THE UNIVERSITY of TENNESSEE KNOXVILLE

Lecture 18: Linear Methods for Spectroscopy and Imaging

Instructor: Sergei V. Kalinin

12_CCA_Elements.ipynb

The catch

Feature 1	Feature 2	Feature 3	Feature 4	Feature 5	…	Feature N

Feature 1	Feature 2	Feature 3	Feature 4	Feature 5	…	Feature M

Independent Component Analysis

- **PCA:** orthogonal transformation of possibly correlated variable into a set of linearly uncorrelated variables
    - Analyzes data representing observations described by dependent variables which are inter-correlated
    - Main goal is to find true variables assuming that corrupting noise is Gaussian

- **ICA:** method for separating multivariate signal into additive subcomponents assuming statistical independence of source signals
    - Finds components that are maximally independent and non-Gaussian
    - Blind source separation – cocktail party problem

Compared to PCA, ICA can produce statistically independent non-Gaussian components by decorrelating the higher-order moments in addition to the first- and second-order moments of the statistical distribution

Mathematics of ICA

The Goal is to transform observed data into maximally independent components measured through some function $F(s_1, \ldots, s_n)$ of independence.

- Data as a set of vectors $\longrightarrow$ $x = (x_1, \ldots, x_m)^T$
- Components of the data $\longrightarrow$ $s = (s_1, \ldots, s_n)^T$.

$$s = Wx$$

An observed data vector *x* can then be represented as a sum of independent components *s* weighted by some mixing weight *a*:

$$x_i = a_{i,1}s_1 + \ldots + a_{i,k}s_k + \ldots + a_{i,n}s_n \quad \text{or} \quad x = \sum_{k=1}^{n} s_k a_k$$

In other words, data vector *x* is represented by basis vectors $a_k = (a_{1,k}, \ldots, a_{m,k})^T$ that can form columns of a mixing matrix such that: $x = As$

https://en.wikipedia.org/wiki/Independent_component_analysis

Mathematics of ICA - 2

$$x = As$$

Given that our data is set of vectors x, we want to find both, the mixing matrix A and sources s

This can be done by calculating w vectors and a cost function that can maximize the Non-Gaussianity, or minimize mutual information of

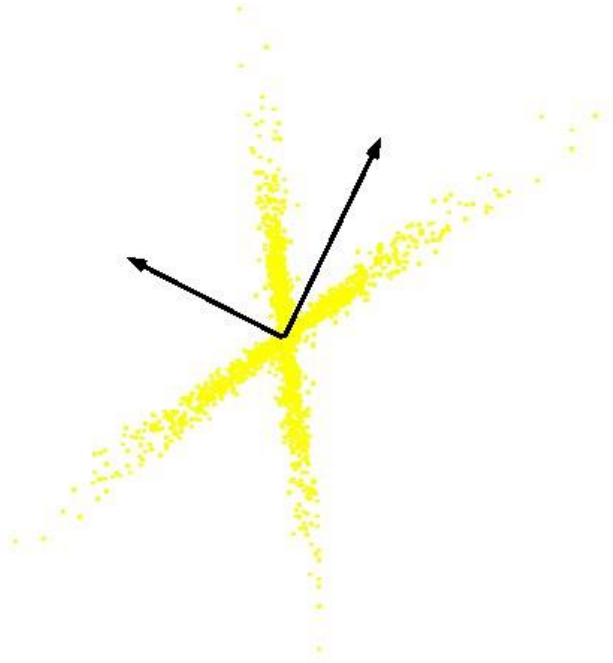
$$s_k = (w^T * x)$$

Original sources can then be recovered by multiplying observed data vectors x with the inverse of the mixing matrix: $W = A^{-1}$

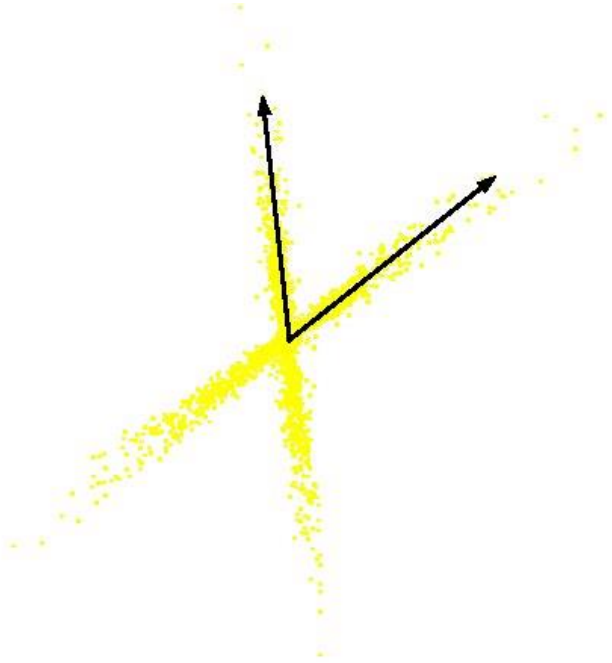
- Not as easy to utilize as PCA, but excellent premade algorithms are readily available, e.g. Aapo Hyvärinen – FastICA*

http://cis.legacy.ics.tkk.fi/aapo/papers/IJCNN99_tutorialweb/IJCNN99_tutorial3.html
http://research.ics.aalto.fi/ica/fastica/

PCA vs. ICA

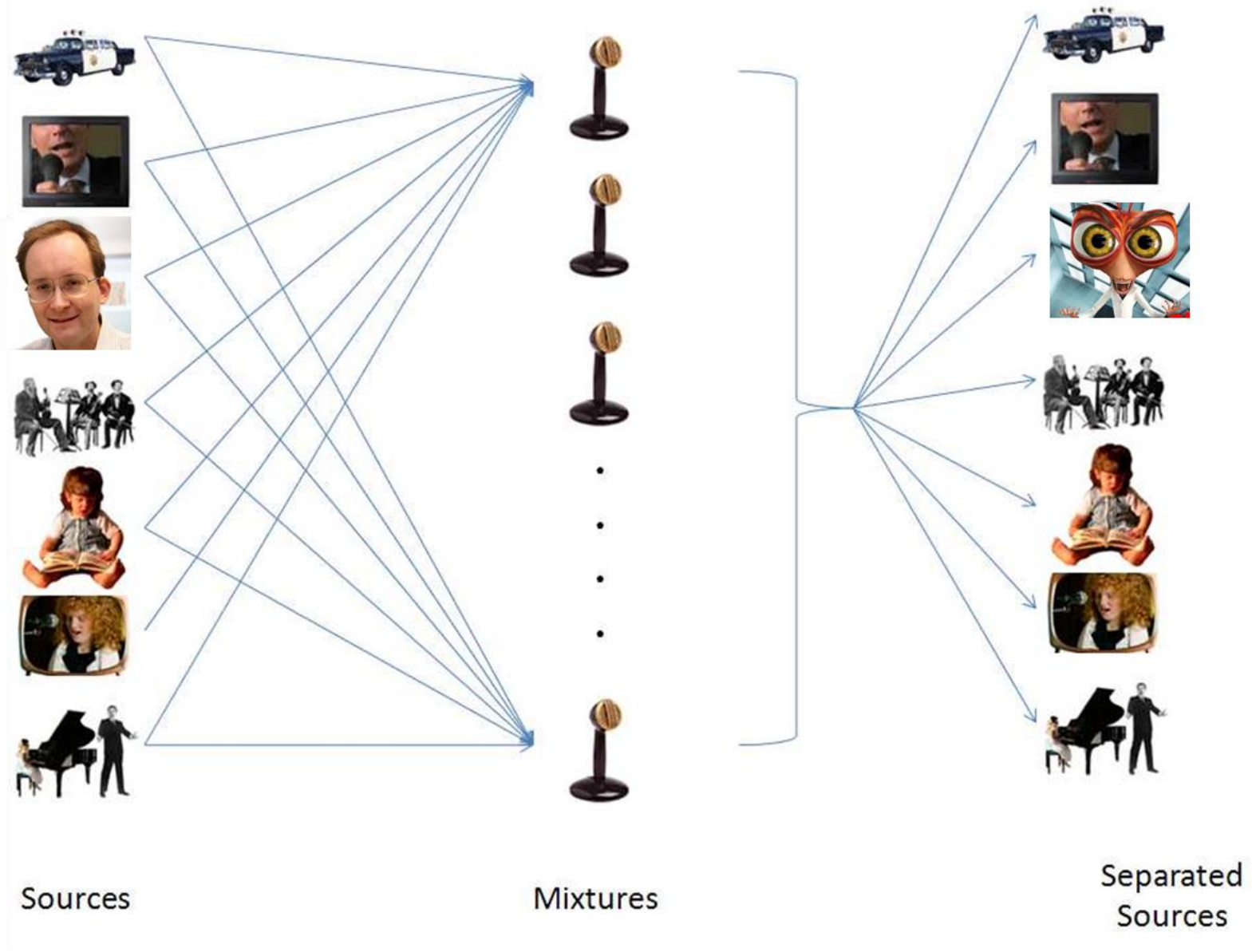


PCA
(orthogonal coordinate)



