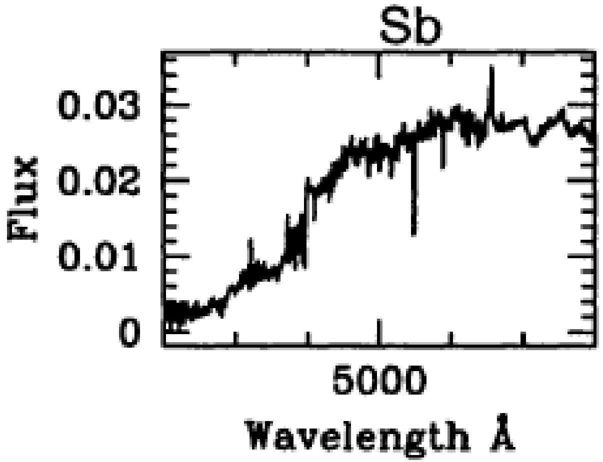
MAPPING GALAXY SPECTRA

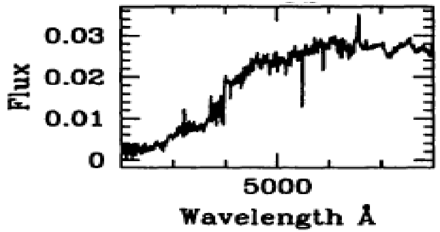
Tamas Budavari / Johns Hopkins University

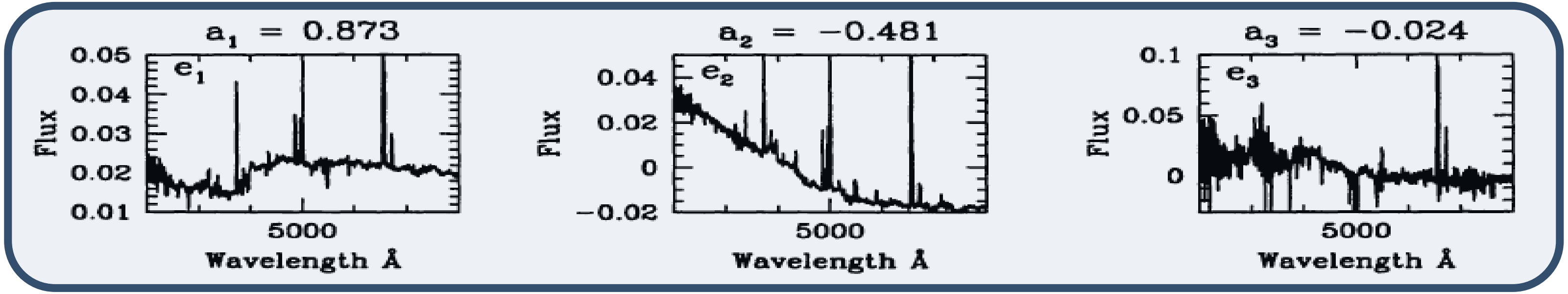
Understanding Spectra

- Photometric redshifts
  - Spectral templates
- Diversity of spectra
  - Census of variations
- Semi-analytic models

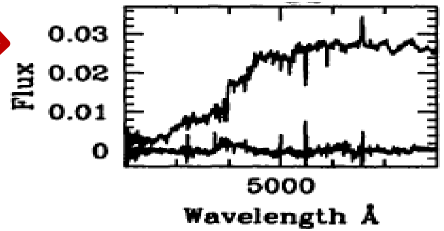


Galaxy Light ~ Linear Combination





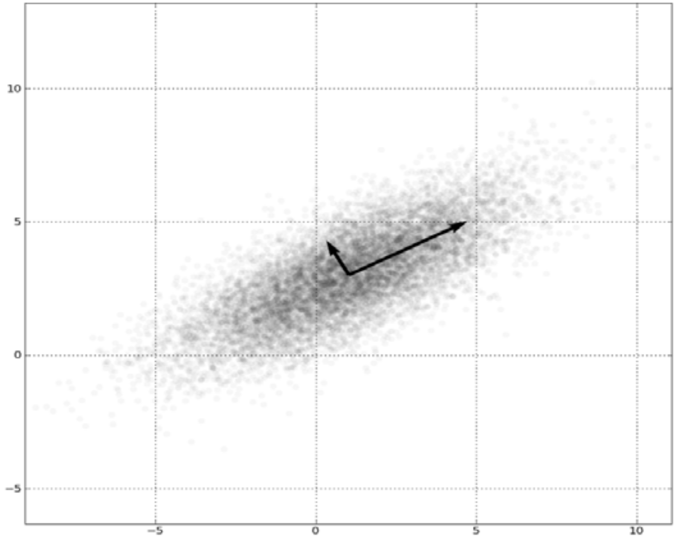
Connolly, TB+ (1999)



Principal Component Analysis

- Principal directions
  - Directions of largest variations
  - Eigenproblem of covariances
  - Singular Value Decomposition

- Problems
  - Needs lots of memory
  - Only need largest ones
  - Very sensitive to outliers



Streaming PCA

- Initialization
  - Eigensystem of a small, random subset
  - Truncate at *p* largest eigenvalues

$$C \approx E_p \Lambda_p E_p^{\mathrm{T}}$$

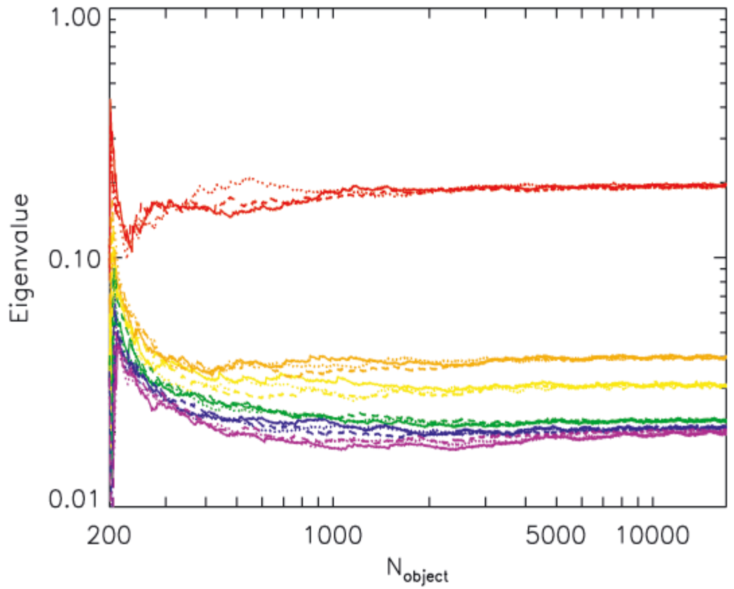
- Incremental updates
  - Mean and the low-rank *A* matrix
  - SVD of *A* yields new eigensystem

$$C \approx \gamma E_p \Lambda_p E_p^{\mathrm{T}} + (1-\gamma) y y^{\mathrm{T}}$$
$$\approx A A^{\mathrm{T}}$$

- Randomized algorithm!

