Topological Data Analysis

Jordan Britton

What is TDA?

- How one infers high-dimensional structure from low-dimensional representations.
- How one assembles discrete points into global structures.

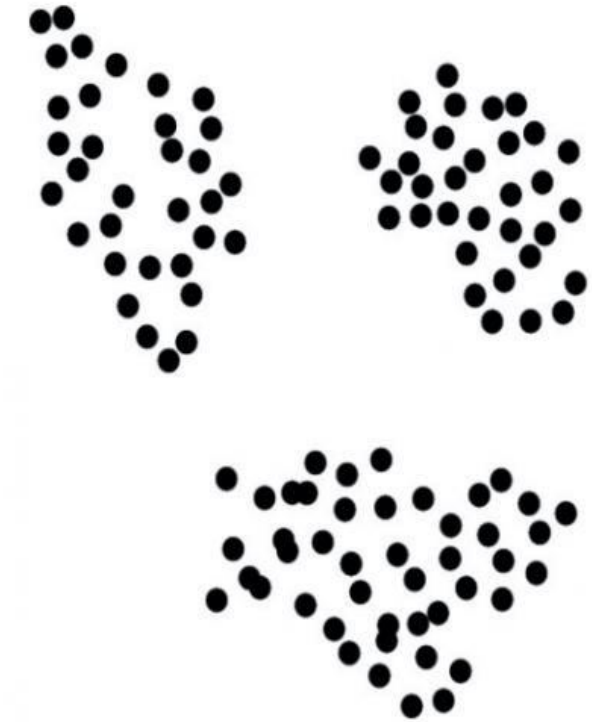


Figure 1: A data set with three clusters

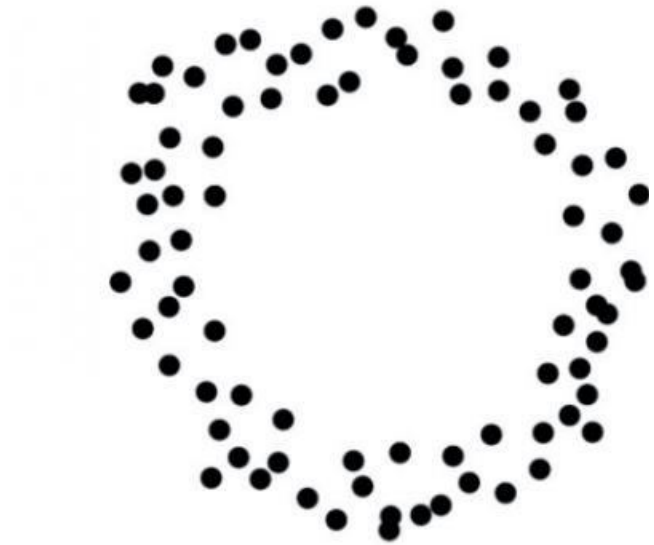


Figure 3: A data set with a loop

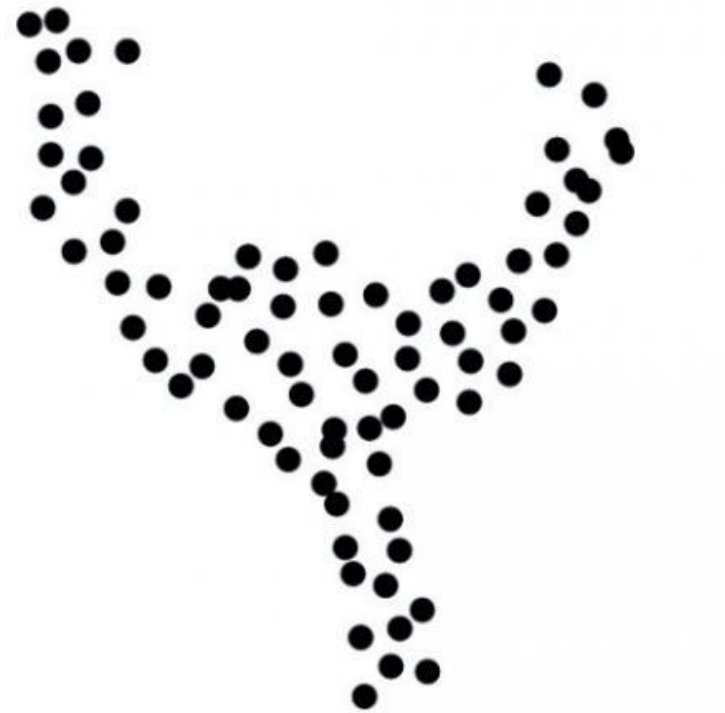
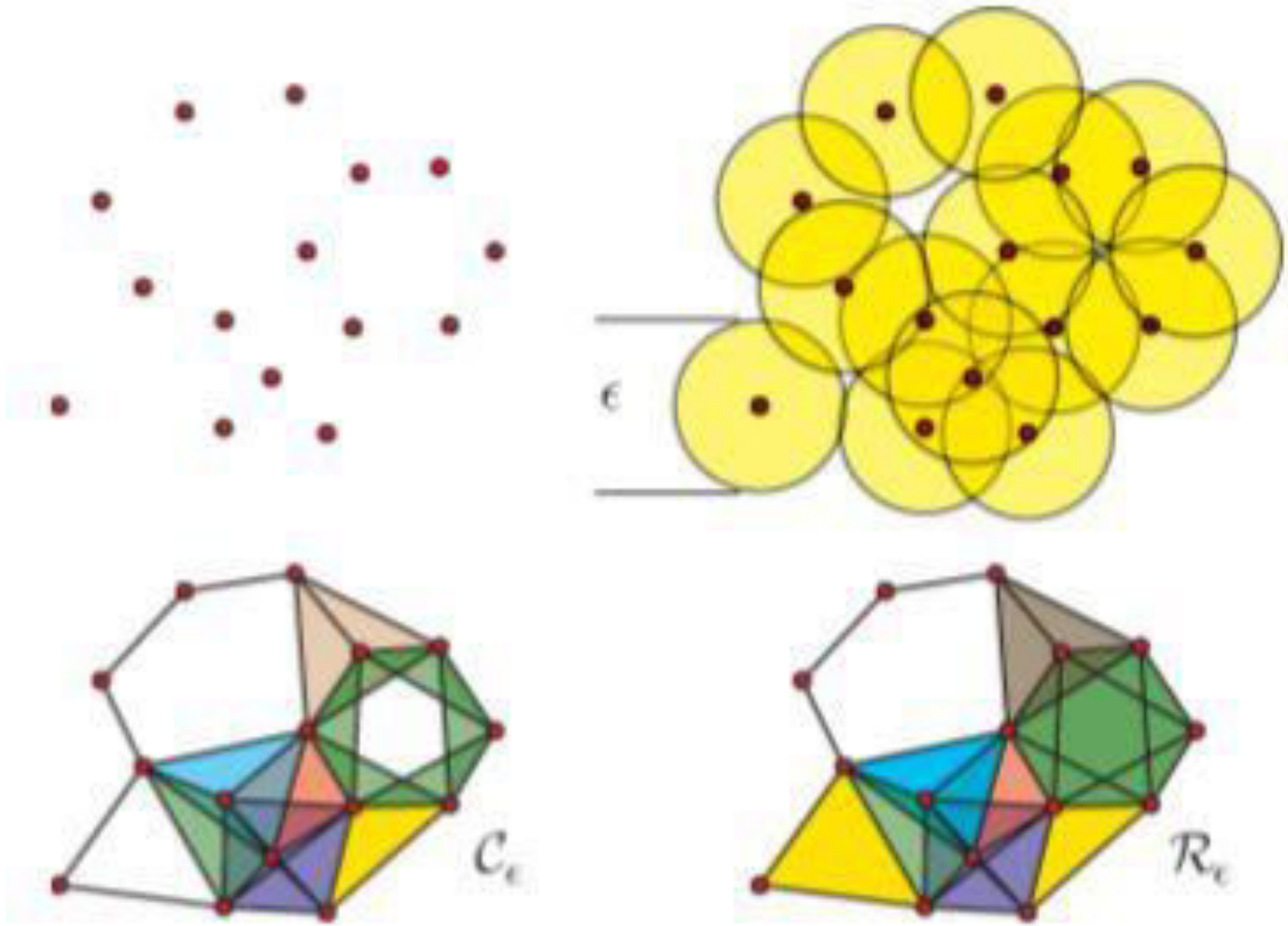


Figure 5: A data set with three tendrils emanating from a central core

Natural Methods

- The Cech complex: “Given a collection of points{xα} in Euclidean space, C, is the abstract simplicial complex whose k-simplices are determined by unordered (k + 1)-tuples of points $\{x\alpha\}^k_0$ whose closed e/2-ball neighborhoods have a point of common intersection.”
- Definition 1.2. Given a collection of points{xα}in Euclidean space, the Rips complex, R, is the abstract simplicial complex whose k-simplices correspond to unordered (k +1)-tuples of points $\{x\alpha\}^k_0$ that are pairwise within distance