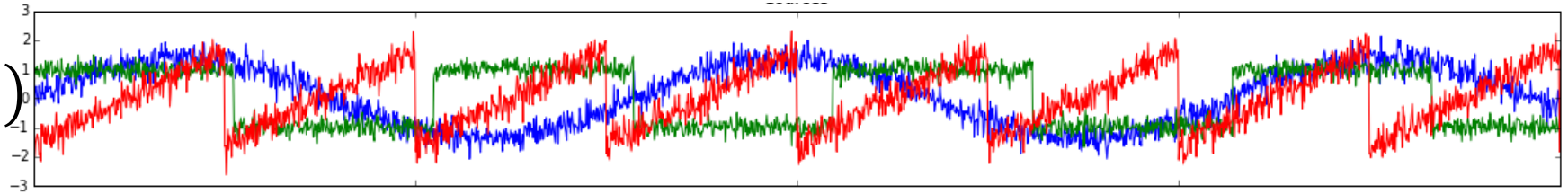
ICA
(non-orthogonal coordinate)

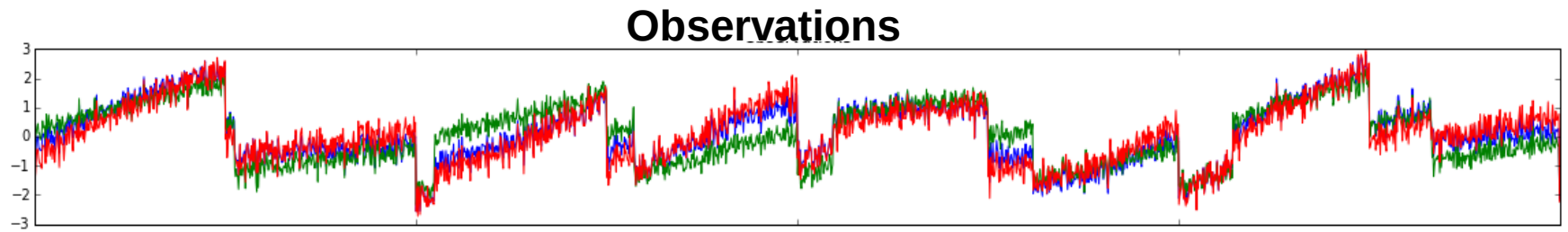
Cocktail Party Problem



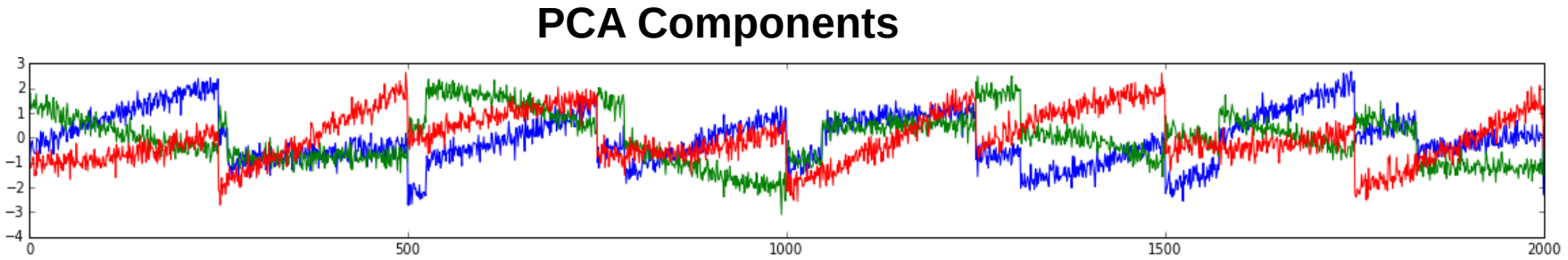
ICA Example

$$\left(R_1(t), R_2(t), \ldots R_n(t)\right)$$

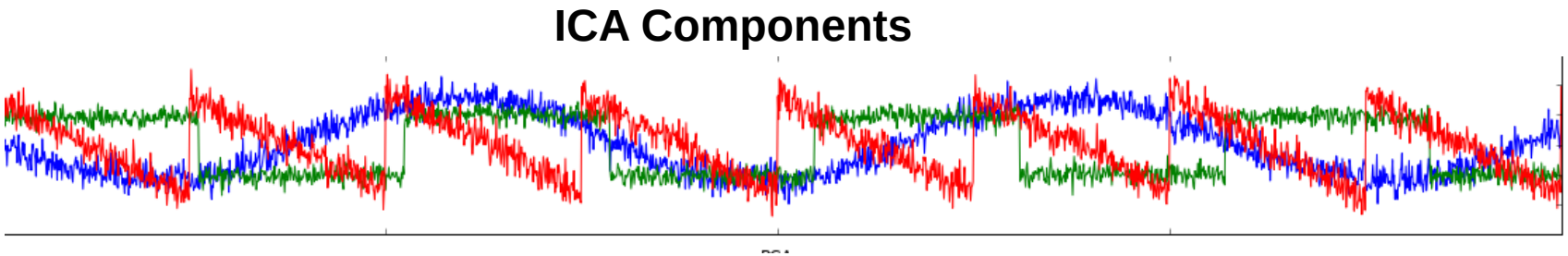




Components are maximally independent



$$\begin{pmatrix} R_1(t) \\ \ldots \\ R_n(t) \end{pmatrix} = \boldsymbol{A} \begin{pmatrix} s_1(t) \\ \ldots \\ s_n(t) \end{pmatrix}$$