Galaxy Spectra

- Incremental updates
  (TB, Wild+ 2008 MNRAS)
  - From 3 days on big computer
  - To 15 minutes on a desktop
- Mix in robust statistics
  - Deal w/outliers *(Maronna 2005)*



Robust Statistics

In a nutshell

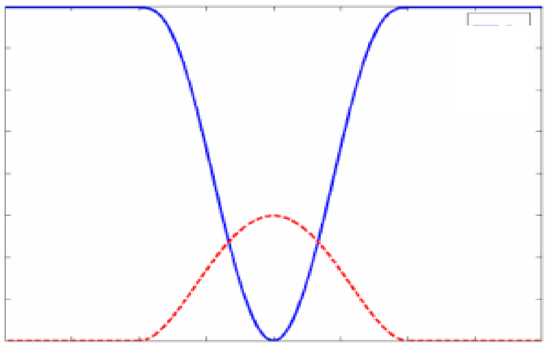
Location

- M-estimates of the location

$$L(x_1, \ldots, x_n; \mu) = \prod_{i=1}^{n} f_0(x_i - \mu)$$

$$\widehat{\mu} = \arg\min_{\mu} \sum_{i=1}^{n} \rho\left(x_i - \mu\right)$$ with $$\rho = -\log f_0$$

$$\sum_{i=1}^{n} \rho'(x_i - \widehat{\mu}) = 0$$



  - E.g., mean if x^2, median if $|x|$

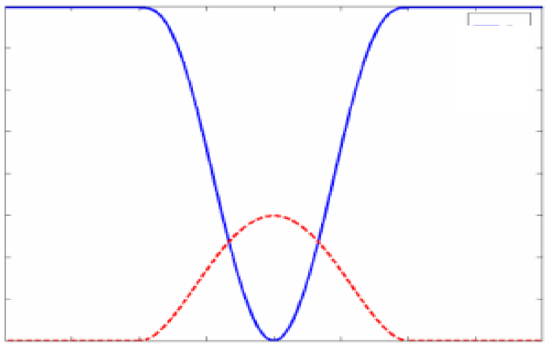
Location

- M-estimates of the location

$$L(x_1, \ldots, x_n; \mu) = \prod_{i=1}^{n} f_0(x_i - \mu)$$

$$\widehat{\mu} = \arg\min_{\mu} \sum_{i=1}^{n} \rho\,(x_i - \mu)$$ with $$\rho = -\log f_0$$

$$\sum_{i=1}^{n} \rho'(x_i - \widehat{\mu}) = 0$$



  - E.g., mean if x^2, median if $|x|$
- Intuitive
  - Weights: $\sum_{i=1}^{n} W(x_i - \widehat{\mu})(x_i - \widehat{\mu}) = 0$

Dispersion

- M-estimates of the scale

$$\frac{1}{\sigma} f_0\left(\frac{x}{\sigma}\right)$$

$$\widehat{\sigma} = \arg\max_{\sigma} \frac{1}{\sigma^n} \prod_{i=1}^{n} f_0\left(\frac{x_i}{\sigma}\right)$$

$$\rho(t) = t\psi(t)$$
$$\psi = -f_0'/f_0$$

$$\frac{1}{n}\sum_{i=1}^{n} \rho\left(\frac{x_i}{\widehat{\sigma}}\right) = 1$$

  - E.g., rms if x^2

- Intuitive

$$\widehat{\sigma}^2 = \frac{1}{n\delta}\sum_{i=1}^{n} W\left(\frac{x_i}{\widehat{\sigma}}\right) x_i^2 \quad \text{with} \quad W(x) = \begin{cases} \rho(x)/x^2 & \text{if} \quad x \neq 0 \\ \rho''(0) & \text{if} \quad x = 0 \end{cases}$$

Robust PCA

- PCA minimizes σ_{RMS} of the residuals $r = y - Py$
  - Quadratic formula: Σr^2 extremely sensitive to outliers
- We optimize a robust M-scale σ^2 (Maronna 2005)
  - Implicitly given by

$$\frac{1}{N}\sum_{n=1}^{N}\rho\left(\frac{r_n^2}{\sigma^2}\right)=\delta$$

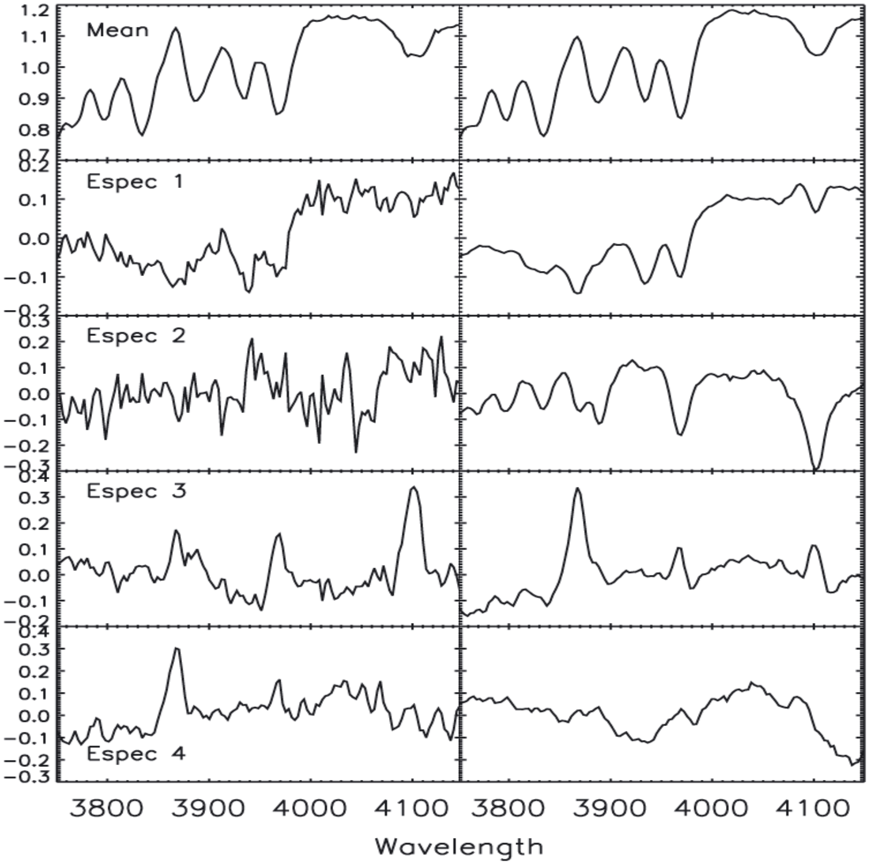
$$\mu = \left(\sum w_n x_n\right) \Big/ \left(\sum w_n\right)$$

$$C = \sigma^2 \left[\sum w_n (x_n - \mu)(x_n - \mu)^{\mathrm{T}}\right] \Big/ \left(\sum w_n r_n^2\right)$$

- Fits in with the iterative method!

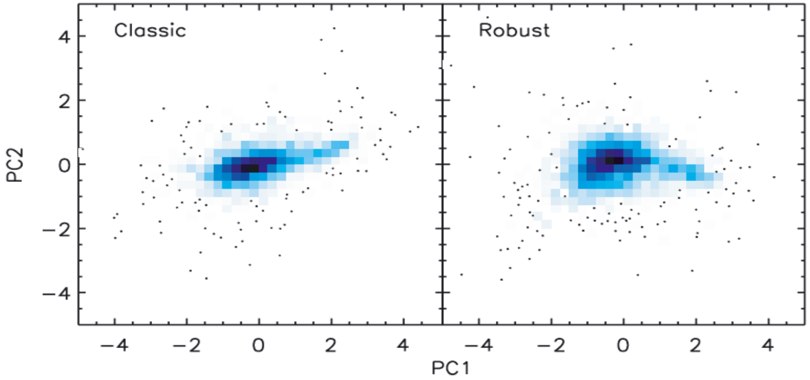
Galaxy Spectra

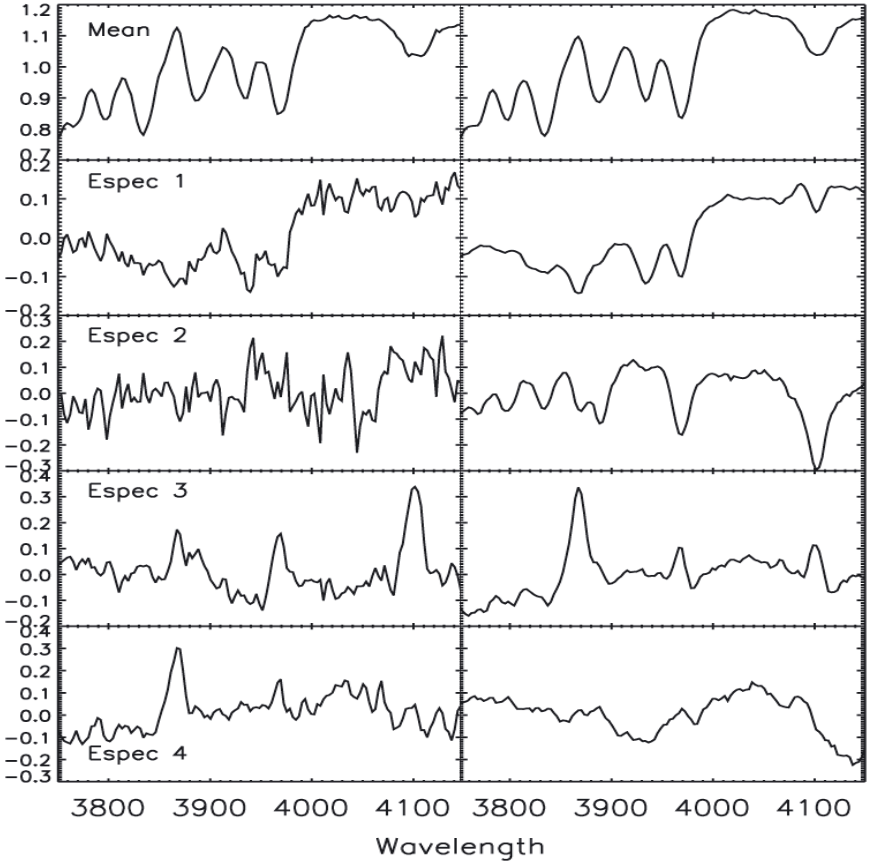
- High SNR eigenfunctions
  - Sign of robustness



