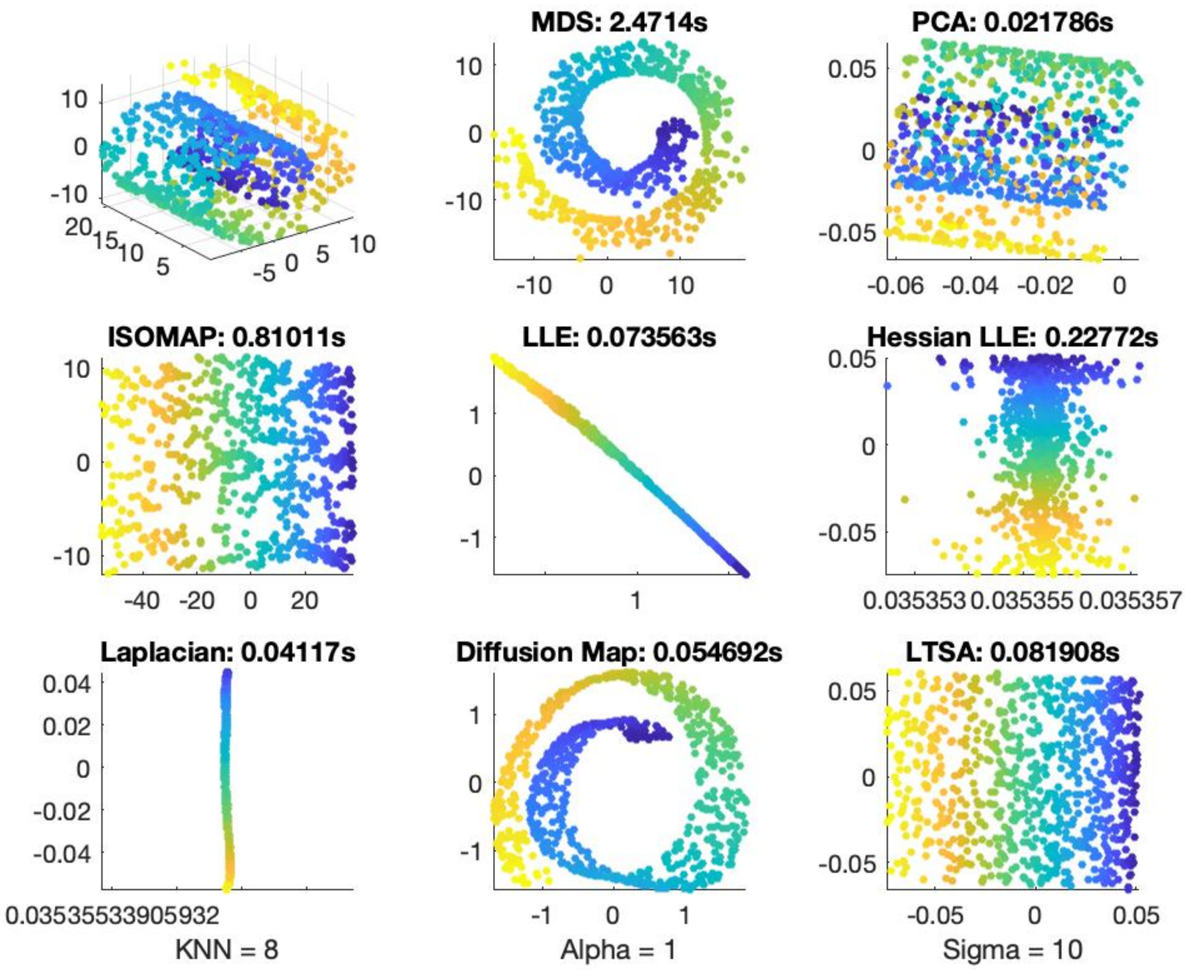
Comparison of manifold learning methods

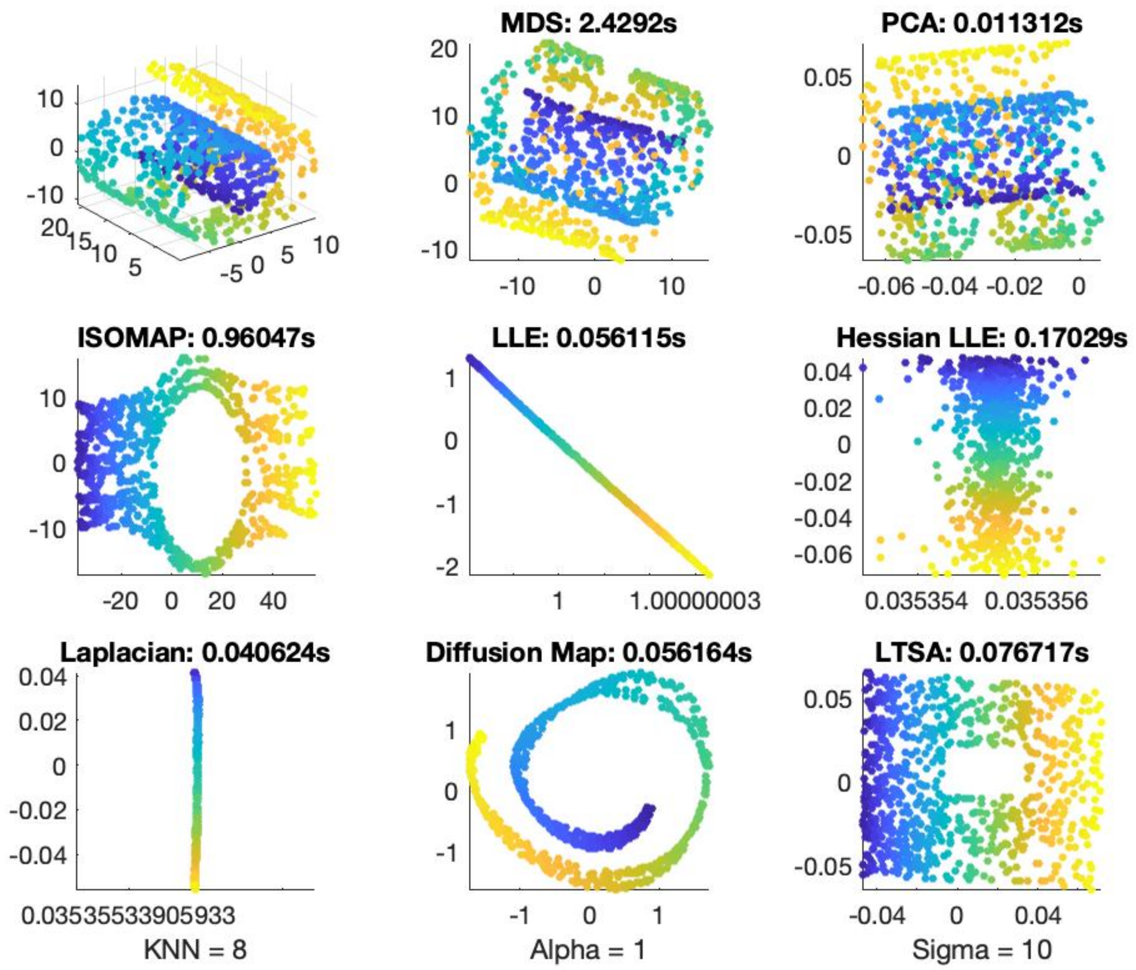
(a) Swiss Roll



We can see that MDS and its manifold version ISOMAP are very slow, this is because they need the global information instead of local ones. PCA cannot unroll the swiss roll at all and MDS also cannot, this is because they are linear methods