ON THE TOPOLOGICAL SCULPTING FOR VESSEL-LIKE STRUCTURE SEGMENTATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Thin 3D structures such as vessels, airways, and other branching tubular anatomies demand not only voxel-level accuracy but also correct global connectivity. Standard overlap-driven objectives often tolerate small breaks or spurious connections that severely damage topology, while many topology-aware losses are local, patch-based, or only influence training and thus cannot reliably repair errors at inference time. We present TopoSculpt, a generic refinement framework that improves the topology of predicted 3D tubular segmentations by explicitly enforcing global topological priors. TopoSculpt introduces a Topological Integrity Betti (TIB) loss that combines (i) Betti-number guidance toward a desired topology and (ii) an integrity regularizer that discourages destructive deviations from the initial prediction. To make persistent-homology computations tractable, we propose a curriculum refinement strategy that progressively transitions from coarse to fine corrections with adaptive sampling. Experiments on three challenging 3D benchmarks (pulmonary airway, Circle of Willis, and coronary artery) show consistent gains in both geometric accuracy and topology fidelity, reducing large connectivity errors by an order of magnitude and improving tree/branch detection rates. These results indicate that TopoSculpt is an effective, broadly applicable mechanism for topology-preserving refinement of 3D thin-structure segmentation.

1 INTRODUCTION

Segmenting thin, branching 3D structures is a persistent challenge in computer vision, where the goal is not only accurate delineation but also topological completeness of the prediction. For tubular and filamentary objects, a small local error can disconnect an entire branch, while a thin spurious bridge can create an incorrect shortcut; both substantially alter the implied graph and can invalidate downstream analysis. Medical tubular anatomies, such as pulmonary airways and cerebral/coronary vasculature, provide particularly demanding testbeds because they form tree- or graph-like networks with strong cross-subject structural regularities and clinically meaningful connectivity (Zhang et al., 2023; Yang et al., 2024; Li et al., 2025; Gu et al., 2022). A central difficulty is that prevailing objectives and metrics for segmentation emphasize pixel/voxel level overlap. However, modern 3D backbones (Çiçek et al., 2016; Milletari et al., 2016) achieve strong overlap scores even when predictions contain disconnections, missing terminal branches, or unintended merges that are critical for thin structures. Therefore, a critical issue is raised:

"Can we preserve the topology of shapes during the model training and inference phases?"

This motivates topology-sensitive formulations that explicitly account for connectivity, beyond what overlap alone can capture. Recent work has explored topology-aware learning to reduce such failures by considering the explicit loss functions and network design (Hu et al., 2022; Zhou et al., 2025; Wu et al., 2024; Qi et al., 2023), yet practical gaps remain for 3D thin-structure segmentation: topology constraints are often applied only during training, providing limited capability to repair case-specific topological errors at inference time. In this paper, we propose a novel framework, TopoSculpt, for the topological refinement of 3D fine-grained tubular structure modeling, as demonstrated in Figure 1. The contributions of TopoSculpt are threefold:

First, TopoSculpt introduces a novel topological constraint that combines Betti number guidance with topological integrity preservation, referred to as Topological Integrity Betti (TIB). TIB is en-

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.

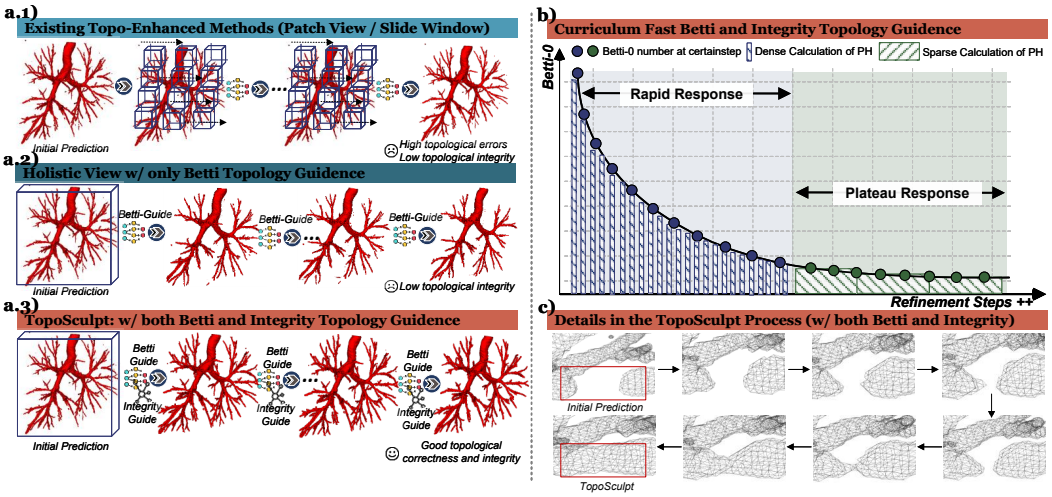

Figure 1: Overview of TopoSculpt. (a.1) Patch-wise methods lack global context and cannot guarantee topological integrity. (a.2) Holistic modeling captures the complete structure, but Betti guidance alone may compromise global integrity. (a.3) Combining Betti guidance with an integrity constraint improves local correctness while preserving global topology. (b) Curriculum refinement uses dense persistent-homology calculations to correct large-scale errors, then sparser sampling to address finer-scale errors. (c) Betti and integrity objectives jointly supervise refinement.

forced on the basis of the holistic-view modeling strategy, since only by considering the entire structure can its global topology be accurately computed and preserved. After $\mathcal{F}$ is trained, it generates initial predictions on unseen test images, which are then iteratively refined. The Betti number guidance encourages the segmentation to conform to known topological priors (e.g., $\beta_0 = 1$ for airways). However, enforcing Betti number constraints alone may induce undesired structural changes and undermine the topological integrity, as shown in Figure 1(a.2). To mitigate this, we introduce a topological integrity term that penalizes deviations from the original topology of the predicted segmentation. By jointly optimizing these two terms, TopoSculpt preserves the global topology of the entire structure while enhancing local topological correctness, as illustrated in Figure 1(a.3) and detailed in Figure 1(c).

Second, we propose a curriculum refinement strategy to efficiently implement TopoSculpt. It is well known that persistent homology (PH) calculation for whole 3D structures is computationally expensive (Otter et al., 2017; Zomorodian, 2005). To mitigate this, we design a progressive refinement paradigm that gradually shifts the focus from large-scale to fine-scale topological errors, thereby substantially reducing the computational burden. As illustrated in Figure 1(b), Betti numbers change rapidly in the early refinement stages but stabilize in later stages. Accordingly, we initially compute PH features densely with fewer steps to correct large-scale errors, and then progressively increase the sampling interval to concentrate on finer-scale corrections. This curriculum design not only accelerates the refinement process but also enables the model to efficiently learn to correct topological inaccuracies at multiple scales.

Third, TopoSculpt was evaluated on three challenging 3D fine-grained tubular structures: the pulmonary airway (Zhang et al., 2023), the Circle of Willis (CoW) (Yang et al., 2024), and the coronary arteries from the CTCA dataset (Gharleghi et al., 2023). Experimental results demonstrate that TopoSculpt significantly enhances both topological integrity and geometric accuracy across all datasets. For the pulmonary airway dataset, TopoSculpt reduces the β_0 error from 69.00 to 3.40, while improving the Tree Detection Rate (TD) from 82.23% to 91.56% and the Branch Detection Rate (BD) from 82.06% to 91.33%. Additionally, clDice and NSDice increase from 90.48% to 92.27% and from 96.96% to 98.11%, respectively. Similar improvements are observed on the CoW and CTCA datasets. For the CoW dataset, which has a relatively lower initial β_0 error, TopoSculpt further refines the topology, reducing the β_0 error from 1.65 to 0.30. This correction of critical connectivity errors translates into substantial gains in TD and BD (from 82.84% and 84.26% to 96.73% and 96.72%, respectively). Concurrently, clDice and NSDice increase from 96.67% to 97.98% and

from 99.10% to 99.26%, respectively. For the CTCA dataset, TopoSculpt demonstrates robustness by reducing the β_0 error from 8.75 to 0.88, resulting in notable improvements in both TD (from 78.98% to 91.68%) and BD (from 75.17% to 91.33%). These consistent advancements across airway, CoW, and coronary anatomies highlight the efficacy of TopoSculpt in preserving global topological integrity while ensuring precise geometric fidelity in fine-grained 3D tubular reconstruction.

2 Related Work

2.1 Evaluation Metrics for Tubular Shapes

For general medical anatomy segmentation, the Dice Similarity Coefficient (DSC) (Milletari et al., 2016) is the most commonly used evaluation metric, and its differentiable form is often adopted as a loss function. In contrast, three-dimensional tubular structures with complex branching patterns, varying radii, and topology-sensitive connectivity pose additional challenges, where DSC alone cannot adequately capture geometric fidelity or topological integrity (Zhang et al., 2023). Recent studies (Reinke et al., 2021; Maier-Hein et al., 2024) have emphasized the limitations of overlap-based measures and suggested geometry-aware alternatives. The centerline Dice (clDice) (Shit et al., 2021) and normalized surface Dice (NSDice) (Nikolov et al., 2018) jointly evaluate centerline continuity and boundary accuracy, while the Hausdorff Distance (HD) (Huttenlocher et al., 2002) captures the maximum boundary deviation. Furthermore, Betti numbers (Otter et al., 2017; Wasserman, 2018) are employed to quantify topological correctness by evaluating the number of connected components, while the Tree Length Detected (TD) and Branch Detected Rate (BD) (Zhang et al., 2023) measure topological completeness by assessing the recovery of major branches. We emphasize that these topology-oriented metrics are critical for assessing the modeling quality of medical tubular structures, as they better represent the true anatomical structure and overall connectivity than conventional overlap-based metrics.

2.2 Topology-aware Segmentation

Recent studies have proposed various topology-aware loss functions (Kervadec et al., 2019; Qin et al., 2019; Zheng et al., 2021; Zhang et al., 2023; Shit et al., 2021; Kirchhoff et al., 2024) to better preserve the topologically critical regions of tubular structures. Kervadec et al. (2019) introduced a boundary loss formulated as a distance metric in the contour space to improve segmentation accuracy along object boundaries. The centerline Dice (clDice) loss (Shit et al., 2021) was designed to explicitly enforce connectivity preservation in thin structure segmentation, and Shi et al. (2024) extended it to a centerline-boundary Dice (cbDice) loss by incorporating boundary-aware components. However, the differentiable skeletonization process in clDice tends to produce jagged or unstable centerlines. To address this issue, Kirchhoff et al. (2024) and Zhang & Gu (2023) employed more robust skeletonization strategies and developed skeleton-based weighted loss functions to further enhance topological accuracy. In parallel, Persistent Homology (PH) (Otter et al., 2017) based approaches have been explored to directly enforce topological correctness in segmentation results (Hu et al., 2019; 2021). Hu et al. (2019) introduced a topological loss that penalizes discrepancies between the persistence diagrams of predictions and ground truth. Subsequently, Discrete Morse theory was incorporated into PH (Hu et al., 2021) to localize critical points and further improve the topological accuracy of segmentation. Despite these advances, several challenges remain. Owing to the high dimensionality and structural complexity of 3D fine-grained tubular anatomy, both surrogate loss and PH-based methods are often applied in a patch-wise manner, which limits their ability to ensure global topological consistency. In addition, PH computation is computationally expensive for high-dimensional data, reducing its practicality for large 3D volumes.

3 Method

3.1 Problem Formulation

As illustrated in Figure 2(a), instead of conventional patch-wise training and sliding-window testing, we employ a whole-region strategy by feeding entire image regions containing complete target structures into the network $\mathcal{F}$. This allows the model to capture the full spatial context, which is crucial for preserving the topological characteristics of 3D tubular shapes during both training and

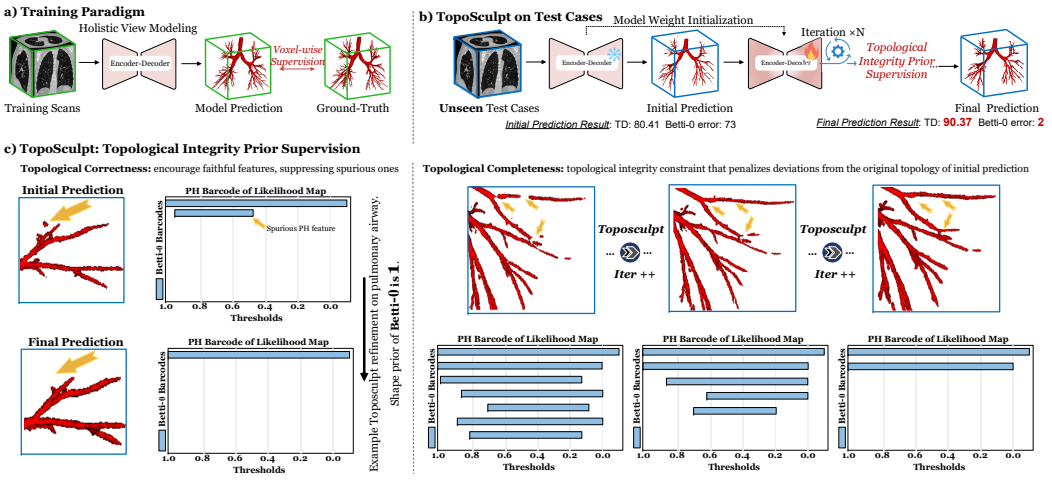

Figure 2: Overview of TopoSculpt. (a) Holistic training on full anatomical regions with voxel-wise supervision. (b) Iterative refinement of predictions on unseen scans under a topological integrity prior. (c) Persistent-homology barcodes guide Betti-based correction to suppress spurious components, while a topological integrity constraint preserves the global topology of the initial prediction.

inference. During training, the whole image region $\mathcal{X}$ and its corresponding ground-truth mask $\mathcal{Y}$ are provided as input, and the network parameters θ are optimized by minimizing a segmentation loss function $\mathcal{L}_{\text{seg}}$ (e.g., Dice loss):

$$\theta^* = \arg\min_{\theta} \mathcal{L}_{\text{seg}}(\mathcal{F}(\mathcal{X};\theta), \mathcal{Y}). \tag{1}$$

When the training process finished, the model with the parameters of θ^* can directly infer the entire image region during testing, producing a segmentation mask $\hat{\mathcal{Y}} = \mathcal{F}(\mathcal{X};\theta)$ without the need for patch-wise aggregation.

