UMONS
Université de Mons

Multivariate analysis in ellipsometry data processing

A review with examples of applications

BPS General Scientific Meeting
Ghent - 29/04/2016

C. Guyot, C.O. Zogning, M. Voué (*)
Physique des Matériaux et Optique
Université de Mons

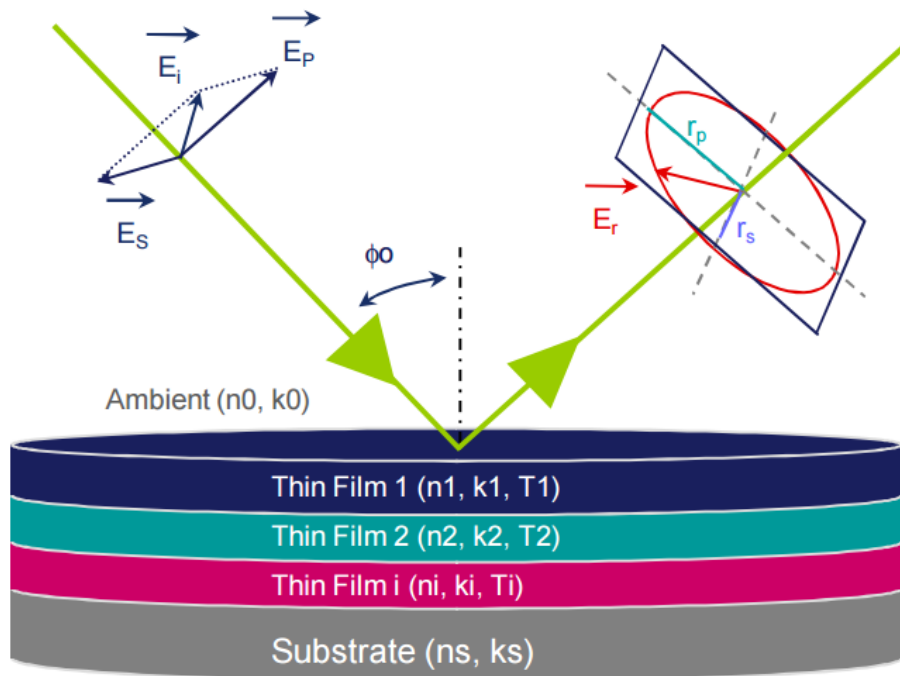
Outline

- **Ellipsometry and optical properties**
- **Multivariate analysis**
- **Recents application of multivariate analysis to ellipsometry data**
  - **Hybrid clustering in SIE data analysis**
  - **Support Vector Machines (SVM) and plasmonic nanocomposites**
  - **Complex Principal Components Analysis (CPCA) for organic monolayers**

Experimental and multivariate techniques

Spectroscopic Ellipsometry (SE)

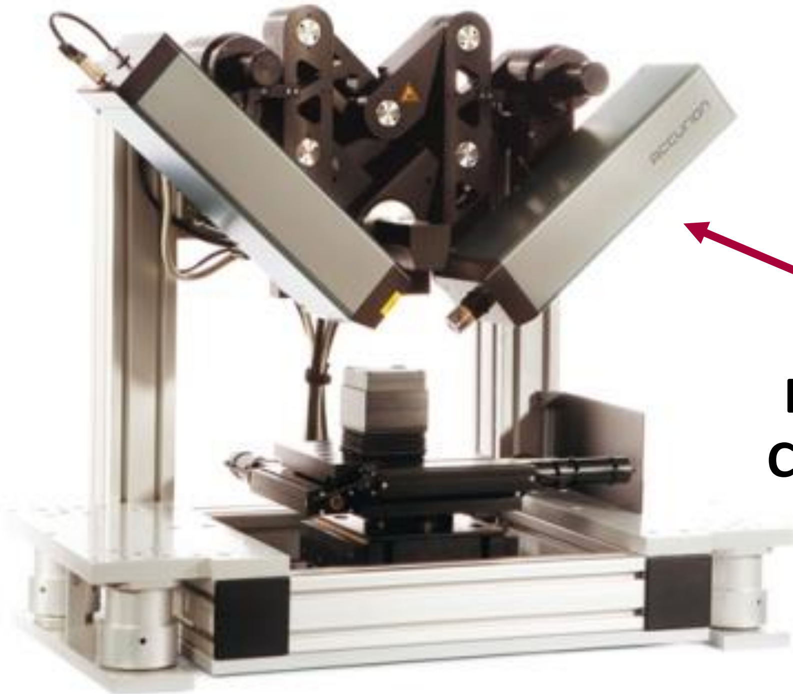
$$\rho = \frac{r_p}{r_s} = \tan\Psi e^{i\Delta}$$



(From : SOPRA R&D)

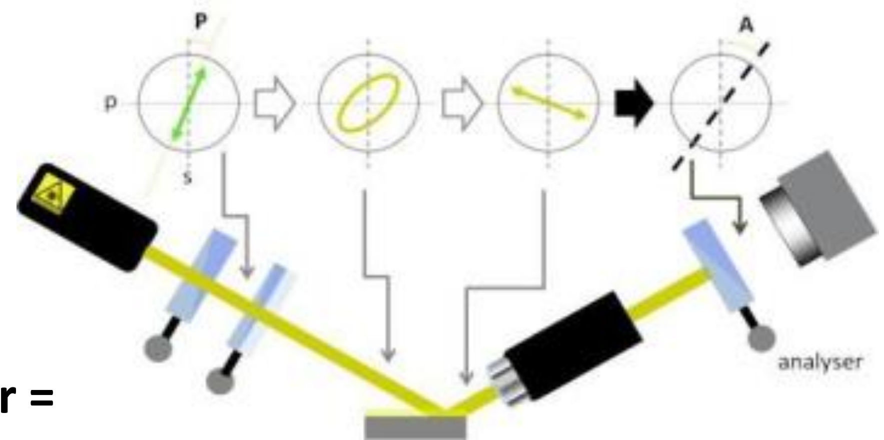
- Non destructive optical method
- Change of polarisation state upon reflexion
- Optical properties (n, k) and thicknesses of the layers
- Ellipticity ρ
- Optical model with 2 unknowns allowed per wavelength
- SE or VASE

Imaging ellipsometry



(From : Accurion GmbH)