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

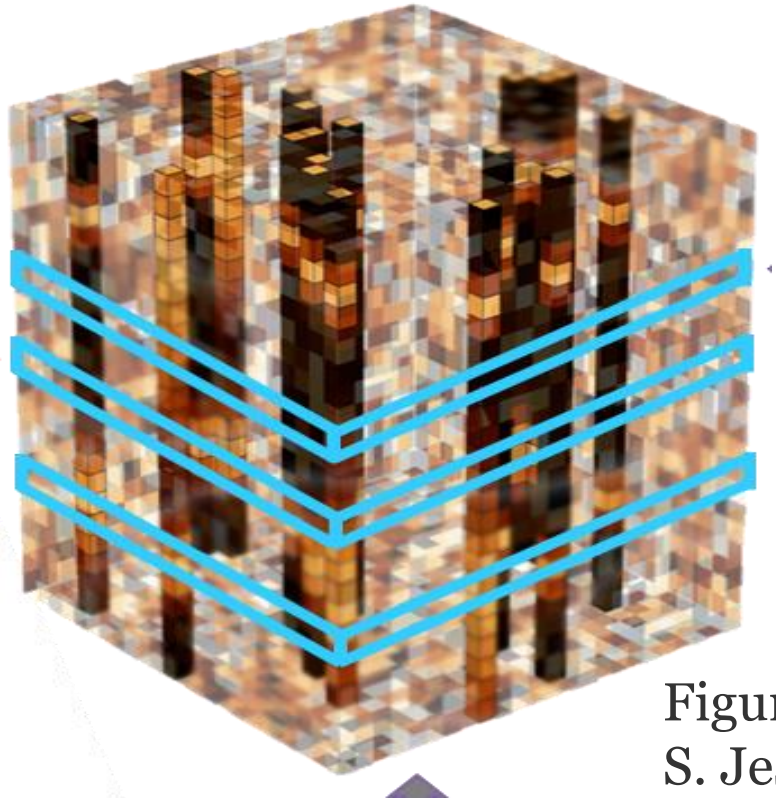
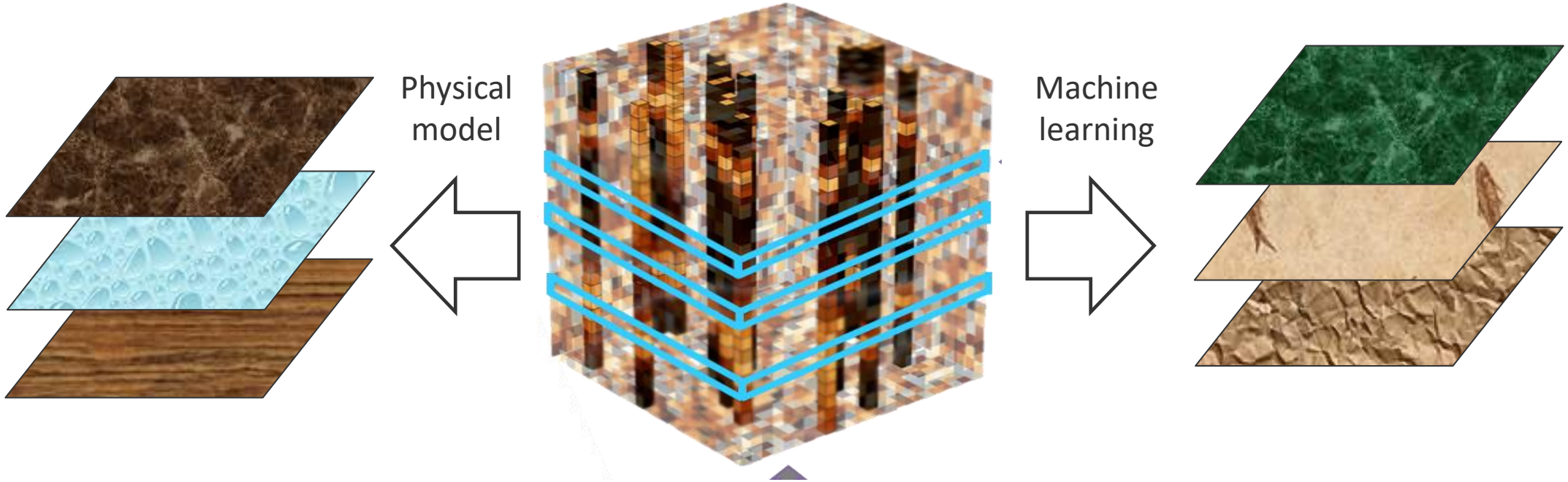


Figure by S. Jesse

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

Physics vs. ML based analysis



- If we have physical model, we can extract relevant parameters from data
- Imperfect model: epistemic uncertainty
- Noisy data: aleatoric uncertainty
- Analysis results do not depend on sampling of data in x,y

- If we don't have physical model, we can learn intrinsic structure of data
- **Unsupervised learning:** based on data only
  - But not really (definition of distance)
  - Analysis can depend on sampling of data
- **Supervised learning:** based on prior examples
  - Out of distribution shifts

Physics-informed ML: Combines strengths (and limitations) of both

General linear unmixing

$$S(\mathbf{x}, \mathbf{R}) = \sum_i a_i(\mathbf{x})\, w_i(\mathbf{R}) + \mathrm{N}$$

We start with:

- **x** is the spatial variable, x = (x,y)
- R is the (vector) parameter variable

Overall, for MxM image and P point in spectra, we have M^2P data points

We aim to get:

- $a_i(\mathbf{x})$ are loading maps
- $w_i(\mathbf{R})$ are endmembers/eigenvectors
- N is noise

Overall, we can have (maximum) P loading maps of M^2 size. However, not all of them will have useful information

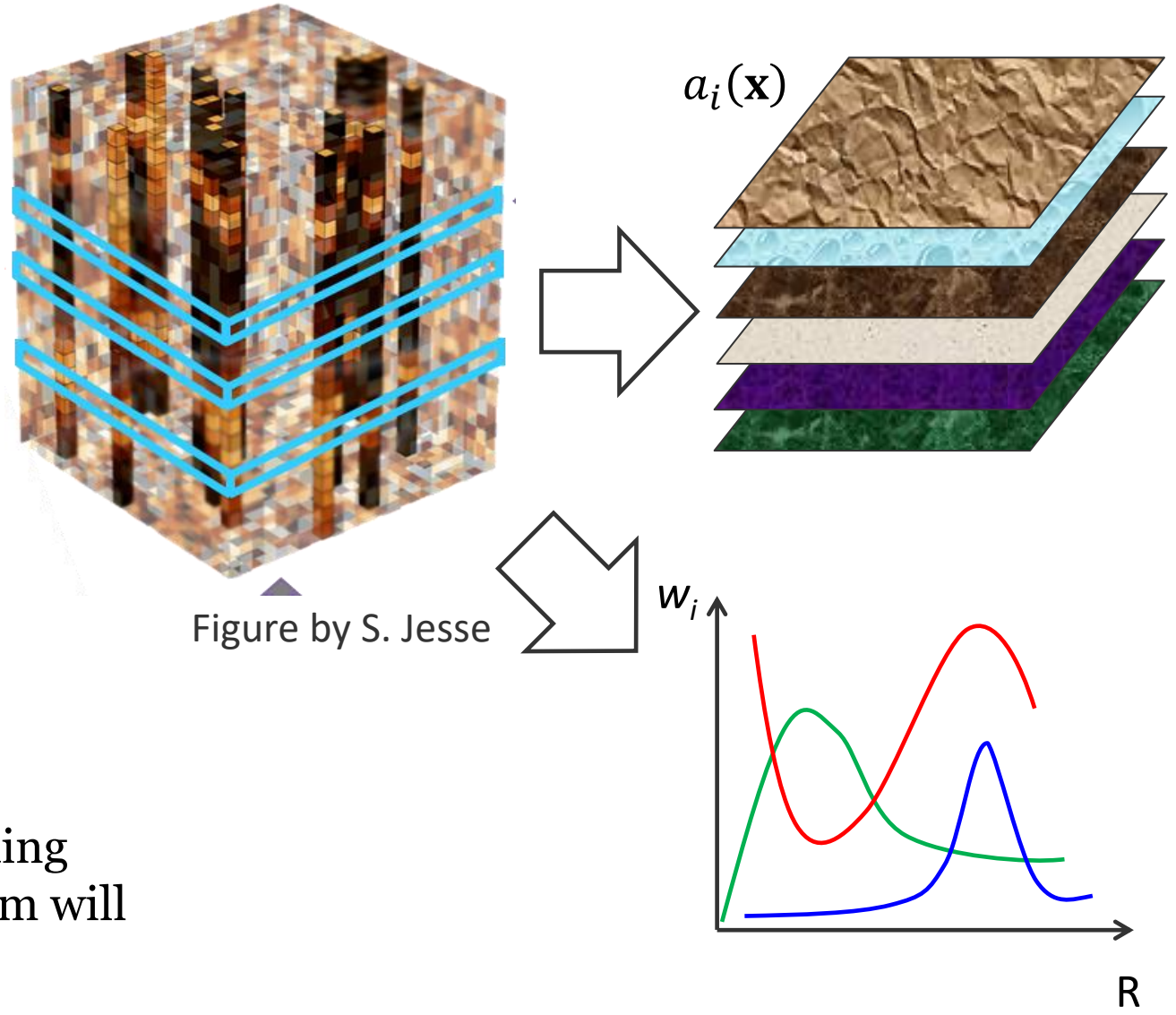
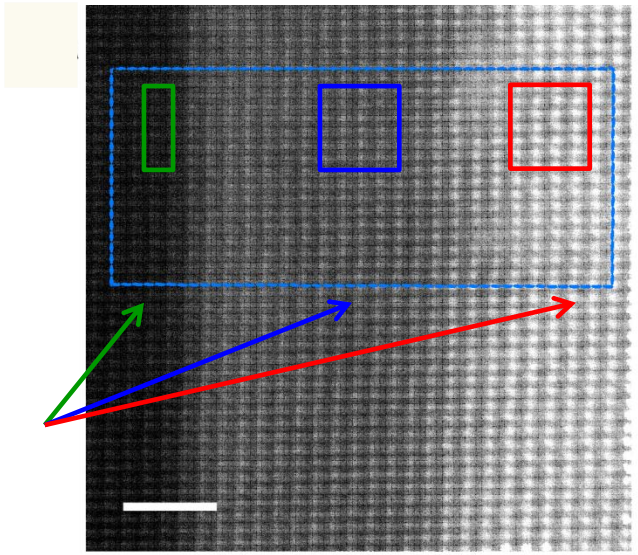


