgm}(\overline{X}_N) = \lim_{N\to\infty} \mathit{dgm}(\overline{Y}_N) = \mathit{dgm}_\infty(f, T, w).$$

Proof. To prove Theorem 6.6 we will use the Approximation Theorem 4.5 to show that

$$\lim_{N\to\infty} d_B\Big(\mathsf{dgm}(\overline{X}_N), \mathsf{dgm}(\overline{Y}_N)\Big) = 0$$

and combine this with Corollary 6.2 to obtain the result.

Assume without loss of generality that f satisfies $\widehat{f}(0) = 0$ and $\|f\|_2 = 1$. Let X_N and Y_N be the resulting sets from pointwise centering the point clouds $SW_{2N,\tau_N} f(T)$ and $SW_{2N,\tau_N} S_N f(T)$, respectively. Using the uniform convergence in equation 5 we get that

$$\lim_{N\to\infty} d_H\left(\overline{X}_N, \frac{X_N}{\sqrt{2N+1}}\right) = 0$$

and moreover, since $\lim_{N\to\infty} \|S_N f\|_2 = \|f\|_2 = 1$ then

$$\lim_{N\to\infty} d_H\left(\frac{X_N}{\sqrt{2N+1}}, \frac{X_N}{\sqrt{2N+1}\|S_N f\|_2}\right) = 0.$$

Now, from the Structure Theorem 5.6(2) we have the identity

$$\overline{Y}_N = \frac{Y_N}{\sqrt{2N+1}\|S_N f\|_2}$$

and using the Approximation Theorem 4.5(1), along with the fact that C is distance non-increasing, we conclude that

$$\lim_{N\to\infty} d_H\left(\frac{X_N}{\sqrt{2N+1}\|S_N f\|_2}, \overline{Y}_N\right) \leq \lim_{N\to\infty} \frac{\sqrt{2}\cdot\|R_N f'\|_2}{\|S_N f\|_2 \cdot \sqrt{N+1}} = 0.$$

The triangular inequality then implies that $\lim_{N\to\infty} d_H(\overline{X}_N, \overline{Y}_N) = 0$ and the result follows from combining the stability of d_B with respect to d_H, and Corollary 6.2. □

The first convergence theorem asserts that for each choice of discretization $T \subset \mathbb{T}$ one obtains a limiting diagram $\mathsf{dgm}_\infty(f, T, w)$, by letting $N \to \infty$ in the pointwise centered and normalized versions of either $SW_{2N,\tau_N} S_N f(T)$ or $SW_{2N,\tau_N} f(T)$. Next we will show that there is also convergence when T tends to $\mathbb{T}$, with respect to the Hausdorff distance on subspaces of $\mathbb{T}$.

Theorem 6.7 (**Convergence II**). *Let $T, T' \subset \mathbb{T}$ be finite, and let $f \in C^1(\mathbb{T})$ be L-periodic with modulus of continuity $\omega : [0,\infty] \longrightarrow [0,\infty]$. If $w = \frac{2\pi}{L}$ then*

$$d_B\big(\mathit{dgm}_\infty(f, T, w), \mathit{dgm}_\infty(f, T', w)\big) \leq 2\left\|f - \widehat{f}(0)\right\|_2 \omega\left(d_H\big(T, T'\big)\right)$$

and thus there exists a persistence diagram $\mathit{dgm}_\infty(f, w)$ so that

$$\lim_{T\to\mathbb{T}} \mathit{dgm}_\infty(f, T, w) = \mathit{dgm}_\infty(f, w).$$

Proof. Fix $t \in T$ and $t' \in T'$. If we let $\mathbf{x}_N = SW_{2N,\tau_N} f(t)$, $\mathbf{x}'_N = SW_{2N,\tau_N} f(t')$, $\tau_N = \frac{2\pi}{L(2N+1)}$ and $\lambda = \left\| f - \widehat{f}(0) \right\|_2$ then

$$\begin{aligned} \left\| \frac{C(\mathbf{x}_N)}{\|C(\mathbf{x}_N)\|} - \frac{C(\mathbf{x}'_N)}{\|C(\mathbf{x}'_N)\|} \right\| &\leq \left\| \frac{C(\mathbf{x}_N)}{\|C(\mathbf{x}_N)\|} - \frac{\lambda C(\mathbf{x}_N)}{\sqrt{2N+1}} \right\| + \frac{\lambda \left\| C(\mathbf{x}_N) - C(\mathbf{x}'_N) \right\|}{\sqrt{2N+1}} \\ &\quad + \left\| \frac{C(\mathbf{x}'_N)}{\|C(\mathbf{x}'_N)\|} - \frac{\lambda C(\mathbf{x}'_N)}{\sqrt{2N+1}} \right\| \end{aligned}$$

It follows from Proposition 6.4 that both the summand

$$\left\| \frac{C(\mathbf{x}_N)}{\|C(\mathbf{x}_N)\|} - \frac{\lambda C(\mathbf{x}_N)}{\sqrt{2N+1}} \right\| = \frac{\|C(\mathbf{x}_N)\|}{\sqrt{2N+1}} \cdot \left| \frac{\sqrt{2N+1}}{\|C(\mathbf{x}_N)\|} - \lambda \right|$$

and its version with $\mathbf{x}'_N$, go to zero as $N \to \infty$. Thus given $\epsilon > 0$ there exists $N_0 \in \mathbb{N}$ so that $N \geq N_0$ implies

$$\begin{aligned} \left\| \frac{C(\mathbf{x}_N)}{\|C(\mathbf{x}_N)\|} - \frac{C(\mathbf{x}'_N)}{\|C(\mathbf{x}'_N)\|} \right\| &\leq \frac{\epsilon}{2} + \frac{\lambda \left\| C(\mathbf{x}_N) - C(\mathbf{x}'_N) \right\|}{\sqrt{2N+1}} \\ &\leq \frac{\epsilon}{2} + \frac{\lambda \|\mathbf{x}_N - \mathbf{x}'_N\|}{\sqrt{2N+1}} \\ &= \frac{\epsilon}{2} + \lambda \left(\sum_{n=0}^{2N} \frac{\left| f(t + n\tau_N) - f(t' + n\tau_N) \right|^2}{2N+1} \right)^{1/2} \\ &\leq \frac{\epsilon}{2} + \lambda \omega(|t - t'|). \end{aligned}$$

Let $\overline{X}_N$ and $\overline{X'}_N$ be the sets resulting from pointwise centering and normalizing $SW_{2N,\tau_N} f(T)$ and $SW_{2N,\tau_N} f(T')$, respectively. Since the estimates above are uniform in t and t' (by Proposition 6.4), it follows that whenever $N \geq N_0$ then

$$d_H\big(\overline{X}_N, \overline{X'}_N\big) \leq \frac{\epsilon}{2} + \lambda \omega\big(d_H(T, T')\big).$$

Notice that the Hausdorff distance on the left hand side is for subspaces of $\mathbb{R}^{2N+1}$, while the one on the right is between subspaces of $\mathbb{T}$.

