

A142 · Consistent Reeb Graph Estimation for Unsupervised Cell-State Topology Discovery

Presenter: Andrew Steindl — Yale

Authors: Andrew J. Steindl, João Felipe Rocha, Dhananjay Bhaskar, Smita Krishnaswamy

Session: Day 2 · Fri Oct 2, 2026 · 2:15–4:15 PM

Single-cell measurements are high-dimensional samples from latent biological state spaces whose organization may contain branches, cycles, disconnected populations, or mixtures of these structures. Detecting this organization without first assuming clusters or a tree remains a central challenge for automated analysis. We introduce a Reeb graph estimator that summarizes the evolution of connected level sets across a sampled data manifold. The estimator requires only a neighborhood graph and scalar filter; in our unsupervised implementation these are a k -nearest-neighbor graph and its first nontrivial diffusion eigenfunction. The resulting Reeb edges index coherent cellular trajectories, including branching and cyclic organization, without assuming either structure in advance. The neighborhood size k is the only user-selected hyperparameter; midpoint thresholds and degree-2 reduction are canonical.

We prove a deterministic consistency theorem. If increasingly dense samples approximate a compact manifold, the observed scalar converges to a Morse function, and the neighborhood graphs approximate the underlying manifold, then the estimated filtered Reeb graphs converge to the manifold Reeb graph in functional Gromov-Hausdorff distance. Thus, once the neighborhood graph and filter are adequate, no additional Reeb-specific statistical estimation problem remains. In included experiments, the implementation recovers the Reeb graph of a sampled torus and a multi-branching organization in the expert-annotated single-cell dataset SCD-0001; its cellular decomposition assigns cells directly to recovered Reeb edges. The method provides an interpretable, topology-flexible interface between diffusion geometry and downstream analysis of developmental trajectories and cyclic biological processes.

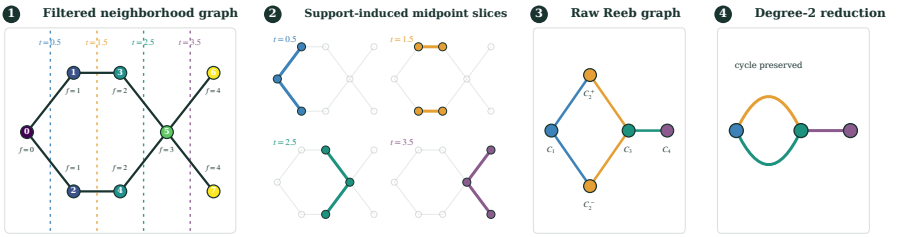
Consistent Reeb Graph Estimation for Unsupervised Cell-State Topology Discovery

Andrew J. Steindl^{1,*}, João Felipe Rocha¹, Dhananjay Bhaskar², and Smita Krishnaswamy¹

¹Yale University, New Haven, Connecticut, USA; ²University of Wisconsin–Madison, Madison, Wisconsin, USA

*Presenting author: drew.steindl@yale.edu

Single-cell measurements are high-dimensional samples from latent biological state spaces whose organization may contain branches, cycles, disconnected populations, or mixtures of these structures. Detecting this organization without first assuming clusters or a tree remains a central challenge for automated analysis. We introduce a Reeb graph estimator that summarizes the evolution of connected level sets across a sampled data manifold. It requires only a neighborhood graph and scalar filter; in our unsupervised implementation these are a k -nearest-neighbor graph and its first nontrivial diffusion eigenfunction. Reeb edges index coherent cellular trajectories, including branching and cyclic organization, without assuming either structure in advance. The neighborhood size k is the only user-selected hyperparameter.



Canonical thresholds are midpoints between observed filter values. At each threshold, connected components of the induced graph on crossing-edge endpoints become Reeb nodes.

Figure 1. Midpoint slices of a filtered neighborhood graph produce a raw Reeb graph whose degree-2 reduction preserves cycles.

No cover resolution, overlap, bin count, level-set bandwidth, or simplification cutoff is introduced. Thresholds are canonical, degree-2 reduction is parameter-free, and neighborhood-graph construction is the remaining modeling choice.

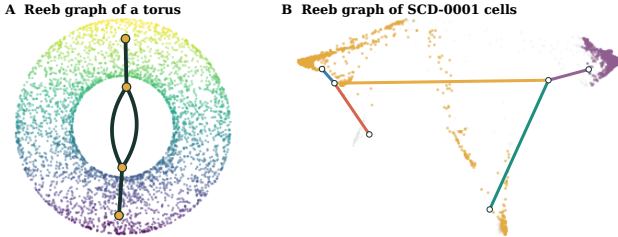


Figure 2. (A) A sampled upright torus under a generic vertical height, with its recovered Reeb graph overlaid; the two middle edges encode the loop. (B) On a fixed 4,000-cell subsample of expert-annotated multi-branching dataset SCD-0001, a $k = 28$ graph and its first nontrivial diffusion eigenfunction yield a five-edge Reeb tree. Cells and graph edges share colors by Reeb-edge assignment.

Theorem 1 (Consistency). *Let $\mathcal{M}$ be a compact manifold, let $f: \mathcal{M} \rightarrow \mathbb{R}$ be Morse, let $\mathcal{R}_f(\mathcal{M})$ denote its Reeb graph, and let d_{GH} denote functional Gromov–Hausdorff distance. If X_n are increasingly dense samples of $\mathcal{M}$, $\hat{f}_n \rightarrow f$, and neighborhood graphs G_n become locally faithful to $\mathcal{M}$, then $d_{\text{GH}}(\hat{\mathcal{R}}(G_n, \hat{f}_n), \mathcal{R}_f(\mathcal{M})) \rightarrow 0$. Here $\hat{\mathcal{R}}$ is our estimated Reeb graph. Thus, once the scalar filter is consistent, Reeb estimation reduces to constructing a locally faithful neighborhood graph; kNN is one such construction.*

Context. This result supplies a consistency foundation for the scReebTower method introduced with scShapeBench. Code and formal proof artifacts are available at <https://github.com/KrishnaswamyLab/scReebTower>.