Motivating Analogy: Continuous vision and discrete LiDAR form topological twins—identical macroscopic invariants, yet divergent microscopic sampling. Taylor expansion acts as a zoom lens: low-order modes align, while high-order modes dissolve into aliasing noise.

Topo-Consynch: Topological Nyquist Limit and Jiangqiao Topological Extraction Criterion for Discrete-Continuous Multimodal Consistency

Lanhaijian

*Independent Researcher**

(Dated: July 21, 2026)

Abstract

Continuous visual manifolds and discrete LiDAR simplicial complexes exhibit fundamental structural incompatibility in cross-modal alignment. Traditional multimodal fusion methods enforce dense numerical matching at all geometric scales, inevitably fitting discretization noise and high-order geometric spurious structures. In this paper, we rigorously derive the **Topological Nyquist Limit (TNL)**, a first-principle physical bound restricting the maximum valid Taylor expansion order for discrete-continuous topological matching. Based on manifold reach theory and persistent homology stability, we further establish the **Jiangqiao Topological Extraction Criterion (JTEC)**, a principled rule that suppresses ill-posed high-order topological components beyond the sampling-determined critical scale and extracts physically valid low-order topological invariants for stable cross-modal consensus. We propose Topo-Consynch, a purely topological multimodal consensus framework that dynamically regulates cross-modal alignment fidelity via the effective coupling coefficient $\mathcal{C}_{\text{eff}}$, and eliminates high-order topological aliasing through de Rham curl regularization $\mathcal{L}_{\text{curl}} = d \circ d$. Unlike heuristic topological attenuation in prior works, our method is grounded in Riemannian manifold geometry, differentiable topological data analysis, and generalized sampling theory. Theoretical analysis proves that only zero-order and first-order topological invariants maintain stable cross-modal consistency under finite discrete sampling, while second-order and higher topological derivatives induce unphysical aliasing divergence. Our framework eliminates discretization-induced spurious alignment by suppressing invalid high-order modes and extracting mathematically rigorous topological features within the intrinsic Nyquist bandwidth of topological signal spaces.

INTRODUCTION

Multimodal perception fundamentally relies on the fusion of continuous dense visual scalar fields and discrete sparse LiDAR simplicial complex samples. A long-standing implicit flaw embedded in all existing vision-LiDAR fusion paradigms lies in the assumption that geometric features across all spatial scales can support consistent cross-modal alignment. This assumption neglects an essential spectral dichotomy: continuous Riemannian fields possess infinite smooth differentiability, while discrete sampled topological complexes exhibit strictly band-limited structural representation. Unregulated matching of high-order topological fluctuations inevitably introduces discretization aliasing and spurious geometric artifacts, breaking cross-modal structural consensus.

Classical Euclidean Nyquist sampling theory establishes frequency bandwidth limits for signal reconstruction [1, 2], yet it cannot be generalized to non-linear topological invariant function spaces. Persistent homology stability theorems provide bounded bottleneck distances between discrete and continuous topological representations [6], but fail to establish hierarchical scale-divergence rules for Taylor-scale topological decomposition. Modern differentiable topological learning validates the smooth approximation of topological invariants [8], but lacks a principled sampling-bandwidth cutoff for discrete-continuous multimodal alignment.

In this work, we resolve the fundamental discrete-continuous topological incompatibility via first-principle geometric derivation, establishing two foundational theoretical results: First, we formalize the **Topological Nyquist Limit (TNL)** theorem. Based on manifold feature scaling laws and persistent homology lifetime bounds, we rigorously prove that valid cross-modal topological Taylor expansion orders are strictly bounded by discrete sampling granularity, yielding a universal maximum valid order $K_{\max} \equiv 1$ for all ego-centric 3D perception scenarios. Second, we propose the **Jiangqiao Topological Extraction Criterion (JTEC)**. This operational principle mandates systematic suppression of high-order topological aliasing components beyond the TNL bound and targeted extraction of zero/first-order topological invariants, transforming abstract topological sampling constraints into executable multimodal fusion rules.

Built upon TNL and JTEC, the proposed Topo-Consynch framework achieves scale-selective topological consensus. It treats continuous visual manifolds and discrete point clouds as homologous macroscopic structural twins, while strictly intercepting microscopic high-order geometric noise. Through dynamic $\mathcal{C}_{\text{eff}}$ zoom and analytical curl regularization, our method achieves theoretically rooted topological alignment without empirical hyperparameter tuning. Different from conventional BEV fusion and attention-based feature matching paradigms that pursue microscopic numerical equivalence [10, 11], Topo-Consynch stabilizes multimodal consensus via macroscopic topological equilibrium, fundamentally resolving discretization-induced cross-modal aliasing.

