Diameter-Normalized Significance Thresholds Cannot Separate Topological Signal from Sampling Noise

Semen A. Kurkin
Plekhanov Russian University of Economics
Moscow, Russia
0000-0002-3438-5717

Nikita M. Smirnov
Plekhanov Russian University of Economics
Moscow, Russia
0009-0002-7593-4345

Marina V. Khramova
Plekhanov Russian University of Economics
Moscow, Russia
0000-0002-6392-4580

Sofia A. Tkach
Plekhanov Russian University of Economics
Moscow, Russia

Zolzaya Dashdorj
Mongolian University of Science and Technology
Ulaanbaatar, Mongolia
0000-0001-6367-9039

2026 10th Scientific School Dynamics of Complex Networks and their Applications (DCNA) | 979-8-3195-2599-4/26/$31.00 ©2026 IEEE | DOI: 10.1109/DCNA72004.2026.11733405

***Abstract*—Persistent homology is increasingly applied to functional connectomes and other weighted networks, but the step that actually produces the reported Betti numbers — deciding which bars of a barcode are features and which are sampling noise — has no validated default. A widespread convention declares a bar significant when its persistence exceeds a fixed fraction of the diameter of the data. We show by measurement on synthetic objects with analytically known homology that this convention is not merely poorly tuned but unusable: on a sphere the longest *noise* bar in H_1 reaches 0.123 of the diameter, whereas on a torus the genuine H_2 bar reaches only 0.045 at the same sample size, so the signal and noise ranges overlap and no single constant can be correct on both objects. We describe a scale-free replacement based on the largest gap in the sorted persistence spectrum, calibrated to zero errors on 54 object–seed–noise combinations where the diameter rule makes nine, and, more importantly, we report its residual false-positive rate, which is 6.0% at $n = 200$ and 1.5% at $n = 400$ and does not vanish with sample size. We also show that random-graph null models cannot certify the absence of features at any density and should be read as measuring a false-positive rate rather than asserting zero.**

***Index Terms*—persistent homology, topological data analysis, dynamic networks, functional connectome, validation, false positives**

I. Introduction

Topological data analysis (TDA) has become a standard addition to the network analyst's toolbox, and in neuroscience in particular it is used to describe functional connectomes through cavities and cycles that classical graph metrics do not capture [1]–[3]. It thereby joins the broader set of quantitative methods — graph-theoretic, spectral, statistical and machine-learning — already applied to brain activity recorded with fMRI [4], [5], EEG [6], [7], MEG [8] and fNIRS [9], in pipelines that are increasingly expected to support decisions — from clinical diagnostics [10]–[13] to the monitoring of learning [14] and of skill acquisition [15], as well as the broader landscape of AI in medical applications [16], which raises the cost of any unvalidated step. The computational core of the method is well understood: a filtration of simplicial complexes is built over a growing scale parameter, and the birth and death of homology classes is summarized by a persistence diagram [17]–[19]. Efficient implementations are available [20], [21], and the diagram is stable under perturbation of the input [22], [23].

What is far less settled is the last step. A persistence diagram of a noisy sample contains hundreds of bars; converting it into the Betti numbers that appear in a results table requires a rule that separates features from sampling noise. This rule is not an implementation detail. It is where the answer is finally formed, and an error in it is indistinguishable from an error of the whole pipeline. Graph analysis of connectomes meets the same problem in another guise: the topology of thresholded fMRI functional networks switches between scale-free, random and small-world regimes as the connection-strength window is moved [4]. Principled statistical answers exist — confidence sets obtained by bootstrap [24], distance-to-a-measure and kernel-based robust inference [25], or vectorizations amenable to classical statistics [26], but they require a sampling model and considerable computation, and in practice a much simpler convention is used: a bar is significant if its persistence exceeds a fixed fraction of the diameter of the data.

This paper reports what happens when that convention is tested against objects whose homology is known in advance. Our contribution is threefold. First, we show that the diameter-

Authorized licensed use limited to: Kant Baltic Federal Univ. Downloaded on October 23,2026 at 15:25:45 UTC from IEEE Xplore. Restrictions apply.

