Topological Periodicity Test (TopPT) via Confidence Bound of Time-Delay Embeddings

Donghyun Park
Department of Statistics
Seoul National University
joshua8699@snu.ac.kr

Junhyun An
Department of Statistics
Seoul National University
anjunhyun@snu.ac.kr

Taehyoung Kim
Department of Statistics
University of Michigan
taehykim@umich.edu

Jisu Kim*
Department of Statistics
Seoul National University
jkim82133@snu.ac.kr

Abstract

Time-delay embedding is a fundamental technique in Topological Data Analysis (TDA) to reconstruct phase-space dynamics of time-series data, where persistent homology can reveal loops associated with periodicity. However, rigorous statistical uncertainty quantification for these features remains underdeveloped. First, we analyze the topology of time-delay embeddings, showing that the embedded trajectory is homotopy equivalent to a circle (S^1) for periodic signals and contractible for non-periodic ones. We also establish a positive lower bound on embedding reach, ensuring stable topological features. Second, we develop a subsampling approach to construct confidence bounds for persistence diagrams. Under standard manifold regularity conditions, we derive data-dependent bounds with asymptotic guarantees. Finally, we propose Topological Periodicity Test (TopPT), a hypothesis testing framework for periodicity with asymptotically controlled type I and type II error rates. Experiments on bounded-error synthetic data show that raw TDA detects periodic alternatives while avoiding false rejections on structured non-periodic signals, and that the robust rule is conservative after interpolation-error correction. On PhysioNet Fantasia and BIDMC respiratory waveforms, TDA detects most records but only a subset of local windows, unlike scalar periodogram baselines.

1 Introduction

Time series data consist of observations ordered in time, and understanding their underlying dynamics is a central problem in many scientific domains. In particular, detecting periodicity is crucial in applications where signals exhibit approximately repeating patterns. Classical approaches to periodicity detection primarily rely on spectral methods, such as the Generalized Lomb–Scargle periodogram [23], and autocorrelation-based techniques.

While these methods effectively identify frequency components, they may fail to capture geometric structures of underlying data. In contrast, Topological Data Analysis (TDA) provides a complementary perspective by focusing on the geometry of the data. Persistent homology summarizes topological features such as connected components and loops through persistence diagrams [19], and has proven effective in time-series analysis, where one-dimensional features capture cyclic behavior [16].

*Corresponding author.

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

Takens' embedding theorem [22] justifies delay coordinates as a reconstruction of the underlying state space; in particular, a generic delay embedding of a periodic orbit is diffeomorphic to S^1. This is a deterministic statement about noiseless, infinitely sampled trajectories: it does not quantify how persistent the resulting loop is, nor how far a finite noisy sample can deviate from it. Similarly, sliding-window methods such as SW1PerS [15] describe the topology of delay embeddings and its convergence, but do not quantify the uncertainty of the observed persistence diagram. Our work addresses this complementary, inferential question. We certify a positive reach that guarantees a nontrivial persistence for the loop, construct confidence bounds for the persistence diagram from finite samples, and convert the circle-versus-interval distinction into a test with controlled error rates.

Our contributions

First, we provide a theoretical characterization of time-delay embeddings, establishing a fundamental distinction between periodic and non-periodic signals. We prove that the embedded trajectory is homotopy equivalent to a circle (S^1) for periodic signals and to an interval for non-periodic ones. Moreover, we show that the reach of the sliding window embedding admits a positive lower bound, ensuring that persistent homology captures a stable topological feature with a nontrivial lifespan.

Second, we develop a statistical inference framework for persistence diagrams based on subsampling. By analyzing the Hausdorff distance between subsamples and the underlying support, we derive a data-dependent confidence radius with asymptotic guarantees. These results enable principled uncertainty quantification for topological features associated with periodicity.

Third, we propose a statistically valid hypothesis test for periodicity with controlled error rates. We show that the test achieves the prescribed significance level α and that both type I and type II errors are controlled under suitable conditions. Furthermore, we establish that these conditions are satisfied asymptotically, yielding consistency of the proposed test.

2 Preliminaries

In this section, we introduce the mathematical and statistical framework for our analysis. First, we describe time-delay embedding, which constructs point clouds from time series data. Then, we review persistent homology as a tool for extracting topological features.

2.1 Time-delay Embedding

Time-delay embedding is a method for reconstructing the phase space of a time-series, which is particularly effective for periodic data [12, 15, 22].

For a continuous signal $f \in C(\mathbb{R}, \mathbb{R})$, the sliding-window map with m delays, spaced by τ is defined as: $SW_{m,\tau} f(t) = [f(t), f(t+\tau), \cdots, f(t+m\tau)]^\top \in \mathbb{R}^{m+1}$. This process maps the time series to a point cloud X in high dimensional space $\mathbb{R}^{m+1}$. The embedding relies on two main parameters: the embedding dimension m and the time delay τ. Choosing appropriate values for m and τ is crucial to capture properties of time-series such as periodicity.

For a purely periodic signal, the reconstructed point cloud X generically traces a limit loop that is topologically equivalent to a one-dimensional circle S^1 [20, 22, 15]. In practice, noise and finite sampling can thicken or distort this manifold, but its essential topology remains a robust descriptor of the system's periodic nature.

2.2 Topological Data Analysis (TDA)

Topological Data Analysis (TDA) is a framework rooted in algebraic topology and computational geometry that provides structural properties of the data. It relies on the multi-scale analysis of topological features, a concept formalized through persistent homology [5, 24].

2.2.1 Persistent Homology, filtrations, and diagrams

In topological data analysis, algebraic topology is employed to characterize the intrinsic shape of data. For a given space, its topological features are described by homology groups, where the k-th

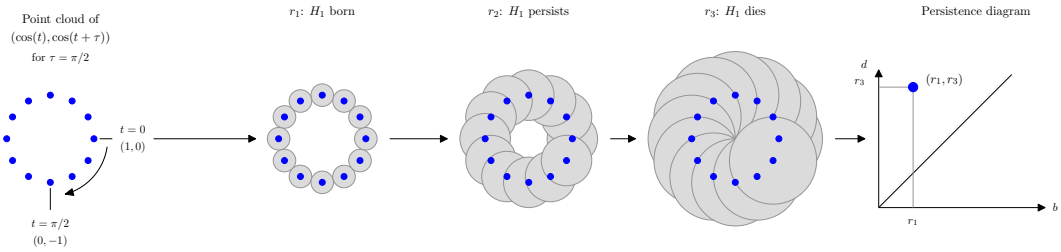


Figure 1: Example of Persistent diagram

Betti number β_k counts the number of k-dimensional features. Specifically, β_0 represents the number of connected components, and β_1 denotes the number of 1-dimensional loops [11].

To infer the underlying topology of a point cloud, we construct a filtration parameterized by a scale $\epsilon \geq 0$. While the Čech filtration $\mathcal{C}(\epsilon)$ accurately captures the exact topology of the ϵ-balls covering the data, it is computationally prohibitive. In practice, we utilize the Vietoris-Rips filtration $\mathcal{VR}(\epsilon)$, which efficiently constructs complexes using only pairwise distances. This substitution is theoretically justified by the tightly interleaved relationship between the two filtrations:

$$\mathcal{C}(\epsilon) \subseteq \mathcal{VR}(2\epsilon) \subseteq \mathcal{C}(2\epsilon).$$

As the scale parameter ϵ increases within a filtration, new topological features form (birth, denoted by b) and eventually merge or fill in (death, denoted by d). Persistent homology algebraically tracks these evolutionary lifespans [5]. The evolution of these homological features is concisely summarized in a persistence diagram $\mathcal{P}$. This diagram is a multiset of points $(b, d) \in \mathbb{R}^2$ with $b < d$, where each point represents a topological feature. The difference $d - b$ is referred to as the persistence of the feature. In practice, points located far from the diagonal $\Delta = \{(x, x) \mid x \in \mathbb{R}\}$ represent robust, prominent topological structures of the underlying space, whereas points close to the diagonal are typically regarded as sampling noise or minor irregularities.

A foundational property of persistence diagrams is their stability under perturbations [4]. For persistence diagrams constructed from Čech or Vietoris-Rips filtrations on discrete point clouds, the stability theorem implies that the bottleneck distance, a distance metric for persistent diagrams, is bounded above by the Hausdorff distance d_H between the underlying sets: $d_B(\mathcal{P}(X), \mathcal{P}(Y)) \leq d_H(X, Y)$.

We refer to Appendix B for formal definitions of persistent homology, filtrations, persistent diagrams, bottleneck distance and theoretical properties.

2.2.2 Confidence bounds for Persistence Diagrams

To provide statistical guarantees for persistence diagrams estimated from finite samples, Fasy et al. [6] developed subsampling-based confidence bounds. Let $S_n = \{X_1, \ldots, X_n\}$ be a sample from a distribution P supported on $\mathbb{M}$. When the subsample size b increases to infinity at a rate $b = o(n/\log n)$, they defined the function $L(t) = \frac{1}{\binom{n}{b}} \sum_{j=1}^{\binom{n}{b}} I(d_H(S_{b,n}^{(j)}, S_n) > t)$ where I is a $0/1$ indicator function and $S_{b,n}^{(j)}$ denotes the j-th subsample of size b of S_n.

Defining $c_\alpha = 2L^{-1}(\alpha)$, let $\hat{\mathcal{P}}$ be the persistence diagram of $\{X_1, \cdots, X_n\}$ and $\mathcal{P}$ be the persistence diagram of $\mathbb{M}$. The following theorem holds. See Section D for formal assumptions and statements.

Theorem 2.1 ([6] Theorem 3)**.** *Under the regularity condition of $\mathbb{M}$, for all large n,*

$$\mathbb{P}(d_B(\hat{\mathcal{P}}, \mathcal{P}) > c_\alpha) \leq \alpha + O\big((b/n)^{1/4}\big).$$

3 Assumptions and Statistical Models

We assume that there exists an underlying signal $f : \mathbb{R} \to \mathbb{R}$. We sample points $t_1, \ldots, t_n$ i.i.d. from a distribution Q on $\mathbb{T} = [T_{\min}, T_{\max}]$. For any m and τ, we assume that we can observe $\{SW_{m,\tau} f(t_j)\}_{1 \leq j \leq n} \subset \mathbb{R}^{m+1}$. Since the window uses values up to time $t + m\tau$, we write $J_{m,\tau} := [T_{\min}, T_{\max} + m\tau]$ for the interval on which the signal must satisfy the analytic assumptions.

Assumption 3.1 (Sampling density assumption)**.** The sampling distribution Q has a density q with respect to Lebesgue measure on $\mathbb{T}$, and there exist constants $0 < q_{\min} \le q_{\max} < \infty$ such that $q_{\min} \le q(t) \le q_{\max}$ for all $t \in \mathbb{T}$. Moreover, q is Lipschitz: $|q(s) - q(t)| \le K_q|s - t|$ for all $s, t \in \mathbb{T}$.

To characterize the recurrent behavior of a function, we distinguish between periodic and non-periodic functions using the phase trajectory $\Phi_f(t) = (f(t), f'(t))$.

Definition 3.2. Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function, and fix $\Xi > 0$. We say f is Ξ-periodic if for any $t \in \mathbb{R}$, $f(t + \Xi) = f(t)$. We say f is non-periodic on an interval I if Φ_f is injective on I; for any $t_1 \neq t_2 \in I$, $f(t_1) \neq f(t_2)$ or $f'(t_1) \neq f'(t_2)$.

Under these definitions, the sliding window image is guaranteed to be homeomorphic to a circle (S^1) for periodic functions, and to an interval for non-periodic functions. Crucially, the non-periodic condition prevents any self-intersection in the phase space, ensuring trivial 1-dimensional homology.

However, even if we have valid confidence interval for the sampling function, sometimes non-periodic functions can behave very similarly to a periodic function or vice versa, and they may not be distinguishable. Therefore, we impose quantitative separation.

Definition 3.3. Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function. We say f is (Ξ, d, ϵ)-periodic if f is Ξ-periodic, and for any $t_1, t_2 \in \mathbb{R}$ with $\min_{n \in \mathbb{Z}} |t_1 - t_2 + n\Xi| \ge d$,

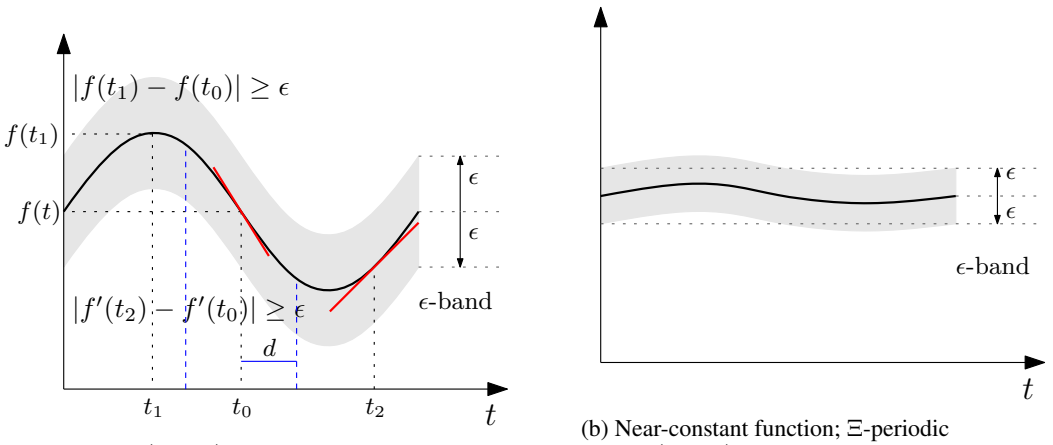
$$|f(t_1) - f(t_2)| \ge \epsilon \text{ or } |f'(t_1) - f'(t_2)| \ge \epsilon.$$

We say f is (d, ϵ)-non-periodic on an interval I if for all $t_1, t_2 \in I$ with $|t_1 - t_2| \ge d$,

$$|f(t_1) - f(t_2)| \ge \epsilon \text{ or } |f'(t_1) - f'(t_2)| \ge \epsilon.$$

Within d, the threshold of gap between times, and ϵ specifies the resolution at which two phase states are treated as separated. Larger ϵ gives stronger separation and a larger certified topology scale, while smaller ϵ allows a broader signal class but weakens the separation margin.