3.2 Topological Integrity Betti

Topological Integrity Betti (TIB) combines Betti number guidance with topological integrity preservation to refine predictions during the test stage. As seen in Figure 2(b), once $\mathcal{F}$ is trained, it produces initial segmentations on unseen test images, which are then iteratively refined by TIB ($\mathcal{L}_{\text{TIB}}$). Generally speaking, the iterative refinement process is formulated as: $\theta^{i+1} := \theta^i - \eta \nabla_\theta \mathcal{L}_{\text{TIB}}\left(\theta^{(i)}\right),\ i \geq 0,\ \theta^0 = \theta^*$. The framework consists of two components, $\mathcal{L}_{\text{TIB}_{cor}}$ and $\mathcal{L}_{\text{TIB}_{com}}$, responsible for correcting topological errors and preserving overall integrity, respectively. The correction term $\mathcal{L}_{\text{TIB}_{cor}}$ leverages known topological priors encoded by Betti numbers. Given the initial probability map $\hat{\mathcal{Y}}$ produced by the trained network $\mathcal{F}_{\theta^*}$, where $\hat{\mathcal{Y}} = \mathcal{F}_{\theta^*}(\mathcal{X}), \hat{\mathcal{Y}} : \mathbb{R}^3 \to [0,1]$, persistent homology (PH) (Otter et al., 2017; Zomorodian, 2005) is applied to characterize the evolving topology of $\hat{\mathcal{Y}}$ across all possible thresholds p. This process yields a dynamic description of the probabilistic segmentation topology and its critical events. The prior topology term ($Topo_P$) can then be incorporated without additional annotation effort; for instance, the pulmonary airway is known to satisfy $\beta_0 = 1$, corresponding to a single connected component. Since the ground-truth topology $\mathcal{Y}$ is deterministic and its PH barcode remains persistent across all thresholds, we distinguish between faithful and superfluous features when analyzing the evolving topology of $\hat{\mathcal{Y}}$. The most persistent feature is defined as the faithful persistent feature ($Topo_F$), whereas all other features are regarded as superfluous persistent features ($Topo_S$). Accordingly, the correction term $\mathcal{L}_{\text{TIB}_{cor}}$ is designed to encourage the persistence of faithful features while suppressing spurious ones. It is formulated as:

$$\mathcal{L}_{\text{TIB}_{cor}} = (Topo_P - Topo_F^i + Topo_S^i),\ i \geq 0, \tag{2}$$

where $Topo_P$ denotes the topology prior, $Topo_F^i$ represents the faithful feature, and $Topo_S^i$ corresponds to the superfluous feature at the i-th refinement step generated from $\hat{\mathcal{Y}}^i = \mathcal{F}(\mathcal{X};\theta^i)$. $Topo_P$ is a constant persistent feature derived from the known topology of the target structure, while $Topo_F^i$

and $Topo_S^i$ are dynamic persistent features that evolve with the segmentation during each refinement iteration. As illustrated in Figure 1(a.2), enforcing only $\mathcal{L}_{\text{TIB}_{cor}}$ may induce shortcut learning, which leads to substantial structural distortions and undermines topological integrity (Appendix G.4). To mitigate this issue, we introduce the integrity preservation term $\mathcal{L}_{\text{TIB}_{com}}$, designed to penalize deviations from the original topology of the predicted segmentation. This term is defined from two complementary perspectives, voxel-wise constraint and structural-wise constraint, as follows:

$$\mathcal{L}_{\text{TIB}_{com}} = \alpha \times \frac{1}{N} \left| \mathcal{F}_{\theta^{i+1}}(\mathcal{X}) - \mathcal{F}_{\theta^i}(\mathcal{X}) \right|^2 - \beta \times 2 \times \frac{Struc_{prec} \times Struc_{sens}}{Struc_{prec} + Struc_{sens}}, \tag{3}$$

where the first term serves as a voxel-wise constraint by minimizing the mean squared error between consecutive predictions, with N denoting the total number of voxels in $\mathcal{X}$. The second term acts as a structural-wise constraint by maximizing structural similarity throughout refinement, thereby promoting the recovery of topological integrity rather than its degradation, as illustrated in Figure 1(a.3) and further detailed in Figure 1(c). We quantify the proposed structural similarity using structural precision ($Struc_{prec}$) and structural sensitivity ($Struc_{sens}$), which are computed from the soft skeletons of the predicted segmentations weighted by the likelihood maps across two consecutive iterations:

$$\text{Struc}_{\text{prec}}(SP^{i+1}, VP^i) = \frac{|SP^{i+1} \cap VP^i|}{|SP^{i+1}|} \; ; \; \text{Struc}_{\text{sens}}(SP^i, VP^{i+1}) = \frac{|SP^i \cap VP^{i+1}|}{|SP^i|}. \tag{4}$$

Unlike Shit et al. (2021), we do not have the access to ground-truth skeletons during test time refinement. Therefore, we compute two consecutive soft skeletons weighted by their respective likelihood maps to approximate the structural similarity, which are defined as:

$$\begin{aligned} &SP^{i+1} = \text{soft-skel}(\mathcal{F}_{\theta^{i+1}}(\mathcal{X}) > 0.5) \odot \mathcal{F}_{\theta^{i+1}}(\mathcal{X}), \quad SP^i = \text{soft-skel}(\mathcal{F}_{\theta^i}(\mathcal{X}) > 0.5) \odot \mathcal{F}_{\theta^i}(\mathcal{X}), \\ &VP^{i+1} = \mathcal{F}_{\theta^{i+1}}(\mathcal{X}), \quad VP^i = \mathcal{F}_{\theta^i}(\mathcal{X}). \end{aligned} \tag{5}$$

In summary, the overall TIB loss ($\mathcal{L}_{\text{TIB}}$) is formulated as:

$$\mathcal{L}_{\text{TIB}} = \mathcal{L}_{\text{TIB}_{cor}} + \mathcal{L}_{\text{TIB}_{com}}. \tag{6}$$

3.3 Curriculum Refinement

It is well known that persistent homology (PH) calculation for whole 3D structures is computationally expensive (Otter et al., 2017; Zomorodian, 2005; Bauer, 2021). We observed that Betti numbers change rapidly in the early refinement stages but stabilize in later stages, as illustrated in Figure 1(b). To mitigate the computational burden, we design a progressive refinement paradigm that gradually shifts the focus from large-scale to fine-scale topological errors. Accordingly, we initially compute PH features densely with fewer steps to correct large-scale errors, and then progressively increase the sampling interval to concentrate on finer-scale corrections. Specifically, the curriculum refinement is defined as:

$$\begin{aligned} &\mathcal{L}_{\text{TIB}} = (Topo_P - Topo_F^i + Topo_S^i) + \mathcal{L}_{\text{TIB}_{com}}, && 0 \le i \le t, \\ &\mathcal{L}_{\text{TIB}} = (Topo_P - Topo_F^{j(i)} + Topo_S^{j(i)}) + \gamma \mathcal{L}_{\text{TIB}_{com}}, \quad t \le i \le T, \quad j(i) = \left\lfloor \tfrac{i}{k} \right\rfloor \cdot k. \end{aligned} \tag{7}$$

When the refinement enters the later stages ($i > t$), Betti numbers become more stable and large-scale corrections are less frequent. To reduce computational cost, we introduce a sampling mechanism through the function $j(i) = \lfloor i/k \rfloor \cdot k$, which selects representative iterations spaced by interval k. At these sampled steps, the topological prior $Topo_P$ is enforced, while the integrity term $\mathcal{L}_{\text{TIB}_{com}}$ is down-weighted by γ to avoid over-regularization. This design enables TopoSculpt to first perform dense and aggressive topological correction, and then progressively shift towards sparse and fine-scale adjustments, thereby achieving both efficiency and accuracy in topological refinement.

4 Experiments

4.1 Experimental Setup

Datasets. We evaluate TopoSculpt on three challenging benchmarks of three-dimensional tubular anatomy. 1) **ATM'22 Pulmonary Airway Dataset** (Zhang et al., 2023; 2024). This dataset

comprises multi-site, multi-domain chest CT scans with detailed airway tree annotations. The airway labels capture complex branching structures, frequent bifurcations, and substantial variability in branch radii. For our experiments, we selected a subset of 150 CT scans, allocating 120 scans for training and validation and 30 scans for testing. 2) **TopCoW'24 Circle of Willis (CoW) Dataset** (Yang et al., 2024). The Circle of Willis is a critical vascular network that provides collateral blood supply between the anterior and posterior circulations of the brain. In this study, we focused on structures satisfying the topological prior $\beta_0 = 1$. Accordingly, we selected 71 CTA cases from the TopCoW24 dataset, with 51 cases used for training and validation and 20 cases reserved for testing. 3) **ASOCA'23 Computed Tomography Coronary Angiography (CTCA) Dataset** (Gharleghi et al., 2023). CTCA contains 40 cases of computed tomography coronary angiography images with annotations of the coronary artery tree. The coronary artery tree is a complex vascular structure with numerous fine branches and high anatomical variability. We randomly selected 32 cases for training and validation and 8 cases for testing.

Implementation details. CT intensities are clamped to [-1000, 400], [-1000, 1800], and [-100, 500] HU for the airway, CoW, and CTCA datasets, respectively, followed by normalization to [0, 1]. The holistic-view model (HVModel) is built upon a standard 3D U-Net architecture (Çiçek et al., 2016) with an initial channel width of 16, operating directly on the full image region. For comparison, we additionally implement a patch-wise variant of the same 3D U-Net trained with the combination of Dice loss and cross-entropy loss (denoted by Dice+CE in Table 1), where a patch size of $128 \times 128 \times 128$ is used for all experiments. All competing methods are implemented using the same 3D U-Net backbone to ensure fairness. During the refinement stage of TopoSculpt, we set $\alpha = 1e4$ and $\beta = 1e3$ in the $\mathcal{L}_{\text{TIB}_{com}}$, and for the curriculum refinement strategy, we set $t = 30$, $T = 90$, $k = 3$, and $\gamma = 0.1$. The AdamW (Loshchilov & Hutter, 2017) optimizer is used with a learning rate of $1e-3$ for pulmonary airway, $1e-4$ for Circle of Willis and $1e-4$ for CTCA. All experiments are conducted on two NVIDIA A800 GPUs each with 80GB memory.

4.2 Quantitative Results

The quantitative performance of TopoSculpt and competing methods was evaluated on three challenging 3D tubular structure datasets, with comprehensive results presented in Table 1. A consistent finding across all datasets is the superior performance of our holistic-view (HVModel) baseline compared to the standard patch-wise approach trained with the combination of Dice loss and cross-entropy loss (denoted by Dice+CE in Table 1), particularly in topological metrics. For instance, the β_0 error is reduced from 96.40 to 69.00 on the airway dataset, from 5.55 to 1.65 on the CoW dataset, and from 33.50 to 8.75 on the CTCA dataset, underscoring the advantage of global context modeling. Further, with the proposed TopoSculpt refinement based on HVModel, we observe substantial improvements across all datasets. Recent competing topology-aware segmentation methods, including clDice (Shit et al., 2021), CAL (Zhang & Gu, 2023), SkeletonRecall (Kirchhoff et al., 2024), cbDice (Shi et al., 2024), BoundaryLoss (Kervadec et al., 2019), TopoLoss (Hu et al., 2019), and DMT (Hu et al., 2021), were included for comparison. On the CTCA coronary angiography dataset, TopoSculpt significantly outperformed all other methods. It demonstrated superior topological correctness, reducing the β_0 error to 0.88 from the baseline's 8.75, while many other methods struggled to reduce this error below 10. This improvement in topological integrity was accompanied by notable gains in completeness metrics, with TD increasing from 78.98% to 91.68% (+12.7%) and BD from 75.17% to 91.33% (+16.2%). It can be observed that TopoSculpt surpassed all other SOTA methods by a large margin. For instance, the next best competitor, SkeletonRecall, only achieved a TD of 80.71%. As for overlap-based geometric accuracy, TopoSculpt achieved a clDice of 90.01%, NSDice of 97.68%, and BD of 91.33%. This represents a substantial improvement over the HVModel baseline, which scored 87.27% clDice and only 75.17% BD. Consistent improvement was observed on the TopCoW'24 Circle of Willis dataset. TopoSculpt achieved high topological completeness with a BD of 96.72% and TD of 96.73%, alongside a clDice of 97.98%. This was a remarkable improvement from the HVModel's BD of 84.26% and TD of 82.84%. TopoSculpt effectively eliminated Betti errors by reducing the β_0 error from 1.65 to 0.30. In contrast, all other SOTA methods, including clDice (Shit et al., 2021) and CAL (Zhang & Gu, 2023), failed to exceed 81% in BD or TD, highlighting the superiority of TopoSculpt in correcting critical vascular disconnection. The most challenging ATM'22 pulmonary airway dataset clearly demonstrates the superiority of TopoSculpt. While most state-of-the-art methods struggled severely with maintaining topological correctness, often yielding β_0 errors exceeding 100 (e.g., CAL (Zhang & Gu, 2023)

Table 1: Quantitative comparison on ATM'22, TopCoW'24, and CTCA (mean±std). HD is measured in mm; clDice, NSDice, BD, and TD are percentages.

Dataset	Method	clDice ↑	NSDice ↑	HD ↓	BD ↑	TD ↑	β_0 error ↓
ATM'22	Dice+CE	86.27±3.44	95.20±2.20	1.56±1.04	73.25±8.52	74.65±7.83	96.40±27.40
	clDice	88.81±4.50	91.69±2.36	4.92±7.31	85.55±8.31	85.63±7.78	102.10±34.80
	CAL	89.64±3.19	97.51±3.19	1.29±1.16	85.23±7.69	85.58±7.41	103.03±32.07
	SkeletonRecall	87.69±3.46	96.57±1.78	3.00±5.77	87.52±5.26	87.88±4.95	104.33±36.15
	cbDice	85.31±6.09	94.65±3.18	2.65±4.33	80.63±11.46	81.81±11.01	64.27±23.62
	BoundaryLoss	88.23±4.67	97.40±1.90	1.58±2.88	80.08±10.18	80.23±9.29	102.87±34.24
	TopoLoss	75.14±30.87	86.99±25.52	40.56±47.99	69.17±31.28	69.58±30.54	114.10±49.67
	DMT	75.54±32.91	83.01±28.97	37.59±46.76	69.50±32.42	69.69±31.92	111.70±53.02
	HVModel	90.48±2.92	96.96±1.50	1.07±0.56	82.06±7.44	82.23±6.37	69.00±19.01
	TopoSculpt	**92.27±1.91**	**98.11±0.82**	**1.02±0.29**	**91.33±4.84**	**91.56±4.09**	**3.40±2.44**
TopCoW'24	Dice+CE	90.39±4.68	94.87±4.16	3.34±3.73	68.06±17.14	64.47±17.50	5.55±3.90
	clDice	93.78±3.24	96.17±3.59	1.89±1.72	79.07±16.15	74.08±17.29	1.80±1.57
	CAL	93.69±3.21	96.06±3.62	1.91±1.68	80.48±15.98	78.21±16.82	2.25±1.51
	SkeletonRecall	92.80±3.87	96.09±3.71	1.75±1.64	79.12±18.22	76.42±19.40	2.80±1.81
	cbDice	91.87±4.03	95.90±3.47	2.03±1.70	63.66±13.65	59.01±12.83	2.25±1.37
	BoundaryLoss	92.96±3.08	95.91±3.27	1.97±1.67	73.33±17.28	70.72±16.92	2.80±1.83
	TopoLoss	91.19±4.31	95.90±3.47	2.03±1.70	63.66±13.65	59.01±12.83	2.25±1.37
	DMT	90.94±4.03	95.27±3.44	2.51±1.92	71.67±15.97	67.85±17.32	5.00±2.79
	HVModel	96.67±1.98	99.10±1.67	0.73±0.69	84.26±16.67	82.84±17.21	1.65±1.19
	TopoSculpt	**97.98±1.00**	**99.26±1.22**	**0.62±0.39**	**96.72±8.62**	**96.73±8.62**	**0.30±0.56**
CTCA	Dice+CE	84.09±4.54	93.18±4.12	6.35±5.67	71.76±12.86	70.49±14.35	33.50±10.34
	clDice	85.82±3.54	95.01±2.66	3.59±2.71	68.52±13.60	70.74±10.58	14.75±4.35
	CAL	85.48±2.35	94.56±2.55	4.52±3.44	67.39±13.21	71.38±13.59	18.12±5.80
	SkeletonRecall	85.08±3.07	94.20±2.69	4.72±3.67	77.86±13.53	80.71±12.22	16.50±6.96
	cbDice	80.07±5.35	91.94±4.10	7.16±4.81	56.96±12.29	59.34±15.26	11.00±3.43
	BoundaryLoss	82.68±6.10	94.63±2.29	4.45±2.70	54.65±11.18	63.29±15.49	13.88±5.73
	TopoLoss	84.26±5.27	94.18±2.91	4.15±2.48	63.93±13.56	64.11±15.18	19.62±7.87
	DMT	85.16±4.68	94.30±2.95	4.40±2.99	70.12±12.81	72.82±14.80	24.25±10.30
	HVModel	87.27±6.89	95.84±5.93	0.78±0.58	75.17±7.65	78.98±10.37	8.75±6.16
	TopoSculpt	**90.01±3.53**	**97.68±2.66**	**0.58±0.26**	**91.33±5.56**	**91.68±5.91**	**0.88±1.36**