TABLE I
TEST OBJECTS WITH ANALYTICALLY KNOWN HOMOLOGY OVER $\mathbb{Z}/2$

Object	β	Stresses	Measured
Circle S^1	(1,1,0)	baseline	(1,1,0)
Two circles	(2,2,0)	H_0: no premature merging	(2,2,0)
Figure eight	(1,2,0)	density drop at the crossing	(1,2,0)
Torus T^2	(1,2,1)	H_2 lives at the small radius	(1,2,1)
Sphere S^2	(1,0,1)	$\beta_1 = 0$: false positives	(1,0,1)
Klein bottle	(1,2,1)	answer depends on the field	(1,2,1)
Hopf link	(2,2,0)	components pass close by	(2,2,0)

normalized rule fails in *both* directions and that the failure is structural rather than a matter of tuning. Second, we give a scale-free replacement and its calibration. Third, and this is the part usually omitted, we measure the residual false-positive rate of the replacement and identify its mechanism, so that users can decide whether it is tolerable for their application.

II. Benchmark Construction and Verification Protocol

The evidence below rests on synthetic objects whose Betti numbers are known analytically. Three methodological commitments make such a benchmark informative rather than self-confirming.

Expected values are input, not output. The expected Betti numbers are fixed in a specification document before any measurement. When a run disagrees, the generator, the sampling scheme or the decision rule is repaired; the expectation table is never adjusted. This is the only arrangement under which a benchmark can falsify anything.

Two independent implementations. Diagrams are computed by `ripser` [20], and, in parallel, Betti numbers of the flag complex at a fixed threshold are computed by an independent reference implementation written for this project from boundary-matrix ranks over $\mathbb{Z}/2$, using no persistence library. Agreement is required between the two implementations *and* the specification. The reference implementation builds all cliques up to dimension $k_{\max}+1$, since without ∂_3 the cavities counted by β_2 are systematically overcounted.

Sampling by the correct measure. Points are drawn uniformly with respect to area, not with respect to the parameters of the embedding: rejection sampling proportional to $(R + r\cos\varphi)$ on the torus, Muller's method on the sphere. Naive parameter-uniform sampling looks plausible and silently biases the density, which is precisely the class of error the benchmark is meant to exclude.

Table I lists the seven point-cloud objects. They are chosen so that each stresses the pipeline differently; in particular the sphere is the control for false positives in H_1, and the torus is the control for false negatives in H_2. The last column is measured with the rule of Section IV.

The Hopf link is included as a geometric stress test, not as a separate topological signature: two linked circles have the homology of two disjoint ones, since linking is a property of the embedding rather than of the homotopy type.

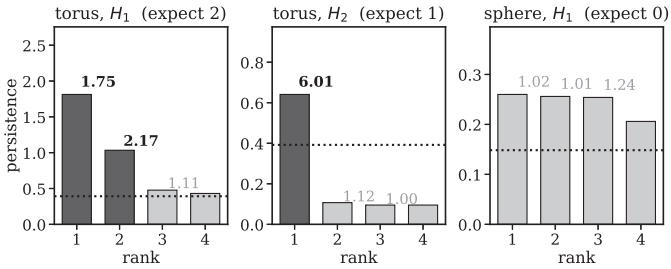


Fig. 1. Sorted persistence spectra (dark bars: genuine features). Numbers between bars are ratios of adjacent persistences; bold marks a ratio above the accepted threshold 1.55. Dotted line: the noise floor $0.07 \cdot \mathrm{diam}$. Torus at $N = 900$, sphere at $N = 400$. Left: the largest ratio sits at rank 2, giving $\beta_1 = 2$. Centre: only one bar survives the floor, and the step into the discarded bars gives $\beta_2 = 1$. Right: no ratio reaches the threshold and the rule correctly returns $\beta_1 = 0$.

III. The Diameter-Normalized Rule Cannot Work

The rule under test declares a bar significant when its persistence is at least $\rho \cdot \mathrm{diam}(X)$. It is attractive because it is scale-invariant and needs one constant. Both of its failure modes appear at $\rho = 0.15$ and at the same sample size, $N = 400$.

False negatives on anisotropic objects. On the torus with $R = 2.0$, $r = 0.8$, the cavity lives at the scale of the small radius, while the threshold is normalized by the diameter $2(R+r) = 5.6$. The genuine H_2 bar measures 0.257 against a threshold of 0.851 and is discarded. The Klein bottle behaves the same way.