Detector = CCD camera

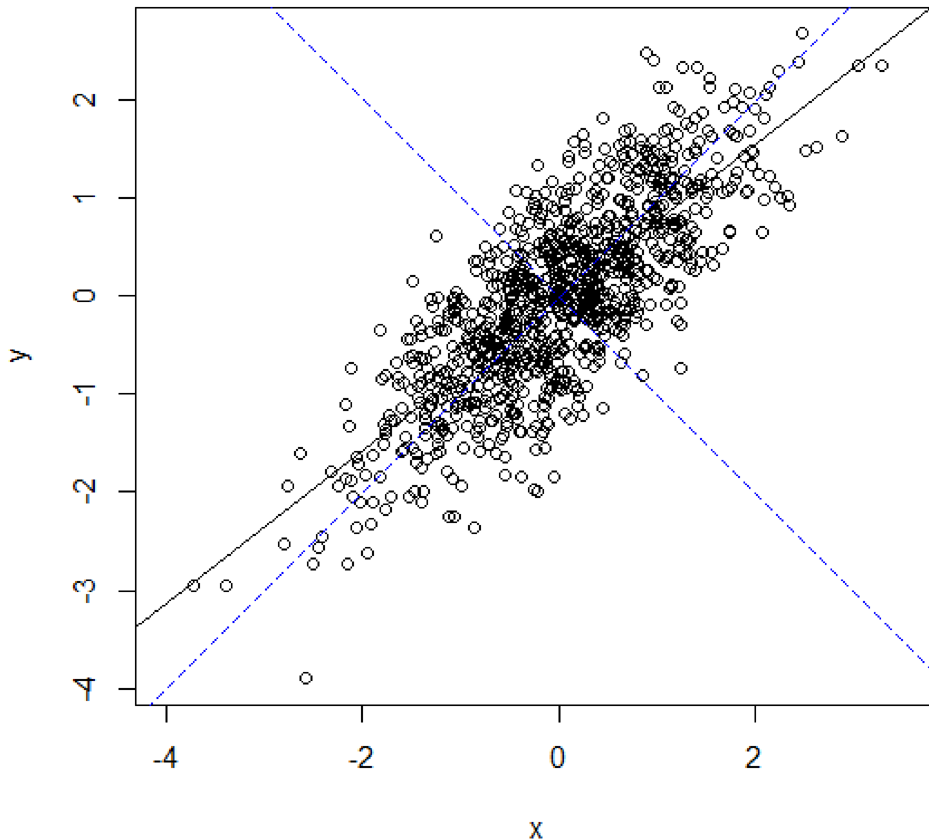


- Nulling ellipsometry
- Same as SE but optical properties at 1µm-scale
- Large number of data (esp. if spectroscopic) : data cube
- Optical model changes from pixel to pixel

Complex Principal Components Analysis (CPCA)

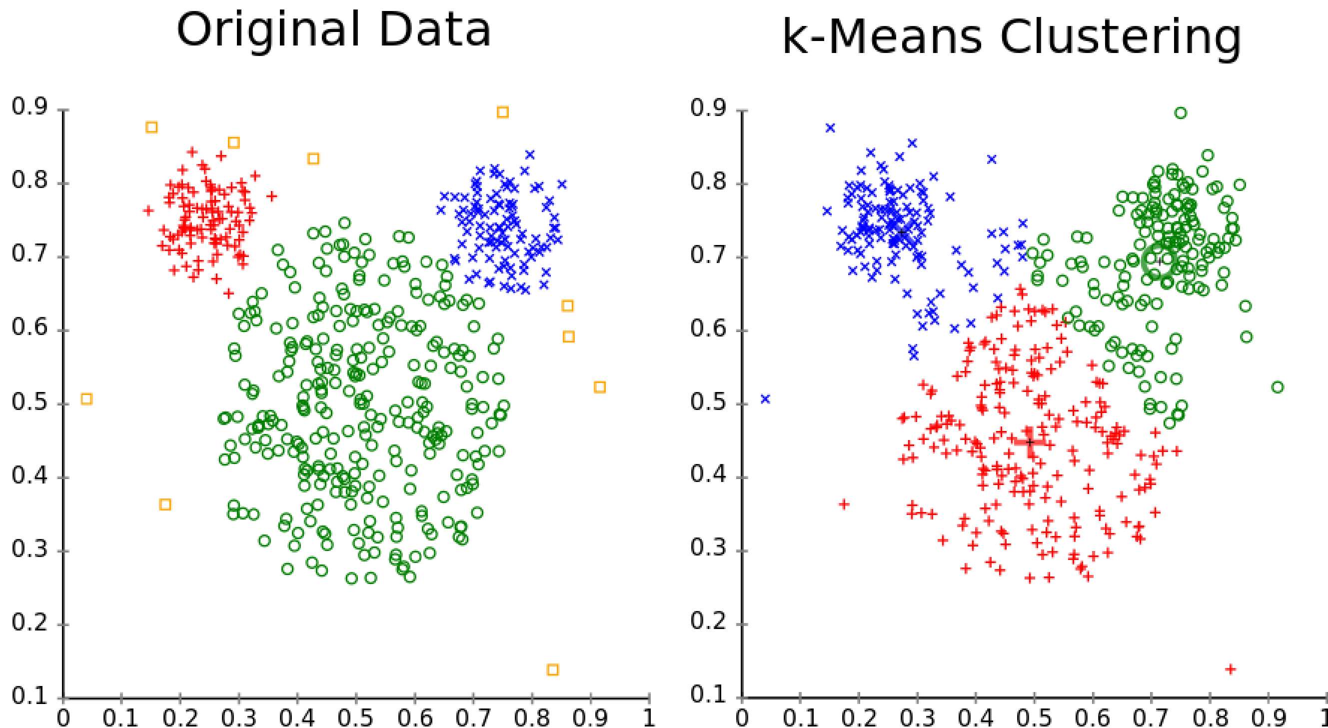
- PCA : statistical method used to reduce the number of variables while preserving the variance of the data
- Diagonalisation of the correlation matrix
- Principal components: projection of the data on the *new* axis
- CPCA : Extension of PCA to complex variables

