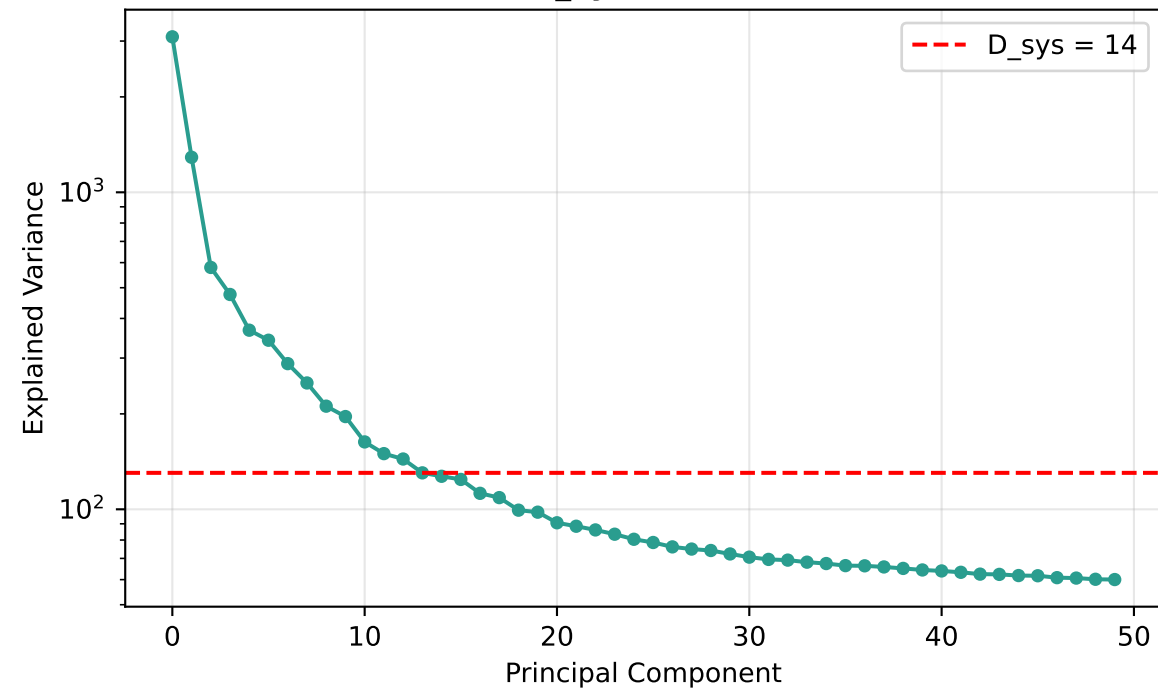
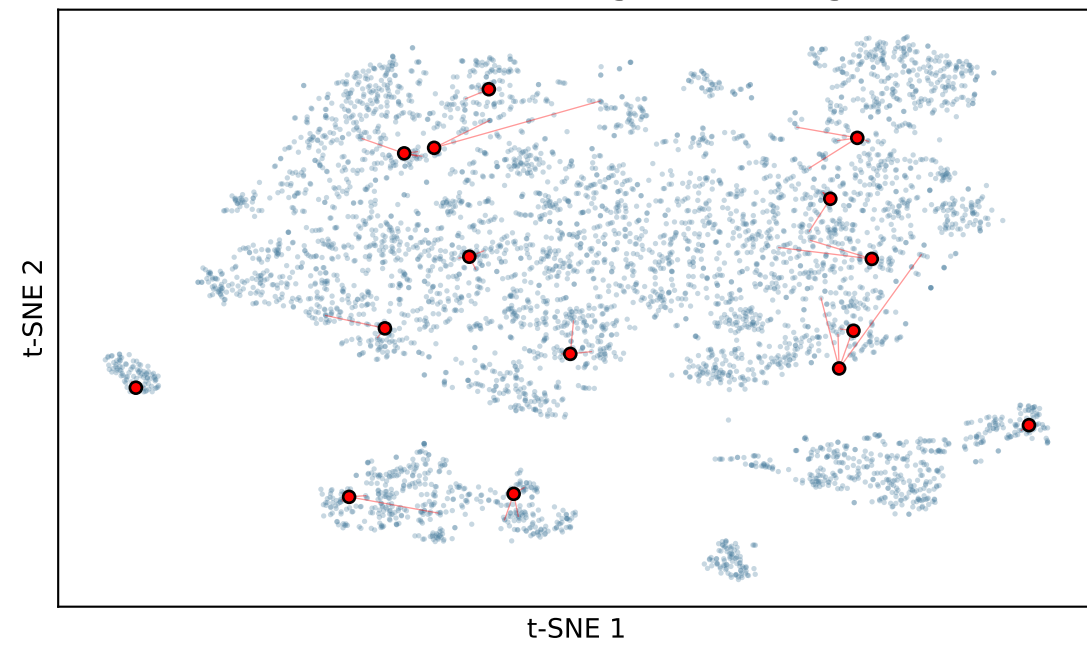