Galaxy Spectra

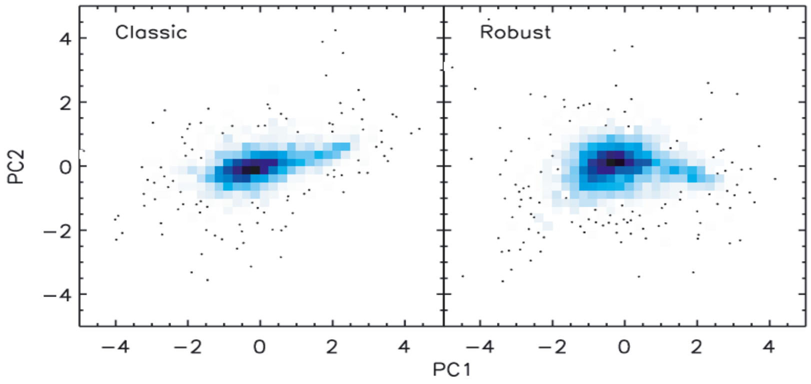
- High SNR eigenfunctions
  - Sign of robustness
- It makes a difference



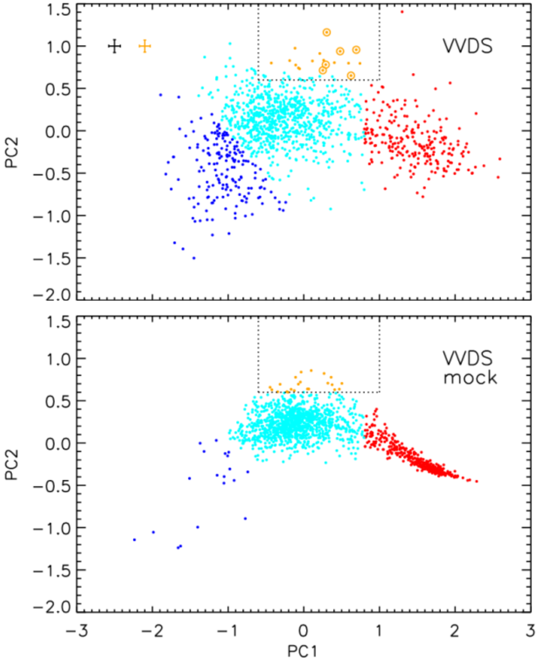


Quenching of Star Formation

- Identify post-starburst galaxies
  - VVDS compared to mock
- Consistent w/being descendants of gas-rich major mergers



TB, Wild+ (2009 MNRAS)

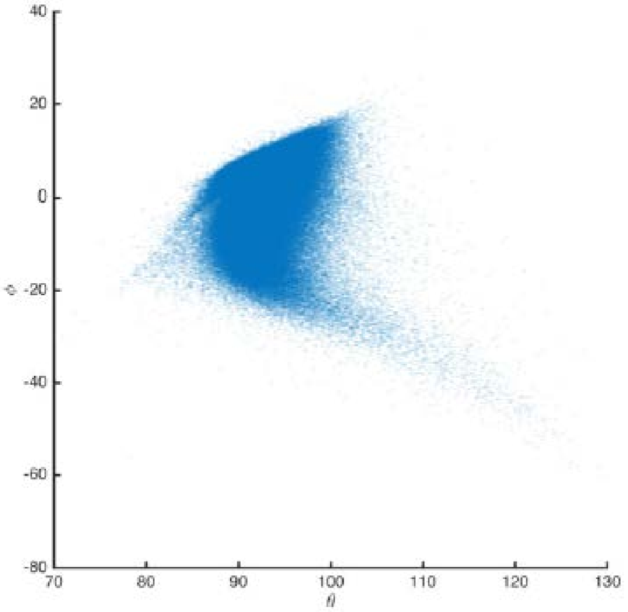


Wild, TB+ (2009)

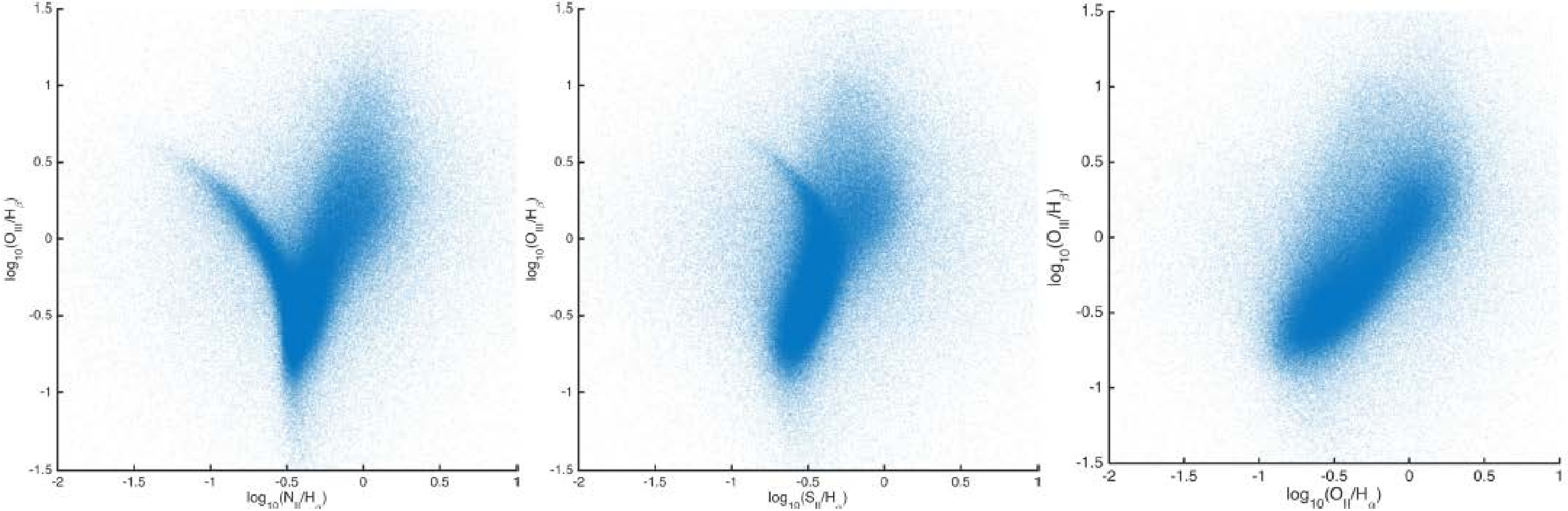
PCA of SDSS DR7 Spectra

- Days on a big-memory machine
- Continuous distribution
- Messy with lots of outliers