These separated classes intentionally exclude degenerate or mixed regimes. Constant and near-constant functions have ill-posed periods and degenerate delay clouds, while partial recurrence can create branching supports. See Figure 2 for illustrations of Ξ-periodic functions that are each (Ξ, d, ϵ)-periodic (Fig 2a) and not (Ξ, d, ϵ)-periodic (Fig 2b). We focus on globally periodic or globally non-periodic behavior, whose delay-space topology separates cleanly as shown in Section 4.



(a) (Ξ, d, ϵ)-periodic function

(b) Near-constant function; Ξ-periodic but not (Ξ, d, ϵ)-periodic function

Figure 2: Example of (Ξ, d, ϵ)-periodic function and not (Ξ, d, ϵ)-periodic function

Assumption 3.4. The underlying signal $f : \mathbb{R} \to \mathbb{R}$ is either (Ξ, d, ϵ)-periodic for some $\Xi, d, \epsilon > 0$, or (d, ϵ)-non-periodic on $J_{m,\tau}$ for some $\epsilon > 0$. In the periodic case, we assume the observation interval contains a full period, that is, $T_{\max} - T_{\min} \ge \Xi$.

We also impose regularity and scale-compatibility conditions, stated precisely in Appendix A. In brief, f is $C^{2,1}$ on the relevant domain, its first two derivatives are uniformly bounded, and the phase

trajectory is prevented from locally flattening in both first and second derivative:

$$\inf_{t\in J_{m,\tau}} \left(f'(t)^2 + f''(t)^2\right) > 0.$$

The separation scale d is required to be smaller than the curvature scale determined by these constants. These conditions, Assumptions A.1 and A.2, are used to certify positive reach of the sliding-window curve.

In practice, assessing the local nondegeneracy condition from noisy, discretely sampled observations requires estimating f' and f''. Such an assessment remains heuristic unless the estimation errors are controlled. Moreover, local nondegeneracy alone does not verify the global phase-separation condition in Definition 3.3.

In practice, instead of observing the exact time-delay embedded samples $\{SW_{m,\tau}f(t_j)\}$, we only have access to a finite sequence of discrete observations $f(s_0), f(s_1), \ldots, f(s_N)$ subject to measurement error. To bound the approximation error of this empirical embedding, we assume the following conditions:

Assumption 3.5 (Bounded-error linear interpolation)**.** Let $\mathcal{G}_N = \{s_0 < s_1 < \cdots < s_N\}$ be an interpolation grid covering $J_{m,\tau}$, and write $\Delta_N := \max_{0\le i<N}(s_{i+1} - s_i)$.

Instead of observing f exactly on this grid, suppose we observe

$$Y_i = f(s_i) + \xi_i, \qquad |\xi_i| \le \sigma_N, \quad 0 \le i \le N.$$

Let $\hat{f}_N$ be the piecewise-linear interpolant of $(s_i, Y_i)_{i=0}^N$. We assume that there exists some constant L_{int} and

$$|f''(t)| \le L_{\mathrm{int}} \qquad \text{for all } t \in J_{m,\tau}.$$

This assumption is only used for the robust confidence and testing statements. The interpolation error may be deterministic or random; if the bound $|\xi_i| \le \sigma_N$ holds only on a high-probability event, the final confidence bound gains the probability of the complement of that event.

4 Topology of Periodicity and non-Periodicity

In this section, we show that the sliding-window support is topologically trivial for non-periodic functions and circle-like for periodic functions. These support-level statements are the topological input for the testing procedure. Proofs of the results in this section are provided in Appendix E.

For a set $A \subset \mathbb{R}^{m+1}$ and $r > 0$, let $A^r := \{x : d(x, A) < r\},\ d(x, A) = \inf_{y\in A} \|x - y\|_2$ be the r-tubular neighborhood of A. We first consider the non-periodic case.

Theorem 4.1. *Suppose f satisfies Assumption 3.4, A.1, and A.2, and is (d, ϵ)-non-periodic on $J_{m,\tau}$. If*

$$m\tau > \max\{\delta/(2L), \epsilon/(4L)\}, \quad \tau < \min\{\delta/(8L), \epsilon/(16L)\}$$

then for every $0 < r < a_{m,\tau}$ where

$$a_{m,\tau} := \sqrt{m+1}\min\Big\{\frac{\epsilon}{32}\min\big\{1, \frac{\epsilon}{16L}\big\}, \frac{1}{2L}\frac{\delta^2}{16^2}\min\big\{1, \frac{\delta}{8L}\big\}^2\Big\},$$

we have that the tubular neighborhood $(SW_{m,\tau}f(\mathbb{T}))^r$ deformation retracts onto $SW_{m,\tau}f(\mathbb{T})$, which is homeomorphic to an interval. In particular, the offset is contractible.

In Theorem 4.1, the assumptions guarantee the delay curve of a non-periodic signal never comes back close to itself. In fact, the theorem quantifies this: thickening the curve by any radius $r < a_{m,\tau}$ creates no hole, since the thickened set can be continuously shrunk onto the curve (a deformation retraction), and the curve, being an interval, can be further shrunk to a point (it is contractible). This property is reflected in the persistence diagram, as the following corollary shows.

Corollary 4.2. *Under the assumptions of Theorem 4.1, let $\mathcal{P}_1(SW_{m,\tau}f)$ denote the 1-dimensional persistence diagram associated with the Čech filtration on $SW_{m,\tau}f(\mathbb{T})$. Then*

$$\mathcal{P}_1(SW_{m,\tau}f(\mathbb{T})) \cap \big\{\,[0, a_{m,\tau}) \times \mathbb{R}\big\} = \emptyset.$$

Next, we look at the periodic case.

Theorem 4.3. *Suppose f satisfies Assumption 3.4, A.1, and A.2, and is (Ξ, d, ϵ)-periodic. If*

$$m\tau > \max\{\delta/(2L), \epsilon/(4L)\}, \quad \tau < \min\{\delta/(8L), \epsilon/(16L)\}$$

then for every $0 < r < a_{m,\tau}$, we have that $(SW_{m,\tau} f(\mathbb{T}))^r$ deformation retracts onto $SW_{m,\tau} f(\mathbb{T})$, which is homeomorphic to S^1.

Corollary 4.4. *Under the assumptions of Theorem 4.3, let $\mathcal{P}_1(SW_{m,\tau} f)$ denote the 1-dimensional persistence diagram associated with the Čech filtration on $SW_{m,\tau} f(\mathbb{T})$. Then*

$$\mathcal{P}_1(SW_{m,\tau} f(\mathbb{T})) \cap \big\{ [0, a_{m,\tau}) \times \mathbb{R} \big\} = \{(0, D)\},$$

where $D \geq a_{m,\tau}$.

5 Subsampling Confidence Bound

While Section 4 showed the underlying support possesses a trivial or circular topology, our actual observations are restricted to finite samples. Accordingly, this section proposes a methodology for deriving confidence bounds from sampled data to infer the true topological features of the support. Proofs of the results in this section are provided in Appendix F and Appendix G.

5.1 Confidence bound

Based on the observed sample $T = \{t_0, t_1, \cdots, t_N\} \subset \mathbb{T}$, we obtain the empirical sliding window embedding $X_{m,\tau} = \mathrm{SW}_{m,\tau}\, f(T)$, with its true underlying support denoted by $M_{m,\tau} = \mathrm{SW}_{m,\tau}\, f(\mathbb{T})$.

To construct confidence bounds, we adapt the subsampling strategy introduced by [6].

We first begin by outlining the algorithmic procedure, followed by a formal derivation of its theoretical guarantees.

Let $X^{(j)}_{m,\tau,b}$ denote the j-th subsample of size b drawn from $X_{m,\tau}$. To estimate the distribution of the Hausdorff distance between the empirical sample and its subsamples, we compute the following empirical estimator over B iterations: $\bar{L}_{m,\tau}(t) = B^{-1} \sum_{j=1}^{B} I\big(d_H(X_{m,\tau}, X^{(j)}_{m,\tau,b}) > t\big)$.

The complete procedure is formalized in the following algorithm.

Algorithm 1 Implementation of Confidence bound

Input: $(t_0, Y_0), (t_1, Y_1), \cdots, (t_N, Y_N)$, m, n, b, τ, α.
1. Compute the sliding window embedding of the sample $X_{m,\tau} = SW_{m,\tau} f(\{t_0, t_1, \ldots, t_N\})$.
(Alternatively, under the error assumptions of Section 5.2, $\hat{X}_{m,\tau} = SW_{m,\tau} \tilde{f}_N(\{t_0, t_1, \ldots, t_N\})$)
2. Compute Monte Carlo subsamples $X^{(j)}_{m,\tau,b}$, for $j = 1, \ldots, B$.
3. Compute upper α quantile of Hausdorff distances $\bar{L}^{-1}_{m,\tau}(\alpha)$.
Output: $c_\alpha = 2\bar{L}^{-1}_{m,\tau}(\alpha)$

Our subsequent theorem demonstrates the construction of valid confidence bounds for the homological features of the support based on the empirical sample. To this end, we define the theoretical tail probability function, computed over all $\binom{n}{b}$ possible subsamples of size b:

$$L_{m,\tau}(t) = \frac{1}{\binom{n}{b}} \sum_{j=1}^{\binom{n}{b}} I\left(d_H(X_{m,\tau}, X^{(j)}_{m,\tau,b}) > t\right).$$

For this monotonically non-increasing tail function, we adopt the standard convention for its generalized inverse:

$$L^{-1}_{m,\tau}(\alpha) := \inf\{t \geq 0 : L_{m,\tau}(t) \leq \alpha\}.$$

Theorem 5.1. *Fix embedding parameters m, τ. Suppose the sampling distribution satisfies Assumption 3.1. Suppose f satisfies Assumption 3.4, A.1, and A.2. Suppose m, τ satisfies*

$$m\tau > \max\{\delta/(2L), \epsilon/(4L)\}, \quad \tau < \min\{\delta/(8L), \epsilon/(16L)\}.$$

Let $b = o(n/\log n)$ be a sequence tending to infinity as $n \to \infty$, and define $c_\alpha = 2(L_{m,\tau})^{-1}(\alpha)$. Then,

$$P\Big(d_B(\mathcal{P}(X_{m,\tau}), \mathcal{P}(M_{m,\tau})) > c_\alpha\Big) \le \alpha + O\big((b/n)^{1/4}\big).$$

Theorem 5.1 says that the distance between the sample and the unknown support can be calibrated by the distances between random subsamples and the sample. With probability at least $1 - \alpha - o(1)$, every point of the empirical diagram can be matched to a point of the true diagram $\mathcal{P}(M_{m,\tau})$, or to the diagonal, with each coordinate moving by at most c_α. In particular, a feature with persistence larger than $2c_\alpha$ cannot be matched to the diagonal and must reflect a genuine feature of the support. Lemma 5.2 adds that this radius vanishes as the sample grows, so the separation of Section 4 eventually becomes visible in the data.

Lemma 5.2. *Under the assumptions of Theorem 5.1, let m, τ and $\alpha \in (0,1)$ be fixed. If $b \to \infty$ and $b \le n$, then*

$$c_\alpha \xrightarrow{P} 0.$$

5.2 Confidence bound under the error assumption

If we assume the bounded error assumption 3.5, we could employ a linear interpolation to approximate the time-delay embedded points. Under the assumption 3.5, linear interpolation error is bounded by

$$\eta_N := \sigma_N + \frac{\Delta_N^2}{8} L_{\text{int}}, \qquad \zeta_{m,N} := \sqrt{m+1}\,\eta_N.$$

Let $\hat{X}_{m,\tau} := SW_{m,\tau}\,\tilde{f}_N(T)$ be the observed sliding-window cloud obtained from the interpolation. For all size b subsamples of $\hat{X}_{m,\tau}$, define $\hat{L}_{m,\tau}(t) = \frac{1}{\binom{n}{b}} \sum_{j=1}^{\binom{n}{b}} I\big(d_H(\hat{X}_{m,\tau}, \hat{X}^{(j)}_{m,\tau,b}) > t\big)$.

Let $\hat{c}_\alpha := 2\hat{L}^{-1}_{m,\tau}(\alpha)$ and $\hat{c}^{\text{err}}_\alpha := \hat{c}_\alpha + 5\zeta_{m,N}$, where $\hat{L}^{-1}_{m,\tau}$ uses the same generalized inverse convention as $L^{-1}_{m,\tau}$. Then, the analogous of Theorem 5.1 and Lemma 5.2 holds in this interpolation settings.

Theorem 5.3 (Confidence bound with bounded interpolation error)**.** *Assume the conditions of Theorem 5.1 and Assumption 3.5. Then,*

$$P\Big(d_B(\mathcal{P}(\hat{X}_{m,\tau}), \mathcal{P}(M_{m,\tau})) > \hat{c}^{\text{err}}_\alpha\Big) \le \alpha + O\big(\,(b/n)^{1/4}\,\big).$$

Each interpolated delay vector is within $\zeta_{m,N}$ of its exact counterpart. Theorem 5.3 shows that adding $5\zeta_{m,N}$ to the estimated subsampling radius yields the stated confidence bound for $d_B(P(\hat{X}_{m,\tau}), P(M_{m,\tau}))$.

