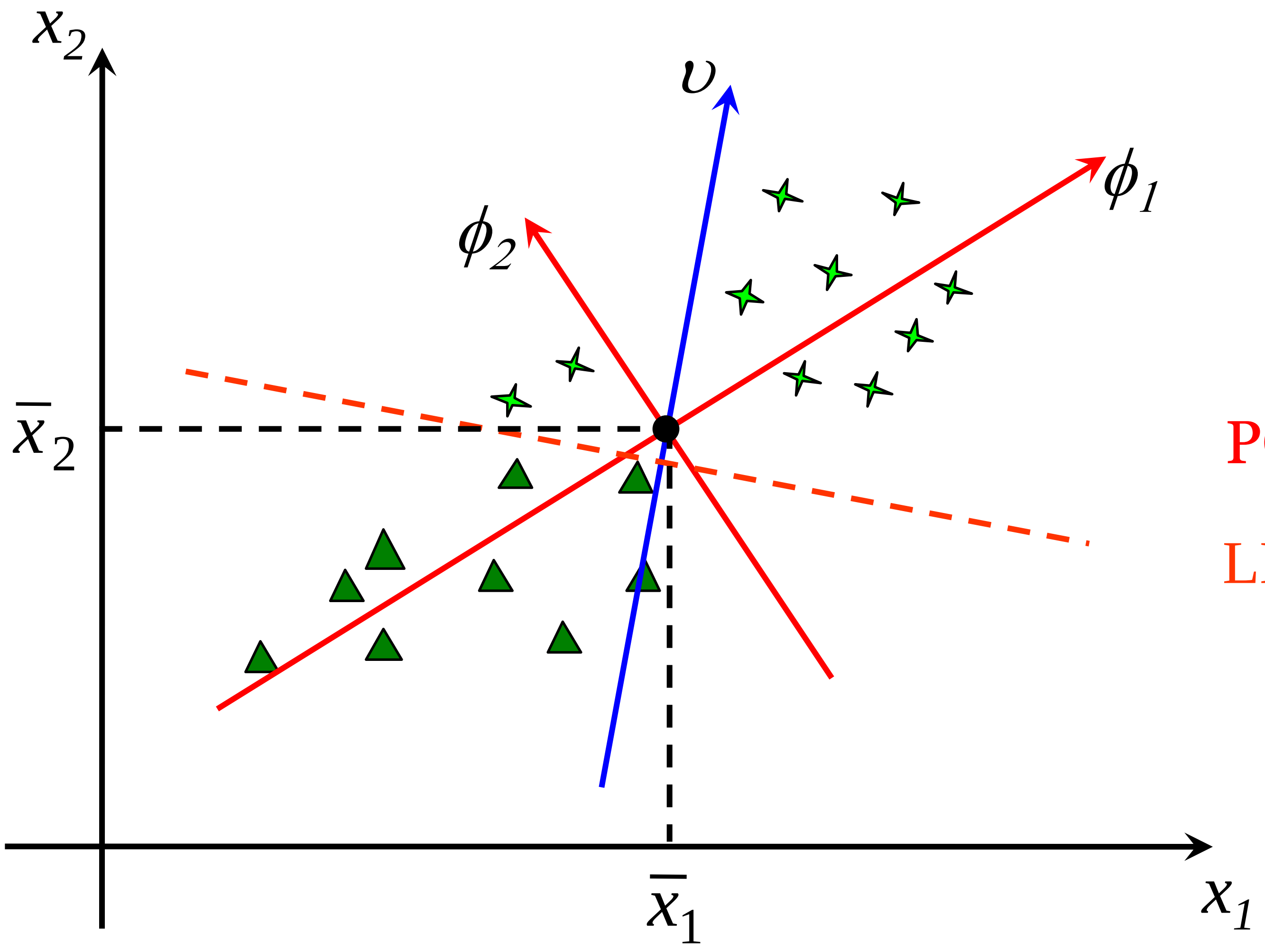
THE UNIVERSITY of TENNESSEE KNOXVILLE

Lecture 14: Linear Dimensionality Reduction for Spectra and Images

Instructor: Sergei V. Kalinin

Geometric Idea of PCA and LDA



PCA: (ϕ_1, ϕ_2)

LDA: υ

From Intelligent Data Analysis and Probabilistic Inference by Longin Jan Latecki Temple University

Spectroscopic Imaging

Scanning probe microscopy:
- Force-distance curve measurements
- Current-voltage measurements
- Piezoresponse force/electrochemical strain spectroscopy

Electron microscopy:
- Electron Energy Loss Spectroscopy

Optical microscopy:
- Hyperspectral imaging
- Time resolved measurements

Mass-spectrometry:
- Secondary ion MS imaging

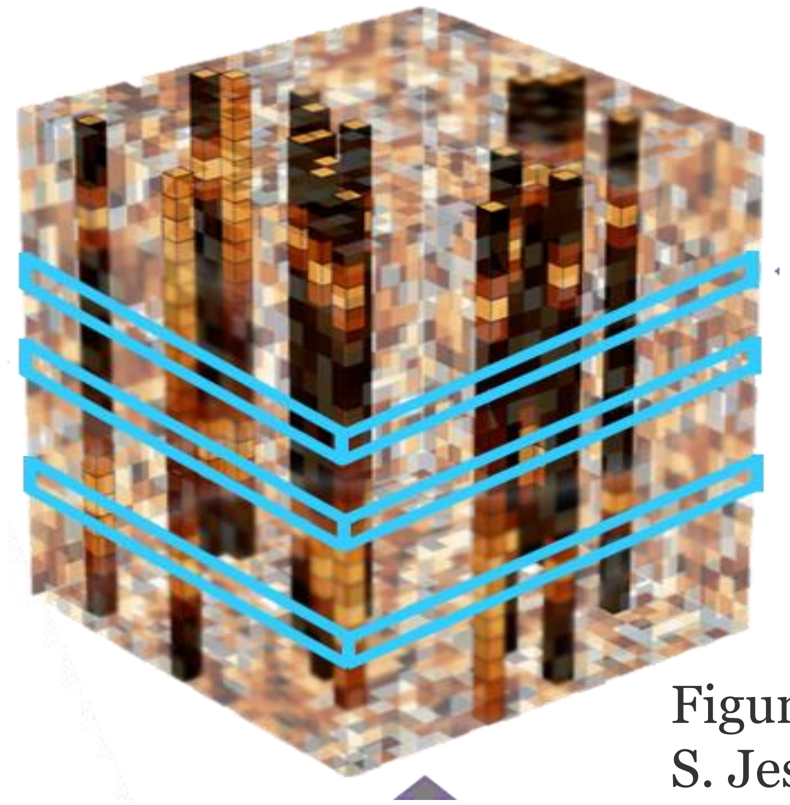


Figure by S. Jesse

In many cases, measured signal can be represented or approximated as a linear combination of signals. However, their functional forms are generally unknown