- Plot of mixing angles:
  - Take first 3 components →

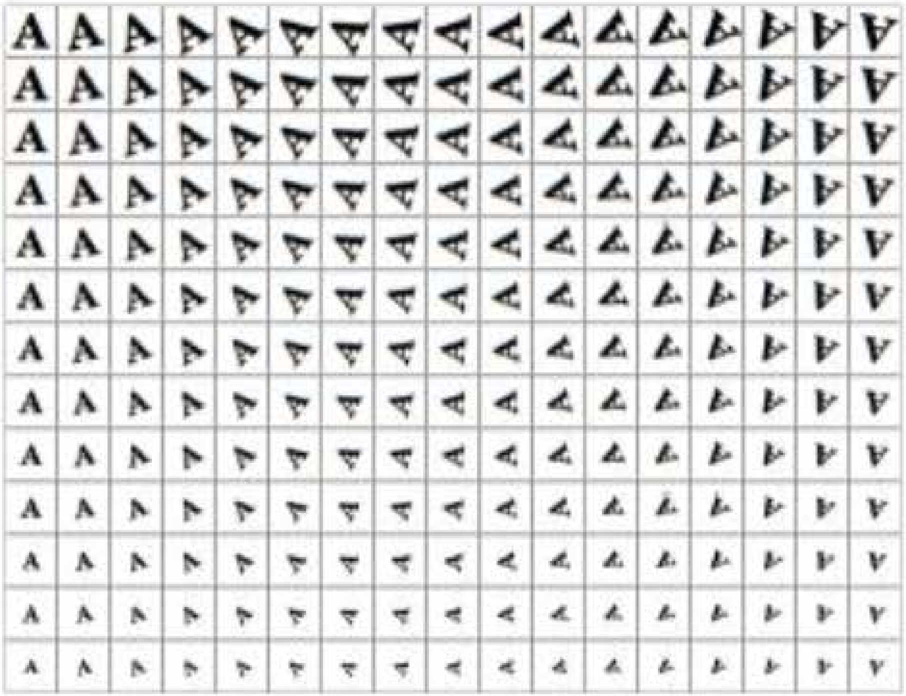


BPT Diagrams

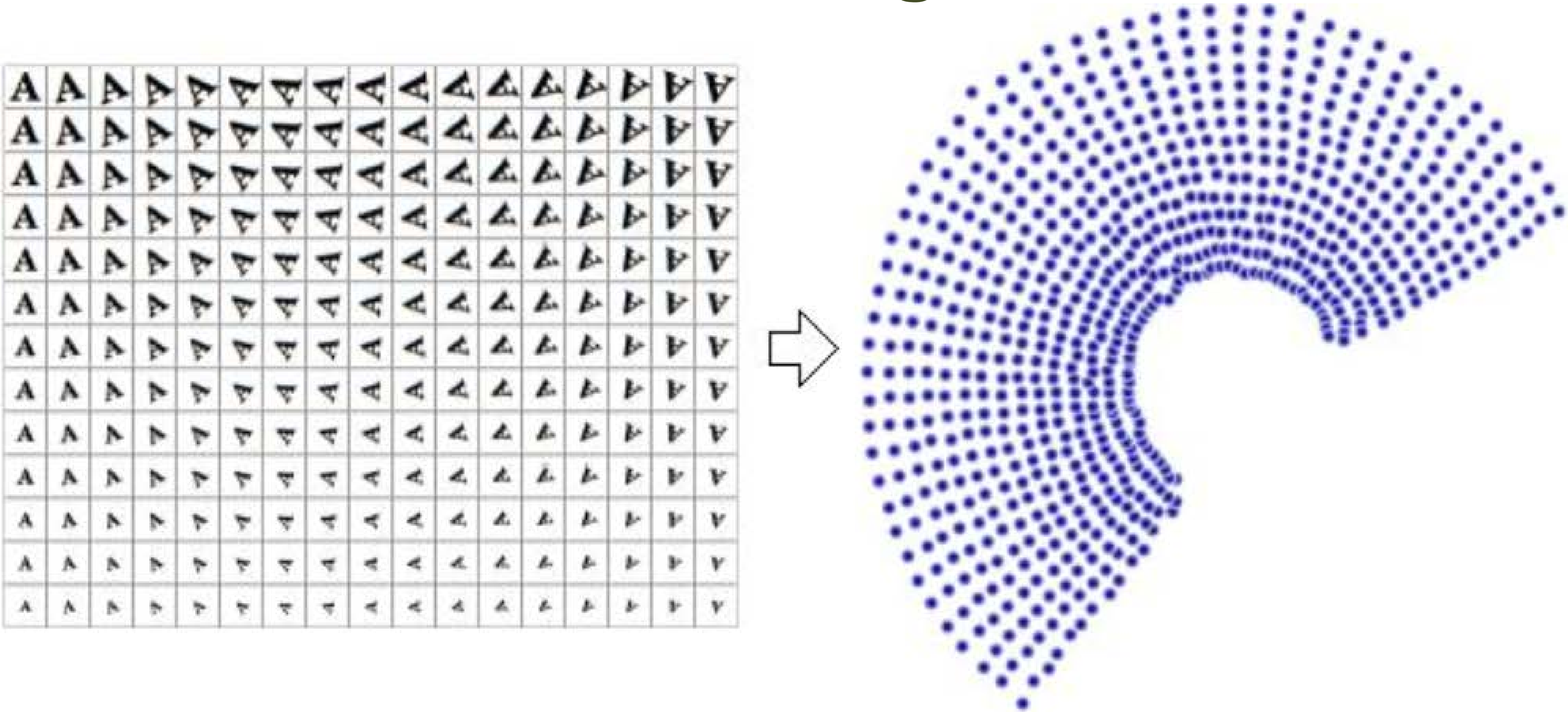


What Parameters?

Nonlinear Embedding



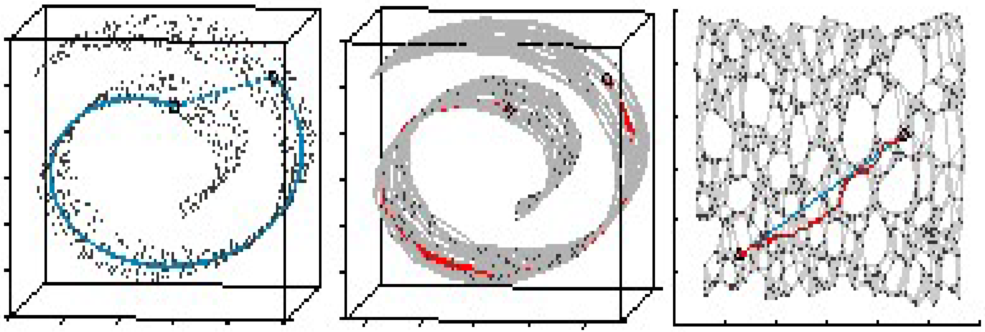
Nonlinear Embedding



Nonlinear Embedding

- ISOMAP and LLE in Science Magazine (2000)
  - ISOMAP – geodesic distance
    - Preserves metric
  - LLE – locally linear
    - Preserves angles