Corollary 5.4. *Under the assumptions of Theorem 5.3, fix m, τ and let $\alpha \in (0,1)$ be fixed. If $b \to \infty$, $b \le n$, and $\zeta_{m,N} \to 0$, then*

$$\hat{c}^{\text{err}}_\alpha \xrightarrow{P} 0.$$

6 Topological Periodicity Test (TopPT)

The quantitatively periodic and non-periodic classes are disjoint but not complements, so we restrict the hypotheses to these two separated regimes. For fixed embedding parameters m, τ,

$$H_0 : f \text{ is } (d,\epsilon)\text{-non-periodic on } J_{m,\tau}, \qquad H_1 : f \text{ is } (\Xi, d, \epsilon)\text{-periodic for some } \Xi > 0.$$

At the support level, Corollary 4.4 gives a genuine 1-cycle born at zero and persisting at least until $a_{m,\tau}$ under H_1, whereas Corollary 4.2 gives no 1-dimensional feature born before $a_{m,\tau}$ under H_0. Thus a confidence radius smaller than $a_{m,\tau}/4$ separates the two cases.

Motivated by this topological separation, we introduce the Topological Periodicity Test (TopPT), which provides a statistical test for periodicity based on persistent homology. Proofs of the results in this section are provided in Appendix H.

Definition 6.1 (Topological Periodicity Test (TopPT))**.** Let Y denote an observed delay point cloud, let $\mathcal{P}(Y) = \mathcal{P}_1(Y)$ be its ambient distance/Čech diagram, and let r_α be a generic confidence radius. Define rejection region and associated separation event

$$R(r_\alpha) := \{(b, d) \in \mathbb{R}^2 : b \le r_\alpha,\ d - b > 2r_\alpha\}, \qquad A^r_\alpha := \{a_{m,\tau} > 4r_\alpha\}.$$

Using the boundary convention that assigns equality in birth to the rejection region and equality in persistence to its complement, we define Topological Periodicity Test (TopPT) as follows:

$$\psi(Y) = 0 \text{ if } \mathcal{P}(Y) \cap R(r_\alpha) = \emptyset, \qquad \psi(Y) = 1 \text{ if } \mathcal{P}(Y) \cap R(r_\alpha) \neq \emptyset.$$

The $\psi(Y) = 1$ corresponds to rejection of H_0 and existence of persistent periodic 1-cycle.

For exact observations, we apply Definition 6.1 with $(Y, r_\alpha, \psi_{m,\tau,r_\alpha}) = (X_{m,\tau}, c_\alpha, \phi)$. For bounded-error interpolation, we instead use $(Y, r_\alpha, \psi_{m,\tau,r_\alpha}) = (\hat{X}_{m,\tau}, \hat{c}^{\text{err}}_\alpha, \hat{\phi})$, where $\hat{c}^{\text{err}}_\alpha$ is defined in Section 5.2. This yields the testing rule used in the bounded interpolation setting.

The following image illustrates the persistence diagram of $X_{m,\tau}$ with a true persistence point $(0, d)$ under H_1.

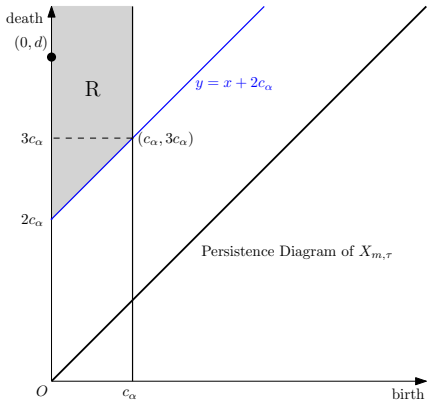


Figure 3: Diagram of Hypothesis Testing

Theorem 6.2 (Testing error bound)**.** *Assume either the exact-observation confidence bound of Theorem 5.1 with $(Y, r_\alpha) = (X_{m,\tau}, c_\alpha)$, or the bounded-interpolation confidence bound of Theorem 5.3 with $(Y, r_\alpha) = (\hat{X}_{m,\tau}, \hat{c}^{\text{err}}_\alpha)$. Then, under H_0,*

$$P(\psi(Y) = 1, A^r_\alpha \mid H_0) \le \alpha + O\big((b/n)^{1/4}\big),$$

and, under H_1,

$$P(\psi(Y) = 0, A^r_\alpha \mid H_1) \le \alpha + O\big((b/n)^{1/4}\big).$$

Consequently,

$$P(\psi(Y) = 1 \mid H_0) \le \alpha + P((A^r_\alpha)^c \mid H_0) + O\big((b/n)^{1/4}\big),$$

and

$$P(\psi(Y) = 0 \mid H_1) \le \alpha + P((A^r_\alpha)^c \mid H_1) + O\big((b/n)^{1/4}\big).$$

Corollary 6.3 (Asymptotic testing)**.** *For fixed m, τ, let $b \to \infty$ with $b = o(n/\log n)$. The exact rule satisfies*

$$P(\phi(X_{m,\tau}) = 1 \mid H_0) \le \alpha + o(1), \qquad P(\phi(X_{m,\tau}) = 0 \mid H_1) \le \alpha + o(1).$$

The same conclusion holds for the bounded-interpolation rule whenever $\zeta_{m,N} \to 0$, with $\phi(X_{m,\tau})$ replaced by $\hat{\phi}(\hat{X}_{m,\tau})$.

Signals outside $H_0 \cup H_1$. The hypotheses are non-exhaustive: rather than partitioning all signals, we restrict H_0 and H_1 to regimes whose delay-space geometries can be rigorously separated. Near-periodic, quasi-periodic, partially recurrent, and drifting signals lie between these regimes and admit no canonical assignment to either hypothesis without additional, arbitrary thresholds, so we do not force them into one.

Table 1: Synthetic rejection rates averaged over bounded-error settings. For periodic signals, entries are detection rates; for non-periodic signals, entries are false rejection rates.

Signal	class	raw TopPT	robust TopPT	Lomb–Scargle	Fisher g	permutation
f_1	periodic	1.000	0.706	1.000	1.000	1.000
f_2	periodic	1.000	0.457	1.000	1.000	1.000
h_1	non-periodic	0.000	0.000	1.000	1.000	1.000
h_2	non-periodic	0.000	0.000	1.000	1.000	1.000
h_3	non-periodic	0.000	0.000	1.000	1.000	1.000
h_4	non-periodic	0.000	0.000	1.000	1.000	1.000

6.1 Finite parameter searches

The guarantees above are stated for fixed m, τ. For a fixed finite grid $\mathcal{G} = \{(m_l, \tau_l) : 1 \le l \le K\}$, run the exact or bounded-error rule at level α/K for each candidate and let ψ_l and $A^r_{\alpha/K,l}$ denote the corresponding decision and separation event. Define

$$\Phi_{\mathcal{G}} = 1 \iff \psi_l = 1 \text{ for at least one } 1 \le l \le K.$$

Corollary 6.4. *Assume K is fixed. If, under H_0, every candidate in $\mathcal{G}$ satisfies the non-periodic assumptions of the corresponding confidence theorem, then*

$$P\left(\Phi_{\mathcal{G}} = 1, \bigcap_{l=1}^{K} A^r_{\alpha/K,l} \mid H_0\right) \le \alpha + O\big(K\,(b/n)^{1/4}\big).$$

If, under H_1, at least one candidate l_* satisfies the periodic assumptions, then

$$P\left(\Phi_{\mathcal{G}} = 0, A^r_{\alpha/K,l_*} \mid H_1\right) \le \alpha/K + O\big((b/n)^{1/4}\big).$$

Consequently, if the relevant separation events have probability tending to one, the finite-grid test controls both error probabilities at level $\alpha + o(1)$.

In practice, candidate pairs (m, τ) can be explored using the fixed finite-grid procedure above. The choice of embedding parameters affects the resulting topological features, and there is no universally optimal selection strategy.

7 Experiments

We evaluate the proposed periodicity test TopPT on synthetic signals and respiratory waveform data. All TopPT experiments use raw sliding-window embeddings, and persistence diagrams are computed with Vietoris–Rips (VR) complexes for computational efficiency. Our guarantees are stated for Čech persistence and transfer to VR only up to the constant factors of the interleaving in Section 2.2.1, so the VR rule is a computational surrogate.

Full signal definitions, preprocessing steps, parameter choices, and reproducibility details are given in Appendix I. For scalar baselines, we use the Generalized Lomb–Scargle periodogram (GLS) [23], Fisher's g-test [8], and a permutation max-periodogram test, all run on the same scalar observations used to build the delay clouds.

7.1 Synthetic Data

We test two periodic signals with known period $\Xi = 2\pi/5$ and four non-periodic signals with trend-like or low-frequency behavior. The goal is to check whether TopPT detects genuine delay-space loops without rejecting structured but non-periodic examples. We use the known-period delay choice $\tau = \Xi/(m+1)$ and compare raw TopPT (radius $\hat{c}_\alpha$), robust TopPT (the bounded-error interpolation rule in Section 5.2), GLS, Fisher's g-test, and a permutation max-periodogram test. Signal definitions, noise models, interpolation details, and parameter choices are in Appendix I.1.

Table 1 shows that raw TopPT detects both periodic signals and does not reject the four structured non-periodic examples. Robust TopPT is more conservative: it keeps the same zero false rejection rate but has lower power because the confidence radius includes the interpolation-error correction.

Table 2: Real-data rejection counts. TopPT columns report the number of rejected records and local windows. Baseline entries are written as record-level rejections followed by window-level rejections.

Dataset	records	windows	TopPT records	TopPT windows	Lomb–Scargle	Fisher g/permutation
Fantasia	40	280	36/40	143/280	40/40, 280/280	40/40, 280/280
BIDMC	53	318	50/53	252/318	53/53, 318/318	53/53, 318/318

The scalar periodogram baselines reject all signals in this experiment, including the structured non-periodic examples. This illustrates the difference between the two notions of periodicity: $h_1, \ldots, h_4$ contain strong low-frequency structure, which the spectral tests detect, but their delay embeddings do not form persistent loops, so TopPT does not reject. Conversely, a noisy periodic signal may still produce a clear spectral peak while its delay-space loop is too distorted to exceed the TopPT confidence radius. TopPT is therefore a complement to spectral methods rather than a replacement.

7.2 Real Data

We evaluate local respiratory periodicity using the Fantasia and BIDMC datasets from PhysioNet [9]. Fantasia contains 40 healthy resting-subject recordings with respiration channels [13, 10], and BIDMC contains 53 critical-care recordings with impedance respiratory signals [18, 17]. Each record is tested over local windows, with candidate periods estimated from autocorrelation and Welch spectral peaks and delay $\tau = \mathrm{period}/(m+1)$. We apply a Bonferroni correction over the realized finite window-period search set; full preprocessing and search-grid details are in Appendix I.3.

Table 2 shows that TopPT rejects for most records while rejecting only a subset of local windows: 36 of 40 Fantasia records and 50 of 53 BIDMC records. The scalar periodogram baselines reject every window in both datasets. Since these respiratory records are expected to contain cyclic components but do not provide verified local periodicity labels, this experiment is a detection and localization study rather than a ground-truth benchmark; the synthetic experiments, with known periodic and non-periodic signals, are the main evaluation of error rates. The difference between TopPT and the scalar baselines suggests that TopPT is more selective: it rejects when the delay embedding contains a persistent loop beyond the subsampling confidence radius, rather than whenever the scalar time series has strong oscillatory content. We report only raw TopPT on real data, since the robust rule requires certified measurement-error and interpolation-smoothness bounds that recordings do not provide.

8 Concluding Remarks

TopPT is a statistical framework for testing periodicity via the topological separation of periodic (S^1-like) and non-periodic (interval-like) sliding-window supports. A quantitative bound locates these features in persistence diagrams. Combining this with a subsampling confidence radius, we construct a hypothesis test with asymptotic error control, extended to interpolated scalar observations. We also provide a multiple testing framework for candidate window sizes and dimensions. Empirical results on synthetic and real data demonstrate that our approach does not simply evaluate whether a scalar time series contains oscillatory energy, but certifies the presence of a coherent persistent loop beyond the subsampling confidence radius. Thus, non-rejection does not imply an absence of oscillation, but rather the absence of a statistically certified loop at the tested window, period, and embedding scale. This provides a useful intermediate description between fully periodic and non-periodic behavior, especially for real signals whose rhythm may drift in amplitude, phase, or waveform shape.

Several directions remain open. A natural extension is to quasi-periodic signals, such as $f(t) = F(\omega_1 t, \ldots, \omega_N t)$ with F 2π-periodic in each argument and incommensurate frequencies $\omega_1, \ldots, \omega_N$. Their delay embeddings may trace out an N-dimensional torus rather than a circle [14], with N independent 1-cycles and higher-dimensional features, so the circle-versus-interval reach and persistence arguments of Section 4 do not apply directly. TopPT may still reject for such signals, but a rejection would then indicate a persistent loop rather than strict periodicity. A test for quasi-periodicity would require new guarantees for toroidal persistence signatures, possibly with an unknown number of frequencies N; partial recurrence and almost-periodic dynamics raise similar questions.

Additionally, establishing error guarantees for adaptive period selection is important since candidate periods for real data are naturally estimated from the same observations being tested.

Acknowledgments and Disclosure of Funding

This work was supported in part by the New Faculty Startup Fund from Seoul National University (No. 0670-20240073), and by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No. RS-2021-II211343, Artificial Intelligence Graduate School Program (Seoul National University)).

References

[1] Eddie Aamari, Jisu Kim, Frédéric Chazal, Bertrand Michel, Alessandro Rinaldo, and Larry Wasserman. Estimating the reach of a manifold. *Electronic Journal of Statistics*, 13(1):1359–1399, 2019.

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

[3] Frédéric Chazal, Vin Silva, and Steve Oudot. Persistence stability for geometric complexes. *Geometriae Dedicata*, 173, 07 2012.

[4] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. In *Proceedings of the Twenty-First Annual Symposium on Computational Geometry*, SCG '05, page 263–271, New York, NY, USA, 2005. Association for Computing Machinery.

[5] Herbert Edelsbrunner, David Letscher, and Afra Zomorodian. Topological persistence and simplification. *Discrete & Computational Geometry*, 28(4):511–533, 2002.

