INTRINSIC MULTIRESOLUTION PERSISTENT HOMOLOGY: DIFFUSION-INDUCED RESOLUTIONS AND ENTROPY-SUMMARY REPRESENTATIONS

Anonymous authors
Paper under double-blind review

ABSTRACT

In topological data analysis, persistent homology (PH) characterizes topological features across filtration scales. Conventional PH, however, analyzes a single geometric representation across these scales and therefore does not organize features by intrinsic geometric resolution. This creates two coupled gaps: resolutions are fixed in advance rather than data-induced, and separate summaries do not jointly retain resolution and filtration-domain information. We introduce DIGIT (Diffusion-Induced Geometry for Intrinsic-Resolution Topology), which constructs ordered diffusion-induced geometries and selects each sample's retained range from its diffusion evolution. Its multiresolution entropy summary function (MESF) integrates the resulting persistence diagrams while retaining diffusion resolution, filtration-domain information, and persistence contribution. We establish conditional stability under discrete diffusion-operator perturbations. Experiments demonstrate progressive topological coarse-graining on synthetic point clouds; on three CATH benchmarks at the Class, Architecture, and Topology levels, DIGIT+MESF attains the highest mean accuracy among the evaluated persistence-based representations on all benchmarks.

1 INTRODUCTION

Topological data analysis (TDA) characterizes the shape of complex data and complements conventional statistical and learning approaches (Su et al., 2025; Dey & Wang, 2022; Chazal & Michel, 2021). Its central tool, *persistent homology* (PH), tracks connected components, loops, and voids across a filtration and records them in compact, stable persistence diagrams (Edelsbrunner et al., 2002; Cohen-Steiner et al., 2005; Otter et al., 2017). Although PH is multiscale in its filtration parameter, conventional PH is typically applied to a single geometric representation and therefore does not organize topology across geometric resolutions. This matters when genuine fine-scale structures coexist with random perturbations and broader geometric organization.

Multiresolution structures, in which fine-scale patterns are embedded within larger organizations, arise in molecular, materials, astronomical, and geophysical data (Xia et al., 2015; Beamer & Cang; Pyrialakos et al., 2025; Maevskiy et al., 2026; Wilding et al., 2021; Prat et al., 2026; Reed et al., 2023). *Multiresolution persistent homology* (MPH) (Xia et al., 2015) partially addresses this limitation by using an additional resolution parameter to construct resolution-dependent geometric representations before PH is computed. However, its resolution values are selected in advance and do not adapt to each sample's characteristic scales.

Figure 1 illustrates two complementary limitations. Under conventional PH, datasets A and B share the same global circular organization but differ in Gaussian perturbation, whereas C contains coherent local circular structures. Yet the 2-Wasserstein distances between the H_1 diagrams of A–C and A–B are comparable; thus, in this example, conventional PH does not by itself separate random perturbations from coherent fine-scale organization. Under MPH, the same resolutions, $\eta \in \{0.04, 0.08\}$, reveal the larger local loops in D at $\eta = 0.04$ but not the finer loops in C. This does not rule out other parameter choices; it shows that fixed resolutions can cover samples with different characteristic scales unevenly.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

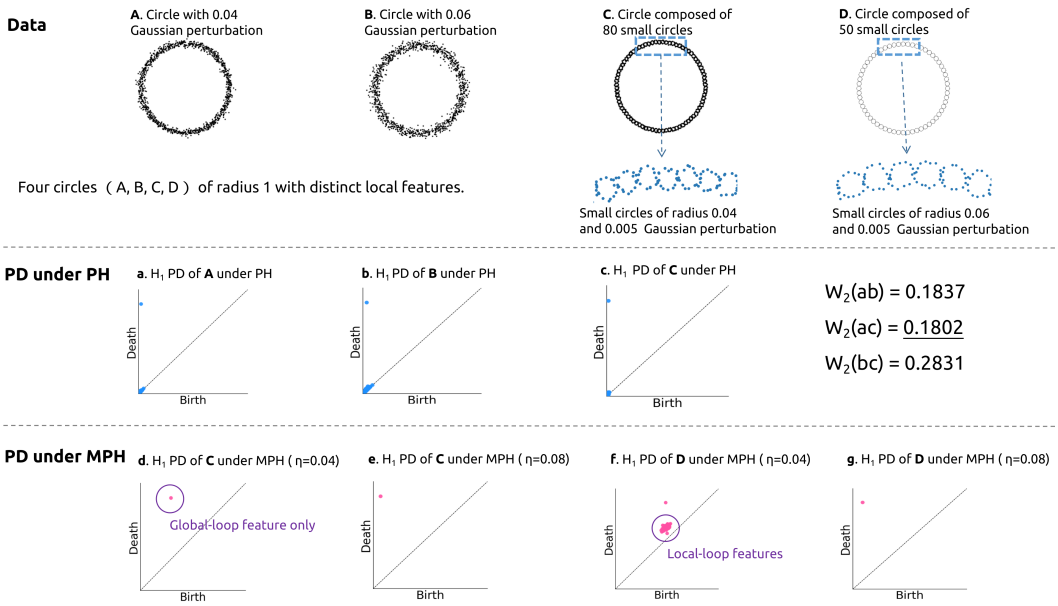


Figure 1: Comparison of conventional PH and MPH on unit-radius circles with distinct local structures. **Top:** A and B are circles with different levels of Gaussian perturbation, whereas C and D are composed of local circles at different scales. **Middle:** conventional PH produces comparable H_1 persistence-diagram distances between A and B and between A and C, despite the different sources of their local variation. **Bottom:** applying the same user-specified MPH resolutions, $\eta \in \{0.04, 0.08\}$, to C and D provides different coverage of their local-loop features.

At the representation stage, MPE (Xia et al., 2018) assigns one entropy value to each prescribed resolution, discarding feature locations, whereas Condensation Homology (Huguet et al., 2023b) integrates degree-0 evolution but not statewise higher-order diagrams across its hierarchy.

Together, these limitations motivate two requirements: (i) the resolution axis and retained range should follow each sample's intrinsic evolution rather than manually selected values; and (ii) the representation should retain both resolution and filtration-domain information, rather than only component evolution or one scalar per diagram.

These requirements motivate DIGIT, which constructs diffusion-time-indexed geometries, selects a sample-dependent retained range from their evolution, and computes PH at each retained time. Its representation component, MESF, organizes entropy summary functions (Atienza et al., 2020) along this axis while retaining resolution, filtration location, and persistence.

The main contributions are:

- **Adaptive intrinsic-resolution topology.** DIGIT traces PH across diffusion-induced geometries and derives its terminal time from the diffusion evolution.
- **Unified multiresolution representation.** MESF jointly retains diffusion resolution, filtration location, and persistence rather than one scalar per diagram.
- **Stability under diffusion-operator perturbations.** Under a prescribed vertex correspondence and nondegeneracy conditions, we establish conditional MESF stability to discrete diffusion-operator perturbations.

2 Related Work and Research Gaps

2.1 Related work

For this problem, methods are organized into two stages: *geometry construction* and *topological representation*. The former asks whether a method provides an ordered resolution family and constructs sample-adaptive intrinsic distances for PH. Here, sample-adaptive means that the realized

range and count depend on each sample under a common rule; it does not imply the absence of global hyperparameters. The latter asks whether a representation retains information along the resolution axis and within the filtration domain. The resolution axis changes the input geometry and is distinct from the PH filtration parameter.

Geometry construction. MPH analyzes topology across geometric resolutions (Xia et al., 2015) through FRI density representations. The kernel scale controls resolution, and PH is computed at each scale. Its range, schedule, and spatial discretization are chosen in advance, so MPH provides explicit multiresolution analysis but not sample-adaptive intrinsic distances.

Diffusion geometry uses random walks on a neighborhood graph (Coifman & Lafon, 2006); increasing time aggregates local connectivity for multiscale manifold analysis (et al., 2021; Huguet et al., 2023a). Diffusion distances improve PH under noise or high dimension (Bendich et al., 2011; Damrich et al., 2024), but are evaluated at prescribed parameters rather than as an adaptive resolution axis.

Diffusion Condensation constructs a data-driven hierarchy (Huguet et al., 2023b), while Condensation Homology records degree-0 evolution. Its statewise positive-degree PH uses ambient Euclidean distance, supplying neither resolution-varying intrinsic distances nor a unified cross-resolution representation.

Topological representation. Persistence diagrams are multisets of varying size rather than fixed-dimensional vectors, which complicates their direct use in statistical learning (Ali et al., 2023). Persistence landscapes encode diagrams as piecewise-linear functions (Bubenik, 2015), whereas persistence images discretize a weighted persistence surface (Adams et al., 2017). Persistent entropy reduces a diagram to a single scalar (Chintakunta et al., 2015), while the entropy summary function (ESF) retains where entropy contributions occur along the filtration (Atienza et al., 2020). These representations operate on individual persistence diagrams and therefore do not encode variation along a separate geometric-resolution axis.

Moving beyond the single-diagram setting, multiresolution persistent entropy (MPE) assigns an entropy scalar to each diagram generated at prescribed MPH resolutions (Xia et al., 2018). It therefore captures variation across resolution, but compresses each diagram to one scalar and discards the filtration locations of its intervals.

2.2 Research gaps

The review reveals two gaps. First, conventional PH fixes one geometry, whereas MPH and Condensation Homology expose multiresolution families; none constructs sample-adaptive, resolution-varying intrinsic distances for PH. Second, single-diagram vectorizations retain filtration-domain information but omit the resolution axis, whereas MPE retains this axis but reduces each diagram to one scalar. These motivate *data-adaptive intrinsic-resolution construction* and an *information-preserving cross-resolution representation*; Table 1 compares their desired properties.

Table 1: Comparison of approaches and desired properties for the two research gaps. "Yes" denotes an original property; "No" does not preclude extensions.