BASED ON LOCAL SIMILARITIES

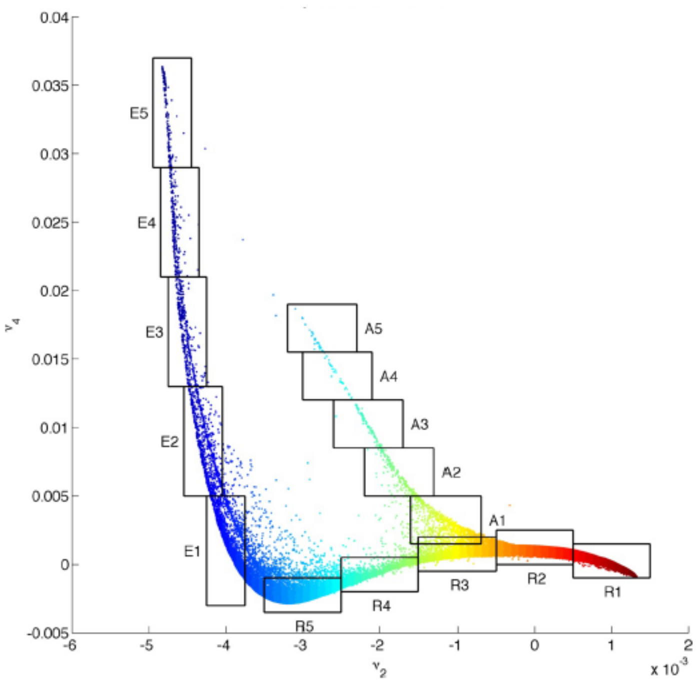


Mapping Galaxy Spectra: Global and Local Views*

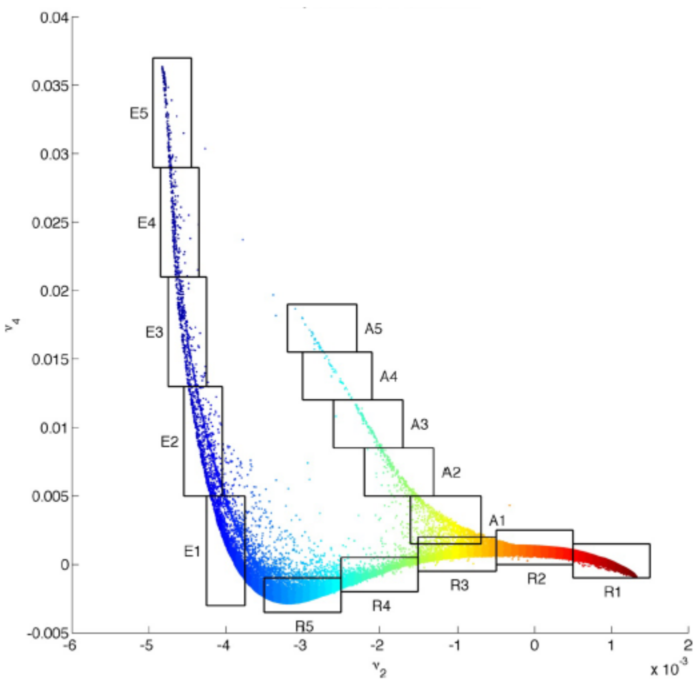
David Lawlor[1,2], Tamás Budavári[3], and Michael Mahoney[4,5]

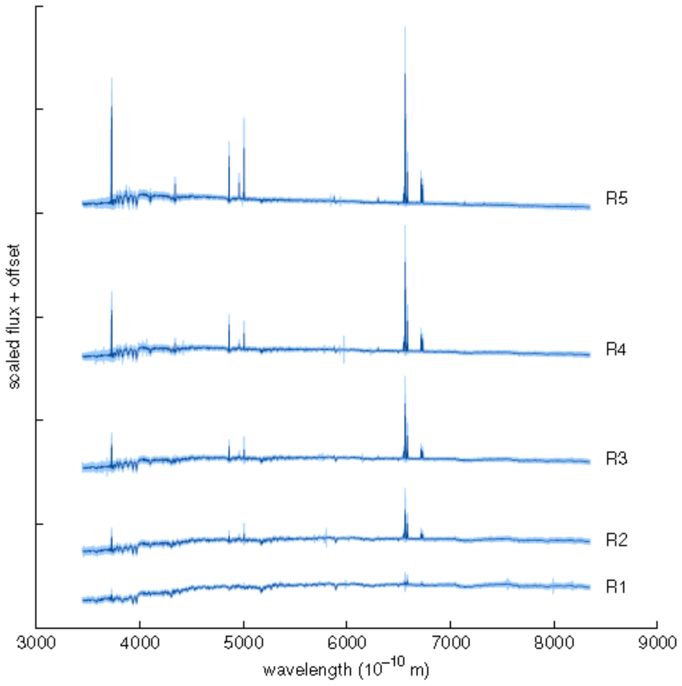
[1]Statistical and Applied Mathematical Sciences Institute
[2]Department of Mathematics, Duke University
[3]Department of Applied Mathematics and Statistics, the Johns Hopkins University
[4]International Computer Science Institute
[5]Department of Statistics, University of California, Berkeley