and SkeletonRecall (Kirchhoff et al., 2024)), TopoSculpt drastically reduced this error from the HVModel baseline of 69.00 to only 3.40. This fundamental correction in topological integrity directly translated into state-of-the-art structural completeness, with BD improving to 91.33% and TD to 91.56%. Such performance markedly surpasses strong competitors, including CAL (85.23% BD) and SkeletonRecall (87.52% BD). Moreover, PH-based approaches such as TopoLoss (Hu et al., 2019) and DMT (Hu et al., 2021) failed to perform well on this dataset, exhibiting clDice scores around 75% alongside high HD/Betti errors. In contrast, TopoSculpt achieved the highest geometric fidelity, reaching 92.27% in clDice, while simultaneously ensuring topological completeness.

In summary, the quantitative results across all three datasets confirm that TopoSculpt not only improves upon its holistic-view baseline but also decisively outperforms all competing SOTA methods. It uniquely demonstrates the ability to correct severe topological errors (lowering β_0) and, crucially, translates these corrections into comprehensive and significant improvements in both geometric accuracy (clDice, NSDice, HD) and topology-based structural integrity (TD, BD).

4.3 Qualitative Results

Figure 3 presents a direct visual comparison between TopoSculpt and several competing methods across all three datasets. The results clearly show that TopoSculpt consistently achieves superior structural integrity and greater consistency to the ground truth. The local regions highlighted by green boxes further emphasize TopoSculpt's ability to preserve fine structural details and maintain topological fidelity, whereas other methods frequently exhibit breakages, discontinuities, or spurious noise in these areas.

Figure 4 visualizes the iterative refinement process of TopoSculpt on unseen test cases of the pulmonary airway. Unlike conventional methods where model parameters are fixed post-training, leav-

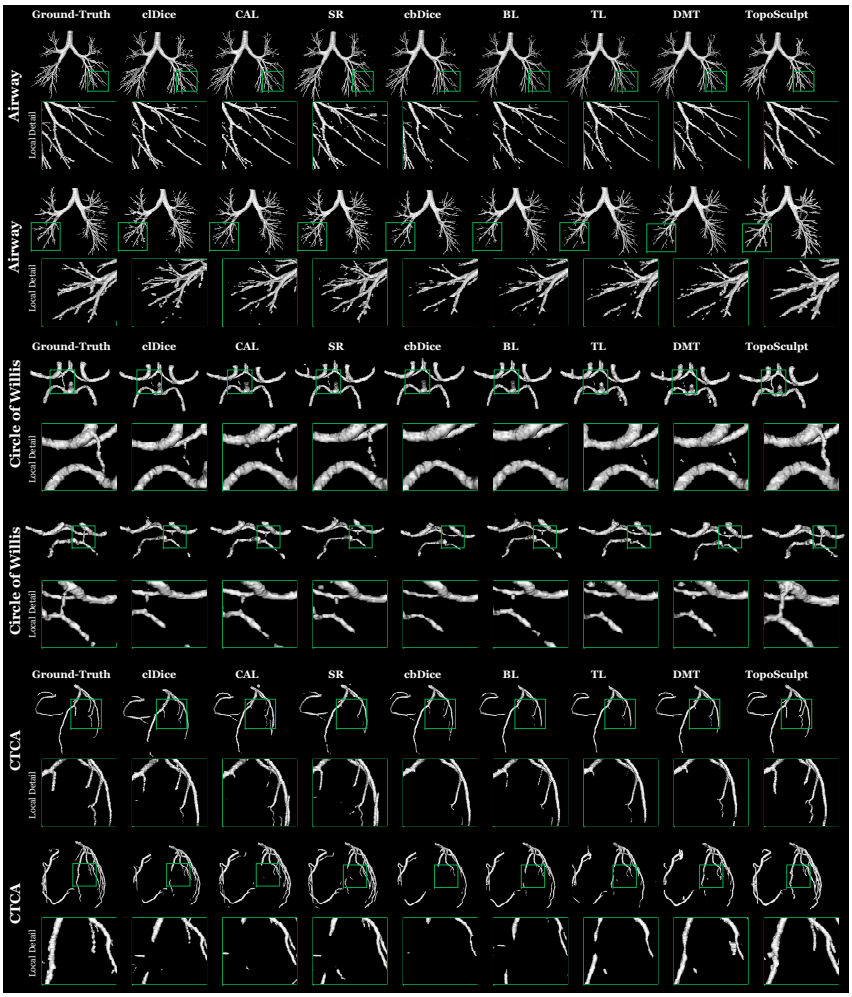

Figure 3: Qualitative results. Green boxes mark improved detail and topology.

ing them unable to correct topological inconsistencies on a case-by-case basis during inference, TopoSculpt performs a test-time refinement supervised by a topological integrity prior. The figure visually confirms that each refinement step progressively improves the reconstruction, steering the initial prediction towards a topologically correct and structurally complete result. Concretely, the initial predictions often suffer from severe fragmentation, resulting in high β_0 errors (e.g., 73, 94) and low TD rates. The figure illustrates TopoSculpt's progressive refinement, showing how it systematically reconnects fragmented subtrees and resolves broken links. The final predictions exhibit dramatically reduced β_0 errors (e.g., 2, 1) and correspondingly high TD scores. The local detail views confirm the restoration of continuity in peripheral branches. The same behaviour is observed on the other two anatomies: the refinement processes on the Circle of Willis and the coronary artery are visualized in Appendix B, and patch-level corrections of breakages and spurious components are shown in Appendix C.

4.4 Ablation Study

Effectiveness of topological completeness constraint. By jointly optimizing $\mathcal{L}_{\text{TIB}_{cor}}$ and $\mathcal{L}_{\text{TIB}_{com}}$, TopoSculpt achieves a balanced refinement process that corrects topological errors while maintaining structural coherence. The qualitative results in Figure 10 (Appendix E) clearly demonstrate this improvement. Without $\mathcal{L}_{\text{TIB}_{com}}$, the refined predictions exhibit residual disconnections in airway and coronary artery structures or missing vessel branches in CoW structures, as highlighted by the

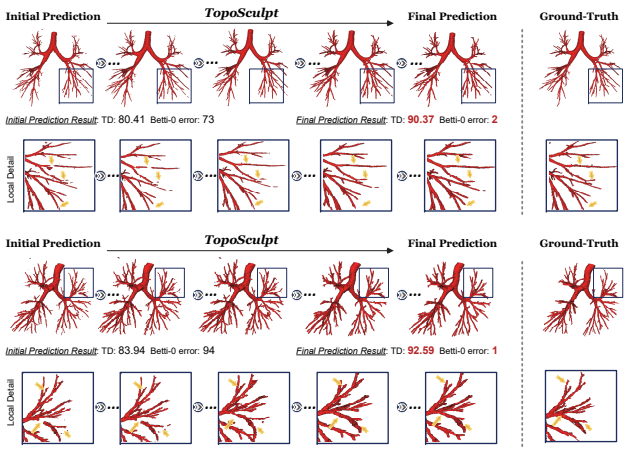

Figure 4: Progressive refinement by TopoSculpt, from the initial airway to the final result. In each case, the top row shows the whole airway with improving global topology, and the bottom row zooms in on refined local branches.

Table 2: Ablation of the effectiveness of topological integrity $\mathcal{L}_{\text{TIB}_{com}}$.

Dataset	Method	clDice ↑	NSDice ↑	HD ↓	BD ↑	TD ↑	β_0 error ↓
ATM'22	TopoSculpt w/o $\mathcal{L}_{\text{TIB}_{com}}$	89.74±2.55	96.88±1.07	1.41±0.75	83.92±5.63	84.48±5.29	12.36±6.33
	TopoSculpt	**92.27±1.91**	**98.11±0.82**	**1.02±0.29**	**91.33±4.84**	**91.56±4.09**	**3.40±2.44**
TopCoW'24	TopoSculpt w/o $\mathcal{L}_{\text{TIB}_{com}}$	96.92±1.56	98.32±6.10	0.71±0.27	91.58±14.20	90.79±14.86	0.45±0.67
	TopoSculpt	**97.98±1.00**	**99.26±1.22**	**0.62±0.39**	**96.72±8.62**	**96.73±8.62**	**0.30±0.56**
CTCA	TopoSculpt w/o $\mathcal{L}_{\text{TIB}_{com}}$	89.06±2.55	96.83±2.59	0.69±0.27	88.17±5.51	90.43±4.61	**0.75±0.83**
	TopoSculpt	**90.01±3.53**	**97.68±2.66**	**0.58±0.26**	**91.33±5.56**	**91.68±5.91**	0.88±1.36

arrows. In contrast, the complete TopoSculpt framework produces continuous and topologically consistent reconstructions that are highly consistent with the ground truth. Quantitative results in Table 2 further validate the importance of $\mathcal{L}_{\text{TIB}_{com}}$. Across all three datasets, incorporating this term significantly improves both topological and geometric metrics. On the ATM'22 pulmonary airway dataset, $\mathcal{L}_{\text{TIB}_{com}}$ reduces the β_0 error from 12.36 to 3.40 and increases BD and TD from 83.92% and 84.48% to 91.33% and 91.56%, respectively. Similar improvements are observed for the CoW and CTCA datasets, where β_0 errors are further reduced and clDice and NSDice consistently increase. These results confirm that $\mathcal{L}_{\text{TIB}_{com}}$ is essential for preventing shortcut learning, preserving meaningful anatomical structures, and achieving globally consistent topological refinement across diverse three-dimensional tubular anatomies, in line with the analysis in Appendix G.5.

Further analyses. The curriculum-like decrease of the β_0 error over refinement iterations and the evolution of the persistent homology barcodes of the likelihood maps are analyzed in Appendix D, and a comparison with nnUNet backbones of substantially larger capacity, which shows that model scaling or holistic input alone cannot attain the topological fidelity restored by TopoSculpt, is reported in Appendix F.

5 Conclusion

In this work, we presented TopoSculpt, a holistic framework for refining three-dimensional tubular shape modeling by jointly leveraging Betti-guided constraints and topological integrity preservation. Unlike prior patch-wise or training-only approaches, TopoSculpt introduces whole-region modeling to ensure global topology consistency, while the Topological Integrity Betti (TIB) constraint balances correctness and structural fidelity. A curriculum refinement strategy further accelerates computation by progressively addressing errors across scales. Extensive experiments on pulmonary airways, the Circle of Willis, and coronary artery confirm that TopoSculpt achieves significant improvements in β_0 error reduction, tree and branch detection, as well as geometric overlap metrics.

AI USE STATEMENT

In this work, we used generative AI tools only to edit the manuscript for readability and to assist with formatting the LaTeX source. We have not used generative AI tools to generate synthetic data, to design or execute the experiments, to implement the method, or to interpret the results. All AI-assisted edits were reviewed by the authors, and all technical claims, experimental results, figures, and tables were produced and verified by the authors.

ETHICS STATEMENT

This work uses only publicly available, de-identified datasets (ATM'22, TopCoW'24, and ASOCA'23) under their respective release licenses, and does not involve any new data collection from human subjects. The proposed method is intended for research on topology-preserving segmentation; it has not been clinically validated and should not be used for diagnostic or treatment decisions without further evaluation. We do not foresee any direct negative societal impact, conflicts of interest, or fairness and privacy concerns arising from this work.

REPRODUCIBILITY STATEMENT

We have made efforts to ensure the reproducibility of our results. The datasets, data splits, and preprocessing (HU clamping and normalization) are described in Section 4, and all hyper-parameters of the TIB loss and the curriculum refinement schedule (α, β, t, T, k, γ), the optimizer settings, and the hardware used are reported in the implementation details. The formulation of the TIB loss and the curriculum refinement is fully specified in Section 3. Additional qualitative results and the comparison with larger-capacity backbones are provided in the appendix.

REFERENCES

Ulrich Bauer. Ripser: efficient computation of vietoris–rips persistence barcodes. *Journal of Applied and Computational Topology*, 5(3):391–423, 2021.

Bea Bleile, Adélie Garin, Teresa Heiss, Kelly Maggs, and Vanessa Robins. The persistent homology of dual digital image constructions. In *Research in Computational Topology 2*, pp. 1–26. Springer, 2022. doi: 10.1007/978-3-030-95519-9_1.

Mathieu Carrière, Frédéric Chazal, Marc Glisse, Yuichi Ike, Hariprasad Kannan, and Yuhei Umeda. Optimizing persistent homology based functions. In *International Conference on Machine Learning*, volume 139, pp. 1294–1303. PMLR, 2021.

Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: learning dense volumetric segmentation from sparse annotation. In *International conference on medical image computing and computer-assisted intervention*, pp. 424–432. Springer, 2016.

Frank H Clarke. *Optimization and Nonsmooth Analysis*, volume 5 of *Classics in Applied Mathematics*. SIAM, 1990.

James R Clough, Nicholas Byrne, Ilkay Oksuz, Veronika A Zimmer, Julia A Schnabel, and Andrew P King. A topological loss function for deep-learning based image segmentation using persistent homology. *IEEE transactions on pattern analysis and machine intelligence*, 44(12): 8766–8778, 2020.

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

Damek Davis, Dmitriy Drusvyatskiy, Sham Kakade, and Jason D Lee. Stochastic subgradient method converges on tame functions. *Foundations of Computational Mathematics*, 20(1):119–154, 2020.

Herbert Edelsbrunner and John L Harer. *Computational topology: an introduction*. American Mathematical Society, 2022.

R Gharleghi, D Adikari, K Ellenberger, M Webster, C Ellis, A Sowmya, S Ooi, and S Beier. Annotated computed tomography coronary angiogram images and associated data of normal and diseased arteries. *Scientific Data*, 10(1):128, 2023.