Principal Component Analysis : exemple



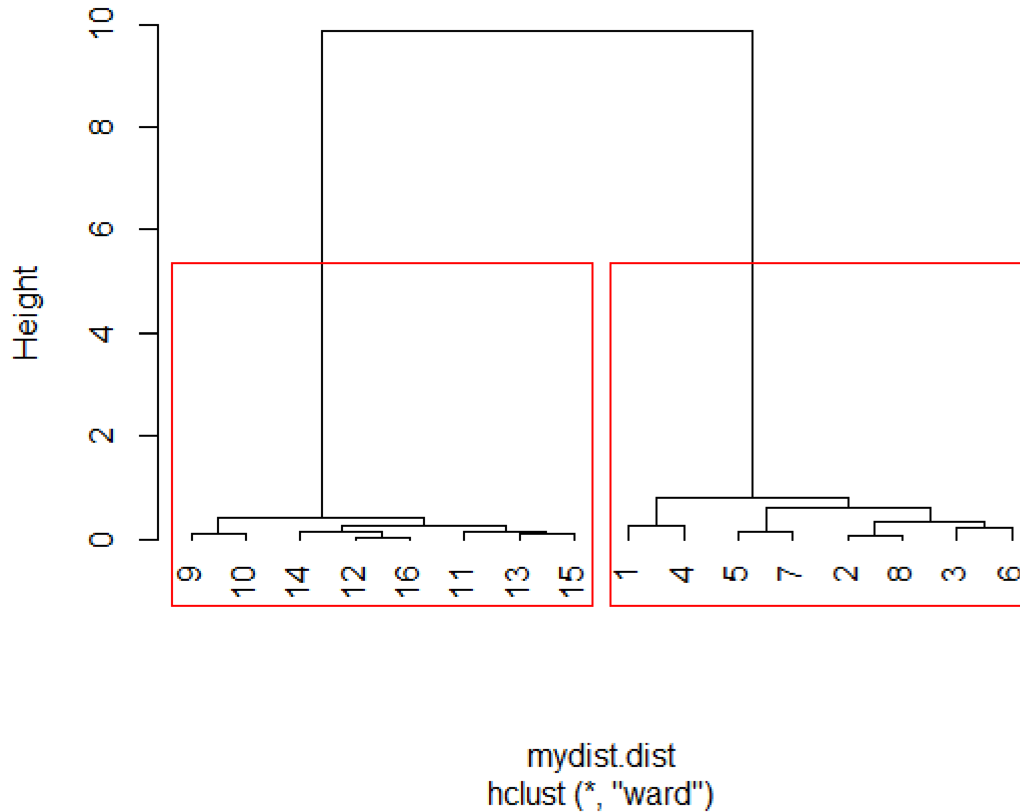
- Randomly generated data with correlation 0.8
- Linear model (black line)
- Principal axis (dashed blue lines)

K-means algorithms

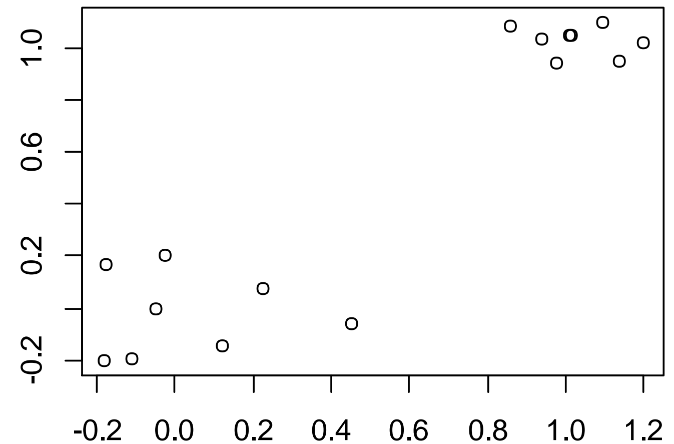


- Iterative method with Random start (local minimum !)
- Element assigned to the nearest cluster
- Number of cluster set at the beginning

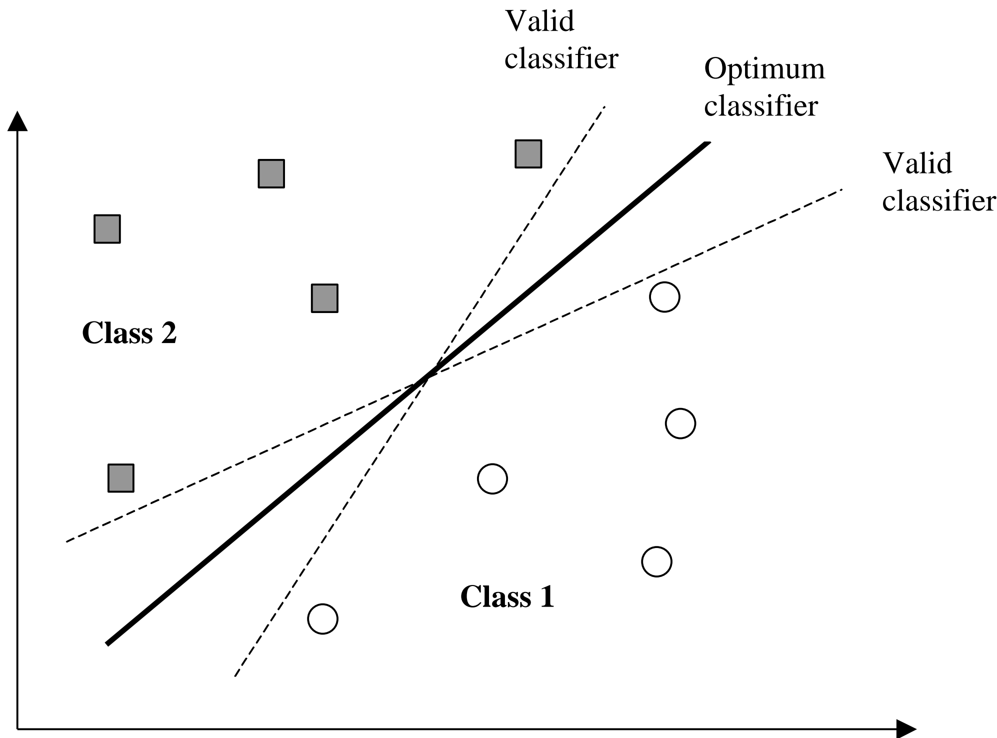
Hierachical clustering



- Distance matrix between elements
- Aggregation of ‘nearest’ elements (variance criterion)
- Hierachy of partitions
- Dendrogram