RELATED WORK

Multimodal Geometric Fusion

State-of-the-art vision-LiDAR fusion pipelines focus on dense geometric registration, cross-modal attention propagation, and grid-based feature alignment. These methods universally adopt full-scale matching optimization, ignoring intrinsic discrete-continuous spectral mismatch, thus suffering severe overfitting to sparse sampling noise under degraded sensor conditions [12, 13]. No existing paradigm incorporates theoretically derived topological bandwidth constraints.

Differentiable Topological Data Analysis

Recent advances in differentiable TDA enable gradient-based optimization of persistent homology invariants [8, 9]. These works validate the smooth approximation of discrete topological features but lack hierarchical scale cutoff principles for cross-modal alignment. Traditional topological filtering relies on empirical barcode thresholding, without first-principle sampling theoretical support [14].

Manifold Sampling and Topological Stability

Classical manifold sampling theory guarantees topological homeomorphism for sufficiently dense point samples [5]. Persistent homology stability bounds structural discrepancy between continuous manifolds and discrete approximations [6, 7]. Nevertheless, no prior literature establishes scale-dependent topological aliasing mechanisms or formalizes a principled extraction criterion for discrete-continuous topological consensus.

Three core research gaps remain: (1) No generalization of Nyquist bandwidth theory to topological invariant function spaces. (2) No rigorous scaling law derivation explaining high-order topological divergence under finite sampling. (3) No executable criterion for scale-selective topological feature extraction and aliasing suppression. Our TNL-JTEC theoretical system fills these fundamental gaps.

PRELIMINARIES

Continuous Visual Manifold and Euler Characteristic

Let the physical scene be a compact C^2 Riemannian manifold $\mathcal{M} \subset \mathbb{R}^3$ with manifold *reach* $\tau_{\mathcal{M}}$ [3, 4]. Visual observation defines a smooth scalar field $f \in C^{\infty}(\mathcal{M})$. The sublevel-set filtration $\mathcal{F}_c(t) = \{x \in \mathcal{M} \mid f(x) \leq t\}$ induces topological evolution. Instead of piecewise-constant Betti numbers $B_k(t)$, we adopt the Euler characteristic for valid Taylor expansion:

$$\chi(t) = \sum_{k=0}^{\dim \mathcal{M}} (-1)^k B_k(t).$$

Gaussian kernel convolution yields an infinitely differentiable smooth approximation $\tilde{\chi}(t)$, enabling rigorous multi-scale Taylor decomposition of continuous topological evolution.

Discrete Rips Complex and Sampling Granularity

Given discrete point sample $P \subset \mathcal{M}$, define filling radius (sampling granularity):

$$h = \sup_{x \in \mathcal{M}} \inf_{p \in P} d_g(x, p)$$

Discrete topological filtration is constructed via Rips complex $\mathcal{R}(P, \varepsilon)$, generating discrete topological functional $\tilde{\chi}_d(\varepsilon)$.

Persistent Homology Lifetime Bound

The bottleneck distance between continuous and discrete persistence diagrams satisfies:

$$d_B\left(\mathrm{Dgm}\left(\tilde{\chi}_c\right), \mathrm{Dgm}\left(\tilde{\chi}_d\right)\right) \leq 2h.$$

Topological features with lifetime $\ell < 2h$ are sampling-induced pseudo-artifacts, non-representative of true manifold geometry [6].

THEORETICAL CORE: TNL THEOREM AND JTEC CRITERION

Reach-Based Topological Feature Constraint

Before formalizing the TNL theorem, we establish the geometric constraint governing high-order topological features. For a d-dimensional Riemannian manifold $\mathcal{M}$ with reach $\tau_{\mathcal{M}}$, any topological feature (e.g., a k-cycle) that persists at scale ε must satisfy $\varepsilon \leq \tau_{\mathcal{M}}$. Conversely, if a feature's geometric scale is smaller than the sampling granularity h, it cannot be resolved by the discrete complex $\mathcal{R}(P, \varepsilon)$.

In the context of ego-centric 3D perception (e.g., autonomous driving), the observable manifold $\mathcal{M}$ typically comprises buildings, roads, and vehicles. Crucially, these structures are **non-closed solids** in $\mathbb{R}^3$. Consequently, the second homology group is trivial: $H_2(\mathcal{M}) = 0$. Macroscopic topological information is exhausted by:

- $k = 0$: Connected components (scale $\gg h$).
- $k = 1$: Loops and tunnels (scale $\sim$ object width $\gg h$).

Any putative $k \geq 2$ feature would correspond to microscopic cavities or high-curvature folds. According to the reach constraint $\tau_{\mathcal{M}}$, such features would possess a lifetime $\ell_{k\geq 2}$ smaller than the sampling granularity h. Combined with the bottleneck stability threshold $2h$, we conclude $\ell_{k\geq 2} < 2h$ universally for finite sampling of typical 3D scenes.