Yun Gu, Chuanjia Gu, Jie Yang, Jiayuan Sun, and Guang-Zhong Yang. Vision–kinematics interaction for robotic-assisted bronchoscopy navigation. *IEEE Transactions on Medical Imaging*, 41 (12):3600–3610, 2022.

Allen Hatcher. *Algebraic Topology*. Cambridge University Press, 2002.

Xiaoling Hu, Fuxin Li, Dimitris Samaras, and Chao Chen. Topology-preserving deep image segmentation. *Advances in neural information processing systems*, 32, 2019.

Xiaoling Hu, Yusu Wang, Li Fuxin, Dimitris Samaras, and Chao Chen. Topology-aware segmentation using discrete morse theory. *arXiv preprint arXiv:2103.09992*, 2021.

Xiaoling Hu, Dimitris Samaras, and Chao Chen. Learning probabilistic topological representations using discrete morse theory. *arXiv preprint arXiv:2206.01742*, 2022.

Daniel P Huttenlocher, Gregory A. Klanderman, and William J Rucklidge. Comparing images using the hausdorff distance. *IEEE Transactions on pattern analysis and machine intelligence*, 15(9): 850–863, 2002.

Fabian Isensee, Paul F Jaeger, Simon AA Kohl, Jens Petersen, and Klaus H Maier-Hein. nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature methods*, 18(2):203–211, 2021.

Fabian Isensee, Maximilian Rokuss, Lars Krämer, Stefan Dinkelacker, Ashis Ravindran, Florian Stritzke, Benjamin Hamm, Tassilo Wald, Moritz Langenberg, Constantin Ulrich, et al. nninteractive: Redefining 3d promptable segmentation. *arXiv preprint arXiv:2503.08373*, 2025.

Arthur Jacot, Franck Gabriel, and Clément Hongler. Neural tangent kernel: Convergence and generalization in neural networks. In *Advances in Neural Information Processing Systems*, volume 31, 2018.

Shizuo Kaji, Takeki Sudo, and Kazushi Ahara. Cubical ripser: Software for computing persistent homology of image and volume data. *arXiv preprint arXiv:2005.12692*, 2020.

Hoel Kervadec, Jihene Bouchtiba, Christian Desrosiers, Eric Granger, Jose Dolz, and Ismail Ben Ayed. Boundary loss for highly unbalanced segmentation. In *International conference on medical imaging with deep learning*, pp. 285–296. PMLR, 2019.

Yannick Kirchhoff, Maximilian R Rokuss, Saikat Roy, Balint Kovacs, Constantin Ulrich, Tassilo Wald, Maximilian Zenk, Philipp Vollmuth, Jens Kleesiek, Fabian Isensee, et al. Skeleton recall loss for connectivity conserving and resource efficient segmentation of thin tubular structures. In *European Conference on Computer Vision*, pp. 218–234. Springer, 2024.

Chenyu Li, Minghui Zhang, Chuyan Zhang, and Yun Gu. Reflecting topology consistency and abnormality via learnable attentions for airway labeling. *International Journal of Computer Assisted Radiology and Surgery*, pp. 1–9, 2025.

Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*, 2017.

Jun Ma, Yuting He, Feifei Li, Lin Han, Chenyu You, and Bo Wang. Segment anything in medical images. *Nature Communications*, 15(1):654, 2024.

Lena Maier-Hein, Annika Reinke, Patrick Godau, Minu D Tizabi, Florian Buettner, Evangelia Christodoulou, Ben Glocker, Fabian Isensee, Jens Kleesiek, Michal Kozubek, et al. Metrics reloaded: recommendations for image analysis validation. *Nature methods*, 21(2):195–212, 2024.

Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 fourth international conference on 3D vision (3DV)*, pp. 565–571. Ieee, 2016.

Stanislav Nikolov, Sam Blackwell, Alexei Zverovitch, Ruheena Mendes, Michelle Livne, Jeffrey De Fauw, Yojan Patel, Clemens Meyer, Harry Askham, Bernardino Romera-Paredes, et al. Deep learning to achieve clinically applicable segmentation of head and neck anatomy for radiotherapy. *arXiv preprint arXiv:1809.04430*, 2018.

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):17, 2017.

Yaolei Qi, Yuting He, Xiaoming Qi, Yuan Zhang, and Guanyu Yang. Dynamic snake convolution based on topological geometric constraints for tubular structure segmentation. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 6070–6079, 2023.

Yulei Qin, Mingjian Chen, Hao Zheng, Yun Gu, Mali Shen, Jie Yang, Xiaolin Huang, Yue-Min Zhu, and Guang-Zhong Yang. Airwaynet: a voxel-connectivity aware approach for accurate airway segmentation using convolutional neural networks. In *International conference on medical image computing and computer-assisted intervention*, pp. 212–220. Springer, 2019.

Annika Reinke, Minu D Tizabi, Carole H Sudre, Matthias Eisenmann, Tim Rädsch, Michael Baumgartner, Laura Acion, Michela Antonelli, Tal Arbel, Spyridon Bakas, et al. Common limitations of image processing metrics: A picture story. *arXiv preprint arXiv:2104.05642*, 2021.

Pengcheng Shi, Jiesi Hu, Yanwu Yang, Zilve Gao, Wei Liu, and Ting Ma. Centerline boundary dice loss for vascular segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 46–56. Springer, 2024.

Suprosanna Shit, Johannes C Paetzold, Anjany Sekuboyina, Ivan Ezhov, Alexander Unger, Andrey Zhylka, Josien PW Pluim, Ulrich Bauer, and Bjoern H Menze. cldice-a novel topology-preserving loss function for tubular structure segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 16560–16569, 2021.

Nico Stucki, Johannes C Paetzold, Suprosanna Shit, Bjoern Menze, and Ulrich Bauer. Topologically faithful image segmentation via induced matching of persistence barcodes. In *International Conference on Machine Learning*, pp. 32698–32727. PMLR, 2023.

Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Deep image prior. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 9446–9454, 2018.

Larry Wasserman. Topological data analysis. *Annual review of statistics and its application*, 5 (2018):501–532, 2018.

Qian Wu, Yufei Chen, Wei Liu, Xiaodong Yue, and Xiahai Zhuang. Deep closing: Enhancing topological connectivity in medical tubular segmentation. *IEEE Transactions on Medical Imaging*, 43 (11):3990–4003, 2024.

Kaiyuan Yang, Fabio Musio, Yihui Ma, Norman Juchler, Johannes C Paetzold, Rami Al-Maskari, Luciano Höher, Hongwei Bran Li, Ibrahim Ethem Hamamci, Anjany Sekuboyina, et al. Benchmarking the cow with the topcow challenge: Topology-aware anatomical segmentation of the circle of willis for cta and mra. *ArXiv*, pp. arXiv–2312, 2024.

Minghui Zhang and Yun Gu. Towards connectivity-aware pulmonary airway segmentation. *IEEE Journal of Biomedical and Health Informatics*, 28(1):321–332, 2023.

Minghui Zhang, Yangqian Wu, Hanxiao Zhang, Yulei Qin, Hao Zheng, Wen Tang, Corey Arnold, Chenhao Pei, Pengxin Yu, Yang Nan, et al. Multi-site, multi-domain airway tree modeling. *Medical Image Analysis*, 90:102957, 2023.

Minghui Zhang, Chenyu Li, Fangfang Xie, Yaoyu Liu, Hanxiao Zhang, Junyang Wu, Chunxi Zhang, Jie Yang, Jiayuan Sun, Guang-Zhong Yang, et al. Airmorph: Topology-preserving deep learning for pulmonary airway analysis. *arXiv preprint arXiv:2412.11039*, 2024.

Hao Zheng, Yulei Qin, Yun Gu, Fangfang Xie, Jie Yang, Jiayuan Sun, and Guang-Zhong Yang. Alleviating class-wise gradient imbalance for pulmonary airway segmentation. *IEEE transactions on medical imaging*, 40(9):2452–2462, 2021.

Jiale Zhou, Wenhan Wang, Shikun Li, Xiaolei Qu, Xin Guo, Yizhong Liu, Wenzhong Tang, Xun Lin, and Yefeng Zheng. Topotta: Topology-enhanced test-time adaptation for tubular structure segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 24123–24134, 2025.

Afra J Zomorodian. *Topology for computing*, volume 16. Cambridge university press, 2005.

APPENDIX

A BROADER APPLICATIONS AND FUTURE DIRECTIONS

The proposed method can be regarded as a scalable method or an extension to existing interactive foundation segmentation approaches. Beyond the conventional prompts used in current interactive methods (Ma et al., 2024; Isensee et al., 2025), such as points, boxes, or scribbles, TopoSculpt could support the alternative geometric prior as the prompt to refine the initial segmentation results. A promising avenue for future work is the generation of geometric priors as language-formatted prompts and their integration, alongside TopoSculpt, into multimodal large language models.

Although TopoSculpt performs zero-shot test-time refinement, it faces the same stopping-criteria challenge as other iterative methods. A meaningful future direction would be to design case-specific refinement strategies by defining the topological distance between the likelihood map and the expected geometric prior. Such a distance measure could serve as a quantitative indicator of how much refinement is required for each case, thereby providing an adaptive stopping criterion. This would allow the refinement process to terminate once the topological discrepancy has been minimized, ensuring both computational efficiency and reliable topological fidelity.

TopoSculpt could also help alleviate the burden on clinicians when annotating fine tubular structures. As noted in airway datasets such as ATM22 and BAS, the provided labels are often incomplete. In such cases, the neural network may generate additional branches beyond a single connected component. Many of these branches, however, are clinically meaningful. By presenting them as candidate components, clinicians can rapidly review the results and filter out false positives. Then our proposed TopoSculpt can focus on refining the remaining structures. This process reduces manual labeling effort while ensuring both topological fidelity and geometrically accurate modeling.

B REFINEMENT PROCESS ON THE CIRCLE OF WILLIS AND THE CORONARY ARTERY

Figures 5 and 6 visualize the iterative test-time refinement of TopoSculpt on unseen cases of the Circle of Willis and the coronary artery, complementing the pulmonary airway results reported in Section 4. For the CoW dataset, the initial predictions exhibit critical disconnections at key arterial junctions, reflected in low topological integrity and non-zero β_0 errors. TopoSculpt effectively corrects these errors, leading to anatomically faithful vascular networks with higher TD scores and satisfactory β_0 numbers in the final predictions. Similarly, on the CTCA dataset, the initial predictions contain topological errors that the refinement process effectively resolves, reducing the Betti errors and significantly boosting TD scores.

C IMAGE-PATCH VIEWS OF THE CORRECTED TOPOLOGICAL ERRORS

In Figure 7, we provide qualitative results that illustrate how TopoSculpt corrects critical topological inconsistencies, shown in both 3D reconstructions and the corresponding image patch views. For the fine-grained pulmonary airway and CoW structures, the initial predictions often exhibit breakages or spurious isolated components, which directly correspond to erroneous β_0 features, as seen in the first row of Figure 7, pointed out by arrows. Such defects compromise structural connectivity and consequently lower branch and tree detection rates. By contrast, TopoSculpt effectively suppresses these Betti errors, restoring continuous branches while eliminating false positives. Importantly, the refinement is achieved without inducing geometric distortions, thereby preserving local morphology and lumen boundaries, as illustrated in the second row of Figure 7. This demonstrates the complementary strength of TopoSculpt: it explicitly enhances topological integrity while simultaneously maintaining geometric fidelity, yielding reconstructions that are both anatomically consistent and visually coherent with the ground truth.

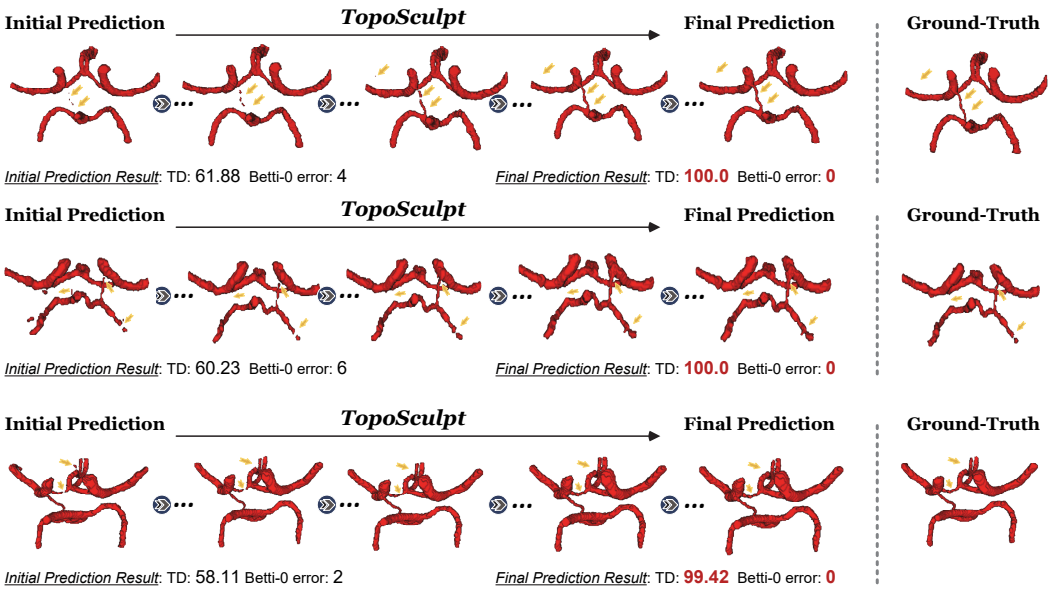

Figure 5: The visualization of TopoSculpt illustrates the refinement process from the initial Circle of Willis (CoW) structure to the final result, exhibiting substantial enhancement in topological fidelity. The arrows indicate the locations of topological errors in the initial segmentation, which are effectively corrected by TopoSculpt.

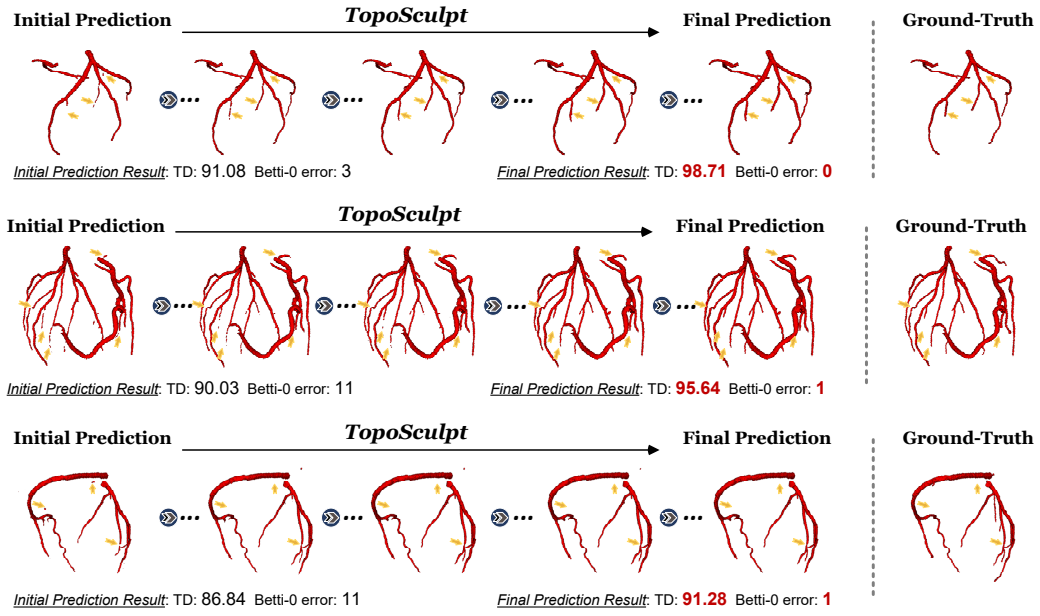

Figure 6: The visualization of TopoSculpt illustrates the refinement process from the initial coronary structure to the final result, exhibiting substantial enhancement in topological fidelity. The arrows indicate the locations of topological errors in the initial segmentation, which are effectively corrected by TopoSculpt.