Applying the Stability Theorem for persistence diagrams yields

$$d_B\big(\mathsf{dgm}(\overline{X}_N), \mathsf{dgm}(\overline{X'}_N)\big) \leq \epsilon + 2\lambda \omega\big(d_H(T, T')\big)$$

which by letting $N \to \infty$ and applying the First Convergence Theorem (6.6), implies

$$d_B\big(\mathsf{dgm}_\infty(f, T, w), \mathsf{dgm}_\infty(f, T', w)\big) \leq \epsilon + 2\lambda \omega\big(d_H(T, T')\big).$$

Since this is true for any $\epsilon > 0$, letting $\epsilon \downarrow 0$ yields the first part of the theorem. The existence of $\mathsf{dgm}_\infty(f, w)$ follows from the fact that the set of generalized persistence diagrams is complete with respect to d_B. □

6.2. **A lower bound for maximum persistence.** The Structure Theorem 5.6(3) and the fact that orthogonal projections are distance non-increasing, allow us to now prove the following:

Theorem 6.8. *Let $f \in C^1(\mathbb{T})$ be an L-periodic function, $N \in \mathbb{N}$, $M \geq 2N$, $L(M+1)\tau = 2\pi$ and let $T \subset \mathbb{T}$ be finite. Furthermore, assume that $d_H(T, \mathbb{T}) < \delta$ for some (see Theorem 5.6(3))*

$$0 < \delta < \max_{1 \leq n \leq N} \frac{\sqrt{3}\widetilde{r}_n}{\kappa_N} \quad , \qquad \textit{where} \quad \kappa_N = \frac{2\sqrt{2}\,\|S_N f'\|_2}{\left\| S_N\left(f - \widehat{f}(0)\right)\right\|_2}$$

Let $\overline{Y} = \overline{Y}_N$ be the set resulting from pointwise centering and normalizing the point cloud

$$SW_{M,\tau} S_N f(T) \subset \mathbb{R}^{M+1},$$

and let $p > N$ be a prime. If $\mathsf{dgm}(\overline{Y})$ denotes the 1-dimensional $\mathbb{F}_p$-persistence diagram for the Rips filtration on $\overline{Y}$, then φ_τ yields an element $\mathbf{x}_\varphi \in \mathsf{dgm}(\overline{Y})$ with

(1) $\mathsf{birth}(\mathbf{x}_\varphi) \leq \delta\kappa_N$

(2) $\mathsf{death}(\mathbf{x}_\varphi) \geq \sqrt{3} \max\limits_{1 \leq n \leq N} \widetilde{r}_n$

and therefore

$$mp\Big(\mathsf{dgm}(\overline{Y})\Big) \geq \left(\sqrt{3} \max_{1 \leq n \leq N} \widetilde{r}_n\right) - \delta\kappa_N \tag{6}$$

Proof. Given the linear decomposition

$$\varphi_\tau(t) \quad = \sum_{\substack{n\,=\,1 \\ n \,\equiv\, 0 \pmod{L}}}^{N} \widetilde{r}_n\big(\cos(nt)\widetilde{\mathbf{x}}_n + \sin(nt)\widetilde{\mathbf{y}}_n\big)$$

of $\varphi_\tau(t)$ with respect to the orthonormal set $\big\{\widetilde{\mathbf{x}}_n, \widetilde{\mathbf{y}}_n \mid 1 \leq n \leq N,\ n \equiv 0 \pmod{L}\big\}$ described in the proof of Theorem 5.6(3), it follows that

$$\begin{array}{rccc} P_n: & \overline{Y} & \longrightarrow & \mathbb{C} \\ & \varphi_\tau(t) & \mapsto & \widetilde{r}_n e^{int} \end{array}$$

can be regarded as the restriction to $\overline{Y}$ of the orthogonal projection from $\mathbb{R}^{M+1}$ onto $Span\{\widetilde{\mathbf{x}}_n, \widetilde{\mathbf{y}}_n\}$. Since orthogonal projections are linear and norm-non-increasing, then $\|P_n(\mathbf{x}) - P_n(\mathbf{y})\| \leq \|\mathbf{x} - \mathbf{y}\|$ for every $\mathbf{x}, \mathbf{y} \in \overline{Y}$. Thus, if

$$S^1(\widetilde{r}_n) = \{\widetilde{r}_n e^{int} \mid t \in T\}$$

it follows that P_n induces simplicial maps

$$\begin{array}{rccc} P_{n\sharp}: & R_\epsilon\left(\overline{Y}\right) & \longrightarrow & R_\epsilon\left(S^1(\widetilde{r}_n)\right) \\ & [\mathbf{x}_0, \ldots, \mathbf{x}_k] & \mapsto & [P_n(\mathbf{x}_0), \ldots, P_n(\mathbf{x}_k)] \end{array}$$

for every $\epsilon > 0$, which in turn yield homomorphisms

$$P_{n_*}: H_k\left(R_\epsilon\left(\overline{Y}\right); \mathbb{F}_p\right) \longrightarrow H_k\left(R_\epsilon\left(S^1(\widetilde{r}_n)\right); \mathbb{F}_p\right)$$

of $\mathbb{F}_p$-vector spaces at the homology level. What we contend is that, via the homomorphisms P_{n_*}, the maximum 1-dimensional persistence of $\overline{Y}$ can be bounded below by that of $S^1(\widetilde{r}_n)$. Indeed, let $\epsilon_1, \epsilon_2 > 0$ be so that

$$\delta\kappa_N < \epsilon_1 < \epsilon_2 < \sqrt{3}\widetilde{r}_m$$

where $m = \arg\max\{\widetilde{r}_n \mid 1 \leq n \leq N\}$. If we write

$$T = \{t_0 < t_2 < \cdots < t_J\}$$

it follows from $d_H(T,\mathbb{T}) < \delta$ that $|t_j - t_{j-1}| < 2\delta$ for all $j = 1,\ldots,J$, and therefore

