

A Stable Measure for Conditional Periodicity of Time Series using Persistent Homology

Abstract

We study how the topology of a pair of time series can be used to quantify how similar their periodicities are. Among the known methods to compare a pair of time series, the widely used measure of percent determinism (%DET) requires multiple input parameters. We experimentally show that its many input parameters reduce the stability of %DET. Persistent homology has been utilized to construct a scoring function with theoretical guarantees of stability that quantifies the periodicity of a single univariate time series f_1 , termed $\text{score}(f_1)$. Building on this concept, we propose a conditional periodicity score to quantify the periodicity similarity of a time series f_1 given another more periodic series f_2 , denoted $\text{score}(f_1|f_2)$, and derive theoretical results on its stability. This measure uniquely requires only one input parameter. For large embedding dimensions, pairwise distances between points may start to concentrate. With this setting in mind, we prove a new stability result for $\text{score}(f_1|f_2)$ under principal component analysis (PCA) when we use the projections of the time series embeddings onto their respective first K principal components. We bound the change in our score by a function of the eigenvalues corresponding to the remaining (unused) $N - K$ principal components, which is small when the first K principal components capture most of the variation in the time series embeddings. We derive a lower bound on the minimum embedding dimension to use in our pipeline which guarantees that any two such higher-valued embeddings give scores that are linearly within ϵ of each other. We use this lower bound to show that small changes in ϵ induce small changes in the resulting minimum embedding dimensions. We experimentally compare our similarity measure to %DET, and show its greater stability under small changes in periodicity and time lag.

Keywords: time series, conditional periodicity, persistent homology, PCA

1 Introduction

A continuous univariate time series $f : T \rightarrow \mathbb{R}$ defined on the real-valued interval T is a collection of points $\{(t, f(t))\}$ that depend on the input measure of time, $t \in T$. Time series analysis is employed in numerous applications and measuring the similarity between pairs of univariate time series is a well-studied problem. The time shift between a pair of time series can be estimated via the cross-correlation coefficients for a range of self-selected lags. Such coefficients have been used to measure the intensity of earthquakes and identify common significant periods between nucleic signals [?]. The correlation between the power spectra of two time series at a given frequency can be measured using coherence, which has been used to estimate the correlation between non-stationary EEG and EMG signals [1], to detect short significant coherence between non-stationary neural signals [2], and to identify the correlation between green investment and environmental sustainability in China [3]. A non-Euclidean distance between a pair of time series can be measured using dynamic time warping (DTW), a method that does not require a one-to-one correspondence when computing pairwise distances. DTW is commonly used for time-series clustering as it is shown to perform better compared to standard Euclidean distance. DTW has specifically been used to cluster suicidal symptoms signals [4], manic and depressive symptoms signals [5], and marine trafficking signals [6]. The correlation between the point-cloud representations of two time series can be estimated using their pairwise Euclidean distances. Denoted cross-recurrence quantification analysis, this method has been used to quantify the structure of utterance signals between children and their parents [7] and to identify different functional movement levels between patients with and without ACL surgery via EEG and EMG signals [8].

1.1 Limitations of Cross-Recurrence

Among the measures previously mentioned, arguably the most popular one is cross-recurrence. For instance, the widely used measure of percent determinism (%DET) is computed using cross-recurrence, but uses the most self-selected parameters. Cross-recurrence is an asymmetric binary measure that indicates whether or not the i -th and j -th states of a pair of time series embeddings are close in Euclidean distance, i.e., cross-recur. See an example of its asymmetry in Figure 1. The collection of all binary cross-recurrence values is summarized in a cross-recurrence matrix, \mathcal{C} . %DET is the proportion of cross-recurring states that belong to the diagonal strips of ones in \mathcal{C} , and essentially measures how correlated the embeddings of two series are. %DET requires four self-selected input parameters for its use: the time lag τ and embedding dimension M used to construct the embeddings of both series, the distance threshold `tol` used to determine whether or not a pair of states cross-recur, and the minimum number of points `minDL` in \mathcal{C} to form a diagonal.

One can also think of %DET as a measure of periodicity similarity between f_1 and f_2 , i.e., how close their cycle lengths are. If a w_1 -periodic time series f_1 and a w_2 -periodic time series f_2 have close periodicities, (i.e., $|(T/w_1) - (T/w_2)|$ is small), then their oscillations sync up more closely. In this case, fixing τ and M should produce time series embeddings that are more correlated (i.e., pairwise distances between the

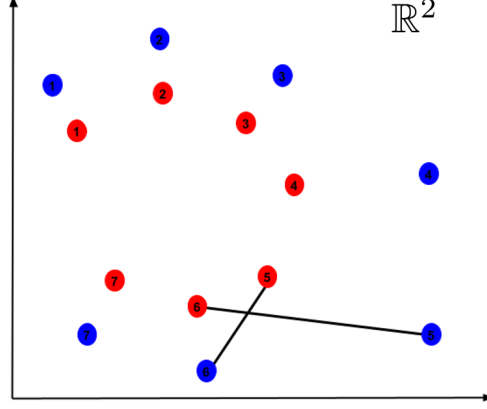


Fig. 1 An example of the asymmetry of cross-recurrence where $\phi_{f_1}(t_i)$ (colored in blue) and $\phi_{f_2}(t_i)$ (colored in red) denote the i -th states of the respective embeddings of time series f_1 and f_2 (motivated by Figure 5 in the work of Marwan et al. [9]). Here, since the pairwise Euclidean distances between states $i = 5$ and $j = 6$ are not the same, the corresponding entries \mathcal{C}_{ij} and \mathcal{C}_{ji} in the cross-recurrence matrix are not necessarily equal. For instance, if we define the distance threshold $\|\phi_{f_1}(t_5) - \phi_{f_2}(t_6)\|_2 < \text{tol} < \|\phi_{f_1}(t_6) - \phi_{f_2}(t_5)\|_2$, then $\mathcal{C}_{56} = 1$ and $\mathcal{C}_{65} = 0$.

same states should be smaller). For fixed `minDL`, one would further expect %DET to be higher for closer periodicities.

Methods have been introduced for computing optimal parameters for %DET. These include using the auto-mutual information (AMI) of a single time series and Cao’s false nearest neighbors (FNN) test to determine the time lag and embedding dimension, and defining τ and M to be either the average or maximum of both lags and dimensions of f_1 and f_2 . As well, setting the distance threshold to be greater than five times the standard deviation of Gaussian noise of the input series is a known method to ensure robustness to noise [9].

A key limitation of computing %DET is its requirement of multiple parametric inputs in order to estimate periodicity similarity. Intuitively, less similar periodicities are thought to produce less correlated point cloud representations. For such time series pairs, small changes in time lag may cause embeddings to be more susceptible to change. Since %DET uses pairwise distances between points in the embeddings of the time series, it may be sensitive to small time lag changes for time series pairs that have periodicities with enough dissimilarity. We show an example in Figure 2 where we fix `minDL` and `tol`, and assess the stability of %DET in detecting periodicity dissimilarity between a fixed pair of time series. To ensure a fair comparison to $\text{score}(f_1|f_2)$, we compute %DET on the top two principal components (PCs) of both time-series embeddings, and compare the results when we vary the time lag from $\tau = 3$ to $\tau = 4$ (see Sections 3.1 and 4.4 for more details). Notice that this small change in time lag produces embeddings (middle column) with similar topology, but whose pairwise distances change substantially. Hence, the resulting cross-recurrence matrices and %DET computed (right column) differ significantly. While %DET detects periodicity dissimilarity in the first case, increasing the time lag by just one substantially changes its performance. We are thus motivated to construct a topologically-computed measure

that quantifies similarity via a single point cloud whose time lag is fixed so that it uniquely captures the relationship between both time-series periodicities. This fixed time lag, coupled with the topology of a single representative embedding, removes the possibility of large changes in similarity measure resulting from small changes in τ .

1.2 Our Contributions

Persistent homology frameworks have proven theoretical stability properties [10, 11]. Due to the instability of %DET in detecting periodicity similarity of time series and its many input parameters required, we are motivated to construct a new periodicity similarity measure that uses persistent homology for guaranteed theoretical stability and requires only one self-selected input parameter. For a pair of time series $f_1, f_2 : T \subset \mathbb{R} \rightarrow \mathbb{R}$ for which f_2 is more periodic (i.e., has shorter cycle-length), we define a new periodicity similarity measure $\text{score}(f_1|f_2)$ (Definition 7). This measure uses the persistent homology of a single point cloud representation (Definition 5) that uniquely captures the relationship between the periodicities of f_1 and f_2 to quantify how similar their periodicities (i.e., cycle-lengths) are. This point cloud can be thought of as the representation of f_1 , but defined via a fixed time lag that is proportional to the cycle length of f_2 . We term this measure the *conditional periodicity score of f_1 given f_2* . See Figure 3 for an illustration of the main difference between how %DET and our score are computed.

The main benefit of our conditional periodicity score as opposed to %DET is its guaranteed theoretical stability under small changes in periodicity (Theorem 9) and Gaussian noise (Lemma 10) as well as requiring only one input parameter, the embedding dimension $M \in \mathbb{N}$. Furthermore, in the context of time series analysis under dimension reduction, we show that our score satisfies a stability result even when one uses truncated versions of the time series embeddings as computed by principal component analysis (PCA) (Theorem 11 and Corollary 12). We derive a lower bound on the embedding dimension used that allows us to control the precision of the conditional periodicity score (Theorem 13), as well as on the number of embedding points to use to capture the difference in periodicities of f_1 and f_2 (Proposition 15). Finally, we show in Corollary 14 that small changes in desired precision induce small changes in the minimum embedding dimension derived in Theorem 13.

Additionally, we highlight the relationship between our stability and convergence results (Section 3.3). The main difference we discuss is that although our periodicity and noise stability results (Theorem 9 and Proposition 10) show that small increases in periodicity of f_2 or Gaussian noise of f_1 induce small changes in conditional periodicity score, this does not guarantee that the scores being produced are “close” to the underlying conditional periodicity score. This closeness is guaranteed by our convergence result for a minimum choice of embedding dimension (Theorem 13). We demonstrate an example of this in Figure 4.

We present an algorithm to quantify the conditional periodicity of two input time series using PCA (Algorithm 1). This algorithm runs in $O(P \log P + NK^2 + N^2)$ time where P is the number of points in the two discrete univariate input signals, N is the number of points in the conditional sliding window embedding (SWE) of

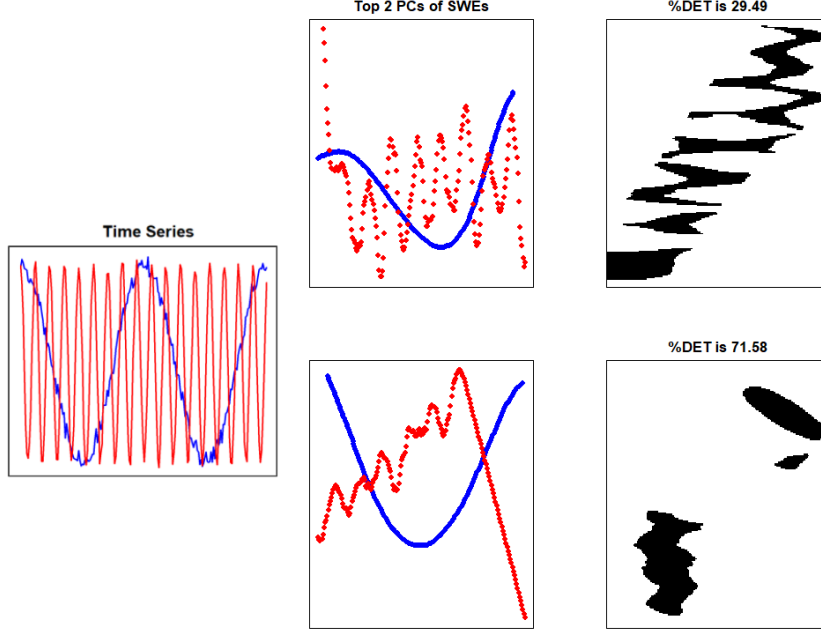


Fig. 2 Instability of %DET due to small changes in parameter τ . The third column shows %DET computed for $M = 16$, $\text{minDL} = 15$, and $\text{tol} = 0.9$, where we change the time lag from $\tau = 3$ (top row) to $\tau = 4$ (bottom row). The time series (first column) are defined with 5% Gaussian noise, where f_1 is 2-periodic (blue) and f_2 is 17-periodic (red). The middle column shows the first two principal components of the SWEs of the time series.

the fitted continuous signals, and K is the number of principal components used ($K \leq M + 1$ for embedding dimension M) (see Remark 16). We present computational evidence that shows our scoring function is robust to input signals of different types including sinusoidals, dampened sinusoidals, triangle-wave series, and square-waves with moderate amounts of Gaussian noise and dampening (Figure 6). In addition, our conditional periodicity score, when compared to %DET, is shown experimentally to maintain greater stability when subject to small changes in periodicity (i.e., cycle length of f_2) and time lag (Figure 7).