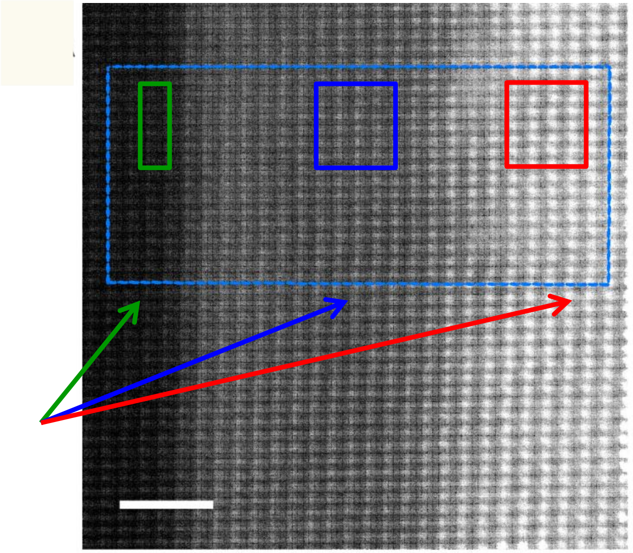
Very important: convolution with resolution function is also mixing

Multiple Linear Regression

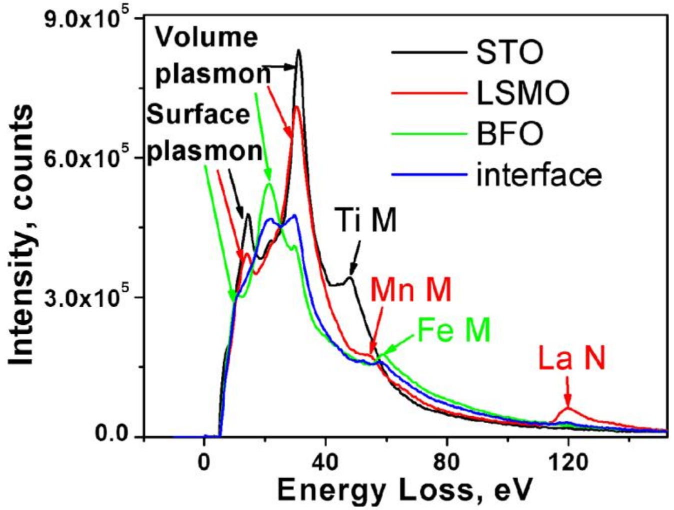
Linear mixing $S(\mathbf{x},\mathbf{R}) = \sum_i \boxed{a_i(\mathbf{x})}\, w_i(\mathbf{R})$ +N but $w_i(\mathbf{R})$ are **known**

STEM of STO/LSMO/BFO interface

Standard spectra



Low-loss EELS spectra of three components

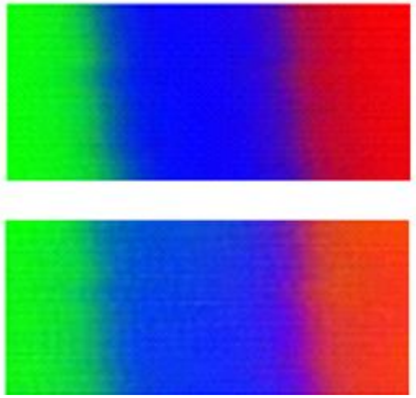


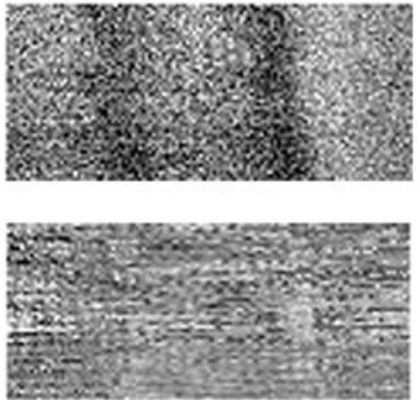
A.Y. Borisevich et al, *Suppression of Octahedral Tilts and Associated Changes in Electronic Properties at Epitaxial Oxide Heterostructure Interfaces*, Phys. Rev. Lett. **105**, 087204 (2010).

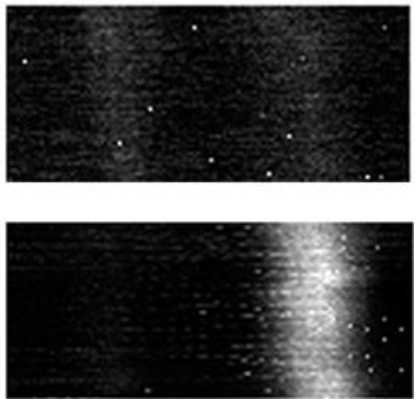
Fit coefficient map **residuals map** **χ2 map**