$$\begin{aligned}
\|\varphi_\tau(t_j) - \varphi_\tau(t_{j-1})\|^2 &= \sum_{\substack{n=1 \\ n \equiv 0 \pmod{L}}}^{N} 2\widetilde{r}_n^2\Big(1 - \cos\big(n(t_j - t_{j-1})\big)\Big) \\
&\leq \sum_{\substack{n=1 \\ n \equiv 0 \pmod{L}}}^{N} \widetilde{r}_n^2\big(n(t_j - t_{j-1})\big)^2 \\
&= (t_j - t_{j-1})^2 \sum_{n=1}^{N} \frac{4n^2\left|\widehat{f}(n)\right|^2}{\|S_N f\|_2^2 - \widehat{f}(0)^2} \\
&= \frac{(t_j - t_{j-1})^2}{\left\|S_N\left(f - \widehat{f}(0)\right)\right\|_2^2} \sum_{1 \leq |n| \leq N} 2\left|\widehat{f'}(n)\right|^2 \\
&\leq 8\delta^2 \frac{\|S_N f'\|_2^2}{\left\|S_N\left(f - \widehat{f}(0)\right)\right\|_2^2} \\
&= (\delta\kappa_N)^2
\end{aligned}$$

The first inequality is a consequence of the Taylor expansion for $\cos(x)$ around zero, and f' denotes the first derivative of f. Therefore

$$\nu = [\varphi_\tau(t_0), \varphi_\tau(t_1)] + \cdots + [\varphi_\tau(t_{J-1}), \varphi_\tau(t_J)] + [\varphi_\tau(t_J), \varphi_\tau(t_0)]$$

is a 1-dimensional cycle on $R_{\epsilon_1}(\overline{Y})$, and we obtain the homology class

$$P_{m_*}([\nu]) \in H_1\left(R_{\epsilon_1}\left(S^1(\widetilde{r}_m)\right); \mathbb{F}_p\right).$$

Let $\{\theta_0 < \theta_1 < \cdots < \theta_{J_m}\} = \left\{t \bmod \frac{2\pi}{m} \mid t \in T\right\}$ and let $c_j = \widetilde{r}_m e^{im\theta_j}$. It follows from a similar calculation that

$$\|c_j - c_{j-1}\|^2 \leq (\theta_j - \theta_{j-1})^2 \frac{4\left|\widehat{f'}(m)\right|^2}{\left\|S_N\left(f - \widehat{f}(0)\right)\right\|_2^2} \leq (\delta\kappa_N)^2$$

and therefore the 1-cycle

$$\mu = [c_0, c_1] + \cdots + [c_{J_m - 1}, c_{J_m}] + [c_{J_m}, c_0]$$

is so that its homology class $[\mu] \in H_1\left(R_{\epsilon_1}\left(S^1(\widetilde{r}_m)\right); \mathbb{F}_p\right)$ satisfies $i_*([\mu]) \neq 0$, where i_* is the homomorphism induced by the inclusion

$$i : R_{\epsilon_1}\left(S^1(\widetilde{r}_m)\right) \hookrightarrow R_{\epsilon_2}\left(S^1(\widetilde{r}_m)\right).$$

Since $P_{m_*}([\nu]) = m[\mu]$, and given that $1 \leq m \leq N < p$ implies that m is invertible in $\mathbb{F}_p$, then $i_* \circ P_{m_*}([\nu]) \neq 0$. From the commutativity of the diagram

$$\begin{array}{ccc}
H_1(R_{\epsilon_1}(\overline{Y}); \mathbb{F}_p) & \xrightarrow{\quad i_* \quad} & H_1(R_{\epsilon_2}(\overline{Y}); \mathbb{F}_p) \\
\Big\downarrow{\scriptstyle P_{m_*}} & & \Big\downarrow{\scriptstyle P_{m_*}} \\
H_1\left(R_{\epsilon_1}\left(S^1(\widetilde{r}_m)\right); \mathbb{F}_p\right) & \xrightarrow[i_*]{} & H_1\left(R_{\epsilon_2}\left(S^1(\widetilde{r}_m)\right); \mathbb{F}_p\right)
\end{array}$$

we conclude that $i_*([\nu]) \neq 0$, and thus $[\nu]$ yields an element $\mathbf{x}_\varphi \in \mathsf{dgm}(\overline{Y})$ so that

$$\mathsf{birth}(\mathbf{x}_\varphi) \leq \epsilon_1 \quad \text{and} \quad \mathsf{death}(\mathbf{x}_\varphi) \geq \epsilon_2.$$

Given that this is true for every $\epsilon_1 > \delta\kappa_N$ and every $\epsilon_2 < \sqrt{3}\widetilde{r}_m$, letting $\epsilon_1 \downarrow \delta\kappa_N$ and $\epsilon_2 \uparrow \sqrt{3}\widetilde{r}_m$ concludes the proof. □

Remark 6.9. It is worth noting that in the proof of Theorem 6.8 one can replace $\mathbb{F}_p$, $p > N$, by the field of rational numbers $\mathbb{Q}$. That is, the estimated bound for maximum persistence is valid for all $N \in \mathbb{N}$ and homology with $\mathbb{Q}$ coefficients.

Equation 6, together with the Convergence Theorems I (6.6) and II (6.7), imply:

Corollary 6.10. *Let $f \in C^1(\mathbb{T})$ be an L-periodic function so that $\widehat{f}(0) = 0$ and $\|f\|_2 = 1$. Let $T \subset \mathbb{T}$ be finite and so that $d_H(T, \mathbb{T}) < \delta$ for some*

$$0 < \delta < \frac{\sqrt{3}}{\sqrt{2}\|f'\|_2} \max_{n\in\mathbb{N}} \left|\widehat{f}(n)\right|$$

Then with $\mathbb{Q}$ coefficients, the 1-dimensional persistence diagram $\mathsf{dgm}_\infty(f, T, w)$ satisfies

$$\frac{1}{2}mp\Big(\mathsf{dgm}_\infty(f, T, w)\Big) \geq \sqrt{3}\max_{n\in\mathbb{N}} \left|\widehat{f}(n)\right| - \sqrt{2}\delta\|f'\|_2$$

and therefore

$$mp\Big(\mathsf{dgm}_\infty(f, w)\Big) \geq 2\sqrt{3}\max_{n\in\mathbb{N}} \left|\widehat{f}(n)\right|.$$

6.3. **The field of coefficients.** One question worth asking is whether the lower bound for maximum persistence presented in Theorem 6.8, is in fact dependent on the field of coefficients. More generally, one would like to determine if the full persistence diagram has such dependency. To this end, let us consider the functions

$$\begin{aligned} g_1(t) &= 0.6\cos(t) + 0.8\cos(2t) \\ g_2(t) &= 0.8\cos(t) + 0.6\cos(2t). \end{aligned}$$

We construct their associated sliding window point clouds, $SW_{M,\tau}g_1(T)$ and $SW_{M,\tau}g_2(T)$, using $M = 4$, $\tau = 2\pi/5$ and $T = \left\{\frac{2\pi k}{150} \mid k = 0, 1, \ldots, 150\right\}$. After pointwise centering and normalizing, we compute their 1-dimensional persistent homology with coefficients in $\mathbb{F}_2$ and $\mathbb{F}_3$. For this, we use a fast implementation of 1-dimensional persistent homology, based on the Union-Find algorithm and the work of Mischaikow and Nanda [20]. Details of this implementation will appear in [23]. We summarize the results in figure 3.

