Blind Source Separation and Applications

Danielle Nuzillard

Danielle.nuzillard@univ-reims.fr

CReSTIC, UFR Sciences exactes et naturelles

Université de Reims Champagne Ardenne

Outline

- Problem
  - Cocktail party

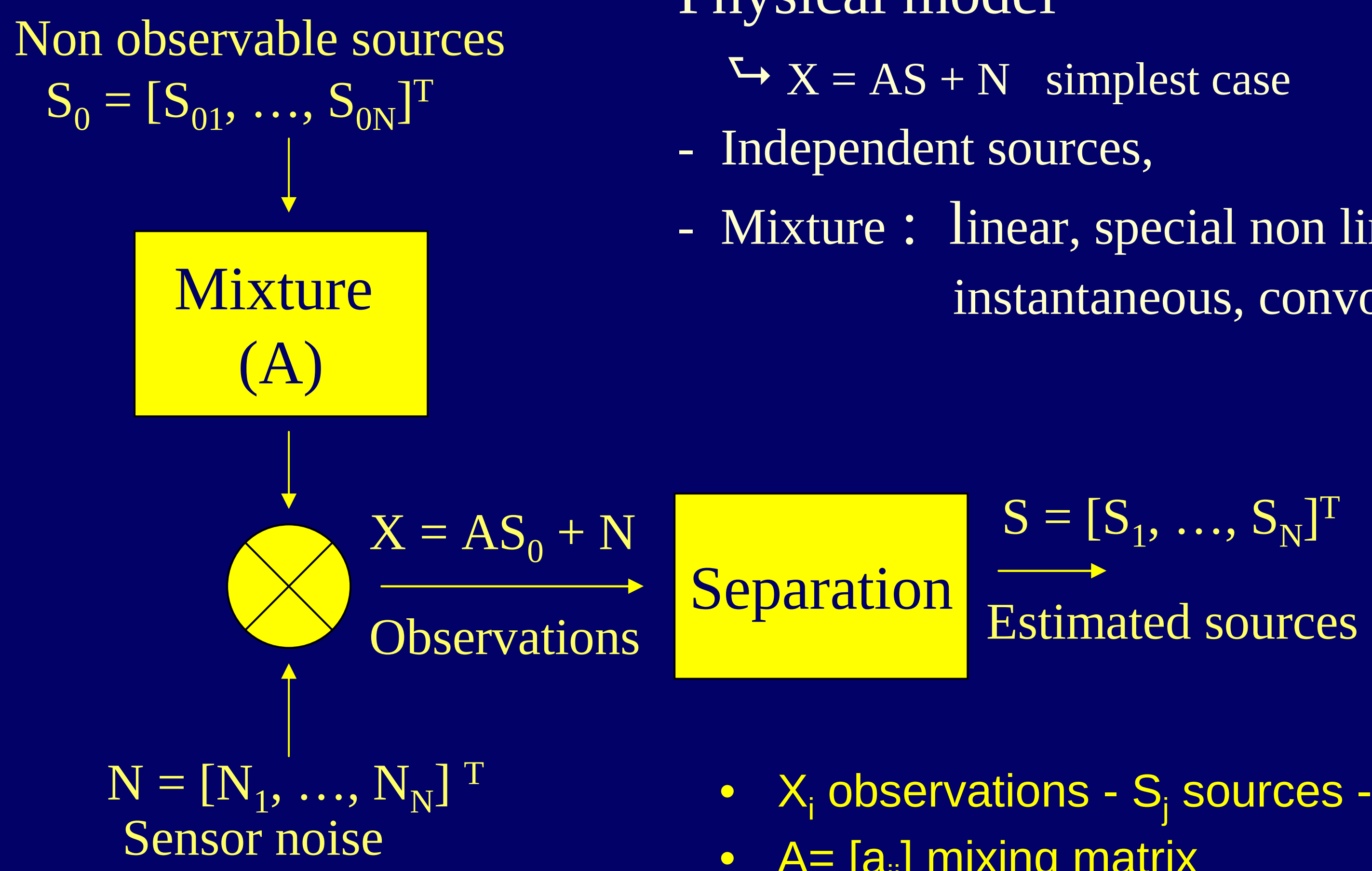
- Decorrelation components
- Independent components
- Applications
  - NMR spectroscopy,
  - Astrophysics,
  - Electron Energy Loss Spectroscopy

Blind Source Separation Methodology

Physical model

↳ $X = AS + N$ simplest case

- Independent sources,
- Mixture : linear, special non linear, instantaneous, convolutiv,



- X_i observations - S_j sources - N_i noise
- $A = [a_{ij}]$ mixing matrix

BSS – Generalities

Mixture → instantaneous linear $x(t) = A \cdot s(t)$

Mixture → post non-linear $x(t) = F(A \cdot s(t))$

Mixture → non-linear $x(t) = F(s(t))$

Separe sources -> Objective fonction + Optimisation

Hypothesis
1. Sources are mutually independent
2. Sources are non gaussian except at least one

