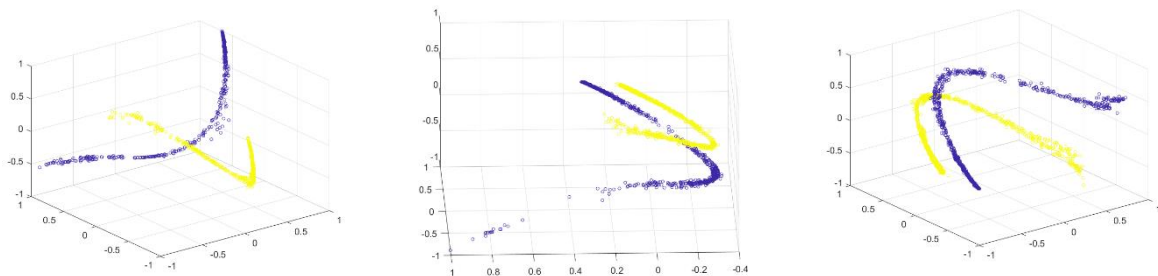


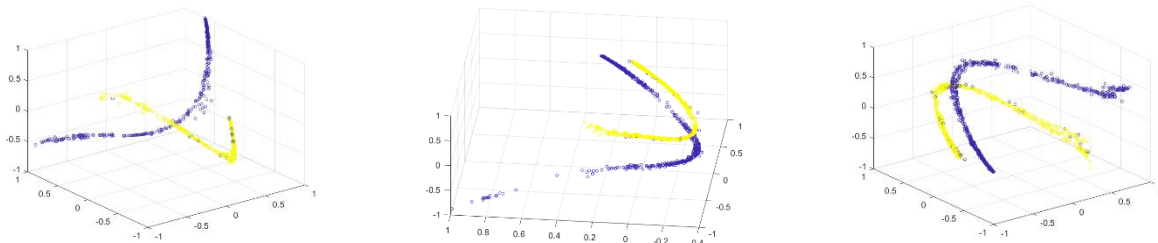
LOWESS (Locally Weighted Scatterplot Smoothing) is a popular tool used in regression analysis that creates a smooth line through a set of points. The most significant advantage LOWESS has over many other methods is the fact that it does not require the specification of a function to fit a model to all of the data in the sample. In addition, this method is very flexible, making it ideal for modeling complex processes for which no theoretical models exist. Non-parametric smoothers like this one try to find a curve of best fit without assuming the data must fit some distribution shape. We use this to determine a smooth curve, which will act as our model for our RANSAC-based procedure.

Review of the problem:

1. These are the original clusters which we'll attempt to classify using a two-step procedure. The first step is based on the moments of the projections followed by a second classification step based on graph Laplacian-based dimensionality reduction.



2. Here we have used merely one metric – the zeroth-order moment to classify the projections and displayed how well it classifies projections. On an average, we see, that the moment-based classification performs well. However, there lie a few points that are misclassified. Incorrectly classifying one projection cluster implies, all the projections associated with that cluster are misclassified, which deteriorates reconstruction quality.

