

Point-by-Point Response to Referee One

Thank you very much for your careful reading our paper as well as your invaluable comments. We list our response to your comments as follows:

“This paper is on manifold learning. It is very confusing to read.”

Answer:

To the best of our knowledge, the reserved terminology manifold learning refers to an approach to non-linear dimensionality reduction (Isomap and a number of others), which is a method trying to find the new representation in a lower-dimensional space (R^d) by approximately keeping the geometry properties of the input which resides in a high-dimensional ambient space (R^D). This line of research is mainly about learning the embedding.

In this paper, we aim to learn the actual manifold in the ambient space. Therefore, our work is not on manifold learning. We call this manifold fitting. There are not many research in this direction. We provided the origin of this line of work in the beginning of paper and also cited the previous of pioneering work from Charles Fefferman et al. 2016, 2018, 2019 with a few others.

“For example, realistically the data generated by the manifold model are generated with noise, and the distribution and properties of the noise in conjunction with the reach of the manifold are critical for the recovery. However, these critical issues are bypassed by making strong assumptions whereby only the local behavior of ridges matters.”

Answer:

We do **not** agree. In fact, in our work, the model in Equation (1) on page 2 assumes the data are generated as

$$x_i = \tilde{x}_i + \epsilon_i, i = 1, \dots, n, \quad (1)$$

where $\{\tilde{x}_i, i = 1, \dots, n\}$ are the noiseless signal which belong to some manifold, and $\{\epsilon_i, i = 1, \dots, n\}$ are the i.i.d noise from some distribution, such as the multivariate Gaussian distribution. We use the observations x_i to recover the manifold. Therefore, we did not bypass consideration of the noise in our model.

The reach is a global concept which describes the degree of curvature for the manifold. A larger “reach” indicates the samples resides further from the manifold can also has a unique projection onto \mathcal{M} . Therefore, the manifold is more flat globally.

We do not prefer to model the properties of the noise in conjunction with the reach in the first place. The reason are two folded:

- Reach is one way to describe the curvature; if we model the noise in conjunction with the reach, we simply reduce the level of the problem,

- We only assume the noise (Gaussian) is added to the pure data in the ambient space; this is a more general setup in manifold fitting just like the one in Charles Fefferman et al. 2016, 2018, 2019 (see our response to your 4th comment below).

“This approach may be interesting in its own rights, but the impact in the bigger picture is rather limited due to the limited perspective and the strong assumptions that are needed. The best part are the computational perspectives provided in the paper, while the theoretical developments are also not very interesting.”

Answer:

In fact, Assumption 2.2 is more likely to be regarded as a “Claim” rather than an Assumption. The main intention of Assumption 2.2 is to say that any low-dimensional manifold can be thought as the ridge of some unknown density function. In this new version of the manuscript, we modify the presentation of Assumption 2.2 to avoid the confusion.

“A key problem is that the errors in the sampling of the data generated by the manifold are not modeled and the noise-contaminated data are not studied. The paper is not well written, the presentation is substandard and is unfocused.”

Answer:

We do **not** agree. We work with noisy data from the beginning. As in Equation (1), we model the data as the noiseless $\tilde{x}_i \in \mathcal{M}$ plus the noise ϵ_i . We consider Gaussian noise added in the ambient space directly and assume the noisy data lie in a lower dimensional manifold. This is a very standard setup for manifold fitting. We have considered and relies on this setup for the noisy data in all the theories and all the cases of the numerical experiments in this paper.

“Problematic assumptions include Assumption 2.2, where parts 1 and 3 are unclear and ambiguous and Assumption 3.4 which is equally unmotivated.”

Answer:

Assumption 2.2 only wants to show that any noiseless manifold can be thought as the ridge of some background density function $p(x)$. This is to say that, on the population level, what we can understand the ridge and the density function. We work with the empirical one for the rest of our paper, based on samples.

Because of $\mu_s(x, r, h) \in (0, 1), \mu_{s,t}(x, r, h) \in (0, 1)$, we know

$$\mu(r, h) = \sup_x \max\{\max_s \mu_s(x, r, h), \max_{s,t} \mu_{s,t}(x, r, h)\} \in [0, 1].$$

Assumption 3.4 is a bit stronger by requiring $\mu(r, h)$ to be upper-bounded by a number ℓ which is less than 1 for all h . This assumption can be neglected and discussed depending on the different occasions of the upper-bound of $\mu(r, h)$.

“The simulations are too limited as there are many other competing methods to recover the manifolds that are considered, and these manifolds are too simple to be of much interest.”

Answer:

We almost compared all methods that we know of in this filed on the one-dimensional manifold to show the good recovery property of our method under the criteria of average margin and Hausdorff distance. The manifold fitting results (theory and numerical results) do not rely on the data set too much. Note that we coded the abstract fitting algorithm from Charles Fefferman et al. 2016, 2018, 2019 as well.

In our revised paper, we have added another experiment corresponding to a 2-dimensional sphere in the 3D ambient space. We can add more experiments by increase the D and d . We choose not to do so since we have decided to add a real data analysis where we roughly know the truth, to support our theory instead.

Our method can be applied to general input of data with a manifold hypothesis. We used the above two cases since we do know the manifold hypothesis is true in the data and we completely know the pure manifold so that we can evaluate the distance for comparison.