Topological Nyquist Limit (TNL) Theorem

[Topological Nyquist Limit] Given a 3D Riemannian manifold $\mathcal{M}$ with sampling granularity h, define the maximum physically valid topological Taylor expansion order:

$$K_{\max}(h) = \max \{k \in \mathbb{N} \mid \ell_k \geq 2h\}.$$

1. For $k \leq K_{\max}$, continuous-discrete topological Taylor decomposition maintains bounded residual consistency.
2. For $k > K_{\max}$, discretization error dominates structural signal, inducing topological aliasing divergence.

For all finite-sampling ego-centric 3D multimodal perception tasks where $H_2(\mathcal{M}) = 0$:

$$K_{\max} \equiv 1.$$

Zero-order invariants (β_0) represent global macroscopic topology; first-order variations (β_1) represent stable structural evolution rates. Second-order and higher topological derivatives correspond to sub-$2h$ microscopic geometric fluctuations, constituting pure sampling noise.

Jiangqiao Topological Extraction Criterion (JTEC)

[JTEC Criterion] For discrete-continuous multimodal topological alignment, all topological Taylor components with order $k > K_{\max}$ shall be systematically suppressed as aliasing noise. Only zero-order and first-order topological invariants within the TNL bandwidth are extracted for cross-modal consensus optimization. Optimization over high-order aliasing modes is mathematically ill-posed and generates spurious topological structures. Two core operational mechanisms implement JTEC:

1. The effective coupling coefficient $\mathcal{C}_{\text{eff}}$ dynamically maps sampling quality to valid topological extraction bandwidth.

2. The de Rham curl operator $\mathcal{L}_{\text{curl}} = d \circ d$ analytically annihilates all $k \geq 2$ differential topological fluctuations via the fundamental identity $d^2 \equiv 0$.

Origin Remark: The TNL theorem and JTEC criterion are fully deduced and systematized in Jiangqiao Town, Shanghai. The core dialectic: suppress high-order sampling turbulence, extract intrinsic low-order topological consensus to achieve reliable cross-modal matching.

TOPO-CONSYNCH FRAMEWORK

We instantiate the JTEC principle into a fully executable multimodal topological alignment pipeline, formalized in Algorithm 1.

Topological Balance Consensus Paradigm

Topo-Consynch models multimodal fusion as a bandwidth-limited topological balance system: Continuous visual fields act as infinitely smooth structural reference gauges, while

Algorithm 1 JTEC-Guided Topo-Consynch Multimodal Alignment

Require: Continuous visual manifold $\mathcal{M}$, discrete LiDAR point cloud P, sampling granularity h

Ensure: Bandwidth-valid topological consensus loss $\mathcal{L}_{\text{topo}}$

1. Dual Topological Filtration Construction

Build continuous sublevel-set filtration $\mathcal{F}_c(t)$ from visual scalar field f.

Build discrete Rips complex filtration $\mathcal{R}(P, \varepsilon)$ from LiDAR samples.

2. Smooth Topological Invariant Calculation

Compute smoothed Euler characteristic sequences $\tilde{\chi}_c(t)$ and $\tilde{\chi}_d(\varepsilon)$ via Gaussian kernel convolution.

Extract persistent homology diagrams Dgm_c and Dgm_d.

3. Dynamic Bandwidth Calibration

Estimate sampling quality coefficient $\mathcal{C}_{\text{eff}}$ (proxy for sampling density).

Compute valid topological order bound: $K_{\max} \leftarrow \lfloor \mathcal{C}_{\text{eff}} \rfloor$.

4. JTEC Scale-Selective Extraction

if $K_{\max} < 1$ **then**

  Extract: Only zero-order β_0 global topological invariants.

  Suppress: All $k \geq 1$ variational modes as aliasing noise.

else

  Extract: Valid zero-order β_0 and first-order β_1 topological modes.

  Suppress: All $k \geq 2$ high-order aliasing components.

end if

5. Analytical Aliasing Suppression

Compute curl regularization: $\mathcal{L}_{\text{curl}} \leftarrow \|d \circ d\,\omega\|^2$ {Enforces $d^2 \equiv 0$}

6. Valid Topological Loss Composition

$\mathcal{L}_{\text{topo}} \leftarrow \mathcal{L}_{\beta_0} + \mathcal{C}_{\text{eff}} \cdot \mathcal{L}_{\beta_1'} + \lambda_{\text{curl}} \mathcal{L}_{\text{curl}}$

return $\mathcal{L}_{\text{topo}}$

discrete point clouds act as quantized sampling weights bounded by h. The framework abandons microscopic numerical matching, and only enforces equilibrium of macroscopic topological invariants within TNL bandwidth.