Figure by S. Jesse

Multiple Linear Regression

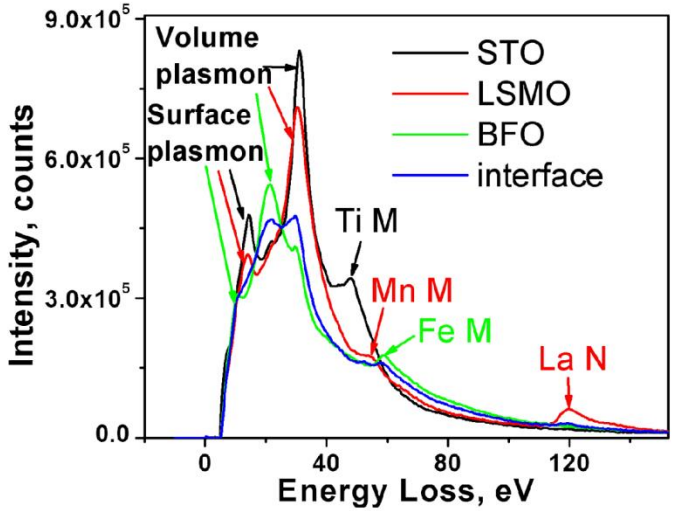
Linear mixing $S(\mathbf{x}, \mathbf{R}) = \sum_i a_i(\mathbf{x}) w_i(\mathbf{R}) + N$ but $w_i(\mathbf{R})$ are **known**

STEM of STO/LSMO/BFO interface



Standard spectra

Low-loss EELS spectra of three components

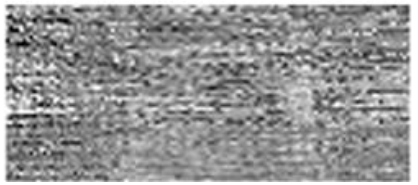


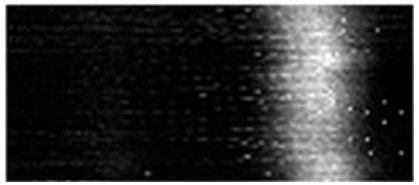
A.Y. Borisevich et al, *Suppression of Octahedral Tilts and Associated Changes in Electronic Properties at Epitaxial Oxide Heterostructure Interfaces*, Phys. Rev. Lett. **105**, 087204 (2010).

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

"Chemistry": 35 to 125 eV

"Plasmons" 5 to 35 eV





Principal Component Analysis

$$S(\mathbf{x}, \mathbf{R}) = \sum_i a_i(\mathbf{x})\, w_i(\mathbf{R})$$

- In PCA, the eigenvectors $w_i(\mathbf{R})$ are orthonormal and are arranged such that corresponding eigenvalues are placed in descending order by variance
- Can be used to separate “real data” from “noise” – but needs cut-off/selection criteria
- PCA eigenvectors generally do not have defined physical meaning
- PCA is a starting point for many other unmixing methods

Ultramicroscopy 34 (1990) 165–178
North-Holland

EELS elemental mapping with unconventional methods I. Theoretical basis: image analysis with multivariate statistics and entropy concepts

Pierre Trebbia *
Laboratoire de Physique des Solides, Bâtiment 510, F-91405 Orsay Cedex, France

and

Noël Bonnet
Unité INSERM 314 et Université de Reims, 21 rue Clément Ader, F-51100 Reims, France

Received 7 June 1990

Electron energy loss filtered images recorded within a transmission analytical electron microscope are now widely used for the mapping of the elemental distribution of a given atomic species in a specimen prepared as a thin film. Such an image processing may produce both valuable results and artifacts if a careful inspection of all the hypotheses needed by the calculation is not carried out. This paper presents some general statistical methods for a contrast information analysis of a noisy image data set. After a brief introduction of different concepts such as contrast, variance, information and entropy, two unconventional approaches for image analysis are explained: the relative entropy computed with respect to a pure random and signal-free image and the factorial analysis of correspondence (a branch of multivariate statistics). In the companion article (part II), these concepts are applied to real experiments and the results compared with those obtained with a conventional method. Although electron energy loss spectroscopy is the only technique considered here, these methods for image analysis can be applied to a wide variety of noisy data sets (spectra, images, ...) recorded from various sources (electrons, photons, ...).

Why historical papers matter:

1. Often contain elementary introductions
2. Deep insights into principles
3. Surprisingly prescient predictions
4. Comparison with the present: see the big picture

““Those who cannot remember the past are condemned to repeat it.”

George Santayana,
The Life of Reason, 1905

Available online at www.sciencedirect.com
ScienceDirect
Ultramicroscopy 106 (2006) 1024–1032
ultramicroscopy
www.elsevier.com/locate/ultramic

Mapping chemical and bonding information using multivariate analysis of electron energy-loss spectrum images

M. Bosman[a,*], M. Watanabe[b], D.T.L. Alexander[c], V.J. Keast[a]

[a]*Australian Key Centre for Microscopy and Microanalysis, University of Sydney, Sydney, NSW 2006, Australia*
[b]*Department of Materials Science and Engineering, Lehigh University, Bethlehem, PA 18015, USA*
[c]*Department of Materials Science and Metallurgy, University of Cambridge, Pembroke St. CB2 3QZ, UK*

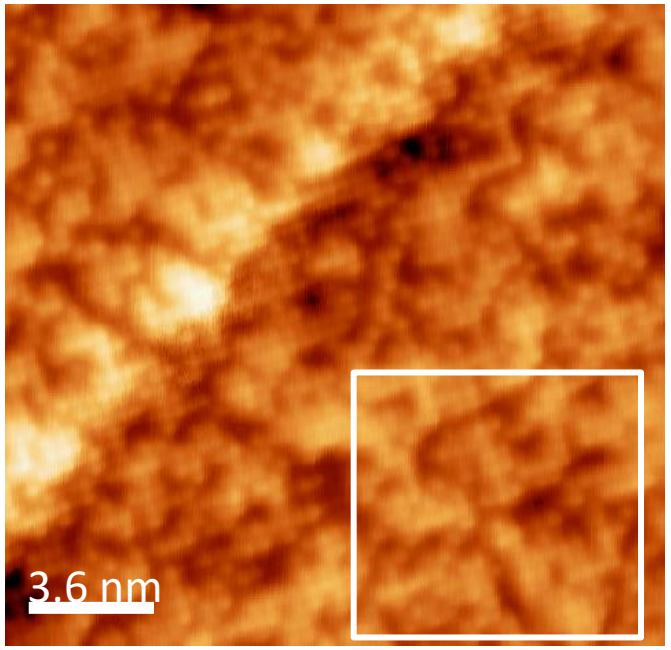
Received 23 June 2005; received in revised form 26 October 2005; accepted 18 April 2006