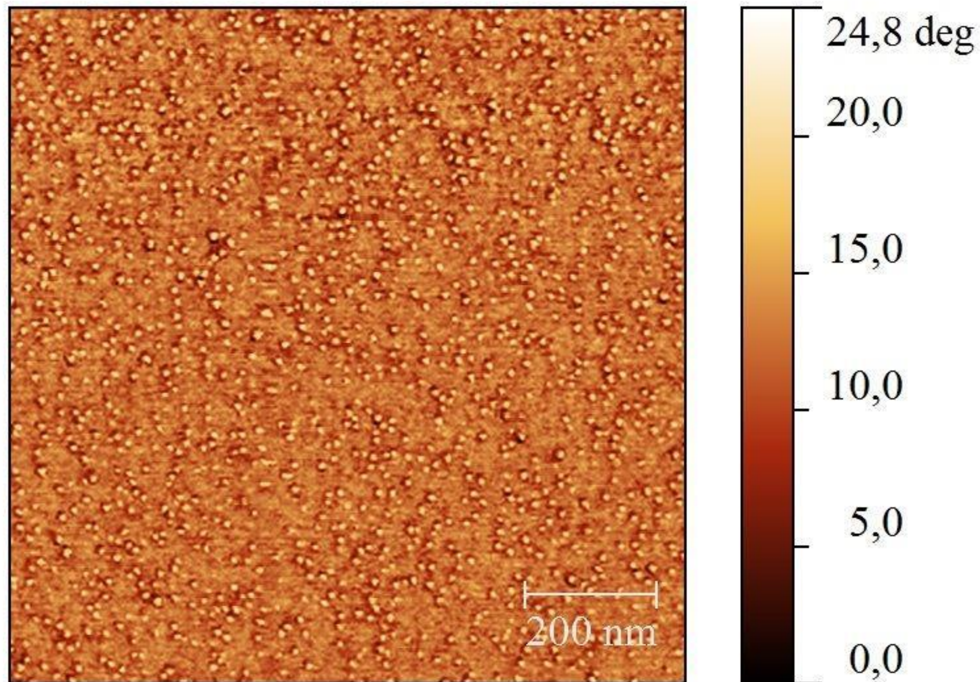
Support Vector Machines



- Linear classifiers
- « Best » hyperplane to separate (overlaping) data sets
- Not linear in real space ? Probably linear in a space of HIGHER dimension
- Maximize the margin