"Chemistry":
35 to 125 eV

"Plasmons"
5 to 35 eV







Eigenvectors and loadings

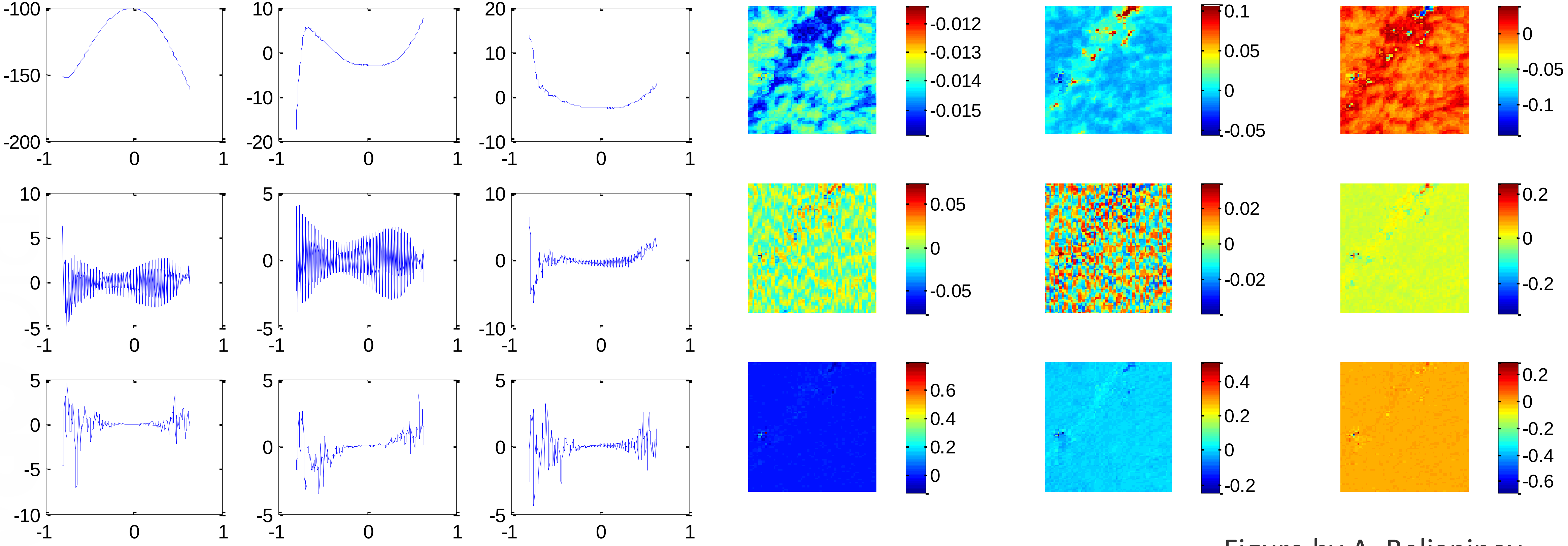


Figure by A. Belianinov

Scree plot and correlations

- Semi log plot indicating the “weight” of each component as a function of all components
- Only the first few components contain useful info, while others are dominated by noise

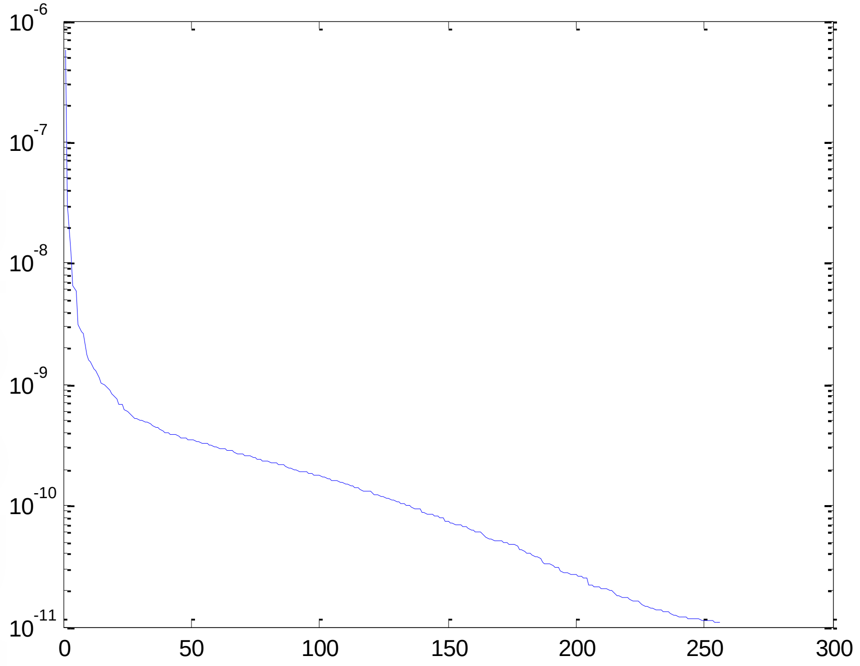
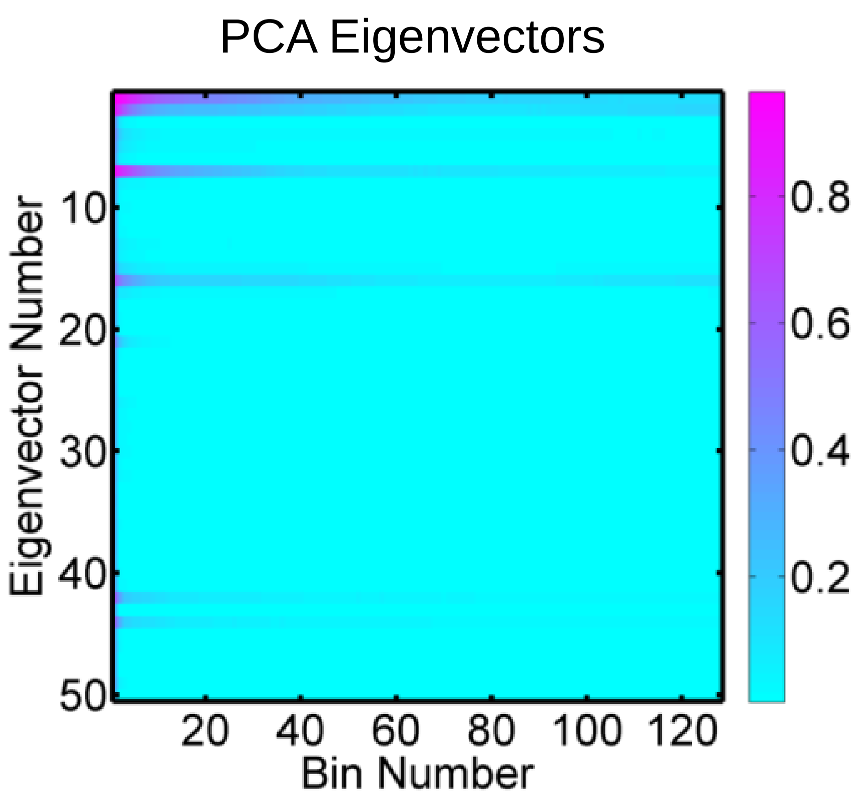