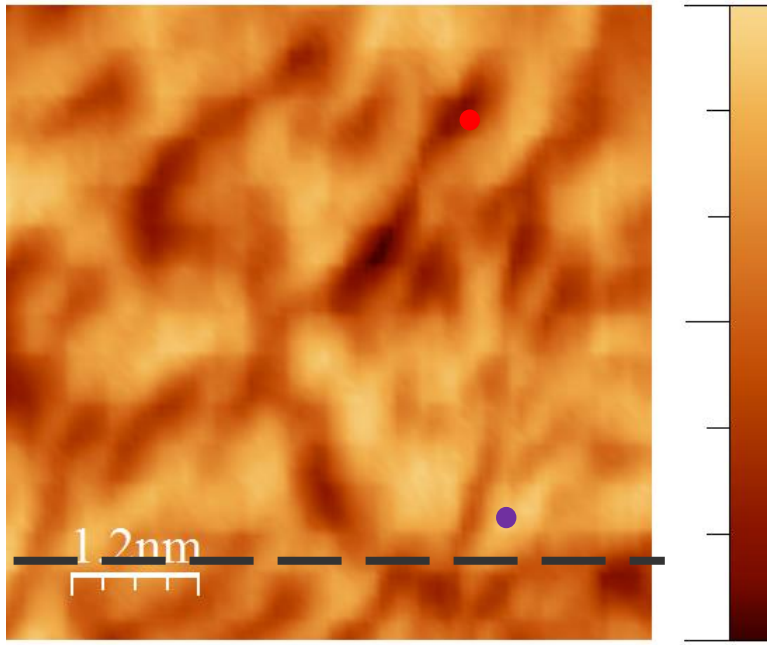
Abstract

Electron energy-loss spectroscopy (EELS) in the transmission electron microscope (TEM) is used to obtain high-resolution information on the composition and the type of chemical bonding of materials. Spectrum imaging, where a full EEL spectrum is acquired and stored at each pixel in the image, gives an exact correlation of spatial and spectral features. However, determining and extracting the important spectral components from the large amount of information contained in a spectrum image (SI) can be difficult. This paper demonstrates that principal component analysis of EEL SIs can be used to extract chemically relevant components. With weighted or two-way scaled principal component analysis, both compositional and bonding information can be extracted. Mapping of the chemical variations in a partially reduced titanium dioxide sample and the orientation-dependent bonding in boron nitride and carbon nanotubes are given as examples.

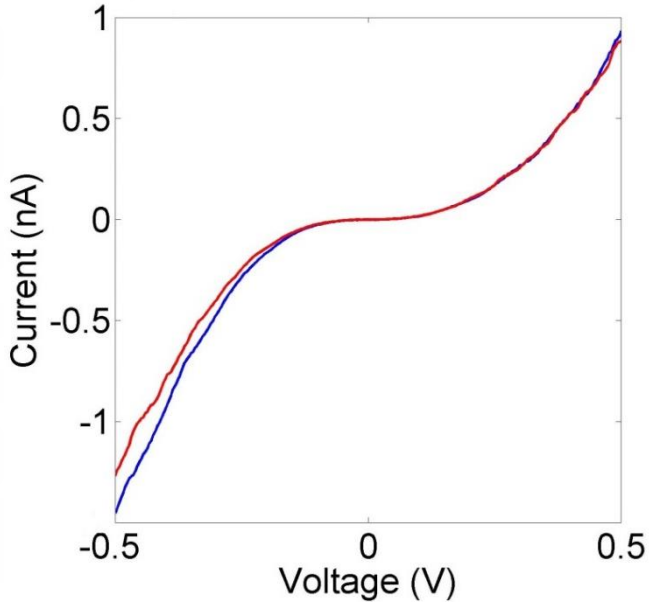
Why did the PCA on EELS started to grow in 2005 – 2010?

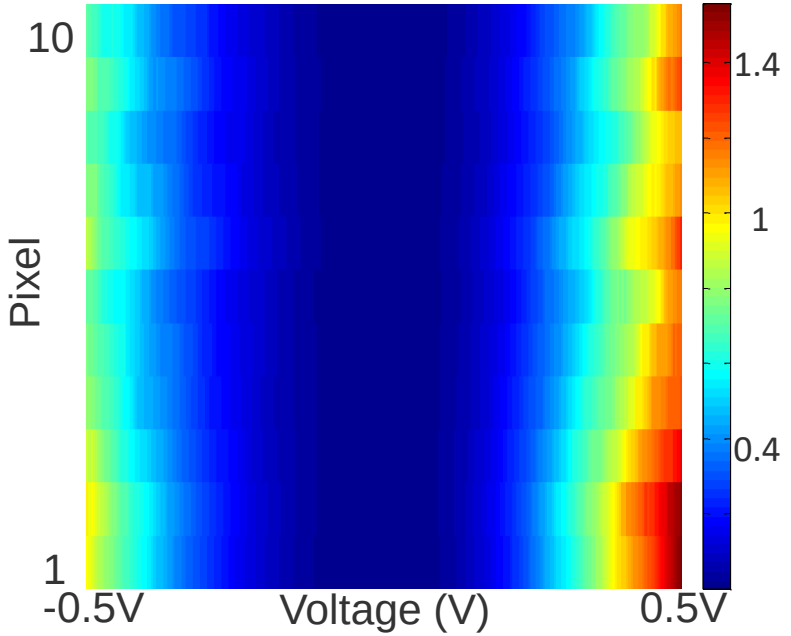
Grain boundary by STM





Topographical STM of FeSeTe, T = 82K 15x15nm^2, 50mV, 100pA, white rectangle represents area where CITS was performed; (b) CITS 80x80 pixel graphical average of the spectrographic data.





M. Ziatdinov, A. Maksov, L. Li, A. Sefat, P. Maksymovych, and S.V. Kalinin, *Deep data mining in a real space: Separation of intertwined electronic responses in a lightly-doped* $BaFe_2As_2$, Nanotechnology **27**, 475706 (2016).