and do not use manifold information. Local approaches, such as LLE, Laplacian and LTSA can unroll this dataset and preserve the local geometry, and among them, LTSA is the fastest one, Hessian LLE is also slow. Diffusion map cannot unroll it for any value of sigma.

(b) Swiss Hole

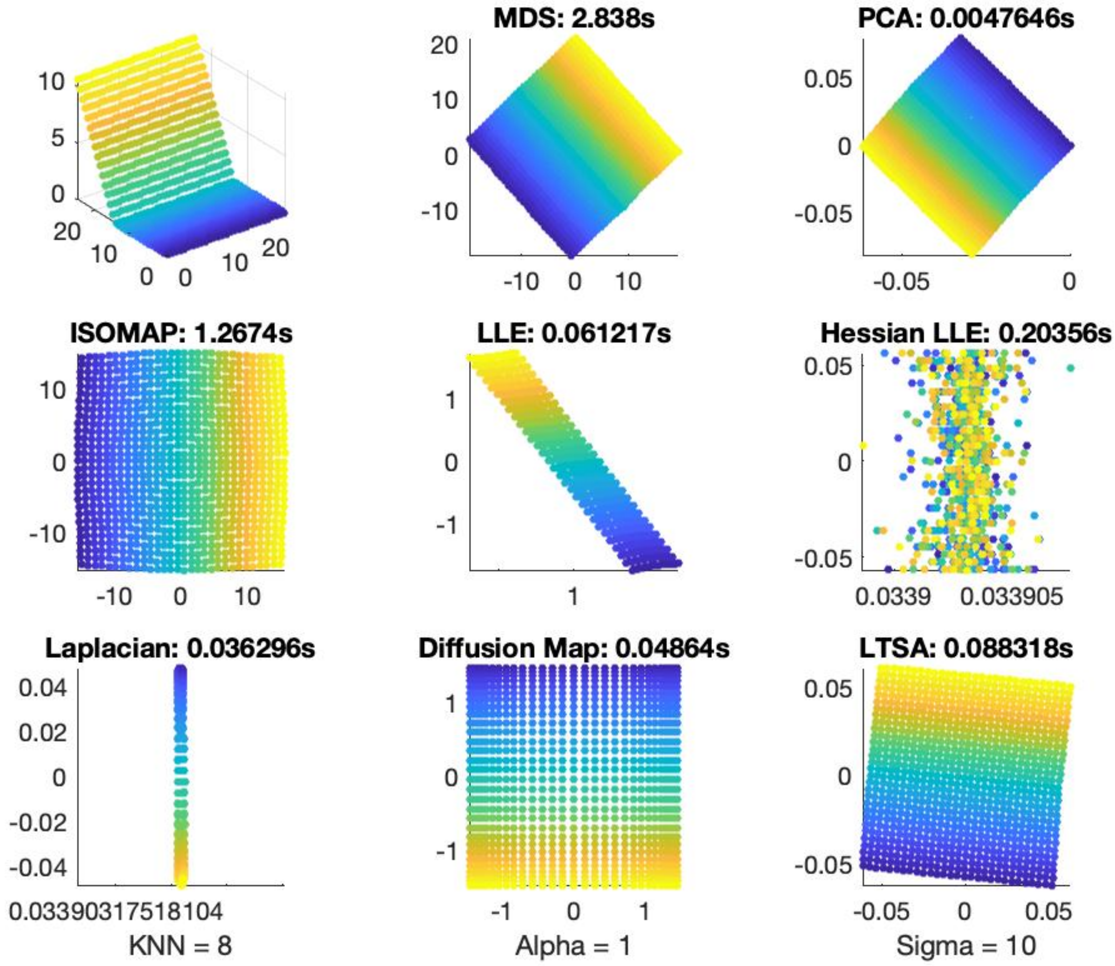
It is swiss role with a hole, hence non-convex.