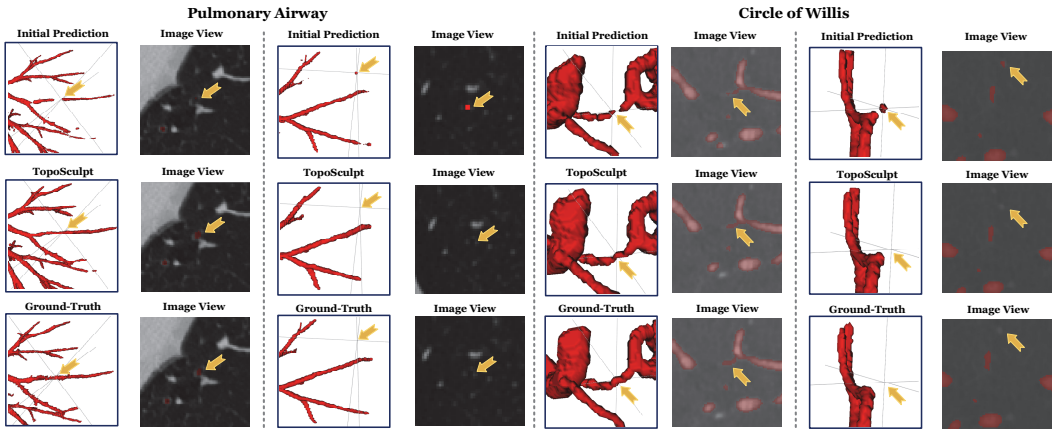

Figure 7: Qualitative comparisons of topological refinement results on the pulmonary airway (left) and Circle of Willis (right) datasets. Each block shows the 3D structure (left) and the corresponding image view (right). The initial predictions (top row) exhibit typical topological errors, such as branch discontinuities, spurious components, or breakages (highlighted by arrows). After refinement with TopoSculpt (middle row), these errors are effectively suppressed, leading to restored connectivity and improved structural integrity that more closely matches the ground truth (bottom row). The results confirm that TopoSculpt not only eliminates erroneous Betti-0 components but also preserves the fine-scale geometry of complex tubular structures across different anatomical regions.

D Refinement Dynamics: Curriculum Process and PH Barcodes

Curriculum learning process. Figure 8 illustrates the evolution of β_0 error throughout the TopoSculpt refinement process across three datasets. The solid line represents the mean β_0 error, while the shaded area denotes the min-max range across all test cases. A clear curriculum-like learning behavior can be observed. At the early refinement stages, the β_0 error decreases sharply, indicating that TopoSculpt efficiently eliminates large-scale connectivity errors through dense persistent homology (PH) sampling. As refinement continues, the rate of reduction gradually slows, and the curves stabilize, reflecting the transition from coarse-level correction to fine-grained structural refinement. This trend confirms the effectiveness of the proposed progressive refinement scheme in improving both convergence efficiency and topological stability. By focusing on large-scale errors first and progressively shifting toward finer-scale corrections, TopoSculpt achieves efficient topological recovery while maintaining global structural integrity across diverse anatomical domains.

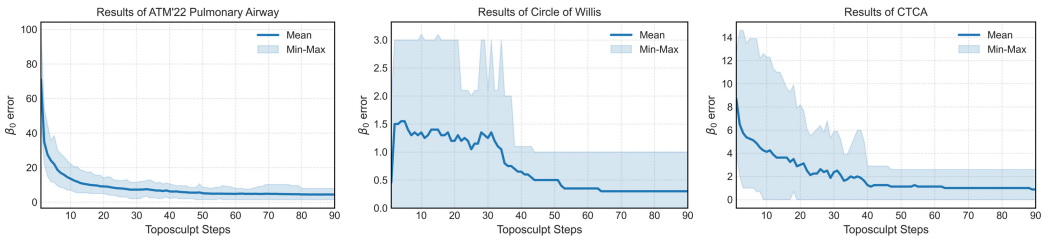

Figure 8: Evolution of β_0 error during the TopoSculpt refinement process across three representative datasets: ATM'22 Pulmonary Airway (left), Circle of Willis (middle), and CTCA (right). The solid line denotes the mean β_0 error, while the shaded region indicates the min-max range across all test cases. Betti numbers change rapidly in the early refinement stages, reflecting the correction of large-scale connectivity errors, but gradually stabilize in later stages as the structures converge toward topologically consistent configurations.

Visualization of PH barcode variation. Figure 9 visualizes the evolution of persistent homology (PH) barcodes for the likelihood maps during the refinement process of TopoSculpt across all three

datasets. Each horizontal line in the barcode represents a topological component feature and its persistence over threshold variations. These visualizations clearly demonstrate that TopoSculpt not only reduces the overall β_0 error but also promotes topological stability in the PH space. In the initial predictions, a large number of short-lived PH bars are observed, indicating the presence of numerous spurious components and local discontinuities. As the refinement proceeds, TopoSculpt progressively suppresses these unstable Betti features while preserving only the dominant, persistent structures corresponding to the true anatomical topology. This trend is particularly evident in the airway dataset, where the dense clusters of short Betti-0 bars gradually diminish, converging to several long-lived components that reflects the enhancement of connectivity. Similar stabilization can be observed in the CoW and CTCA datasets, where the number of spurious bars rapidly decreases and the remaining features exhibit longer persistence and higher structural consistency.

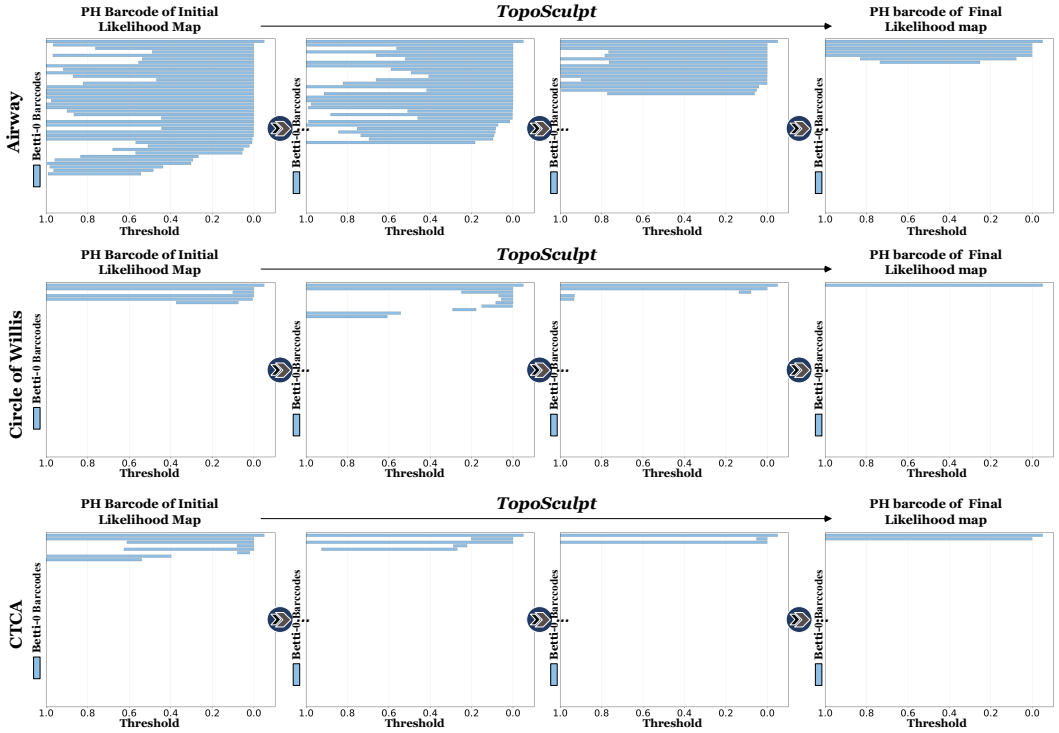

Figure 9: Evolution of PH barcodes of the likelihood map during the TopoSculpt refinement process across three representative datasets: ATM'22 Pulmonary Airway (top), Circle of Willis (middle), and CTCA (bottom). Each horizontal line in the barcode represents a topological component feature and its persistence over threshold variations.

E Qualitative Ablation of the Topological Integrity Term

Figure 10 provides the qualitative counterpart of Table 2, comparing TopoSculpt with and without the topological integrity term $\mathcal{L}_{\text{TIB}_{com}}$.

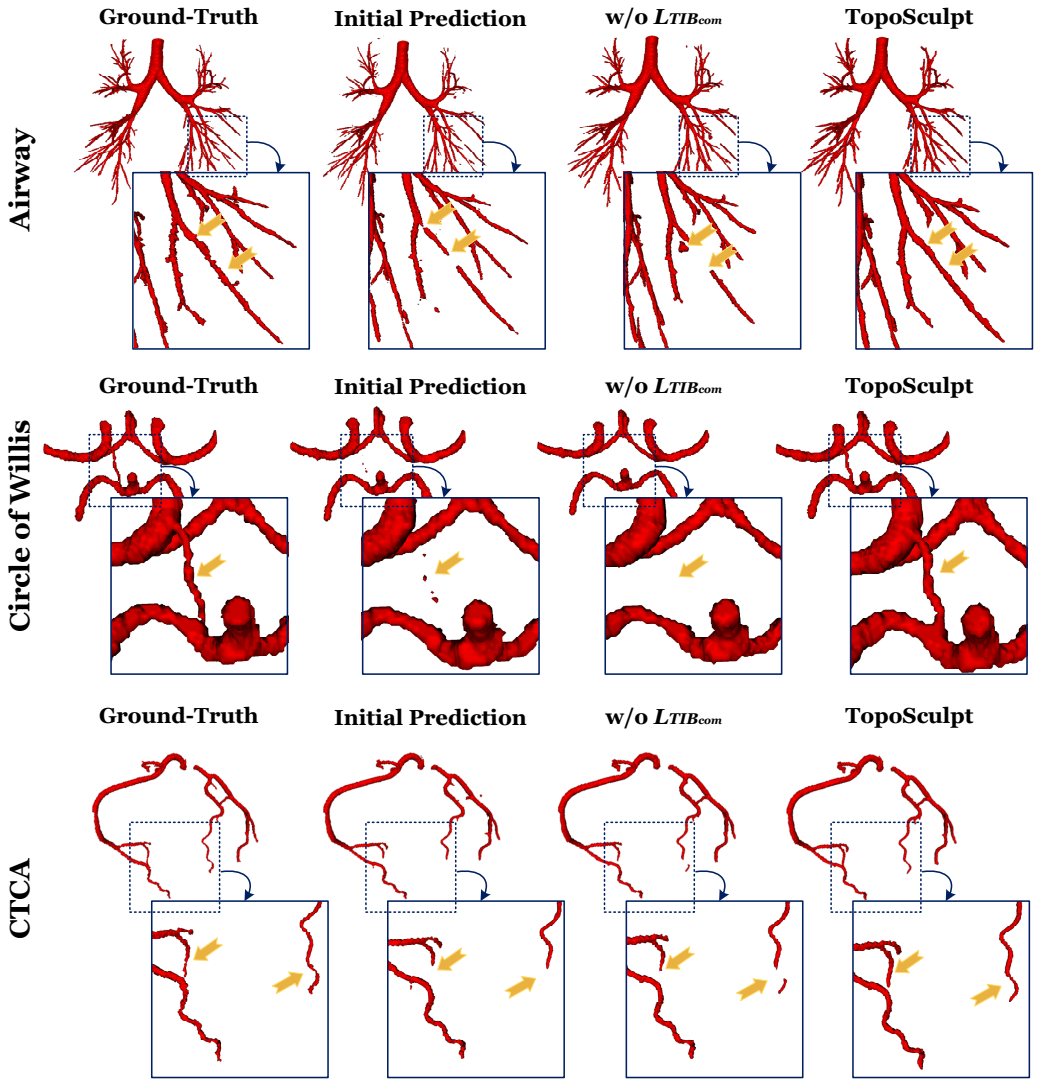

Figure 10: Ablation study on the TIB_{com} component of the proposed TopoSculpt framework across three representative datasets, ATM'22 Pulmonary Airway (top), Circle of Willis (middle), and CTCA (bottom). TIB_{com} demonstrates its contribution to improving topological integrity and structural consistency.

F Comparison with Larger-capacity Backbones

Table 3 presents a comprehensive comparison among TopoSculpt, its holistic-view baseline, and several nnUNet (Isensee et al., 2021) variants with substantially larger model capacities on the ATM'22 pulmonary airway dataset. Although adopting a holistic-view input yields improvements over patch-wise training, the resulting model still exhibits a high β_0 error of 69.00, indicating that global spatial context alone is insufficient to ensure reliable topological consistency. Despite achieving satisfactory clDice scores, all three nnUNet configurations demonstrate weak performance on BD, TD, and β_0 error, revealing that larger backbones provide limited capability in correcting connectivity defects in fine-grained tubular anatomy. In contrast, TopoSculpt achieves a substantial improvement while using the same lightweight backbone as the holistic-view baseline. It reduces the β_0 error from 69.00 to 3.40 and enhances both BD and TD by approximately 10%, outperforming all larger nnUNet configurations by a considerable margin. These results indicate that TopoSculpt plays a critical role in restoring the topological fidelity of tubular structures, which cannot be attained through model scaling or holistic input alone. The scaling behavior of large models on full-resolution volumetric data remains underexplored in this work. Nonetheless, a promising direction is to combine higher-capacity backbones with holistic full-volume processing and TopoSculpt-based refinement, particularly as high-memory GPUs and advanced computational resources become increasingly available.

Table 3: Comparison of different configurations on ATM'22 pulmonary airway. clDice (%), NSDice (%), HD95 (mm), BD (%), TD (%), and β_0 error are reported. Best results in bold, second best underlined.