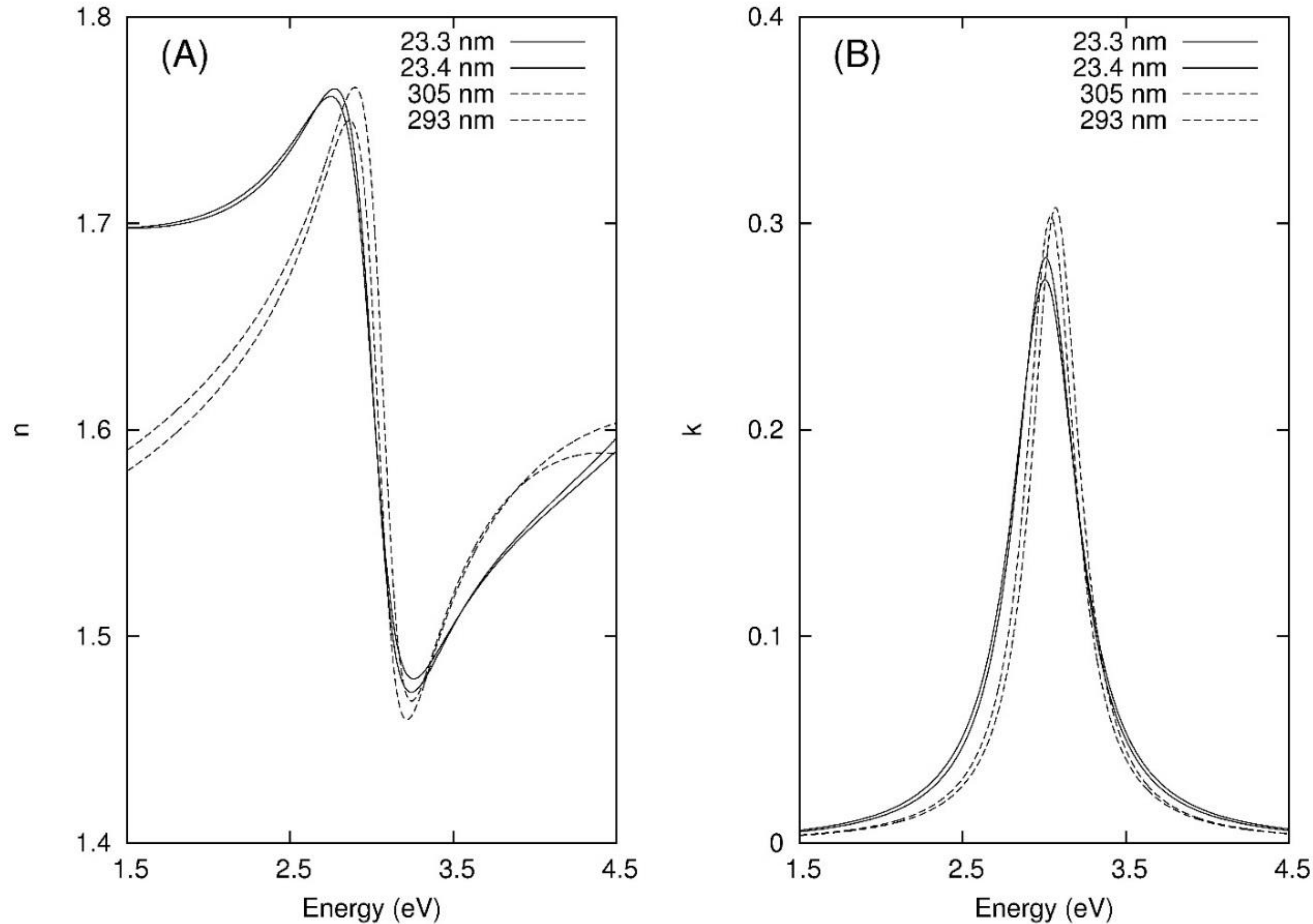
Applications to ellipsometric data

Ag-PVA nanocomposites

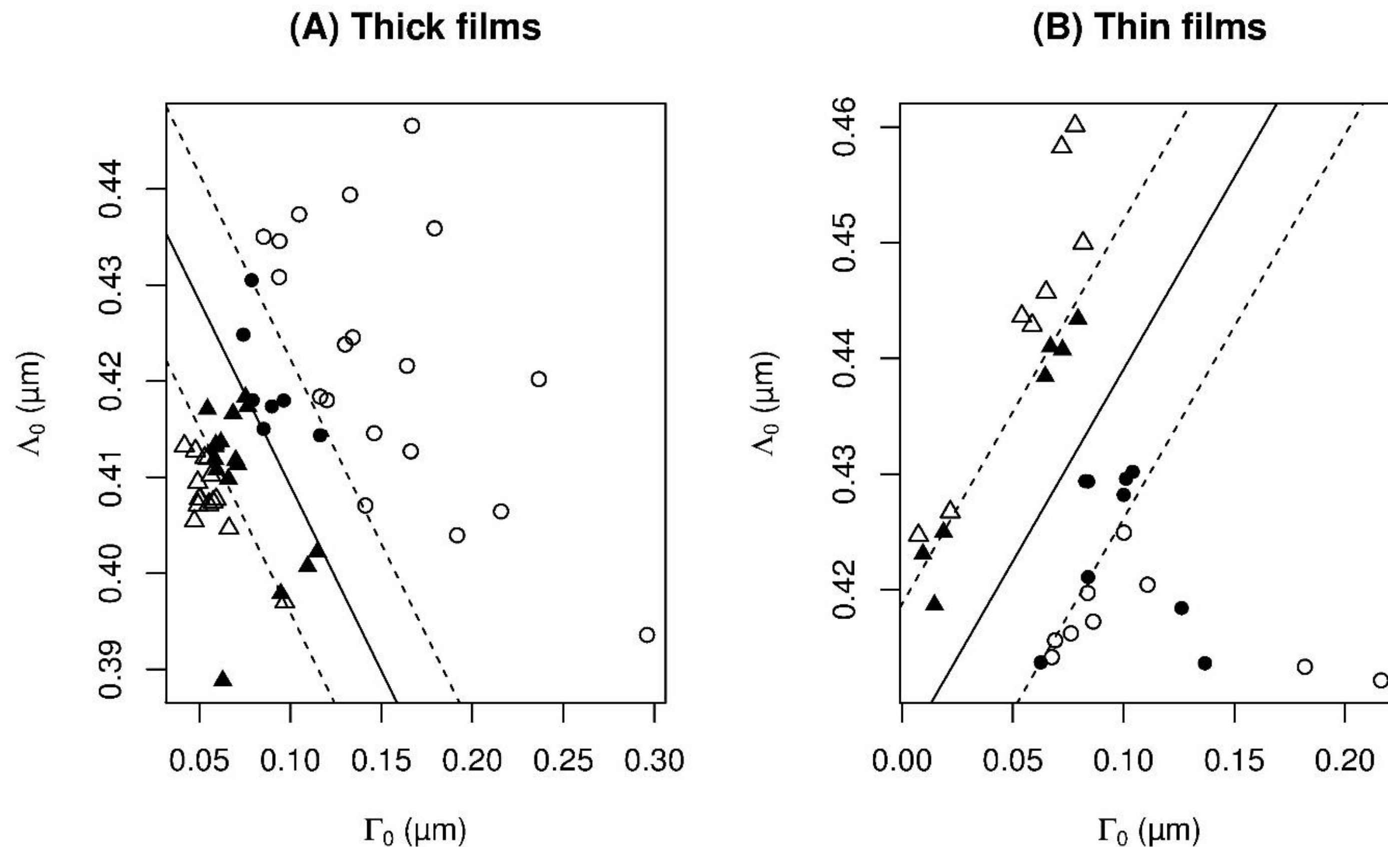


- Ag-PVA 25% 300nm (Phase image)
- Local difference in the elasticity response

Spectroscopic ellipsometry



Link between plasmon resonance parameters

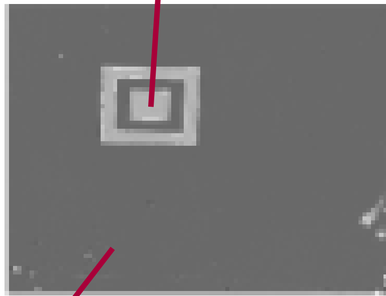
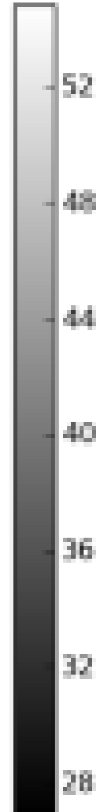


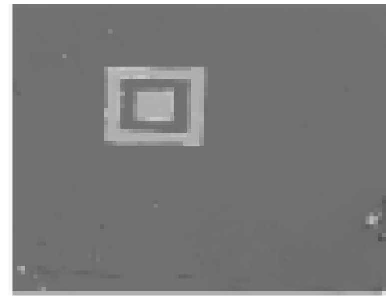
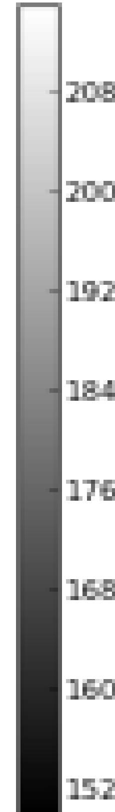
Imaging Ellipsometry