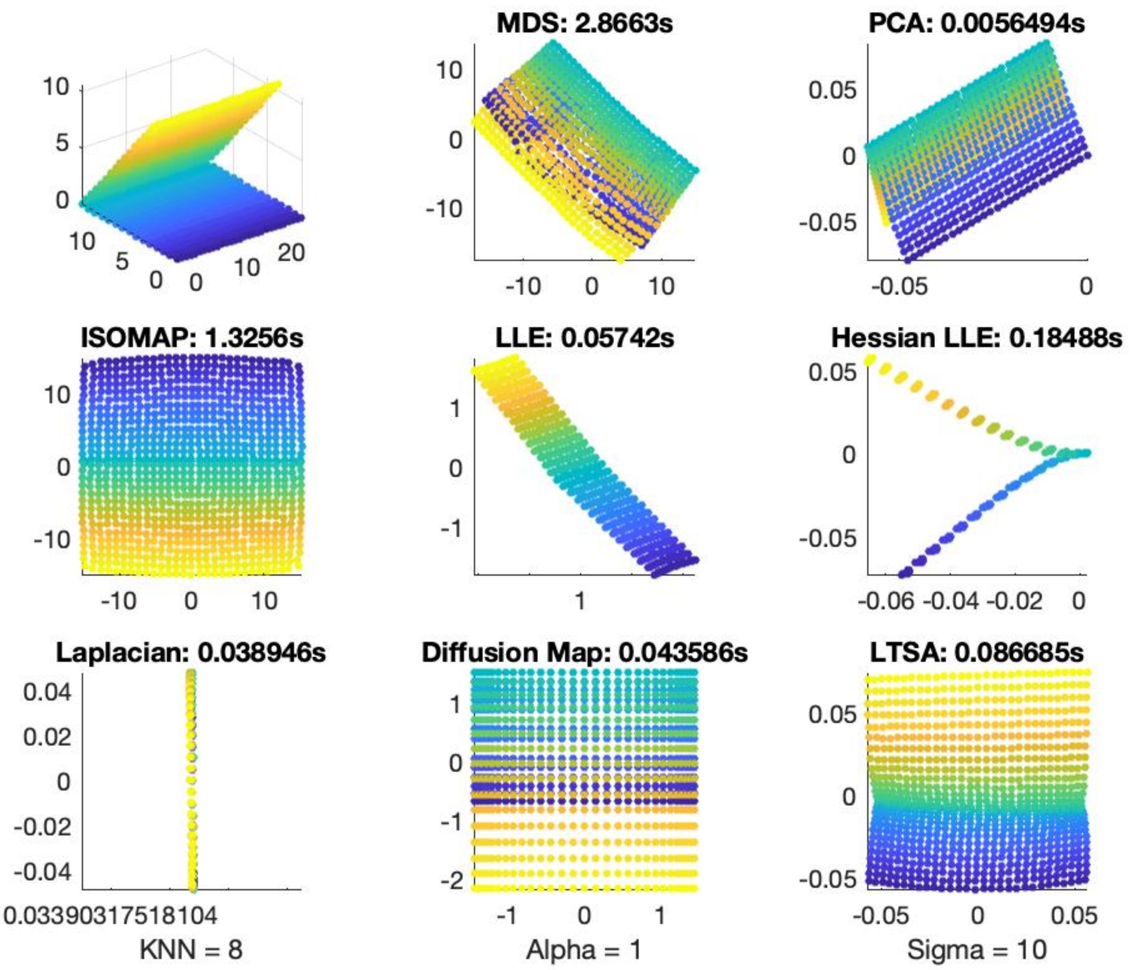
We can see that similar as last example, MDS, PCA and Diffusion Map fail totally. For the non-convexity, only LTSA can handle it, while LLE, Hessian LLE and Laplacian fail. ISOMAP also finds the hole but the result is distorted.

(c) Corner Planes

We bend a plane with a lift angle A. We want to bend it back down to a plane.



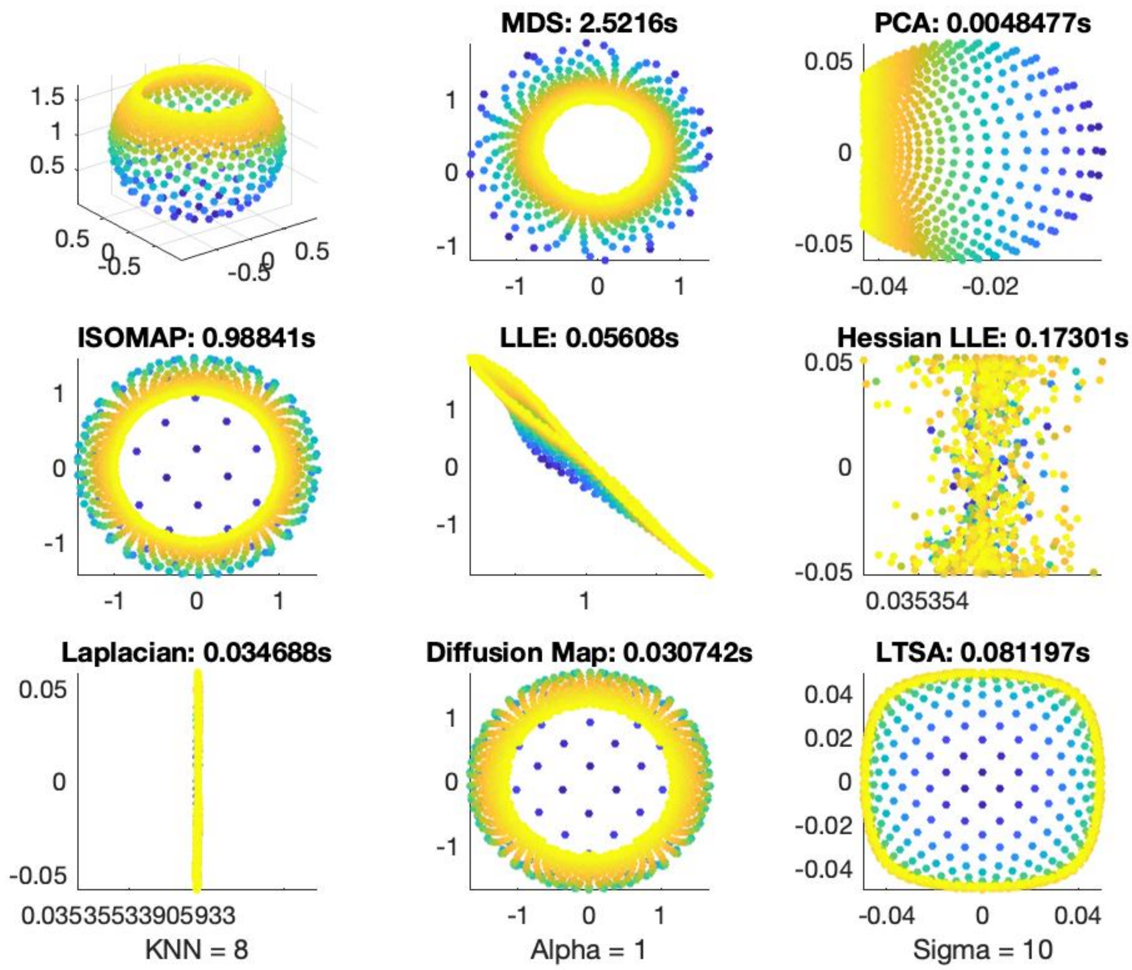
When the lift angle is 45, Hessian LLE and Laplacian fail totally while LLE also do not perform well. However, MDS, PCA, ISOMAP, Diffusion Map and LTSA all do a good job.



When the lift angle is 135, MDS and PCA write the data on the top of itself. LTSA is a little distorted. ISOMAP works perfectly.