Gap 1: Data-adaptive intrinsic-resolution construction

Approach	Explicit multiresolution	Adaptive intrinsic distance for PH
Conventional PH	No	No
MPH	Yes	No
Condensation Homology	Yes	No
Desired	Yes	Yes

Gap 2: Information-preserving cross-resolution representation

Approach	Resolution-axis information	Filtration-domain information
PD vectorizations	No	Yes
MPE	Yes	No (entropy only)
Desired	Yes	Yes

3 INTRINSIC MULTIRESOLUTION PERSISTENT HOMOLOGY

Let $X = \{x_i\}_{i=1}^n \subset \mathbb{R}^d$. Figure 2 summarizes the pipeline, and Algorithm 1 in Appendix B provides complete pseudocode. We construct a connected undirected weighted graph with symmetric affinity matrix W. Although the construction only requires a connected reversible diffusion, our implementation augments a symmetrized k-nearest-neighbor graph with the Euclidean minimum spanning tree. For adjacent vertices we use the self-tuning affinity (Zelnik-Manor & Perona, 2004)

$$w_{ij} = \exp\left(-\frac{\|x_i - x_j\|_2^2}{\sigma_i \sigma_j}\right), \quad w_{ii} = 1, \tag{1}$$

where σ_i is the kth-neighbor distance and nonedge weights are zero. With $Q = \mathrm{diag}(q_1, \ldots, q_n)$ and $q_i = \sum_j w_{ij}$, the random-walk operator is $P = Q^{-1}W$ and its stationary distribution is $\pi_i = q_i / \sum_j q_j$.

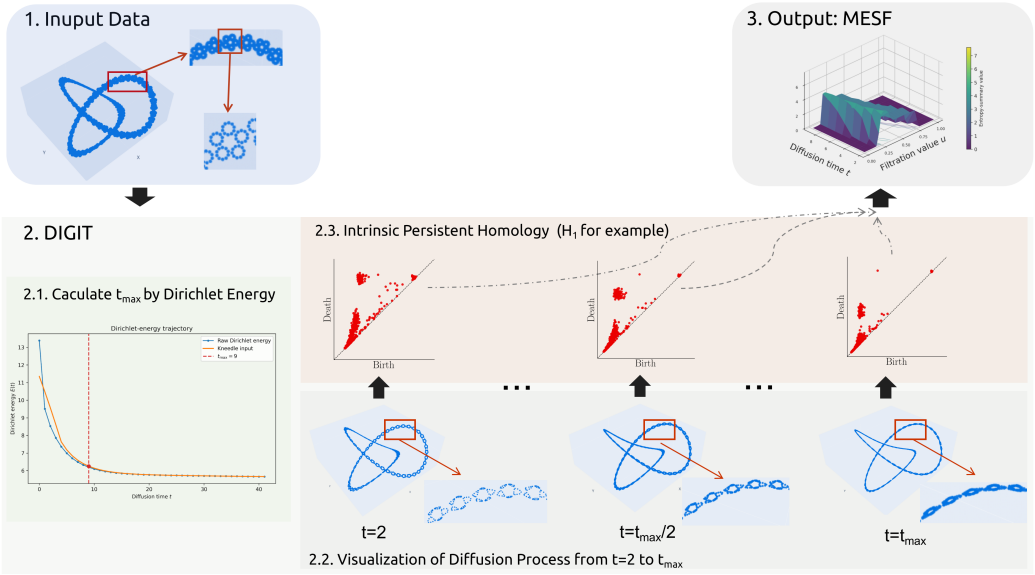


Figure 2: Overview of DIGIT+MESF. The energy trajectory selects the retained diffusion times. Reference-normalized diffusion distances induce persistence diagrams, whose entropy summary functions are arranged jointly over diffusion time and filtration position.

3.1 ADAPTIVE DIFFUSION-RESOLUTION RANGE

Let $\mathbf{X} \in \mathbb{R}^{n \times d}$ be the coordinate matrix and $\mathbf{X}^{(t)} = P^t \mathbf{X}$. We quantify remaining graph variation by the graph Dirichlet energy (Shuman et al., 2013)

$$E_t = \frac{1}{2} \sum_{i,j} w_{ij} \|\mathbf{X}_i^{(t)} - \mathbf{X}_j^{(t)}\|_2^2. \tag{2}$$

This graph-signal smoothness measures how much coordinate variation remains after t diffusion steps (Kalofolias, 2016). For a reversible walk, E_t is nonincreasing and discretely convex (Guan & Shi, 2025): spectrally it is a nonnegative mixture of modes proportional to $(1 - \lambda_\ell)\lambda_\ell^{2t}$; see Appendix C.1. The transition from rapid dissipation to diminishing change therefore supplies a sample-specific terminal resolution.

We evaluate E_t for $t = 0, \ldots, t_{\mathrm{cap}}$, smooth it with an odd-width moving average with edge padding, and apply a cumulative minimum. Kneedle (Satopaa et al., 2011) detects a knee under the convex-decreasing setting. With t_{knee} denoting a detected knee, we define

$$t_{\max}(X) = \begin{cases} \max\{2, t_{\mathrm{knee}}\}, & \text{if a knee is detected,} \\ t_{\mathrm{cap}}, & \text{otherwise,} \end{cases} \qquad \mathcal{T}_X = \{2, \ldots, t_{\max}(X)\}. \tag{3}$$

The lower limit $t = 2$ retains topology only after multistep denoising (Huguet et al., 2023a); $t = 0, 1$ remain available for estimating the energy trajectory. The smoothing width, Kneedle sensitivity, and cap are shared hyperparameters, whereas the realized terminal time follows each sample.

3.2 Diffusion-indexed persistent topology

At time t, the diffusion distance is (Coifman & Lafon, 2006)

$$D_t^2(x_i, x_j) = \sum_{\ell=1}^{n} \frac{\left[(P^t)_{i\ell} - (P^t)_{j\ell}\right]^2}{\pi_\ell}. \tag{4}$$

It compares complete transition profiles rather than direct displacement. For a Q-orthonormal eigensystem $(\lambda_\ell, \psi_\ell)$ of P,

$$D_t^2(x_i, x_j) = \mathrm{vol} \sum_{\ell=1}^{n-1} \lambda_\ell^{2t} \big(\psi_\ell(i) - \psi_\ell(j)\big)^2, \quad \mathrm{vol} = \sum_j q_j. \tag{5}$$

Thus increasing t attenuates nonstationary modes and moves the induced geometry from local organization toward coarser connectivity.

When comparing diffusion-indexed persistence diagrams across samples, their absolute filtration ranges may differ. We therefore adopt sample-wise reference normalization. Unlike normalizing each time-indexed diagram by its own range, this choice retains the relative contraction of filtration coordinates induced by diffusion. We put the retained filtrations of one sample on a common domain by anchoring their scale to the earliest retained geometry:

$$\Delta_X = \max_{i,j} D_2(x_i, x_j), \qquad \widehat{D}_t = D_t / \Delta_X. \tag{6}$$

We assume $\Delta_X > 0$. Since $|\lambda_\ell| \le 1$, Equation (5) gives $D_t \le D_2$ for $t \ge 2$; hence every pairwise value of $\widehat{D}_t$ lies in $[0, 1]$. Vietoris–Rips PH on $\widehat{D}_t$ yields the diagram $\widehat{\mathrm{PD}}^h_{X,t} = \{(\widehat{b}^h_{t,i}, \widehat{d}^h_{t,i})\}_i$. Positive homogeneity of Rips filtrations means this is equivalent to computing PH from D_t and dividing every birth and death by Δ_X. We do not match persistence points between adjacent times: each diagram describes the topology of one induced geometry.

3.3 Multiresolution entropy summary functions

The entropy summary function is defined for finite barcodes (Atienza et al., 2020). Accordingly, we discard intervals with infinite death from the unnormalized diagram $\mathrm{PD}^h_{D_t}$ and apply the ESF to the resulting barcode. For its intervals $(b^h_{t,i}, d^h_{t,i})$, let $\ell^h_{t,i} = d^h_{t,i} - b^h_{t,i}$ and $p^h_{t,i} = \ell^h_{t,i} / \sum_j \ell^h_{t,j}$. The ESF at time t is

$$Z^h_{\mathrm{ESF}}(t, \alpha) = -\sum_i p^h_{t,i} \log p^h_{t,i} \mathbf{1}_{[b^h_{t,i}, d^h_{t,i}]}(\alpha). \tag{7}$$

An empty finite barcode maps to the zero function. The degree-h MESF is

$$Z^h_{\mathrm{MESF}}(X; t, u) = Z^h_{\mathrm{ESF}}(X; t, u\Delta_X), \quad (t, u) \in \mathcal{T}_X \times [0, 1]. \tag{8}$$

Thus t records intrinsic resolution, u records relative filtration location, and the function value records the entropy-weighted persistence contribution. The same Δ_X anchors the retained diffusion-time axis. If all within-sample diffusion distances are multiplied by $s > 0$, then persistence coordinates and Δ_X both scale by s, while every $p^h_{t,i}$ and $u = \alpha/\Delta_X$ remain unchanged. MESF is therefore invariant to common sample-level scaling while preserving relative filtration positions.

For two samples represented on a common finite time set $\mathcal{T}$, we use the degree-h dissimilarity

$$d^h_{\mathrm{MESF}}(X, \widetilde{X}; \mathcal{T}) = \frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} \int_0^1 \left| Z^h_{\mathrm{MESF}}(X; t, u) - Z^h_{\mathrm{MESF}}(\widetilde{X}; t, u) \right| du. \tag{9}$$

For vector-valued learning, we sample u on an N_u-point uniform grid and concatenate time and homology channels. Within each outer training partition, sample-specific terminal times are averaged and rounded to obtain T_f; every training and test sample is then directly evaluated on $\{2, \ldots, T_f\}$. Thus test labels and geometry do not determine the shared domain, and no row is created by padding or time interpolation.

4 STABILITY ANALYSIS

We track a perturbation from the discrete diffusion operator through distance, reference normalization, persistence, and MESF. Consider samples with the same number of points and a prescribed vertex correspondence. Let P and $\widetilde{P}$ be irreducible, reversible, row-stochastic operators, with stationary distributions π and $\widetilde{\pi}$, and set $\delta_P = \|P - \widetilde{P}\|_\infty$, where $\|A\|_\infty = \max_i \sum_j |A_{ij}|$. Assume

$$\min_i\{\pi_i, \widetilde{\pi}_i\} \geq \pi_{\min} > 0, \qquad \|\pi - \widetilde{\pi}\|_\infty \leq C_\pi \delta_P. \tag{10}$$

The second condition follows under standard quantitative ergodicity assumptions for finite Markov chains (Mitrophanov, 2005; Rudolf & Schweizer, 2018).