This example shows that, at least in low dimensions, the persistent homology of sliding window point clouds is coefficient-dependent. Let us see why this is the case. If $(r_1, r_2) \in \mathbb{R}^2$ is so that $r_1^2 + r_2^2 = 1$ and $r_1 r_2 \neq 0$, it follows from the Structure Theorem 5.6(3) that if $\alpha_1, \alpha_2 \in [0, 2\pi]$ and

$$g(t) = r_1\cos(t - \alpha_1) + r_2\cos(2t - \alpha_2)$$

then for every $t \in [0, 2\pi]$ and $M \geq 4$,$\tau = \frac{2\pi}{M+1}$, one has that

$$\varphi_\tau(t) = \frac{C\left(SW_{M,\tau}g(t)\right)}{\|C\left(SW_{M,\tau}g(t)\right)\|}$$

can be isometrically identified with

$$\widetilde{\varphi}(t) = \left(r_1 e^{it}, r_2 e^{2it}\right) \in \mathbb{C}^2.$$

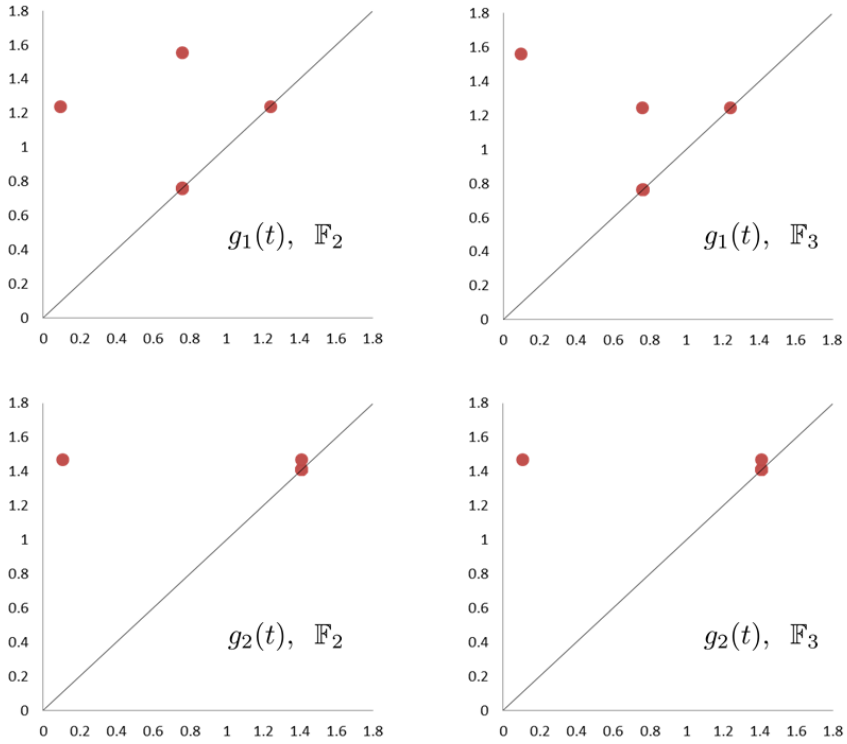

FIGURE 3. 1-dimensional $\mathbb{F}_p$-persistence diagrams for the centered and normalized sliding window point clouds on g_i. Here the columns correspond to $p = 2, 3$ and the rows to $i = 1, 2$.

Let us use $\widetilde{\varphi}$ instead of φ_τ for the persistent homology computation. The first thing to notice is that the image of $\widetilde{\varphi}$ can be realized as the boundary of a Möbius strip. Indeed, consider the map

$$\begin{array}{rccc} \mathcal{M}: & [0,\pi]\times[-1,1] & \longrightarrow & \mathbb{C}^2 \\ & (t,s) & \mapsto & \left(-sr_1e^{it}, r_2e^{2it}\right). \end{array}$$

It follows that $\mathcal{M}$ is a continuous injection on $[0,\pi)\times[-1,1)$, since $r_1r_2 \neq 0$, and that it descends to an embedding of the quotient space

$$\widetilde{\mathcal{M}} : \left([0,\pi]\times[-1,1]/\sim\right) \longrightarrow \mathbb{C}^2$$

where $(0,s) \sim (\pi,-s)$ for every $s \in [-1,1]$. Notice that $[0,\pi]\times[-1,1]/\sim$ serves as the usual model for the Möbius strip, and that $\partial\Big(Img(\widetilde{\mathcal{M}})\Big) = Img(\widetilde{\varphi})$.

Let $T = \{t_0 < t_2 < \cdots < t_J\}$ be a δ-dense subset of $[0, 2\pi]$, $X = \widetilde{\varphi}(T)$, and let

$$[\nu] \in H_1\big(R_r(X); \mathbb{F}_2\big)$$

for $r > 4\delta$, be the homology class of the 1-cycle

$$\nu = [\widetilde{\varphi}(t_0), \widetilde{\varphi}(t_1)] + \cdots + [\widetilde{\varphi}(t_{J-1}), \widetilde{\varphi}(t_J)] + [\widetilde{\varphi}(t_J), \widetilde{\varphi}(t_0)].$$