1.3 Related Work

As previously mentioned, several similarity measures have been used that indirectly quantify the closeness of periodicities, but we have yet to find any known theoretical stability results for these. These include cross-correlation, coherence, cross-recurrence, and DTW. Persistent homology, on the other hand, has been utilized to quantify the periodicity of a single univariate time series with theoretical stability guarantees [12–15]. Inspired by the Takens embedding [16], one selects an embedding dimension $M \in \mathbb{N}$ and a time lag $\tau > 0$ to map each time-series point $(t, f(t))$ from a univariate

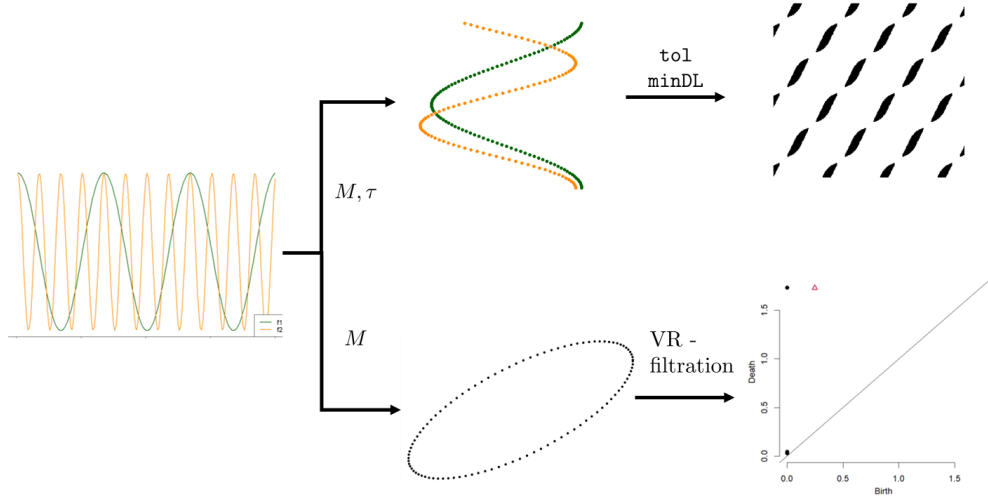


Fig. 3 Computation of %DET (top) compared with that of $\text{score}(f_1|f_2)$ (bottom). The score is computed using the persistence diagram (bottom right) obtained from the VR-filtration of a single point cloud representation (bottom middle) that captures the relationship between the periodicities of both input series (left). This measure requires only one input parameter, the embedding dimension M . %DET is computed using pairwise Euclidean distances between points in the point cloud representations of both input series (top middle). This measure requires four input parameters, the embedding dimension and time lag (M and τ) for both time series embeddings, the distance threshold and a minimum number of points to form a diagonal (tol and minDL) for the cross-recurrence matrix computation (top right).

series to an $(M + 1)$ -dimensional vector via the map

$$\text{SW}_{M,\tau} f(t) = \left(f(t), f(t + \tau), \dots, f(t + M\tau) \right)^T.$$

In other words, one maps windows of size $M\tau$ of the input signal to one vector in a point cloud in \mathbb{R}^{M+1} . Further, if the series is periodic on $[0, 2\pi]$, the resulting point cloud will be an elliptic curve that is roundest when the sliding window size is proportional to the underlying periodicity of the time series. One then performs Vietoris-Rips (VR) filtration on the sliding windows embedding, $\text{SW}_{M,\tau} f(t)$. The more periodic the input signal is, the more rounded the sliding windows embedding is. This yields a higher lifetime of the longest-surviving hole, called the maximum 1D-persistence. Dividing this maximum lifetime by the square root of three (assuming the sliding windows embedding (SWE) is centered and normalized) yields a periodicity score between 0 and 1, denoted $\text{score}(f)$, where $\text{score}(f)$ is closer to 1 when f is more periodic. This periodicity score and its aforementioned properties were introduced by Perea and Harer [12], and further used to quantify the periodicity of gene-expressions data by Perea, Deckard, Haase, and Harer [13].

The authors derive the normalization factor $1/\sqrt{3}$ for their scoring function in their proof of a lower bound on 1D max persistence [Theorem 6.8, [12]]. That is, they first define a finite collection of discrete time points $T = \{t_1, \dots, t_{J-1}, t_J\}$ for a time

series $f : T \rightarrow \mathbb{R}$, and the pointwise centered and normalized SWE of its N -truncated Fourier series $S_N f$ at time $t \in T$, a vector in \mathbb{R}^{M+1} denoted by $\psi_\tau(t)$. They further define $\kappa_N = \frac{2\sqrt{2}\|S_N f'(t)\|_2}{\|S_N(f - \hat{f}(0))\|_2}$. They then show that if the Hausdorff distance between T and the set of all input times corresponding to the Torus, $\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z})$, is strictly less than δ for $0 < \delta < \frac{\sqrt{3}\tilde{r}_n}{\kappa_N}$ (i.e., $|t_j - t_{j-1}| < 2\delta$), then the Euclidean distance between any consecutive pair of vectors $\psi_\tau(t_{j-1})$ and $\psi_\tau(t_j)$ in $\phi_\tau(T)$ is at most $\delta\kappa_N < \max_{1 \leq n \leq N} \sqrt{3}\tilde{r}_n \leq \sqrt{3}$, where \tilde{r}_n the n -th (normalized) Fourier coefficient. Since the VR-complex of $\phi_\tau(T)$ is determined by such pairwise Euclidean distances, they fix the maximum filtration scale to $\epsilon = \sqrt{3}$. The stability of this scoring function is proven by the authors [12] using the well-known stability result for persistent homology [10, 11]:

$$d_B\left(\text{dgm}(\text{VR}(X)), \text{dgm}(\text{VR}(Y))\right) \leq 2 d_{\text{GH}}(X, Y) \leq 2 d_H(X, Y) \quad (1)$$

where the finite data sets of points X and Y lay in a common metric space. Here, $\text{dgm}(\text{VR}(X))$ and $\text{dgm}(\text{VR}(Y))$ denote the persistence diagrams obtained from the VR filtrations on the point clouds X and Y , respectively, $d_B\left(\text{dgm}(\text{VR}(X)), \text{dgm}(\text{VR}(Y))\right)$ defines the bottleneck distance between persistence diagrams $\text{dgm}(\text{VR}(X))$ and $\text{dgm}(\text{VR}(Y))$, $d_{\text{GH}}(X, Y)$ denotes the Gromov-Hausdorff distance between X and Y , and $d_H(X, Y)$ denotes the Hausdorff distance between X and Y . Since we use the VR filtration by default in this work, we will write $\text{dgm}(X)$ in short to denote $\text{dgm}(\text{VR}(X))$ when there is no cause for confusion.

2 Definitions

Here we introduce standard definitions of distances used in this paper. See, for instance, the book by Burago, Bugaro, and Ivanov [17] for details.

Definition 1 (Hausdorff Distance) *Given two sets of points X and Y in a common metric space, the Hausdorff distance between them is given by*

$$d_H(X, Y) = \inf \{\epsilon > 0 : X \subseteq Y^\epsilon, Y \subseteq X^\epsilon\}, \quad \text{where } X^\epsilon = \bigcup_{x \in X} B_\epsilon(x) \text{ and } Y^\epsilon = \bigcup_{y \in Y} B_\epsilon(y)$$

denote the union of all ϵ -balls centered at each point in either set.

Definition 2 (Hausdorff Definition of Gromov-Hausdorff Distance) *Given two sets of points X and Y , the Gromov-Hausdorff distance between them is given by*

$$d_{\text{GH}}(X, Y) = \inf \{d_H(f(X), g(Y)) : f : X \rightarrow S, g : Y \rightarrow S\},$$

where f and g are isometric embeddings of X and Y into a common metric space S . If X and Y lay in a shared metric space S , then $d_{\text{GH}}(X, Y) \leq d_H(X, Y)$ [18].

Definition 3 (Distortion Definition of Gromov-Hausdorff Distance) *An alternative definition of the Gromov-Hausdorff distance between two sets of points X and Y is given by*

$$d_{\text{GH}}(X, Y) = \frac{1}{2} \inf \{ \text{dis}(R) : R : X \rightarrow Y \in \mathcal{R}(X, Y) \},$$

where R is a relation between X and Y whose distortion is defined by

$$\text{dis}(R) = \sup \{ |d_X(x, x') - d_Y(y, y')| : (x, y), (x', y') \in R \},$$

where d_X and d_Y are the corresponding metrics for X and Y , respectively.

Definition 4 (Bottleneck Distance) *Given two finite sets of points X and Y , let $\text{dgm}(X)$ and $\text{dgm}(Y)$ denote the persistence diagrams of a chosen dimension obtained from the Vietoris-Rips (VR) filtration on X and Y , respectively. Then the Bottleneck distance between $\text{dgm}(X)$ and $\text{dgm}(Y)$ is given by*

$$d_{\text{B}}(\text{dgm}(X), \text{dgm}(Y)) = \inf_{\phi} \sup_x \|x - \phi(x)\|_{\infty},$$

where $\phi : \text{dgm}(X) \rightarrow \text{dgm}(Y)$ denotes a bijection between $\text{dgm}(X)$ and $\text{dgm}(Y)$, including points along the diagonal in either diagram when they both do not share the same cardinality.

2.1 The Conditional Periodicity Score

We now present the definitions of our measure for the conditional periodicity of two univariate time series.

Definition 5 (Conditional Sliding Windows Embedding) *Let $f_1, f_2 : [0, 2\pi] \rightarrow \mathbb{R}$ be two continuous, periodic, univariate time series with cycle-lengths $\frac{2\pi}{w_1}$ and $\frac{2\pi}{w_2}$, respectively. Assume $w_1, w_2 \in \mathbb{N}$, and that $\frac{2\pi}{w_2} \leq \frac{2\pi}{w_1}$. Then the conditional sliding windows embedding (SWE) of f_1 given f_2 is defined by*

$$\text{SW}_{M,\tau} f_{1|2}(t) = \left(f_1(t), f_1(t + \tau), \dots, f_1(t + M\tau) \right)^T$$

where $M \in \mathbb{N}$ is a selected embedding dimension and the time lag $\tau = \frac{2\pi}{w_2(M+1)}$ is proportional to the length of one cycle of f_2 .

Definition 6 (Periodic Functions) *A continuous time series $f : [0, 2\pi] \rightarrow \mathbb{R}$ is L -periodic if $f(t + \frac{2\pi}{L}) = f(t)$, $L \in \mathbb{N}$ [12]. Throughout this paper, we denote the cycle length of a time series as its periodicity. We further denote one time series to be more periodic than another if its cycle length is smaller.*

Definition 7 (Conditional Periodicity Score) *Let $\text{mp}(\text{dgm}_1(\text{SW}_{M,\tau} f_{1|2}(T)))$ denote the lifetime of the longest surviving one-dimensional homology feature (i.e., hole) in the VR filtration on the conditional SWE of f_1 given f_2 . Then the conditional periodicity score of f_1 given f_2 is defined as*

$$\text{score}(f_1|f_2) = \frac{\text{mp}(\text{dgm}_1(\text{SW}_{M,\tau} f_{1|2}(T)))}{\sqrt{3}}$$

for $T = \left[0, \frac{2\pi}{w_1} \right]$.

3 Stability of the Conditional Periodicity Score

To obtain $\text{score}(f_1|f_2)$, we first compute the conditional SWE of f_1 given a more periodic series f_2 (Definition 5), and then compute the VR filtration of this embedding to obtain the conditional periodicity score (Definition 7). Assuming that f_2 is more periodic than f_1 , we ultimately deduce that small changes in the periodicity of f_2 yield small changes in the conditional periodicity score. We assume that f_1 and f_2 are continuous series defined on $[0, 2\pi]$ where f_2 is more periodic than f_1 , i.e., $\frac{2\pi}{w_2} \leq \frac{2\pi}{w_1}$ for $w_1, w_2 \in \mathbb{N}$. We also assume that any conditional SWE contains $N \in \mathbb{N}$ points. We first observe that as the periodicity of f_2 approaches that of f_1 , the conditional periodicity score reduces to the periodicity score of f_1 .