Theorem 1 (Diffusion-induced persistence stability)**.** *For every fixed integer $t \geq 1$,*

$$\|D_t - \widetilde{D}_t\|_{\max} \leq C_{D,t}\delta_P, \quad C_{D,t} = \frac{2t}{\sqrt{\pi_{\min}}} + \frac{C_\pi}{\pi_{\min}^{3/2}}, \tag{11}$$

and $d_B(\mathrm{PD}^h_{D_t}, \mathrm{PD}^h_{\widetilde{D}_t}) \leq C_{D,t}\delta_P$.

The proof uses the telescoping identity for $P^t - \widetilde{P}^t$, giving $\|P^t - \widetilde{P}^t\|_\infty \leq t\delta_P$, followed by common-set Rips stability. The linear dependence on t reflects perturbation accumulation under repeated diffusion.

Proposition 2 (Reference-normalization stability)**.** *If additionally $\min\{\Delta_X, \Delta_{\widetilde{X}}\} \geq \Delta_{\min} > 0$, then for $t \geq 2$,*

$$\begin{aligned} \|\widehat{D}_t - \widehat{\widetilde{D}}_t\|_{\max} &\leq C_{\widehat{D},t}\delta_P, \\ d_B(\widehat{\mathrm{PD}}^h_{X,t}, \widehat{\mathrm{PD}}^h_{\widetilde{X},t}) \leq C_{\widehat{D},t}\delta_P, &\qquad C_{\widehat{D},t} = \frac{C_{D,t} + C_{D,2}}{\Delta_{\min}}. \end{aligned} \tag{12}$$

All finite coordinates of both normalized diagrams lie in $[0, 1]$. The same bound holds for the finite barcodes obtained after discarding the matched intervals with infinite death.

This follows because the diameter is 1-Lipschitz in matrix max norm and $D_t \leq D_2$ for retained times.

We finally aggregate ESF stability over time. Fix a finite common time set $\mathcal{T}$. For each t, let A_t, B_t be the normalized finite barcodes used by the corresponding ESF rows, and define

$$\begin{aligned} n_{\max,t} &= \max\{|A_t|, |B_t|\}, \qquad L_{\max,t} = \max\{L(A_t), L(B_t)\}, \\ r_{\infty,t} &= \frac{2n_{\infty,t}}{L_{\max,t}} d_B(A_t, B_t), \qquad L(C) = \sum_{(b,d)\in C} (d - b), \end{aligned} \tag{13}$$

where $n_{\infty,t}$ is the cardinality of a finite bottleneck matching.

Theorem 3 (MESF stability)**.** *Assume Proposition 2 on $\mathcal{T}$. At every t, suppose either $A_t = B_t = \emptyset$, or both have positive total persistence and $r_{\infty,t} \leq 2/(3e)$. Writing $\mathcal{T}_+$ for the nonempty times,*

$$d^h_{\mathrm{MESF}}(X, \widetilde{X}; \mathcal{T}) \leq \frac{\ln 2}{|\mathcal{T}|} \sum_{t\in\mathcal{T}_+} r_{\infty,t} L_{\max,t} \left[\frac{\log_2 n_{\max,t}}{n_{\max,t}} - \frac{3}{2} \log_2\left(\frac{3}{2} r_{\infty,t}\right) \right], \tag{14}$$

$$r_{\infty,t} \leq \frac{2n_{\infty,t} C_{\widehat{D},\mathcal{T}}}{L_{\max,t}} \delta_P, \quad C_{\widehat{D},\mathcal{T}} = \max_{t\in\mathcal{T}} C_{\widehat{D},t}. \tag{15}$$

Empty rows contribute zero.

Consequently, if reference diameters and nonzero total persistences are uniformly bounded away from zero and matching cardinalities are bounded, then

$$d^h_{\mathrm{MESF}}(X, \widetilde{X}) = O\left(\delta_P \log \frac{1}{\delta_P}\right). \tag{16}$$

Proofs are in Appendix C. These are deliberately conditional, operator-level results: they do not claim that discrete graph construction is stable under arbitrary coordinate perturbations, that Kneedle selects the same terminal time, or that the bound is uniform on an unbounded time axis. Computational complexity is deferred to Appendix D.

5 EXPERIMENTS

We study three questions: whether diffusion time induces interpretable topological coarse-graining, whether MESF improves prediction over single-geometry and scalar multiresolution summaries, and whether the result is sensitive to its filtration-grid resolution. Persistent homology is computed with GUDHI using the complete Vietoris–Rips filtration through the maximum pairwise dissimilarity; no quantile truncation or sparse-Rips approximation is used. Additional implementation settings and fold-level time-selection statistics are in Appendix E.

5.1 SYNTHETIC MULTIRESOLUTION GEOMETRY

We construct Circle, Eyeglasses, Linked circles, and Knot point clouds with nested circular structures at three spatial scales. Planar and three-dimensional global geometries test whether the behavior is geometry-specific. All use the same configuration ($k = 15$ (Damrich et al., 2024)). We compare persistence diagrams at $t = 2$, the retained-time midpoint, and $t = t_{\max}$, corresponding to early, intermediate, and terminal diffusion stages. Table 2 summarizes the construction parameters of the four multiresolution point clouds.

Table 2: Construction parameters of the four synthetic multiresolution datasets. Scale entries give points per local circle / radius.

Dataset	# Points	Coarse-scale locations	Intermediate circle (# pts / radius)	Fine circle (# pts / radius)
Circle	2160	45	6 / 0.05	8 / 0.02
Eyeglasses	2160	60	6 / 0.04	6 / 0.01
Linked circles	2880	80	6 / 0.06	6 / 0.01
Knot	1800	60	5 / 0.06	6 / 0.01

Before PH computation, each diffusion-distance matrix is normalized by the sample-specific reference diffusion diameter defined at $t = 2$. As a single-geometry intrinsic reference, we additionally compute PH from effective-resistance distance, whose distance matrix is normalized by its own diameter.

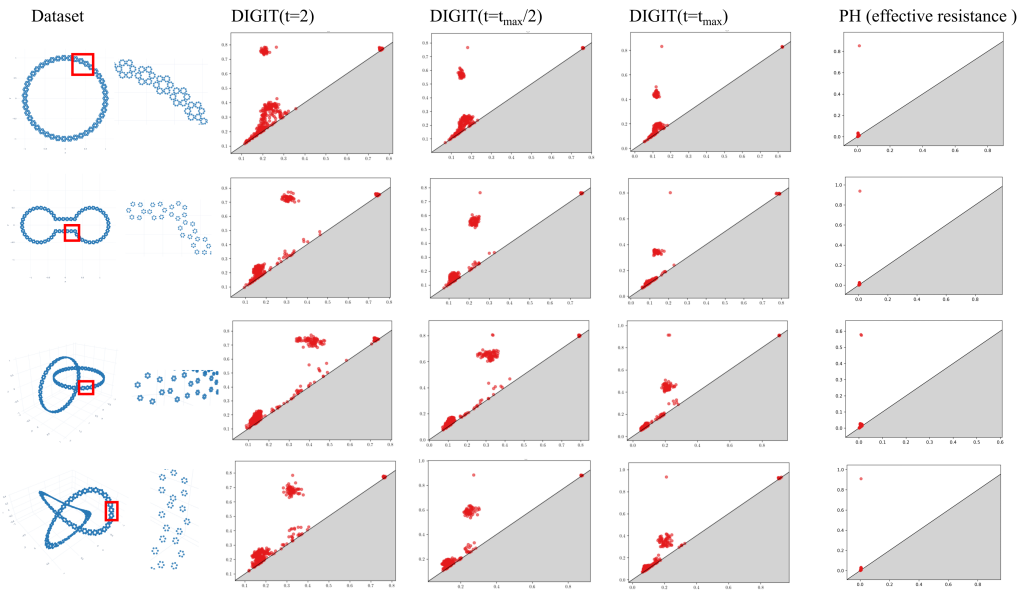


Figure 3: Resolution-dependent H_1 persistence on four hierarchical point clouds. Columns show the input, early, intermediate, and terminal DIGIT diagrams, and an effective-resistance PH reference. Across all geometries, diffusion progressively concentrates the persistence distribution.

Table 3: Persistent entropy (PE) and total persistence (TP) of the synthetic H_1 diagrams. These characterize the persistence distribution and are not performance scores.

Dataset	Statistic	DIGIT ($t = 2$)	DIGIT ($t = t_{\max}/2$)	DIGIT ($t = t_{\max}$)	PH (effective resistance)
Circle	PE	4.824	4.693	4.504	3.487
	TP	36.262	23.964	13.858	2.461
Eyeglasses	PE	4.958	4.730	4.346	3.087
	TP	34.576	22.206	9.934	1.945
Linked circles	PE	5.219	4.947	4.694	3.608
	TP	35.029	31.229	15.681	1.792
Knot	PE	5.169	4.864	4.824	3.509
	TP	33.896	30.426	16.515	2.388

Figure 3 shows a common progression from many broadly distributed finite features toward a smaller, concentrated set. Consistently, Table 3 shows that TP and PE decrease monotonically from $t = 2$ to $t_{\max}$ for every geometry. TP records the cumulative lifetime, whereas PE records how diffusely persistence is distributed across intervals. Their joint decrease indicates a systematic transition from a detail-rich short-time representation toward a more concentrated topology at larger diffusion times. The three displayed diffusion times are not interpreted as one-to-one recoveries of the three constructed spatial scales.

5.2 CATH PROTEIN-STRUCTURE CLASSIFICATION

Data and configuration. We use public CATH point-cloud benchmarks (Orengo et al., 1997) with predefined superfamily-separated 10-fold splits; their sizes and point-cloud statistics are summarized in Table 4. Each protein is represented by the three-dimensional coordinates of its C_α atoms. The tasks progress from broad secondary-structure composition (Class), through its spatial arrangement (Architecture), to connectivity of secondary-structure elements (Topology/Fold). We use $k = 5$, $H_0 \oplus H_1$, a seven-step moving-average window, Kneedle sensitivity $S = 10$, $t_{\text{cap}} = 81$, and $N_u = 20$.

Table 4: CATH benchmark statistics. Sample $t_{\max}$ reports descriptive full-benchmark mean $\pm$ standard deviation and is not used for outer-fold terminal-time selection.