Method	Patch size	Params (M)	clDice ↑	NSDice ↑	HD95 ↓	BD ↑	TD ↑	β_0 error ↓
3D U-Net	$128 \times 128 \times 128$	4.083	86.27±3.44	95.20±2.20	1.56±1.04	73.25±8.52	74.65±7.83	96.40±27.40
3D U-Net	Holistic view	4.083	90.48±2.92	96.96±1.50	<u>1.07±0.56</u>	<u>82.06±7.44</u>	82.23±6.37	69.00±19.01
nnUNet	$160 \times 128 \times 112$	30.786	<u>90.50±3.77</u>	<u>97.52±1.60</u>	1.14±1.21	83.88±8.69	<u>84.26±7.86</u>	<u>43.77±13.04</u>
nnUNet	$64 \times 64 \times 64$	30.786	88.59±3.51	96.55±2.30	4.00±9.73	81.22±8.50	82.10±8.021	79.10±20.96
nnUNet	$128 \times 128 \times 128$	30.786	90.39±3.82	97.17±3.45	1.28±1.86	83.47±9.98	83.78±9.54	49.70±14.50
3D U-Net w/ TopoSculpt	Holistic view	4.083	**92.27±1.91**	**98.11±0.82**	**1.02±0.29**	**91.33±4.84**	**91.56±4.09**	**3.40±2.44**

G Theoretical Analysis of the TIB Objective

TopoSculpt rests on three design choices: persistent homology is computed on the whole region rather than on patches, the Betti-prior correction is coupled with a topological integrity term in the TIB objective, and persistence is recomputed on a curriculum schedule. This section analyzes these choices theoretically. Beyond characterizing the guarantees of the correction term, the analysis exposes a degeneracy that is intrinsic to Betti-prior persistence losses, namely that deleting a disconnected branch and reconnecting it are rewarded equally, and shows how the integrity term of TIB breaks this tie through an explicit, scale-aware preference for reconnection. The main results are summarized below.

1. **Global nature of Betti constraints.** Patch-wise Betti numbers do not determine the global one, and correct patch-wise targets cannot be specified without the global structure (Proposition G.2 and the discussion following it). Enforcing the prior therefore calls for the whole structure, which the holistic view of TopoSculpt provides.
2. **Topological guarantees.** $\mathcal{L}_{\mathrm{TIB}_{cor}}$ vanishes only if the prediction has the prior topology at every threshold (for a single component, if and only if), and it upper-bounds a threshold-robust version of the β_0 error (Propositions G.3 and G.4). The annotation-free objective used at test time is thus a faithful surrogate of the Betti prior.
3. **Degeneracy of the Betti prior.** The gradient of $\mathcal{L}_{\mathrm{TIB}_{cor}}$ is supported on critical voxels and rewards two moves equally for every superfluous component: *carving* it away or *bridging* it to an older component (Lemma G.6 and Corollary G.8). A Betti prior alone therefore cannot distinguish a false positive from a true branch that has been disconnected.
4. **Integrity-guided selection.** In an idealized setting, the integrity term breaks this tie: it favors bridging over carving, protects the main structure, and its weight β sets the scale

below which fragments may still be erased (Proposition G.10 and Corollary G.11). This is consistent with the higher branch and tree detection rates of the full TIB objective in the ablation study (Table 2).

5. **Network-mediated and lazy refinement.** Refining the network weights rather than the likelihood map can spread the sparse topological gradient over whole regions through the tangent kernel of the network (Proposition G.12). Reused critical voxels give the exact loss and gradient at every recomputation, and the diagram drifts by at most the accumulated change of the likelihood map in between (Proposition G.13), which motivates the curriculum refinement.

G.1 Preliminaries and Formalization of the Correction Term

Likelihood maps and superlevel sets. Let V be the voxel grid of the input region ($|V| = N$), which contains the whole target structure, and fix an adjacency on V (6- or 26-connectivity) that is used both for the persistence computation and for counting components. Let $f = \mathcal{F}_\theta(\mathcal{X}) \in [0,1]^V$ denote the likelihood map, and write $f^i = \mathcal{F}_{\theta^i}(\mathcal{X})$ for the likelihood map at refinement step i. For a threshold τ, the superlevel set $U_\tau(f) = \{v \in V : f(v) \geq \tau\}$ is the segmentation obtained by thresholding f at τ; these sets are nested, $U_\tau(f) \subseteq U_{\tau'}(f)$ for $\tau' \leq \tau$. We write $\beta_0(U)$ for the number of connected components of a voxel set U under the chosen adjacency, i.e., the rank of its zeroth homology group $H_0(U)$ with coefficients in $\mathbb{F}_2$; equivalently, U can be represented as a cubical complex (Kaji et al., 2020), using the V-construction for 6-connectivity and the T-construction for 26-connectivity (Bleile et al., 2022).

Barcodes. As τ decreases from 1 to 0, components of $U_\tau(f)$ appear at local maxima of f and merge. By the elder rule, when two components merge at level τ, the one that appeared later (the one with the lower birth value) dies. Each component thus defines a bar (b_l, d_l) with $1 \geq b_l \geq d_l \geq 0$ that is alive for $\tau \in (d_l, b_l]$; following common practice (Clough et al., 2020), the component that never dies (the essential bar) is assigned $d_l = 0$. The multiset of bars is the 0-dimensional persistence diagram $\mathrm{Dgm}(f)$, and $p_l = b_l - d_l \in [0,1]$ is the persistence of bar l. We index bars by decreasing persistence, $p_1 \geq p_2 \geq \cdots \geq p_n$, ranking the essential bar first in case of ties. Every bar has a birth voxel v_l^{b} with $f(v_l^{\mathrm{b}}) = b_l$ and, unless it is essential, a death voxel v_l^{d} with $f(v_l^{\mathrm{d}}) = d_l$, namely the voxel whose insertion merges the component of v_l^{b} into an older one. Processing the voxels in decreasing order of f with a union–find structure computes $\mathrm{Dgm}(f)$ together with all critical voxels in $\mathcal{O}(N \log N)$ time (Edelsbrunner & Harer, 2022). We use three standard facts (Zomorodian, 2005; Edelsbrunner & Harer, 2022; Cohen-Steiner et al., 2007). First, for every $\tau \in (0,1]$, the Betti number of the segmentation at threshold τ is the number of bars alive at τ,

$$\beta_0\big(U_\tau(f)\big) = \#\{l : d_l < \tau \leq b_l\}. \tag{8}$$

Second, for $0 < \tau' < \tau$, the persistent Betti number $\beta_0^{\tau,\tau'}(f)$, defined as the rank of the map $H_0(U_\tau(f)) \to H_0(U_{\tau'}(f))$ induced by the inclusion, equals the number of components of $U_{\tau'}(f)$ that intersect $U_\tau(f)$ and satisfies

$$\beta_0^{\tau,\tau'}(f) = \#\{l : b_l \geq \tau,\ d_l < \tau'\}. \tag{9}$$

Third, persistence diagrams are stable: for any two likelihood maps f and g, the bottleneck distance satisfies

$$d_B\big(\mathrm{Dgm}(f), \mathrm{Dgm}(g)\big) \leq \|f - g\|_\infty. \tag{10}$$

The correction term. Let $\beta_0^\star$ denote the prior number of connected components of the target structure, e.g., $\beta_0^\star = 1$ for the pulmonary airway (Section 3) and for the CoW cases selected in Section 4. The ideal barcode consists of $\beta_0^\star$ bars of unit persistence, whose total persistence $\beta_0^\star$ is the constant prior $Topo_P$ of Section 3. The $\beta_0^\star$ most persistent bars are the faithful features, with total persistence $Topo_F(f) = \sum_{l \leq \beta_0^\star} p_l$, and all remaining bars are superfluous, with total persistence $Topo_S(f) = \sum_{l > \beta_0^\star} p_l$, so that

$$\mathcal{L}_{\mathrm{TIB}_{cor}}(f) = Topo_P - Topo_F(f) + Topo_S(f) = \sum_{l \leq \beta_0^\star} (1 - p_l) \ + \ \sum_{l > \beta_0^\star} p_l. \tag{11}$$

This is the Betti-prior form of persistence losses (Clough et al., 2020), written with the linear persistence of Section 3. In contrast to losses that match the predicted diagram to the ground-truth

diagram (Hu et al., 2019; Stucki et al., 2023), it requires no annotation and can therefore be used at test time. TIB goes further by coupling this correction with the integrity term $\mathcal{L}_{\text{TIB}_{com}}$; Sections G.4 and G.5 show that this coupling addresses a limitation inherent to the correction term itself. For $\beta_0^\star = 1$, the essential bar is one of the most persistent bars, since its persistence $\max_v f(v)$ satisfies $\max_v f(v) \geq b_l \geq p_l$ for every l; with the ranking convention above, the faithful feature is the essential bar, i.e., the most persistent component, as stated in Section 3.

Remark G.1 (Squared persistence). Clough et al. (2020) use squared persistences, $\sum_{l\leq\beta_0^\star}(1-p_l^2)+\sum_{l>\beta_0^\star} p_l^2$. All results below carry over to this variant with the following changes. The term of each bar in Eq. (13) is multiplied by $2p_l$, the same weights enter Proposition G.12, and the frozen correction term of Proposition G.13 is squared as well. The loss is piecewise quadratic rather than piecewise linear, and its continuity in Lemma G.6(a) follows from the fact that the zero-padded vector of sorted persistences is 2-Lipschitz in the sup-norm by Eq. (10). In Lemma G.6(c), the gradients at f_ϵ converge to, rather than equal, the weighted right-hand side. In Proposition G.4, $\tau-\tau'$ is replaced by $(\tau-\tau')^2$, so that the constant $\frac{1}{3}$ in its last claim becomes $\frac{1}{9}$. In Proposition G.10(a), $1-\lambda$ becomes $(1-\lambda)^2$, and in Corollary G.11(i), the rate -1 becomes -2, so that the carving thresholds discussed after it double.

G.2 Global Nature of Betti Constraints

Patch-wise pipelines evaluate topological quantities on restrictions $Y \cap P$ of a segmentation $Y \subseteq V$ to patches $P \subset V$. The next statement shows that such quantities do not determine the global Betti number.

Proposition G.2 (Non-locality of Betti numbers). *Let a box-shaped region V be covered by two overlapping box-shaped patches, $V = P_1 \cup P_2$, such that $P_1 \setminus P_2$ and $P_2 \setminus P_1$ are non-empty and any two adjacent voxels lie in a common patch. For $Y \subseteq V$, write $Y_j = Y \cap P_j$.*

- *(a) For every $Y \subseteq V$, $\beta_0(Y) = \beta_0(Y_1) + \beta_0(Y_2) - \operatorname{rank}\varphi_Y$, where $\varphi_Y : H_0(Y_1 \cap Y_2) \to H_0(Y_1) \oplus H_0(Y_2)$ is induced by the two inclusions.*
- *(b) There exist $Y, Y' \subseteq V$ with $\beta_0(Y_j) = \beta_0(Y'_j) = 1$ for $j = 1, 2$, but $\beta_0(Y) = 1 \neq 2 = \beta_0(Y')$. Consequently, no function Φ satisfies $\beta_0(Y) = \Phi\big(\beta_0(Y_1), \beta_0(Y_2)\big)$ for all $Y \subseteq V$.*

Proof. (a) Regard Y, Y_1, and Y_2 as graphs whose edges join adjacent voxels. Since every edge of Y lies in Y_1 or in Y_2, the Mayer–Vietoris sequence (Hatcher, 2002) contains the exact segment

$$H_0(Y_1 \cap Y_2) \xrightarrow{\varphi_Y} H_0(Y_1) \oplus H_0(Y_2) \xrightarrow{\psi} H_0(Y) \to 0.$$

Hence $\beta_0(Y) = \operatorname{rank}\psi = \beta_0(Y_1) + \beta_0(Y_2) - \dim\ker\psi$ with $\ker\psi = \operatorname{im}\varphi_Y$. (b) Let $Y = V$, so that $Y_j = P_j$ and all three sets are connected boxes. Pick $a \in P_1 \setminus P_2$ and $c \in P_2 \setminus P_1$; they are not adjacent because they share no patch. For $Y' = \{a, c\}$ we have $Y'_1 = \{a\}$ and $Y'_2 = \{c\}$, whereas $\beta_0(Y') = 2$.

By part (a), patch-wise Betti numbers only bound the global count from above, $\beta_0(Y) \leq \beta_0(Y_1) + \beta_0(Y_2)$; the exact value depends on the gluing term $\operatorname{rank}\varphi_Y$, which encodes how the components seen in the two patches are connected through their overlap. Moreover, the patch-wise target is itself unknown: if the restriction of a connected tree to a patch consists of m peripheral branches that are connected only through ancestors outside the patch, then $\beta_0(Y \cap P) = m$, so imposing $\beta_0 = 1$ on every patch would penalize a correct prediction. TopoSculpt sidesteps both issues: its holistic model predicts the entire structure in a single forward pass (Section 3), so the persistence of the whole likelihood map is computed and backpropagated through one network evaluation, and the global prior is imposed directly rather than through patch-wise proxies.

G.3 Topological Guarantees of the Correction Term

Proposition G.3 (Exactness of the correction term). *$\mathcal{L}_{\text{TIB}_{cor}}(f) \geq 0$, with equality if and only if $\operatorname{Dgm}(f)$ consists of exactly $\beta_0^\star$ bars equal to $(1, 0)$ together with bars of zero persistence. In that case $\beta_0(U_\tau(f)) = \beta_0^\star$ for every $\tau \in (0, 1]$; in particular, the binarized predictions $U_{1/2}(f)$ and $\{f > \frac{1}{2}\}$ have exactly $\beta_0^\star$ connected components. Conversely, for $\beta_0^\star = 1$, if $\beta_0(U_\tau(f)) = 1$ for every $\tau \in (0, 1]$, then $\mathcal{L}_{\text{TIB}_{cor}}(f) = 0$.*

Proof. Every summand in Eq. (11) is non-negative because $p_l \in [0,1]$, so the loss vanishes if and only if $p_l = 1$ for $l \le \beta_0^\star$ and $p_l = 0$ for $l > \beta_0^\star$. Since $0 \le d_l \le b_l \le 1$, $p_l = 1$ forces $(b_l, d_l) = (1,0)$. By Eq. (8), a bar with $b_l = d_l$ is never counted, whereas a bar $(1,0)$ is counted for every $\tau \in (0,1]$; the set $\{f > \frac{1}{2}\}$ equals $U_\tau(f)$ for the smallest value $\tau > \frac{1}{2}$ taken by f. For the converse with $\beta_0^\star = 1$, $U_1(f) \neq \emptyset$ forces $\max_v f(v) = 1$, so the essential bar is $(1,0)$; any further bar with positive persistence would be alive at some $\tau \in (0,1]$ together with the essential bar, contradicting $\beta_0(U_\tau(f)) = 1$.

For $\beta_0^\star \ge 2$ the converse fails. On a path of five voxels with values $(1, \tau_0, 1, 0, \tau_0)$, $\tau_0 \in (0,1)$, the voxel of value 0 keeps the last voxel separate down to level 0, so $\beta_0(U_\tau(f)) = 2$ for every $\tau \in (0,1]$. Yet the two non-essential bars $(1,\tau_0)$ and $(\tau_0, 0)$ have persistences $1-\tau_0$ and τ_0, neither of which equals 1, so the loss with $\beta_0^\star = 2$ is positive; adding $\beta_0^\star - 2$ isolated voxels of value 1 gives a counterexample for any $\beta_0^\star > 2$.