Diffusion Maps of SDSS

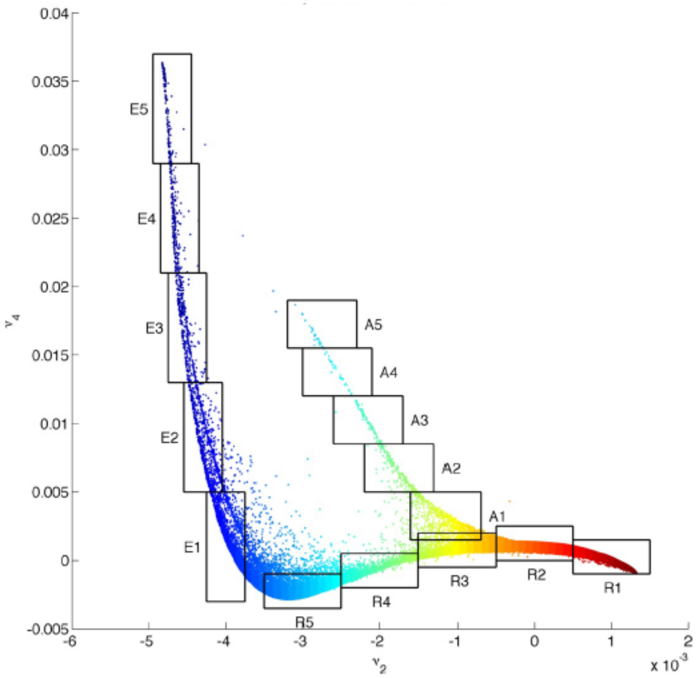


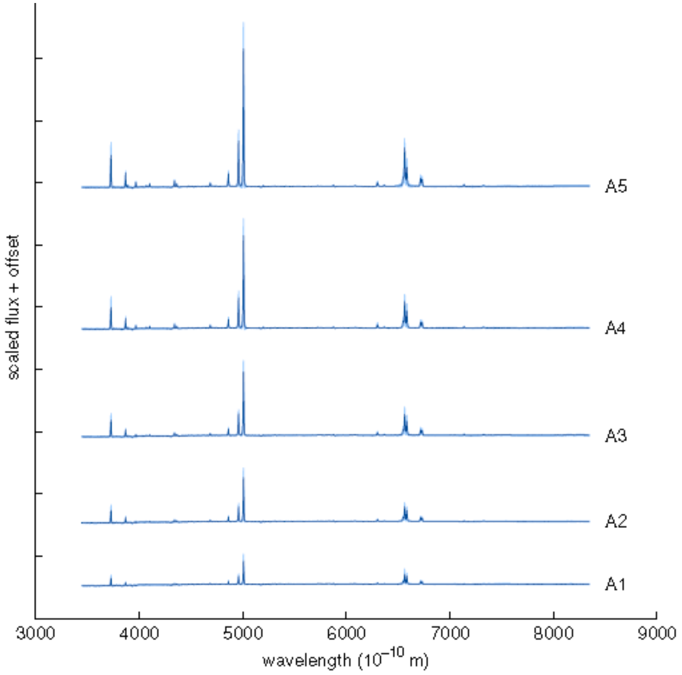
Diffusion Maps of SDSS



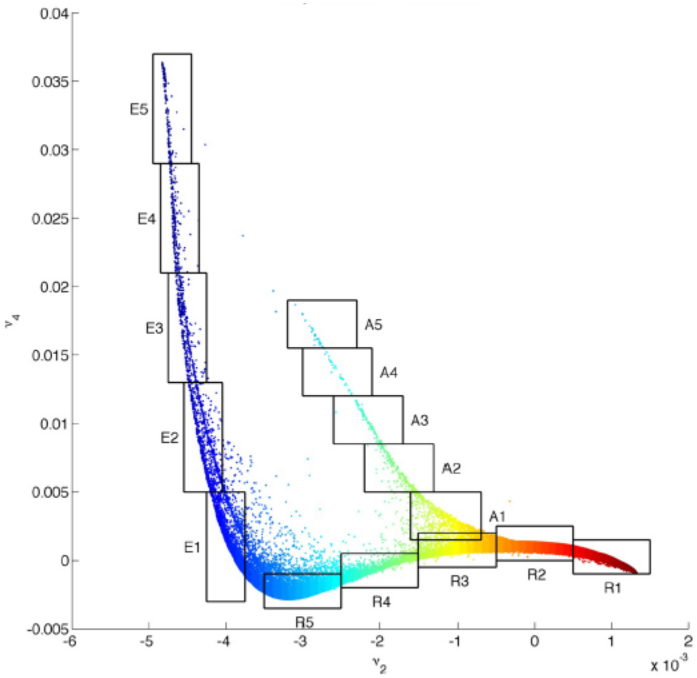


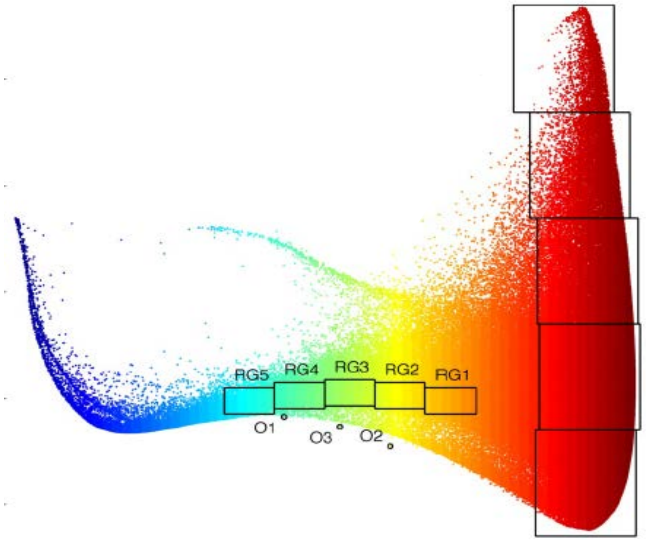
Diffusion Maps of SDSS





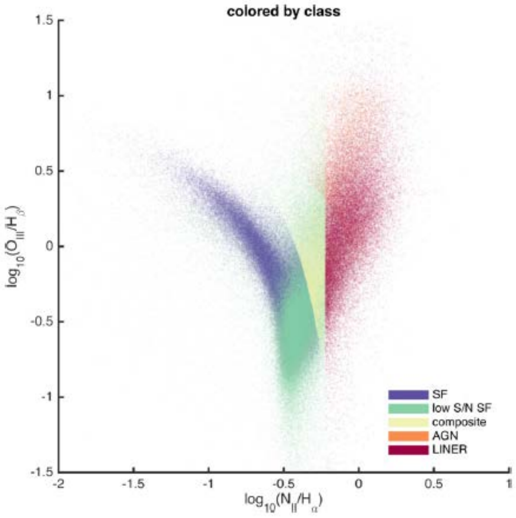
Diffusion Maps of SDSS

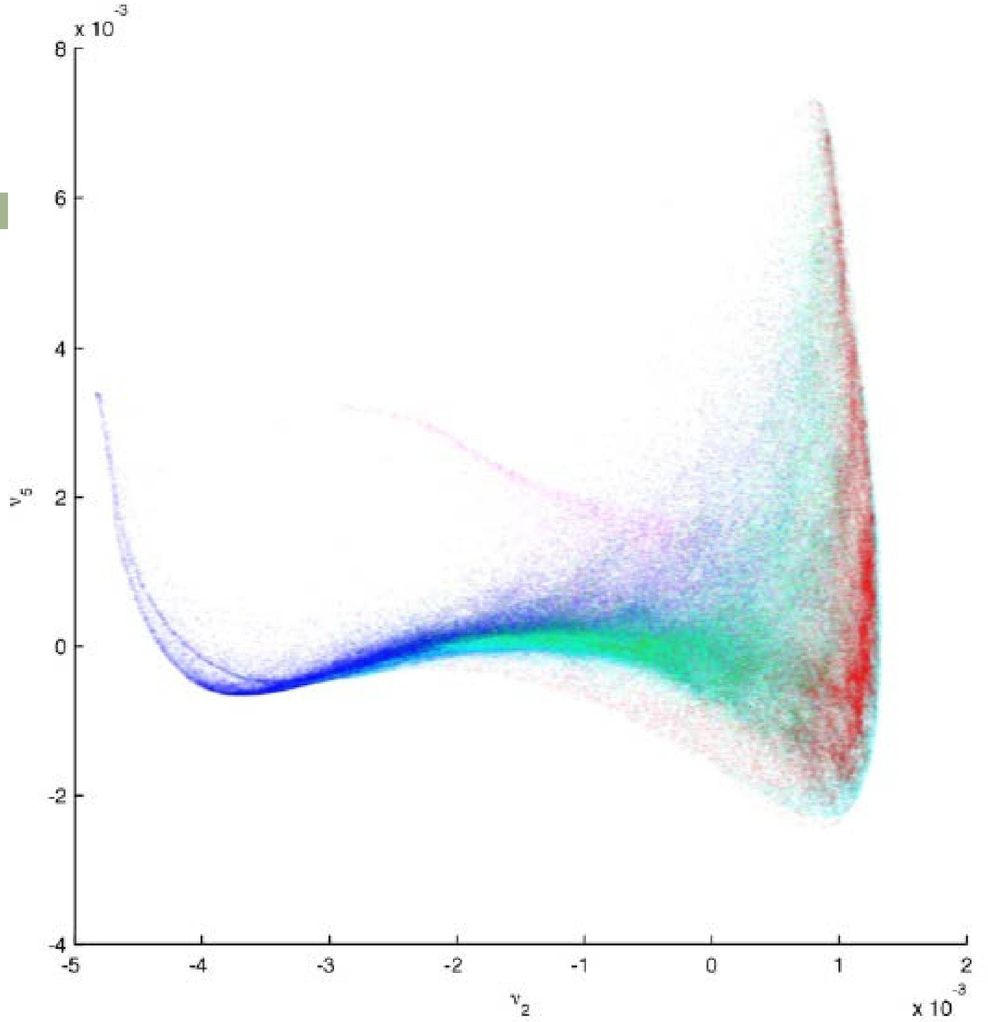




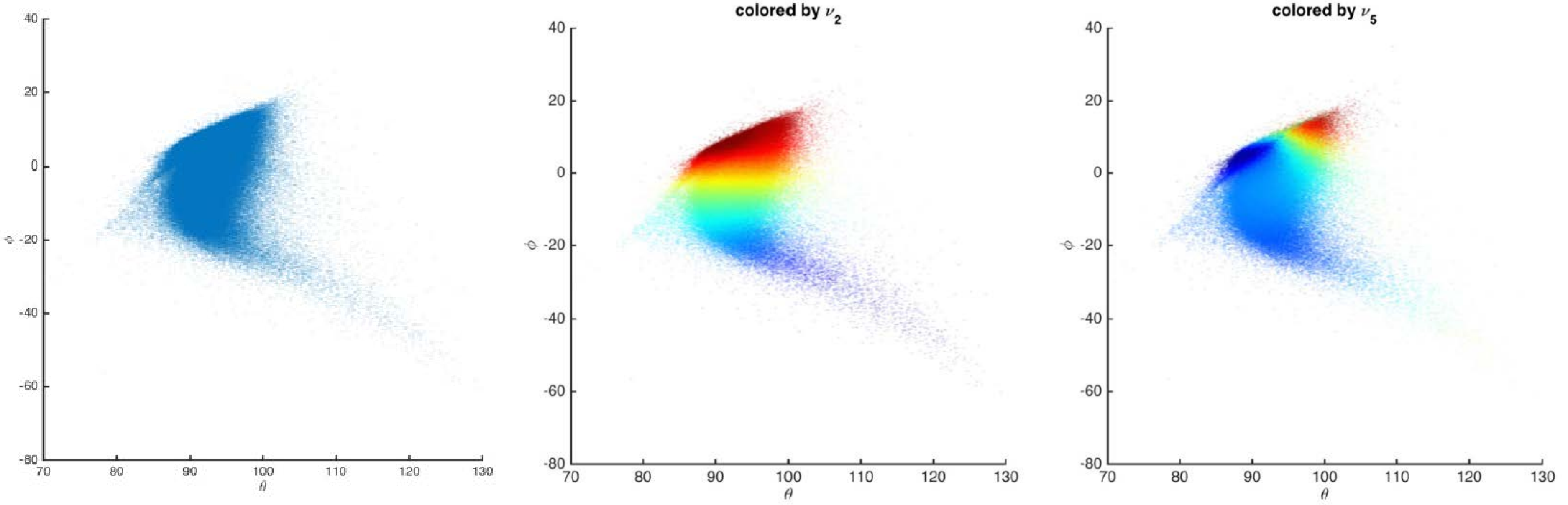
Spectral Classes

- SDSS classification