It can be readily checked that if we let

$$V = \Big\{(t,s) \;\Big|\; (t,s) \in \big(T \cap [0,\pi)\big)\times\{-1\} \quad \text{or} \quad (t+\pi, s) \in \big(T\cap(\pi, 2\pi]\big)\times\{1\}\Big\},$$

then there exists a triangulation of $Img(\widetilde{\mathcal{M}})$ having $\widetilde{\mathcal{M}}(V)$ as vertex set, and so that if we take coefficients in $\mathbb{F}_2$, then the formal sum of its triangles yields a 2-chain Σ with $\partial_2(\Sigma) = \nu$. Moreover, since T is δ-dense in $[0, 2\pi]$ and for all $t \in [0, \pi]$

$$\begin{aligned}\left\|\widetilde{\mathcal{M}}(t,-1) - \widetilde{\mathcal{M}}(t \pm \delta, 1)\right\|^2 &= 2\big[r_1^2(1+\cos(\delta)) + r_2^2(1-\cos(2\delta))\big] \\ &\leq 2\left[r_1^2\left(2 - \frac{\delta^2}{2}\right) + 2r_2^2\delta^2\right] \\ &= r_1^2(4 - 5\delta^2) + 4\delta^2\end{aligned}$$

if $\delta > 0$ is small, then we can choose Σ so that

$$\Sigma \in C_2\big(R_{r'}(X); \mathbb{F}_2\big) \quad , \quad r' = r_1\sqrt{4 - 5\delta^2} + 2\delta.$$

In summary, if

$$r_1\sqrt{4 - 5\delta^2} + 2\delta < \sqrt{3}\, r_2 \tag{7}$$

then the death-time of the class $[\nu]$ is less than or equal to $r_1\sqrt{4 - 5\delta^2} + 2\delta$ with coefficients in $\mathbb{F}_2$, but larger than $\sqrt{3}\, r_2$ (by Theorem 6.8) with coefficients in $\mathbb{F}_p$ for any prime $p \geq 3$. Moreover, with coefficients in $\mathbb{F}_2$ and provided equation 7 holds (e.g. for g_1), the first edge across the Möbius band prompts the birth of a new class corresponding to the equator

$$t \mapsto \widetilde{\mathcal{M}}(t, 0) = (0, r_2 e^{2it})$$

of the embedded Möbius strip. This class, in turn, survives up to $\sqrt{3}r_2$. With coefficients in $\mathbb{F}_3$, on the other hand, the equatorial and boundary classes will be in the same persistence class once all the 2-simplices in the Möbius band have been added. This results in the death of the class which was born later, i.e. the one represented by the equator.

7. Examples: Quantifying Periodicity of Sampled Signals

We present in this section two experiments to test our ideas: First, the ranking of signals by periodicity alone, in a way which is invariant to the shape of the periodic pattern; and second, the accurate classification of a signal as periodic or non-periodic at different noise levels. A detailed description of our methods is provided below, but roughly speaking, we associate to each sampled signal $S = [s_1, \ldots, s_J]$ a real valued function f_S by cubic spline interpolation, construct its centered and normalized sliding window point cloud X_S, and let

$$\frac{mp\big(\mathsf{dgm}(X_S)\big)}{\sqrt{3}} = Score(S)$$

be its periodicity score. We then compare it to those obtained with the JTK_CYCLE [15], Lomb-Scargle [13, 18, 24] and Total Persistent Homology [7] algorithms.

7.1. Shape Independence.

For this experiment we construct ten different shapes: A 2-periodic pure cosine-like curve, a 2-periodic cosine-like function plus three levels of gaussian noise (variances at 25%, 50%, and 75% of the signal's amplitude), a noisy saw-tooth (noise level at 25% of the signal's amplitude), a function of the form $\cos(\phi(t))$ for $\phi(t) = e^{at+b}$, a noisy and damped cosine-like curve with three periods, a spiky signal with three periods, a noisy square wave with two

periods, and a 1-periodic function of the form $Re\left(\sum_{n=1}^{5} \widehat{f}(n)e^{2int}\right)$ for $\widehat{f}(n)$ drawn randomly and uniformly from the unit disk in $\mathbb{C}$. Each function is then evaluated at 50 evenly spaced time points, yielding the sampled signals $[s_1, \ldots, s_{50}]$ which we input into the algorithms. For constructing the sliding window point clouds we use $N = 10$, coefficients in $\mathbb{F}_{11}$, $L = 2, 3, 4$ and report the best score. In all the other algorithms we set the parameters to their suggested or default values. The results are summarized in figure 4.

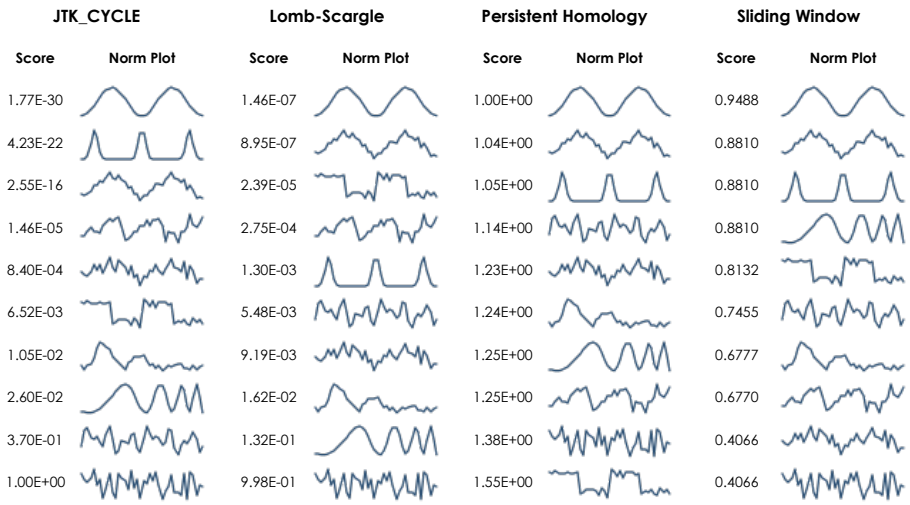

FIGURE 4. Ranking of signals by periodicity. For each algorithm we provide the score and a normalized plot of the signal. The ranking goes from top (highest score) to bottom (lowest score).

Two things are worth noting: First, except for the Sliding Windows method (SW1PerS), all other algorithms have clear preferences for the type of shape they consider to be most periodic. These biases are of course part of the wiring of the algorithms, and were to be expected. The second thing to notice has to do with the distribution of scores and their relative differences. Methods such as JTK or Lomb-Scargle define their periodicity score in terms of p-values, which are extremely difficult to interpret. Our scoring method, by way of contrast, has a clear geometric interpretation and a reasonable distribution.

7.2. **Classification Rates.** We compare the different algorithms by their ability to separate periodic from non-periodic signals. The performance of this type of binary classification can be visualized using a Receiver Operator Characteristic (ROC) plot, which compares the True Positive Rate (TPR) to the False Positive Rate (FPR) as a cutoff on the scores is varied. Here the TPR is the proportion of correctly identified positive cases out of all positives, and FPR is the proportion of negative cases incorrectly identified as positives out of all the negatives. The line TPR=FPR is the performance of random guessing; the higher the ROC curve is above this line, the better its classification performance. An algorithm that is able to perfectly separate all positive from negative test cases would have a ROC curve

that passes through the point TPR=1 and FPR=0. It follows that a reasonable measure of classification success for a particular method, is the area under its ROC curves.