489 nm

Region 2

Region 1

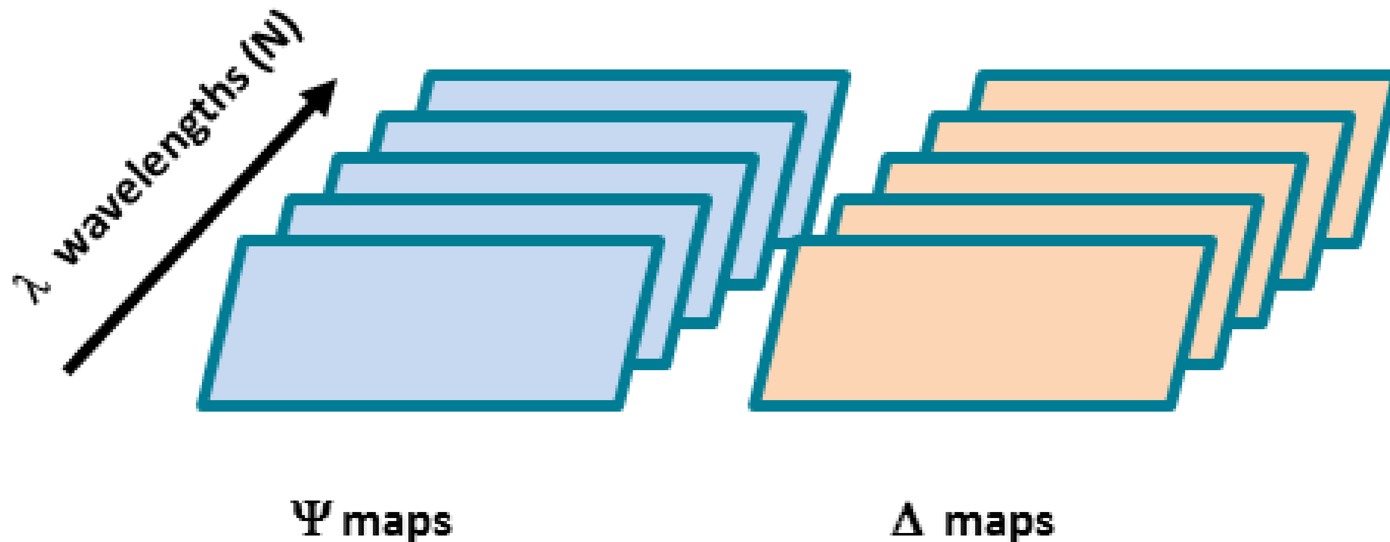



Ψ

Δ

- SiO2 box on native oxide

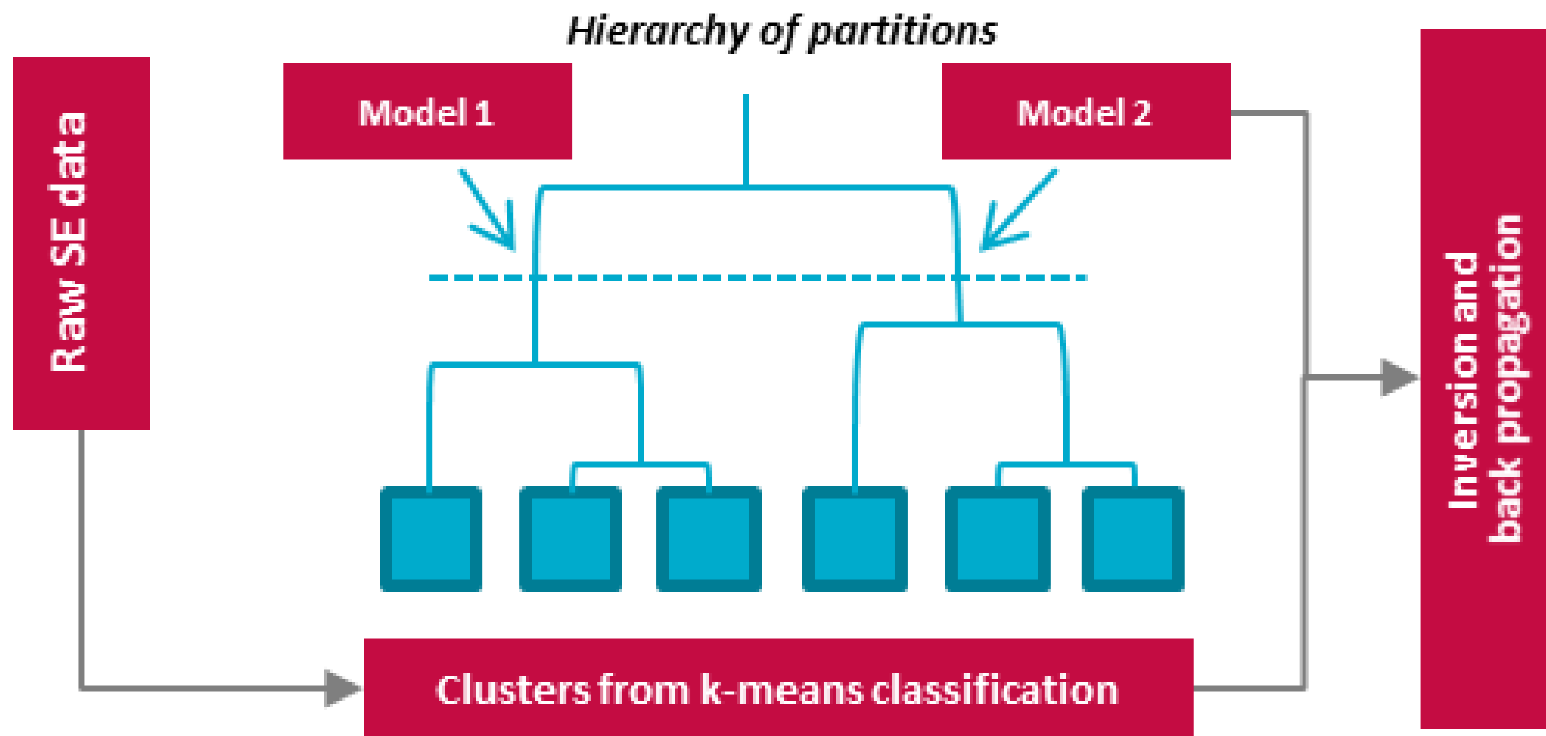
(S)IE data cube



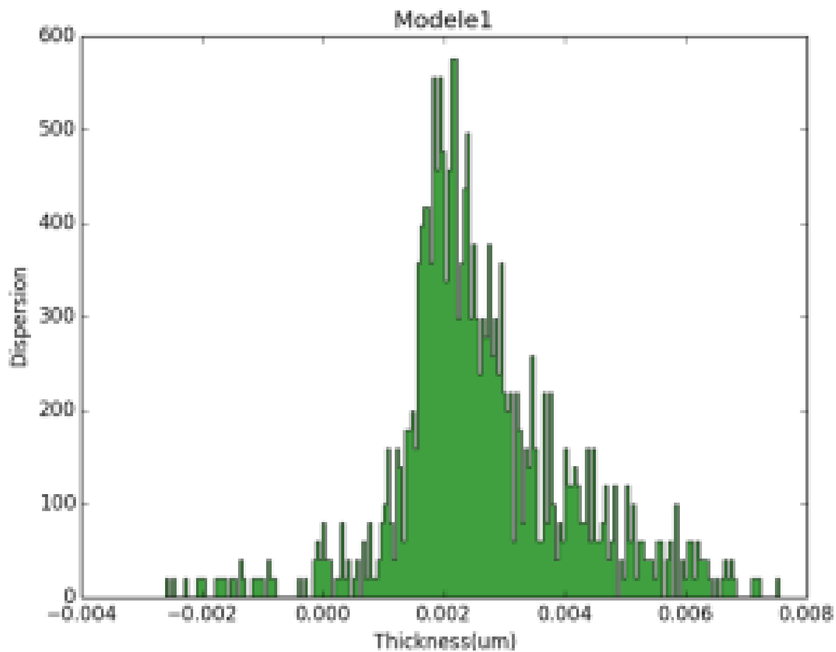
Data cube size : **2N x L x W** with N the number of wavelengths, L the length and W the width of the mapped region of interest