Figure by A. Belianinov

- We can also analyze correlations in images

For AFM data

Spatial correlations

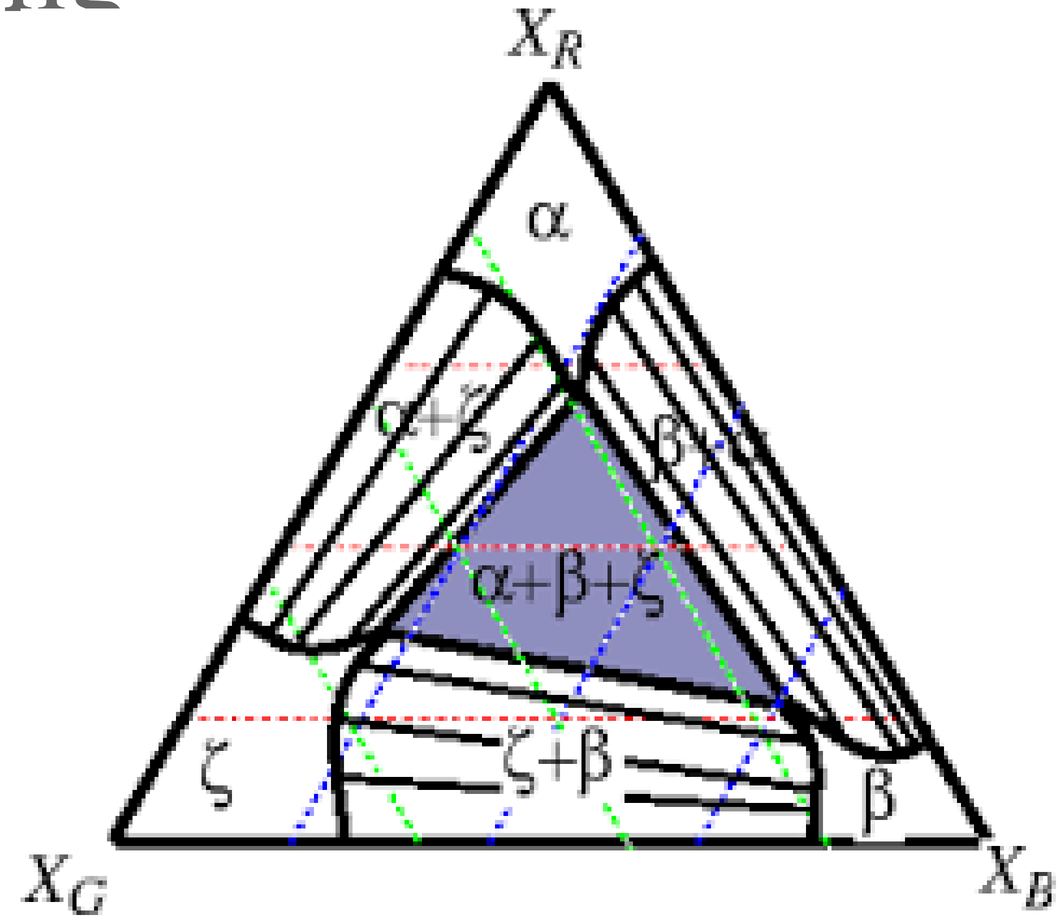


Bayesian Linear Unmixing

$$S(\mathbf{x},\mathbf{R}) = \sum_{i=1}^{K} a_i(\mathbf{x}) w_i(\mathbf{R}) + \mathrm{N}$$

$$\sum_{i=1}^{K} a_i(\mathbf{x}) = 1$$

- The eigenvectors $w_i(\mathbf{R})$ are non-negative, $w_i(\mathbf{R}) \geq 0$
- The loading coefficients sum to 1
- The number of eigenvectors, K, is a priori unknown



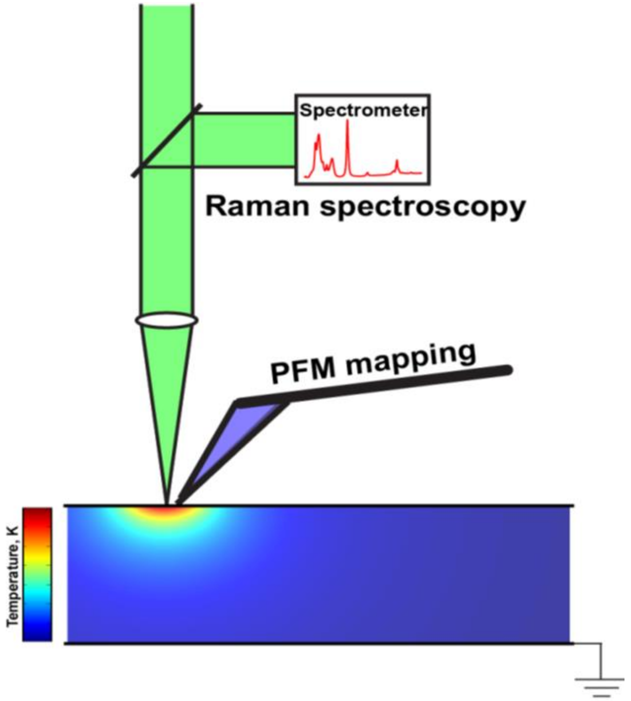
BLU is ideally suited for certain classes of problems, e.g. conduction through parallel channels, optical or electronic spectra of mixtures, etc

Laser heating induced phase transitions

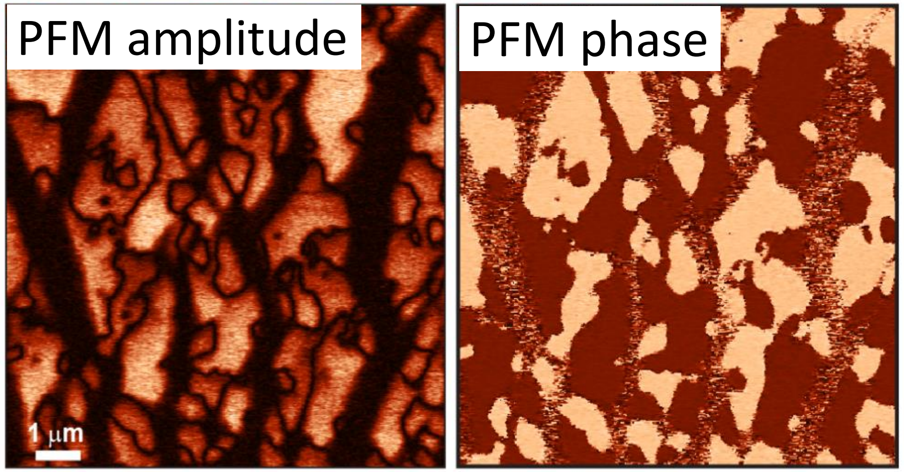
- Copper indium thiophosphate ($Cu_{0.77}In_{1.12}P_2S_6$) layered ferroelectric
  - Ferroelectric state at room temperature
  - Curie temperature T_c = 320 K
  - Non-polar $In_{4/3}P_2S_6$ inclusions

- Combined Atomic Force Microscopy (AFM) and confocal Raman spectroscopy investigative approach
  - AFM – topography measurements
  - Piezoresponse force microscopy (PFM) – static ferroelectric domain structure
  - Raman – crystallographic structure via Raman spectra