Eigenvectors and loadings

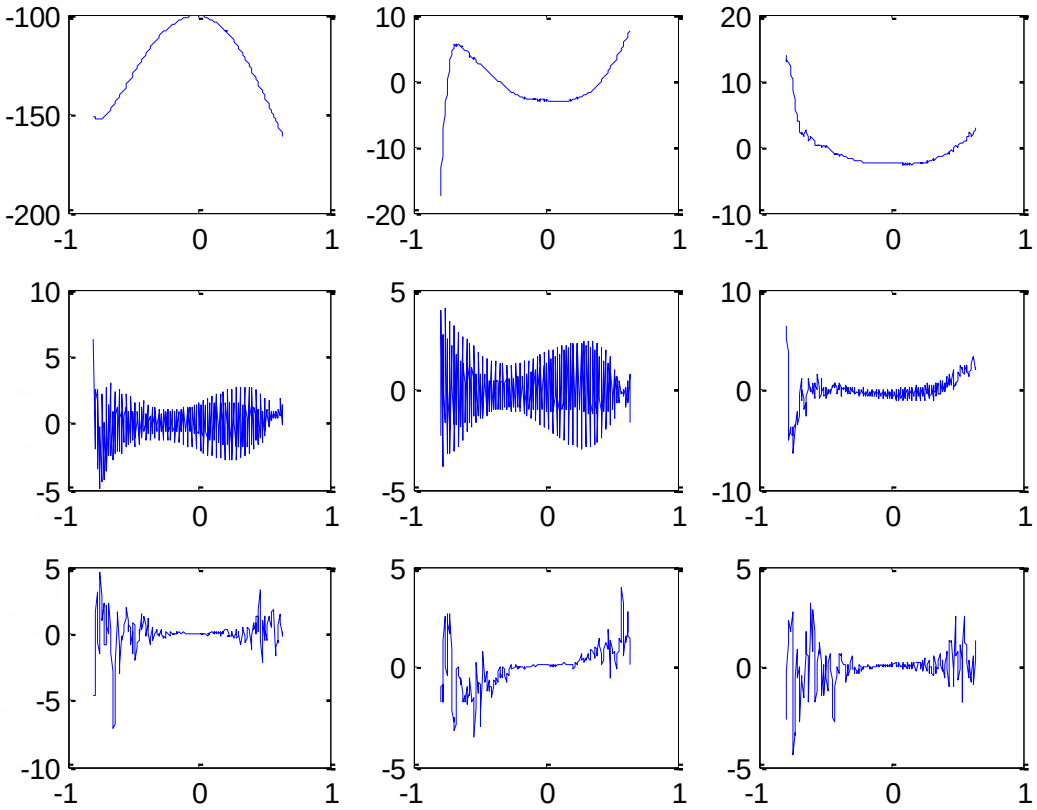


Figure by A. Belianinov

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

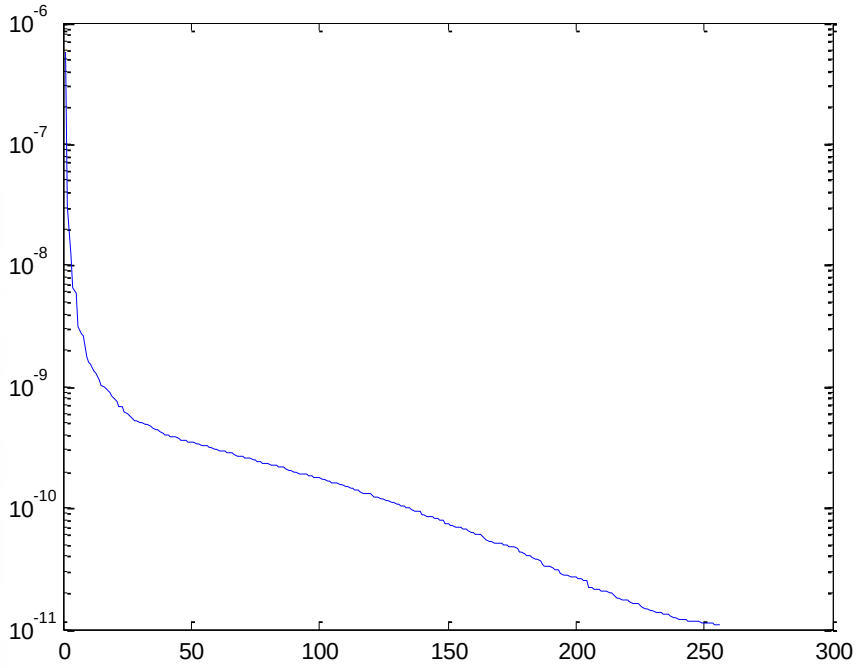
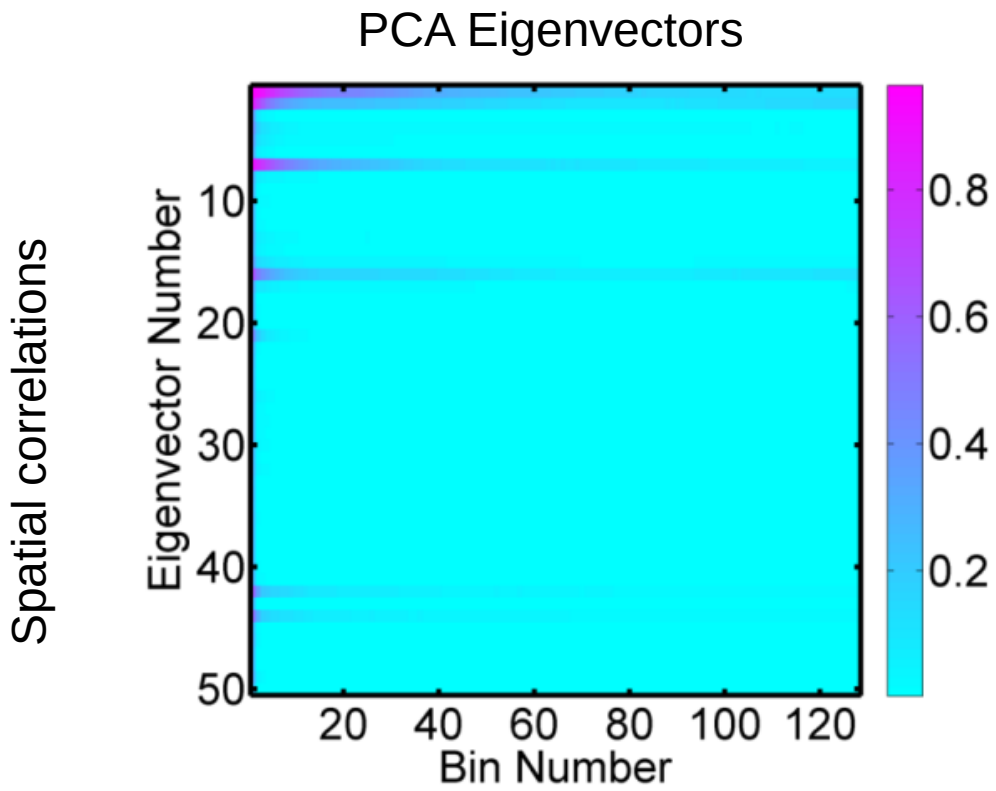