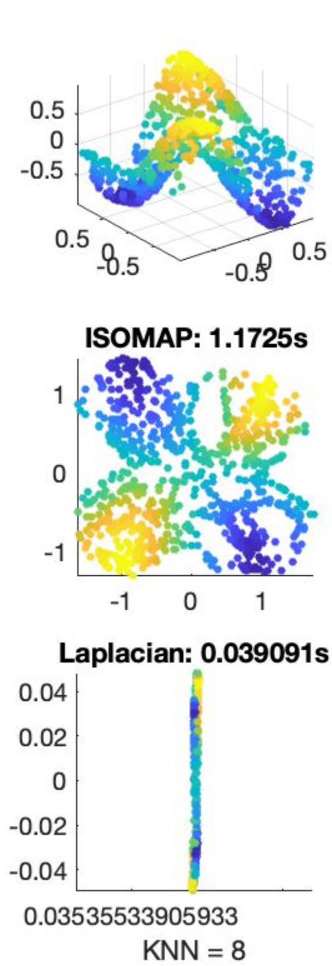
(d) Punctured Sphere

The sampling is very sparse at the bottom and dense at the top.



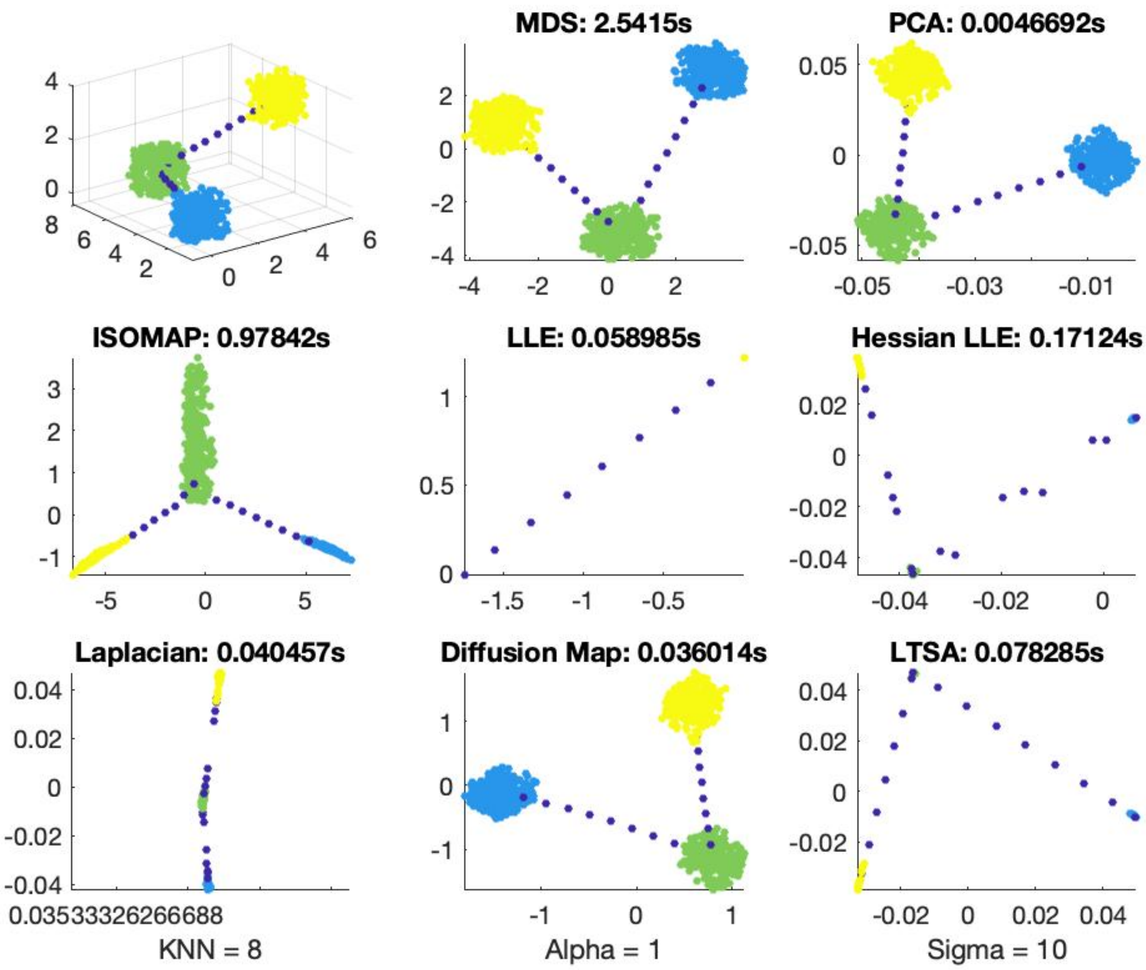
We can see that here ISOMAP, Diffusion Map and LTSA give the right shape, but they all give too much space to the sparse part at the bottom of the sphere. MDS turns the inside out. PCA projects the sphere from the side. LLE, Hessian LLE, and Laplacian fail.