The synthetic data is generated as follows: The periodic signals (positive cases) span two periods and include a cosine, cosine with trending, cosine with damping, and cosine with increased peak steepness. The non-periodic signals (negative cases) include a constant and a linear function. We generate 100 profiles from each shape by adjusting its phase. For instance, in the case of the cosine shape we let

$$f_i(t) = \cos\left(2t - \frac{j\pi}{50}\right), \quad j = 0, \ldots, 99$$

be the profiles. We sample each of the 600 profiles at 50 evenly spaced time points $t \in [0, 2\pi]$, and add gaussian noise with standard deviation at 0%, 25% and 50% of the signal's amplitude. Please refer to Figure 5 for examples.

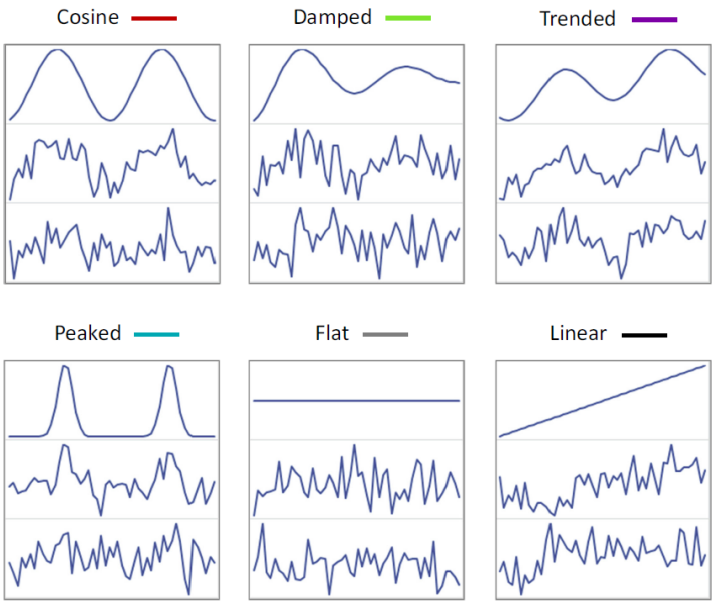

FIGURE 5. Examples of signals in the synthetic data. We show one signal from each profile type at noise levels 0%, 25% and 50%.

Remark 7.1. There are two reason why we regard constant functions as non-periodic. On the one hand, the intended application for SW1PerS (Sliding Windows and 1-Persistence Scoring) is to identify genes that are both relevant and exhibit a periodic expression pattern with respect to time. Relevance in this case means that changes in expression-level translate into physiological phenomena. The second reason has to do with the philosophy of the proposed method: We quantify periodicity as the prominence of 1-homology classes in the sliding window point cloud. Since this point cloud for a constant function is only a point, it does not have 1-homology and hence is interpreted as coming from a non-periodic function.

For the Sliding Windows + 1D $\mathbb{F}_p$-Persistence computation we let $N = 10$, $L = 2$ and $p = 11$. In order to address noise, we include a layer of (simple) moving average

at the sampled signal level, and one iteration of mean-shift [8] at the sliding window point cloud level. For the moving average we fix a window size with 7 data points, and use a cubic spline of this denoised signal to populate the point cloud. Mean-shift on a pointwise centered and normalized point cloud X_S was implemented as follows: Given a point $x \in X_S$, we let $\overline{x}$ be the mean of the set

$$\{y \in X_S : 1 - (x \cdot y) < \epsilon\}$$

where $\epsilon = \cos\left(\frac{\pi}{16}\right)$ and $x \cdot y$ denotes the Euclidean inner product of x and y. In other words, $\overline{x}$ is the mean of the ϵ-neighbors of x if distance is measured with cosine similarity. We obtain the mean-shifted point cloud

$$\overline{X}_S = \left\{ \frac{\overline{x}}{\|\overline{x}\|} \ : \ x \in X_S \right\}$$

which we now use for the persistent homology computation. We report our results in figure 6.

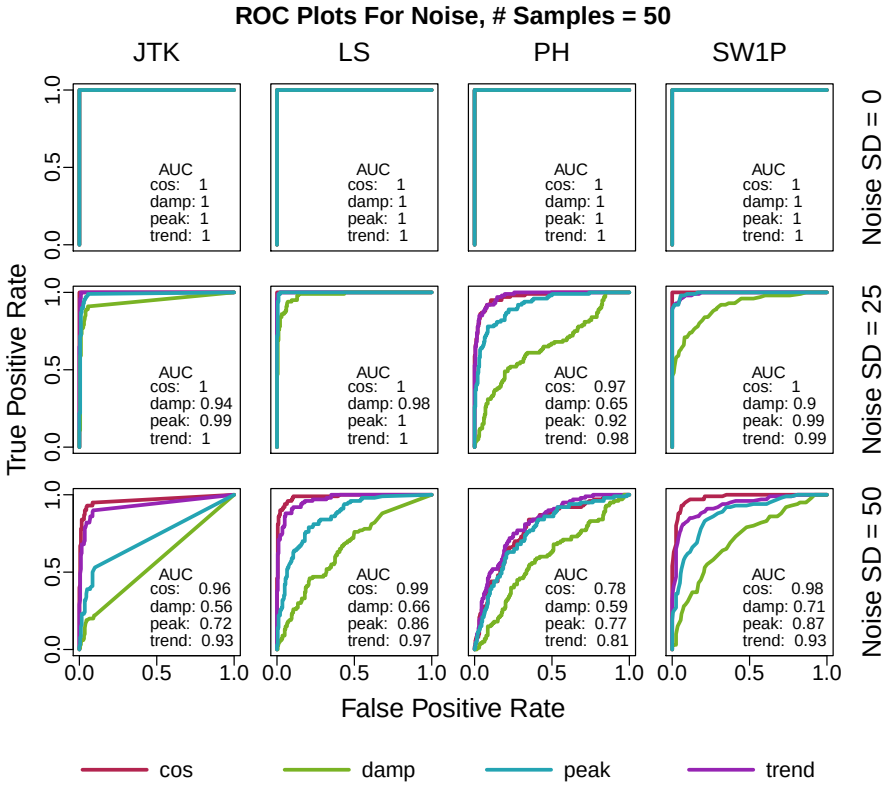

FIGURE 6. ROC curves: True Positive Rate vs False Positive Rate. We compare the classification success of each algorithm on the synthetic data set using the area under its ROC curves. Curves are colored according to the type of periodic shape, and the area under the curve (AUC) is reported. Please refer to an electronic version for colors.

