

## **Improving an Intrinsic Estimator of Scalar Curvature for Application to High-Dimensional Biological Data Sets**

*Sruthi Parthasarathi (Mentor: Andrew Blumberg, 2022 IICD SRP)*

Scalar curvature provides a measure of how volumes of geodesic balls on a Riemannian manifold differ from their analogs in Euclidean space. For instance, in two-dimensions, a unit disk on the surface of a 2-sphere (which has constant positive curvature) is smaller in area than the unit disk in the zero-curvature plane. Computing curvature for data sets is useful because given a point cloud sampled from a manifold, it can inform a choice of embedding in which the geodesic distances between points are minimally distorted – spherical and hyperbolic embeddings, for example, are better suited for regions of positive and negative curvature, respectively – allowing us to preserve the underlying relationships arising from the manifold in the embedded form of the data as well.

However, existing methods for computing scalar curvature rely on extrinsic information, which requires the data to already be embedded in an ambient space. In contrast, we look at a method of computing scalar curvature using ball volume growth estimates that rely solely on intrinsic information and is applicable regardless of the dimensionality of the data. We primarily aim to determine the optimal values for the parameters in this estimator by testing accuracy on toy models with known scalar curvatures, increase the accuracy of the estimator on boundaries, and apply the tuned model to real-life data sets. The last of these objectives specifically includes analyzing single cell RNA sequencing (scRNA-seq) data – which describes levels of gene expression in individual cells using their transcriptomes – to determine whether or not areas of significant nonzero curvature are present, as well as graphs to determine the extent to which our curvature computation is meaningful in non-manifold contexts.