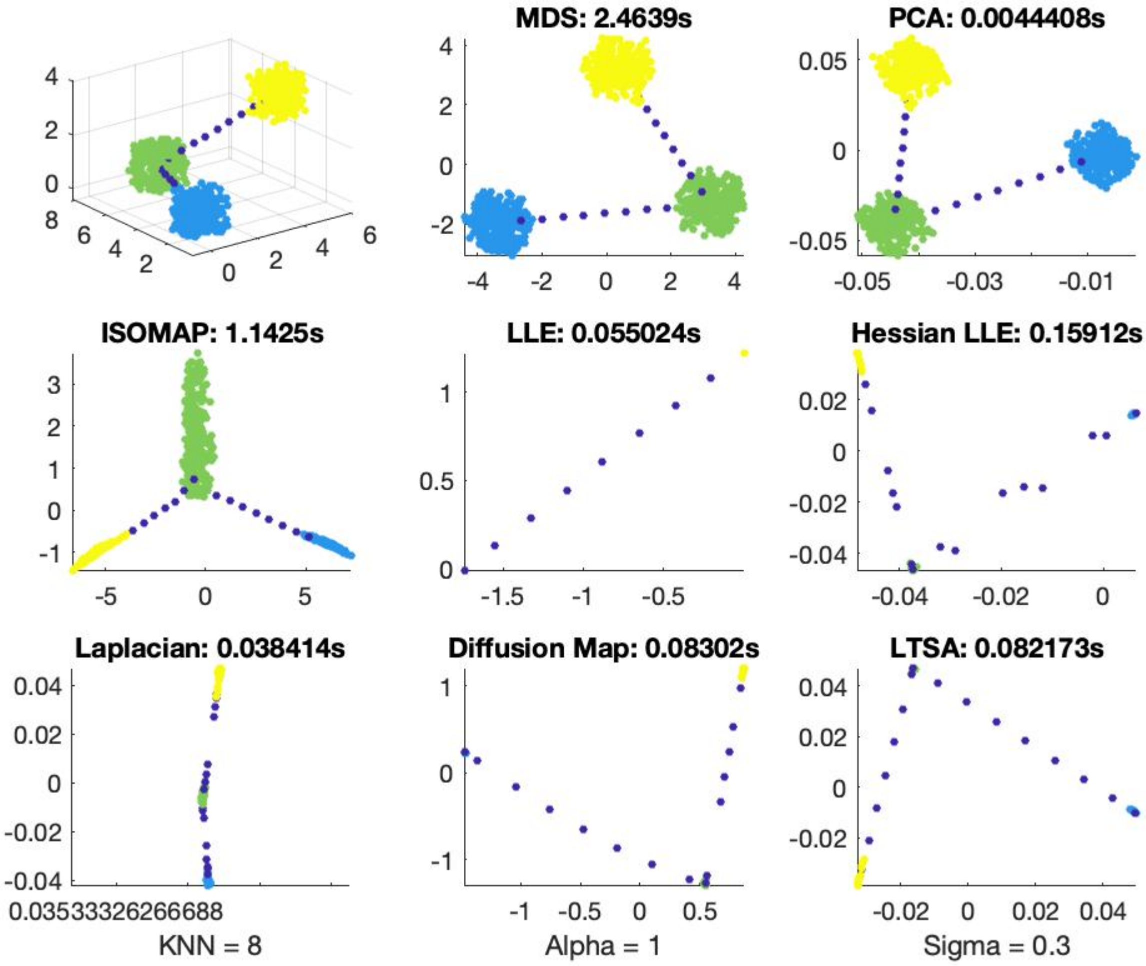
(e) Twin Peaks



We can see that PCA,LLE and Laplacian fail. LLE has trouble because it introduces curvature to plane. MDS and LTSA distorts it, and Diffusion Map gives too much emphasis on the inner spasity. ISOMAP has the best performance.