Figure by A. Belianinov

- We can also analyze correlations in images

For AFM data

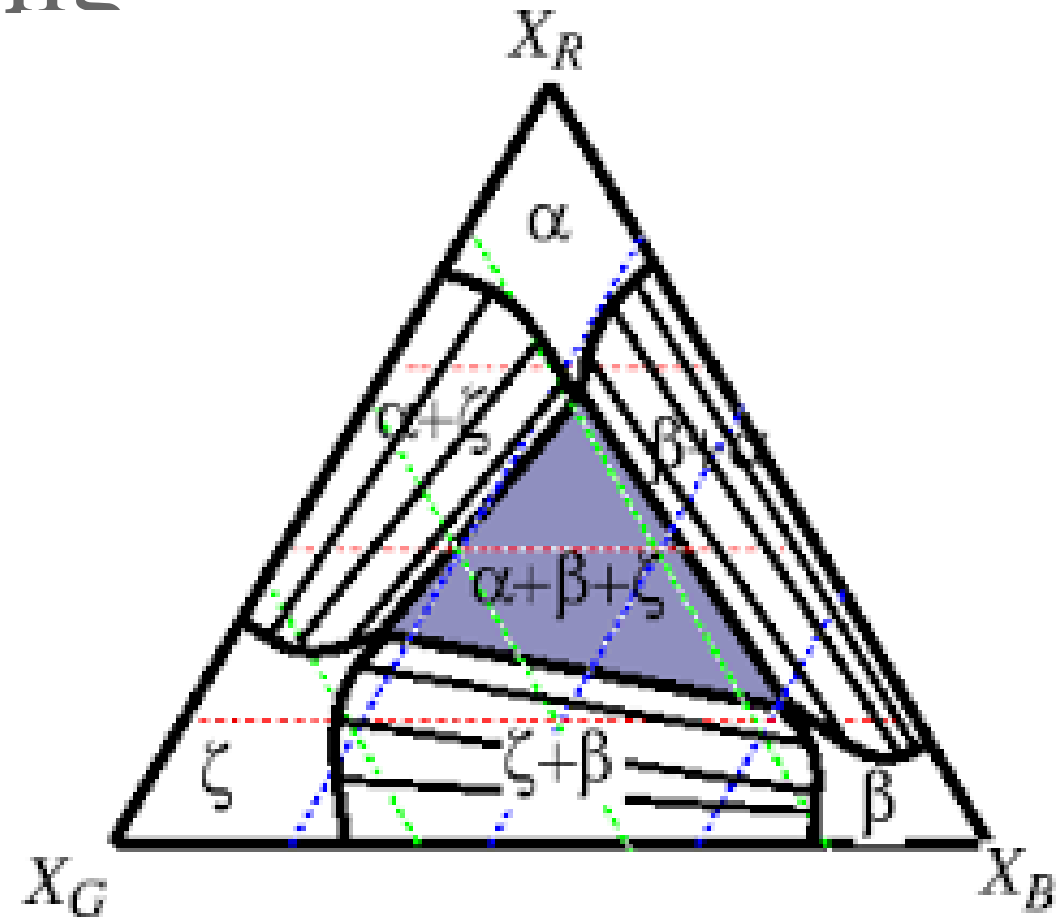


Bayesian Linear Unmixing

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

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

- 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



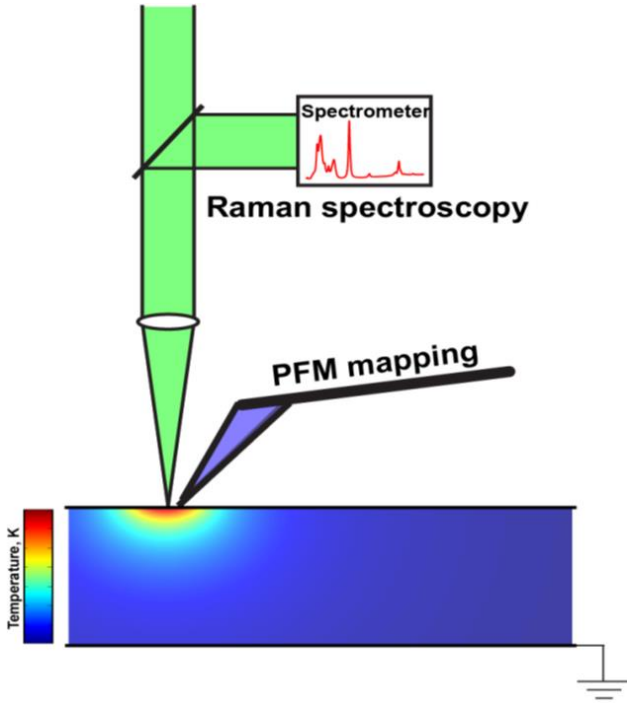
BLU is ideally suited for certain classes of problems, e.g. conduction through parallel channels, optical or electronic spectra of mixtures, etc

Laser heating induced phase transitions

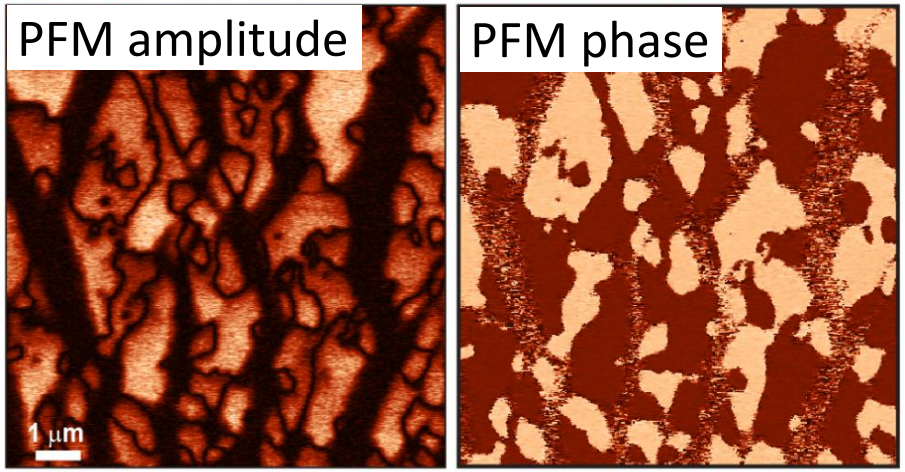
- Copper indium thiophosphate ($Cu_{0.77}In_{1.12}P_2S_6$) layered ferroelectric
  - Ferroelectric state at room temperature
  - Curie temperature T_c = 320 K
  - Non-polar $In_{4/3}P_2S_6$ inclusions

- 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

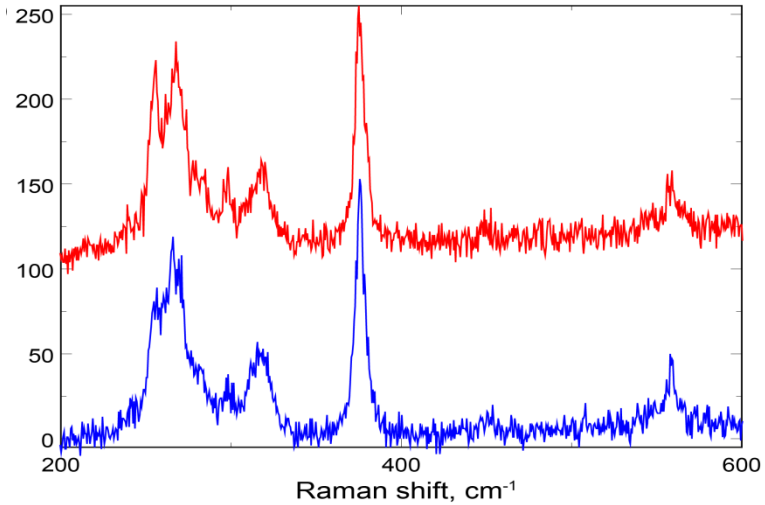
Experimental scheme



Ferroelectric domain structure



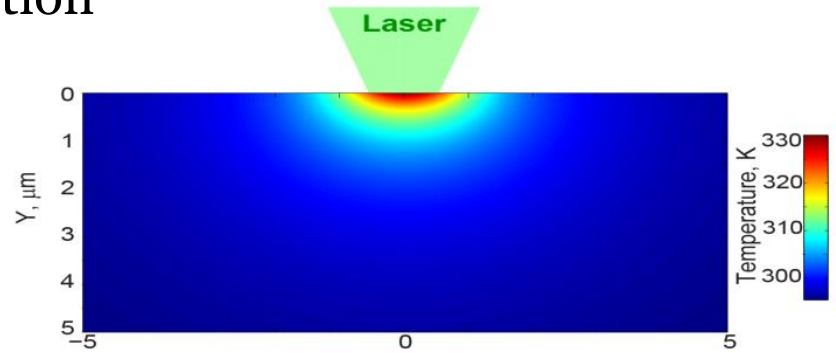
Single point Raman spectra



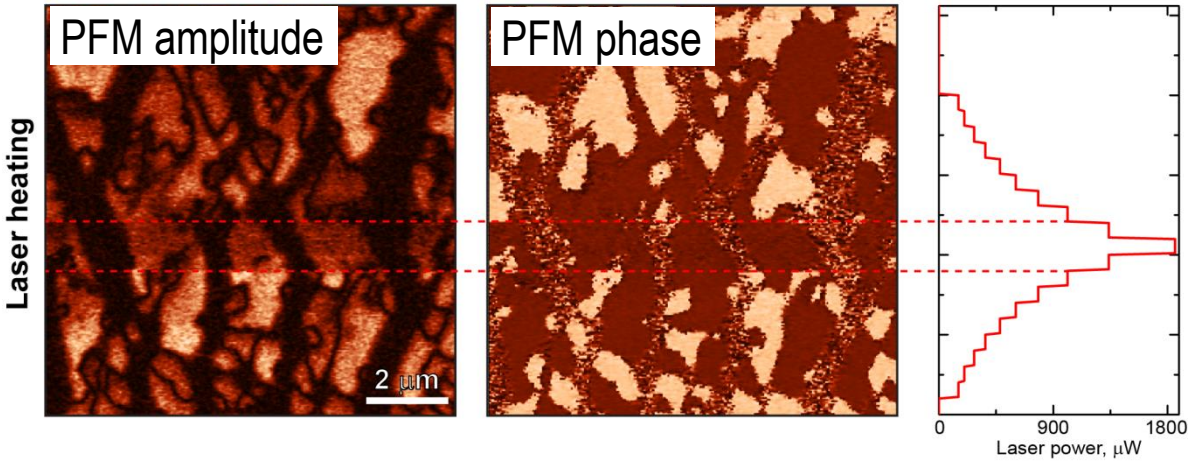
A. Ievlev, ACS Nano **9**, 12442 (2015).