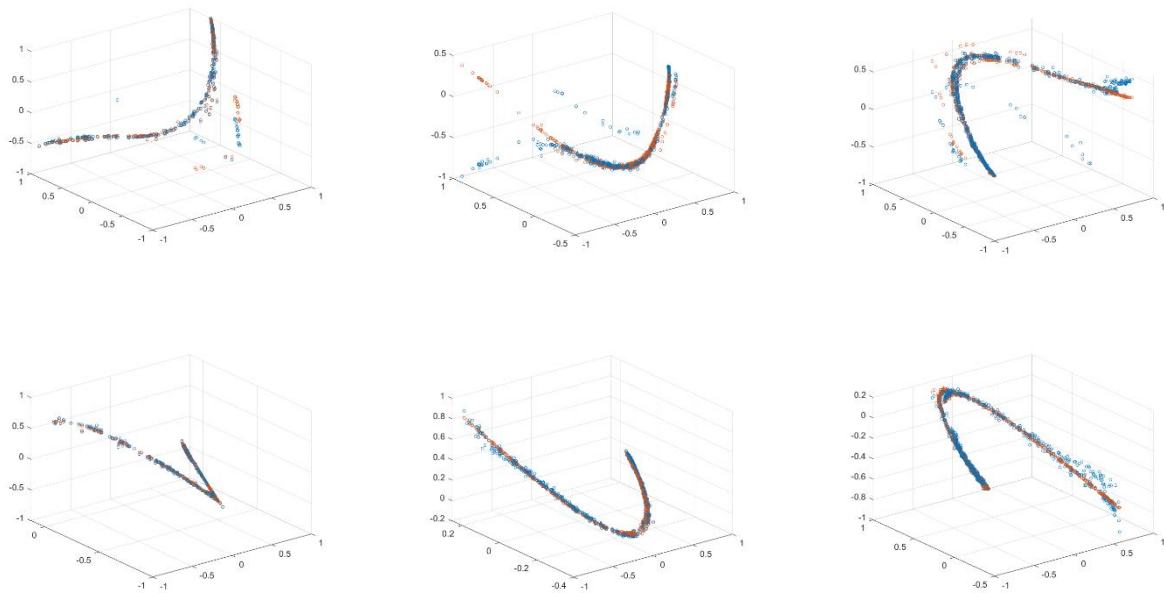


3. We will try to find a smooth curve between the points using the LOWESS procedure. The LOWESS procedure doesn't assume the curves are in the shape of polynomials or any model for that instance. It does this by fitting simple low degree polynomials to subsets of data at a time and smoothly joins those corresponding pieces.

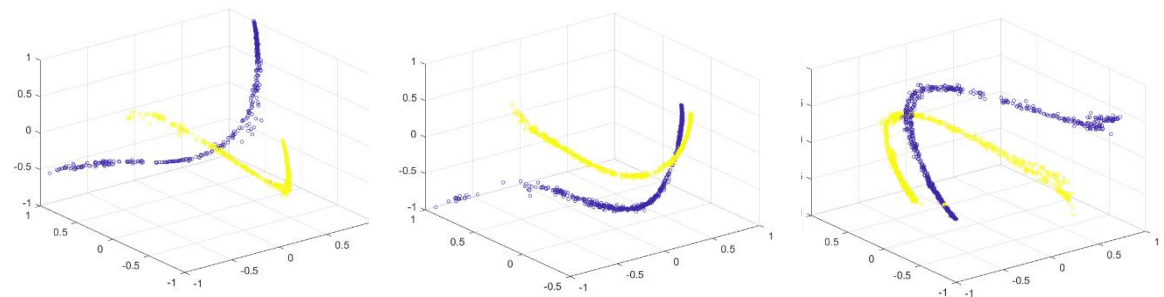
We use the moments-based classification procedure as an initialization as it already gives pretty good classification results. For each class, we follow two steps.

- In the first step, a sample subset of points is randomly selected. A fitting smooth curve model is computed using only the elements of this sample subset.
- In the second step, we check which points of the entire class are consistent with the smooth curve model instantiated in the first step. A point will be considered as an

outlier if it does not fit the smooth curve within some error threshold (user-defined, but remains consistent across multiple objects)



We will repeat the steps mentioned above a few times and select the model which is able to best explain the maximum number of points. As we can see, from the above images, the smooth curves fit nicely to the set of points belonging to each class, while leaving out the points which were misclassified by the zeroth-moment based classification procedure. Using these robust models, we reclassify the points – assign each point to the curve it is closest to and thus succeed in accurately classifying the points.



Advantages of using LOWESS

As discussed above, the biggest advantage LOWESS has over many other methods is the fact that it does not require the specification of a function to fit a model to all of the data in the sample. Instead the analyst only has to provide a smoothing parameter value and the degree of the local polynomial. In addition, LOESS is very flexible, making it ideal for modeling complex processes for which no theoretical models exist.