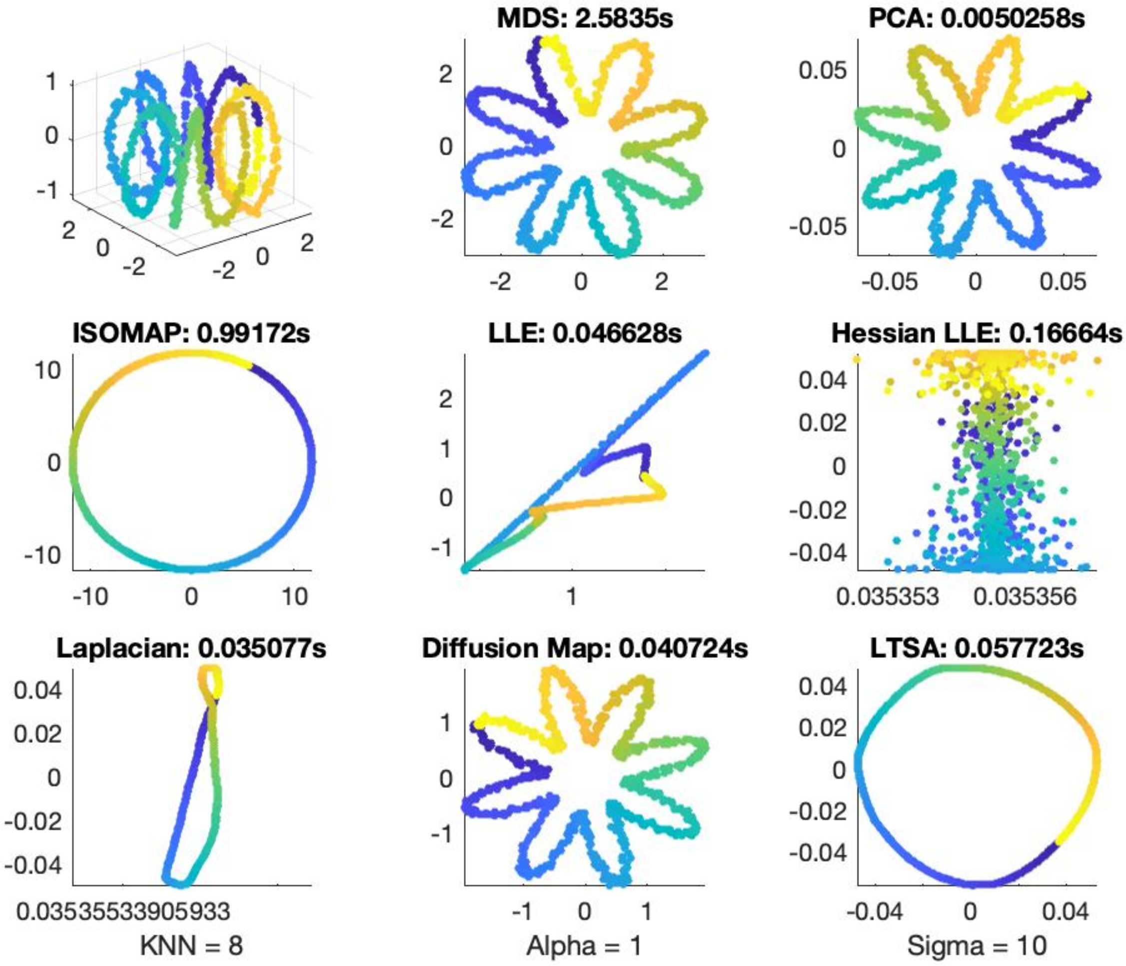
(f) 3D Clusters

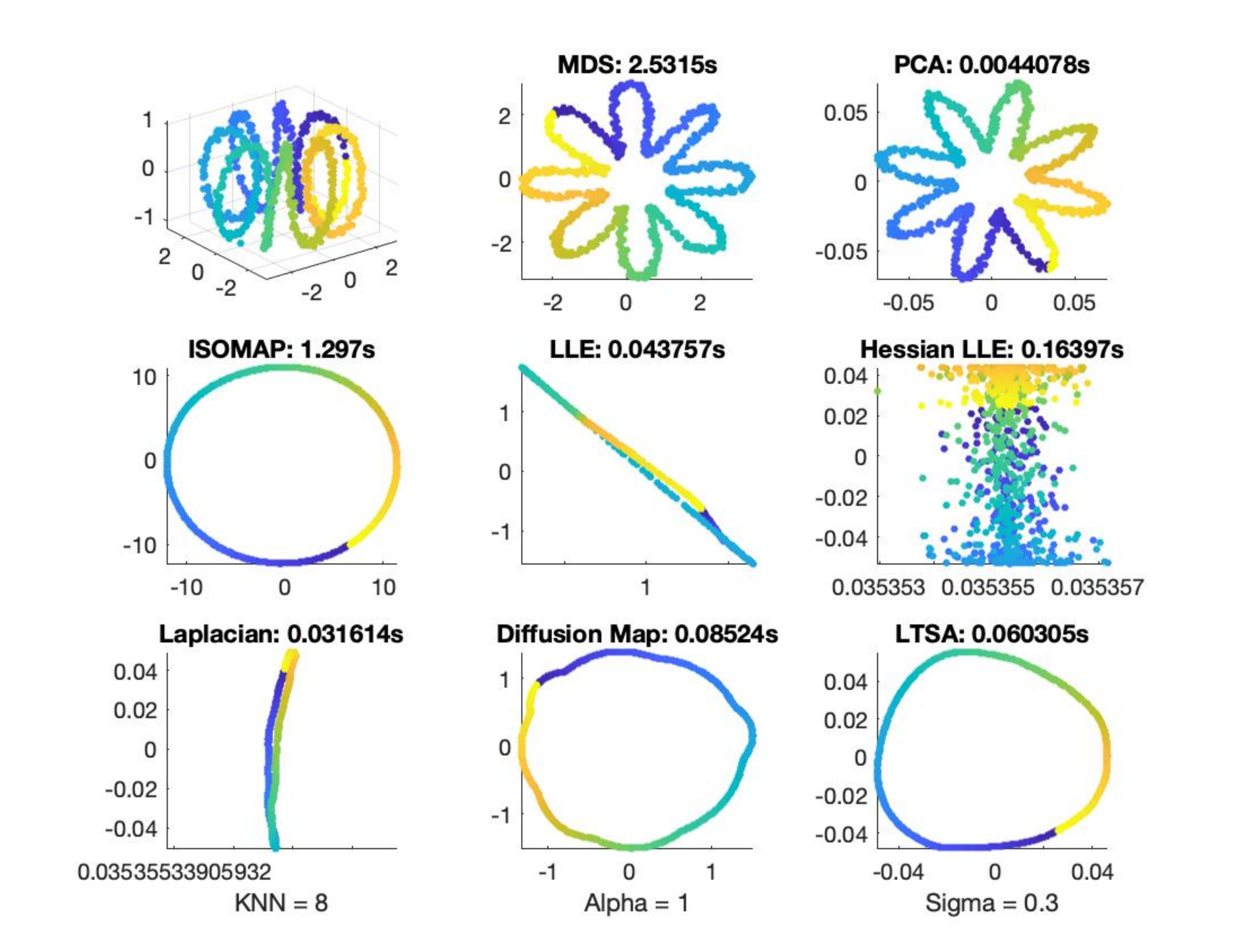

MDS: 2.5415s
PCA: 0.0046692s
ISOMAP: 0.97842s
LLE: 0.058985s
Hessian LLE: 0.17124s
Laplacian: 0.040457s
Diffusion Map: 0.036014s
LTSA: 0.078285s
KNN = 8
Alpha = 1
Sigma = 10



We can see that MDS and PCA work well for this dataset. ISOMAP also works although it distorts the data. LLE fails totally. Hessian LLE, Laplacian and LTSA project each cluster to a point. Diffusion Map is sensitive to sigma, when sigma is small, it cannot work well; but when sigma is large, it works well.

