FACULTY OF ENGINEERING
AND ARCHITECTURE

DEPARTMENT OF ELECTRONICS AND INFORMATION SYSTEMS – AIDA/DAMBI

Robin Vandaele, Yvan Saeys, and Tijl De Bie

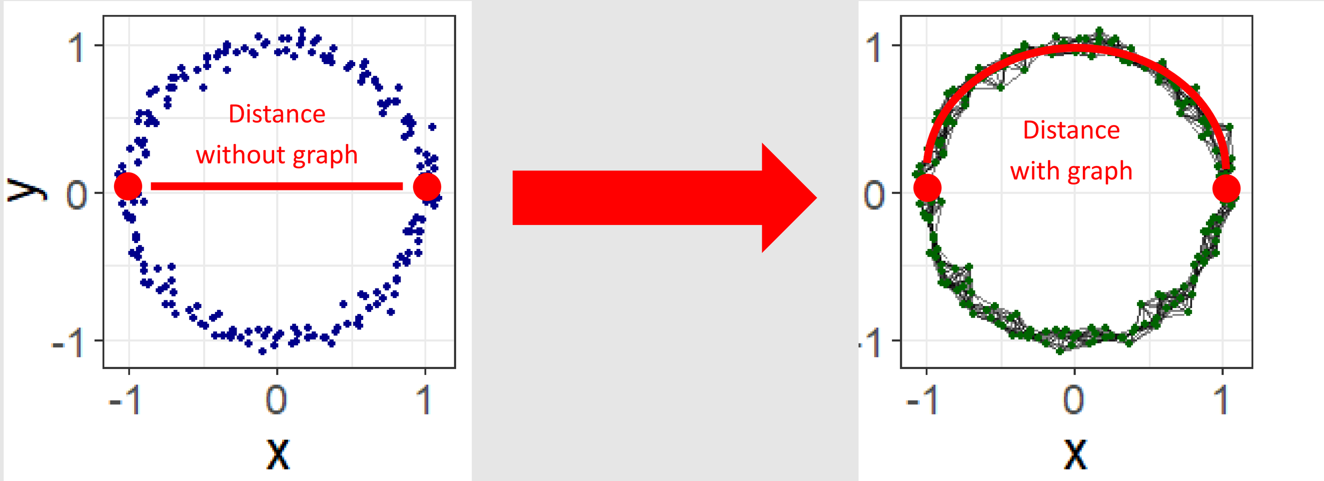
Graph Approximations to Geodesics on Metric Graphs

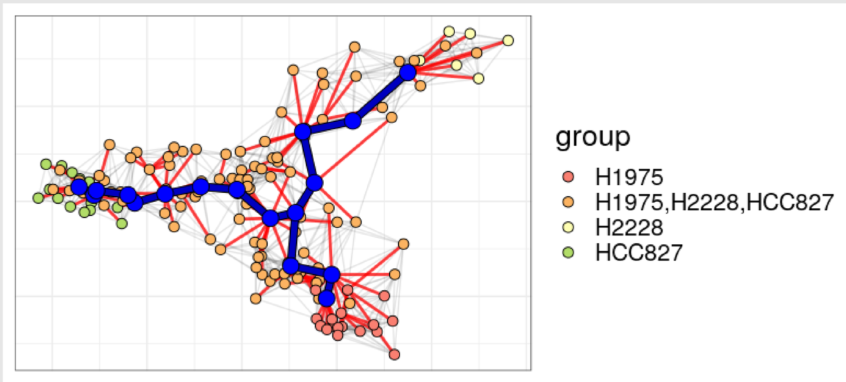
MOTIVATION

Proximity graphs are often used to approximate geodesics

Used for various dimensionality reduction methods

- Laplacian Eigenmaps
- Locally Linear Embedding
- ISOMAP
- Maximum Variance Unfolding
- Local Tangent Space Alignment
- …