False positives on compact objects. On the sphere the diameter is 2.12 and the threshold correspondingly small, 0.32, while the three longest noise bars in H_1 measure 0.260, 0.256 and 0.254. They sit immediately below the threshold, and on a subset of seeds the rule returns $(1, 1, 1)$ instead of $(1, 0, 1)$—a loop where none exists.

Lowering ρ repairs the torus and worsens the sphere; raising it does the reverse. The two demands are not merely in tension, they are contradictory, and the decisive numbers are these: the sphere's longest noise bar is 0.123 of the diameter, whereas the torus's genuine H_2 bar is 0.045 of the diameter—roughly three times *shorter* than the noise it must be distinguished from. Signal and noise ranges, expressed in units of diameter, overlap. No constant ρ can therefore be correct on both objects simultaneously, and the failure is not a question of tuning.

The underlying reason is that the diameter of a point cloud carries no information about the scale at which its homology lives. For an anisotropic object these two scales differ by a factor that depends on the object, not on the sampling.

IV. A Spectral-Gap Rule and Its Calibration

The property that does separate features from noise is not the absolute length of a bar but its position in the sorted spectrum of persistences. A genuine feature is followed by an abrupt drop; noise descends as an even staircase. Fig. 1 shows the three decisive cases.

Within each homological dimension the rule proceeds as follows. Persistences are sorted in decreasing order; bars below

Authorized licensed use limited to: Kant Baltic Federal Univ. Downloaded on October 23,2026 at 15:25:45 UTC from IEEE Xplore. Restrictions apply.

a noise floor of $0.07 \cdot \mathrm{diam}(X)$ are set aside as dust; the largest ratio between adjacent values is located, *including the step from the last survivor to the first discarded bar*; and β_k is the rank at which this ratio occurs, provided it exceeds 1.55. Otherwise no features are reported.

The step into the dust is not a technicality. Without it the rule always returns at least one feature, so the case $\beta_k = 0$ is unreachable and the sphere can never be answered correctly. The same step resolves the two disjoint circles, where the two surviving bars measure 1.473 and 1.453—a ratio of 1.01, with no gap between them—and the only informative gap is the one below them.

Calibration used 9 cases $\times$ 3 seeds $\times$ 2 noise levels $=$ 54 combinations. The rule is correct on all 54; the diameter-normalized rule makes nine errors on the same table. The two constants were taken from the interior of the region of zero errors, which extends over a floor of 0.05–0.10 and a gap threshold of 1.45–1.62, rather than from its edge; the accepted values are 0.07 and 1.55.

Relation to statistical significance methods

We do not claim that a gap heuristic is preferable to the statistical machinery cited in the introduction. Bootstrap confidence bands [24] and distance-to-a-measure inference [25] are better founded, and where a sampling model is available they should be used. Our claim is narrower and, we think, more immediately useful: a heuristic of this kind is what is in fact applied in most reported pipelines, usually implicitly and without any statement of its error rate, and the two contributions of this paper are to show that the most common variant of it is not repairable by tuning, and to attach a measured error rate to the variant we adopt. In the connectome setting the choice is additionally constrained, since resampling a subject's correlation matrix does not correspond to resampling from a fixed underlying distribution in any straightforward way.

Two limits of the rule should be stated explicitly. It is scale-free but not invariant under reparametrization of the filtration: a monotone map from similarity to distance preserves the structure of the diagram but distorts persistences, and the rule compares exactly those. And its constants are calibrated on point clouds; Section VII explains why they must not be inherited by connectome data.

V. The Residual False-Positive Rate

A benchmark that reports only successes is of limited use. The rule of Section IV has a non-zero false-positive rate which does not vanish with sample size, and we measured it on uniform clouds in $\mathbb{R}^3$, where the correct answer is $(1, 0, 0)$, over 200 trials per size (Table II).

The mechanism is specific and worth naming. When exactly one bar survives the noise floor, there is nothing to compare it against, and the lone survivor is significant by construction. Increasing n makes this configuration rarer but does not remove it. Treating such a rate as a primary result rather than a footnote is routine in event detection on neurophysiological recordings: in automatic seizure detection, rare events are sought against a heavy-tailed background, and the balance between recall and precision, not accuracy alone, decides whether a detector is usable [10], [11].