[6] Brittany Terese Fasy, Fabrizio Lecci, Alessandro Rinaldo, Larry Wasserman, Sivaraman Balakrishnan, and Aarti Singh. Confidence sets for persistence diagrams. *The Annals of Statistics*, 42(6), December 2014.

[7] Herbert Federer. Curvature measures. *Transactions of the American Mathematical Society*, 93(3):418–491, 1959.

[8] R. A. Fisher. Tests of significance in harmonic analysis. *Proceedings of the Royal Society of London. Series A*, 125(796):54–59, 1929.

[9] Ary L. Goldberger, Luis A. N. Amaral, Leon Glass, Jeffrey M. Hausdorff, Plamen Ch. Ivanov, Roger G. Mark, Joseph E. Mietus, George B. Moody, Chung-Kang Peng, and H. Eugene Stanley. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220, 2000.

[10] N. Iyengar, Chung-Kang Peng, R. Morin, Ary L. Goldberger, and Lewis A. Lipsitz. Age-related alterations in the fractal scaling of cardiac interbeat interval dynamics. *American Journal of Physiology*, 271:R1078–R1084, 1996.

[11] James R. Munkres. *Elements of Algebraic Topology*. Addison-Wesley Publishing Company, Menlo Park, California, 1984.

[12] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw. Geometry from a time series. *Phys. Rev. Lett.*, 45:712–716, Sep 1980.

[13] Chung-Kang Peng and Lewis A. Lipsitz. Fantasia Database. PhysioNet, 2003.

[14] Jose A. Perea. Persistent homology of toroidal sliding window embeddings. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6435–6439, 2016.

[15] Jose A. Perea and John Harer. Sliding windows and persistence: An application of topological methods to signal analysis. *Foundations of Computational Mathematics*, 15(3):799–838, 2015.

[16] Cássio M.M. Pereira and Rodrigo F. de Mello. Persistent homology for time series and spatial data clustering. *Expert Systems with Applications*, 42(15):6026–6038, 2015.

[17] Marco A. F. Pimentel, Alistair E. W. Johnson, Peter H. Charlton, Daniel Birrenkott, Peter J. Watkinson, Lionel Tarassenko, and David A. Clifton. Toward a robust estimation of respiratory rate from pulse oximeters. *IEEE Transactions on Biomedical Engineering*, 64(8):1914–1923, 2017.

[18] Marco A. F. Pimentel, Alistair E. W. Johnson, Peter H. Charlton, and David A. Clifton. BIDMC PPG and Respiration Dataset. PhysioNet, 2018.

[19] Chi Seng Pun, Si Xian Lee, and Kelin Xia. Persistent-homology-based machine learning: a survey and a comparative study. *Artificial Intelligence Review*, 55:5169–5213, 2022.

[20] James C. Robinson. The takens time-delay embedding theorem. In *Dimensions, Embeddings, and Attractors*, pages 145–159. Cambridge University Press, 2011.

[21] Vin Silva and Robert Ghrist. Coverage in sensor networks via persistent homology. *Algebraic & Geometric Topology*, 7, 04 2007.

[22] Floris Takens. Detecting strange attractors in turbulence. In David A. Rand and Lai-Sang Young, editors, *Dynamical Systems and Turbulence, Warwick 1980*, volume 898 of *Lecture Notes in Mathematics*, pages 366–381. Springer Berlin Heidelberg, Berlin, Heidelberg, 1981.

[23] M. Zechmeister and M. Kürster. The generalised Lomb-Scargle periodogram. A new formalism for the floating-mean and Keplerian periodograms. *Astronomy & Astrophysics*, 496:577–584, 2009.

[24] Afra Zomorodian and Gunnar Carlsson. Computing persistent homology. *Discrete & Computational Geometry*, 33(2):249–274, 2005.

Impact Statement

This paper presents work whose goal is to advance the field of machine learning. There are many potential societal consequences of our work, none of which we feel must be specifically highlighted here.

A Technical Assumptions

This appendix records the regularity and scale-compatibility assumptions used to control the reach of the sliding-window curve in Section 4.

Assumption A.1 (Sliding-window regularity and nondegeneracy)**.** The true sampling function $f : \mathbb{R} \to \mathbb{R}$ is $C^{2,1}$ on the relevant domain, meaning that f is C^2 and f'' is Lipschitz. There exists some $\delta, L > 0$ such that $|f'(t)| \geq \delta$ or $|f''(t)| \geq \delta$ for every relevant t, and $|f'(t)|, |f''(t)| \leq L$, $|f''(s) - f''(t)| \leq L\,|s - t|$ for every relevant s, t.

Define

$$\nu_{\delta,L} := \frac{\delta}{16} \min\left\{1, \frac{\delta}{8L}\right\}, \qquad d_{\delta,L} := \frac{\pi}{2L^2}\nu_{\delta,L}^2.$$

Assumption A.2 (Scale compatibility)**.** The separation scale d in Assumption 3.4 satisfies $d < d_{\delta,L}$.

Assumption A.1 prevents the sliding-window trajectory from locally collapsing, while Assumption A.2 ensures that the global separation condition activates before curvature alone can allow a near self-intersection.

B Simplicial complex, Filtration, and Persistent homology

All homology groups use coefficients in the same fixed field $\mathbb{F}$. In the Euclidean setting of this paper, Čech balls are taken in the ambient Euclidean space.

Given a set V, an *(abstract) simplicial complex* is a set K of subsets of V such that $\alpha \in K$ and $\beta \subset \alpha$ implies $\beta \in K$. Each set $\alpha \in K$ is called its *simplex*. The *dimension* of a simplex α is

$\dim \alpha = \text{card } \alpha - 1$, and the dimension of the simplicial complex is the maximum dimension of any of its simplices.

Assume we observe finite samples X from a metric space $(\mathbb{X}, d)$. *Čech complex* is one method for approximating the topology of the underlying space $\mathbb{X}$ by observed samples X.

Definition B.1 (Čech complex)**.** Let $X \subset \mathbb{X}$ be finite and $r > 0$. The Čech complex is the simplicial complex

$$\mathcal{C}(r) \coloneqq \{\sigma \subset X : \cap_{x \in \sigma} B_{\mathbb{X}}(x, r) \neq \emptyset\}.$$

Another common choice is the *Vietoris-Rips complex*, also referred to as *Rips complex*, where simplexes are built based on pairwise distances among its vertices.

Definition B.2 (Vietoris-Rips complex)**.** Let $X \subset \mathbb{X}$ be finite and $r > 0$. The *Vietoris-Rips complex* $\mathcal{VR}(r)$ is the simplicial complex defined as

$$\mathcal{VR}(r) \coloneqq \{\sigma \subset X : d(x_i, x_j) < 2r, \forall x_i, x_j \in \sigma\}.$$

Note that Čech complex and Vietoris-Rips complex have the following interleaving inclusion relationship

$$\mathcal{C}(r) \subset \mathcal{VR}(r) \subset \mathcal{C}(2r).$$

In particular, when $\mathbb{X} \subset \mathbb{R}^d$ is a subset of a Euclidean space of dimension d, then the constant 2 can be tightened to $\sqrt{\frac{2d}{d+1}}$ (e.g., see Theorem 2.5 in [21]):

$$\mathcal{C}(r) \subset \mathcal{VR}(r) \subset \mathcal{C}\left(\sqrt{\frac{2d}{d+1}} r\right).$$

A *filtration* $\mathcal{F}$ is a collection of subcomplexes, formally defined as follows.

Definition B.3 (Filtration)**.** A *filtration* $\mathcal{F} = \{K_a \subset K\}_{a \in \mathbb{R}}$ is a collection of subcomplexes of K such that $a \leq b$ implies that $K_a \subset K_b$.

By defining a filtration, we can approximate a point cloud at different levels of resolution. *Persistent homology* is the subsequent multiscale tool used to capture the topological features of the resulting complex K. For any filtration $\mathcal{F}$ and non-negative integer k, the persistent homology $PH_k\mathcal{F}$ is simply an ordered sequence of k-dimensional homology groups—one for every step in the filtration $\mathcal{F}$.

Definition B.4 (Persistent homology)**.** Let $\mathcal{F}$ be a filtration and let $k \in \mathbb{N}_0$. The associated k-th *persistent homology* $PH_k\mathcal{F}$ is a collection of groups $\{H_k(K_a)\}_{a \in \mathbb{R}}$ of each subcomplex K_a in $\mathcal{F}$ equipped with homomorphisms $\{\imath_k^{a,b}\}_{a \leq b}$, where $H_k(K_a)$ is the k-th dimensional homology group of K_a and $\imath_k^{a,b} \colon H_k K_a \to H_k K_b$ is the homomorphism induced by the inclusion map $\imath : K_a \hookrightarrow K_b$. If we write $H_k^{a,b} := \text{im}(\imath_k^{a,b})$, the corresponding k-th *persistent Betti numbers* are the ranks of these groups, $\beta_k^a = \text{rank } H_k^a$.

For the k-th persistent homology $PH_k\mathcal{F}$, the set of filtration levels at which a specific homology appears is always an interval $[b, d) \subset [-\infty, \infty]$. We often say that the homological feature α is born at b and dies at d. To be more precise, let γ be a class in $H_k(\mathcal{F}_b)$. We say it is born at $b \in \mathbb{R}$ if $\gamma \notin H_k^{a,b}$ for all $a < b$. Furthermore, if γ is born at b then it dies entering d if it merges with an older class as we go from $\mathcal{F}_c$ to $\mathcal{F}_d$ for any $c \in [b, d)$.

By visualizing these pairs as points in the plane, one obtains the *persistence diagram* as below.

Definition B.5 (Persistence diagram)**.** Let $\mathbb{R}^2_* \coloneqq \{(b, d) \in (\mathbb{R} \cup \infty)^2 : d > b\}$. Let $\mathcal{F}$ be a filtration and let $k \in \mathbb{N}_0$. The corresponding k-th persistence diagram $\mathcal{P}_k(\mathcal{F})$ is a finite multiset of $\mathbb{R}^2_*$, consisting of all pairs (b, d), where $[b, d)$ is the interval of filtration values for which a specific homology class appears in $PH_k\mathcal{F}$. b is called a birth time and d is called a death time.

When topological information of the underlying space is approximated by the observed points, it is often needed to compare two sets with respect to their metric structures. We first endow the space of persistence diagrams with a metric.

Definition B.6 (Bottleneck distance)**.** The bottleneck distance between two persistence diagrams $\mathcal{P}_k(\mathcal{F})$ and $\mathcal{P}_k(\mathcal{G})$ is defined by

$$d_B(\mathcal{P}_k(\mathcal{F}), \mathcal{P}_k(\mathcal{G})) = \inf_{\gamma \in \Gamma} \sup_{p \in \mathcal{P}_k(\mathcal{F})} \|p - \gamma(p)\|_\infty,$$

where the set Γ consists of all the bijections $\gamma : \mathcal{P}_k(\mathcal{F}) \cup \Delta \to \mathcal{P}_k(\mathcal{G}) \cup \Delta$, and Δ is the diagonal $\{(x, x) : x \in \mathbb{R}\} \subset \mathbb{R}^2$ with infinite multiplicity.

The *Hausdorff distance* [2, Definition 7.3.1] is defined on sets embedded in the same metric spaces. This distance measures how two sets are close to each other in the embedded metric space. When $S \subset \mathbb{X}$, we denote by S^r the r-neighborhood of a set S in a metric space, i.e. $S^r = \bigcup_{x \in S} B_{\mathbb{X}}(x, r)$.

Definition B.7 (Hausdorff distance)**.** Let $\mathbb{X}$ be a metric space, and $X, Y \subset \mathbb{X}$ be subsets. The *Hausdorff distance* between X and Y, denoted by $d_H(X, Y)$, is defined as

$$d_H(X, Y) = \inf\{r > 0 : \ X \subset Y^r \text{ and } Y \subset X^r\}.$$

Finally, one of the most crucial properties of persistent homology for practical data analysis is its robustness to noise and minor perturbations in the data. This robustness is formally guaranteed by the Stability Theorem. In essence, the theorem states that if two point clouds are geometrically close to each other, their corresponding topological signatures—represented as persistence diagrams—will also be strictly close.

Theorem B.8 (Stability of Persistence Diagrams [3])**.** *Let X and Y be finite subsets of a metric space. Let $\mathcal{P}_k(X)$ and $\mathcal{P}_k(Y)$ denote the k-dimensional persistence diagrams of the Čech or Vietoris-Rips filtrations built on X and Y, respectively. Then, the bottleneck distance between the two persistence diagrams is bounded by the Hausdorff distance between the original sets:*

$$d_B(\mathcal{P}_k(X), \mathcal{P}_k(Y)) \leq d_H(X, Y)$$

C Reach

The reach of a set, introduced by [7], is a fundamental measure of its geometric regularity. For a closed subset $A \subset \mathbb{R}^d$, its medial axis, denoted by $\text{Med}(A)$, is defined as the set of points in $\mathbb{R}^d$ that have at least two distinct nearest neighbors in A:

$$\text{Med}(A) = \left\{x \in \mathbb{R}^d \mid \exists q_1 \neq q_2 \in A, \|q_1 - x\| = \|q_2 - x\| = d(x, A)\right\}$$

where $d(x, A) = \inf_{q \in A} \|q - x\|$.

The reach of A is then defined as the minimum distance between A and its medial axis.

Definition C.1 ([7], Definition 4.1)**.** The reach of a closed subset $A \subset \mathbb{R}^d$ is defined as

$$\tau_A = \inf_{q \in A} d(q, \text{Med}(A)).$$

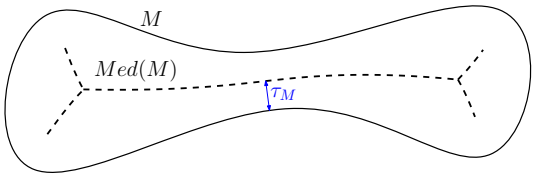


Figure 4: The medial axis and the reach [1]

A strictly positive reach serves as a standard regularity assumption in various geometric and topological contexts. In particular, it is a crucial prerequisite for homological inference and manifold reconstruction; for instance, the r-tubular neighborhood A^r deformation retracts to A for any $r < \tau_A$. In light of this, Section 4 is essentially dedicated to verifying the lower bound for the reach of the time-delay embedded image.