Proposition 8 (Reduction to Periodicity Score)

$$\lim_{\frac{2\pi}{w_2} \rightarrow \frac{2\pi}{w_1}^-} \text{SW}_{M,\tau} f_{1|2}(t) = \text{SW}_{M,\tau} f_1(t) \text{ for } t \in \left[0, \frac{2\pi}{w_1}\right].$$

Proof Let $i \in \{1, 2, \dots, M+1\}$ for fixed $M \in \mathbb{N}$. Since f_1 is continuous and differentiable on $[0, 2\pi]$ and hence on $(0, 2\pi)$, it is continuous and differentiable on any subset $[a, b] \subset [0, 2\pi]$ and $(a, b) \subset (0, 2\pi)$. Let $t \in [0, 2\pi]$. Consider the subinterval $I_i = \left[t + \frac{2(i-1)\pi}{w_2(M+1)}, t + \frac{2(i-1)\pi}{w_1(M+1)}\right]$ of $[0, 2\pi]$. Then f_1 is also continuous and differentiable on I_i . By the Mean Value Theorem (MVT), there exists some $c_i \in I_i$ such that

$$f'_1(c_i) \left[\left(t + \frac{2(i-1)\pi}{w_1(M+1)}\right) - \left(t + \frac{2(i-1)\pi}{w_2(M+1)}\right) \right] = f_1\left(t + \frac{2(i-1)\pi}{w_1(M+1)}\right) - f_1\left(t + \frac{2(i-1)\pi}{w_2(M+1)}\right).$$

Then we get the following equality:

$$\begin{aligned} & \left\| \text{SW}_{M,\tau} f_{1|2}(t) - \text{SW}_{M,\tau} f_1(t) \right\|_2 \\ &= \left(\sum_{i=1}^{M+1} \left| f_1\left(t + \frac{2(i-1)\pi}{w_2(M+1)}\right) - f_1\left(t + \frac{2(i-1)\pi}{w_1(M+1)}\right) \right|^2 \right)^{\frac{1}{2}} \\ &= \left(\sum_{i=1}^{M+1} |f'_1(c_i)|^2 \left| \left(t + \frac{2(i-1)\pi}{w_2(M+1)}\right) - \left(t + \frac{2(i-1)\pi}{w_1(M+1)}\right) \right|^2 \right)^{\frac{1}{2}} \\ &= \left(\sum_{i=2}^{M+1} |f'_1(c_i)|^2 \left| \frac{i-1}{M+1} \right|^2 \left| \frac{2\pi}{w_1} - \frac{2\pi}{w_2} \right|^2 \right)^{\frac{1}{2}}. \quad \square \end{aligned}$$

A direct consequence of Proposition 8 is that $\text{score}(f_1|f_2) \rightarrow \text{score}(f_1)$ as $\frac{2\pi}{w_2} \rightarrow \frac{2\pi}{w_1}^-$.

Theorem 9 (Stability of Conditional Periodicity Score) *Let $f_1, f_{21}, f_{22} : [0, 2\pi] \rightarrow \mathbb{R}$ be three continuous univariate time series such that $\frac{2\pi}{w_{22}} < \frac{2\pi}{w_{21}} \leq \frac{2\pi}{w_1}$ and $w_1, w_{21}, w_{22} \in \mathbb{N}$. Define the conditional SWE of f_1 given f_{21} as X_1 and the conditional SWE of f_1 given f_{22} as X_2 , where the sliding window sizes are defined using time lags $\tau_1 = \frac{2\pi}{w_{21}(M+1)}$ and*

$\tau_2 = \frac{2\pi}{w_{22}(M+1)}$, respectively. Similarly, define the 1D persistence diagrams from the VR filtrations on X_1 and X_2 as $\text{dgm}_1(X_1)$ and $\text{dgm}_1(X_2)$, respectively. Let the max 1D persistence in each diagram be denoted by $\text{mp}(\text{dgm}_1(X_1))$ and $\text{mp}(\text{dgm}_1(X_2))$, and the resulting conditional periodicity scores of f_1 given f_{21} and f_1 given f_{22} be denoted by $\text{score}(f_1|f_{21})$ and $\text{score}(f_1|f_{22})$, respectively. Then the following results hold:

$$d_H(X_1, X_2) \leq \sqrt{M+1} \left| \frac{2\pi}{w_{21}} - \frac{2\pi}{w_{22}} \right| \sqrt{\sum_{i=2}^{M+1} |f'_1(c_i)|^2} \quad (2)$$

$$d_B(\text{dgm}_1(X_1), \text{dgm}_1(X_2)) \leq 2\sqrt{M+1} \left| \frac{2\pi}{w_{21}} - \frac{2\pi}{w_{22}} \right| \sqrt{\sum_{i=2}^{M+1} |f'_1(c_i)|^2} \quad (3)$$

$$|\text{mp}(\text{dgm}_1(X_1)) - \text{mp}(\text{dgm}_1(X_2))| \leq 4\sqrt{M+1} \left| \frac{2\pi}{w_{21}} - \frac{2\pi}{w_{22}} \right| \sqrt{\sum_{i=2}^{M+1} |f'_1(c_i)|^2} \quad (4)$$

$$|\text{score}(f_1|f_{21}) - \text{score}(f_1|f_{22})| \leq 4\sqrt{\frac{M+1}{3}} \left| \frac{2\pi}{w_{21}} - \frac{2\pi}{w_{22}} \right| \sqrt{\sum_{i=2}^{M+1} |f'_1(c_i)|^2} \quad (5)$$

for some $c_i \in (t + i\tau_2, t + i\tau_1)$ and $i = 1, \dots, M+1$.

Proof Bound on Hausdorff distance in Equation (2):

We first find an upper bound on the Euclidean distance between respective pairs of points in X_1 and X_2 , and then use it to find an upper bound on the Hausdorff distance between the two point clouds. Similar to the proof of Proposition 8, there exists some $c_i \in (t + (i-1)\tau_2, t + (i-1)\tau_1)$ where

$$\begin{aligned} \left\| \text{SW}_{M,\tau} f_{1|21}(t) - \text{SW}_{M,\tau} f_{1|22}(t) \right\|_2 &= \left(\sum_{i=2}^{M+1} |f'_1(c_i)|^2 \left| \frac{i-1}{M+1} \right|^2 \left| \frac{2\pi}{w_{21}} - \frac{2\pi}{w_{22}} \right|^2 \right)^{\frac{1}{2}} \\ &\leq \frac{1}{M+1} \left| \frac{2\pi}{w_{21}} - \frac{2\pi}{w_{22}} \right| \sqrt{\sum_{i=2}^{M+1} |f'_1(c_i)|^2} \sqrt{\sum_{i=2}^{M+1} |i-1|} \\ &\quad \text{(by Cauchy-Schwartz inequality)} \\ &< \sqrt{M+1} \left| \frac{2\pi}{w_{21}} - \frac{2\pi}{w_{22}} \right| \sqrt{\sum_{i=2}^{M+1} |f'_1(c_i)|^2}. \end{aligned}$$

Let $\epsilon > \sqrt{M+1} \left| \frac{2\pi}{w_{21}} - \frac{2\pi}{w_{22}} \right| \sqrt{\sum_{i=2}^{M+1} |f'_1(c_i)|^2}$. Then $X_1 \subset X_2^\epsilon$ and $X_2 \subset X_1^\epsilon$, so by Definition 1, $d_H(X_1, X_2) \leq \epsilon$. Taking the infimum of both sides yields the relation

$$d_H(X_1, X_2) \leq \sqrt{M+1} \left| \frac{2\pi}{w_{21}} - \frac{2\pi}{w_{22}} \right| \sqrt{\sum_{i=2}^{M+1} |f'_1(c_i)|^2}.$$

Bound on bottleneck distance in Equation (3):

By the bound on stability of persistence diagrams in Equation (1), we have that $d_B(\text{dgm}_1(X_1), \text{dgm}_1(X_2)) \leq 2 d_H(X_1, X_2)$, since X_1 and X_2 both lay in \mathbb{R}^{M+1} .

Bound on max persistence in Equation (4):

This proof is motivated by a similar result obtained for difference in the max persistence between embeddings of time series and their truncated Fourier approximations by Perea and Harer [12, Theorem 4.5 (2)]. Let (b_1^{\max}, d_1^{\max}) and (b_2^{\max}, d_2^{\max}) be the points corresponding to the 1D-features of max persistence in $\text{dgm}_1(X_1)$ and $\text{dgm}_1(X_2)$, respectively. Let $\epsilon^* = d_B(\text{dgm}_1(X_1), \text{dgm}_1(X_2))$. Then by Definition 4, any pairs of points (b_1, d_1) and (b_2, d_2) in the respective diagrams satisfy $\|(b_1, d_1) - (b_2, d_2)\|_\infty = \sup\{|b_1 - b_2|, |d_1 - d_2|\} \leq \epsilon^*$. Hence, $|b_1^{\max} - b_2^{\max}| \leq \epsilon^*$ and $|d_1^{\max} - d_2^{\max}| \leq \epsilon^*$. Therefore

$$|\text{mp}(\text{dgm}_1(X_1)) - \text{mp}(\text{dgm}_1(X_2))| \leq |d_1^{\max} - d_2^{\max}| + |b_1^{\max} - b_2^{\max}| \leq 2 d_B(\text{dgm}_1(X_1), \text{dgm}_1(X_2)).$$

Bound on the score in Equation (5): By Definition 7,

$$\begin{aligned} |\text{score}(f_1|f_{21}) - \text{score}(f_1|f_{22})| &= \frac{1}{\sqrt{3}} |\text{mp}(\text{dgm}_1(X_1)) - \text{mp}(\text{dgm}_1(X_2))| \\ &\leq \frac{2}{\sqrt{3}} d_B(\text{dgm}_1(X_1), \text{dgm}_1(X_2)) \\ &\leq \frac{4}{\sqrt{3}} d_H(X_1, X_2) \\ &\leq 4\sqrt{\frac{M+1}{3}} \left| \frac{2\pi}{w_{21}} - \frac{2\pi}{w_{22}} \right| \sqrt{\sum_{i=2}^{M+1} |f'_1(c_i)|^2}. \square \end{aligned}$$

We now derive lower bounds on the likelihood that our conditional periodicity score is stable under small amounts of Gaussian noise in f_1 and f_2 .

Lemma 10 (Stability of score under added Gaussian noise) *Let $f_1, f_2 : [0, 2\pi] \rightarrow \mathbb{R}$ be continuous, real-periodic time series. Let $\sigma > 0$ denote the standard deviation of added Gaussian noise to f_1 . Consider*

$$f_1^\sigma(t) = f_1(t) + \epsilon_t,$$

where $\epsilon_t \sim N(0, \sigma^2)$ is a randomly sampled scalar from the associated Gaussian distribution. Define the constant $a = \frac{\sigma}{\sqrt{\delta}}$ for some $\delta \in (0, 1)$. Denote the conditional periodicity score under this added noise at σ by $\text{score}_\sigma(f_1|f_2)$. Then it is at least $(1 - \delta) \cdot 100$ % likely that

$$|\text{score}(f_1|f_2) - \text{score}_\sigma(f_1|f_2)| \leq 4\sigma\sqrt{\frac{M+1}{3\delta}}.$$

Proof Define the time series f_1 at noise level σ by $f_1^\sigma(t) = f_1(t) + \epsilon_t$ for $\epsilon_t \sim N(0, \sigma^2)$. Then the conditional embedding at noise level σ is denoted

$$\begin{aligned} \text{SW}_{M,\tau}^\sigma f_{1|2}(t) &= \left(f_1^\sigma(t), f_1^\sigma(t + \tau), \dots, f_1^\sigma(t + M\tau) \right)^T \\ &= \left(f_1(t) + \epsilon_t, f_1(t + \tau) + \epsilon_{t+\tau}, \dots, f_1(t + M\tau) + \epsilon_{t+M\tau} \right)^T. \end{aligned}$$

By our assumption, we have that $\delta = \left(\frac{\sigma}{a}\right)^2$. By Chebyshev's inequality, the probability that any given draw ϵ_t deviates from the mean of zero by at least a is given by $P[|\epsilon_t| \geq a] \leq \delta$. Hence, the probability that ϵ_t is within a of the mean is given by $P[|\epsilon_t| \leq a] \geq 1 - \delta$. Put another way, it is at least $(1 - \delta) \cdot 100\%$ likely that $|\epsilon_t| \leq a$. Hence

$$\left\| \text{SW}_{M,\tau} f_{1|2}(t) - \text{SW}_{M,\tau}^\sigma f_{1|2}(t) \right\|_2 = \sqrt{\sum_{i=1}^{M+1} |\epsilon_{t+(i-1)\tau}|^2} \leq a\sqrt{M+1}$$

with at least $(1 - \delta) \cdot 100\%$ likelihood. Hence, by similar arguments as in the proof of Theorem 9, it is at least $(1 - \delta) \cdot 100\%$ likely that

$$|\text{score}(f_1|f_2) - \text{score}_\sigma(f_1|f_2)| \leq \frac{4a\sqrt{M+1}}{\sqrt{3}} = 4\sigma\sqrt{\frac{M+1}{3\delta}}. \quad \square$$

3.1 Stability of Score under PCA

The main bottleneck when using the periodicity score in practice is the computation of the dimension 1 persistence diagram of the Vietoris-Rips filtration of the SWE. In general, for a point cloud X with N points, the computation of $\text{dgm}_1(X)$ runs in $O(N^6)$ time, although faster approaches may be available in lower dimensions [19]. Moreover, in Theorem 13 below we prove that greater precision of $\text{score}(f_1|f_2)$ warrants larger embedding dimensions. For sufficiently large enough embedding dimensions (i.e. $M > 10$), pairwise Euclidean distances used to compute VR-complexes during filtration start to converge [20]. This could lead to inaccurate topological summaries when computing $\text{score}(f_1|f_2)$. As such, we study the conditional periodicity score under principal component analysis (PCA), a widely used dimension reduction technique [21].

