3. Topological Dimensionality Reduction (TDR)

3.1. **Circular Coordinates for Non-Uniform Data.** The circular coordinates algorithm implemented in DREiMac may exhibit strange behavior when the input data does not follow a uniform distribution. The goal of this project is to remedy said situation. Figure 1 below shows an example of two data sets sampled around the circle; the data on the left column, X, was sampled uniformly at random along and near the unit circle, while the data on the right, Y, does not follow a uniform distribution – there are many more points near $(1, 0)$ than near $(-1, 0)$. The persistence diagrams for both data sets are shown in the second row; the most persistent 1-dimensional class was used to construct circular coordinates $f : X \longrightarrow S^1 = [0, 2\pi]/0 \sim 2\pi$ and $g : Y \longrightarrow [0, 2\pi]/0 \sim 2\pi$ in each case, and each point of $x \in X$ (resp $y \in Y$) is colored according to the value $f(x) \in [0, 2\pi]$ (resp $g(y)$). The desired behavior would be for the color in the points of Y to change slowly and uniformly as one goes around the circle – like the colors in X – instead of changing very rapidly near $(-1, 0)$ and then going to almost constant for the rest of the points.

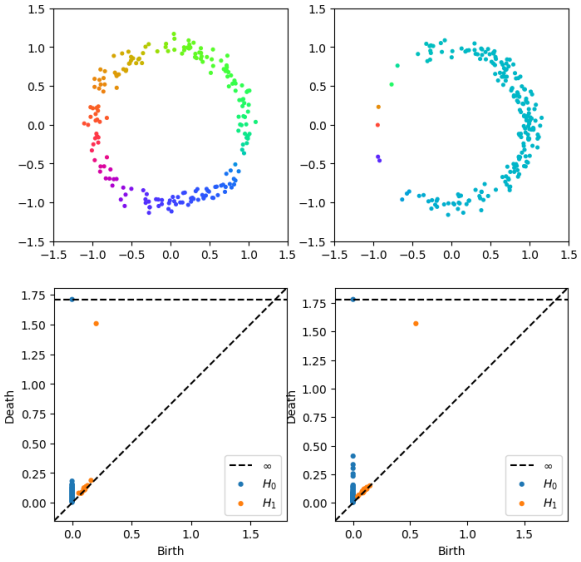

Figure 1. Example where non-uniform density makes a difference in DREiMac's circular coordinates

A place to start:

(1) Use the DREiMac library to code up a small example replicating the behavior seen in Figure 1.

References.

(1) Sparse Circular Coordinates via Principal $\mathbb{Z}$-Bundles, J.A. Perea, https://arxiv.org/pdf/1809.09269.pdf
(2) DREiMAc: Dimensionality Reduction with Eilenberg-MacLane Coordinates, J. A. Perea, L. Scoccola and C. Tralie, https://github.com/scikit-tda/DREiMac/tree/experimental