Dataset	Proteins	Min. C_α	Avg. C_α	Sample $t_{\max}$
`cath_3class_ca`	16,962	13	142.9 ± 82.3	15.17 ± 2.14
`cath_10arch_ca`	7,110	13	144.5 ± 83.4	15.41 ± 1.99
`cath_20topo_ca`	4,240	13	131.0 ± 77.0	14.76 ± 2.37

For each outer fold f, sample-specific terminal times are determined using only the nine training folds and averaged to obtain a common terminal time T_f. All outer splits select $T_f = 15$ on all three tasks. Training and held-out samples are therefore evaluated directly at $t = 2, \ldots, 15$; the held-out fold does not influence time-range selection, and no shorter representation is padded or interpolated.

For each protein, the reference diffusion diameter is computed at $t = 2$. All persistence coordinates over $t = 2, \ldots, 15$ and both H_0 and H_1 are divided by this common sample-specific reference before MESF discretization on $[0, 1]$.

Baselines and evaluation. We compare with 12 persistence-diagram statistics (PD statistics)[1], persistence images (PI), persistence landscapes (PL), Betti curves (BC), single-diagram ESF, and DIGIT+MPE. The conventional baselines are computed from persistence diagrams obtained directly from the Euclidean geometry of each protein. All representations use both H_0 and H_1.

PI uses bandwidth 0.2 and a 20×20 grid per homology degree; PL uses five layers sampled at 100 coordinates; BC uses 100 coordinates; and single-diagram ESF uses 100 coordinates without maximum-amplitude normalization. Intervals with infinite death and numerical zero lifetime are removed before functional vectorization. DIGIT+MPE uses the same diffusion-indexed persistence diagrams as DIGIT+MESF, but reduces each diagram to a persistent-entropy scalar.

[1] `Anonymous link` (double-blind placeholder); the public notebook will be disclosed after review.

Table 5: Complete CATH results (mean ± standard deviation over ten predefined outer folds). The highest mean is bold and the second highest is underlined within each task and metric.

Task	Method	Accuracy	Precision	Recall	Macro-F1
3-Class	PD statistics	$0.8382 \pm .0278$	$0.8389 \pm .0229$	$0.8384 \pm .0252$	$0.8377 \pm .0243$
	PI	$\underline{0.8438 \pm .0223}$	$\underline{0.8429 \pm .0198}$	$\underline{0.8435 \pm .0210}$	$\underline{0.8425 \pm .0200}$
	PL	$0.6845 \pm .0294$	$0.6854 \pm .0235$	$0.6851 \pm .0309$	$0.6839 \pm .0277$
	BC	$0.7670 \pm .0264$	$0.7622 \pm .0234$	$0.7660 \pm .0261$	$0.7626 \pm .0236$
	ESF	$0.7821 \pm .0226$	$0.7789 \pm .0172$	$0.7826 \pm .0215$	$0.7804 \pm .0196$
	DIGIT+MPE	$0.7811 \pm .0257$	$0.7806 \pm .0222$	$0.7819 \pm .0251$	$0.7803 \pm .0239$
	DIGIT+MESF	$\mathbf{0.8507 \pm .0197}$	$\mathbf{0.8490 \pm .0164}$	$\mathbf{0.8508 \pm .0200}$	$\mathbf{0.8491 \pm .0183}$
10-Arch.	PD statistics	$\underline{0.4397 \pm .0274}$	$\underline{0.4256 \pm .0205}$	$\underline{0.4425 \pm .0197}$	$\underline{0.4239 \pm .0193}$
	PI	$0.4283 \pm .0243$	$0.4237 \pm .0139$	$0.4298 \pm .0152$	$0.4185 \pm .0138$
	PL	$0.3301 \pm .0223$	$0.3099 \pm .0171$	$0.3320 \pm .0192$	$0.3119 \pm .0168$
	BC	$0.3624 \pm .0270$	$0.3491 \pm .0143$	$0.3623 \pm .0180$	$0.3464 \pm .0131$
	ESF	$0.3930 \pm .0273$	$0.3751 \pm .0209$	$0.3934 \pm .0223$	$0.3760 \pm .0193$
	DIGIT+MPE	$0.3987 \pm .0238$	$0.3787 \pm .0146$	$0.4020 \pm .0139$	$0.3723 \pm .0136$
	DIGIT+MESF	$\mathbf{0.4755 \pm .0322}$	$\mathbf{0.4663 \pm .0224}$	$\mathbf{0.4779 \pm .0207}$	$\mathbf{0.4654 \pm .0213}$
20-Topo.	PD statistics	$\underline{0.3154 \pm .0553}$	$\underline{0.3225 \pm .0313}$	$\underline{0.3576 \pm .0345}$	$\underline{0.3063 \pm .0228}$
	PI	$0.2923 \pm .0433$	$0.2971 \pm .0194$	$0.3146 \pm .0341$	$0.2774 \pm .0168$
	PL	$0.2260 \pm .0496$	$0.2198 \pm .0447$	$0.2609 \pm .0403$	$0.2077 \pm .0293$
	BC	$0.2607 \pm .0489$	$0.2592 \pm .0342$	$0.2885 \pm .0338$	$0.2475 \pm .0232$
	ESF	$0.2838 \pm .0543$	$0.2819 \pm .0351$	$0.3176 \pm .0238$	$0.2685 \pm .0258$
	DIGIT+MPE	$0.2734 \pm .0426$	$0.2626 \pm .0323$	$0.2999 \pm .0224$	$0.2504 \pm .0217$
	DIGIT+MESF	$\mathbf{0.3248 \pm .0491}$	$\mathbf{0.3259 \pm .0418}$	$\mathbf{0.3597 \pm .0321}$	$\mathbf{0.3105 \pm .0251}$

An RBF SVM is selected inside each outer training set using stratified five-fold cross-validation over $C \in \{0.1, 1, 10, 100\}$ and $\gamma \in \{0.001, 0.01, 0.1, 1\}$; standardization is fitted within the training pipeline. We use random seed 50 and report accuracy, macro-precision, macro-recall, and macro-F1 as mean ± standard deviation over the ten identical outer folds.

Results. Table 5 shows that DIGIT+MESF attains the highest mean for all four metrics on every task. Relative to the strongest conventional baseline for each task, mean accuracy is higher by 0.69, 3.58, and 0.94 percentage points at the Class, Architecture, and Topology/Fold levels, respectively; mean macro-F1 is higher by 0.66, 4.15, and 0.42 points. The largest separation occurs at Architecture rather than increasing monotonically with hierarchy depth.

Because DIGIT+MPE and DIGIT+MESF use the same diffusion-indexed persistence diagrams, their comparison controls for the underlying diffusion-induced geometry. DIGIT+MESF achieves higher mean accuracy and macro-F1 on all three tasks, indicating that retaining information along the filtration is more informative here than reducing each diffusion time to a single entropy scalar. These are differences between outer-fold means; without a paired test, we do not claim statistical significance.

Filtration-grid sensitivity. Varying $N_u \in \{10, 15, 20, 25, 30\}$ changes mean accuracy by only 0.69, 1.25, and 2.22 percentage points across the three tasks (Table 6). The best values occur at 15, 20, and 25, respectively, so performance is not monotone in grid density. We use $N_u = 20$ as a compact common setting: best on Architecture, near-best on Class, and second-best on Topology/Fold.

Table 6: Sensitivity to the filtration-grid resolution N_u. Each cell gives mean accuracy / macro-F1 over the predefined outer folds. Bold denotes the highest mean and underlining the second highest for each task and metric.

N_u	3-Class	10-Architecture	20-Topology
10	0.8456/0.8441	0.4630/0.4495	0.3142/0.3000
15	**0.8525/0.8508**	0.4733/0.4623	0.3232/0.3097
20	<u>0.8507</u>/<u>0.8491</u>	**0.4755/0.4654**	<u>0.3248</u>/<u>0.3105</u>
25	0.8499/0.8483	<u>0.4751</u>/<u>0.4638</u>	**0.3364/0.3211**
30	0.8475/0.8459	0.4678/0.4563	0.3206/0.3081

6 Conclusion and Discussion

DIGIT turns diffusion time into an intrinsic resolution axis, while MESF retains topology over resolution and normalized filtration position. Synthetic data show consistent coarse-graining, and DIGIT+MESF attains the highest mean for every metric on all three CATH tasks.

The largest conventional-baseline gap occurs at Architecture and is not monotone in hierarchy depth. Against DIGIT+MPE, the results favor retaining filtration location over one entropy scalar per time. Results are stable across the tested grids but remain fold-mean comparisons pending paired tests.

Future work will address two main directions: (i) reducing DIGIT's main efficiency bottleneck, repeated Rips computations through adaptive time subsampling, and approximate diffusion; and (ii) integrating DIGIT with generative AI, for example by conditioning molecular generators on MESF or using it as a topology-aware objective for drug design. Theory beyond prescribed vertex correspondence and fixed adaptive-knee behavior also remains open.

Reproducibility Statement

The source code is provided in the supplementary materials and will be publicly released upon acceptance. The protein datasets used in our experiments are publicly available, with the corresponding sources cited in the paper. Theoretical results and their proofs are provided in the Appendix.

AI Use Statement

Generative AI tools were used to polish the manuscript and to assist in reviewing the coverage of related work identified by the authors. All literature references, technical content, and scientific claims were independently reviewed and verified by the authors. The authors take full responsibility for the final content of the manuscript.

References

Henry Adams, Tegan Emerson, Michael Kirby, Rachel Neville, Chris Peterson, Patrick Shipman, Sofya Chepushtanova, Eric Hanson, Francis Motta, and Lori Ziegelmeier. Persistence images: A stable vector representation of persistent homology. *Journal of Machine Learning Research*, 18 (8):1–35, 2017. URL https://www.jmlr.org/papers/v18/16-337.html.

Dashti Ali, Aras Asaad, Maria-Jose Jimenez, Vidit Nanda, Eduardo Paluzo-Hidalgo, and Manuel Soriano-Trigueros. A survey of vectorization methods in topological data analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(12):14069–14080, 2023. doi: 10.1109/TPAMI.2023.3308391.