TABLE II
False positives of the spectral-gap rule on uniform clouds in $\mathbb{R}^3$ (200 trials, correct answer $(1, 0, 0)$)

n	Errors	Rate	95% upper bound
200	12/200	6.0%	10.2%
400	3/200	1.5%	4.3%

TABLE III
Flag-complex Betti numbers of Erdős–Rényi graphs $G(24, m)$; m edges, density $p = m/276$; five realizations per row

m	p	β_0	β_1	β_2	clean $(1, 0, 0)$
12	0.043	12–14	0–1	0	0/5
24	0.087	4–5	2–5	0	0/5
36	0.130	1–3	8–12	0	0/5
60	0.217	1	13–22	0	0/5
100	0.362	1	3–7	3–6	0/5
180	0.652	1	0	0–2	3/5
220	0.797	1	0	0	5/5

This measurement also carries a methodological lesson about calibration budgets. On three seeds the rule showed zero errors at $n = 200$; by the rule of three, zero out of three bounds the error rate only at 63%, which is no evidence at all. The sixty trials that would have been a comfortable budget for tuning a constant are inadequate for certifying a rate, and the threshold in our configuration was therefore set above the upper confidence bound of the measured rate rather than above an observed zero.

VI. Null Models Measure, They Do Not Certify

A common formulation of the control for false positives is that a null model must exhibit no significant long-lived bars. Taken literally this is unachievable, and the demonstration is elementary. Table III gives Betti numbers of the flag complex of Erdős–Rényi graphs on $n = 24$ nodes, computed by the reference implementation. Graphs are drawn in the uniform model $G(n, m)$: m edges are chosen without replacement among the $\binom{24}{2} = 276$ possible pairs, so that $p = m/276$ is the resulting edge density. Five realizations are reported per value of m, as ranges rather than means, because a null model has no typical value.

Below the connectivity threshold, which for $n = 24$ sits at $p \approx \ln n/n = 0.13$, that is at $m \approx 36$, the graph is fragmented; above it the graph is connected but carries eight to twenty-two independent cycles; in between, β_2 appears as well. The answer becomes reliably clean only at $p \gtrsim 0.8$, on an almost complete graph that controls nothing. There is no density window in which the null model is topologically empty.

The conclusion is not that the null model is badly specified but that its ground truth is of a different kind. A random

Authorized licensed use limited to: Kant Baltic Federal Univ. Downloaded on October 23,2026 at 15:25:45 UTC from IEEE Xplore. Restrictions apply.

graph genuinely has cycles; they are not an artifact to be thresholded away, they are the null distribution against which everything else is measured. The verifiable ground truth of such a dataset is therefore that no event schedule exists and the event log is empty—a statement about the construction—while the observed Betti numbers are a measurement of the false-positive rate of whatever detector is applied downstream.

VII. Limitations and Transfer

The constants of Section IV are calibrated on point clouds in $\mathbb{R}^3$ and $\mathbb{R}^4$ sampled from manifolds. They must not be inherited by functional connectomes, for two reasons: the distribution of pairwise distances obtained from a correlation matrix is different, and, more fundamentally, there is no known β_k against which a cut could be calibrated on real data. There significance has to be established differently—by permutation tests and comparison against an ensemble of random graphs—and the null models discussed above are the natural instrument for that, precisely in their measuring rather than certifying role.

Two further caveats apply to any benchmark of this kind. Betti numbers of a noisy sample are recoverable only as persistent features, never as the homology of a fixed complex; and the Vietoris–Rips complex of a circle is homotopy equivalent to $S^1, S^3, S^5, \dots$ as the scale grows [27], so a loop that disappears at large scale may be a regime change rather than a defect. Neither is a bug, and both are easily mistaken for one.

Finally, a caution specific to dynamic networks, which is the setting this benchmark was built for. Ground truth for a temporal sequence is usually validated by comparing a Betti curve with a schedule of planned events. A Betti curve is blind to *pairing*: if the birth–death pairing in the ground-truth log is inverted, the curve is unchanged, every consistency test passes, and the error surfaces only later as an apparent defect of the zigzag algorithm [28] being validated. Pairing must be verified algebraically, not through the Betti curve.