Theorem 11 (Stability of Conditional Periodicity Score Under PCA) *Let $K \leq M + 1$ for $K \in \mathbb{N}$. Suppose f_2 is more periodic than f_1 on $[0, 2\pi]$ with cycle lengths $\frac{2\pi}{w_2} \leq \frac{2\pi}{w_1}$, respectively. For $T = \left[0, \frac{2\pi}{w_1}\right]$, define the orthogonal projection of $X = \text{SW}_{M,\tau} f_{1|2}(T)$ onto its top K principal components by $\phi(X) = (\langle \mathbf{c}_1, X \rangle, \dots, \langle \mathbf{c}_K, X \rangle)^T$ for orthonormal eigenvectors and corresponding eigenvalues $\{\mathbf{c}_k, \lambda_k\}_{k=1}^N$ produced by PCA. Suppose X contains $N \in \mathbb{N}$ points. Denote the conditional periodicity score under ϕ , $\text{score}_\phi(f_1|f_2)$, as the maximum 1-d persistence from the VR-filtration on $\phi(X)$ divided by $\sqrt{3}$. Then*

$$|\text{score}(f_1|f_2) - \text{score}_\phi(f_1|f_2)| \leq \sqrt{\frac{8}{3}} \sqrt[4]{\sum_{i=K+1}^N \lambda_i^2}.$$

Proof Our proof is inspired by methods used in [22]. Notice that ϕ is a relation in $\mathcal{R}(X, Y)$ for $Y = \phi(X)$. Then by Definition 3 we have that $d_{\text{GH}}(X, Y) \leq \frac{1}{2} \text{dis}(\phi)$, where $\text{dis}(\phi)^2 = \|D_X - D_Y\|_{\max}^2$ with D_X being the matrix of pairwise distances in X . We get that [22, Lemma 3.9]

$$\|D_X - D_Y\|_{\max}^2 \leq \left\| D_X^{\circ 2} - D_Y^{\circ 2} \right\|_{\max},$$

where $D_X^{\circ^2}$ denotes the matrix of squared pairwise Euclidean distances in X .

Recall that for a matrix A , $\|A\|_{\max}^2 = (\max_{ij} |A_{ij}|)^2 = \max_{ij} |A_{ij}|^2 \leq \sum_{ij} |A_{ij}|^2 = \|A\|_F^2$, where $\|\cdot\|_F$ denotes the Frobenius norm. Then we have

$$\left\| D_X^{\circ^2} - D_Y^{\circ^2} \right\|_{\max} \leq \left\| D_X^{\circ^2} - D_Y^{\circ^2} \right\|_F.$$

Define the eigendecomposition of the covariance of X as $XX^T = Q\Lambda Q^T$, where $Q = [\mathbf{c}_1, \dots, \mathbf{c}_N]$ is the $N \times N$ matrix of orthonormal eigenvectors and $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_N)$ is the $N \times N$ diagonal matrix of eigenvalues corresponding to the orthogonal projection of X . Then the eigendecomposition of the K -dimensional subspace containing the top K principal components of X can be defined by $YY^T = Q\Lambda|_K Q^T$, where $\Lambda|_K = \text{diag}(\lambda_1, \dots, \lambda_K, \underbrace{0, \dots, 0}_{N-K})$.

Then, if we center the squared distances in $D_X^{\circ^2}$ and $D_Y^{\circ^2}$, we obtain the relations [23]

$$XX^T = -\frac{1}{2}C_N D_X^{\circ^2} C_N \quad \text{and}$$

$$YY^T = -\frac{1}{2}C_N D_Y^{\circ^2} C_N,$$

where C_N is the $N \times N$ centering matrix with diagonal entries $1 - \frac{1}{N}$ and off-diagonal entries $-\frac{1}{N}$. Then $D_X^{\circ^2} = -2C_N(Q\Lambda Q^T)C_N$ and $D_Y^{\circ^2} = -2C_N(Q\Lambda|_K Q^T)C_N$. Since $Q^T Q = QQ^T = I$, we obtain that

$$\begin{aligned} \left\| D_X^{\circ^2} - D_Y^{\circ^2} \right\|_F &= 2 \left\| C_N(Q\Lambda Q^T - Q\Lambda|_K Q^T)C_N \right\|_F \\ &= 2 \left\| Q(\Lambda - \Lambda|_K)Q^T \right\|_F \quad (\text{since } C_N \approx I \text{ for sufficiently large } N) \\ &= 2 \sqrt{\text{tr}(Q(\Lambda - \Lambda|_K)^2 Q^T)} \\ &= 2 \sqrt{\text{tr}((\Lambda - \Lambda|_K)^2)} \quad (\text{since trace is cyclically invariant}) \\ &= 2 \sqrt{\sum_{i=K+1}^N \lambda_i^2}. \end{aligned}$$

Thus we get that $\text{dis}(\phi) \leq \sqrt{2} \sqrt[4]{\sum_{i=K+1}^N \lambda_i^2}$ and hence $d_{\text{GH}}(X, Y) \leq \frac{\sqrt{2}}{2} \sqrt[4]{\sum_{i=K+1}^N \lambda_i^2}$. Then again by the standard result on stability of persistence diagrams (Equation (1)), we get that

$$d_{\text{B}}(\text{dgm}_1(X), \text{dgm}_1(Y)) \leq 2 d_{\text{GH}}(X, Y) \leq \sqrt{2} \sqrt[4]{\sum_{i=K+1}^N \lambda_i^2}.$$

Hence by similar arguments to those in the proof of Theorem 9, we have that

$$\begin{aligned} |\text{score}(f_1|f_2) - \text{score}_{\phi}(f_1|f_2)| &= \frac{1}{\sqrt{3}} |\text{mp}(\text{dgm}_1(X)) - \text{mp}(\text{dgm}_1(\phi(X)))| \\ &\leq \frac{2}{\sqrt{3}} d_{\text{B}}(\text{dgm}_1(X), \text{dgm}_1(\phi(X))) \\ &\leq 2 \sqrt{\frac{2}{3}} \sqrt[4]{\sum_{i=K+1}^N \lambda_i^2} \\ &= \sqrt{\frac{8}{3}} \sqrt[4]{\sum_{i=K+1}^N \lambda_i^2}. \square \end{aligned}$$

Theorem 11 shows that if the top K principal components of the conditional SWE capture most of its structure, then the conditional periodicity score will not change much under the associated orthogonal projection. We now reveal a direct consequence of this stability under small changes in periodicity of f_2 .

Corollary 12 (Consequence of Stability of Conditional Periodicity Score Under PCA) *Let $K \leq M + 1$ for $K \in \mathbb{N}$. Suppose f_2 is more periodic than f_1 on $[0, 2\pi]$ with cycle lengths $\frac{2\pi}{w_2} \leq \frac{2\pi}{w_1}$, respectively. For $T = \left[0, \frac{2\pi}{w_1}\right]$, define the orthogonal projections of $X_1 = \text{SW}_{M, \tau_1} f_{1|2}(T)$ and $X_2 = \text{SW}_{M, \tau_2} f_{1|2}(T)$ onto their top K principal components by the relations $\phi_1(X_1) = (\langle \mathbf{c}_1, X_1 \rangle, \dots, \langle \mathbf{c}_K, X_1 \rangle)^T$ and $\phi_2(X_2) = (\langle \mathbf{d}_1, X_2 \rangle, \dots, \langle \mathbf{d}_K, X_2 \rangle)^T$ for orthonormal eigenvectors and corresponding eigenvalues $\{\mathbf{c}_k, \lambda_k\}_{k=1}^N$, $\{\mathbf{d}_k, \gamma_k\}_{k=1}^N$ produced by PCA. Suppose X_1 and X_2 each contain $N \in \mathbb{N}$ points. Define the conditional periodicity score of f_1 given f_{21} and f_1 given f_{22} under ϕ_1 and ϕ_2 , respectively, as $\text{score}_{\phi_1}(f_1|f_{21})$ and $\text{score}_{\phi_2}(f_1|f_{22})$. Then the following inequality holds:*

$$\begin{aligned} |\text{score}_{\phi_1}(f_1|f_{21}) - \text{score}_{\phi_2}(f_1|f_{22})| &\leq \sqrt{\frac{8}{3}} \left(\sqrt[4]{\sum_{i=K+1}^N \lambda_i^2} + \sqrt[4]{\sum_{i=K+1}^N \gamma_i^2} \right) \\ &\quad + |\text{score}(f_1|f_{21}) - \text{score}(f_1|f_{22})|. \end{aligned}$$

Proof By Theorem 11, we have that

$$\begin{aligned} |\text{score}_{\phi_1}(f_1|f_{21}) - \text{score}(f_1|f_{21})| &\leq \sqrt{\frac{8}{3}} \sqrt[4]{\sum_{i=K+1}^N \lambda_i^2} \quad \text{and} \\ |\text{score}_{\phi_2}(f_1|f_{22}) - \text{score}(f_1|f_{22})| &\leq \sqrt{\frac{8}{3}} \sqrt[4]{\sum_{i=K+1}^N \gamma_i^2}. \end{aligned}$$

Hence we get

$$\begin{aligned} &|\text{score}_{\phi_1}(f_1|f_{21}) - \text{score}_{\phi_2}(f_1|f_{22})| \\ &\leq |\text{score}_{\phi_1}(f_1|f_{21}) - \text{score}(f_1|f_{21})| + |\text{score}(f_1|f_{21}) - \text{score}(f_1|f_{22})| \\ &\quad + |\text{score}(f_1|f_{22}) - \text{score}_{\phi_2}(f_1|f_{22})| \\ &\leq \sqrt{\frac{8}{3}} \left(\sqrt[4]{\sum_{i=K+1}^N \lambda_i^2} + \sqrt[4]{\sum_{i=K+1}^N \gamma_i^2} \right) + |\text{score}(f_1|f_{21}) - \text{score}(f_1|f_{22})|. \square \end{aligned}$$

Corollary 12 reveals that small changes in periodicity of f_2 still yield small changes in the $\text{score}(f_1|f_2)$ under orthogonal projection if the top K principal components of the conditional SWE capture most of its structure.

3.2 A Minimum Embedding Dimension for Convergence

We define a minimum embedding dimension \mathcal{M} that we can use to control the convergence behavior of the conditional periodicity score at dimension $M \in \mathbb{N}$, denoted $\text{score}_M(f_1|f_2)$, near the limiting score, $\text{score}(f_1|f_2)$.