D Statement on Section 2.2.2

In Section 2.2.2 we addressed the results from [6]. We describe the formal statement here. Assume the observed sample $\{X_1, \cdots, X_n\}$ from a distribution P concentrated on the set $\mathbb{M}$. Fasy et al.[6] make the following assumptions.

Assumption A1. $\mathbb{M}$ is a d-dimensional compact manifold without boundary, embedded in $\mathbb{R}^D$ and $\text{reach}(\mathbb{M}) > 0$

Assumption A2. For each $x \in \mathbb{M}$, $\rho(x,t) = \frac{P(B(x,t/2))}{t^d}$ is a bounded continuous function of t, differentiable for $t \in (0, t_0)$ and right differentiable at zero. $\rho(x,t)$ is bounded above and away from zero and infinity and there exists $t_0 > 0$ and some C that

$$\sup_x \sup_{0 \le t \le t_0} \left| \frac{\partial \rho(x,t)}{\partial t} \right| \le C < \infty.$$

Then, it is possible to construct the subsample bound from the Hausdorff distance between the subsample and the original sample. Define $L(t) = \frac{1}{\binom{n}{b}} \sum_{j=1}^{\binom{n}{b}} I(d_H(S_{b,n}^j, S_n) > t)$, which is the summation of the indicator function over all possible $\binom{n}{b}$ subsamples of size b. As remarked in [6], in practice we shall use the Monte Carlo method to approximate this function. Denote the persistence diagram of $\{X_1, \cdots, X_n\}$ by $\hat{\mathcal{P}}$ and the persistence diagram of $\mathbb{M}$ by $\mathcal{P}$. For b increasing to infinity as $n \to \infty$ and satisfying $b = o(\frac{n}{\log n})$,

Theorem D.1 ([6] Theorem 3)**.** *For all large n,*

$$\mathbb{P}(d_B(\hat{\mathcal{P}}, \mathcal{P}) > c_\alpha) \le \alpha + O\left(\frac{b}{n}\right)^{1/4}.$$

For the non-periodic sliding-window support in Section 4, the support is an embedded interval rather than a manifold without boundary. We use the following one-dimensional version of the same subsampling argument. The point is that the "without boundary" assumption in [6] is used in Method I, Section 4.1, to obtain uniform covering and local-mass estimates for the Hausdorff-distance subsampling argument. In our setting the same estimates remain valid for compact embedded intervals with positive reach. Indeed, if $M = \Gamma([a,b])$ is parametrized by a regular C^2 curve and the sampling density along the parameter interval is bounded above and away from zero, then for an interior point

$$P(B(\Gamma(s), t/2)) = \frac{q(s)}{\|\Gamma'(s)\|} t + O(t^2),$$

whereas at an endpoint the local mass is one-sided and

$$P(B(\Gamma(s), t/2)) = \frac{q(s)}{2\|\Gamma'(s)\|} t + O(t^2).$$

The factor $1/2$ changes only constants, not the order or differentiability of the local mass function. Thus the Method I proof of [6] applies with modified constants once the following local mass condition is imposed.

Proposition D.2 (Subsampling for embedded curves, possibly with boundary)**.** *Let $M \subset \mathbb{R}^D$ be a compact embedded C^2 curve, either a circle or an interval, with positive reach. Let P be a probability distribution supported on M. Suppose that, for*

$$\rho(x,t) := \frac{P(B(x,t/2))}{t}, \qquad x \in M,$$

there exists $t_0 > 0$ such that $\rho(x,t)$ is bounded above and away from zero uniformly over $x \in M$ and $0 < t < t_0$, is differentiable in t for $0 < t < t_0$, and is right differentiable at $t = 0$, with

$$\sup_{x \in M} \sup_{0 \le t \le t_0} \left| \frac{\partial}{\partial t} \rho(x,t) \right| < \infty.$$

Let $S_n = \{X_1, \ldots, X_n\}$ be i.i.d. from P. For all size-b subsamples $S_{b,n}^{(j)}$ of S_n, define

$$L_b(t) = \frac{1}{\binom{n}{b}} \sum_j I\{d_H(S_{b,n}^{(j)}, S_n) > t\}, \qquad c_\alpha = 2L_b^{-1}(\alpha).$$

If $b \to \infty$ and $b = o(n/\log n)$, then, for all large n,

$$\mathbb{P}\left(d_B(\mathcal{P}(S_n), \mathcal{P}(M)) > c_\alpha\right) \le \alpha + O\left(\frac{b}{n}\right)^{1/4}.$$

E Proofs for Section 4

Claim E.1. Let $h : I \supset [u, u + m\tau] \to \mathbb{R}$ be C^1. Suppose that for all $s \in I$,

$$|h(s)| \ge \eta \qquad \text{or} \qquad |h'(s)| \ge \eta,$$

and for all $r, s \in I$,

$$|h'(s)| \le L \qquad \text{and} \qquad |h'(s) - h'(r)| \le L|s - r|.$$

If

$$m\tau \ge \frac{\eta}{2L}, \qquad \tau \le \frac{\eta}{8L},$$

then

$$\|SW_{m,\tau}h(u)\|_2 \ge \sqrt{m+1}\,\frac{\eta}{16}\min\left\{1, \frac{\eta}{8L}\right\}.$$

Proof. Let $\ell = \eta/(8L)$ and write the grid as

$$u_j = u + j\tau, \quad 0 \le j \le m.$$

We first prove the bound on any consecutive block of grid points whose time span is between $\eta/(2L)$ and $2\eta/L$. Let $I_B = [u_p, u_q]$ be such a block and let $c = (u_p + u_q)/2$.

If $|h(c)| \ge \eta$, then $|h'| \le L$ gives $|h(s)| \ge \eta/2$ whenever $|s - c| \le \eta/(2L)$. Since the block has half-span at least $\eta/(4L)$, the intersection of this interval with I_B contains an interval of length at least ℓ.

If $|h'(c)| \ge \eta$, then the Lipschitz bound on h' implies that h' has constant sign and

$$|h'(s)| \ge \frac{\eta}{2} \qquad \text{whenever } |s - c| \le \frac{\eta}{2L}.$$

If $h(c)h'(c) \ge 0$ we take $[c + \ell, c + 2\ell]$, while if $h(c)h'(c) < 0$ we take $[c - 2\ell, c - \ell]$. In either case the interval is contained in I_B, and on it

$$|h(s)| \ge \frac{\eta^2}{16L}.$$

Thus in either case there is a subinterval of I_B of length ℓ on which

$$|h(s)| \ge \eta_* := \min\left\{\frac{\eta}{2}, \frac{\eta^2}{16L}\right\}.$$

Because $\tau \le \ell$, this subinterval contains at least $\ell/(2\tau)$ grid points: if $\tau \le \ell/2$, then the count is at least $\lfloor \ell/\tau \rfloor \ge \ell/\tau - 1 \ge \ell/(2\tau)$; if $\ell/2 < \tau \le \ell$, the subinterval contains at least one grid point and $1 \ge \ell/(2\tau)$. The full block contains at most $2\eta/(L\tau) + 1 = 16\ell/\tau + 1 \le 17\ell/\tau$ grid points. Hence at least a $1/64$ fraction of the grid points in the block have $|h(u_j)| \ge \eta_*$. If the block contains N_B grid points, then

$$\sum_{u_j \in I_B} |h(u_j)|^2 \ge \frac{N_B}{64}\eta_*^2 \ge N_B\left(\frac{\eta}{16}\min\left\{1, \frac{\eta}{8L}\right\}\right)^2.$$

It remains only to cover the whole grid by such blocks. Greedily split the grid into consecutive blocks with time span at least $\eta/(2L)$ and stop each block as soon as this is achieved. Each completed block has span at most $\eta/(2L) + \tau \le 5\eta/(8L)$. If the final leftover block has shorter span, merge it with the preceding block; the merged block has span at most $2\eta/L$. Applying the preceding estimate block by block and summing the squares gives the claimed bound on the full sliding window. □

Claim E.2. Suppose $|f'|, |f''| \le L$ and

$$m\tau \ge \frac{\epsilon}{4L}, \qquad \tau \le \frac{\epsilon}{16L}.$$

(a) If f is (d, ϵ)-non-periodic on $J_{m,\tau}$, then for any $t_1, t_2 \in \mathbb{T}$ with $|t_1 - t_2| \ge d$,

$$\|SW_{m,\tau} f(t_1) - SW_{m,\tau} f(t_2)\|_2 \ge \sqrt{m+1}\, \frac{\epsilon}{16} \min\left\{1, \frac{\epsilon}{16L}\right\}.$$

(b) If f is (Ξ, d, ϵ)-periodic, then for any $t_1, t_2 \in \mathbb{R}$ with $\min_{n\in\mathbb{Z}} |t_1 - t_2 + n\Xi| \ge d$,

$$\|SW_{m,\tau} f(t_1) - SW_{m,\tau} f(t_2)\|_2 \ge \sqrt{m+1}\, \frac{\epsilon}{16} \min\left\{1, \frac{\epsilon}{16L}\right\}.$$

Proof. Set $g(u) = f(t_1 + u) - f(t_2 + u)$. In the non-periodic case, $|t_1 - t_2| \ge d$ implies that, for every $u \in [0, m\tau]$,

$$|g(u)| \ge \epsilon \quad \text{or} \quad |g'(u)| \ge \epsilon.$$

In the periodic case, $\min_{n\in\mathbb{Z}} |(t_1 + u) - (t_2 + u) + n\Xi| = \min_{n\in\mathbb{Z}} |t_1 - t_2 + n\Xi|$, so again, for every $u \in [0, m\tau]$,

$$|g(u)| \ge \epsilon \qquad \text{or} \qquad |g'(u)| \ge \epsilon.$$

Also $|g'| \le 2L$, and g' is $2L$-Lipschitz. Applying Claim E.1 to g gives

$$\|SW_{m,\tau} g(0)\|_2 \ge \sqrt{m+1}\, \frac{\epsilon}{16} \min\left\{1, \frac{\epsilon}{16L}\right\}.$$

Since $SW_{m,\tau} g(0) = SW_{m,\tau} f(t_1) - SW_{m,\tau} f(t_2)$, the claim follows. □

Claim E.3. Suppose Assumption A.1 holds. If

$$m\tau \ge \frac{\delta}{2L}, \qquad \tau \le \frac{\delta}{8L},$$

then

$$\|SW f'_{m,\tau}(t)\|_2 \ge \sqrt{m+1}\, \nu_{\delta,L} \qquad \text{for all relevant } t.$$

where $\nu_{\delta,L} = \frac{\delta}{16} \min\left\{1, \frac{\delta}{8L}\right\}$ as defined in Assumption A.1.

Proof. Since $(SW_{m,\tau} f(t))' = SW_{m,\tau}(f')(t)$, apply Claim E.1 to $h = f'$ with $\eta = \delta$. The assumptions needed for that claim are exactly $|f'| \ge \delta$ or $|f''| \ge \delta$, $|f''| \le L$, and f'' is L-Lipschitz. □

Claim E.4. Let γ be the arc-length parametrization of $SW_{m,\tau} f$. Under the assumptions of Claim E.3,

$$\|\gamma''\|_2 \le \frac{2L}{\sqrt{m+1}\nu_{\delta,L}^2}.$$

Proof. For any regular parametrized curve $\Gamma(t)$, its arc-length parametrization satisfies

$$\|\gamma''\|_2 \le \frac{2\|\Gamma''(t)\|_2}{\|\Gamma'(t)\|_2^2}.$$

Here

$$\|\Gamma''_{m,\tau}(t)\|_2 = \left(\sum_{j=0}^{m} f''(t + j\tau)^2\right)^{1/2} \le L\sqrt{m+1}.$$

Together with Claim E.3, this gives

$$\|\gamma''\|_2 \le \frac{2L\sqrt{m+1}}{(m+1)\nu_{\delta,L}^2} = \frac{2L}{\sqrt{m+1}\nu_{\delta,L}^2}.$$

□

For a curve $\gamma : I \subset \mathbb{R} \to \mathbb{R}^{m+1}$, let $d_\gamma : I \times I \to \mathbb{R}$ be geodesic distance on γ. That is, for $s_1, s_2 \in I$,

$$d_\gamma(s_1, s_2) = \inf_{s_1'=s_1, s_2', \cdots, s_{n-1}', s_n'=s_2} \sum_{i=1}^{n-1} \left\|\gamma(s_{i+1}') - \gamma(s_i')\right\|_2 .$$

If we assume that γ is C^1 and unit speed, i.e., $\|\gamma'(s)\|_2 = 1$ for all $s \in I$, then this is equivalent to that

$$d_\gamma(s_1, s_2) = \min_{\gamma(s_1')=\gamma(s_1), \gamma(s_2')=\gamma(s_2)} |s_1' - s_2'| .$$

Lemma E.5 (Short-arc chord bound)**.** *Let γ be a unit-speed C^2 curve with $\|\gamma''\| \le 1/R$. If $0 < d_\gamma(s_1, s_2) < \pi R$, then*

$$\|\gamma(s_1) - \gamma(s_2)\|_2 \ge 2R \sin\left(\frac{d_\gamma(s_1, s_2)}{2R}\right) > 0.$$

Proof. This is the standard higher-dimensional Schur comparison for curves with curvature bounded by $1/R$. Among unit-speed curves with curvature at most $1/R$ and fixed intrinsic length less than πR, the circular arc of radius R has the smallest endpoint chord. □

Lemma E.6 (Curve reach criterion)**.** *Let γ be a unit-speed C^2 embedded compact curve, either an interval or a circle, in Euclidean space. Let d_γ denote intrinsic distance, with circular intrinsic distance in the closed case. Suppose that*

$$\|\gamma''\| \le \frac{1}{R}$$

and that every pair with $d_\gamma(s_1, s_2) \ge \pi R$ satisfies

$$\|\gamma(s_1) - \gamma(s_2)\|_2 \ge 2R'.$$

Then

$$\operatorname{reach}(\gamma) \ge \min\{R, R'\}.$$

Proof. We use the standard reach characterization for C^2 curves: the reach is the minimum of the reciprocal maximal curvature and one half of the minimal Euclidean distance over doubly critical pairs; see, for example, [1]. For compact arcs with boundary, the same criterion is applied with one-sided tangent cones at the endpoints; equivalently, one may extend the C^2 arc slightly past its endpoints and then restrict the resulting tubular neighborhood to the original compact arc. Endpoints therefore do not introduce a smaller reach scale than the curvature and global normal-chord scales below. The curvature condition gives the first lower bound R. A doubly critical pair with intrinsic separation strictly smaller than πR is excluded by the comparison theorem for curves with curvature bounded by $1/R$; the model case is a circle of radius R, where the first normal chord occurs at intrinsic separation πR. Hence any doubly critical pair has intrinsic separation at least πR, and therefore has Euclidean distance at least $2R'$. The reach is thus at least $\min\{R, R'\}$. □

Lemma E.7. *Under the assumptions of Theorem 4.1, $\Gamma_{m,\tau}(\mathbb{T})$ is an embedded interval and*

$$\operatorname{reach}(\Gamma_{m,\tau}(\mathbb{T})) \ge a_{m,\tau}.$$

Proof. Let $\gamma : [0, S_{\max}] \to \mathbb{R}^{m+1}$ be the arc-length parametrization of $SW_{m,\tau} f(\mathbb{T})$.