The Lomb-Scargle periodogram is considered to be one of the best methods for detecting periodicity, and its ROC curves support this belief. The fact that it is attuned to favoring cosine-like curves makes it very resilient to dampening, trending and noise. It was thus a great surprise to see that our method performs comparably well in all cases, except for trended cosines and cosines, and that outperforms it for peaked and damped profiles at high noise levels.

A final point we would like to make, is that denoising really is a crucial element of the SW1PerS pipeline. We show in Figure 7 the results of quantifying periodicity on the raw synthetic data, i.e. without applying denoising. As one can see, in the absence of pre-processing, the results degrade considerably.

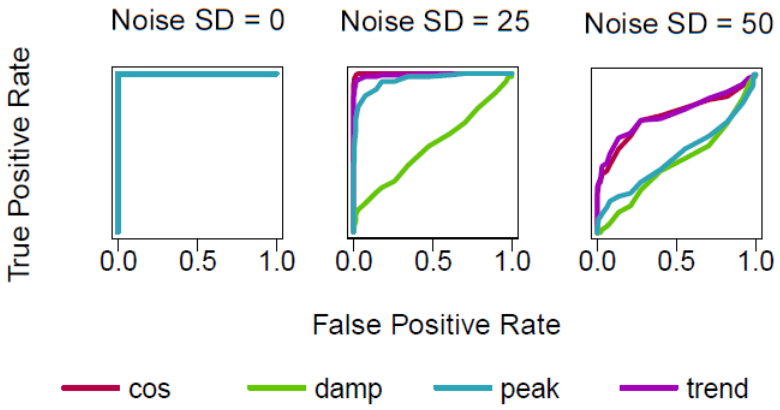

FIGURE 7. ROC curves for the SW1PerS analysis on the raw synthetic data. That is, we do not apply moving average or mean-shift. For a comparison on the effect of denoising, please refer to the rightmost column in Figure 6.

8. Final Remarks

We prove in this paper results which describe the structure of persistence diagrams obtained by sliding window embeddings of time series. The main tools for the analysis were a Fourier series approximation argument and the Stability Theorem for persistence diagrams. These results we obtained, provide explicit information about how diagrams from sliding window point clouds depend on the embedding dimension, window size and field of coefficients. We then present examples of the effectiveness of our method for quantifying periodicity of time series data. The experimental side of this framework will be explored more in depth in future work.

This paper also presents the first full *theoretical analysis* of the use of persistent homology to find structure in time series data. Time delay embeddings as a means to analyze signals is not new, rather it is a well-established method in dynamical systems and in image analysis. And the use of computation topology methods to find structure in transformed data has already been considered experimentally before, notably in [4] and [10]. This paper, however, is the first to provide a theoretical analysis of the dependency of persistence on embedding dimension and window size.

There are some interesting new aspects of the use of persistence in this method. It provides one of the first examples where persistent **homology** with coefficients other than $\mathbb{F}_2$ is required. Other notable examples include [4], where coefficients in $\mathbb{F}_3$ were essential to discover the embedded Klein bottle, and [9, 10] where the authors start with a 1D persistent **cohomology** class mod $\mathbb{F}_p$, and lift it to an integer 1D class to get a map to the circle. They do this by choosing p so that the relevant homomorphism in the Bockstein long exact sequence is surjective. Their approach for real data is to choose a prime p at random, and evaluate if $H^2(X, \mathbb{Z})$ has p-torsion. If it does, they then choose another prime. By contrast, we have established exactly which primes could be problematic and can avoid them in advance.

It was highlighted in Remark 4.4 that maximum persistence, the main feature for periodicity we study in this paper, satisfies $mp(\mathsf{dgm}) = 2d_B(\mathsf{dgm}, \mathsf{dgm}_\Delta)$. This is in fact part of a bigger picture: Indeed, the q-Wasserstein distance between two persistence diagrams dgm_1 and dgm_2 is defined by

$$W_q(\mathsf{dgm}_1, \mathsf{dgm}_2) = \min_{\phi} \Big(\sum_{\mathbf{x} \in \mathsf{dgm}_1} ||\mathbf{x} - \phi(\mathbf{x})||_\infty^q \Big)^{\frac{1}{q}}$$

where $\phi : \mathsf{dgm}_1 \to \mathsf{dgm}_2$ is a matching of dgm_1 with dgm_2. As with the Bottleneck distance (Remark 4.4) one can show that

$$W_q(\mathsf{dgm}, \mathsf{dgm}_\Delta) = \frac{1}{2} \Big(\sum_{(x,y) \in \mathsf{dgm}} (y - x)^q \Big)^{\frac{1}{q}}$$

and since $W_q \to d_B$ as $q \to \infty$, then $2W_q(\mathsf{dgm}, \mathsf{dgm}_\Delta)$ can be regarded as a smoother version of $mp(\mathsf{dgm})$. When dgm comes from a sliding window point cloud, $2W_q(\mathsf{dgm}, \mathsf{dgm}_\Delta)$ can be interpreted as a sequence of signatures, or features, for periodicity and other phenomena at the signal level. Here the parameter q serves as the level of smoothing, and as it gets lager, the emphasis in what $W_q(\mathsf{dgm}, \mathsf{dgm}_\Delta)$ measures shifts from topological noise and fine attributes to large topological events.

The ring of algebraic functions on persistence diagrams, as a source of features for machine learning purposes, has been recently studied by Adcock et. al. [1]. We believe these and other signatures such as $W_q(\mathsf{dgm}, \mathsf{dgm}_\Delta)$ should uncover non-trivial signal properties captured by their sliding window point clouds. We have devoted this paper to exploring the use of $mp(\mathsf{dgm})$, but we hope that in future work the list of useful features from persistence diagrams on sliding window point clouds can be extended.

We also mention that Section 6.3 is the first explicit computation of the persistence diagram of a parametrized space. The method of Fourier Approximation presented here is one of the first in a much needed toolbox for explicit computations of persistence diagrams.

Our final comment is to point out that the fact that the size of the sliding window should match the period searched for was not obvious in advance. Knowing this provides powerful information on sampling density to scientists planning an experiment that looks for periodic data, and lays the ground work for the use of SW1PerS as a *filter* for time series data.