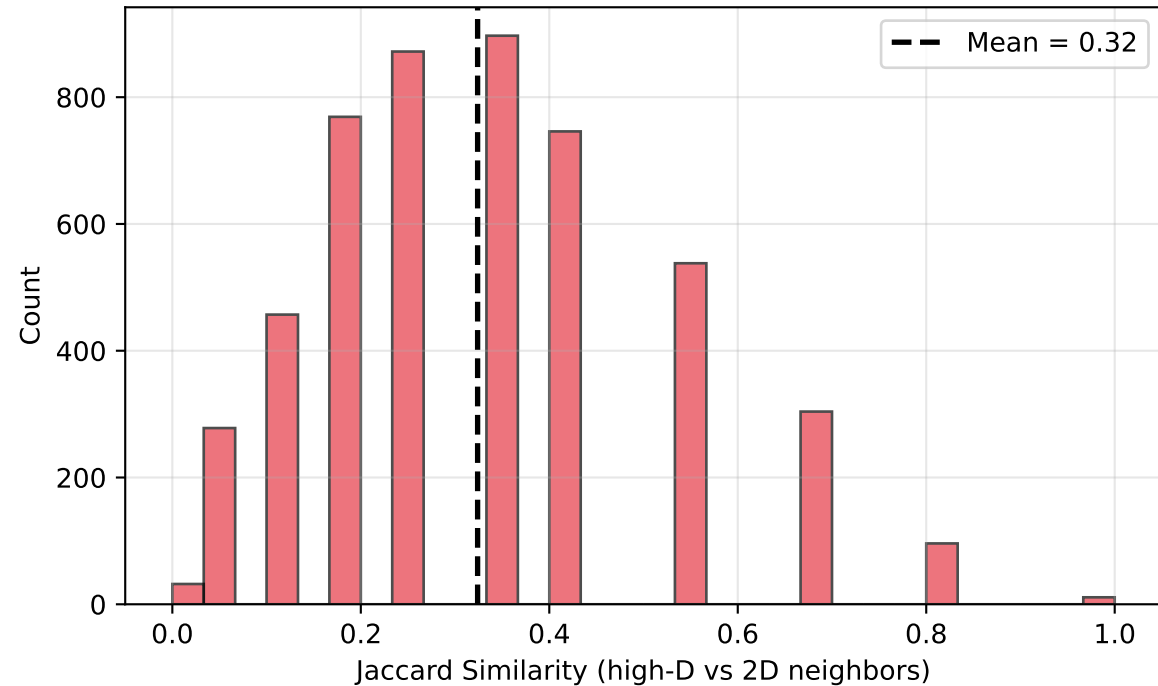
A. Eigenvalue Spectrum
($D_{\text{sys}} = 13.6$)



B. The Shadow Lie
(red lines = TRUE neighbors in high-D)



C. Neighbor Preservation
(Aliasing = 67.6%)



D. The Epistemological Point

TOPOLOGICAL ALIASING IN scRNA-seq

Dataset: GSE120575 (Sade-Feldman)
Cells: 16,291
Genes: 55,738 \rightarrow 39,254 (filtered)

SYSTEM (D_{sys}):
Participation Ratio = 13.6
 \rightarrow The gene expression manifold has
~14 effective dimensions

SHADOW (D_{obs}):
t-SNE dimension = 2
 \rightarrow We project to a 2D visualization

ALIASING:
Rate = 67.6%
 \rightarrow 68% of high-D neighbors are
"lost" in the 2D projection

COVERAGE:
 2.17×10^{-6} of 3^{20} cells occupied
 \rightarrow The data covers a vanishing fraction
of the high-D space

IMPLICATION:

When $D_{\text{sys}} \gg D_{\text{obs}}$, the 2D "shadow"
introduces systematic distortions.
Clusters that appear close may be
far apart in reality, and vice versa.

This is topological aliasing applied
to real biological data.