  - By Jarle Brinchmann

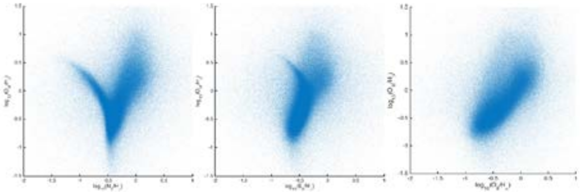


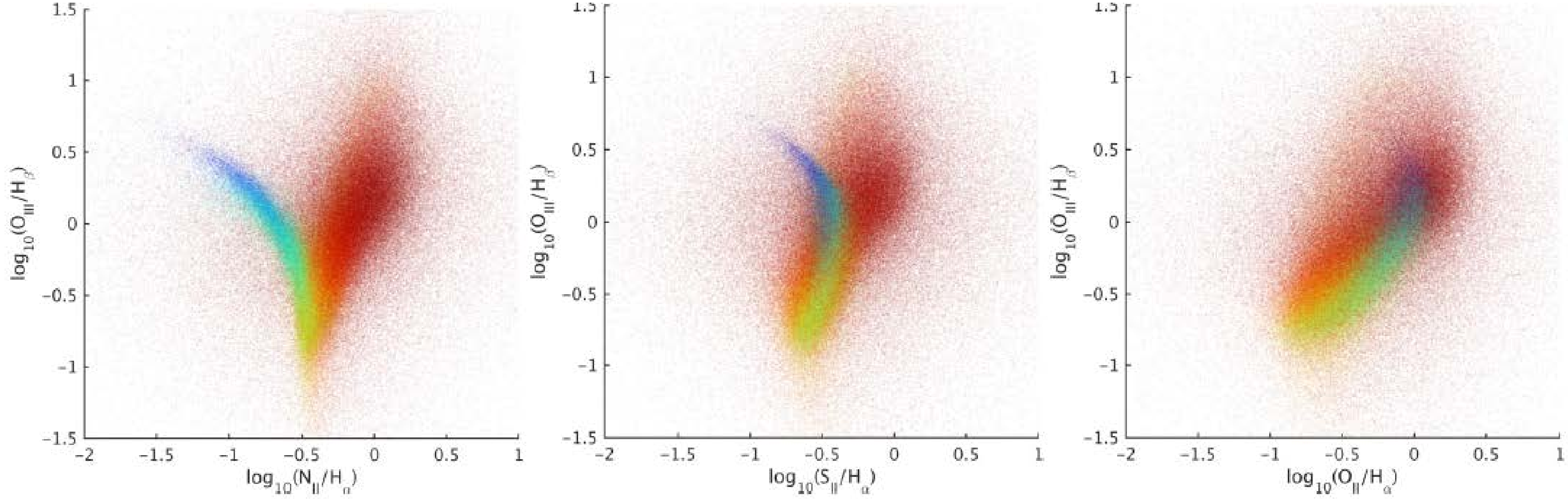


SDSS eCoeff 3D→2 angles

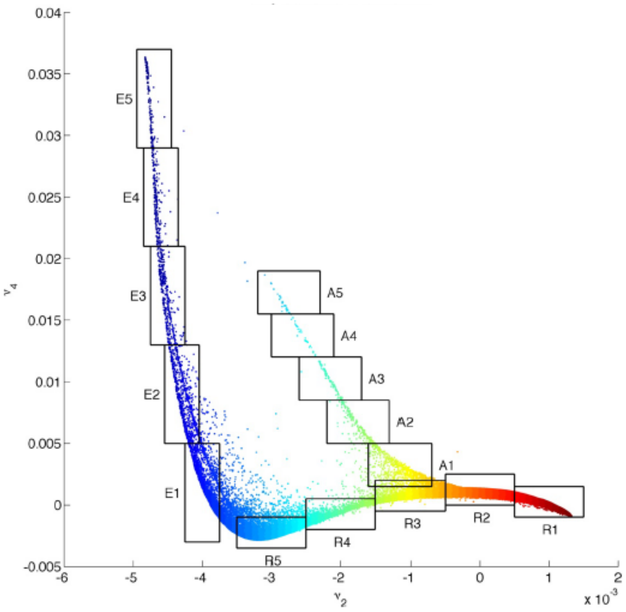


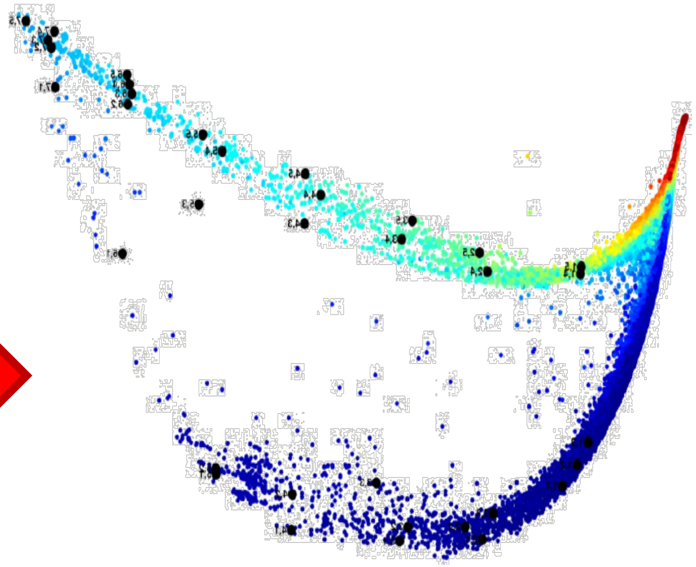
BPT Diagrams





Locally-Biased Embedding



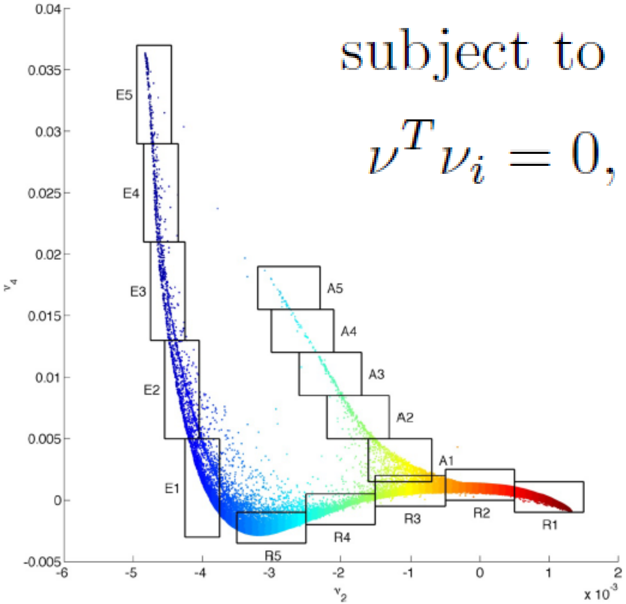


Locally-Biased Embedding

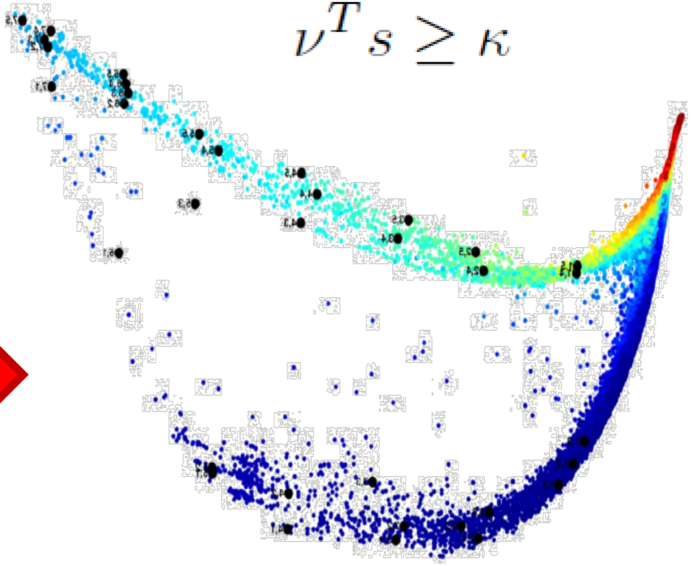
$$\nu_j = \text{argmax} \quad \nu^T M \nu$$