Laser heating induced phase transition

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

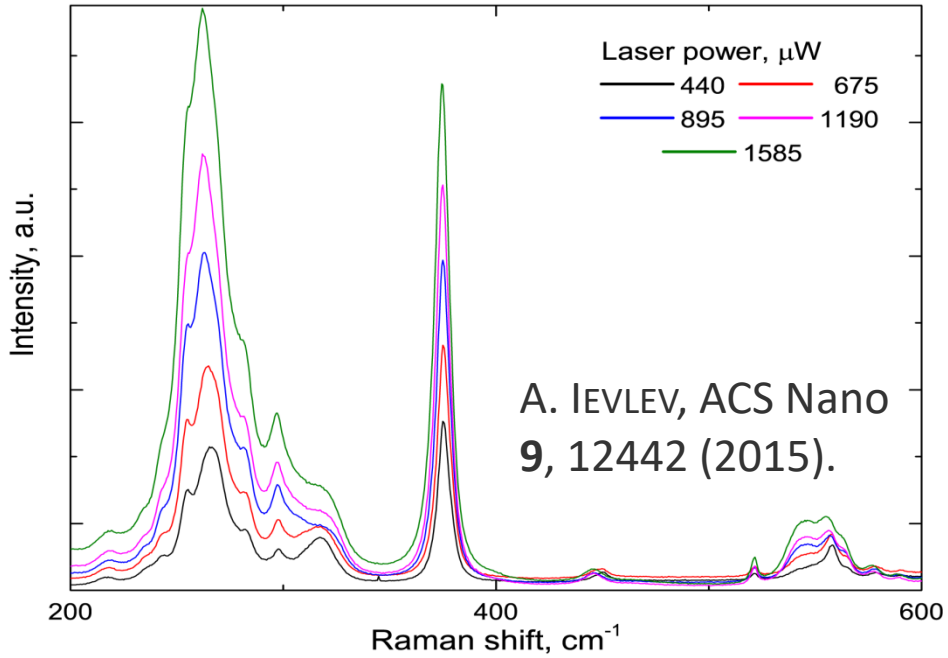


Domain structure evolution **Laser power**



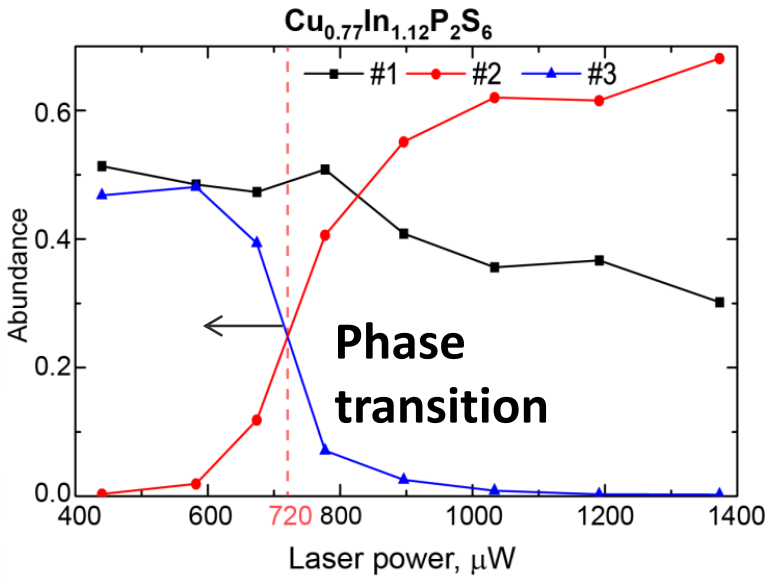
- Measurements with variation of the laser power
  - PFM – *in-situ* change in the domain structure above T_c
  - Raman – evolution of the Raman spectra through the phase transition
- Comprehensive analysis of Raman spectra is complicated by inhomogeneous chemical composition and high noise level
- Bayesian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

Raman spectra evolution (averaged)



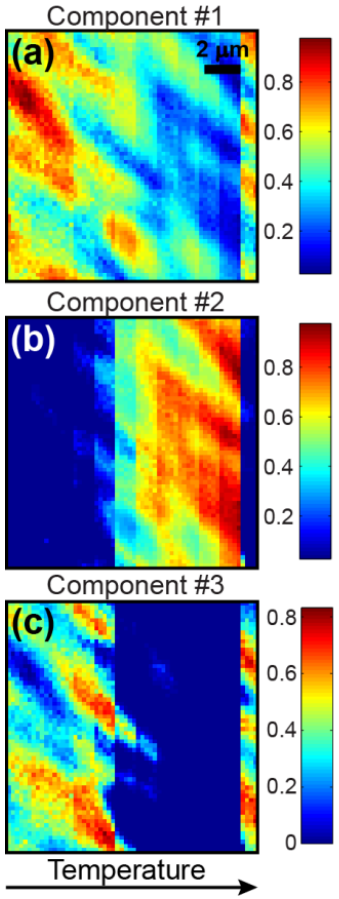
BLU separation of components

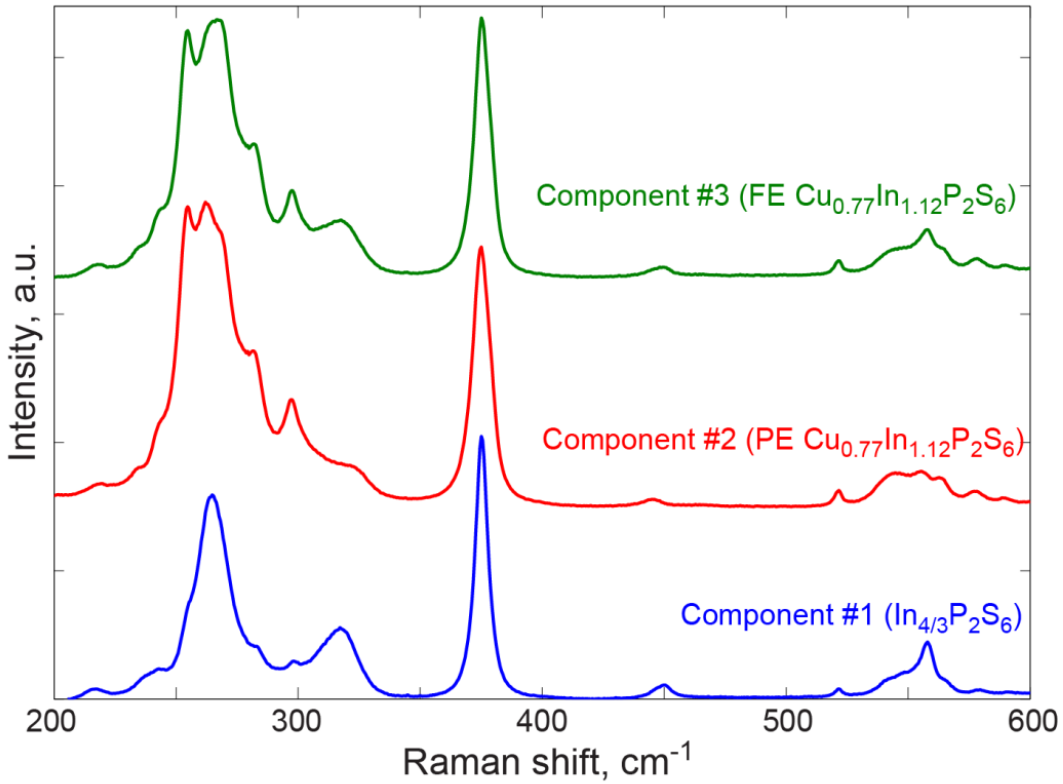
Spatial concentration of components



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

13_PCA_CL.ipynb