Theorem 13 (Minimum embedding dimension for convergence) *Let $\epsilon > 0$. Suppose f_2 is more periodic than f_1 on $T = [0, 2\pi]$ with cycle lengths $\frac{2\pi}{w_2} \leq \frac{2\pi}{w_1}$, respectively. Then, any embedding dimension $M_2 > M_1 \geq \mathcal{M} \in \mathbb{N}$ for $\mathcal{M} = \left\lceil \frac{2\pi}{w_2 \epsilon} \right\rceil$ guarantees that the change in respective conditional periodicity scores $\text{score}_{M_2}(f_1|f_2)$ and $\text{score}_{M_1}(f_1|f_2)$ are linearly bounded above by ϵ .*

Proof We first show that $\tau = \frac{2\pi}{w_2(M+1)}$ is a Cauchy sequence of M . By the definition of \mathcal{M} , $\mathcal{M} \geq \frac{2\pi}{w_2 \epsilon}$ and hence $\frac{2\pi}{w_2 \mathcal{M}} \leq \epsilon \implies \frac{2\pi}{w_2} \left(\frac{1}{\mathcal{M}} \right) \leq \epsilon$. Since $M_2 > M_1 \geq \mathcal{M}$, $M_2 + 1 > M_1 + 1 \geq \mathcal{M} + 1$ and hence $\frac{1}{M_2+1} < \frac{1}{M_1+1} \leq \frac{1}{\mathcal{M}+1} < \frac{1}{\mathcal{M}}$. Thus we get

$$\begin{aligned} |\tau(M_1) - \tau(M_2)| &= \frac{2\pi}{w_2} \left| \frac{1}{M_1+1} - \frac{1}{M_2+1} \right| \\ &= \frac{2\pi}{w_2} \left(\frac{1}{M_1+1} - \frac{1}{M_2+1} \right) \\ &\leq \frac{2\pi}{w_2} \left(\frac{1}{\mathcal{M}+1} - \frac{1}{M_2+1} \right) \\ &< \frac{2\pi}{w_2} \left(\frac{1}{\mathcal{M}} \right) \leq \epsilon. \end{aligned} \quad (6)$$

Define the zero-padded conditional SWE of f_1 given f_2 for the smaller embedding dimension M_1 by $\text{SW}_{M_1, \tau(M_1)} f_{1|2}(t) = \left(f_1(t), \dots, f_1(t + M_1 \tau(M_1)), \underbrace{0, \dots, 0}_{M_2 - M_1} \right)^T$. Then for

$i = 1, \dots, M_1 + 1$, by the MVT there exists some $c_i \in I_i = \left(t + (i-1)\tau(M_2), t + (i-1)\tau(M_1) \right)$ such that

$$|f'_1(c_i)| \left| (i-1)(\tau(M_1) - \tau(M_2)) \right| = \left| f_1(t + (i-1)\tau(M_1)) - f_1(t + (i-1)\tau(M_2)) \right|. \quad (7)$$

Applying the results in Equations (6) and (7), we obtain the following relation.

$$\begin{aligned} &\left\| \text{SW}_{M_1, \tau(M_1)} f_{1|2}(t) - \text{SW}_{M_2, \tau(M_2)} f_{1|2}(t) \right\|_2^2 \\ &= \sum_{i=1}^{M_1+1} \left| f_1(t + (i-1)\tau(M_1)) - f_1(t + (i-1)\tau(M_2)) \right|^2 + \sum_{M_1+1}^{M_2+1} \left| f_1(t + (i-1)\tau(M_2)) \right|^2 \\ &= \sum_{i=1}^{M_1+1} |f'_1(c_i)|^2 |i-1|^2 |\tau(M_1) - \tau(M_2)|^2 + \sum_{M_1+1}^{M_2+1} \left| f_1(t + (i-1)\tau(M_2)) \right|^2 \\ &\leq \epsilon^2 (M_1 + 1)^3 \sqrt{\sum_{i=1}^{M_1+1} |f'_1(c_i)|^2} + \sum_{M_1+1}^{M_2+1} \left| f_1(t + (i-1)\tau(M_2)) \right|^2 \\ &= \epsilon^2 \cdot g(M_1, f_1) + h(M_1, M_2, f_1) \\ &\leq \left(\epsilon \cdot \sqrt{g(M_1, f_1)} + \sqrt{h(M_1, M_2, f_1)} \right)^2, \end{aligned}$$

since $g(M_1, f_1)$ and $h(M_1, M_2, f_1)$ are nonnegative constants for any given comparison, and $\epsilon > 0$.

Then $\text{d}_H \left(\text{SW}_{M_1, \tau(M_1)} f_{1|2}(t), \text{SW}_{M_2, \tau(M_2)} f_{1|2}(t) \right) \leq \epsilon \cdot \sqrt{g(M_1, f_1)} + \sqrt{h(M_1, M_2, f_1)},$

which implies that

$$d_B\left(\text{dgm}_1(\text{SW}_{M_1, \tau(M_1)} f_{1|2}(t)), \text{dgm}_1(\text{SW}_{M_2, \tau(M_2)} f_{1|2}(t))\right) \leq \epsilon \cdot 2\sqrt{g(M_1, f_1)} + 2\sqrt{h(M_1, M_2, f_1)}.$$

By a similar argument as in the proof of Theorem 9 (Equation (3)), this means that

$$\begin{aligned} & \left| \text{mp}\left(\text{dgm}_1(\text{SW}_{M_1, \tau(M_1)} f_{1|2}(t))\right) - \text{mp}\left(\text{dgm}_1(\text{SW}_{M_2, \tau(M_2)} f_{1|2}(t))\right) \right| \\ & \leq \epsilon \cdot 4\sqrt{g(M_1, f_1)} + 4\sqrt{h(M_1, M_2, f_1)} \\ \implies & \quad |\text{score}_{M_1}(f_1|f_2) - \text{score}_{M_2}(f_1|f_2)| \leq \frac{4}{\sqrt{3}} \left(\epsilon \cdot \sqrt{g(M_1, f_1)} + \sqrt{h(M_1, M_2, f_1)} \right). \end{aligned}$$

Hence the difference between the scores under M_1 and M_2 is bounded above by a linear function of ϵ with positive slope. Therefore, the smaller the ϵ -precision is, the closer these scores are to each other. \square

Theorem 13 tells us that the time lag is a Cauchy sequence of the embedding dimension, where smaller values of ϵ yield larger input embedding dimensions which increase the precision of convergence of $\text{score}_M(f_1|f_2)$ to $\text{score}(f_1|f_2)$. In other words, the greater the precision of convergence wanted, the smaller the value of ϵ we must choose, the greater the minimum embedding dimension required, and the closer the conditional periodicity scores for any pair of higher-valued dimensions will be. This result can be useful if we want to ensure that the conditional scores we are computing for a given input dimension M are similar to the underlying conditional periodicity score, albeit at the cost of higher run times and more embedding points required due to a larger embedding dimension. We next prove that small changes in ϵ induce small changes in minimum embedding dimension \mathcal{M} .

Corollary 14 (Stability of Minimum Embedding Dimension) *Let $f_1, f_2 : [0, 2\pi] \rightarrow \mathbb{R}$ be w_1 - and w_2 -periodic time series such that $w_2 > w_1$. Let $\epsilon > 0$ and fix the minimum embedding dimension for $\text{score}(f_1|f_2)$ to $M(\epsilon) = \mathcal{M}$ as specified in Theorem 13. Without loss of generality, let $\epsilon_1 > \epsilon_2 > 0$. Then for some $c \in (\epsilon_2, \epsilon_1)$, we have*

$$|M(\epsilon_1) - M(\epsilon_2)| \leq \frac{2\pi}{w_2} \cdot \frac{1}{c^2} \cdot |\epsilon_1 - \epsilon_2| + 2.$$

Proof By the Triangle inequality,

$$|M(\epsilon_1) - M(\epsilon_2)| \leq \left| \left\lceil \frac{2\pi}{w_2\epsilon_1} \right\rceil - \frac{2\pi}{w_2\epsilon_1} \right| + \left| \frac{2\pi}{w_1\epsilon_1} - \frac{2\pi}{w_2\epsilon_2} \right| + \left| \frac{2\pi}{w_2\epsilon_2} - \left\lceil \frac{2\pi}{w_2\epsilon_2} \right\rceil \right|.$$

Since $\left\lfloor \frac{2\pi}{w_2\epsilon} \right\rfloor \leq \frac{2\pi}{w_2\epsilon} \leq \left\lceil \frac{2\pi}{w_2\epsilon} \right\rceil$, the distance $\left| \frac{2\pi}{w_2\epsilon} - \left\lceil \frac{2\pi}{w_2\epsilon} \right\rceil \right|$ is at most 1. Hence,

$$|M(\epsilon_1) - M(\epsilon_2)| \leq \left| \frac{2\pi}{w_2\epsilon_1} - \frac{2\pi}{w_2\epsilon_2} \right| + 2.$$

Let $f(\epsilon) = \frac{1}{\epsilon}$. Then $f'(\epsilon) = -\frac{1}{\epsilon^2}$ and by MVT there exists some $c \in (\epsilon_2, \epsilon_1)$ such that

$$\left| \frac{1}{\epsilon_1} - \frac{1}{\epsilon_2} \right| = \frac{1}{c^2} |\epsilon_1 - \epsilon_2|.$$

Combining the above two expressions gives

$$|M(\epsilon_1) - M(\epsilon_2)| \leq \frac{2\pi}{w_2} \left| \frac{1}{\epsilon_1} - \frac{1}{\epsilon_2} \right| + 2 \leq \frac{2\pi}{w_2} \cdot \frac{1}{c^2} \cdot |\epsilon_1 - \epsilon_2| + 2. \quad \square$$

Hence, if we want higher precision of our score, then we must use larger embedding dimensions. In these higher-dimension cases, Corollary 14 ensures that small changes in the desired precision ϵ of $\text{score}(f_1|f_2)$ induce small changes in minimum embedding dimension \mathcal{M} . Intuitively, greater desired precision induces smaller distances $|\epsilon_1 - \epsilon_2|$ and $c \in (\epsilon_2, \epsilon_1)$, and hence yields smaller changes in \mathcal{M} . Oppositely, less desired precision increases $|\epsilon_1 - \epsilon_2|$ and c , and hence produces larger changes in \mathcal{M} .

Proposition 15 (A minimum number of embedding points) *Let $f_1, f_2 : T \subset [0, 2\pi] \rightarrow \mathbb{R}$ be two discrete univariate time series on $[0, 2\pi]$. Suppose $\frac{2\pi}{w_2} < \frac{2\pi}{w_1}$ for $w_1, w_2 \in \mathbb{N}$. Let $T = \{t_0 = 0 < t_1 < \dots < t_P = 2\pi\}$ for $P \in \mathbb{N}$. Let $M = M_1 \in \mathbb{N}$. Suppose there exists $\delta \in \mathbb{N}$ points in $[0, \frac{\pi}{w_2}]$. Then the minimum number of embedding points $N \in \mathbb{N}$ required to capture the relationship between the periodicities of f_1 and f_2 in the conditional SWE satisfies*

$$N \geq \frac{P}{w_1} - \delta.$$

Proof Let $M = M_1 \in \mathbb{N}$. Suppose there exists $\delta_1 < P$ points on $[0, M_1\tau]$ for $M_1\tau = \frac{2\pi}{w_2} \cdot \frac{M_1}{(M_1+1)}$. Then there exists $\frac{P}{w_1} - \delta_1$ points in $\text{SW}_{M_1, \tau} f_{1|2}([0, M_1\tau])$. Suppose we increase M to $M_2 > M_1$. Then there exists $\delta_2 > \delta_1$ points on $[0, M_2\tau]$, since

$$M_2\tau = \frac{2\pi}{w_2} \cdot \frac{M_2}{(M_2+1)} > \frac{2\pi}{w_2} \cdot \frac{M_1}{(M_1+1)} = M_1\tau.$$

Hence, there exists $\frac{P}{w_1} - \delta_2 < \frac{P}{w_1} - \delta_1$ points in $\text{SW}_{M_2, \tau} f_{1|2}([0, M_2\tau])$, since $\delta_2 > \delta_1$. Therefore, increasing M decreases the number of points in the conditional SWE. Hence, to maximize the number of points in $\text{SW}_{M, \tau} f_{1|2}(T)$, N , we must let $N \geq \frac{P}{w_1} - \delta_1$. The smallest such δ occurs when $M = 1$, since any larger dimension increases $[0, M\tau]$ and thus decreases $\frac{P}{w_1} - \delta$. Therefore, we let

$$N \geq \frac{P}{w_1} - \delta,$$

where δ is the number of points in $[0, \tau] = [0, \frac{\pi}{w_2}]$. □

3.3 Stability versus Convergence of $\text{score}(f_1|f_2)$

We highlight the difference between the stability and convergence results we have proven. The main difference is that although small changes in periodicity of f_2 and Gaussian noise in f_1 guarantee small changes in $\text{score}(f_1|f_2)$ (Theorem 9 and Proposition 10), this does not necessarily guarantee that the conditional scores computed are close to the underlying score. This closeness (i.e., convergence) is guaranteed whenever $M = \mathcal{M} = \left\lceil \frac{2\pi}{w_2\epsilon} \right\rceil$, as proven in Theorem 13. We show an example of this main difference in Figure 4, where we can see that although increased embedding dimensions may produce larger changes in the conditional periodicity score, this score remains closer to the underlying score when we choose $M = \mathcal{M}$.