In future work we plan to establish our conjecture that $mp(\mathsf{dgm})$ is maximized by our choice of window size, and the main ingredient will be strengthening the lower bound presented in Theorem 6.8. We also plan to establish the filtering properties

of SW1PerS and apply it to a variety of data, including biological data like that from gene expression and physiology, astronomical data, and weather. Finally, we plan to extend these methods by using other tools from Topological Data Analysis to find structure and features in time series.

References

[1] A. Adcock, E. Carlsson and G. Carlsson, *The Ring of Algebraic Functions on Persistence Bar Codes*, Preprint available at http://comptop.stanford.edu/u/preprints/multitwo.pdf, 2012.

[2] A. J. Blumberg, I. Gal, M. A. Mandell and M. Pancia, *Persistent homology for metric measure spaces, and robust statistics for hypothesis testing and confidence intervals*, arXiv preprint arXiv:1206.4581, 2012.

[3] G. Carlsson, *Topology and Data*, Bulletin of the American Mathematical Society, vol 46(2), p.p. 255–308, 2009.

[4] G. Carlsson, T. Ishkhanov, V. de Silva and A. Zomorodian, *On the local behavior of spaces of natural images*, International Journal of Computer Vision, vol 7(1), p.p. 1-12, 2008.

[5] F. Chazal, D. Cohen-Steiner, M. Glisse, L. J. Guibas, and S. Y. Oudot, *Proximity of persistence modules and their diagrams*, In *SCG*, p.p. 237-246, 2009.

[6] D. Cohen-Steiner, H. Edelsbrunner and J. Harer, *Stability of persistence diagrams*, Discrete and Computational Geometry, vol 37(1), p.p. 103-120, 2007.

[7] D. Cohen-Steiner, H. Edelsbrunner, J. Harer and Y. Mileyko, *Lipschitz Functions Have L^p-Stable Persistence*, Foundations of Computational Mathematics, vol 10(2), p.p. 127–139, 2010.

[8] D. Comaniciu and P. Meer, *Mean shift: A robust approach toward feature space analysis*, Pattern Analysis and Machine Intelligence, vol 24(5), p.p. 603–619, 2002.

[9] V. de Silva, D. Morozov and M. Vejdemo-Johansson, *Persistent Cohomology and Circular Coordinates*, Discrete & Computational Geometry, vol 45(4), p.p. 737–759, 2011.

[10] V. de Silva, P. Skraba and M. Vejdemo-Johansson, *Topological Analysis of Recurrent Systems*, Workshop on Algebraic Topology and Machine Learning, NIPS 2012, Preprint available at http://sites.google.com/site/nips2012topology/contributed-talks.

[11] A. Deckard, R. Analfi, D. Orlando, J. Hogenesch, S. Haase and J. Harer, *Design and Analysis of Large-Scale Biological Rhythm Studies: A Comparison of Algorithms for Detecting Periodic Signals in Biological Data*, Bioinformatics, btt541v1-btt541, 2013.

[12] H. Edelsbrunner and J. Harer **Computational Topology, an Introduction**, American Mathematical Society, (2010) (241 pages).

[13] E. F. Glynn, J. Chen and A. Mushegian, *Detecting periodic patterns in unevenly spaced gene expression time series using Lomb–Scargle periodograms*, Bioinformatics, vol 22(3), p.p. 310–316, 2006.

[14] A. Hatcher **Algebraic Topology.** Cambridge Univ. Press, England, 2002.

[15] M. E. Hughes, J. B. Hogenesch and K. Kornacker, *JTK_CYCLE: An Efficient Nonparametric Algorithm for Detecting Rhythmic Components in Genome-Scale Data Sets*, Journal of Biological Rhythms, vol 25(372), p.p. 372–380, 2010.

[16] H. Kantz and T. Schreiber, **Nonlinear Time Series Analysis**, Cambridge University Press, 2003.

[17] H. S. Kim, R. Eykholt and J. D. Salas, *Nonlinear dynaimcs, delay times, and embedding windows*, Physica D: Nonlinear Phenomena, vol 127(1), p.p. 48-60, 1999.

[18] N. R. Lomb, *Least-squares frequency analysis of unequally spaced data*, Astrophysics and Space Science, vol 39, p.p. 447–462, 1976.

[19] Y. Mileyko, S. Mukherjee, and J. Harer, *Probability measures on the space of persistence diagrams*, Inverse Problems, 27(12), p.p. 124007, 2011.

[20] K. Mischaikow and V. Nanda, *Morse Theory for Filtrations and Efficient Computation of Persistent Homology*, To appear on Discrete and Computational Geometry, 2013.

[21] J. R. Munkres **Elements of Algebraic Topology.** Addison-Wesley, Redwood City, California, 1984.

[22] M. A. Pinsky, **Introduction to Fourier Anlysis and Wavelets**, The Brooks/Cole Series in Advanced Mathematics, USA, 2003.

[23] J. A. Perea, A. Deckard, S. B. Haase and J. Harer, *SW1PerS: Sliding Windows and 1-Persistence Scoring; Discovering Periodicity in Gene Expression Time Series Data*, preprint (2013).

[24] J. D. Scargle, *Studies in astronomical time series analysis. II-Statistical aspects of spectral analysis of unevenly spaced data*, Astrophysical Journal, vol 263, p.p. 835–853, 1982.

[25] C. J. Stam, *Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field*, Clinical Neurophisiology , 116, p.p. 2266-2301, 2005.

[26] N. Hundewale, *The application of methods of nonlinear dynamics for ECG in Normal Sinus Rythm*, Int. J. of Computer Science, 9, p.p. 458-467, 2012.

[27] C. E. Shannon, *Communication in the presence of noise*, Proc. Inst. Radio Eng., vol. 37, no. 1, pp. 10-21, 1949.

[28] F. Takens, *Detecting strange attractors in turbulence*. in D. A. Rand and L. -S. Young. **Dynamical Systems and Turbulence**, Lecture Notes in Mathematics, vol. 898. Springer-Verlag. pp. 366–381.

[29] A. Tausz, M. Vejdemo-Johansson and H. Adams, *JavaPlex: A research software package for persistent (co)homology*, 2011, Software available at http://code.google.com/p/javaplex.

[30] A. Zomorodian and G. Carlsson, *Computing Persistent Homology*, Discrete & Computational Geometry, vol 33(2), p.p. 249–274, 2005.

Department of Mathematics, Duke University, Durham, North Carolina, USA.
E-mail address: joperea@math.duke.edu

Departments of Mathematics, Computer Science and Electrical and Computer Engineering, Duke University, Durham, North Carolina, USA.
E-mail address: john.harer@duke.edu