# Mathematical Sciences Colloquium

Friday, October 18
Fisher Hall 327B
1:05 P.M.

- **Anna Little**
- **Postdoctoral Research Associate**
- **Department of Computational Mathematics, Science and Engineering**
- **Michigan State University**

## Robust Statistical Procedures for Clustering in High Dimensions

**Abstract:** This talk addresses multiple topics related to robust statistical procedures for clustering in high dimensions, including path-based spectral clustering (a new method), classical multidimensional scaling (an old method), and clustering in signal processing. Path-based spectral clustering is a novel approach which combines a data driven metric with graph-based clustering. Using a data driven metric allows for fast algorithms and strong theoretical guarantees when clusters concentrate around low-dimensional sets. Another approach to high-dimensional clustering is classical multidimensional scaling (CMDS), a dimension reduction technique widely popular across disciplines due to its simplicity and generality. CMDS followed by a simple clustering algorithm can exactly recover all cluster labels with high probability when the signal to noise ratio is high enough. However, scaling conditions become increasingly restrictive as the ambient dimension increases, illustrating the need for robust unbiasing procedures in high dimensions. Clustering in signal processing is the final topic; in this context each data point corresponds to a corrupted signal. The classic multireference alignment problem is generalized to include random dilation in addition to random translation and additive noise, and a wavelet based approach is used to define an unbiased representation of the target signal(s) which is robust to high frequency perturbations.

**Michigan Tech**