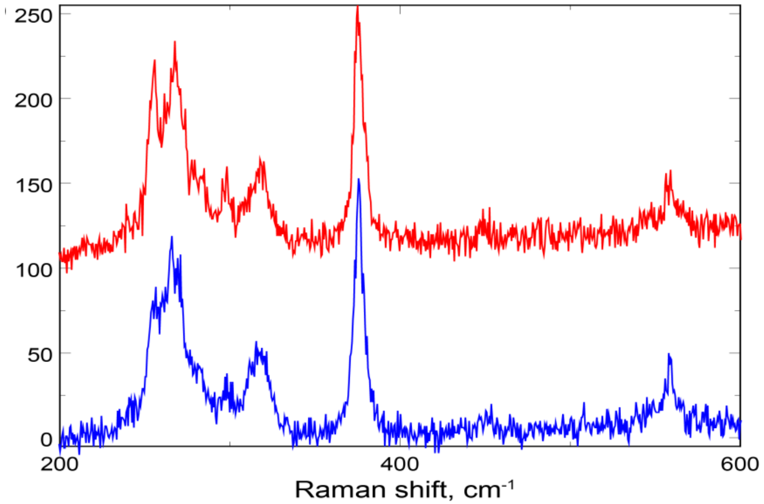
Experimental scheme



Ferroelectric domain structure



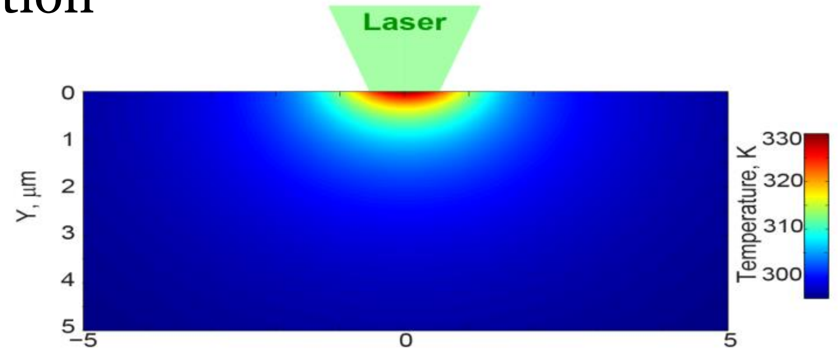
Single point Raman spectra

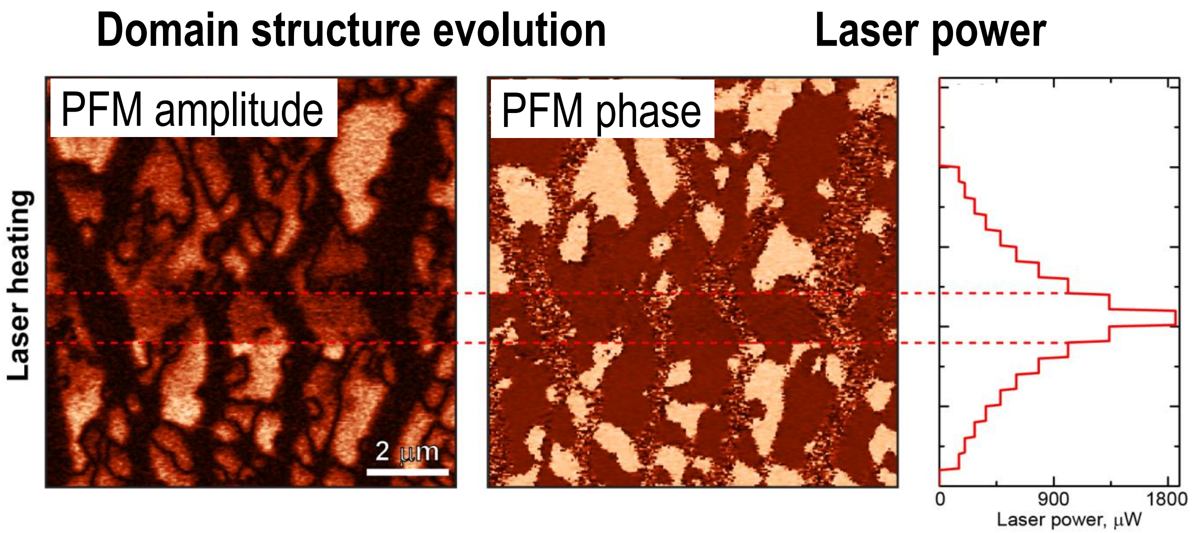


A. Ievlev, ACS Nano **9**, 12442 (2015).

Laser heating induced phase transition

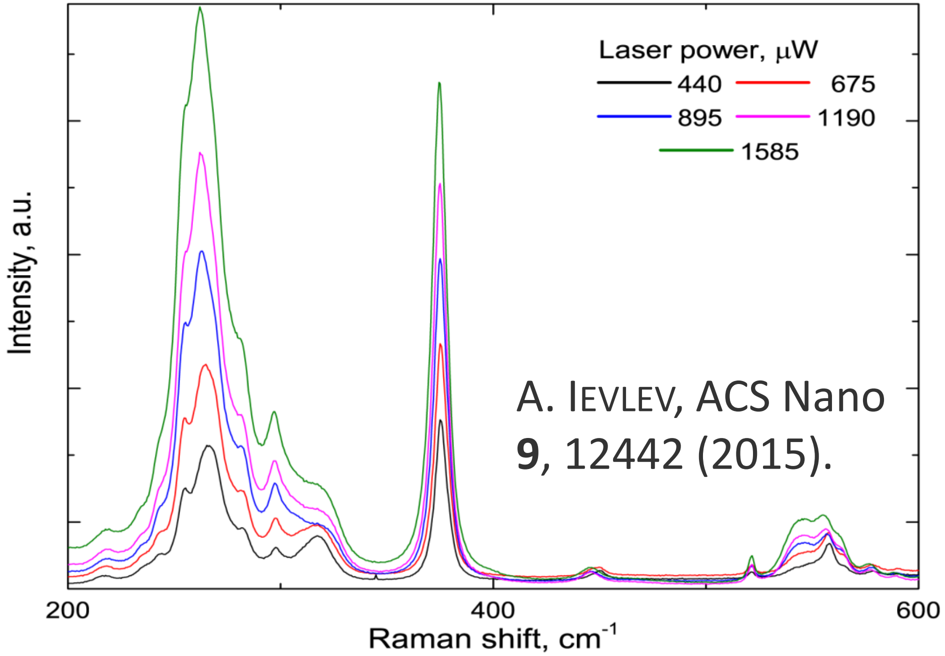
Laser can be used for local heating to induce ferroelectric- paraelectric phase transition





- Measurements with variation of the laser power
  - PFM – *in-situ* change in the domain structure above T_c
  - Raman – evolution of the Raman spectra through the phase transition
- Comprehensive analysis of Raman spectra is complicated by inhomogeneous chemical composition and high noise level
- Bayesian Linear Unmixing can be used for automated identification of spectra evolution