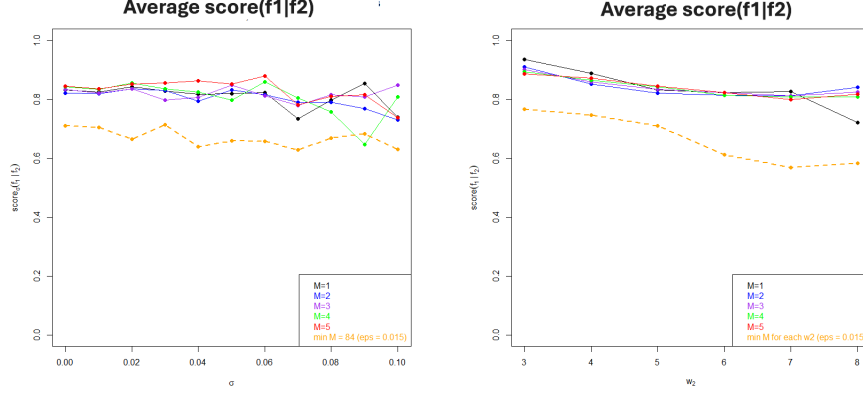


Fig. 4 An example of the stability versus convergence behavior of the conditional periodicity score. The left image shows the stability of our score against small changes in Gaussian noise versus the increased ‘closeness’ of our score to the underlying score when $M = \left\lceil \frac{2\pi}{w_2 \epsilon} \right\rceil$. The right image shows the same but for small changes in cycle length of f_2 , $\frac{2\pi}{w_2}$.

4 Computational Results

We present a framework for computing the conditional periodicity score using PCA. We then apply this framework on periodic signals of multiple types. We also compare the performance of our conditional periodicity score with that of percent determinism.

4.1 Procedure for Quantifying Conditional Periodicity

We introduce a procedure for computing the conditional periodicity score of a discrete time series $\{f_i(t_p)\}_{p=1}^P$ given another $\{f_j(t_p)\}_{p=1}^P$, $P \in \mathbb{N}$ (see Algorithm 1). We first fit two continuous time series f_i^{cts} and f_j^{cts} onto each discrete signal over the interval $[0, 2\pi]$ via cubic spline interpolation [13]. We then estimate the length of one cycle of f_i and f_j using the discrete fast Fourier transform (FFT) on the continuously-fitted series. More specifically, we compute the discrete fast Fourier spectra of f_i and f_j up to their Nyquist frequency of $\frac{P}{2}$. We then estimate w_i and w_j as the frequency indices corresponding to the maximum spectral densities corresponding to f_i and f_j , respectively. We assign f_1 as the series with the larger cycle-length and f_2 as that with the smaller length. We then compute the top K principal components of the conditional SWE.

Assuming the embedding dimension $M + 1$ is at least 3, we choose $K = 2$. This is because, assuming we are working with L -periodic time series, such embeddings will be elliptic curves and hence we presume that 2 principal components will suffice. As an example to motivate our choice of $K = 2$, see Figure 5. In this example, we fix the periodicity of f_1 so that $w_1 = 3$, the embedding dimension to $M = 10$, and increase the periodicity of f_2 so that w_2 varies from w_1 to 15. We set the standard deviation of noise to $\sigma = 0.15$, $P = 300$ time points, the simple moving average (SMA) window size to 19 points, and $N = \frac{P}{w_1} - 2 = 198$ (see Proposition 15). We compute the average variance captured by all principal components among 100 samples for each w_2 . Notice

that if w_2 is close to w_1 , then the conditional SWE is quite round, and hence two principal components capture roughly the same majority of variance in the data. If w_2 is much larger than w_1 , then the conditional SWE is more “skinny” and hence two principal components should still capture a majority of the variance, but the first component will capture more than the second.

We finally perform VR filtration on the SWE, obtaining the conditional periodicity score termed $\text{score}_\phi(\cdot)$ under the PCA projection ϕ .

Algorithm 1 Procedure for quantifying conditional periodicity

Inputs: Embedding dimension $M \in \mathbb{N}$, $N \in \mathbb{N}$ points in the conditional SWE, and two discrete univariate time series f_i, f_j defined on $\{t_1, \dots, t_P : P \in \mathbb{N}\}$.

Fit a continuous signal $f_k^{\text{cts}} : [0, 2\pi] \rightarrow \mathbb{R}$ to the discrete signal f_k via cubic spline interpolation such that $f_k^{\text{cts}}(0) = f_k(t_1)$ and $f_k^{\text{cts}}(2\pi) = f_k(t_P)$ for $k = i, j$.

Estimate $\frac{2\pi}{w_i}$ and $\frac{2\pi}{w_j}$ via spectral analysis using the (Discrete) Fast Fourier Transform. Assign f_1 to be the less periodic and f_2 to be the more periodic signal of the pair $\{f_i^{\text{cts}}, f_j^{\text{cts}}\}$. That is, define :

$$\frac{2\pi}{w_1} = \max \left\{ \frac{2\pi}{w_i}, \frac{2\pi}{w_j} \right\}, \quad \frac{2\pi}{w_2} = \min \left\{ \frac{2\pi}{w_i}, \frac{2\pi}{w_j} \right\}$$

$$f_1 = \left\{ f_k^{\text{cts}} : \frac{2\pi}{w_k} = \frac{2\pi}{w_1} \right\}, \quad f_2 = \left\{ f_k^{\text{cts}} : \frac{2\pi}{w_k} = \frac{2\pi}{w_2} \right\}$$

Define $\tau = \frac{2\pi}{w_2(M+1)}$.

Compute $X = \text{SW}_{M,\tau} f_{1|2}(T)$ for N evenly-spaced time points in $T = \left[0, \frac{2\pi}{w_1}\right]$.

For $K = \min\{2, M+1\}$, compute $\phi : \mathbb{R}^{M+1} \rightarrow \mathbb{R}^K$, the orthogonal projection of X onto its first K principal components ($K = 2$, typically).

Center and normalize $\phi(X)$ to obtain Y .

Compute $\text{dgm}_1(Y)$ from VR filtration on Y and $\text{mp}(\text{dgm}_1(Y))$ from $\text{dgm}_1(Y)$.

Compute $\text{score}_\phi(f_1|f_2)$ using $\text{mp}(\text{dgm}_1(Y))$.

Return: $\{\text{score}_\phi(f_1|f_2), w_1, w_2, f_i \text{ assignment}, f_j \text{ assignment}\}$.

Remark 16 (Computational Complexity of Conditional Periodicity Score) *Algorithm 1 runs in $O(P \log P + NK^2 + N^6)$ time, where P is the number of points in the discrete univariate input signals f_i and f_j , N is the number of points in the conditional SWE of the fitted continuous signals f_1 given f_2 , and $K < M+1$ is the number of principal components used for the conditional SWE where M is the embedding dimension. We take $K = 2$ by default. The cubic spline interpolation on f_i and f_j runs in $O(P)$ time [24] and the discrete FFT on $f_i^{\text{cts}}, f_j^{\text{cts}}$ can be computed in $O(P \log P)$ time [25]. The PCA computations run in $O(NK^2)$ time [21, 26]. The bottleneck step is usually the computation of the 1D persistence diagram using the VR filtration of Y , which runs in $O(N^6)$ time [19].*

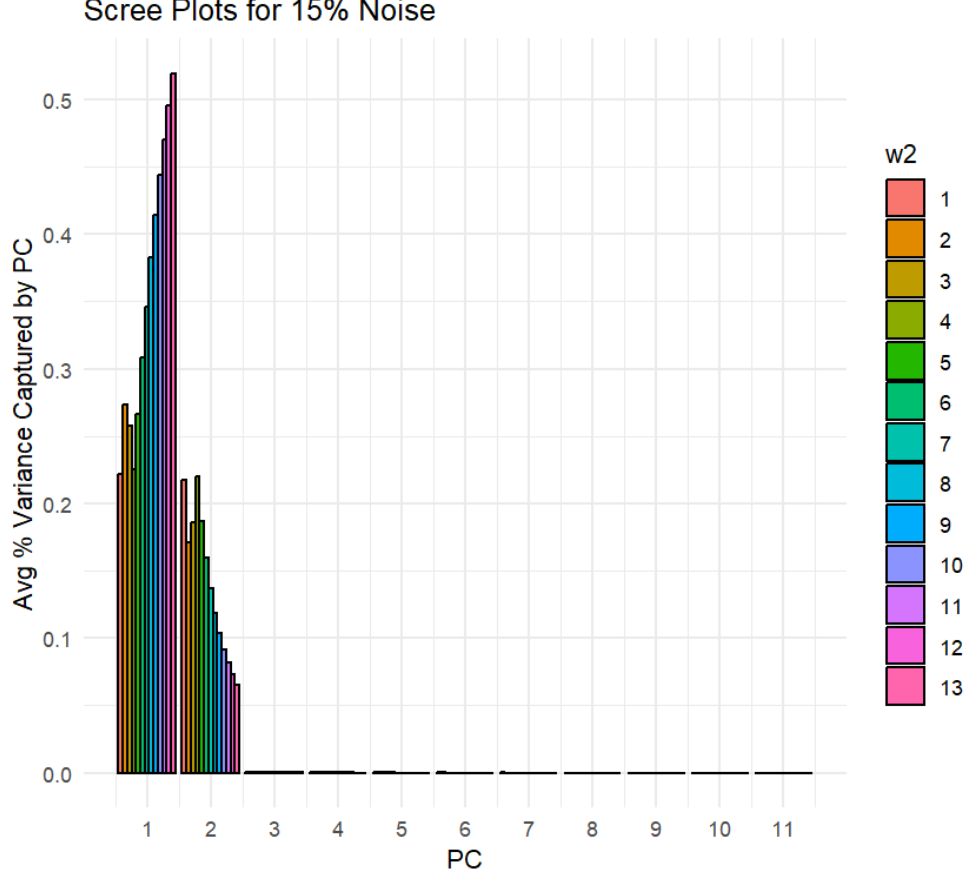


Fig. 5 Example of scree plots depicting the average variance captured by all principal components in the conditional SWE of a 1-periodic series given several more-periodic series (as shown by the w_2 values) over 100 samples.

4.2 Handling Noise

To handle noise when computing %DET, we fix the distance threshold to be greater than five times the standard deviation of Gaussian noise in each case of f_1 and f_2 [9]. To handle noise when computing $\text{score}(f_1|f_2)$, we perform two denoising methods on the discrete time series and on the resulting conditional SWE, respectively. To denoise the input series, we locally average every point in f_1 and f_2 by taking the mean of a window of points around it, and repeat this for each point. This process is called a simple moving average (SMA) and results in two averaged discrete signals. For our experiments, we pick $P = 300$ points in f_1 and f_2 . Since f_1 is w_1 -periodic and we compute the conditional SWE on one cycle of f_1 , we choose the SMA window size so that it is at most one-third the number of points in one cycle of f_1 (P/w_1). In this

way, we maintain a balance of enough points for denoising while also maintaining the periodic structure of the time series.

To denoise the conditional SWE of f_1 given f_2 , we apply a similar process called mean shifting. That is, for each point $\mathbf{v}_i \in \mathbb{R}^{M+1}$ in the point cloud, we average it with all of its neighbors \mathbf{v}_j , where \mathbf{v}_j is a neighbor of \mathbf{v}_i if the angle between these vectors is less than $\pi/16$ [13, 27].

4.3 Shape and Noise Robustness of Our Score

We conduct similar experiments to those performed by Perea and Harer [12, §7.1] to analyze the stability of $\text{score}(f_1|f_2)$ under small changes in Gaussian noise and periodic signal type of f_1 and f_2 . We consider periodic cosine, dampened cosine, triangle wave, and square wave signals with varying levels of dampening (5%–80%) and Gaussian noise (0%–75%) applied to f_1 and f_2 . We plot one example of graphs for each of these cases (see Figure 6). For this experiment, we produce a 4×3 grid plot, where each row corresponds to one signal type among cosine, dampened cosine, square wave, and sawtooth signals from top to bottom, respectively. The first column is a plot of a 3-periodic series f_1 and a 7-periodic series f_2 with 5% Gaussian noise.

The second column in Figure 6 is a plot of $\text{score}(f_1|f_2)$ against increasing embedding dimensions. This chart illustrates the application of Theorem 13. We vary M from 2 to 200, and include vertical lines at three embedding dimensions produced at three decreasing epsilon values. We define $P = 300$ time series points, the simple-moving average (SMA) window to 10 points, the standard deviation of Gaussian noise to $\sigma = 0.05$, and $M = \left\lceil \frac{2\pi}{w_2\epsilon} \right\rceil$ for $\epsilon \in \{0.1, 0.05, 0.02\}$. We maximize the number of embedding points N by fixing $N = P$.

