Title: Robust clustering tools based on optimal transportation

Abstract: A robust clustering method for probabilities in Wasserstein space is introduced. This new 'trimmed k-barycenters' approach relies on recent results on barycenters in Wasserstein space that allow intensive computation, as required by clustering algorithms. The possibility of trimming the most discrepant distributions results in a gain in stability and robustness, highly convenient in this setting. As a remarkable application we consider a parallelized estimation setup in which each of m units processes a portion of the data, producing an estimate of k-features, encoded as k probabilities. We prove that the trimmed k-barycenter of the $m\times k$ estimates produces a consistent aggregation. We illustrate the methodology with simulated and real data examples.