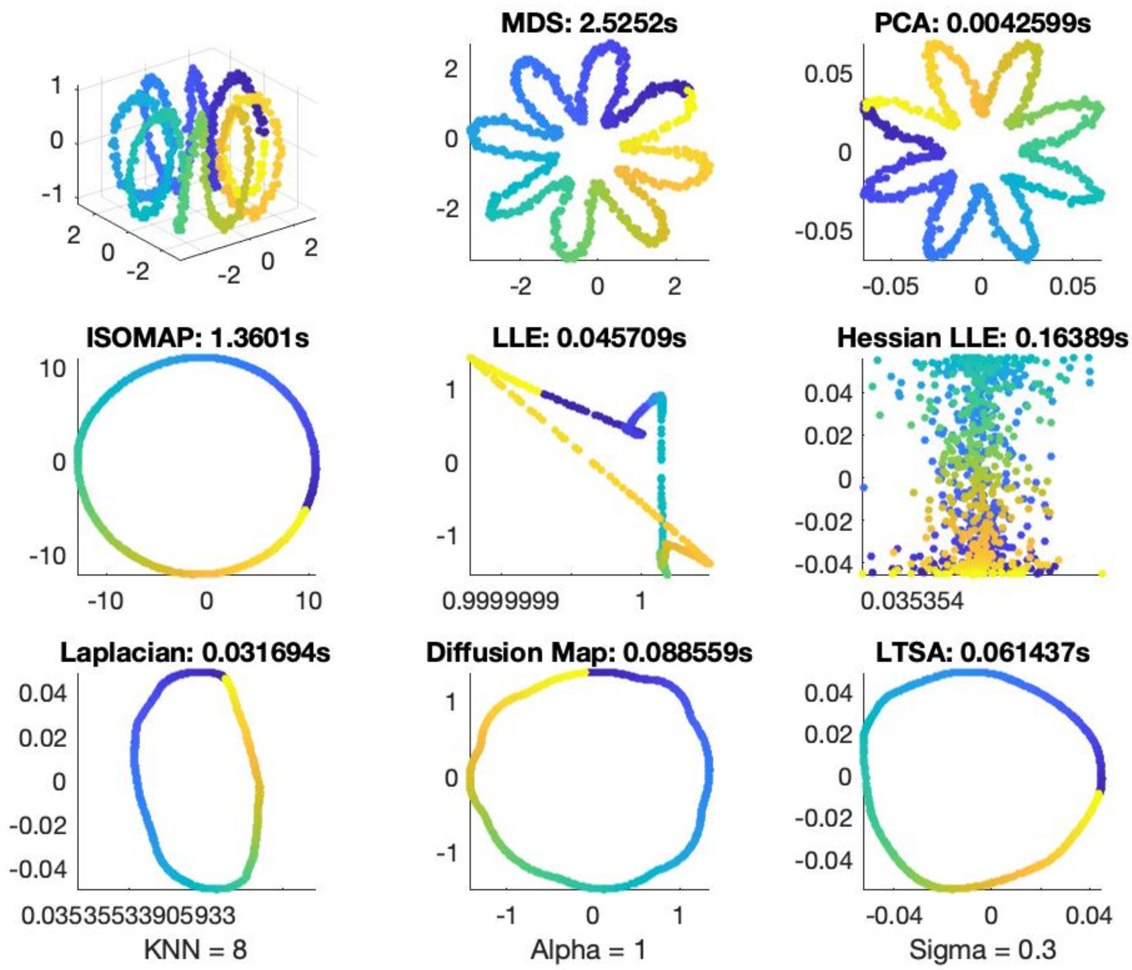
(g) Toroidal Helix


MDS: 2.5835s
PCA: 0.0050258s
ISOMAP: 0.99172s
LLE: 0.046628s
Hessian LLE: 0.16664s
Laplacian: 0.035077s
Diffusion Map: 0.040724s
LTSA: 0.057723s
0.03535533905933
0.035353
0.035356
KNN = 8
Alpha = 1
Sigma = 10



ISOMAP and LTSA are correct. MDS and PCA project to an flower-looking thing. Laplacian distorts it largely. LLE and Hessian LLE fail. Diffusion Map shows decent results when sigma is small, and projects to a flower when sigma is large.

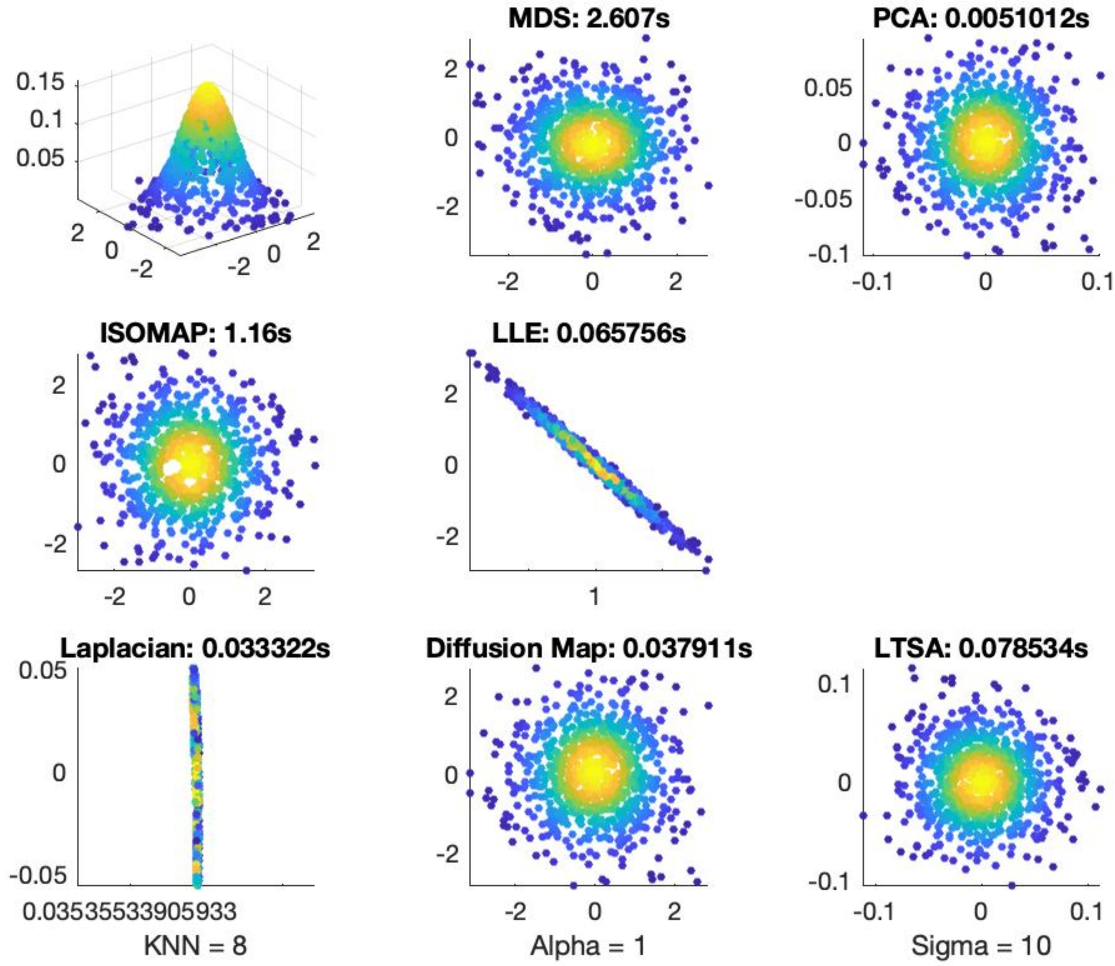
Change the sample rate to 1.2 below.



We can see that Laplacian shows less distorted result and LLE also seems to work a little bit.