The third column in Figure 6 shows the plots of average measures of $\text{score}(f_1|f_2)$ against increasing values of w_2 from $w_1 = 3$ to 15. We average over 100 samples and plot the average measures along with their 95% confidence intervals (CIs). The p -values for all CIs corresponding to each value of w_2 were less than 0.001. We obtain these using student’s t-testing. We define the parameters for this plot the same as for that in column two, except we fix the SMA window to 19 points and increase the precision of our score by setting $\epsilon = 0.0005$.

4.4 Comparing Stability of Our Score and Percent Determinism

The goal of this experiment is to show experimentally that $\text{score}(f_1|f_2)$ is more stable than %DET with respect to small changes in periodicity of f_2 (i.e., $2\pi/w_2$) and time lag (i.e., τ). For this experiment, we compare the average stability of $\text{score}(f_1|f_2)$ to %DET over 100 randomly generated time-series pairs $\{f_1(t), f_2(t)\}_{t=0}^{P=199}$ with 5% Gaussian noise when subject to small changes in time lag. We define the periodicity of f_1 to $w_1 = 2$ and vary the periodicity of f_2 from $w_2 = 2$ to $w_2 = 20$. By fixing $\epsilon = 0.2$ and selecting $w_2 = 2$, we use Theorem 13 to define $\mathcal{M} = 16$. We then let $M = \{16, 17, 18\}$. We further denoise the respective embeddings used (i.e., the conditional SWE and the SWEs of f_1 and f_2) with a simple moving average window of 12 points. To ensure an equal comparison between $\text{score}(f_1|f_2)$ and %DET, we compute both measures on the

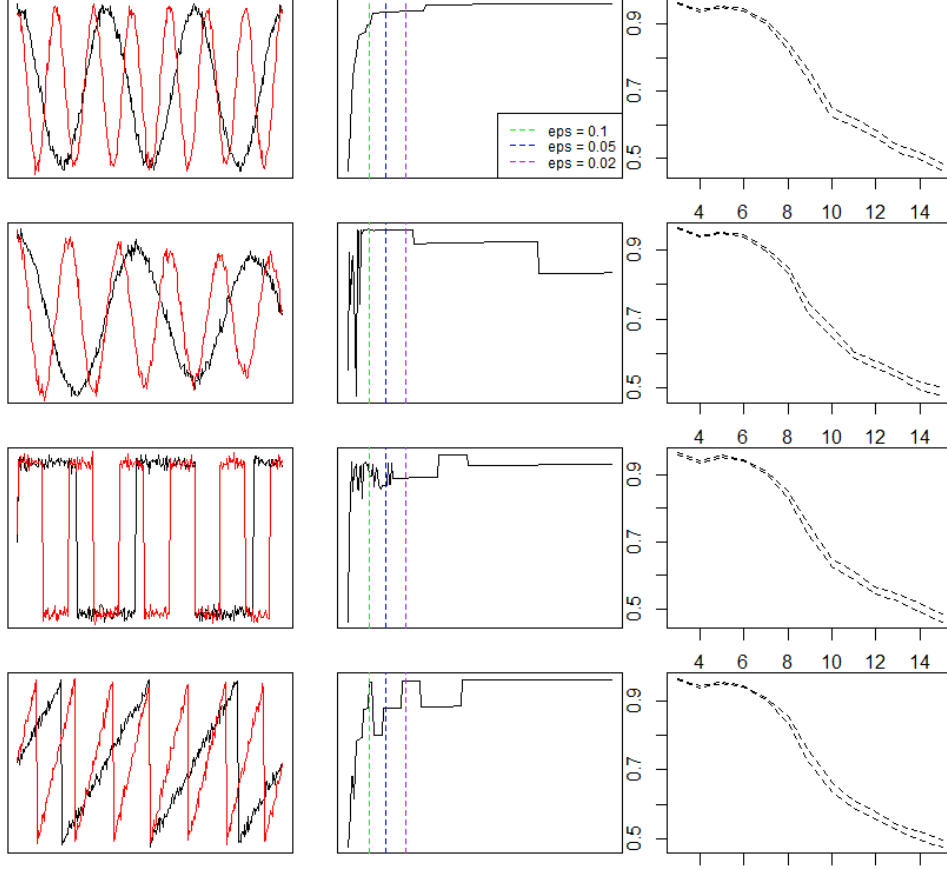


Fig. 6 An example showing the robustness of $\text{score}(f_1|f_2)$ to varying signal shapes and noise levels. We show 5% noisy time-series pairs (left column) of sinusoidal (top row), dampened sinusoidal (second row), square wave (third row), and triangle wave (fourth row) structure. We illustrate an application of Theorem 13 in column two, where we plot $\text{score}(f_1|f_2)$ for the pair of signals in the left column for increasing values of embedding dimension. We plot the minimum dimensions required to produce scores of ϵ precision for decreasing values of epsilon as vertical lines in the plots (green, blue, and violet for $\epsilon = 0.1, 0.05$, and 0.02 , respectively). In column three, we plot the 95% confidence intervals produced from 100 randomly sampled scores for each periodicity of f_2 as it diverges from that of f_1 . We fix the minimum embedding dimension using $\epsilon = 0.0005$.

top two PCs of these embeddings. For each case of embedding dimension, we fix the parameters for %DET to $\text{minDL} = 15$ and $\text{tol} = 0.9$. See Figure 7 for more details.

This experiment shows that for small changes in time lag and periodicity of f_2 , percent determinism is unstable for the same embedding dimensions in which our score is stable. When $\tau = 3$, %DET detects periodicity similarity well, but small changes in τ to 2 or 4 substantially decrease its stability with respect to small changes in periodicity of f_2 . These results experimentally illustrate the increased robustness

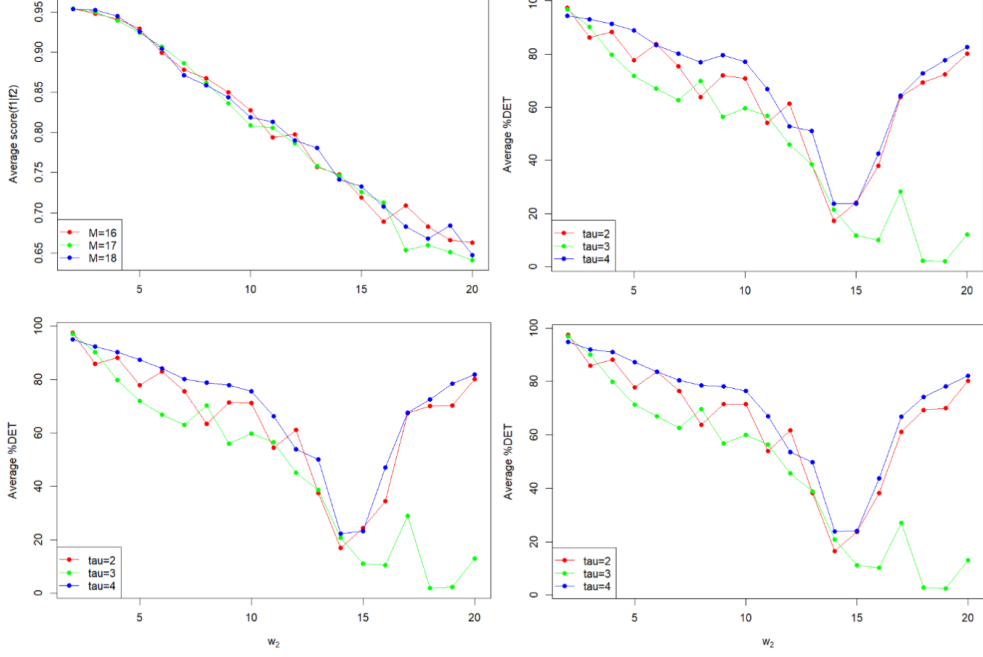


Fig. 7 An example of the increased stability of $\text{score}(f_1|f_2)$ (top left) when compared to %DET for embedding dimensions $M = 16$ (top right), $M = 17$ (bottom left), and $M = 18$ (bottom right) when subject to small changes in periodicity of f_2 (i.e., w_2) and time lag (i.e., ‘tau’). While %DET behaves well for $\tau = 3$, its behavior becomes much less stable for small changes in this time lag.

gained by our conditional scoring function resulting from the uniquely defined time lag used to compute it.

4.5 Periodicity Similarity of Incoming Police Agency Calls

We compute the periodicity similarity between different original call types to the Fife Police Department in the state of Washington in USA per month between 2019 and 2024 (see Table 1). This data is obtained from the Washington State Data Exchange for Public Safety [28]. We specifically compare the proportion of yearly incoming calls per month in this five-year time span. Hence, each time series from our data set contains 72 points. We first compute the periodicity score, $\text{score}(f|f) = \text{score}(f)$ (by Proposition 8), for each time series independently in order to determine if it is periodic enough (i.e., if its periodicity score is greater than 0.5). We consider time series with large enough periodicity scores for our analysis.

For each time series f , we vary the simple moving average window (SMA window) between 1 and 10 points, and select a value that produces the highest periodicity score, $\text{score}(f|f)$. We report this window size in column 2 of Table 1, as well as the number of cycles estimated by our conditional periodicity scoring algorithm (Algorithm 1), w_f , in column 3. We finally report $\text{score}(f|f)$ in column 4, as well as the description of each time series in column 5. To meet our assumptions for the conditional periodicity

f	SMA Window	w_f	$\text{score}(f f)$	Description
MVT	4	5	0.856	motor vehicle theft
THEV	5	6	0.847	theft from vehicle
CIT	5	8	0.778	citizen assist
UNW	6	3	0.793	unwanted person
MED	3	6	0.821	medical aid (call transferred to fire ASAP)

Table 1 Time series describing different types of incoming calls to Fife police agency per month.

score, we compare $\text{score}(\text{UNW}|f_2)$ to %DET for $f_2 = \{\text{MVT}, \text{THEV}, \text{CIT}, \text{MED}\}$, since UNW is the least-periodic time series.

Using Theorem 13, we choose $\epsilon = 0.75$ and $w_2 = 3$ (i.e., the estimated cycle number in UNW from Table 1) to define our embedding dimension as $M = 25$. To ensure a fair comparison, we compute both $\text{score}(f_1|f_2)$ and %DET on the top 2 principal components of the embeddings used for each measure. We additionally fix $\tau = 0.15$ for %DET. To ensure stability of our score in comparison with %DET, we vary $\text{tol} \in \{0.95, 1.00, 1.05\}$ and $\text{minDL} \in \{9, 10, 11\}$. We compute %DET values in $[0.879, 0.921]$ for UNW and MVT, in $[0.875, 0.978]$ for UNW and THEV, in $[0.189, 0.77]$ for UNW and CIT, and in $[0.891, 0.925]$ for UNW and MED. This indicates a drop in the periodicity similarity when UNW is compared to CIT for all combinations of tol and minDL . We report results in Figure 8 for the median values of these parameters, $\text{tol} = 1$ and $\text{minDL} = 10$.

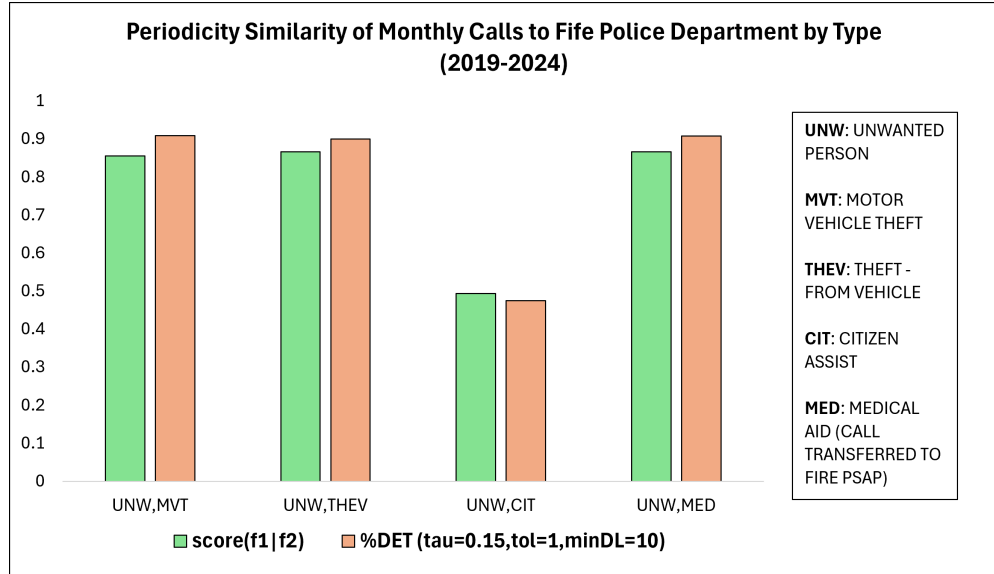


Fig. 8 The conditional periodicity score and percent determinism of different pairs of time series describing the proportion of yearly incoming calls to Fife police per month.