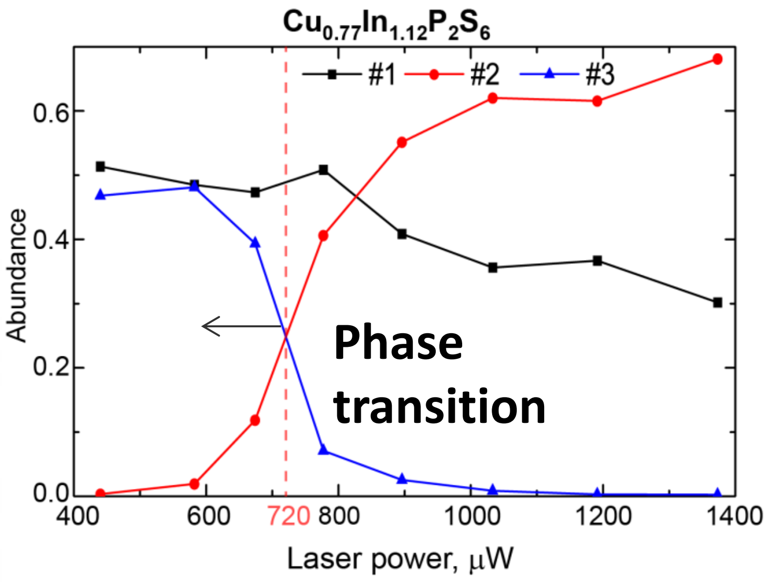
Raman spectra evolution (averaged)



A. Ievlev, ACS Nano **9**, 12442 (2015).

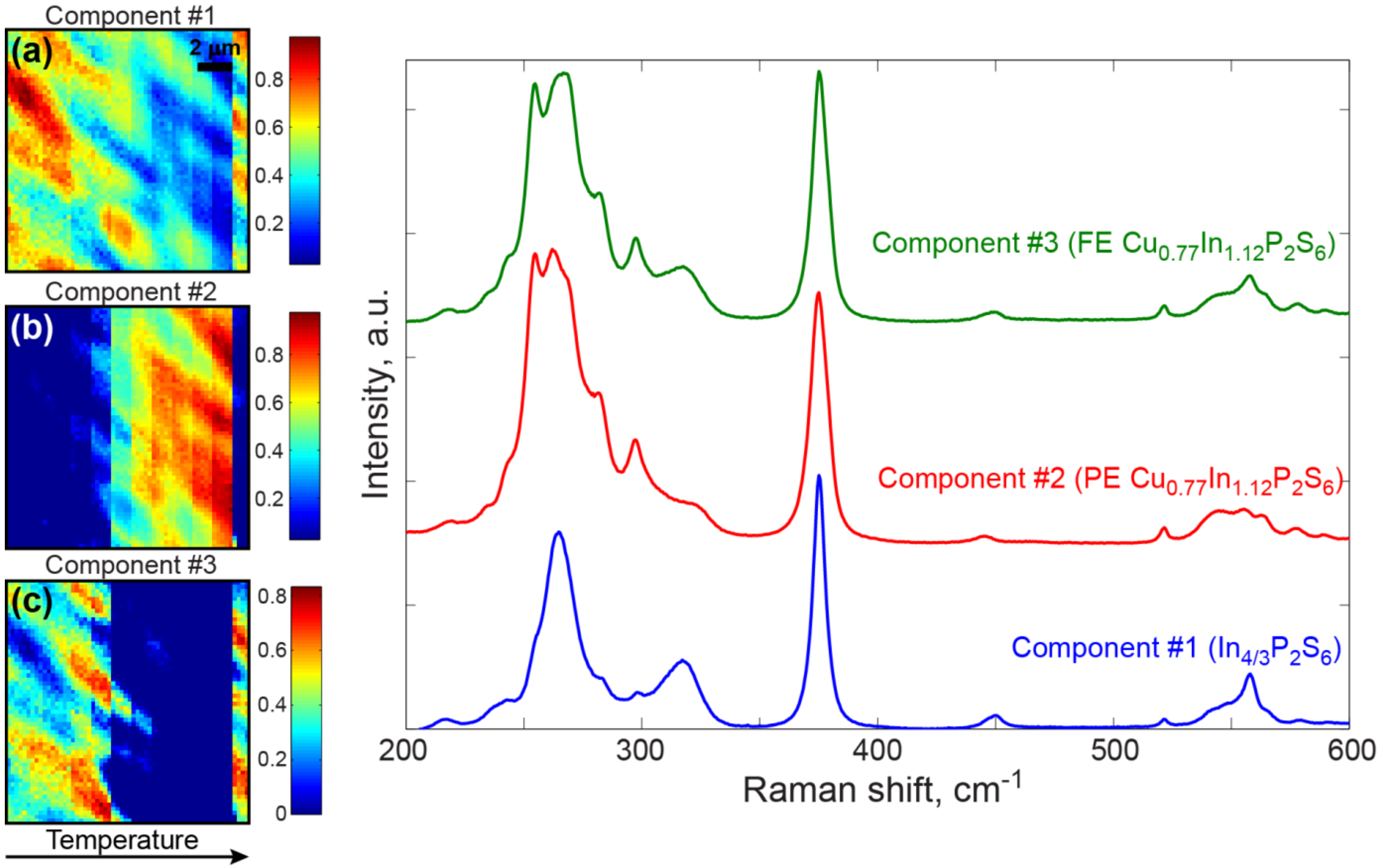
BLU separation of components

Spatial concentration of components



A. IEVLEV, ACS Nano **9**, 12442 (2015).

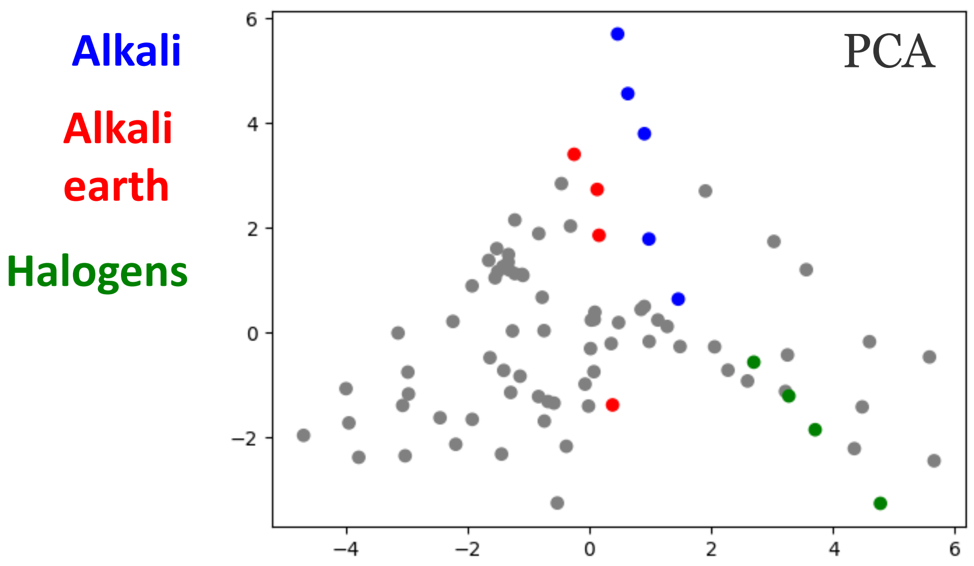
Results of BLU: components and loading maps