Claim E.4 gives the curvature condition with

$$R_0 = \sqrt{m+1}\, \frac{\nu_{\delta,L}^2}{2L}.$$

Also $\|SW_{m,\tau} f'(t)\|_2 \le L\sqrt{m+1}$. We first show that the curve is embedded. If $0 < |s_1 - s_2| < \pi R_0$, then Lemma E.5 gives

$$\|\gamma(s_1) - \gamma(s_2)\|_2 > 0.$$

If $|s_1 - s_2| \ge \pi R_0$, then their original time parameters satisfy

$$|t(s_1) - t(s_2)| \ge \frac{|s_1 - s_2|}{L\sqrt{m+1}} \ge \frac{\pi \nu_{\delta,L}^2}{2L^2} = d_{\delta,L} > d.$$

By Claim E.2,

$$\|\gamma(s_1) - \gamma(s_2)\|_2 \geq 2\sqrt{m+1}\,\frac{\epsilon}{32}\min\left\{1, \frac{\epsilon}{16L}\right\}.$$

Thus distinct arc-length parameters have distinct images.

Lemma E.6 gives

$$\operatorname{reach}(\Gamma_{m,\tau}(\mathbb{T})) \geq \sqrt{m+1}\min\left\{\frac{\epsilon}{32}\min\left\{1, \frac{\epsilon}{16L}\right\}, \frac{\nu_{\delta,L}^2}{2L}\right\} = a_{m,\tau}.$$

Hence the image is an embedded copy of the interval $\mathbb{T}$. □

Proof of Theorem 4.1. By Lemma E.7, the support is an embedded interval with reach at least $a_{m,\tau}$. For $0 < r < a_{m,\tau}$, the nearest-point projection onto the support is uniquely defined on the offset. The straight-line homotopy from each point to its nearest point is a deformation retraction. Hence the offset deformation retracts to an interval and is contractible. □

Lemma E.8. *Under the assumptions of Theorem 4.3,* $\Gamma_{m,\tau}([0,\Xi])$ *is an embedded circle and*

$$\operatorname{reach}(\Gamma_{m,\tau}([0,\Xi])) \geq a_{m,\tau}.$$

Proof. Because f is Ξ-periodic, $\Gamma_{m,\tau}(0) = \Gamma_{m,\tau}(\Xi)$. Let $\gamma : [0, S_{\max}] \to \mathbb{R}^{m+1}$ be the arc-length parametrization over one period. The curvature bound is the same as in the non-periodic case, with $R_0 = \sqrt{m+1}\,\frac{\nu_{\delta,L}^2}{2L}$.

We first prove injectivity on $\mathbb{R}/\Xi\mathbb{Z}$. If the circular arc-length distance between two parameters is in $(0, \pi R_0)$, then Lemma E.5 gives positive Euclidean distance. If the circular arc-length distance is at least πR_0, then the corresponding circular time distance satisfies

$$d_\Xi(t(s_1), t(s_2)) \geq \frac{\pi R_0}{L\sqrt{m+1}} = d_{\delta,L} > d.$$

Claim E.2 then gives Euclidean separation at least $2\sqrt{m+1}\cdot(\epsilon/32)\cdot\min\{1, (\epsilon/16L)\}$. Hence the map from $\mathbb{R}/\Xi\mathbb{Z}$ to its image is injective, so the image is an embedded circle. Lemma E.6 gives reach at least $a_{m,\tau}$. □

Proof of Theorem 4.3. Since $T_{\max} - T_{\min} \geq \Xi$ and f is Ξ-periodic, $\Gamma_{m,\tau}$ is Ξ-periodic in its time parameter. Hence the image of any observation interval of length at least Ξ equals the image over one period, so $\Gamma_{m,\tau}(\mathbb{T}) = \Gamma_{m,\tau}([0,\Xi])$. By Lemma E.8, this support is an embedded circle with reach at least $a_{m,\tau}$. For $0 < r < a_{m,\tau}$, the nearest-point projection gives a deformation retraction of the offset onto the circle. □

F Proof of Theorem 5.1

To prove Theorem 5.1 we verify the assumptions required by Proposition D.2. This proposition covers both the periodic circle case and the non-periodic interval case.

Proof of Proposition D.2. The proof is the Method I subsampling argument of [6], with the "without boundary" manifold assumption replaced by the stated local mass condition and the one-dimensional packing estimate for compact curves. The stability reduction in Section 4 of [6] only uses the inequality between bottleneck distance and Hausdorff distance for distance functions, and therefore does not depend on whether the support has boundary. The boundary issue enters only through the small-ball estimates used in the subsampling step; for a compact embedded interval these estimates are one-sided at the two endpoints and hence differ from the circle or interior case only by constants. Indeed, because M is a compact C^2 embedded curve with positive reach, there are constants $0 < C_1 < C_2 < \infty$ such that, for all sufficiently small $r > 0$, every maximal r-packing and every minimal r-covering of M have cardinality between $C_1 r^{-1}$ and $C_2 r^{-1}$. This follows from finite curve length and the uniform chord–arc control implied by the positive reach.

The local mass condition gives

$$\rho := \liminf_{t\downarrow 0} \inf_{x\in M} \frac{P(B(x,t/2))}{t} > 0.$$

Consequently, for $t_n = C\log n/n$ with C large enough,

$$\mathbb{P}\{d_H(S_n, M) > t_n\} = o(1).$$

On the event $\{d_H(S_n, M) < t_n\}$, the same packing argument used in the proof of Theorem 3 of [6] gives

$$d_H(S_n, M) \le d_H(S_{b,n}^{(j)}, S_n)$$

simultaneously for every size-b subsample, since $b = o(n/\log n)$ while the number of t_n-packing cells is of order $n/\log n$. The remainder of the subsampling proof compares the empirical subsample tail with the corresponding i.i.d. b-sample tail, giving the same $O((b/n)^{1/4})$ error term. Finally, bottleneck stability for distance functions gives

$$d_B(\mathcal{P}(S_n), \mathcal{P}(M)) \le d_H(S_n, M),$$

which yields the claimed confidence bound. □

Lemma F.1. *Under the assumptions of Theorem 5.1, $M_{m,\tau}$ satisfies:*

(1) $\mathrm{reach}(M_{m,\tau}) > 0$

(2) For

$$\rho(x,t) = \frac{P(B(x,t/2))}{t}, \quad x \in M_{m,\tau}$$

there exists $t_0 > 0$ such that $\rho(x,t)$ is differentiable for $t \in (0,t_0)$ and right differentiable at zero. Moreover, $\rho(x,t)$ is bounded above and away from zero, and there exists $C < \infty$ such that

$$\sup_x \sup_{0\le t\le t_0} \left|\frac{\partial \rho(x,t)}{\partial t}\right| \le C < \infty$$

Proof. (1) By our results in Section 4, for the functions satisfying Assumption 3.4, A.1, and A.2, and m, τ satisfying inequalities in Theorem 5.1, $M_{m,\tau}$ is a manifold with positive reach.

(2) Let $\Gamma(t) = SW_{m,\tau}f(t)$. Claim E.3 and Assumption A.1 give,

$$\nu_{\delta,L}\sqrt{m+1} \le \|\Gamma'(t)\|_2 \le L\sqrt{m+1}, \qquad \|\Gamma''(t)\|_2 \le L\sqrt{m+1}.$$

Choose $t_0 <$ reach$(M_{m,\tau})$ small enough that every $B(\Gamma(s), t/2) \cap M_{m,\tau}$ is a single connected arc for $0 < t < t_0$, with a one-sided arc allowed when $\Gamma(s)$ is an endpoint in the non-periodic interval case. For $x = \Gamma(s)$, write

$$A_{s,t} := \{u : \Gamma(u) \in B(x, t/2)\}.$$

In the periodic case the parameter u is taken on $\mathbb{T}/\Xi\mathbb{Z}$; in the non-periodic case it is taken on the interval $\mathbb{T} = [T_{\min}, T_{\max}]$.

Then the endpoints of these arcs are C^1 functions of t by the inverse function theorem applied to $u \mapsto \|\Gamma(s+u) - \Gamma(s)\|_2$ because the speed is bounded below and the curve has positive reach.

The lower and upper speed bounds imply that the length of $A_{s,t}$ is comparable to t, uniformly in s and $t < t_0$. Since the sampling density is assumed to be bounded above and below by Assumption 3.1, the P-mass of the set $A_{s,t}$ is comparable to t. Thus $\rho(x,t)$ is bounded above and away from zero.

It remains to control the derivative of ρ. Each local mass has the form

$$P(B(\Gamma(s), t/2)) = \int_{a_-(s,t)}^{a_+(s,t)} q(u)du$$

where we collapse the sampling density to $\mathbb{T}/\Xi\mathbb{Z}$ in the periodic case. At an interior point of the interval, and everywhere in the periodic quotient, the endpoint functions $a_\pm(s,t)$ are C^1 in t for $0 \le t \le t_0$, and the C^2 curve bound gives the uniform expansions

$$a_\pm(s,t) = s \pm \frac{t}{2\|\Gamma'(s)\|} + O(t^2), \qquad \partial_t a_\pm(s,t) = \pm\frac{1}{2\|\Gamma'(s)\|} + O(t).$$

Since q is Lipschitz by Assumption 3.1,

$$P(B(\Gamma(s), t/2)) = \frac{q(s)}{\|\Gamma'(s)\|}t + O(t^2), \qquad \partial_t P(B(\Gamma(s), t/2)) = \frac{q(s)}{\|\Gamma'(s)\|} + O(t).$$

At the left endpoint $s = T_{\min}$, the lower endpoint is fixed and

$$a_+(s,t) = s + \frac{t}{2\|\Gamma'(s)\|} + O(t^2), \qquad \partial_t a_+(s,t) = \frac{1}{2\|\Gamma'(s)\|} + O(t),$$

so

$$P(B(\Gamma(s), t/2)) = \frac{q(s)}{2\|\Gamma'(s)\|}t + O(t^2).$$

The right endpoint is identical with the one-sided interval reversed. Thus the same boundedness and differentiability conclusions hold at the boundary, with the leading local mass constant divided by two. Therefore,

$$\frac{\partial}{\partial t}\rho(x,t) = \frac{\partial_t P(B(\Gamma(s), t/2))}{t} - \frac{P(B(\Gamma(s), t/2))}{t^2}$$

is uniformly bounded for $0 < t < t_0$ and has a one-sided limit at $t = 0$. □

Proof of Lemma 5.2. Let $X_n = X_{m,\tau}$ and $X_b^{(I)}$ denote a uniformly chosen b-subsample of X_n, where the index set I is chosen independently of the sample after conditioning on X_n. Let

$$P(r) = P_I(d_H(X_n, X_b^{(I)}) > r \mid X_n).$$

Taking expectations over the original sample gives

$$\mathbb{E}(P(r)) = P(d_H(X_n, X_b^{(I)}) > r).$$

Although X_b^I is not independent of X_n, its unconditional joint distribution is the same as that of b i.i.d. samples from the pushforward distribution on $M_{m,\tau}$.

Fix $r > 0$. Since $M_{m,\tau}$ is compact and sampling density is bounded below, there exist finitely many balls $B(y_1, r/4), \cdots, B(y_K, r/4)$ covering $M_{m,\tau}$ and $p_r > 0$ such that $P(B(y_k, r/4)) \ge p_r$ for every k. For N i.i.d. samples Z_N from P,

$$P(d_H(Z_N, M_{m,\tau}) > r/2) \le K(1 - p_r)^N \to 0.$$

Therefore by triangle inequality,

$$P(d_H(X_n, X_b^{(I)}) > r) \to 0$$

as $n, b \to \infty$.

Thus $\mathbb{E}(P(r)) \to 0$ for every $r > 0$. By Markov's inequality, $P(r) \xrightarrow{P} 0$, and therefore

$$P(c_\alpha > \eta) \le P(P(\eta/2) > \alpha) \to 0.$$

Hence $c_\alpha \xrightarrow{P} 0$. □

Proof of Theorem 5.1. The sampled sliding-window points $X_{m,\tau}$ are sampled from the support $M_{m,\tau}$. By Lemma F.1, the support satisfies the regularity assumptions needed for Proposition D.2. Therefore, by the definition of c_α,

$$P\left(d_B(\mathcal{P}(X_{m,\tau}), \mathcal{P}(M_{m,\tau})) > c_\alpha\right) \le \alpha + O\left(\frac{b}{n}\right)^{1/4}.$$

□

G Proof of Section 5.2

To prove Theorem 5.3 first check the error term based on the assumption 3.5.