VIII. Conclusion

The step that converts a barcode into Betti numbers deserves the same scrutiny as the homology computation itself. We have shown by measurement on objects with known topology that the common diameter-normalized significance threshold cannot be made to work, because on our test set the noise range and the signal range overlap in units of diameter; that a spectral-gap rule with an explicit noise floor and a step into the discarded bars is correct on all 54 calibration combinations where the former makes nine errors; and that this replacement nevertheless retains a false-positive rate of 6.0% at $n = 200$ and 1.5% at $n = 400$, with an identifiable mechanism. We regard the last of these as the most useful result: a decision rule whose error rate is unmeasured is not a validated rule, and the budgets normally used to tune such constants are an order of magnitude too small to certify a rate. This matters most where topological summaries are passed on as features to classifiers, as graph measures of functional networks already are [5], [7], [12], [13]: an unmeasured false-positive rate of the feature extractor then becomes an unmeasured error of the diagnosis.

Acknowledgment

The study was funded by Russian Science Foundation (grant 26-41-03005).

References

[1] G. Petri, P. Expert, F. Turkheimer, R. Carhart-Harris, D. Nutt, P. J. Hellyer, and F. Vaccarino, "Homological scaffolds of brain functional networks," *Journal of the Royal Society Interface*, vol. 11, no. 101, p. 20140873, 2014.

[2] C. Giusti, E. Pastalkova, C. Curto, and V. Itskov, "Clique topology reveals intrinsic geometric structure in neural correlations," *Proceedings of the National Academy of Sciences*, vol. 112, no. 44, pp. 13 455–13 460, 2015.

[3] A. E. Sizemore, C. Giusti, A. Kahn, J. M. Vettel, R. F. Betzel, and D. S. Bassett, "Cliques and cavities in the human connectome," *Journal of Computational Neuroscience*, vol. 44, no. 1, pp. 115–145, 2018.

[4] A. N. Pisarchik, A. V. Andreev, S. A. Kurkin, D. Stoyanov, A. A. Badarin, R. Paunova, and A. E. Hramov, "Topology switching during window thresholding fMRI-based functional networks of patients with major depressive disorder: Consensus network approach," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 33, no. 9, p. 093122, 2023.

[5] A. V. Andreev, S. A. Kurkin, D. Stoyanov, A. A. Badarin, R. Paunova, and A. E. Hramov, "Toward interpretability of machine learning methods for the classification of patients with major depressive disorder based on functional network measures," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 33, no. 6, p. 063140, 2023.

[6] M. V. Khramova, A. K. Kuc, V. A. Maksimenko, N. S. Frolov, V. V. Grubov, S. A. Kurkin, A. N. Pisarchik, N. N. Shusharina, A. A. Fedorov, and A. E. Hramov, "Monitoring the cortical activity of children and adults during cognitive task completion," *Sensors*, vol. 21, no. 18, p. 6021, 2021.

[7] M. S. Kabir, S. Kurkin, G. Portnova, O. Martynova, Z. Wang, and A. Hramov, "Contrastive machine learning reveals in EEG resting-state network salient features specific to autism spectrum disorder," *Chaos, Solitons & Fractals*, vol. 185, p. 115123, 2024.

[8] S. Kurkin, A. Hramov, P. Chholak, and A. Pisarchik, "Localizing oscillatory sources in a brain by MEG data during cognitive activity," in *Proc. 2020 4th International Conference on Computational Intelligence and Networks (CINE)*, Kolkata, India, 2020, pp. 1–4.

[9] A. A. Badarin, V. M. Antipov, V. V. Grubov, A. V. Andreev, E. N. Pitsik, S. A. Kurkin, and A. E. Hramov, "Brain compensatory mechanisms during the prolonged cognitive task: fNIRS and eye-tracking study," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 17, no. 2, pp. 303–314, 2025.

[10] O. E. Karpov, V. V. Grubov, V. A. Maksimenko, S. A. Kurkin, N. M. Smirnov, N. P. Utyashev, D. A. Andrikov, N. N. Shusharina, and A. E. Hramov, "Extreme value theory inspires explainable machine learning approach for seizure detection," *Scientific Reports*, vol. 12, p. 11474, 2022.

[11] V. V. Grubov, S. I. Nazarikov, S. A. Kurkin, N. P. Utyashev, D. A. Andrikov, O. E. Karpov, and A. E. Hramov, "Two-stage approach with combination of outlier detection method and deep learning enhances automatic epileptic seizure detection," *IEEE Access*, vol. 12, pp. 122 168–122 182, 2024.