Nieves Atienza, Rocío González-Díaz, and Manuel Soriano-Trigueros. On the stability of persistent entropy and new summary functions for topological data analysis. *Pattern Recognition*, 107: 107509, 2020. doi: 10.1016/j.patcog.2020.107509.

Perry Beamer and Zixuan Cang. Multiscale domain identification for spatial transcriptomics via persistent homology. *Cell Reports Methods*.

Paul Bendich, Taras Galkovskyi, and John Harer. Improving homology estimates with random walks. *Inverse Problems*, 27(12):124002, 2011.

Peter Bubenik. Statistical topological data analysis using persistence landscapes. *Journal of Machine Learning Research*, 16(3):77–102, 2015. URL https://www.jmlr.org/papers/v16/bubenik15a.html.

Frédéric Chazal and Bertrand Michel. An introduction to topological data analysis: fundamental and practical aspects for data scientists. *Frontiers in artificial intelligence*, 4:108, 2021.

Frédéric Chazal, David Cohen-Steiner, Leonidas J. Guibas, Facundo Mémoli, and Steve Y. Oudot. Gromov–hausdorff stable signatures for shapes using persistence. *Computer Graphics Forum*, 28 (5):1393–1403, 2009. doi: 10.1111/j.1467-8659.2009.01516.x.

Harish Chintakunta, Thanos Gentimis, Rocio Gonzalez-Diaz, Maria-Jose Jimenez, and Hamid Krim. An entropy-based persistence barcode. *Pattern Recognition*, 48(2):391–401, 2015. doi: 10.1016/j.patcog.2014.06.023.

David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. In *Proceedings of the twenty-first annual symposium on Computational geometry*, pp. 263–271, 2005.

David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. *Discrete & Computational Geometry*, 37(1):103–120, 2007. doi: 10.1007/s00454-006-1276-5.

Ronald R. Coifman and Stéphane Lafon. Diffusion maps. *Applied and Computational Harmonic Analysis*, 21(1):5–30, 2006. doi: 10.1016/j.acha.2006.04.006.

Sebastian Damrich, Philipp Berens, and Dmitry Kobak. Persistent homology for high-dimensional data based on spectral methods. *Advances in Neural Information Processing Systems*, 37:41954–42014, 2024.

Tamal Krishna Dey and Yusu Wang. *Computational topology for data analysis*. Cambridge University Press, 2022.

Edelsbrunner, Letscher, and Zomorodian. Topological persistence and simplification. *Discrete Comput. Geom.*, 28(4):511–533, November 2002. ISSN 0179-5376. doi: 10.1007/s00454-002-2885-2. URL https://doi.org/10.1007/s00454-002-2885-2.

Burkhardt et al. Multiscale phate identifies multimodal signatures of covid-19. *Nature Biotechnology*, 2021.

Weiqi Guan and Zihao Shi. Measuring over-smoothing beyond dirichlet energy. *arXiv preprint arXiv:2512.06782*, 2025. doi: 10.48550/arXiv.2512.06782.

Guillaume Huguet, Alexander Tong, Edward De Brouwer, Yanlei Zhang, Guy Wolf, Ian Adelstein, and Smita Krishnaswamy. A heat diffusion perspective on geodesic preserving dimensionality reduction. *Advances in Neural Information Processing Systems*, 36:6986–7016, 2023a.

Guillaume Huguet, Alexander Tong, Bastian Rieck, Jessie Huang, Manik Kuchroo, Matthew Hirn, Guy Wolf, and Smita Krishnaswamy. Time-inhomogeneous diffusion geometry and topology. *SIAM Journal on Mathematics of Data Science*, 5(2):346–372, 2023b.

Vassilis Kalofolias. How to learn a graph from smooth signals. In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, volume 51 of *Proceedings of Machine Learning Research*, pp. 920–929. PMLR, 2016. URL https://proceedings.mlr.press/v51/kalofolias16.html.

Artem Maevskiy, Vitalii Kapitan, and Andrey Ustyuzhanin. Artificial intelligence for multiscale modeling in solid-state physics and chemistry: A comprehensive review. *Advanced Intelligent Systems*, pp. e202501219, 2026.

Alexander Yu. Mitrophanov. Sensitivity and convergence of uniformly ergodic markov chains. *Journal of Applied Probability*, 42(4):1003–1014, 2005. doi: 10.1239/jap/1134587812.

Christine A Orengo, Alex D Michie, Susan Jones, David T Jones, Mark B Swindells, and Janet M Thornton. Cath–a hierarchic classification of protein domain structures. *Structure*, 5(8):1093–1109, 1997.

Nina Otter, Mason A Porter, Ulrike Tillmann, Peter Grindrod, and Heather A Harrington. A roadmap for the computation of persistent homology. *EPJ Data Science*, 6:1–38, 2017.

Judit Prat, M Gatti, C Doux, P Pranav, C Chang, N Jeffrey, L Whiteway, D Anbajagane, S Sugiyama, A Thomsen, et al. Dark energy survey year 3 results: w cdm cosmology from simulation-based inference with persistent homology on the sphere. *Monthly Notices of the Royal Astronomical Society*, 545(3):staf2152, 2026.

Stefanos Pyrialakos, Ioannis Kalogeris, and Vissarion Papadopoulos. An efficient hierarchical bayesian framework for multiscale material modeling. *Composite Structures*, 351:118570, 2025.

Colorado J Reed, Ritwik Gupta, Shufan Li, Sarah Brockman, Christopher Funk, Brian Clipp, Kurt Keutzer, Salvatore Candido, Matt Uyttendaele, and Trevor Darrell. Scale-mae: A scale-aware masked autoencoder for multiscale geospatial representation learning. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 4088–4099, 2023.

Daniel Rudolf and Nikolaus Schweizer. Perturbation theory for markov chains via wasserstein distance. *Bernoulli*, 24(4A):2610–2639, 2018. doi: 10.3150/17-BEJ938.

Ville Satopaa, Jeannie Albrecht, David Irwin, and Barath Raghavan. Finding a "kneedle" in a haystack: Detecting knee points in system behavior. In *2011 31st International Conference on Distributed Computing Systems Workshops*, pp. 166–171, 2011. doi: 10.1109/ICDCSW.2011.20.

David I. Shuman, Sunil K. Narang, Pascal Frossard, Antonio Ortega, and Pierre Vandergheynst. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Processing Magazine*, 30(3):83–98, 2013. doi: 10.1109/MSP.2012.2235192.

Zhe Su, Xiang Liu, Layal Bou Hamdan, Vasileios Maroulas, Jie Wu, Gunnar Carlsson, and Guo-Wei Wei. Topological data analysis and topological deep learning beyond persistent homology: a review. *Artificial Intelligence Review*, 2025.

Georg Wilding, Keimpe Nevenzeel, Rien van de Weygaert, Gert Vegter, Pratyush Pranav, Bernard JT Jones, Konstantinos Efstathiou, and Job Feldbrugge. Persistent homology of the cosmic web–i. hierarchical topology in λcdm cosmologies. *Monthly Notices of the Royal Astronomical Society*, 507(2):2968–2990, 2021.

Kelin Xia, Zhixiong Zhao, and Guo-Wei Wei. Multiresolution persistent homology for excessively large biomolecular datasets. *The Journal of chemical physics*, 143(13), 2015.

Kelin Xia, Zhiming Li, and Lin Mu. Multiscale persistent functions for biomolecular structure characterization. *Bulletin of mathematical biology*, 80(1):1–31, 2018.

Lihi Zelnik-Manor and Pietro Perona. Self-tuning spectral clustering. In *Advances in Neural Information Processing Systems*, volume 17, 2004.

Afra Zomorodian. Fast construction of the vietoris-rips complex. *Computers & Graphics*, 34(3): 263–271, 2010.

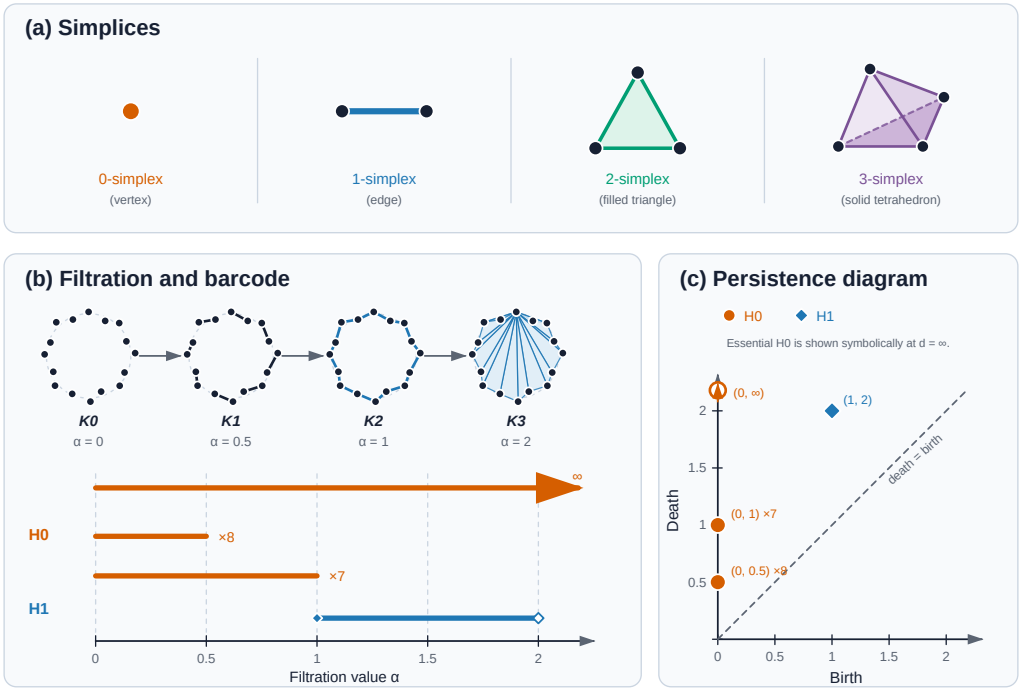


Figure 4: From simplices to persistence. **(a)** Aligned examples of 0-, 1-, 2-, and 3-simplices, ending with a solid tetrahedron. **(b)** A filtration on sixteen irregularly spaced points sampled around a circle. Eight pairs merge at $\alpha = 0.5$, the remaining components merge and the circular H_1 class appears at $\alpha = 1$, and this hole is filled at $\alpha = 2$. **(c)** Each finite interval $[b, d)$ maps to the persistence-diagram point (b, d). The essential H_0 interval is shown symbolically at $d = \infty$ and is excluded from the finite functional summaries used by ESF and MESF.