Lemma G.1. *Under the assumption 3.5, for the linear interpolation $\hat{f_N}$,*

$$\left\|\hat{f_N} - f\right\|_\infty \leq \eta_N$$

Consequently,

$$d_H(\hat{X}_{m,\tau}, X_{m,\tau}) \leq \zeta_{m,N}, \qquad d_H(\hat{X}^{(j)}_{m,\tau}, X^{(j)}_{m,\tau}) \leq \zeta_{m,N}$$

for every subsample, and

$$d_B(\mathcal{P}(\hat{X}_{m,\tau}), \mathcal{P}(X_{m,\tau})) \leq \zeta_{m,N}$$

Proof. The linear interpolation $\hat{f_N}$ of the sample (s_i, Y_i) is

$$\begin{aligned}\hat{f_N}(t) &= \frac{s_{i+1}-t}{s_{i+1}-s_i}Y_i + \frac{t-s_i}{s_{i+1}-s_i}Y_{i+1} \\ &= \frac{s_{i+1}-t}{s_{i+1}-s_i}f(s_i) + \frac{t-s_i}{s_{i+1}-s_i}f(s_{i+1}) + \frac{(s_{i+1}-t)\xi_i + (t-s_i)\xi_{i+1}}{s_{i+1}-s_i}.\end{aligned}$$

For the error $|\hat{f_N}(t) - f(t)|$, we get an error term from each interpolation and noise.

The noise error is bounded by assumption 3.5

$$\left|\frac{(s_{i+1}-t)\xi_i + (t-s_i)\xi_{i+1}}{s_{i+1}-s_i}\right| \leq \max_j |\xi_j| \leq \sigma_N$$

For the interpolation error, Taylor series gives the following

For some $s_i < r_i < t < r_{i+1} < s_{i+1}$,

$$f(s_i) = f(t) + f'(t)(s_i - t) + \frac{f''(r_i)}{2}(s_i - t)^2$$

and

$$f(s_{i+1}) = f(t) + f'(t)(s_{i+1} - t) + \frac{f''(r_{i+1})}{2}(s_{i+1} - t)^2$$

$$\begin{aligned}&\left|\frac{s_{i+1}-t}{s_{i+1}-s_i}f(s_i) + \frac{t-s_i}{s_{i+1}-s_i}f(s_{i+1}) - f(t)\right| \\ &= \left|\left(\frac{s_{i+1}-t}{s_{i+1}-s_i} + \frac{t-s_i}{s_{i+1}-s_i} - 1\right)f(t)\right. \\ &+ \left(\frac{(s_{i+1}-t)(s_i-t) + (t-s_i)(s_{i+1}-t)}{s_{i+1}-s_i}\right)f'(t) \\ &\left.+ \frac{(s_{i+1}-t)f''(r_i)(t-s_i)^2 + (t-s_i)f''(r_{i+1})(s_{i+1}-t)^2}{2(s_{i+1}-s_i)}\right| \\ &= \left|\frac{(s_{i+1}-t)(t-s_i)}{2}\frac{f''(r_i)(t-s_i) + f''(r_{i+1})(s_{i+1}-t)}{s_{i+1}-s_i}\right| \\ &\leq \left|\frac{(s_{i+1}-t)(t-s_i)}{2}\right| \max_a |f''(a)|\end{aligned}$$

Under the assumption 3.5, we know that $\Delta_N = \max_{0\leq i<N}(s_{i+1} - s_i)$ and $|f''(t)| \leq L_{int}$ for all $t \in J_{m,\tau}$. Hence,

$$\|\hat{f_N} - f\|_\infty \leq \sup_{t\in J_{m,\tau}} |\hat{f_N}(t) - f(t)| \leq \frac{\Delta_N^2}{8}L_{\text{int}} + \sigma_N = \eta_N$$

gives all $m+1$ components bounded by η_N, so

$$\|SW_{m,\tau}\hat{f_N}(t) - SW_{m,\tau}f(t)\|_2 \leq \sqrt{m+1}\eta_N = \zeta_{m,N}.$$

Since each point is perturbed by at most $\zeta_{m,N}$, the Hausdorff distance satisfies

$$d_H(\hat{X}_{m,\tau}, X_{m,\tau}) \le \zeta_{m,N}, \qquad d_H(\hat{X}^{(j)}_{m,\tau}, X^{(j)}_{m,\tau}) \le \zeta_{m,N}.$$

Also, by stability theorem, the bottleneck distance is bounded by

$$d_B(\mathcal{P}(\hat{X}_{m,\tau}), \mathcal{P}(X_{m,\tau})) \le d_H(\hat{X}_{m,\tau}, X_{m,\tau}) \le \zeta_{m,N}.$$

□

Lemma G.2. *Under the assumption 3.5,*

$$c_\alpha \le \hat{c_\alpha} + 4\zeta_{m,N}, \qquad \hat{c_\alpha} \le c_\alpha + 4\zeta_{m,N}.$$

Proof. By the triangle inequality, for every subsample index j,

$$\begin{aligned}&\left|d_H(\hat{X}_{m,\tau}, \hat{X}^{(j)}_{m,\tau,b}) - d_H(X_{m,\tau}, X^{(j)}_{m,\tau,b})\right| \\ &\le d_H(\hat{X}_{m,\tau}, X_{m,\tau}) + d_H(\hat{X}^{(j)}_{m,\tau,b}, X^{(j)}_{m,\tau,b}) \le 2\zeta_{m,N}.\end{aligned}$$

For fixed $t > 0$, $d_H(\hat{X}_{m,\tau}, \hat{X}^{(j)}_{m,\tau,b}) > t + 2\zeta_{m,N}$ implies that $d_H(X_{m,\tau}, X^{(j)}_{m,\tau,b}) > t$. Hence,

$$I(d_H(\hat{X}_{m,\tau}, \hat{X}^{(j)}_{m,\tau,b}) > t + 2\zeta_{m,N}) \le I(d_H(X_{m,\tau}, X^{(j)}_{m,\tau,b}) > t)$$

Taking average over the subsample index j, we obtain a bound for shifted $\hat{L}$ as

$$\hat{L}_{m,\tau}(t + 2\zeta_{m,N}) \le L_{m,\tau}(t),$$

Similarly, we can say $d_H(X_{m,\tau}, X^{(j)}_{m,\tau,b}) > t + 2\zeta_{m,N}$ implies $d_H(\hat{X}_{m,\tau}, \hat{X}^{(j)}_{m,\tau,b}) > t$, so

$$L_{m,\tau}(t + 2\zeta_{m,N}) \le \hat{L}_{m,\tau}(t).$$

Let $\delta = 2\zeta_{m,N}$ and let $\epsilon > 0$. By the generalized inverse convention, there exists $s \le \hat{L}^{-1}_{m,\tau}(\alpha) + \epsilon$ such that $\hat{L}_{m,\tau}(s) \le \alpha$. Hence $L_{m,\tau}(s + \delta) \le \alpha$, and therefore

$$L^{-1}_{m,\tau}(\alpha) \le \hat{L}^{-1}_{m,\tau}(\alpha) + \epsilon + \delta.$$

Sending $\epsilon \to 0$ gives

$$L^{-1}_{m,\tau}(\alpha) \le \hat{L}^{-1}_{m,\tau}(\alpha) + \delta.$$

The reverse inequality follows identically from the other shifted tail bound: for every $\epsilon > 0$, choose $s \le L^{-1}_{m,\tau}(\alpha) + \epsilon$ such that $L_{m,\tau}(s) \le \alpha$, then $\hat{L}_{m,\tau}(s + \delta) \le \alpha$. Thus

$$\hat{L}^{-1}_{m,\tau}(\alpha) \le L^{-1}_{m,\tau}(\alpha) + \delta.$$

It follows that $|\hat{L}^{-1}_{m,\tau}(\alpha) - L^{-1}_{m,\tau}(\alpha)| \le \delta = 2\zeta_{m,N}$. Thus, $|\hat{c}_\alpha/2 - c_\alpha/2| \le 2\zeta_{m,N}$ i.e. $|\hat{c}_\alpha - c_\alpha| \le 4\zeta_{m,N}$.

□

Proof of Theorem 5.3. Let

$$E_\alpha := \{d_B(\mathcal{P}(X_{m,\tau}), \mathcal{P}(M_{m,\tau})) \le c_\alpha\}.$$

By Theorem 5.1,

$$P(E^c_\alpha) \le \alpha + O\left(\frac{b}{n}\right)^{1/4}.$$

On E_α, Lemmas G.1 and G.2 give

$$\begin{aligned}d_B(\mathcal{P}(\hat{X}_{m,\tau}), \mathcal{P}(M_{m,\tau})) &\le d_B(\mathcal{P}(\hat{X}_{m,\tau}), \mathcal{P}(X_{m,\tau})) + d_B(\mathcal{P}(X_{m,\tau}), \mathcal{P}(M_{m,\tau})) \\ &\le \zeta_{m,N} + c_\alpha \\ &\le \hat{c}_\alpha + 5\zeta_{m,N} = \hat{c}^{\text{err}}_\alpha.\end{aligned}$$

Thus the failure probability is bounded by $P(E^c_\alpha)$.

□

Proof of Corollary 5.4. By Lemma G.2,

$$\hat{c}_\alpha^{\mathrm{err}} = \hat{c}_\alpha + 5\zeta_{m,N} \le c_\alpha + 9\zeta_{m,N}.$$

Lemma 5.2 gives $c_\alpha \to 0$ in probability, and $\zeta_{m,N} \to 0$ by assumption. Hence $\hat{c}_\alpha^{\mathrm{err}} \to 0$ in probability. □

H Proofs for Section 6

Proof of Theorem 6.2. Let

$$E_\alpha := \{d_B(\mathcal{P}(Y), \mathcal{P}(M_{m,\tau})) \le r_\alpha\}.$$

In the exact case, Theorem 5.1 gives $P(E_\alpha^c) \le \alpha + O((b/n)^{1/4})$; in the bounded-interpolation case, the same bound follows from Theorem 5.3.

Under H_0, suppose $E_\alpha \cap A_\alpha^r$ holds and $p = (b_p, d_p) \in \mathcal{P}(Y) \cap R(r_\alpha)$. Then $b_p \le r_\alpha$ and $d_p - b_p > 2r_\alpha$. Corollary 4.2 implies that every off-diagonal support point $q = (b_q, d_q) \in \mathcal{P}(M_{m,\tau})$ satisfies $b_q \ge a_{m,\tau} > 4r_\alpha$, so $b_q - b_p > 3r_\alpha$. Thus p cannot be matched to any off-diagonal support point within r_α. It also cannot be matched to the diagonal within r_α, since $d_\infty(p, \Delta) = (d_p - b_p)/2 > r_\alpha$. This contradicts E_α. Hence

$$\{\psi(Y) = 1, A_\alpha^r\} \subseteq E_\alpha^c$$

under H_0.

Under H_1, Corollary 4.4 gives a support-level point $q_0 = (0, D)$ with $D \ge a_{m,\tau}$. On A_α^r, $D \ge a_{m,\tau} > 4r_\alpha$, so q_0 is farther than r_α from the diagonal. Hence, on E_α, it must be matched to some $p = (b_p, d_p) \in \mathcal{P}(Y)$ with $|b_p| \le r_\alpha$ and $|d_p - D| \le r_\alpha$. Since births are nonnegative for the distance filtration, $b_p \le r_\alpha$, and

$$d_p - b_p \ge D - 2r_\alpha \ge a_{m,\tau} - 2r_\alpha > 2r_\alpha.$$

Thus $p \in R(r_\alpha)$ and $\psi(Y) = 1$. Therefore

$$\{\psi(Y) = 0, A_\alpha^r\} \subseteq E_\alpha^c$$

under H_1. The two conditional bounds follow from the bound on $P(E_\alpha^c)$, and the unconditional consequences follow by decomposing according to A_α^r and its complement. □

Proof of Corollary 6.3. For the exact rule, Lemma 5.2 gives $c_\alpha \xrightarrow{P} 0$. For the bounded-interpolation rule, Corollary 5.4 gives $\hat{c}_\alpha^{\mathrm{err}} \xrightarrow{P} 0$ whenever $\zeta_{m,N} \to 0$. Since m, τ are fixed, the topological results give $a_{m,\tau} > 0$, and therefore $P((A_\alpha^r)^c) = P(a_{m,\tau} \le 4r_\alpha) \to 0$ in either case. Combining this with Theorem 6.2 gives the stated bounds. □

Proof of Corollary 6.4. Under H_0, on the event $\bigcap_{l=1}^K A_{\alpha/K,l}^r$, the aggregate test rejects only if at least one individual test rejects. The union bound and Theorem 6.2 applied at level α/K give

$$\begin{aligned} P\left(\Phi_{\mathcal{G}} = 1, \bigcap_{l=1}^K A_{\alpha/K,l}^r \mid H_0\right) &\le \sum_{l=1}^K P(\psi_l = 1, A_{\alpha/K,l}^r \mid H_0) \\ &\le \sum_{l=1}^K \left\{\frac{\alpha}{K} + O\left(\frac{b}{n}\right)^{1/4}\right\} \\ &= \alpha + O\left(K\left(\frac{b}{n}\right)^{1/4}\right). \end{aligned}$$

Under H_1, if candidate l_* satisfies the periodic assumptions and the aggregate rule fails to reject, then the individual rule for l_* also fails to reject. Hence

$$\{\Phi_{\mathcal{G}} = 0, A_{\alpha/K,l_*}^r\} \subseteq \{\psi_{l_*} = 0, A_{\alpha/K,l_*}^r\}.$$

Theorem 6.2 applied at level α/K gives

$$P\left(\Phi_{\mathcal{G}} = 0, A^{r}_{\alpha/K,l_*} \mid H_1\right) \leq \frac{\alpha}{K} + O\left(\frac{b}{n}\right)^{1/4}.$$

The final asymptotic statement follows when the corresponding separation events have probability tending to one. □

I Additional explanation for Section 7

The synthetic code supports several scalar error models. For the bounded-error experiments used with the robust TDA rule, the observed scalar values are

$$Y_i = f(s_i) + \xi_i, \qquad |\xi_i| \leq \sigma_N.$$

We implement this either as bounded uniform noise

$$\xi_i \sim \mathrm{Unif}[-\sigma_N, \sigma_N]$$

or as a truncated Gaussian supported on the same interval. These are the cases where the deterministic scalar error bound σ_N is available. The interpolation theorem also requires a smoothness constant for the underlying signal; in the simulations below we use a dense-grid numerical estimate of this constant when forming the reported error-inflated radius.

