1086-VI-417   **Anna V Little*** (`alittle2@ju.edu`), **Mauro Maggioni** and **Lorenzo Rosasco**. *Estimating the Intrinsic Dimension of High-Dimensional Data Sets.*

This talk introduces a novel approach for estimating the intrinsic dimension of noisy, high-dimensional point clouds. A general class of sets which are locally well-approximated by k dimensional planes but which are embedded in a $D \gg k$ dimensional Euclidean space are considered. The dimension is estimated via a new multiscale algorithm that generalizes principal component analysis (PCA). The classical PCA approach recovers the dimension when the data is linear but fails when the data is non-linear, overestimating the intrinsic dimension. This new multiscale algorithm exploits the low-dimensional structure of the data, so that its power depends on k rather than D, and is robust to small sample size, noise, and non-linearities in the data. (Received August 30, 2012)

1