Instability of Manifold Reach and a Stable Combinatorial Reach

Michael Rawson
Department of Mathematics
University of Maryland at College Park
rawson@umd.edu

Michael Robinson
Department of Mathematics
American University
michaelr@american.edu

***Abstract*—In application and theory, reconstruction of an embedded manifold from samples hinges on the manifold's $curvature$ and $reach$. The $reach$ of a manifold is the smallest distance to the medial axis. Reach is a function of curvature but contains more information for reconstruction. We show instability, theoretically and practically, in Federer's manifold reach formula [1]. To compensate for this defect, we introduce a $combinatorial\ reach$ for stratifications. We show the reach and the combinatorial reach are equivalent in the limit of sampling on a manifold. With high probability, the sampling rate necessary to detect reach via combinatorial reach is a function of reach class [2] and multiplicity. With noisy samples, we show that there is a trade-off in estimation error of the $combinatorial\ reach$ based on the noise covariance matrix and the curvature of the manifold. Noisy samples corrupt estimates of the curvature so we are forced to estimate both the reach and noise level by cross validation. We compare these noise level estimates with those obtained by piece-wise linear methods. We find the combinatorial reach noise estimates far more suitable for high dimensional, high curvature, or sparse data problems. We showcase these results on samples of conic sections, since conic sections are second order approximations of smooth manifolds locally.**

I. Introduction

In applications, reconstruction of an embedded manifold from samples hinges on the manifold's $curvature$ and $reach$. Given a fixed set of samples, the higher the manifold curvature can be, the more under-determined the manifold is. For example, a connected manifold with no curvature is determined with just dimension+1 samples while a quadratic or cubic surface is under-determined with dimension+1 samples. The $medial\ axis$ of a manifold are the points outside of it (in the embedding space) that do not have a unique nearest neighbor inside of the manifold. The $reach$ of a manifold is the smallest distance to the medial axis. Reach is a function of curvature, but gives more information about reconstruction.

II. Main Results

We show instability, theoretically and practically, in the manifold reach formula $\tau_M = \inf_{x\neq y; x,y\in M} \frac{\|y-x\|_2^2}{2d(y-x,T_xM)}$ for manifold M, distance d, and norm $\|\cdot\|$ (Theorem 4.18 in [1]). Given a set of samples $S \subset M$, let estimate $\hat{\tau}_M = \inf_{x\neq y; x,y\in S} \frac{\|y-x\|_2^2}{2d(y-x,T_xM)}$

Theorem 2.1: For N samples of $M \subset \mathbb{R}^m$ and perturbation p of T_xM and $\eta(r) = \inf_\gamma \|proj_{T_xM}(\gamma(r)-x)\|$ where γ is a geodesic and some $0 \le \xi \le \|y-x\|$,

$$\hat{\tau}_M \le \inf_{x\neq y} \frac{\|y-x\|}{2\|p\|(\sup \|\gamma''|_x\| + \eta''(\xi)\|y-x\|)}$$

The proof works by Taylor expansion of the embedding map.

Remark 2.2: Thus the instability shows up the flatter the manifold, pushing $\hat{\tau}_M$ to 0. We can see that a perturbation of the tangent space is equivalent to perturbing a sample since the samples estimate the tangent space.

To compensate for this defect, we introduce a combinatorial reach for stratifications which is a robust multiscale generalization of reach.

A. Combinatorial Reach

Let simplicial complex $S = \{[v_{i_1}, ..., v_{i_n}]\}_{\{i\}_1^n \in I} \subset \mathbb{S}$, where v_i are labels of vertices embedded in metric space $(X \subset \mathbb{R}^m, d, \|\cdot\|)$. We always will assume the faces of S are linear in $\mathbb{R}^m$. Let $\{V_k\}_{k\in I}$ be the Voronoi complex of the vertices of S. The cells

$$V_k = \{x \in X : d(x, v_k) < d(x, v_j)\ \forall j : v_k \neq v_j\}.$$

Now, we'll set V_δ as the complex formed by the boundaries of the cells V_k that are δ separated from S.

$$V_\delta = \{\sigma \in complex\,(\cup_k\ cl(V_k) - V_k) : d(\sigma, S) > \delta\}$$

where $d(\sigma, S) = \inf_{x\in\sigma, y\in S} d(x, y)$.