A TECHNICAL BACKGROUND

A.1 VIETORIS–RIPS PERSISTENCE

Persistent homology tracks how homology changes along a filtration. All homology groups are computed over a fixed coefficient field $\mathbb{F}$; our experiments use $\mathbb{F} = \mathbb{Z}_2$. An abstract simplicial complex on a vertex set V is a collection of nonempty finite subsets of V that is closed under taking nonempty subsets. An element containing $q + 1$ vertices is a q-simplex.

Definition 4 (Vietoris–Rips filtration)**.** *Let X be finite and let $\delta : X \times X \to \mathbb{R}_{\geq 0}$ be symmetric with $\delta(x, x) = 0$. At filtration value $\alpha \geq 0$, define*

$$\mathrm{VR}(X, \delta, \alpha) = \{\sigma \subseteq X : \delta(x_i, x_j) \leq \alpha \text{ for all } x_i, x_j \in \sigma\}. \tag{17}$$

Because $\mathrm{VR}(X, \delta, \alpha_1) \subseteq \mathrm{VR}(X, \delta, \alpha_2)$ whenever $\alpha_1 \leq \alpha_2$, these complexes form a filtration (Zomorodian, 2010).

Applying degree-h homology records intervals $[b_i^h, d_i^h)$, where a class appears at b_i^h and disappears at d_i^h. The corresponding persistence diagram is the multiset $\mathrm{PD}^h(X, \delta) = \{(b_i^h, d_i^h)\}_i$. Following Atienza et al. (2020), the ESF is applied to the finite barcode obtained after discarding intervals with infinite death.

A.2 BOTTLENECK AND COMMON-SET STABILITY

For diagrams A and B, each augmented by diagonal points with infinite multiplicity, the bottleneck distance is

$$d_B(A, B) = \inf_{\gamma} \sup_{z \in A} \|z - \gamma(z)\|_\infty, \tag{18}$$

where γ ranges over bijections between the augmented diagrams. On a common finite vertex set, uniform perturbations of the pairwise dissimilarity directly control the diagrams (Cohen-Steiner et al., 2007; Chazal et al., 2009).

Proposition 5 (Common-set Rips stability)**.** *For two pairwise dissimilarities δ and $\widetilde{\delta}$ on the same finite set and every degree h,*

$$d_B\Big(\mathrm{PD}^h_\delta, \mathrm{PD}^h_{\widetilde{\delta}}\Big) \le \|\delta - \widetilde{\delta}\|_{\max}. \tag{19}$$

Indeed, if the max-norm difference is ε, then each Rips complex at α under either dissimilarity is contained in the other at $\alpha + \varepsilon$. The resulting ε-interleaving gives Equation (19). This is the persistence-stability step used below after bounding the diffusion dissimilarities.

B Algorithmic Summary

Algorithm 1 DIGIT with the MESF representation

Require: Point cloud X; neighborhood size k; degrees $\mathcal{H}$; t_{cap}; grid size N_u.
Ensure: Multiresolution representation $\{Z^h_{\mathrm{MESF}}\}_{h\in\mathcal{H}}$.

```
Construct a connected undirected weighted graph from X using k; define W by Equation (1), and form Q and P = Q^{-1}W.
Form X^{(t)} = P^t X and evaluate E_t in Equation (2) for t = 0, ..., t_cap.
Smooth {E_t}, apply the cumulative minimum, and detect t_knee with Kneedle.
Set t_max(X) and T_X by Equation (3).
Compute D_2 by Equation (4) and set Δ_X by Equation (6).
for t ∈ T_X do
    Compute D_t by Equation (4) and normalize it by Equation (6).
    for h ∈ H do
        Compute the Rips diagram of D̂_t.
        Discard intervals with infinite death and evaluate the ESF row in Equation (7) on the N_u-point grid.
    end for
end for
For each h ∈ H, stack the ESF rows across time as in Equation (8).
return {Z^h_MESF}_{h∈H}.
```

C Proofs and Additional Analysis

C.1 Dirichlet-energy decay

We first justify the monotonicity and discrete convexity used for terminal-time selection. Because W is symmetric, $P = Q^{-1}W$ is self-adjoint under the weighted inner product $\langle f, g\rangle_Q = f^\top Q g$. Let $\{(\lambda_\ell, \psi_\ell)\}_{\ell=0}^{n-1}$ be a Q-orthonormal eigensystem of P. For coordinate signal $x^{(r)}$, write

$$x^{(r)} = \sum_{\ell=0}^{n-1} c_{\ell r}\psi_\ell. \tag{20}$$

Since $Q - W = Q(I - P)$, the energy in Equation (2) satisfies

$$E_t = \sum_{r=1}^{d}\sum_{\ell=0}^{n-1}(1-\lambda_\ell)\lambda_\ell^{2t}c_{\ell r}^2. \tag{21}$$

Reversibility gives $\lambda_\ell \in [-1, 1]$. Hence every coefficient in Equation (21) is nonnegative and $\lambda_\ell^{2(t+1)} \le \lambda_\ell^{2t}$, so

$$E_{t+1} \le E_t. \tag{22}$$

Moreover,

$$E_{t+2} - 2E_{t+1} + E_t = \sum_{r=1}^{d} \sum_{\ell=0}^{n-1} (1-\lambda_\ell)\lambda_\ell^{2t} \times (1-\lambda_\ell^2)^2 c_{\ell r}^2 \geq 0. \tag{23}$$

Thus the exact energy trajectory is nonincreasing and discretely convex, which supports applying Kneedle in the convex-decreasing setting after the numerical smoothing described in Section 3.

C.2 Proof of diffusion-induced persistence stability

We prove Theorem 1. Identify the vertices of X and $\widetilde{X}$ through the prescribed correspondence and index them by the common finite set $V = \{1, \dots, n\}$. Recall that $\delta_P = \|P - \widetilde{P}\|_\infty$, where $\|A\|_\infty = \max_i \sum_j |A_{ij}|$.

The standard matrix-power identity gives

$$P^t - \widetilde{P}^t = \sum_{s=0}^{t-1} P^s (P - \widetilde{P}) \widetilde{P}^{t-1-s}. \tag{24}$$

By submultiplicativity of the induced infinity norm,

$$\|P^t - \widetilde{P}^t\|_\infty \leq \sum_{s=0}^{t-1} \|P^s\|_\infty \|P - \widetilde{P}\|_\infty \|\widetilde{P}^{t-1-s}\|_\infty. \tag{25}$$

Every power of a row-stochastic matrix is row-stochastic and therefore has induced infinity norm one. Consequently,

$$\|P^t - \widetilde{P}^t\|_\infty \leq t\delta_P. \tag{26}$$

For corresponding vertices (i, j), define

$$a_{ij}^{(t)} = P^t(i,\cdot) - P^t(j,\cdot), \qquad \widetilde{a}_{ij}^{(t)} = \widetilde{P}^t(i,\cdot) - \widetilde{P}^t(j,\cdot). \tag{27}$$

Let $\Pi = \mathrm{diag}(\pi)$ and $\widetilde{\Pi} = \mathrm{diag}(\widetilde{\pi})$. The diffusion distances can be written as

$$D_t(i,j) = \|a_{ij}^{(t)} \Pi^{-1/2}\|_2, \qquad \widetilde{D}_t(i,j) = \|\widetilde{a}_{ij}^{(t)} \widetilde{\Pi}^{-1/2}\|_2. \tag{28}$$

The reverse triangle inequality, followed by adding and subtracting $\widetilde{a}_{ij}^{(t)} \Pi^{-1/2}$, yields

$$\begin{aligned} |D_t(i,j) - \widetilde{D}_t(i,j)| \leq{}& \|(a_{ij}^{(t)} - \widetilde{a}_{ij}^{(t)})\Pi^{-1/2}\|_2 \\ &+ \|\widetilde{a}_{ij}^{(t)} (\Pi^{-1/2} - \widetilde{\Pi}^{-1/2})\|_2. \end{aligned} \tag{29}$$

We first control the perturbation of the transition profiles. From Equation (10),

$$\begin{aligned} \|(a_{ij}^{(t)} - \widetilde{a}_{ij}^{(t)})\Pi^{-1/2}\|_2 &\leq \frac{1}{\sqrt{\pi_{\min}}} \|a_{ij}^{(t)} - \widetilde{a}_{ij}^{(t)}\|_2 \\ &\leq \frac{1}{\sqrt{\pi_{\min}}} \|a_{ij}^{(t)} - \widetilde{a}_{ij}^{(t)}\|_1. \end{aligned} \tag{30}$$

Furthermore,

$$\begin{aligned} \|a_{ij}^{(t)} - \widetilde{a}_{ij}^{(t)}\|_1 &\leq \|P^t(i,\cdot) - \widetilde{P}^t(i,\cdot)\|_1 \\ &\quad + \|P^t(j,\cdot) - \widetilde{P}^t(j,\cdot)\|_1 \\ &\leq 2\|P^t - \widetilde{P}^t\|_\infty. \end{aligned} \tag{31}$$

Combining Equations (26), (30) and (31) gives

$$\|(a_{ij}^{(t)} - \widetilde{a}_{ij}^{(t)})\Pi^{-1/2}\|_2 \leq \frac{2t}{\sqrt{\pi_{\min}}} \delta_P. \tag{32}$$

We next control the perturbation of the stationary weights. The scalar map $f(x) = x^{-1/2}$ has derivative $f'(x) = -\frac{1}{2}x^{-3/2}$. Hence the mean-value theorem and Equation (10) give

$$\|\Pi^{-1/2} - \widetilde{\Pi}^{-1/2}\|_{2,\mathrm{op}} \le \frac{1}{2\pi_{\min}^{3/2}} \|\pi - \widetilde{\pi}\|_\infty \le \frac{C_\pi}{2\pi_{\min}^{3/2}} \delta_P. \tag{33}$$

Because $\widetilde{a}_{ij}^{(t)}$ is the difference of two probability vectors, $\|\widetilde{a}_{ij}^{(t)}\|_2 \le \|\widetilde{a}_{ij}^{(t)}\|_1 \le 2$. Therefore,

$$\|\widetilde{a}_{ij}^{(t)}(\Pi^{-1/2} - \widetilde{\Pi}^{-1/2})\|_2 \le \frac{C_\pi}{\pi_{\min}^{3/2}} \delta_P. \tag{34}$$

Substitution into Equation (29) yields, for every pair (i, j),

$$|D_t(i,j) - \widetilde{D}_t(i,j)| \le \left(\frac{2t}{\sqrt{\pi_{\min}}} + \frac{C_\pi}{\pi_{\min}^{3/2}} \right) \delta_P. \tag{35}$$

Taking the maximum over all corresponding pairs proves Equation (11).