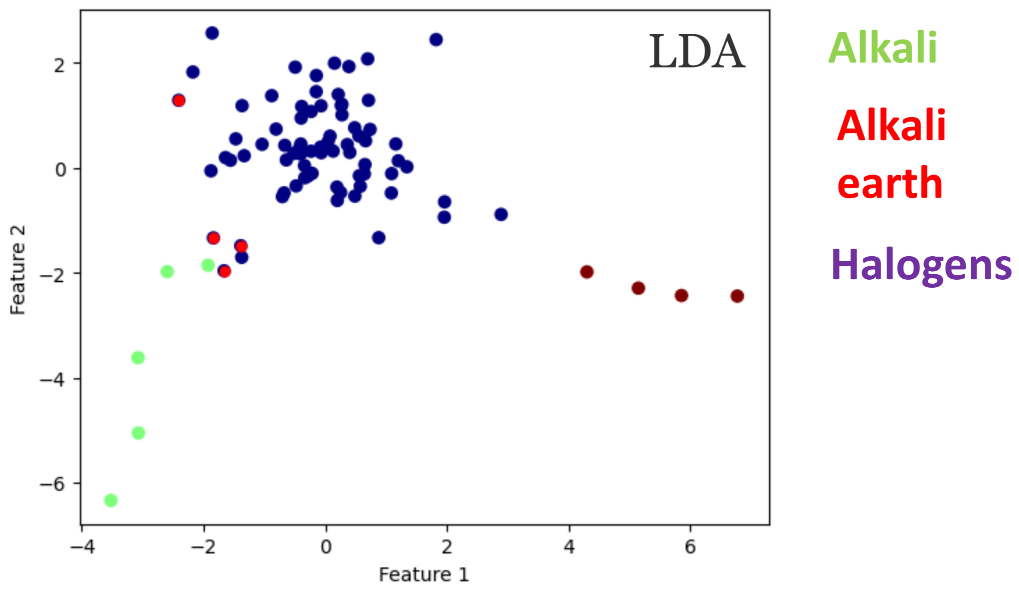


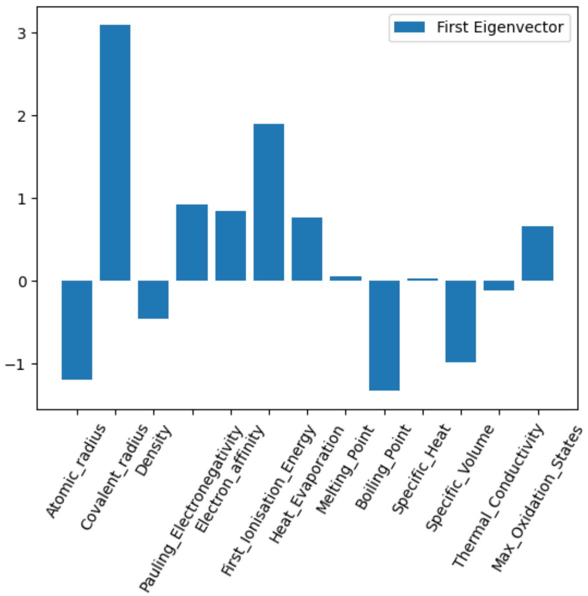
Unmixing showed presence of three independent components in Raman spectra:

1. Non-polar $In_{4/3}P_2S_6$ – weak changes in intensity with temperature
2. Paraelectric $CuInP_2S_6$ above T_c – appears at higher laser powers
3. Ferroelectric $CuInP_2S_6$ below T_c – disappears at higher temperatures