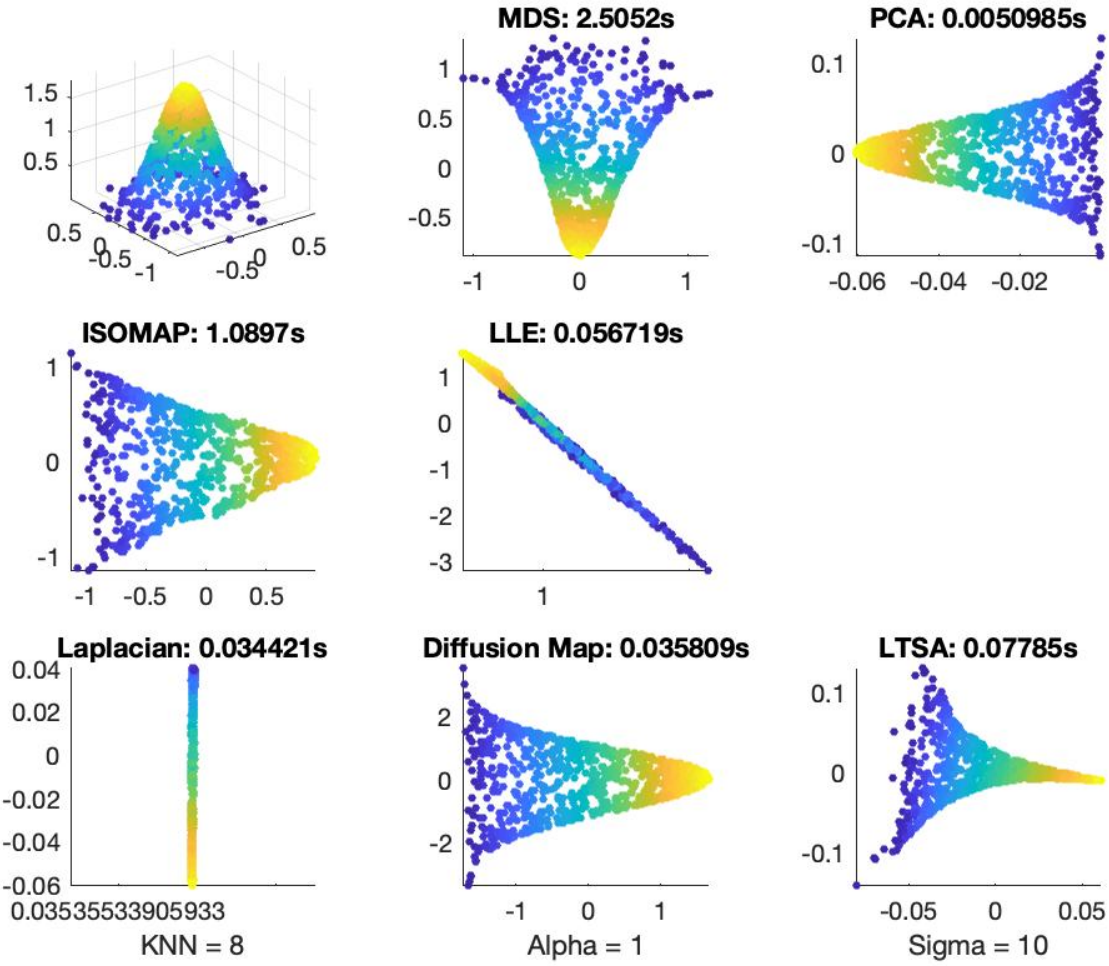
(h) Gaussian

When std=1 (low curvature):



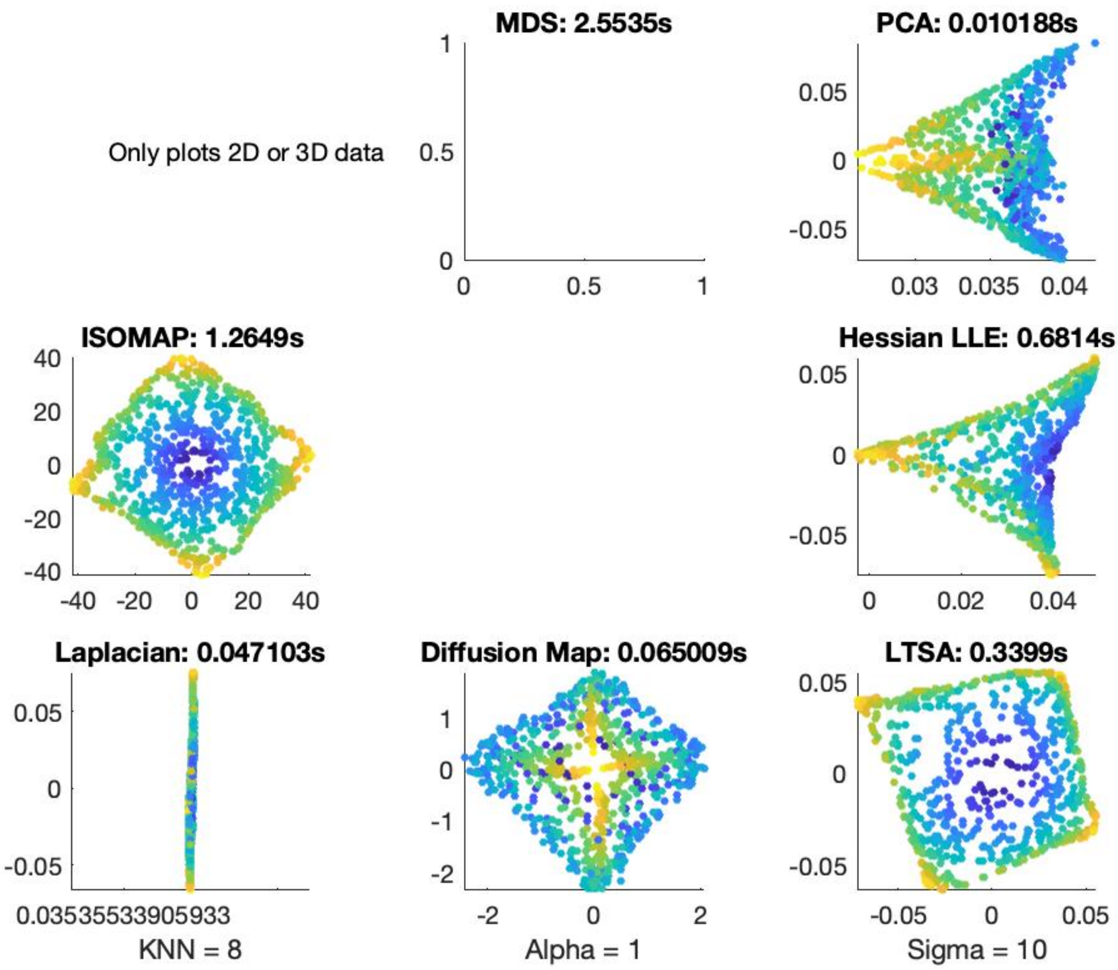
We can see that LLE and Laplacian cannot handle this properly while other methods work well. MDS works extremely slow while ISOMAO works slow. PCA is fast. Hessian LLE cannot work at all.

When std = 0.3 (high curvature):



We can see all of MDS, PCA, ISOMAP, Diffusion Map and LTSA which worked well before now project from the side rather than top-down.

(i) Occluded Disks



We can see that MDS and LLE crash for this, and Laplacian fail totally. However, ISOMAP do a good job and LTSA is also good. So ISOMAP is the best for high-dimensional data.

In conclusion:

Speed: ISOMAP is extremely slow, MDS is slow. PCA is very fast, LLE, Laplacian, LTSA and Diffusion map are fast.

Non-convexity: LTSA can handle non convexity perfectly.

Infer geometry: All manifold learning methods except the linear MDS and PCA can infer the geometry of data to a certain degree.

Sensitive to parameter: Diffusion map is extremely sensitive to parameters.

Handle corners: ISOMAP, Diffusion map and LTSA are all good.

Handle curvature: MDS, PCA, ISOMAP, Diffusion Map and LTSA can all do it but may fail when the curvature is rather high.