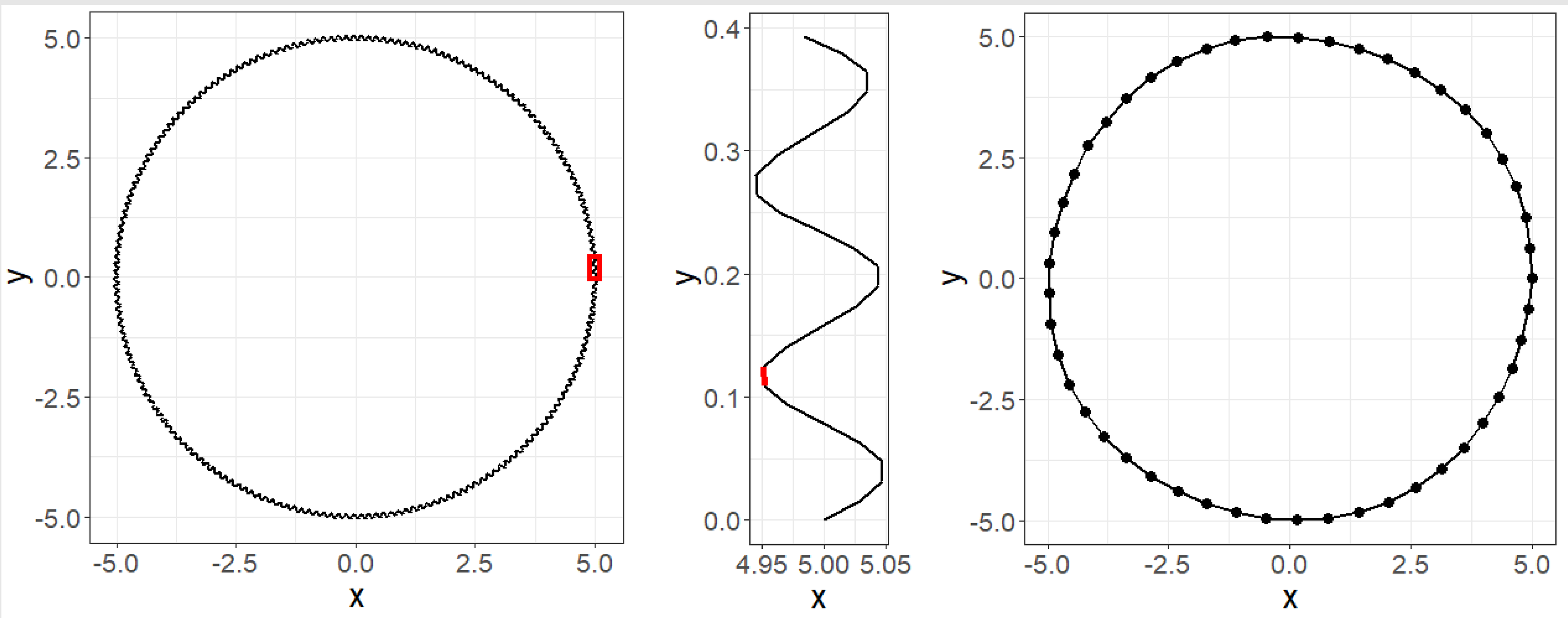


Used as intermediate representations for topological inference [1, 2, 3]

THE PROBLEM

Existing results for approximating geodesic distances [4]

- Only apply to smooth submanifolds of $\mathbb{R}^n$
  - No singularities (such as bifurcations) allowed
- Conditions under which approximated and true geodesic distances are close are stringent
  - Must approximate all local patterns well
  - However, global patterns may be approximated well without local patterns



OUR APPROACH

New geometric characteristics for metric graphs

Definition 3. *Given a connected metric graph M. For any $\epsilon > 0$ we define the* branch separation $s_\epsilon(M) \in \mathbb{R}^+ \cup \{\infty\}$ of M at resolution $1/\epsilon$ *as*

$$s_\epsilon(M) := \sup\{s \in \mathbb{R} : \|x - y\| < s \implies d_M(x, y) \leq \epsilon\}.$$

Definition 4. *Given a connected metric graph M. For any $0 \leq \epsilon' < \epsilon$ we define the* linearity of M between resolutions $1/\epsilon$ and $1/\epsilon' \in \mathbb{R}^+ \cup \{\infty\}$ *as*

$$\lambda_{\epsilon',\epsilon}(M) := \sup\{\lambda \in \mathbb{R} : \epsilon' \leq d_M(x, y) \leq \epsilon \implies \lambda d_M(x, y) \leq \|x - y\|\}.$$

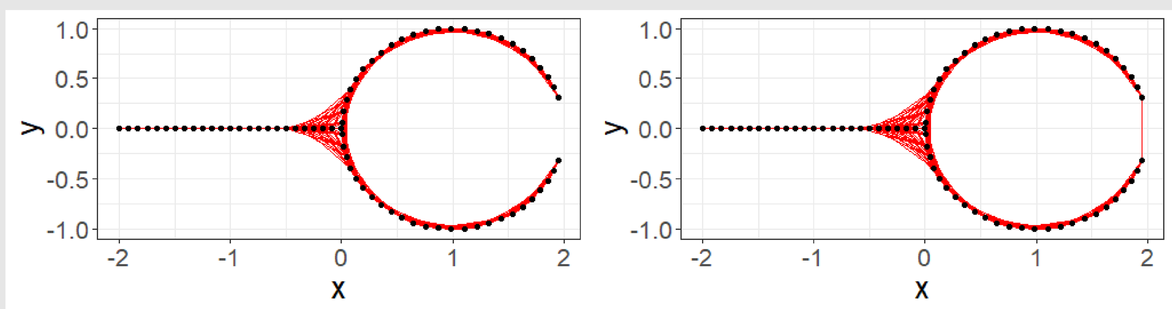
- More flexibility (account for data at hand)
- Is also necessary to extend to singularities (proof in paper)

Theorem 3. *Let M be a connected metric graph in $\mathbb{R}^d$ and let X be a finite set of data points in M. Suppose a graph $G = (X, E)$ is given, defining the following three thresholds:*

1) $\|x - y\| \geq \epsilon'$ for all $\{x, y\} \in E$,
2) $\|x - y\| < s_\epsilon(M)$ for all $\{x, y\} \in E$, with $\epsilon > 0$,
3) for all $x, y \in X$ with $\|x - y\| \leq \tau$, we have $\{x, y\} \in E$.

If for $0 < 4\delta < \tau$, X satisfies the δ-sampling condition, *i.e., for every $m \in M$ there is $x \in X$ with $d_M(m, x) \leq \delta$, then for all $x, y \in M$*

$$\lambda_{\epsilon,\epsilon'}(M) d_M(x, y) \leq d_G(x, y) \leq (1 + 4\delta/\tau) d_M(x, y). \quad (1)$$



REFERENCES

[1] R. Vandaele et al, "Mining Topological Structure in Graphs through Forest Representations," Journal of Machine Learning Research, 21(215):1–68, 2020.

[2] R. Vandaele. "Topological Data Analysis of Metric Graphs for Evaluating Cell Trajectory Data Representations," Master's thesis, Ghent University, 2020.

[3] R. Vandaele. "Topological Inference in Graphs and Images," Doctoral thesis, Ghent University, 2020.

[4] M. Bernstein et al, "Graph approximations to geodesics on embedded manifolds," 2000.

Contact

Robin.Vandaele@UGent.be
https://users.ugent.be/~rvdaele

Universiteit Gent
@ugent
Ghent University