A Betti number at a single threshold is discontinuous in f: raising one isolated voxel from $\frac{1}{2} - \epsilon$ to $\frac{1}{2} + \epsilon$ adds a component to $U_{1/2}(f)$ while changing f by only 2ϵ. Hence no continuous loss can both vanish on all predictions with the correct number of components at $\tau = \frac{1}{2}$ and upper-bound the corresponding β_0 error. The appropriate robust notion is the persistent Betti number of Eq. (9), which counts only components that exist at a threshold τ and remain distinct down to $\tau' < \tau$, i.e., components that do not depend on the choice of the threshold within $[\tau', \tau]$. The correction term controls exactly this quantity.

Proposition G.4 (Robust Betti error bound)**.** *For all* $0 < \tau' < \tau < 1$,

$$\beta_0^\star - \frac{Topo_P - Topo_F(f)}{\min\{\tau', 1-\tau\}} \;\le\; \beta_0^{\tau,\tau'}(f) \;\le\; \beta_0^\star + \frac{Topo_S(f)}{\tau - \tau'}. \tag{12}$$

Consequently, $\left|\beta_0^{\tau,\tau'}(f) - \beta_0^\star\right| \le \mathcal{L}_{\mathrm{TIB}_{cor}}(f) / \min\{\tau - \tau', \tau', 1-\tau\}$. *In particular,* $\mathcal{L}_{\mathrm{TIB}_{cor}}(f) < \frac{1}{3}$ *implies* $\beta_0^{2/3,1/3}(f) = \beta_0^\star$.

Proof. By Eq. (9), $\beta_0^{\tau,\tau'}(f) = \beta_0^\star - m_F + m_S$, where m_F is the number of faithful bars ($l \le \beta_0^\star$) that are not counted and m_S is the number of superfluous bars ($l > \beta_0^\star$) that are counted. A counted bar has $b_l \ge \tau$ and $d_l < \tau'$, hence $p_l > \tau - \tau'$, so that $m_S(\tau - \tau') \le Topo_S(f)$. A faithful bar that is not counted has either $b_l < \tau$, whence $1 - p_l \ge 1 - b_l > 1 - \tau$, or $d_l \ge \tau'$, whence $1 - p_l \ge d_l \ge \tau'$ because $b_l \le 1$; therefore $m_F \min\{\tau', 1-\tau\} \le Topo_P - Topo_F(f)$. This proves Eq. (12). The remaining claims follow from $|\beta_0^{\tau,\tau'}(f) - \beta_0^\star| = |m_S - m_F| \le \max\{m_F, m_S\}$, the left-hand side being an integer.

Remark G.5. Proposition G.4 and the forward direction of Proposition G.3 use only the structure of the barcode. They therefore hold verbatim in higher homology dimensions (e.g., for loops, β_1), with $\beta_0^\star$ replaced by the corresponding prior. The converse in Proposition G.3 relies on the essential bar of H_0 and does not extend in this way.

G.4 Critical Voxels and the Carve–bridge Degeneracy

We call f *generic* if its voxel values are pairwise distinct and $p_{\beta_0^\star} > p_{\beta_0^\star + 1}$, with $p_{\beta_0^\star + 1} := 0$ if no further bar exists; for $\beta_0^\star = 1$, the second condition follows from the first. Let $\mathbf{e}_v \in \mathbb{R}^V$ denote the indicator vector of voxel v.

Lemma G.6 (Differential structure of the correction term)**.** *The correction term* $\mathcal{L}_{\mathrm{TIB}_{cor}}$ *has the following properties.*

(a) $\mathcal{L}_{\mathrm{TIB}_{cor}}$ *is continuous and piecewise linear on* $[0,1]^V$*; in particular, it is Lipschitz continuous.*

(b) If f *is generic, the critical voxels* $v_l^{\mathrm{b}}, v_l^{\mathrm{d}}$ *are constant on a neighborhood of* f*, and on this neighborhood* $\mathcal{L}_{\mathrm{TIB}_{cor}}$ *is differentiable with*

$$\nabla_f \mathcal{L}_{\mathrm{TIB}_{cor}}(f) \;=\; \sum_{l \le \beta_0^\star} \left(\mathbf{e}_{v_l^{\mathrm{d}}} - \mathbf{e}_{v_l^{\mathrm{b}}}\right) \;+\; \sum_{l > \beta_0^\star} \left(\mathbf{e}_{v_l^{\mathrm{b}}} - \mathbf{e}_{v_l^{\mathrm{d}}}\right), \tag{13}$$

where $\mathbf{e}_{v_l^{\mathrm{d}}} := 0$ *for the essential bar.*

(c) *Let $\beta_0^\star = 1$ and let f have ties. Fix a strict total order $\prec$ on V in which u precedes v whenever $f(u) > f(v)$ (e.g., ties broken by voxel index). If the union–find sweep and the elder rule follow $\prec$, all pairs are retained (including zero-length pairs, which may now have $v_l^{\mathrm{b}} \neq v_l^{\mathrm{d}}$), and the essential bar is taken as the faithful one, then the right-hand side of Eq. (13) belongs to the Clarke subdifferential of $\mathcal{L}_{\mathrm{TIB}_{cor}}$ at f (Clarke, 1990).*

(d) *For every non-essential bar with $p_l > 0$ of a generic f, the birth voxel v_l^{b} is a local maximum of f, and*

$$d_l = \max_{\rho} \min_{v \in \rho} f(v), \tag{14}$$

where ρ ranges over the voxel paths from v_l^{b} to the set $\{f > b_l\}$; the death voxel v_l^{d} is the minimizer of f on a max–min path, i.e., the bottleneck of the most plausible connection of the component to an older one.

Proof. (b) The union–find computation of $\mathrm{Dgm}(f)$ depends on f only through the order of its voxel values, which is locally constant at a generic f; hence so are the critical voxels, and each persistence $p_l = f(v_l^{\mathrm{b}}) - f(v_l^{\mathrm{d}})$ is locally a fixed linear function of f. Since $p_{\beta_0^\star} > p_{\beta_0^\star+1}$, the split into faithful and superfluous bars is locally constant as well, and differentiating Eq. (11) yields Eq. (13). At a generic f, bars of zero persistence have $v_l^{\mathrm{b}} = v_l^{\mathrm{d}}$ and contribute nothing. (a) By the argument in (b), $\mathcal{L}_{\mathrm{TIB}_{cor}}$ is linear on each of the finitely many polyhedral cells on which the order of the voxel values and the ranking of the persistences are fixed. It is continuous because $\mathcal{L}_{\mathrm{TIB}_{cor}}(f) = \beta_0^\star + \sum_l p_l - 2\sum_{l \le \beta_0^\star} p_l$, where the total persistence $\sum_l p_l = \sum_v f(v) - W(f)$ is continuous, $W(f)$ being the weight of a maximum spanning tree of the adjacency graph with edge weights $\min\{f(u), f(v)\}$ (the union–find pass is Kruskal's algorithm), and where the sum of the $\beta_0^\star$ largest persistences is continuous by the stability bound in Eq. (10). A continuous piecewise-linear function with finitely many pieces is Lipschitz. (c) Let $h : V \to [0,1]$ be strictly decreasing along $\prec$ and $f_\epsilon = (1-\epsilon)f + \epsilon h \in [0,1]^V$. For small $\epsilon > 0$, the values of f_ϵ are pairwise distinct and decrease along $\prec$, so the union–find sweep of f_ϵ produces exactly the pairs of the tie-broken computation at f. For $\beta_0^\star = 1$, f_ϵ is generic and its faithful bar is the essential one, so $\nabla\mathcal{L}_{\mathrm{TIB}_{cor}}(f_\epsilon)$ equals the right-hand side of Eq. (13). As $f_\epsilon \to f$, this vector is a limit of gradients at points of differentiability converging to f and hence lies in the Clarke subdifferential. (d) For $\tau \in (d_l, b_l]$, the component of v_l^{b} in $U_\tau(f)$ is alive, so its oldest voxel is v_l^{b} and it contains no voxel of $\{f > b_l\}$. At $\tau = d_l$ it merges with an older component, which contains such a voxel. Hence d_l is the largest τ for which v_l^{b} is connected to $\{f > b_l\}$ within $U_\tau(f)$, which is the right-hand side of Eq. (14); the merge happens through the voxel v_l^{d}, the lowest voxel of a max–min path. Finally, if v_l^{b} had a neighbor of higher value, its component would merge at birth and p_l would vanish.

Remark G.7 (Well-posedness). Under the conditions of Lemma G.6(c), i.e., for $\beta_0^\star = 1$ with consistent tie-breaking and all zero-length pairs retained (which some libraries discard by default), the gradient of Eq. (13) is a valid Clarke subgradient of $\mathcal{L}_{\mathrm{TIB}_{cor}}$ with respect to the likelihood map, even when f has ties, as is common for saturated sigmoid outputs. When several components are saturated at exactly 1, however, the tie-break decides which of them is treated as faithful, and the subgradient may then lower the peak of the main structure instead of that of a fragment. Composed with a network built from standard layers (convolutions, normalization, ReLU or sigmoid activations, pooling), $\theta \mapsto \mathcal{L}_{\mathrm{TIB}_{cor}}(\mathcal{F}_\theta(\mathcal{X}))$ is locally Lipschitz and definable in an o-minimal structure, which is the setting in which subgradient methods have been analyzed (Davis et al., 2020), including for persistence-based objectives (Carrière et al., 2021). We do not claim a convergence guarantee for the full refinement, which additionally involves the hard threshold of the structural term, a reference prediction that changes during refinement, and lazily recomputed critical voxels.

Corollary G.8 (Carve–bridge degeneracy)**.** *Let f be generic and let $l > \beta_0^\star$ be a superfluous bar with $p_l > 0$. Its term in Eq. (13) is $\mathbf{e}_{v_l^{\mathrm{b}}} - \mathbf{e}_{v_l^{\mathrm{d}}}$, so a descent step on $\mathcal{L}_{\mathrm{TIB}_{cor}}$ simultaneously (i)* carves*: it lowers the peak $f(v_l^{\mathrm{b}})$ of the fragment born at v_l^{b}, driving the fragment towards deletion; and (ii)* bridges*: it raises $f(v_l^{\mathrm{d}})$, the bottleneck of the most plausible path linking the fragment to an older component (Lemma G.6(d)), driving the fragment towards reconnection. Both moves shorten bar l at the same unit rate. Hence $\mathcal{L}_{\mathrm{TIB}_{cor}}$ distinguishes the two moves only through their side effects on other bars; it does not prefer reconnecting a disconnected branch to deleting it (see Proposition G.10(a) for an exact statement).*

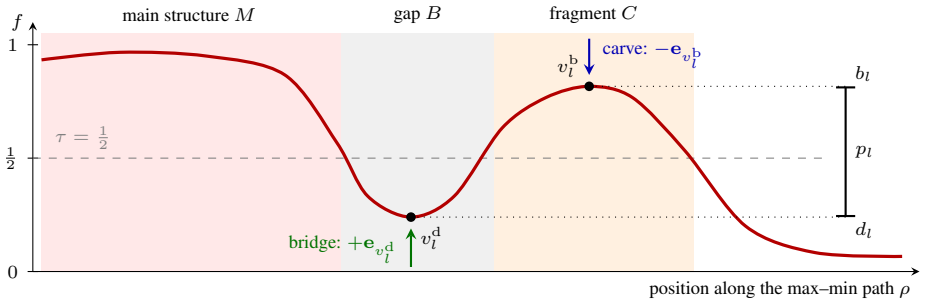

Figure 11: Illustration of Corollary G.8 along the max–min path ρ of Lemma G.6(d). The superfluous bar l is born at the peak v_l^{b} of the fragment C and dies at the bottleneck v_l^{d} of the gap B. Its gradient term offers two moves that shorten the bar at the same rate: carving (lowering the peak, which drives C towards deletion) and bridging (raising the bottleneck, which drives C towards reconnection with M). $\mathcal{L}_{\mathrm{TIB}_{cor}}$ is indifferent between them, whereas the integrity term penalizes region-level carving at first order and therefore favors bridging (Proposition G.10 and Corollary G.11(i)).

Corollary G.8, illustrated in Figure 11, gives a formal account of the shortcut learning that arises when only the correction term is enforced (Section 3 and Figure 1(a.2)): nothing in $\mathcal{L}_{\mathrm{TIB}_{cor}}$ prevents the refinement from erasing genuine but disconnected branches. The degeneracy is a property of the correction term itself and is therefore shared by Betti-prior persistence losses of the form in Eq. (11), such as that of Clough et al. (2020). The integrity term of TIB is designed to break this tie, as shown next.

G.5 Integrity-guided Selection between Carving and Bridging

The integrity term $\mathcal{L}_{\mathrm{TIB}_{com}}$ breaks the tie of Corollary G.8 in favor of reconnection. To obtain closed-form statements, we analyze $\mathcal{L}_{\mathrm{TIB}}$ as a function of an edited likelihood map F for a fixed reference $F^i = f^i$, as in Section 3, in the following idealized configuration, which Figure 11 illustrates.

Assumption G.9. (i) The reference prediction is binary, $F^i = \mathbf{1}_{M\cup C}$, where the main structure M and a fragment C are disjoint, each connected, and not 26-adjacent to each other, and $\beta_0^\star = 1$. (ii) The gap $B \subseteq V \setminus (M \cup C)$ is a set of voxels such that $M \cup B \cup C$ is connected. (iii) The structural term uses the soft skeleton of clDice (Shit et al., 2021) with a fixed number of iterations and soft intersections $|A_1 \cap A_2| := \sum_v A_1(v)A_2(v)$ without smoothing constants, as in Section 3. (iv) The skeleton masses S_M, s_C, and S' defined below are positive.

Because the soft skeleton is composed of min/max pooling operations within $3 \times 3 \times 3$ neighborhoods and ReLU operations, on a binary mask it returns a binary map supported inside the mask, and it acts separately on components that are not 26-adjacent. Let $K = \text{soft-skel}(\mathbf{1}_{M\cup C})$, $S = \sum_v K(v)$, $s_C = \sum_{v\in C} K(v)$, and $S_M = S - s_C$ be the skeleton masses of the reference, and let $K' = \text{soft-skel}(\mathbf{1}_{M\cup B\cup C})$, $S' = \sum_v K'(v)$, and $s'_B = \sum_{v\in B} K'(v)$ be those of the bridged mask. The smoothing constant used in some clDice implementations changes the expressions below by a relative amount of order $1/S$. For $\lambda \in [0, 1]$ we compare three edits of the reference:

$$\text{carve}_\lambda(C):\ F^i - \lambda \mathbf{1}_C, \qquad \text{bridge}_\lambda(B):\ F^i + \lambda \mathbf{1}_B, \qquad \text{carve}_\lambda(M):\ F^i - \lambda \mathbf{1}_M. \tag{15}$$

For an edited likelihood map F, we denote by

$$\Delta\mathcal{L}_{\text{vox}} = \frac{\alpha}{N}\|F - F^i\|^2, \qquad \Delta\mathcal{L}_{\text{str}} = \beta\left(1 - \frac{2\,Struc_{prec}\,Struc_{sens}}{Struc_{prec} + Struc_{sens}}\right) \tag{16}$$

the two parts of $\mathcal{L}_{\mathrm{TIB}_{com}}$, evaluated with F in place of $\mathcal{F}_{\theta^{i+1}}(\mathcal{X})$ and F^i in place of $\mathcal{F}_{\theta^i}(\mathcal{X})$, and measured relative to the unedited prediction, for which $Struc_{prec} = Struc_{sens} = 1$.