ϵ
$\mathcal{C}_\epsilon$
$\mathcal{R}_\epsilon$

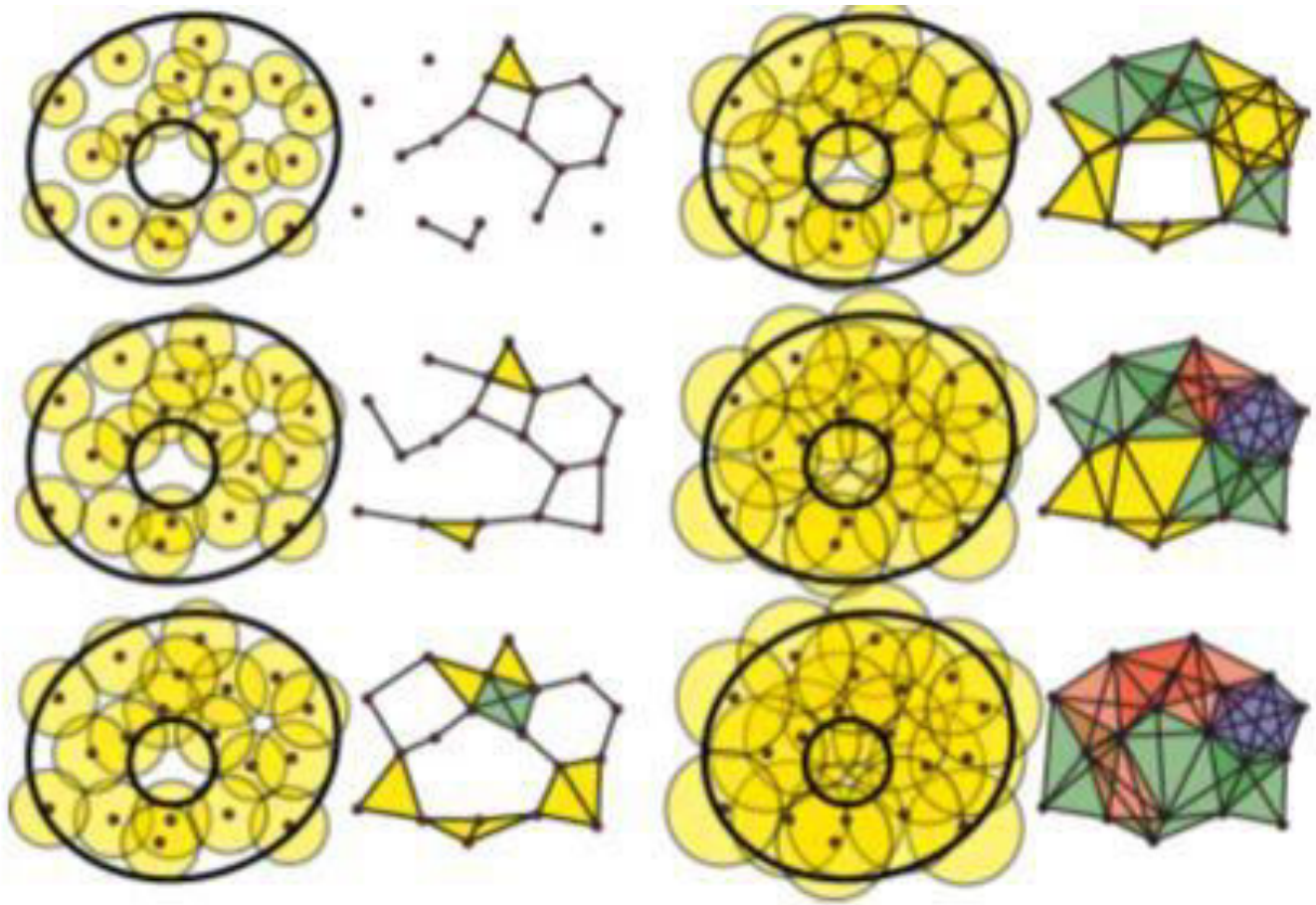


FIGURE 3. A sequence of Rips complexes for a point cloud data set representing an annulus. Upon increasing ϵ, holes appear and disappear. Which holes are real and which are noise?

Persistence

- Definition: “Given a parameterized family of spaces, those topological features which persist over a significant parameter range are to be considered as signal with short-lived features as noise”
- Notable function: $T(X, \delta)$