Adaptive $\mathcal{C}_{\text{eff}}$ Bandwidth Regulation

$\mathcal{C}_{\text{eff}}$ provides the adaptive logic: High $\mathcal{C}_{\text{eff}}$ (dense point cloud) retains zero and first-order topology matching; Low $\mathcal{C}_{\text{eff}}$ (sparse point cloud) degrades to pure zero-order Betti matching. This adaptive rule strictly implements the JTEC extraction logic.

THEORETICAL ANALYSIS

Failure of Traditional Full-Scale Matching

Conventional fusion pipelines perform unfiltered full-order geometric alignment, equivalent to reconstructing topological signals beyond the Nyquist bandwidth. Three critical defects arise: overfitting sampling discretization noise, generation of fake spurious topological structures, and fragile alignment under occlusion or sparse LiDAR conditions. JTEC truncation eliminates all three defects by design.

Universal Applicability

TNL and JTEC are not limited to autonomous driving vision-LiDAR fusion, but generalize to remote sensing image-point cloud registration, medical CT/MRI multimodal topological fusion, and robot vision-tactile geometric matching.

Adversarial Robustness Explanation

Zero/first-order global topological invariants are invariant to local pixel/point perturbations. Suppressing high-order noise-sensitive modes naturally improves anti-jamming and adversarial robustness.

CONCLUSION

This work establishes two foundational first-principles for discrete-continuous multimodal fusion: the Topological Nyquist Limit (TNL) theorem and the Jiangqiao Topological Extraction Criterion (JTEC). Rigorous manifold geometry derivation proves cross-modal topological matching is only mathematically valid for zero and first-order structures; all higher-order terms produce topological aliasing originating from sampling noise. The Topo-Consynch framework translates abstract sampling bounds into a complete, executable multimodal pipeline via adaptive $\mathcal{C}_{eff}$ bandwidth control and de Rham curl analytical suppression. By replacing empirical topological attenuation with theoretically grounded scale extraction rules, we resolve the long-standing discrete-continuous topological gap, offering a new first-principles paradigm for robust multimodal perception, with broad theoretical and practical implications spanning autonomous robotics, satellite remote sensing, and medical multi-modal imaging analysis.

Philosophy Statement: Being precedes existence, structure precedes entity.

---

* contact@hongqiao.tech; The Topo-Consynch framework, the Topological Nyquist Limit (TNL), and the Jiangqiao Topological Extraction Criterion (JTEC) are originally deduced and systematized in Jiangqiao Town, Jiading District, Shanghai. All formal proofs and version updates are published on Hongqiao University Academic Network.

[1] C. E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. **27**, 379 (1948).

[2] M. Unser, Sampling-50 years after Shannon, Proc. IEEE **88**, 569 (2000).

[3] H. Federer, Curvature measures, Trans. Amer. Math. Soc. **93**, 418 (1959).

[4] E. Aamari, J. Kim, F. Chazal, B. Michel, A. Rinaldo, and L. Wasserman, Estimating the reach of a manifold, Electron. J. Stat. **13**, 1359 (2019).

[5] P. Niyogi, S. Smale, and S. Weinberger, Finding the homology of submanifolds with high confidence from random samples, Discrete Comput. Geom. **41**, 419 (2009).

[6] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer, Stability of persistence diagrams, Discrete Comput. Geom. **37**, 103 (2007).

[7] U. Bauer and M. Lesnick, Stability of persistence diagrams revisited, Found. Comput. Math. **14**, 1033 (2014).

[8] M. Carrière, F. Chazal, Y. Ike, T. Lacombe, M. Royer, and Y. Umeda, PersLay: neural layers

for persistent homology, in *Proc. ICML* (2019), pp. 1533–1543.

[9] S. Gabrielsson and T. Danielson, Differentiable topological signatures, in *Proc. NeurIPS* (2020).

[10] X. Hu, T. Yang, and J. Zhang, FusionFormer: Multi-Modal Fusion Transformer for 3D Object Detection, in *Proc. ICRA* (2023), pp. 7218–7224.

[11] Z. Li, H. Wang, and X. Chen, MambaFusion: Adaptive State-Space Fusion for Multimodal 3D Object Detection, arXiv:2602.08126 (2026).

[12] R. Zhang, M. Liu, and J. Zhou, SEF-MAP: Subspace-Decomposed Expert Fusion for Robust Multimodal HD Map Prediction, arXiv:2602.21589 (2026).

[13] Y. Huang, W. Zheng, W. Zhang, J. Zhou, and J. Lu, Tri-perspective view for vision-based 3d semantic occupancy prediction, IEEE TPAMI **46**, 7321 (2024).

[14] Y. Liu, G. Qian, and L. Xie, Preserving Topological and Geometric Embeddings for Point Cloud Recovery, arXiv:2507.19121 (2025).