As shown, both measures reveal that the monthly proportion of incoming calls regarding citizen assists occur much more periodically than calls regarding unwanted person(s). As well, monthly calls regarding unwanted person(s) and those involving motor vehicle thefts, thefts from vehicles, and medical aid all occur with similar periodicity.

In Figure 9, we show the instability of %DET in detecting the periodicity dissimilarity between UNW and CIT when varying the time lag from $\tau = 0.15$ to $\tau = 0.14$ and $\tau = 0.16$. Here, the green and orange bars correspond to our original results in Figure 8. The pink and blue bars correspond to the changes in %DET that arise when we alter the time lag. When we fix $M = 25$, $\text{tol} = 1$, and $\text{minDL} = 10$, but vary τ slightly, the drop in periodicity similarity between UNW and CIT compared to all other time-series pairs is not detected by %DET. This again highlights the benefit of the fixed τ used in our conditional periodicity score that avoids the possibility of time lag causing instability as shown here for %DET computations.

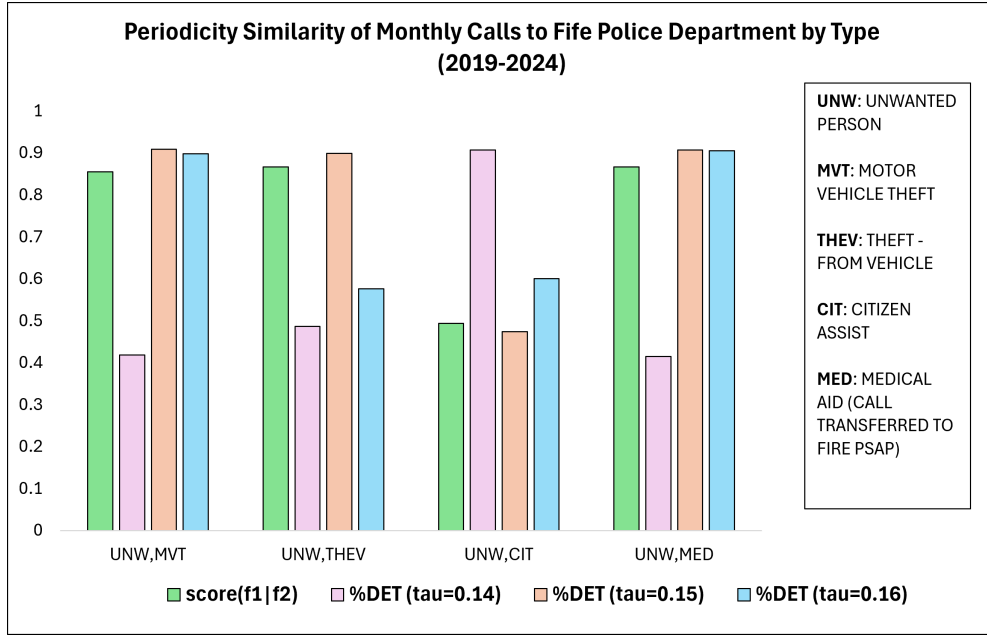


Fig. 9 The instability of %DET in detecting periodicity similarity between UNW and CIT for small changes in time lag.

5 Discussion

Our conditional periodicity score is a similarity measure that is unique in its guaranteed theoretical stability under small changes in periodicity. Complementing our theoretical stability results, our score can be applied efficiently in practice with

increased computational efficiency using PCA for dimension reduction. As well, our score is experimentally shown to be more robust than percent determinism when determining periodicity similarity, and only requires one input parameter compared to the four required by %DET. This parameter reduction is a highly favorable characteristic of $\text{score}(f_1|f_2)$ compared to %DET, as it reduces the number of biased decisions made during computation that can affect quantifications as shown in our experimental results. Our computations highlight the superior robustness of the conditional periodicity score in measuring periodicity similarity with respect to small changes in periodicity and time lag.

While our measure provides a parameter-reduced quantification of periodicity similarity that maintains unique theoretical stability and increased robustness compared to %DET, we plan to obtain results comparing our measure to %DET on synthetic time series of different types; e.g., cosine vs square-wave or triangle-wave vs dampened. We also plan to study how our score behaves when time series are relaxed to be generally periodic rather than L -periodic. In addition, we would like to obtain a stronger lower bound on the minimum number of embedding points required as a function of the optimal embedding dimension proven in Theorem 13. Finally, we would also like to obtain a lower bound on the minimum number of principal components needed so that the conditional periodicity score does not change much as discussed in Corollary 12. This result will help us to theoretically extend our work to other types of times series pairs whose expected conditional SWEs might not contain just two principal components capturing a majority of their variance.

References

- [1] Zhan, Y., Halliday, D., Jiang, P., Liu, X., Feng, J.: Detecting time-dependent coherence between non-stationary electrophysiological signals: A combined statistical and time-frequency approach. *Journal of Neuroscience Methods* **156**(1), 322–332 (2006) <https://doi.org/10.1016/j.jneumeth.2006.02.013>
- [2] Lachaux, J.-P., Lutz, A., Rudrauf, D., Cosmelli, D., Quyen, M.L.V., Martinerie, J., Varela, F.: Estimating the time-course of coherence between single-trial brain signals: An introduction to wavelet coherence. *Neurophysiologie Clinique/Clinical Neurophysiology* **32**(3), 157–174 (2002) [https://doi.org/10.1016/S0987-7053\(02\)00301-5](https://doi.org/10.1016/S0987-7053(02)00301-5)
- [3] Zhang, W., Bakhsh, S., Ali, K., Anas, M.: Fostering environmental sustainability: An analysis of green investment and digital financial inclusion in china using quantile-on-quantile regression and wavelet coherence approach. *Gondwana Research* **128**, 69–85 (2024) <https://doi.org/10.1016/j.gr.2023.10.014>
- [4] Beurs, D., Giltay, E.J., Nuij, C., O’Connor, R., Winter, R.F.P., Kerkhof, A., Ballegooijen, W., Riper, H.: Symptoms of a feather flock together? an exploratory secondary dynamic time warp analysis of 11 single case time series of suicidal ideation and related symptoms. *Behaviour Research and Therapy* **178**, 104572 (2024) <https://doi.org/10.1016/j.brat.2024.104572>

- [5] Mesbah, R., Koenders, M.A., Spijker, A.T., Leeuw, M., Hemert, A.M., Giltay, E.J.: Dynamic time warp analysis of individual symptom trajectories in individuals with bipolar disorder. *Bipolar Disorders* **26**, 44–57 (2024) <https://doi.org/10.1111/bdi.13340>
- [6] Wei, Z., Gao, Y., Zhang, X., Li, X., Han, Z.: Adaptive marine traffic behaviour pattern recognition based on multidimensional dynamic time warping and dbscan algorithm. *Expert Systems with Applications* **238**, 122229 (2024) <https://doi.org/10.1016/j.eswa.2023.122229>
- [7] Duong, S., Davis, T.J., Bachman, H.J., Votruba-Drzal, E., Libertus, M.E.: Dynamic structures of parent-child number talk: An application of categorical cross-recurrence quantification analysis and companion to duong et al. *The Quantitative Methods for Psychology* **20**, 137–155 (2024) <https://doi.org/10.20982/tqmp.20.2.p137>
- [8] Riehm, C.D., Bonnette, S., Rush, J.L., Diekfuss, J.A., Koohestani, M., Myer, G.D., Norte, G.E., Sherman, D.A.: Corticomuscular cross-recurrence analysis reveals between-limb differences in motor control among individuals with acl reconstruction. *Experimental Brain Research* **242**, 355–365 (2024) <https://doi.org/10.1007/s00221-023-06751-1>
- [9] Marwan, N., Romano, M.C., Thiel, M., Kurths, J.: Recurrence plots for the analysis of complex systems. *Physics Reports* **438**(5), 237–329 (2007) <https://doi.org/10.1016/j.physrep.2006.11.001>
- [10] Chazal, F., Silva, V., Glisse, M., Oudot, S.: *The Structure and Stability of Persistence Modules*, 1st edn. SpringerBriefs in Mathematics. Springer, ??? (2016)
- [11] Chazal, F., Silva, V., Oudot, S.: Persistence stability for geometric complexes. *Geometriae Dedicata* **173**, 193–214 (2014) <https://doi.org/10.1007/s10711-013-9937-z>
- [12] Perea, J., Harer, J.: Sliding windows and persistence: An application of topological methods to signal analysis. *Foundations of Computational Mathematics* **15**, 799–838 (2015) <https://doi.org/10.1007/s10208-014-9206-z> 1307.6188
- [13] Perea, J.A., Deckard, A., Haase, S.B., Harer, J.: Swlpers: Sliding windows and 1-persistence scoring; discovering periodicity in gene expression time series data. *BMC Bioinformatics* **16**(1), 257 (2015) <https://doi.org/10.1186/s12859-015-0645-6>
- [14] Tralie, C.J., Perea, J.A.: (quasi)periodicity quantification in video data, using topology. *SIAM Journal on Imaging Sciences* **11**(2), 1049–1077 (2018) <https://doi.org/10.1137/17M1150736>

- [15] Tymochko, S., Munch, E., Dunion, J., Corbosiero, K., Torn, R.: Using persistent homology to quantify a diurnal cycle in hurricanes. *Pattern Recognition Letters* **133**, 137–143 (2020) <https://doi.org/10.1016/j.patrec.2020.02.022>
- [16] Takens, F.: Detecting strange attractors in turbulence. In: Rand, D., Young, L.-S. (eds.) *Dynamical Systems and Turbulence*, Warwick 1980, pp. 366–381. Springer, Berlin, Heidelberg (1981)
- [17] Burago, D., Burago, Y., Ivanov, S.: *A Course in Metric Geometry*. Graduate Studies in Mathematics, vol. 33. American Mathematical Society, ??? (2001)
- [18] Adams, H., Frick, F., Majhi, S., McBride, N.: Hausdorff vs gromov-hausdorff distances (2024). <https://arxiv.org/abs/2309.16648>
- [19] Koyama, M.A., Memoli, F., Robins, V., Turner, K.: Faster computation of degree-1 persistent homology using the reduced Vietoris-Rips filtration. arXiv:2307.16333. <https://arxiv.org/abs/2307.16333> (2024)
- [20] Aggarwal, C.C., Hinneburg, A., Keim, D.A.: On the surprising behavior of distance metrics in high dimensional space. In: Bussche, J.V., Vianu, V. (eds.) *Database Theory — ICDT 2001*, pp. 420–434. Springer, Berlin, Heidelberg (2001)
- [21] Jolliffe, I.T.: *Principal Component Analysis* vol. 89. Springer, ??? (2002)
- [22] May, N.H., Krishnamoorthy, B., Gambill, P.: A normalized bottleneck distance on persistence diagrams and homology preservation under dimension reduction. *La Matematica*, 1–23 (2024) <https://doi.org/10.1007/s44007-024-00130-0>
- [23] Anselin, L.: *Dimension Reduction Methods (2): Distance Preserving Methods*. Available as part of *GeoDa: An Introduction to Spatial Data Science* (2020). https://geodacenter.github.io/workbook/7ab_mds/lab7ab.html
- [24] Dyer, S.A., Dyer, J.S.: Cubic-spline interpolation. 1. *IEEE Instrumentation & Measurement Magazine* **4**(1), 44–46 (2001) <https://doi.org/10.1109/5289.911175>
- [25] Brigham, E.O., Morrow, R.E.: The fast Fourier transform. *IEEE Spectrum* **4**(12), 63–70 (1967) <https://doi.org/10.1109/MSPEC.1967.5217220>
- [26] Yiğit, H.: Time Complexity of PCA 1. General Time Complexity of PCA. Posted on ResearchGate (2024). <https://doi.org/10.13140/RG.2.2.12847.34728>
- [27] Comaniciu, D., Meer, P.: Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**(5), 603–619 (2002) <https://doi.org/10.1109/34.1000236>
- [28] Public Safety (WADEPS), W.S.D.E.: What is CAD data? <https://wadeeps.org/the-data/cad-data/>