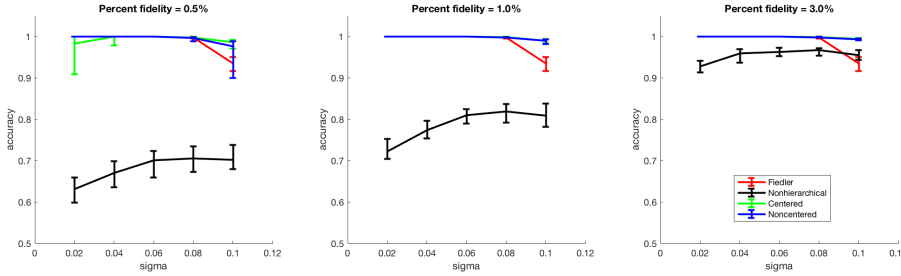


COMPARING HIERARCHICAL METHODS

V. CHEN

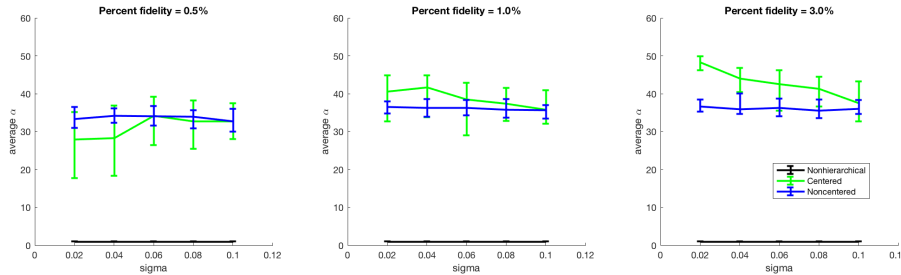
1. Introduction. Inspired by the experiment comparing the different clustering algorithms given in [1], we ran a similar set of experiments to determine if a hierarchical approach was beneficial to clustering. The clustering algorithms we want to compare are: Fiedler thresholding, nonhierarchical pCN, and two hierarchical algorithms for learning τ, α as well as the classifying function u . For each of the four algorithms tested, 50 random realizations of the two moons data set were generated with $r = 1, N = 2000, d = 100$, and varying σ . We used the same RNG seed for the trials across different algorithms so that these 50 realizations are consistent over the different algorithms tested. With each two moons realization, we tested the classification accuracy with varying fidelity percents for the labeled data. Again, the chosen labeled set Z' is consistent across the trials for different algorithms. Our results are summarized in Figure 1.

FIG. 1. Classification accuracy of different algorithms for two moons dataset compared with σ and percent fidelity. Plotted is the median classification accuracy with error bars that indicate the 25 and 75-th quantiles over the 50 trials for each parameter combination. For the pCN and hierarchical algorithms, we used spectral projection up to the first 50 eigenvectors. We ran 100000 iterations with a burn-in period of 1000, and we fixed $\gamma = 0.1$.



2. Learning τ and α . In these experiments, the hierarchical algorithms performed much better in classifying the two moons dataset than the nonhierarchical algorithm. We fixed $\tau = \alpha = 1$ in the nonhierarchical pCN, while we set $\tau^{(0)} = \alpha^{(0)} = 1$ for the hierarchical algorithms. In the hierarchical algorithms, we allowed $\tau \in [0.01, 60]$ and $\alpha \in [0.1, 60]$ as the uniform priors.

FIG. 2. Average values of α from the MCMC. Plotted is the median of the averages of α with error bars that indicate the 25 and 75-th quantiles over the 50 trials for each parameter combination. Notice that α is on average much larger in the hierarchical algorithms.



It appears that fixing $\alpha = 1$ in the nonhierarchical algorithm is partially responsible for the discrepancy in performance. With small α , the prior on u drops off more slowly with the higher eigenvectors, while larger α enforces that these higher eigenvectors have less influence on u . This suggests that for the two moons data set, information for binary classification is concentrated in the lower eigenvectors, as expected.

REFERENCES

- [1] A. L. BERTOZZI, X. LUO, A. M. STUART, AND K. C. ZYGALAKIS, *Uncertainty quantification in the classification of high dimensional data*, ArXiv e-prints, (2017), <https://arxiv.org/abs/1703.08816>.