It remains to transfer this estimate to persistent homology. Under the prescribed correspondence, both dissimilarities are defined on the common vertex set V. For every simplex $\sigma \subseteq V$, their Rips filtration values are

$$F_t(\sigma) = \max_{i,j \in \sigma} D_t(i,j), \qquad \widetilde{F}_t(\sigma) = \max_{i,j \in \sigma} \widetilde{D}_t(i,j). \tag{36}$$

Thus

$$|F_t(\sigma) - \widetilde{F}_t(\sigma)| \le \|D_t - \widetilde{D}_t\|_{\max} \tag{37}$$

for every simplex. Applying Proposition 5 gives

$$d_B(\mathrm{PD}_{D_t}^h, \mathrm{PD}_{\widetilde{D}_t}^h) \le \|D_t - \widetilde{D}_t\|_{\max} \le C_{D,t} \delta_P, \tag{38}$$

which completes the proof of Theorem 1.

C.3 Reference-normalization stability

We prove Proposition 2. For the reversible diffusion operator, Equation (5) gives, for every integer $t \ge 2$,

$$\begin{aligned} D_t^2(i,j) &= \mathrm{vol} \sum_{\ell=1}^{n-1} \lambda_\ell^{2t} (\psi_\ell(i) - \psi_\ell(j))^2 \\ &\le \mathrm{vol} \sum_{\ell=1}^{n-1} \lambda_\ell^4 (\psi_\ell(i) - \psi_\ell(j))^2 = D_2^2(i,j), \end{aligned} \tag{39}$$

because $|\lambda_\ell| \le 1$. Consequently,

$$0 \le \widehat{D}_t(i,j) = \frac{D_t(i,j)}{\Delta_X} \le 1, \tag{40}$$

and the same conclusion holds for $\widehat{\widetilde{D}}_t$. Since a Rips simplex enters at the maximum pairwise dissimilarity among its vertices, all finite filtration coordinates of both normalized diagrams lie in $[0, 1]$.

The maximum functional is 1-Lipschitz in the max norm, so

$$|\Delta_X - \Delta_{\widetilde{X}}| \le \|D_2 - \widetilde{D}_2\|_{\max}. \tag{41}$$

For every corresponding pair (i, j) and every $t \ge 2$, the triangle inequality gives

$$\begin{aligned} &\left| \frac{D_t(i,j)}{\Delta_X} - \frac{\widetilde{D}_t(i,j)}{\Delta_{\widetilde{X}}} \right| \\ &\quad \le \frac{|D_t(i,j) - \widetilde{D}_t(i,j)|}{\Delta_X} + \widetilde{D}_t(i,j) \frac{|\Delta_X - \Delta_{\widetilde{X}}|}{\Delta_X \Delta_{\widetilde{X}}}. \end{aligned} \tag{42}$$

By Equation (39), $\widetilde{D}_t(i,j) \leq \Delta_{\widetilde{X}}$. Combining this fact with Equations (11) and (41) and $\min\{\Delta_X, \Delta_{\widetilde{X}}\} \geq \Delta_{\min}$ yields

$$\begin{aligned}\|\widehat{D}_t - \widehat{\widetilde{D}}_t\|_{\max} &\leq \frac{\|D_t - \widetilde{D}_t\|_{\max} + \|D_2 - \widetilde{D}_2\|_{\max}}{\Delta_{\min}} \\ &\leq \frac{C_{D,t} + C_{D,2}}{\Delta_{\min}}\delta_P = C_{\widehat{D},t}\delta_P. \end{aligned} \tag{43}$$

Applying Proposition 5 to the normalized dissimilarities gives

$$d_B(\widehat{\mathrm{PD}}^h_{X,t}, \widehat{\mathrm{PD}}^h_{\widetilde{X},t}) \leq C_{\widehat{D},t}\delta_P. \tag{44}$$

For a complete Rips filtration on a nonempty finite vertex set, each degree-0 diagram has one interval with infinite death, born at zero, whereas diagrams in positive degree have none. The two infinite intervals therefore match at zero cost. Removing these matched intervals, together with any diagonal zero-length intervals added for the matching, leaves a matching of the finite barcodes A_t and B_t with no larger cost. Hence

$$d_B(A_t, B_t) \leq d_B(\widehat{\mathrm{PD}}^h_{X,t}, \widehat{\mathrm{PD}}^h_{\widetilde{X},t}) \leq C_{\widehat{D},t}\delta_P, \tag{45}$$

which proves Proposition 2.

C.4 MESF STABILITY

We prove Theorem 3. Fix the common finite time set $\mathcal{T}$. For degree h and time t, let A_t, B_t be the normalized finite barcodes used by the corresponding MESF rows.

For an interval $(b^h_{X,t,i}, d^h_{X,t,i})$, reference normalization changes its lifetime from

$$\ell^h_{X,t,i} = d^h_{X,t,i} - b^h_{X,t,i} \quad \text{to} \quad \ell^h_{X,t,i}/\Delta_X. \tag{46}$$

Its normalized persistence weight is unchanged because

$$\frac{\ell^h_{X,t,i}/\Delta_X}{\sum_j \ell^h_{X,t,j}/\Delta_X} = \frac{\ell^h_{X,t,i}}{\sum_j \ell^h_{X,t,j}}. \tag{47}$$

Thus reference normalization changes the filtration coordinates but not the entropy weights. If $S(A_t)$ and $S(B_t)$ denote the ESFs of the normalized barcodes, then by Equation (8),

$$Z^h_{\mathrm{MESF}}(X; t, \cdot) = S(A_t), \qquad Z^h_{\mathrm{MESF}}(\widetilde{X}; t, \cdot) = S(B_t). \tag{48}$$

By Equation (45),

$$d_B(A_t, B_t) \leq C_{\widehat{D},t}\delta_P. \tag{49}$$

Recall the quantities

$$\begin{aligned} &n_{\max,t} = \max\{|A_t|, |B_t|\}, \\ &L_{\max,t} = \max\{L(A_t), L(B_t)\}, \qquad L(C) = \sum_{(b,d)\in C}(d-b). \end{aligned} \tag{50}$$

Let $n_{\infty,t}$ be the cardinality of a bottleneck matching, with zero-length intervals added when necessary to obtain a finite bijection. For $L_{\max,t} > 0$, define

$$r_{\infty,t} = \frac{2n_{\infty,t}}{L_{\max,t}} d_B(A_t, B_t). \tag{51}$$

Writing $C_{\widehat{D},\mathcal{T}} = \max_{t\in\mathcal{T}} C_{\widehat{D},t}$, Equation (49) implies

$$r_{\infty,t} \leq \frac{2n_{\infty,t}C_{\widehat{D},\mathcal{T}}}{L_{\max,t}}\delta_P. \tag{52}$$

Suppose now that both barcodes have positive total persistence and $r_{\infty,t} \le 2/(3e)$. The ESF stability theorem (Atienza et al., 2020, Theorem 4.2), adjusted from base-2 to natural logarithms, gives

$$\|S(A_t) - S(B_t)\|_{L^1} \le (\ln 2) r_{\infty,t} L_{\max,t} \times \left[\frac{\log_2 n_{\max,t}}{n_{\max,t}} - \frac{3}{2} \log_2 \left(\frac{3}{2} r_{\infty,t} \right) \right]. \tag{53}$$

All finite endpoints of A_t and B_t lie in $[0, 1]$. Their ESFs therefore vanish outside this interval, so Equation (53) is equivalently

$$\int_0^1 \left| Z^h_{\mathrm{MESF}}(X; t, u) - Z^h_{\mathrm{MESF}}(\widetilde{X}; t, u) \right| du \le (\ln 2) r_{\infty,t} L_{\max,t} \left[\frac{\log_2 n_{\max,t}}{n_{\max,t}} - \frac{3}{2} \log_2 \left(\frac{3}{2} r_{\infty,t} \right) \right]. \tag{54}$$

If $A_t = B_t = \emptyset$, both MESF rows are zero and their contribution is zero. At every other time covered by the theorem, both barcodes have positive total persistence, so the preceding bound applies.

Summing Equation (54) over the nonempty times $\mathcal{T}_+$ and dividing by $|\mathcal{T}|$ gives

$$d^h_{\mathrm{MESF}}(X, \widetilde{X}; \mathcal{T}) \le \frac{\ln 2}{|\mathcal{T}|} \sum_{t \in \mathcal{T}_+} r_{\infty,t} L_{\max,t} \times \left[\frac{\log_2 n_{\max,t}}{n_{\max,t}} - \frac{3}{2} \log_2 \left(\frac{3}{2} r_{\infty,t} \right) \right], \tag{55}$$

which is Equation (14). When $r_{\infty,t} = 0$, the corresponding term is understood by continuity and equals zero. Together with Equation (52), this proves Theorem 3.

Finally, suppose over the fixed finite domain $\mathcal{T}$ that reference diameters and nonzero total persistences are uniformly bounded away from zero and matching cardinalities are uniformly bounded. Then Equation (52) gives $r_{\infty,t} = O(\delta_P)$. Substitution into the logarithmic ESF bound yields

$$d^h_{\mathrm{MESF}}(X, \widetilde{X}) = O\left(\delta_P \log \frac{1}{\delta_P} \right) \quad \text{as } \delta_P \to 0, \tag{56}$$

which proves Equation (16).

The argument is stated on a fixed common time domain and therefore does not establish stability of the adaptive stopping decision. It also requires the reference diameters to remain uniformly bounded away from zero and does not assert robustness to arbitrary outlier contamination. Thus the result is a conditional diffusion-to-MESF stability statement rather than an unconditional end-to-end guarantee under coordinate perturbations.