PCA vs. LDA for elements

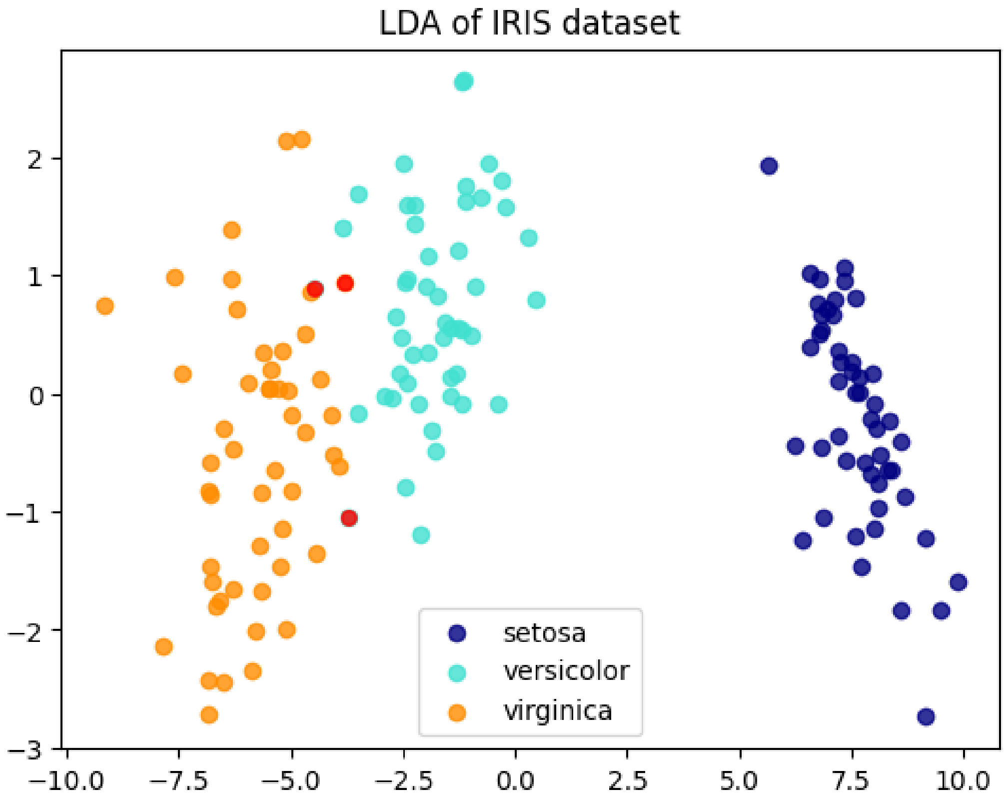




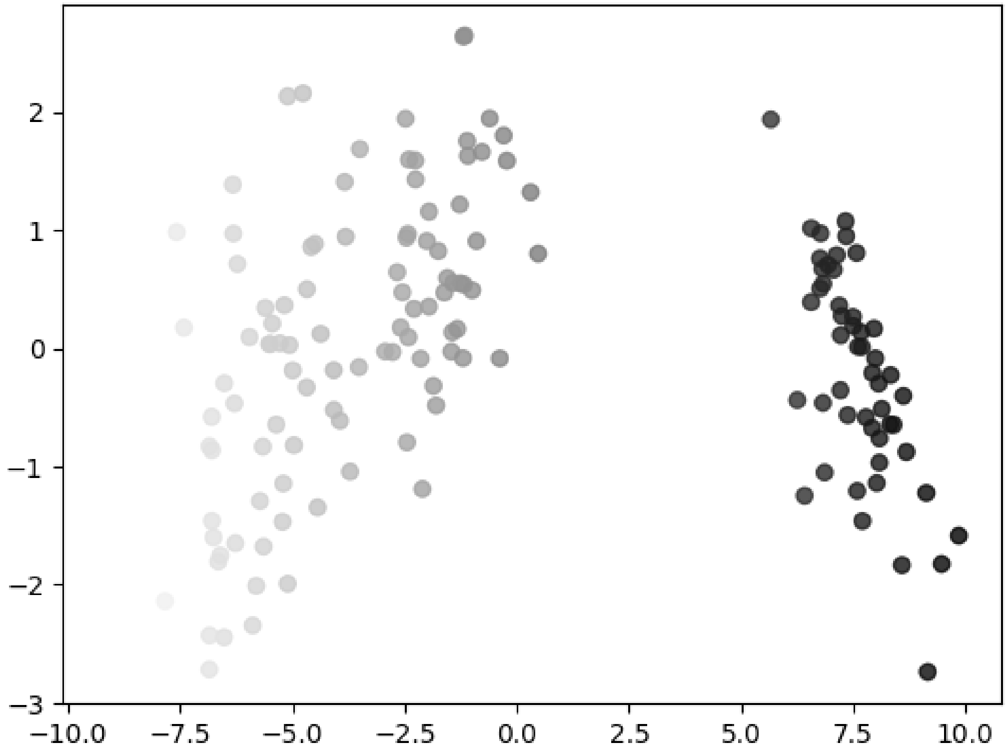


- Our element space is described by 13 descriptors
- In PCA, we found 2 linear combinations of these descriptors that describe this data set best.
- Alkali, alkali-earth, and halogens are close to each other in PCA space
- LDA finds best representation to separate alkali and halogens from everything else
- In LDA representation, alkali earth are close to alkali

What else can LDA give us?

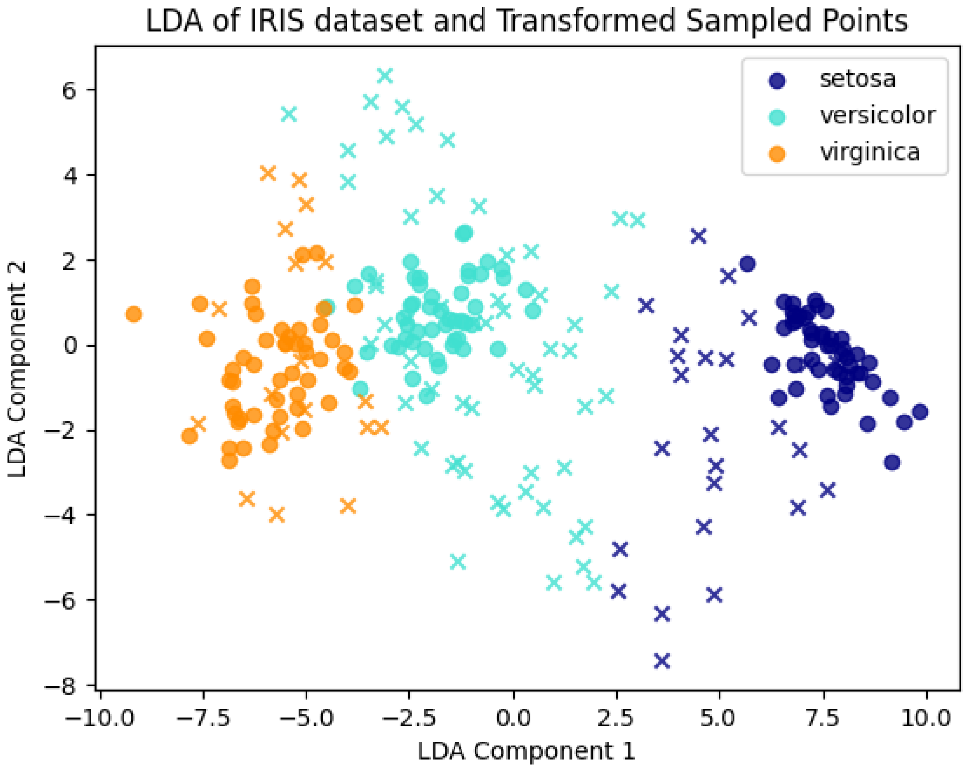


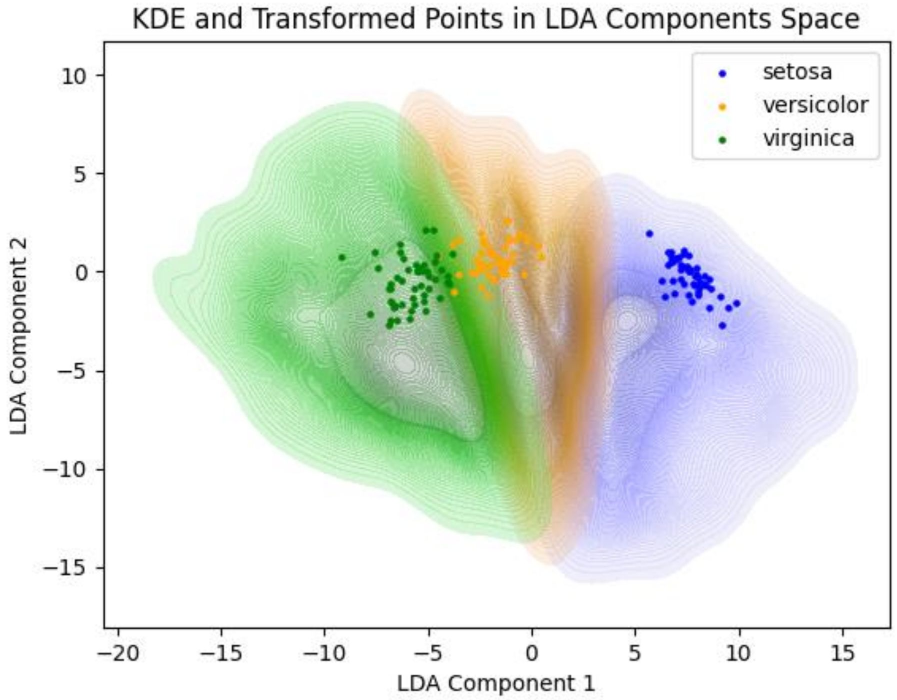
Decision function



- Decision function allows us to quantify how likely is the feature to belong to certain class

Visualizing the decision surfaces





- Generate multiple points uniformly distributed in the original high-dimensional space
- Perform the LDA transform
- Calculate the KDE