[12] S. Zhang, H. Zhao, W. Wang, Z. Wang, X. Luo, A. Hramov, and J. Kurths, "Edge-centric effective connection network based on multimodal MRI for the diagnosis of Alzheimer's disease," *Neurocomputing*, vol. 552, p. 126512, 2023.

[13] W. Wang, H. Zhao, C. He, Y. Cui, Z. Wang, A. Hramov, P. Luan, X. Luo, J. Ouyang, and J. Kurths, "Target localization intervention and prognosis evaluation for an individual with mild cognitive impairment," *Brain-X*, vol. 1, p. e25, 2023.

[14] S. A. Kurkin, V. V. Grubov, V. A. Maksimenko, E. N. Pitsik, M. V. Khramova, and A. E. Hramov, "System for monitoring and adjusting the learning process of primary schoolchildren based on the EEG data analysis," *Informatsionno-upravliaiushchie sistemy [Information and Control Systems]*, no. 5, pp. 50–61, 2020.

Authorized licensed use limited to: Kant Baltic Federal Univ. Downloaded on October 23,2026 at 15:25:45 UTC from IEEE Xplore. Restrictions apply.

[15] A. Badarin, V. Antipov, V. Grubov, N. Grigorev, A. Savosenkov, A. Udoratina, S. Gordleeva, S. Kurkin, V. Kazantsev, and A. Hramov, "Psychophysiological parameters predict the performance of naive subjects in sport shooting training," *Sensors*, vol. 23, no. 6, p. 3160, 2023.
[16] O. E. Karpov, E. N. Pitsik, S. A. Kurkin, V. A. Maksimenko, A. V. Gusev, N. N. Shusharina, and A. E. Hramov, "Analysis of publication activity and research trends in the field of AI medical applications: Network approach," *International Journal of Environmental Research and Public Health*, vol. 20, p. 5335, 2023.
[17] H. Edelsbrunner, D. Letscher, and A. Zomorodian, "Topological persistence and simplification," *Discrete & Computational Geometry*, vol. 28, no. 4, pp. 511–533, 2002.
[18] A. Zomorodian and G. Carlsson, "Computing persistent homology," *Discrete & Computational Geometry*, vol. 33, no. 2, pp. 249–274, 2005.
[19] G. Carlsson, "Topology and data," *Bulletin of the American Mathematical Society*, vol. 46, no. 2, pp. 255–308, 2009.
[20] U. Bauer, "Ripser: efficient computation of Vietoris–Rips persistence barcodes," *Journal of Applied and Computational Topology*, vol. 5, pp. 391–423, 2021.
[21] N. Otter, M. A. Porter, U. Tillmann, P. Grindrod, and H. A. Harrington, "A roadmap for the computation of persistent homology," *EPJ Data Science*, vol. 6, p. 17, 2017.
[22] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer, "Stability of persistence diagrams," *Discrete & Computational Geometry*, vol. 37, no. 1, pp. 103–120, 2007.
[23] F. Chazal, D. Cohen-Steiner, L. J. Guibas, F. Mémoli, and S. Y. Oudot, "Gromov–Hausdorff stable signatures for shapes using persistence," *Computer Graphics Forum*, vol. 28, no. 5, pp. 1393–1403, 2009.
[24] B. T. Fasy, F. Lecci, A. Rinaldo, L. Wasserman, S. Balakrishnan, and A. Singh, "Confidence sets for persistence diagrams," *The Annals of Statistics*, vol. 42, no. 6, pp. 2301–2339, 2014.
[25] F. Chazal, B. Fasy, F. Lecci, B. Michel, A. Rinaldo, and L. Wasserman, "Robust topological inference: distance to a measure and kernel distance," *Journal of Machine Learning Research*, vol. 18, no. 159, pp. 1–40, 2017.
[26] P. Bubenik, "Statistical topological data analysis using persistence landscapes," *Journal of Machine Learning Research*, vol. 16, pp. 77–102, 2015.
[27] M. Adamaszek and H. Adams, "The Vietoris–Rips complexes of a circle," *Pacific Journal of Mathematics*, vol. 290, no. 1, pp. 1–40, 2017.
[28] G. Carlsson and V. de Silva, "Zigzag persistence," *Foundations of Computational Mathematics*, vol. 10, no. 4, pp. 367–405, 2010.

Authorized licensed use limited to: Kant Baltic Federal Univ. Downloaded on October 23,2026 at 15:25:45 UTC from IEEE Xplore. Restrictions apply.