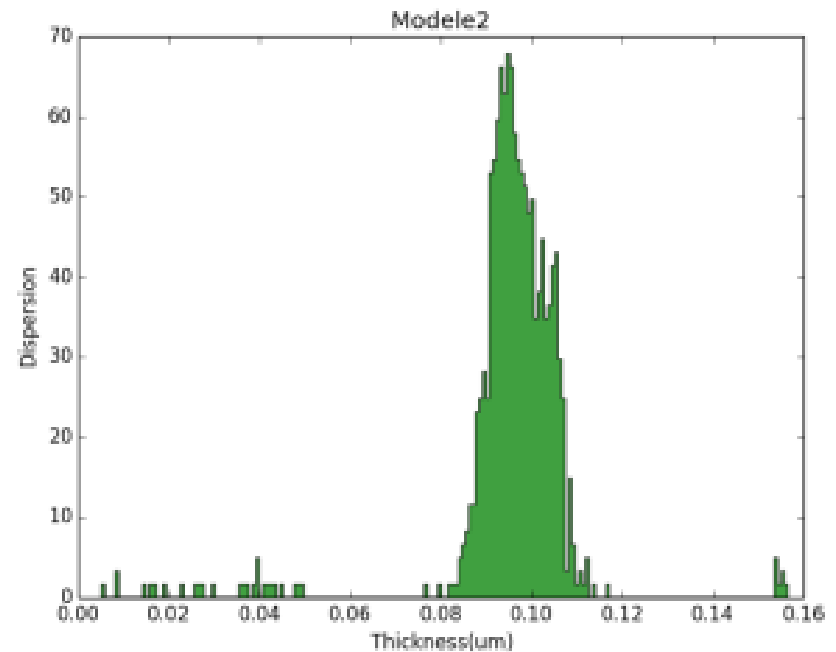
Vector representation : 1 pixel = 1 vector in a 2N-dimensions space

Hybrid clustering method



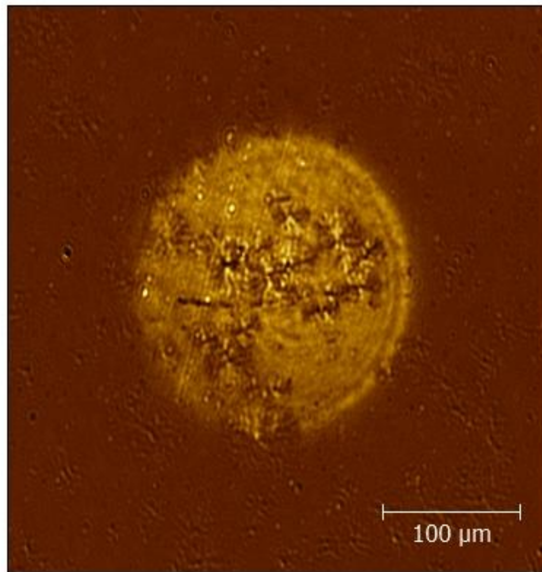
SEI results on SiO2 box

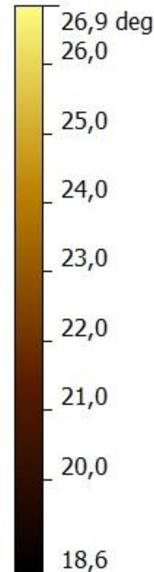


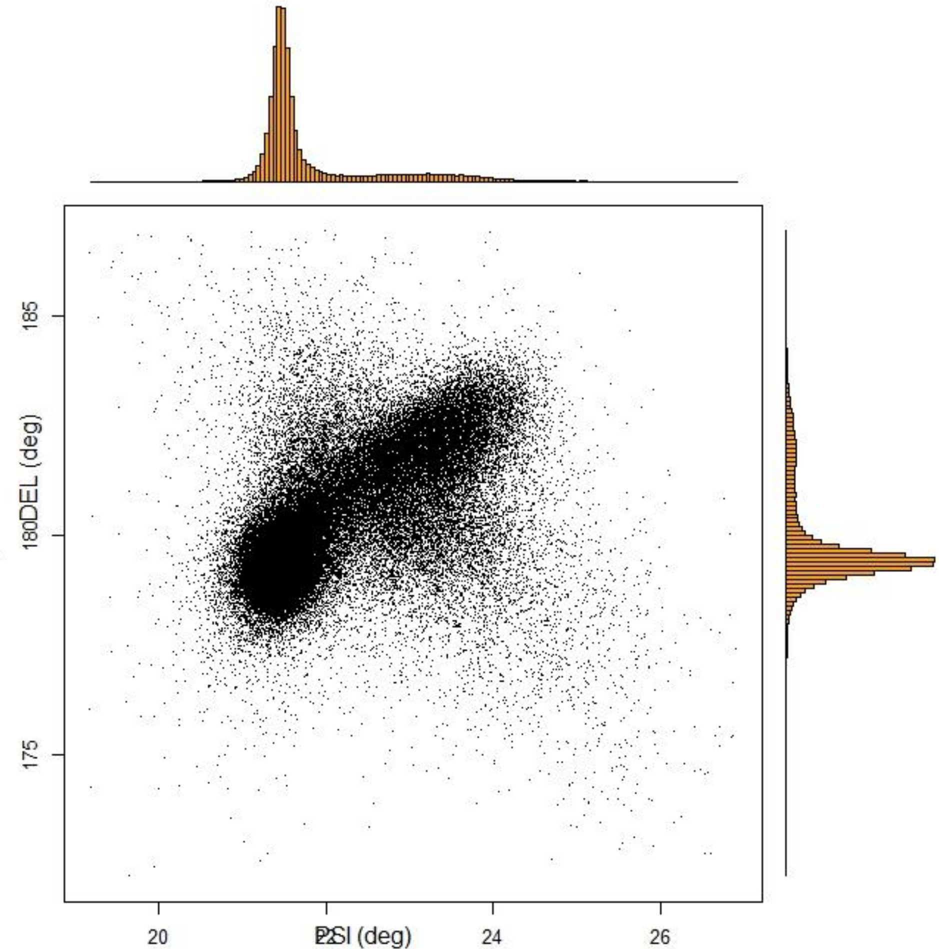


- Inversion of the SIE data and statistics
- Region 1: ~ 2.1 nm
- Region 2: ~ 100 nm