I.1 Synthetic design and result generation

The synthetic experiments are performed on the interval $[0, 2\pi]$. The two periodic alternatives are

$$f_1(t) = \frac{3}{2 - \cos(5t)}, \qquad f_2(t) = 5\log(\sin(10t) + 5) + \exp(\cos(5t)),$$

both with fundamental period $2\pi/5$. The non-periodic null signals are

$$h_1(t) = (t - \pi)/\pi, \qquad h_2(t) = ((t - \pi)/\pi)^2,$$
$$h_3(t) = \{1 + \exp[-5(t - \pi)/\pi]\}^{-1} - 1/2, \qquad h_4(t) = 0.45(t - \pi)/\pi + 0.35\sin(1.7t + 0.2).$$

These null examples contain trend-like or low-frequency scalar structure without forming a closed delay-space loop. For the bounded-error experiments, we observe scalar values on the extended interval $[0, 2\pi + m\tau]$, construct a piecewise-linear interpolant, and use observed timestamps in $[0, 2\pi]$ as delay-vector anchors. The delay is chosen by the known-period rule

$$\tau = \frac{2\pi/5}{m + 1}.$$

The synthetic results in Table 1 were generated by the runner `src/new/run_synthetic.py` with the settings in `configs/synthetic_main.toml`. The runner constructs the noisy scalar observations, builds the linear interpolant, forms the raw delay cloud, computes the subsampling radius, and records both the raw decision and the bounded-error decision based on $\hat{c}^{\mathrm{err}}_{\alpha} = \hat{c}_{\alpha} + 5\zeta_{m,N}$. It writes the replicate-level output to `raw_results.csv` and records the configuration, code version, package versions, and random-seed rule in `metadata.json`; `src/new/report.py` computes the tables.

The code computes

$$\zeta_{m,N} = \sqrt{m + 1}\left(\sigma_N + \widehat{L}_{\mathrm{int}}\delta_N^2/8\right),$$

where δ_N is the maximum scalar observation spacing and $\widehat{L}_{\mathrm{int}}$ is a dense-grid numerical estimate of a second derivative bound on the observed interval. Thus the robust synthetic results should be read as the numerical implementation of the error-inflated rule, not as using a separately certified analytic upper bound for L_{int}.

With $m = 20$, the noise term alone gives $\zeta_{m,N} = \sqrt{21}\,\sigma_N \approx 0.092, 0.229, 0.458$ for $\sigma_N = 0.02, 0.05, 0.10$; including the interpolation term, the realized values average 0.099, 0.236, and 0.466 (and 0.007 in the noiseless case). Since robust TopPT requires persistence above $2(\hat{c}_{\alpha} + 5\zeta_{m,N})$, the

Table 3: Appendix stress-test rejection rates for near-periodic nonstationary signals. Entries are averaged over 100 seeds and the same seven bounded-error noise settings used in Table 1.

Signal	description	TDA raw	TDA robust	GLS	Fisher g	Perm.
g_1	phase drift	0.803	0.141	1.000	1.000	1.000
g_2	amplitude modulation	1.000	0.714	1.000	1.000	1.000

inflation $10\zeta_{m,N}$ grows quickly with σ_N: averaged over the two periodic signals, the robust detection rate is 1.000, 1.000, 0.535, and 0.000 for $\sigma_N = 0, 0.02, 0.05, 0.10$. This explains the lower robust power in Table 1.

For the main bounded-error run, the script was run with the six signals $f_1, f_2, h_1, \ldots, h_4$, 100 seeds, $N = 400$ scalar observations, $m = 20$, $\tau = (2\pi/5)/(m+1)$, $n = 336$ delay-vector anchors, $b = 120$, 500 Monte Carlo subsamples, and level $\alpha = 0.05$. The table averages over seven bounded noise settings: the noiseless case, bounded uniform noise with $\sigma \in \{0.02, 0.05, 0.10\}$, and truncated Gaussian noise with $\sigma \in \{0.02, 0.05, 0.10\}$. The robust TDA rule is reported only for these bounded-error cases, where the deterministic scalar error bound σ_N is available. For unbounded noise models such as Gaussian or Laplace noise, the raw TDA rule can still be computed, but the bounded-error theorem does not apply without an additional high-probability error event.

I.2 Near-periodic stress cases

We also ran an appendix stress test on two nonstationary oscillatory signals,

$$g_1(t) = \sin(5t + 0.18t^2), \qquad g_2(t) = \{1 + 0.18\cos(0.5t) + 0.08t\}\sin(5t).$$

Here, phase drift means that the oscillation phase is not linear in time, so the local frequency changes slowly; amplitude modulation means that a sinusoid is multiplied by a slowly varying, non-periodic envelope. These examples are not part of the main clean-null table because they may produce substantial loop-like sliding-window geometry even though they are not strictly periodic with period $2\pi/5$. Thus, TDA rejection here should be read as detection of recurrent geometric structure, not as a type-I error under the topological null used in Section 6.

Table 3 shows that the scalar spectral tests reject all stress cases. The raw TDA rule also rejects frequently, especially for the amplitude-modulated signal. The robust rule is more conservative because the deterministic interpolation-error correction inflates the rejection radius, but it can still reject when the delay cloud has a persistent loop beyond that correction.

I.3 Real-data protocol and diagnostic plots

The real-data analysis uses the respiration channel in both PhysioNet datasets. It is local in time because respiratory periodicity may drift over long intervals. For Fantasia, we use the first 90 seconds, downsample to 25 Hz, and test seven overlapping 20-second windows with starting times $0, 10, \ldots, 60$ seconds. For BIDMC, we use the first 180 seconds, apply a 0.05–0.8 Hz bandpass filter, downsample to 25 Hz, and test six 30-second windows with starting times $0, 30, \ldots, 150$ seconds. Each recording is linearly detrended and standardized over the whole interval before windows are selected.

For each window, candidate periods are estimated from autocorrelation and Welch spectral peaks. We take the strongest period hint and test up to three distinct local candidates obtained by multiplying it by 0.9, 1, and 1.1, clipped to the allowed period range. The TDA embedding uses $m = 8$ and $\tau = \text{period}/(m+1)$. Each delay cloud is capped at 300 anchor points, and the subsampling confidence radius uses $b = 0.99n$ with 300 Monte Carlo subsamples. This large finite-sample subsample fraction improves sensitivity in local windows but should be read as a practical calibration choice rather than the asymptotic regime of Theorem 5.1. We use global level $\alpha = 0.05$ and apply the finite-search Bonferroni correction from Corollary 6.4 within each record. The TDA correction is over the tested window-period pairs, whereas the scalar baselines are run once per window and are corrected over windows. A record is declared periodic if at least one Bonferroni-corrected local test rejects. In the realized search sets, this gives 825 TDA window-period tests for Fantasia and 954 for BIDMC; the Fantasia count is not exactly 3×280 because clipping and duplicate candidates can reduce the number of distinct local periods. Since the period candidates are estimated from the

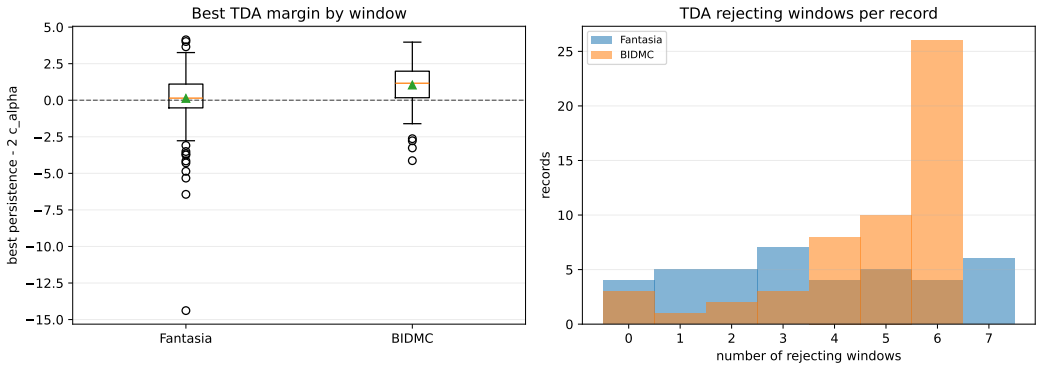


Figure 5: Real-data TDA diagnostics. Left: for each local window, the plotted quantity is the largest value of $\mathrm{pers}(\gamma) - 2\hat{c}_\alpha$ over the tested candidate periods. Positive values show that the persistence part of the TDA rule is satisfied for some candidate period; the full rejection rule also requires the corresponding birth time to be at most $\hat{c}_\alpha$. Right: distribution, over records, of the number of local windows with at least one rejecting TDA test.

observed window, this should be read as a Bonferroni correction over the realized finite search set; the formal statement of Corollary 6.4 applies to fixed finite grids.

For real data, we report only the raw TDA rule. The robust bounded-error rule requires a deterministic or high-probability upper bound on $\zeta_{m,N}$, which depends on both the scalar measurement error and the interpolation smoothness constant. The real recordings alone do not provide either a certified scalar measurement-error bound or a certified interpolation smoothness bound.

The real-data plotting script `src/new/report.py` also produces aggregate diagnostic plots from the record-level and window-level summaries. Figure 5 reports the two diagnostics that are most useful beyond Table 2: the distribution of the best TDA persistence margin over local windows and the number of rejecting windows per record. The plots show that the real-data rejections are local rather than uniform across all windows, and that BIDMC has larger and more frequent positive TDA margins than Fantasia.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction state the scope of the proposed subsampling confidence bound, the resulting periodicity test, and the experimental setting. These claims are supported by the theoretical results in Sections 4–6, the assumptions in Section 3 and Appendix A, and the experiments in Section 7.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

Question: Does the paper discuss the limitations of the work performed by the authors?

Answer: [Yes]

Justification: Section 8 discusses the main limitations and open directions, including the separated periodic/non-periodic signal classes, adaptive period selection, and the conservativeness of the robust interpolation-error rule.

Guidelines:

- The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
- The authors are encouraged to create a separate "Limitations" section in their paper.
- The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
- The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
- The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
- The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
- If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
- While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

   Justification: The assumptions are stated in Section 3 and Appendix A, and the main theoretical statements appear in Sections 4–6. Complete proofs are provided in the appendices.

   Guidelines:

   - The answer [N/A] means that the paper does not include theoretical results.
   - All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
   - All assumptions should be clearly stated or referenced in the statement of any theorems.
   - The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
   - Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
   - Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [Yes]

Justification: Section 7 and Appendix I describe the synthetic signals, real-data preprocessing, search grids, baselines, and decision rules. The experimental code and command-line runners are available in the public repository.

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [Yes]

Justification: The code is available at `https://github.com/taehyoungkim270/TopPT`, with command-line runners, configuration files, pinned package versions, and the saved results used in the paper. The real-data experiments use public PhysioNet datasets, and Appendix I.3 describes the preprocessing.

Guidelines:

- The answer [N/A] means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).

- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification: Section 7 gives the main experimental protocols and Appendix I gives the full signal definitions, preprocessing steps, period grids, embedding parameters, subsampling settings, and baseline settings. There is no model training or optimizer in these experiments.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   Justification: The experimental quantities are rejection decisions from tests run at stated significance levels, with Bonferroni correction for finite real-data searches. Synthetic tables report rejection rates over 100 seeds and seven bounded-error settings, and the method reports subsampling confidence radii rather than conventional estimator error bars.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.
   - It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
   - For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).

- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: All experiments were run on CPU only on an AMD Ryzen AI 7 450 machine with 32 GB RAM. The experiment scripts record wall-clock runtime in the saved configuration files; the complete rerun took approximately 12 hours.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: The authors have reviewed the NeurIPS Code of Ethics and the work conforms to it. The experiments use synthetic data and publicly available de-identified PhysioNet resources.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: The paper includes an Impact Statement. The work is methodological and not a deployed system; the discussion therefore focuses on its role as a general statistical tool and does not identify a direct high-risk application.

    Guidelines:

    - The answer [N/A] means that there is no societal impact of the work performed.
    - If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
    - The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate

to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.

- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: The paper does not release a model or dataset with high risk for misuse. The released asset is experiment code, and the real-data resources are already public de-identified PhysioNet datasets.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: The paper cites PhysioNet and the original Fantasia and BIDMC dataset publications. Both datasets are public PhysioNet resources distributed under the Open Data Commons Attribution License v1.0, and the experiments use only the respiration channels under those terms.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.

- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [Yes]

    Justification: The new asset is available in the public repository. The code includes command-line runners, saved configuration files, and documentation for reproducing the synthetic and real-data experiments.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing or new data collection from human subjects. The real-data experiments are secondary analyses of public de-identified PhysioNet datasets.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: The paper does not conduct new human-subject research. It uses publicly available de-identified PhysioNet data, so no new participant interaction or study-specific IRB approval is involved.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.

- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

    Answer: [N/A]

    Justification: The core method development and experiments do not use LLMs as an important, original, or non-standard component. Any writing or formatting assistance does not affect the methodology, scientific rigor, or originality of the work.

    Guidelines:

    - The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.