$$\text{subject to} \quad \nu^T \nu = 1$$

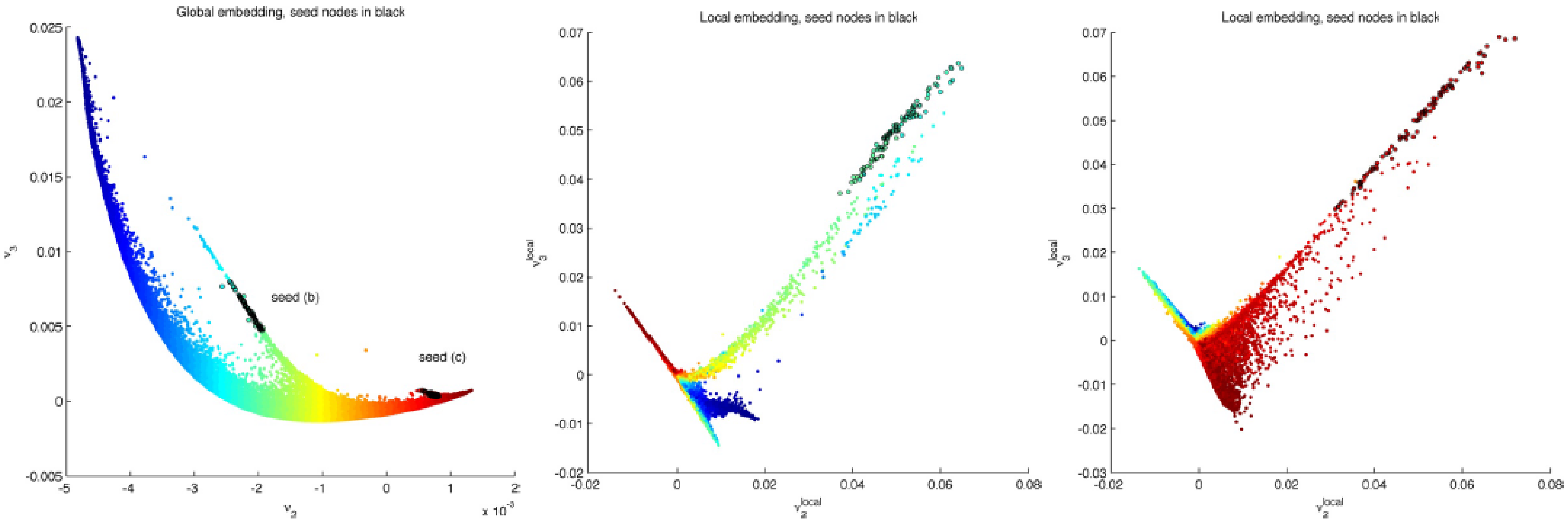
$$\nu^T \nu_i = 0, \; 1 \leq i < j$$



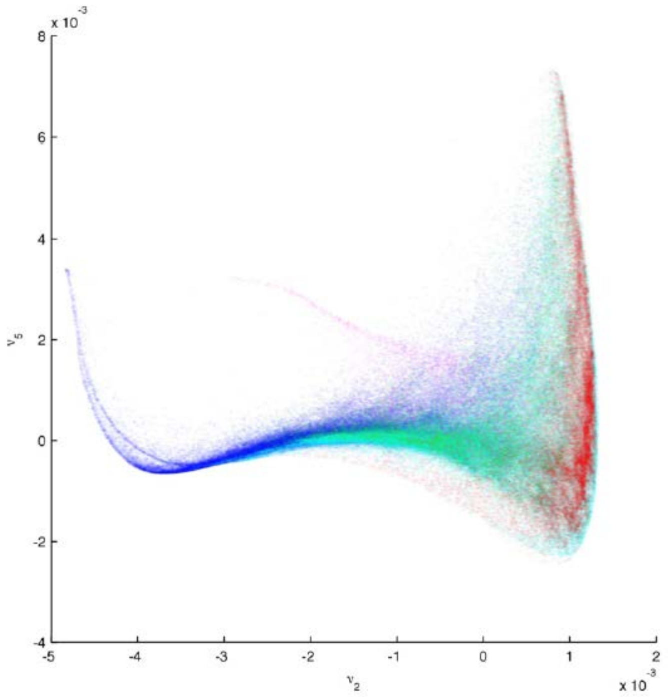
$$\nu^T s \geq \kappa$$



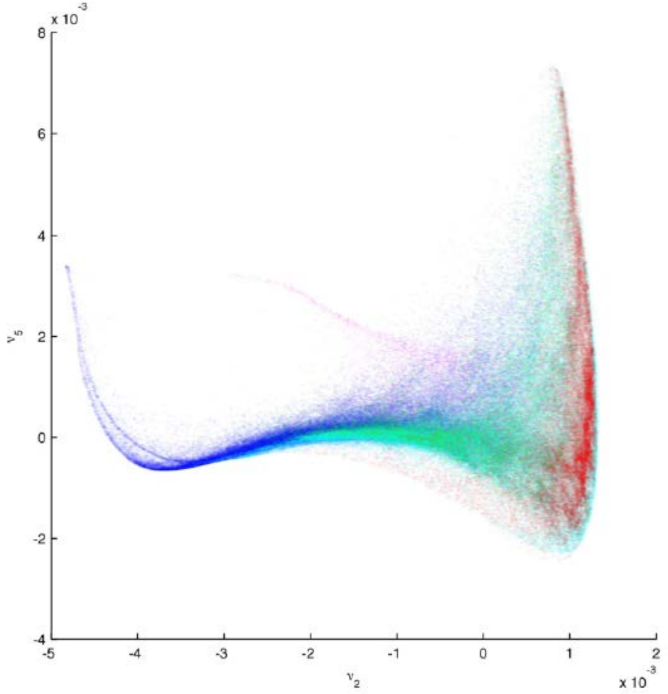
Different Views

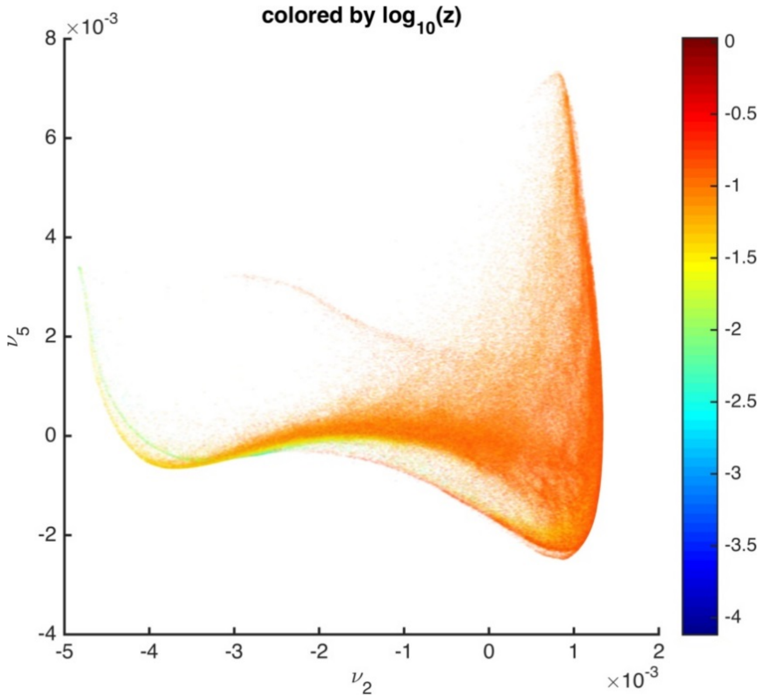


Double Strand of Blue



Double Strand of Blue





Summary

- Clean trends of known features w/ few outliers
- Unknown patterns & new insights
- Toward better spectrum models
  - Lines and Continuum
- Locally biased approach yields new "microscope"
  - Help look for subtleties