D Computational Complexity

The method is defined using the exact diffusion distance in Equation (4). For computational acceleration, a spectral implementation may retain $r \ll n$ nonstationary modes,

$$D^2_{t,r}(x_i, x_j) = \mathrm{vol} \sum_{\ell \in \mathcal{I}_r} \lambda_\ell^{2t} \big(\psi_\ell(i) - \psi_\ell(j) \big)^2, \tag{57}$$

where $\mathcal{I}_r$ is the retained index set. Modes may, for example, be ranked by the magnitude of their time-dependent contribution $|\lambda_\ell|^t$. This truncation is an implementation approximation; the method and stability statements in the paper concern the exact distance.

Let m be the number of edges, d the ambient dimension, $T = |\mathcal{T}_X|$ the retained time count, and t_{cap} the terminal-time search horizon. The symmetric kNN graph has $O(nk)$ edges, and MST augmentation adds at most $n - 1$, so sparse graph storage remains $O(nk)$. Nearest-neighbor search and MST construction are implementation dependent. Once P is available, recursively updating $X^{(t)} = PX^{(t-1)}$ and evaluating the Dirichlet-energy trajectory costs $O(t_{\mathrm{cap}} md)$.

Let $C_{\mathrm{eig}}(m, n, r)$ denote the cost of computing the retained eigenpairs. After eigendecomposition, all pairwise spectral diffusion distances cost $O(n^2 r)$ per time and $O(Tn^2 r)$ over the retained axis. If $C_{\mathrm{VR}}(n, h_{\max})$ is the cost of one Vietoris–Rips computation through degree $h_{\max}$, persistence contributes $O(TC_{\mathrm{VR}}(n, h_{\max}))$. Finally, if N_t^h intervals belong to the finite barcode at (t, h), sampling the entropy summary functions on an N_u-point grid costs

$$O\left(N_u \sum_{t \in \mathcal{T}_X} \sum_{h \in \mathcal{H}} N_t^h\right). \tag{58}$$

This vectorization cost is typically lower order. The dominant spectral pipeline cost is therefore

$$O\big(C_{\mathrm{eig}}(m, n, r) + t_{\mathrm{cap}} md + Tn^2 r + TC_{\mathrm{VR}}(n, h_{\max})\big), \tag{59}$$

in addition to graph construction and MESF vectorization. Materializing a pairwise distance matrix requires $O(n^2)$ memory; the current implementation also uses dense pairwise intermediates. Repeated Rips persistence, whose simplex count can grow combinatorially in n and $h_{\max}$, is the principal scalability bottleneck. Sparse persistence, time subsampling, and truncated spectral representations are natural routes to larger point clouds.

E Additional Experimental Details

E.1 Persistence and representation protocol

All experiments use Python 3.11.1, and persistent homology is computed with GUDHI. The experiments were conducted on a system with a 32-core CPU and 64 GB of RAM, without GPU acceleration. At each reported diffusion time, we build the complete Vietoris–Rips filtration through the maximum pairwise dissimilarity, without quantile truncation or a sparse-Rips approximation. Intervals with infinite death and numerical zero-length intervals are removed before functional vectorization. For the synthetic effective-resistance reference, the dissimilarity is divided by its own diameter. For DIGIT, the single reference diameter Δ_X from $t = 2$ normalizes the retained diffusion-time sequence within each sample; no held-out label or cross-sample statistic enters this normalization.

The CATH experiments use $k = 5$, $H_0 \oplus H_1$, a seven-step moving-average window, Kneedle sensitivity 10, $t_{\mathrm{cap}} = 81$, and $N_u = 20$. PI uses Gaussian bandwidth 0.2 and a 20×20 grid per degree; PL retains five layers at 100 filtration coordinates; BC and single-diagram ESF use 100 coordinates. The ESF is not normalized by its maximum amplitude. All methods concatenate H_0 and H_1 features.

The ten predefined superfamily-separated folds are used as outer splits. In each split, an RBF SVM is selected only from the nine training folds by stratified five-fold cross-validation over $C \in \{0.1, 1, 10, 100\}$ and $\gamma \in \{0.001, 0.01, 0.1, 1\}$. Standardization is fitted within each inner training pipeline, accuracy is the selection criterion, and the random seed is 50. We report mean and population standard deviation over the ten outer test folds.

E.2 Training-only terminal-time selection

For outer fold f, let $\mathcal{I}_{-f}$ be the nine training folds. We use

$$T_f = \left\lfloor \frac{1}{|\mathcal{I}_{-f}|} \sum_{i \in \mathcal{I}_{-f}} t_{\max}(X_i) + \frac{1}{2} \right\rfloor. \tag{60}$$

Thus half-integers are rounded upward. The held-out fold is excluded from the mean and all fold statistics below. Once selected, the same T_f is used to compute both training and test representations directly on $\{2, \ldots, T_f\}$; no padding or time interpolation is performed.

Table 7: Training-only terminal-time statistics for `cath_3class_ca`. Fold identifiers follow the zero-based encoding in the released dataset. Mean, standard deviation, and median are computed only over the nine training folds; the standard deviation uses the training population as its denominator.

Test fold	Train	Test	Mean	Std.	Median	Selected T_f
0	15,266	1,696	15.168	2.131	15	15
1	15,266	1,696	15.186	2.107	15	15
2	15,266	1,696	15.165	2.158	15	15
3	15,266	1,696	15.165	2.139	15	15
4	15,266	1,696	15.172	2.164	15	15
5	15,266	1,696	15.183	2.137	15	15
6	15,266	1,696	15.177	2.099	15	15
7	15,266	1,696	15.162	2.140	15	15
8	15,265	1,697	15.147	2.154	15	15
9	15,265	1,697	15.163	2.170	15	15

Table 8: Training-only terminal-time statistics for `cath_10arch_ca`. The reporting conventions are identical to those in Table 7.

Test fold	Train	Test	Mean	Std.	Median	Selected T_f
0	6,399	711	15.406	1.989	16	15
1	6,399	711	15.418	1.983	16	15
2	6,399	711	15.400	1.996	16	15
3	6,399	711	15.407	1.979	16	15
4	6,399	711	15.415	1.974	16	15
5	6,399	711	15.412	1.980	16	15
6	6,399	711	15.383	1.988	16	15
7	6,399	711	15.390	1.991	16	15
8	6,399	711	15.406	1.984	16	15
9	6,399	711	15.433	1.989	16	15

Table 9: Training-only terminal-time statistics for `cath_20topo_ca`. The reporting conventions are identical to those in Table 7.

Test fold	Train	Test	Mean	Std.	Median	Selected T_f
0	3,817	423	14.805	2.343	15	15
1	3,816	424	14.786	2.372	15	15
2	3,816	424	14.791	2.371	15	15
3	3,816	424	14.781	2.392	15	15
4	3,816	424	14.772	2.369	15	15
5	3,816	424	14.697	2.356	15	15
6	3,816	424	14.726	2.375	15	15
7	3,816	424	14.797	2.344	15	15
8	3,816	424	14.747	2.410	15	15
9	3,815	425	14.696	2.408	15	15

The ranges across outer splits are 0.039, 0.050, and 0.109 diffusion steps for 3-Class, 10-Architecture, and 20-Topology, respectively. All thirty splits select $T_f = 15$. The complete-dataset descriptive sample-level means and standard deviations in Table 4 are not used for selection. Although the training median is 16 for `cath_10arch_ca`, the prespecified rule in Equation (60) rounds the training mean rather than the median. The agreement across splits indicates that the selected terminal time is insensitive to the held-out fold while preserving strict separation between parameter estimation and outer-fold evaluation.

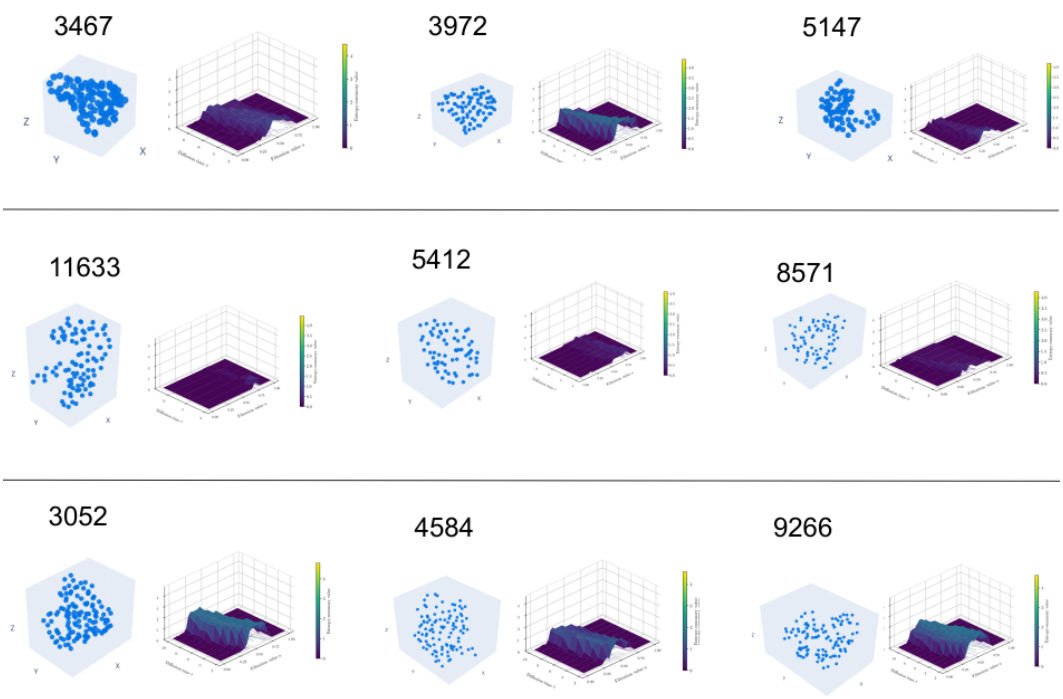


Figure 5: Representative visualizations of protein structures from the three CATH classes. Each row corresponds to one CATH class and contains three representative proteins. For each protein, the Cα point cloud (left) and its corresponding multiresolution topological representation (right) are shown.

Fig. 5 presents representative protein structures and their corresponding MESF representations for the three CATH classes.