Laser annealing of Ag-doped PVA films

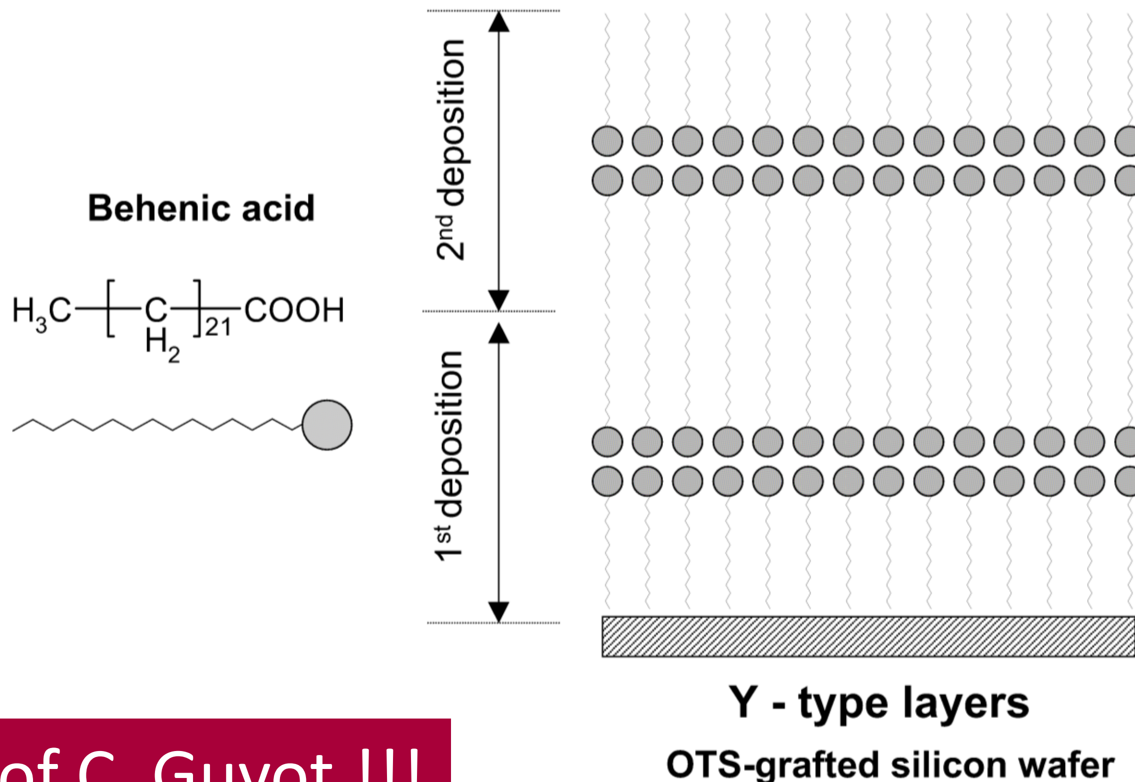






- Two regions (optical models) clearly identified

Optically anisotropic organic multilayers

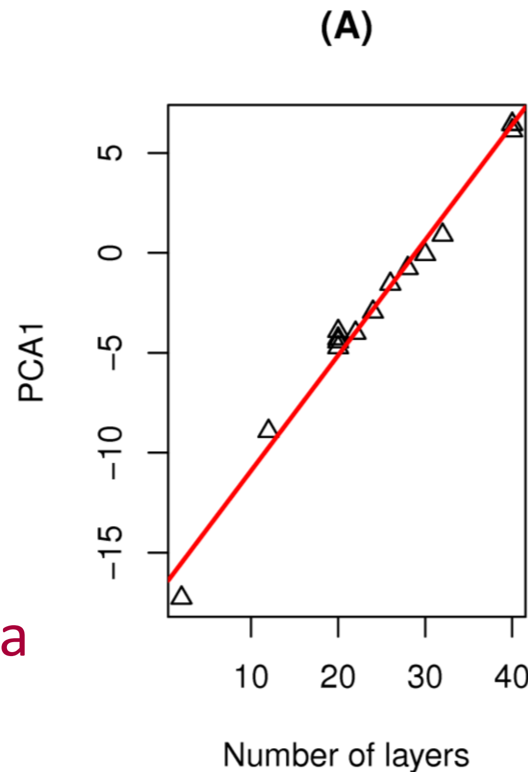


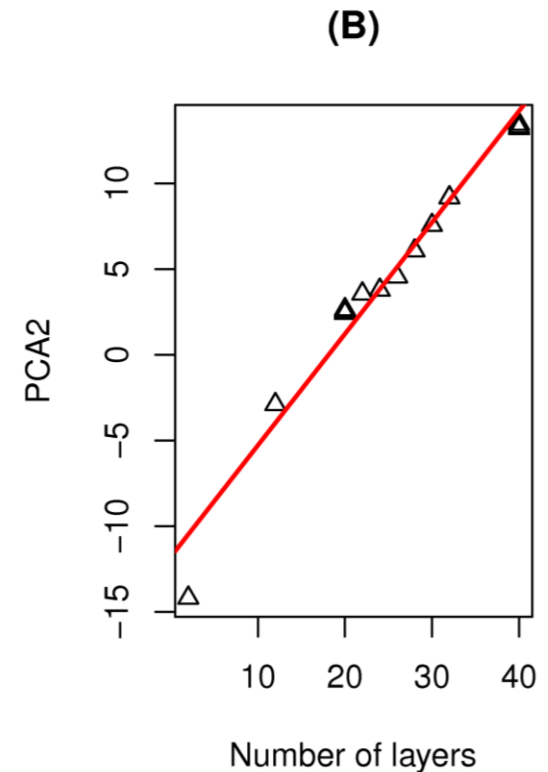
Poster of C. Guyot !!!

Correlation beween PC and multilayer thickness

Transformation of the ellipsometric data
- N wavelengths = 2N variables
- M samples = M individuals

ACP on the (2N x M) data





Take-away message …

- SE and SEI : powerful experimental techniques to investigate locally the optical properties
- Generation of large sets of data on complex samples with latent variables
- Considerable help brought for the data interpretation by multivariate analysis

Acknowledgements and support