Proposition G.10 (Integrity-guided selection)**.** *Under Assumption G.9, the following holds.*

(a) Correction term. *For all three edits and all $\lambda \in [0, 1]$, $\mathcal{L}_{\mathrm{TIB}_{cor}} = 1 - \lambda$. The correction term does not distinguish the edits, not even carving the main structure.*

(b) Partial edits *($0 < \lambda < \frac{1}{2}$). The integrity costs are*

	$\text{carve}_\lambda(C)$	$\text{bridge}_\lambda(B)$	$\text{carve}_\lambda(M)$
$\Delta\mathcal{L}_{\text{vox}}$	$\alpha\lambda^2\lvert C\rvert/N$	$\alpha\lambda^2\lvert B\rvert/N$	$\alpha\lambda^2\lvert M\rvert/N$
$\Delta\mathcal{L}_{\text{str}}$	$\dfrac{\beta\lambda s_C}{2S-\lambda s_C}$	0	$\dfrac{\beta\lambda S_M}{2S-\lambda S_M}$

(c) Complete edits *($\lambda = 1$). $\Delta\mathcal{L}_{\text{vox}}$ is as in (b) with $\lambda = 1$, and $\Delta\mathcal{L}_{\text{str}}$ equals $\beta s_C/(2S - s_C)$, $\beta s'_B/(2S' - s'_B)$, and $\beta S_M/(2S - S_M)$ for the three edits, respectively.*

Proof. (a) In F^i, the sets M and C are two components of value 1 and all other voxels have value 0, so $\text{Dgm}(F^i)$ consists of two bars $(1, 0)$ and bars of zero persistence, and $\mathcal{L}_{\text{TIB}_{cor}}(F^i) = 0 + 1$. Carving C turns its bar into $(1 - \lambda, 0)$. Bridging makes C merge with M through B at level λ, turning one of the two bars into $(1, \lambda)$. Carving M turns the bar of M into $(1 - \lambda, 0)$, while C becomes the essential component. In all three cases, the non-zero persistences are 1 and $1 - \lambda$, and the loss equals $1 - \lambda$. (b) For $\lambda < \frac{1}{2}$, none of the edits changes the binarized mask $\{F > \frac{1}{2}\} = M \cup C$, so the skeleton is K before and after the edit and $SP^{i+1} = K \odot F$. Since K is supported in $M \cup C$, where $F^i = 1$, we get $Struc_{prec} = \langle K \odot F, F^i\rangle/\langle K, F\rangle = 1$ for all three edits, whereas $Struc_{sens} = \langle K, F\rangle/S$ equals $1 - \lambda s_C/S$, 1, and $1 - \lambda S_M/S$, respectively; the bridge lies outside the support of K. Substituting $Struc_{prec} = 1$ and $Struc_{sens} = 1 - x$ into $1 - 2\,Struc_{prec}\,Struc_{sens}/(Struc_{prec} + Struc_{sens}) = x/(2-x)$ gives the structural costs. The voxel-wise costs follow from $\|F - F^i\|^2 = \lambda^2|C|$, $\lambda^2|B|$, and $\lambda^2|M|$. (c) Carving C completely leaves the mask M, whose skeleton is the restriction of K to M because M and C are not 26-adjacent; hence $Struc_{prec} = 1$ and $Struc_{sens} = 1 - s_C/S$, and carving M is symmetric. Bridging completely yields the mask $M \cup B \cup C$ with skeleton K'. Now $Struc_{sens} = \langle K, F\rangle/S = 1$ and $Struc_{prec} = \langle K', F^i\rangle/S' = 1 - s'_B/S' =: 1 - y$ because F^i vanishes on B; the identity used in (b), with precision and sensitivity exchanged, gives the cost $\beta y/(2 - y)$.

Corollary G.11 (Preference for reconnection)**.** *Under Assumption G.9, let $\mathcal{L}_{\text{TIB}} = \mathcal{L}_{\text{TIB}_{cor}} + \mathcal{L}_{\text{TIB}_{com}}$ with $\alpha, \beta > 0$.*

(i) The one-sided directional derivatives of $\mathcal{L}_{\text{TIB}}$ at F^i along carving C, bridging B, and carving M are

$$-1 + \frac{\beta s_C}{2S}, \qquad -1, \qquad -1 + \frac{\beta S_M}{2S}, \tag{17}$$

respectively. Bridging is always a descent direction, whereas carving the fragment (resp. the main structure) is a descent direction if and only if $s_C < 2S/\beta$ (resp. $S_M < 2S/\beta$).

(ii) For $0 < \lambda < \frac{1}{2}$, bridging is strictly preferred, i.e., $\mathcal{L}_{\text{TIB}}(\text{bridge}_\lambda(B)) < \mathcal{L}_{\text{TIB}}(\text{carve}_\lambda(C))$, whenever $\alpha\lambda(|B| - |C|)/N < \beta s_C/(2S)$; in particular, for every such λ if $|B| \le |C|$.

(iii) Among complete edits, bridging is strictly preferred to carving if $s'_B/S' \le s_C/S$ and $|B| \le |C|$, with at least one of the two inequalities strict.

Proof. All three edits change $\mathcal{L}_{\text{TIB}_{cor}}$ by the same amount (Proposition G.10(a)), so the comparisons reduce to those of $\Delta\mathcal{L}_{\text{vox}} + \Delta\mathcal{L}_{\text{str}}$. Claim (i) follows from Proposition G.10(b), since the voxel-wise costs are quadratic in λ and

$$\frac{d}{d\lambda}\left[\frac{\lambda s}{2S - \lambda s}\right]_{\lambda=0} = \frac{s}{2S}.$$

Claim (ii) follows from $\lambda s_C/(2S - \lambda s_C) \ge \lambda s_C/(2S)$, and claim (iii) from Proposition G.10(c) and the monotonicity of $x \mapsto x/(2 - x)$ on $[0, 1]$, applied to $x = s'_B/S'$ and $x = s_C/S$.

Interpretation. Corollary G.11 shows that the integrity term, the component that distinguishes TIB from Betti-prior losses, acts as a selection rule with a clear geometric meaning rather than as a generic regularizer. (1) Without the integrity term ($\alpha = \beta = 0$), deleting a fragment, reconnecting it, and even erasing the main structure decrease the objective at the same rate (Proposition G.10(a)). The correction term alone therefore does not exclude the destructive shortcuts visible

in Figure 1(a.2), consistent with the lower branch and tree detection rates obtained without the integrity term (Table 2 and Figure 10). (2) The structural part of $\mathcal{L}_{\mathrm{TIB}_{com}}$ resists carving at first order, in proportion to the skeleton mass that would be lost, but it exerts no gradient on a bridge before the bridge crosses the binarization threshold. Bridging is therefore favored over carving, and the main structure is protected as soon as it carries more than a fraction $2/\beta$ of the skeleton mass. (3) The weight β acts as a scale parameter: fragments with $s_C < 2S/\beta$ may still be carved. With $\beta = 10^3$ (Section 4), this concerns only fragments that carry less than 0.2% of the total skeleton mass; in the late curriculum stage, where $\mathcal{L}_{\mathrm{TIB}_{com}}$ is scaled by $\gamma = 0.1$, the threshold relaxes to 2%. (4) The voxel-wise part is quadratic in the edit and does not act at first order; it penalizes large changes between consecutive predictions and, at second order, favors the edit that modifies fewer voxels.

These statements describe the objective along three edit directions of interest in an idealized configuration, not the exact trajectory of the refinement. At a binary prediction, for instance, the subgradient of Lemma G.6(c) involves single tie-broken voxels rather than the whole fragment or gap, and region-level edits arise only through the network (Proposition G.12). Likewise, the complete-edit costs in Proposition G.10(c) compare edits relative to a fixed reference, as in a single proximal step; if the reference is updated at every step, the structural cost of a bridge is incurred gradually as its voxels cross the threshold. Two ingredients of the mechanism nevertheless hold for any likelihood map. First, the structural term has zero gradient on voxels that lie outside the skeletons of both the current and the reference prediction, and thus on a bridge that has not yet crossed the threshold. Second, its sensitivity part $Struc_{sens}$ has the positive gradient $SP^i(v)/\sum_u SP^i(u)$ on the reference skeleton; at a binary reference, the precision part has zero gradient there, so the structural term as a whole has gradient $-\beta/(2S)$ on every skeleton voxel and resists any decrease of the likelihood on the skeleton. With several fragments, the first-order costs in Corollary G.11(i) add up over the fragments.

G.6 Network-mediated Refinement and Curriculum Efficiency

TopoSculpt updates the network parameters θ rather than the likelihood map itself. The following first-order statement makes explicit what this parametrization does with the sparse gradient of Lemma G.6.

Proposition G.12 (Function-space view of a refinement step)**.** *Let $f(\theta) = \mathcal{F}_\theta(\mathcal{X})$ be continuously differentiable near θ^i with a locally Lipschitz Jacobian, which for piecewise-smooth networks such as ReLU U-Nets holds for almost every θ^i. Let $J = \partial f/\partial\theta$ at θ^i, let $\mathcal{L}$ be differentiable at $f(\theta^i)$ with gradient g, and let $\theta^{i+1} = \theta^i - \eta D J^\top g$ for a symmetric positive definite matrix D ($D = I$ for gradient descent; a diagonal D for adaptive methods such as AdamW, ignoring momentum and weight decay). Then, with the kernel $\Theta_D = JDJ^\top \succeq 0$,*

$$f(\theta^{i+1}) = f(\theta^i) - \eta\,\Theta_D\,g + \mathcal{O}(\eta^2), \qquad \mathcal{L}\big(f(\theta^{i+1})\big) = \mathcal{L}\big(f(\theta^i)\big) - \eta\,g^\top\Theta_D\,g + o(\eta). \tag{18}$$

In particular, $\mathcal{L}$ does not increase at first order, and it decreases at first order whenever $\Theta_D\,g \neq 0$. For $\mathcal{L} = \mathcal{L}_{\mathrm{TIB}_{cor}}$ at a generic f, Eq. (13) gives

$$\Theta_D\,g = \sum_{l>\beta_0^\star} \Theta_D\big(\mathbf{e}_{v_l^{\mathrm{b}}} - \mathbf{e}_{v_l^{\mathrm{d}}}\big) - \sum_{l\le\beta_0^\star} \Theta_D\big(\mathbf{e}_{v_l^{\mathrm{b}}} - \mathbf{e}_{v_l^{\mathrm{d}}}\big), \tag{19}$$

with $\mathbf{e}_{v_l^{\mathrm{d}}} := 0$ for the essential bar.

Proof. Expand f around θ^i and $\mathcal{L}$ around $f(\theta^i)$ to first order. Since $\Theta_D \succeq 0$, $g^\top\Theta_D\,g \geq 0$, with equality if and only if $\Theta_D\,g = 0$.

The update of the likelihood map is thus a superposition of kernel columns $\Theta_D\mathbf{e}_v$ centered at the critical voxels, with entries

$$(\Theta_D\mathbf{e}_v)(u) = \big\langle \nabla_\theta f_u,\, D\,\nabla_\theta f_v \big\rangle, \qquad u \in V. \tag{20}$$

This clarifies why TopoSculpt refines the network weights rather than the likelihood map. Optimizing the likelihood map directly ($\Theta_D = I$) would move only the critical voxels themselves, whereas the network couples each critical voxel to every voxel u with $(\Theta_D\mathbf{e}_v)(u) \neq 0$, as described by the empirical neural tangent kernel (Jacot et al., 2018). A correction signal at a single bottleneck voxel

can therefore become an edit of a whole region, with the network acting as an image prior (Ulyanov et al., 2018). Because of weight sharing, these couplings are not confined to a neighborhood of the critical voxel; the integrity term complements the network prior by penalizing edits that alter the rest of the structure.

Curriculum refinement. In the late stage of the curriculum refinement ($i \geq t$), persistent homology is recomputed only at the sampled steps $j(i) = \lfloor i/k \rfloor k$. We consider the implementation in which the critical voxels obtained at step $j(i)$ are reused, so that the correction term at step i is evaluated on the current likelihood map at these voxels,

$$\widehat{\mathcal{L}}^{(j)}(f) = \sum_{l \leq \beta_0^\star} \big(1 - f(v_l^{\mathrm{b}}) + f(v_l^{\mathrm{d}})\big) + \sum_{l > \beta_0^\star} \big(f(v_l^{\mathrm{b}}) - f(v_l^{\mathrm{d}})\big), \tag{21}$$

where the critical voxels and the faithful/superfluous split are those of f^j with $j = j(i)$, and $f(v_l^{\mathrm{d}}) := 0$ for the essential bar.

Proposition G.13 (Consistency of lazily recomputed critical voxels)**.** *Let $j = j(i)$ be the most recent recomputation step.*

(a) $\widehat{\mathcal{L}}^{(j)}(f^j) = \mathcal{L}_{\mathrm{TIB}_{cor}}(f^j)$. If f^j is generic, $\widehat{\mathcal{L}}^{(j)}$ and $\mathcal{L}_{\mathrm{TIB}_{cor}}$ coincide on a neighborhood of f^j, so that $\nabla\widehat{\mathcal{L}}^{(j)}(f^j) = \nabla\mathcal{L}_{\mathrm{TIB}_{cor}}(f^j)$.

(b) For $i > j$, $d_B\big(\mathrm{Dgm}(f^i), \mathrm{Dgm}(f^j)\big) \leq \|f^i - f^j\|_\infty \leq \sum_{r=j}^{i-1} \|f^{r+1} - f^r\|_\infty$.

Proof. (a) The first identity holds by construction. At a generic f^j, the critical voxels and the faithful/superfluous split are constant on a neighborhood (Lemma G.6(b)), on which $\mathcal{L}_{\mathrm{TIB}_{cor}}$ therefore has the form of $\widehat{\mathcal{L}}^{(j)}$. (b) This follows from the stability bound in Eq. (10) and the triangle inequality.

Proposition G.13 provides the rationale for the curriculum refinement of TopoSculpt. The reused critical voxels give the exact correction term and gradient at every recomputation, and in between, the persistence diagram can move away from the one they encode by at most the accumulated change of the likelihood map in the sup-norm. Frozen pairs can nevertheless become stale, for example once a fragment has merged with the main structure, and the sampling interval k bounds how long such a pair is reused. Recomputing densely while the likelihood map changes rapidly, and sparsely once the Betti numbers stabilize (Figure 8), is thus a natural trade-off between accuracy and cost. With $t = 30$, $T = 90$, and $k = 3$ (Section 4), this amounts to about $t + (T - t)/k = 50$ persistence computations instead of 90, each of which costs $\mathcal{O}(N \log N)$ on a whole-region volume with millions of voxels.