Definition 2.3: Define the $combinatorial\ reach$ of S as

$$\tau_S(\delta) := \min\left(d(S, V_\delta), \sup_{u,w\in X} d(u, w)\right)$$

which may equal ∞.

The reach and the combinatorial reach are equivalent in the limit of sampling on a manifold. With high probability, the sampling rate necessary to detect reach via combinatorial reach is a function of reach class [2] and multiplicity. For example, see figure 1.

B. Noise

With noisy samples, there is a trade-off in estimation error of the combinatorial reach based on the noise covariance matrix and the curvature of the manifold. Noisy samples corrupt estimates of the curvature so we are forced to estimate both the reach and noise level by cross validation. We compare these noise level estimates with those obtained by piece-wise linear methods. We find the combinatorial reach noise estimates far more suitable for high dimensional, high curvature, or sparse data problems.

We showcase these results on noisy vs. noiseless samples of conic sections, since conic sections are second order approximations of smooth manifolds locally. Conic sections are governed by one free parameter. As that one parameter is varied, the reach varies and jumps while the combinatorial reach is more stable. This allows the combinatorial reach of the manifold to be approximated with manifold sampling. Furthermore, the combinatorial reach contains information about the manifold under different resolutions. For example, a fractal may have almost any reach under some blurring or smoothing filter. Combinatorial reach will detect all of these values for the reach; geometry permitting and sampling permitting.

III. Conclusion

We have shown instability in the reach formula and proposed a multiscale generalization of reach. Combinatorial reach is more useful in characterizing the underlying manifold and approximating reconstruction accuracy.

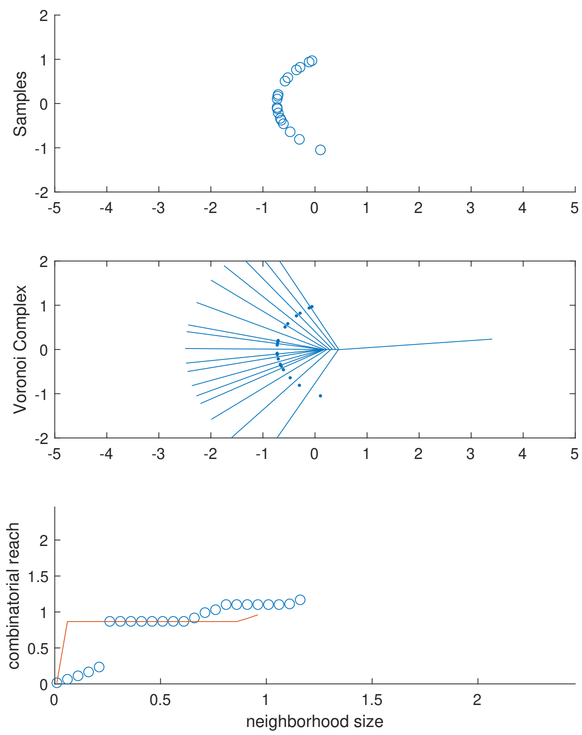

Fig. 1. Top plot: Random sample of 1D manifold. Mid plot: Voronoi complex of samples. Bottom plot: Blue circles are combinatorial reach of samples. Red line is the combinatorial reach of the manifold.

REFERENCES

[1] H. Federer, "Curvature measures," Trans. Amer. Math. Soc., vol. 93, pp. 418–491, 1959, doi: https://doi.org/10.1090/S0002-9947-1959-0110078-1.

[2] E. Aamari, J. Kim, F. Chazal, B. Michel, A. Rinaldo, and L. Wasserman, "Estimating the reach of a manifold," p. 41, 2019.