Estimated Sources Vectors $y = Wx$

Source -> random variable

Second order statistics – insuffisant for the separation

Third order statistics – not really used

Forth order statistics – suffisant for the separation

- *Sources can be reconstructed up to permutation and scaling factors .*
- *At most only one source is gaussian.*
- *Under-determined mixture, requires supplementary hypothesis*

Decorrelation, Y = AX + N

- PCA, searched sources are :
  - Gaussian, zero mean (centered)
- Rotates the data from the centroid (unitary transformation) so that the variance is maximized along the first axis …

  Goal : Dimension reduction, Denoising (weakest variance direction)
- Whitening (sphering),
  - PCA+ normalized sources

  The transformation is not unitary, the variance is the same in all directions
- Factorial Analysis
  - Wanted sources (gaussian, zero-mean, normalized), they have a physical meaning to interpret the data (less sources than observations and non correlated noise neither between them, nor with the sources).
- Projection pursuit
  - Find interesting projections for sources interpretation

•SOBI (Belouchrani & *al*.) : Cross-correlations between data and shifted data, Joint diagonalization
•f-SOBI (Nuzillard), Cross-correlations computed in the Fourier space
•Maximum likelihood from the spectral density (Pham, Cardoso), analysis of CMB (cosmic microwave background)

Independant Component Analysis (ICA)

Hypothesis: (most common assumptions)

- S_i are real processus statistically independent.
- S_i are non-gaussian.
- Noise is negligible
- Matrix A is invertible.

Principe:

Central Limit theorem = the distribution of a sum of independent variables tends toward a gaussian distribution

♦means that mixing variables provide a gaussian mixture

♦idea: maximise the non-gaussianity of variables allow to separe them

♦ Contrast function: measure the non-gaussianity

♦ To Maximise a contrast function = to separe variables or to give independent variables.

Statistical Independence (1/2)

Statistical Independence $p(u_i, u_j) = p(u_i)\,p(u_j)$

Fondamental constraint : non gaussian sources

Non-gaussianity measurement

Kurtosis $kurt(y) = E(y^4) - 3$

Gaussian variable	sub-gaussian variable	super-gaussian variable
$kurt(y) = 0$	$kurt(y) < 0$	$kurt(y) > 0$

Drawback: it is sensitive to the noise ans to aberrant values

---

Negentropy $J(y) = H(y_{gauss}) - H(y)$

Practically we use an approximation of the negentropy :

$$J(y) \approx (E[G^2(y)] - E[G^2(v)])^2$$

$$G_1(y) = \frac{1}{a_1} \log\cosh a_1 u \qquad G_2(y) = -\exp(-u^2/2)$$

Statistical Independence (2/2)

Likelihood function

$$\log L(W) = \sum_{t=1}^{T}\sum_{i=1}^{n} \log p_i(w_i^T x(t)) + T\log\left|\det W\right|$$

Sub-gaussian variables

$$\log p_i^+(s) = \alpha_1 - 2\log\cosh(s)$$

Super-gaussian variables

$$\log p_i^-(s) = \alpha_2 - [s^2/2] - \log\cosh(s)$$

Mutual Information

$$I(y_1, y_2, \ldots, y_m) = \sum_{i=1}^{m} H(y_i) - H(y)$$

Property: for a linear transformation $y = Wx$

$$I(y_1, y_2, \ldots, y_m) = \sum_{i=1}^{m} H(y_i) - H(x) - \log\left|\det W\right|$$

Pre-processings for BSS

1. Centering $\underline{X} = X - m$
2. Whitening $X = ED^{-1/2}E^T X$
3. Filtering $Xf = XF = ASF = ASf$

ICA Algorithms (most popular)

- Comon's approach
  - PDF Edgeworth Approximation
  - Cumulants
- JADE (Cardoso & Souloumiac)
  - Based on 4th order cumulants
  - Joint Diagonalisation
- Infomax (Bell & Sejnowski)
  - ANN (neural network)
- FastICA (Oja & Hyvarynen)

Constraint

Notation: Y data, W weight matrix, H sources, F is the criterion

- No constraint: to Mimimize the Likelihood ⇔ to solve F(W,H) = $1/2\| Y - WH \|^2_F$

$$\frac{\partial F}{\partial W} = (WH - Y)H^T \qquad \frac{\partial F}{\partial H} = W^T(WH - Y)$$

$$w_{ij}^{(t)} = w_{ij}^{(t-1)} - \boldsymbol{\theta_{ij}} \frac{\partial F(W,H)}{\partial w_{ij}} \qquad h_{ij}^{(t)} = h_{ij}^{(t-1)} - \boldsymbol{\phi_{ij}} \frac{\partial F(W,H)}{\partial h_{ij}}$$

Lee and Seung chose the quantities:

$$\theta_{ij}^{t} = \frac{w_{ij}^{(t-1)}}{(W^{(t-1)} H H^T)_{ij}} \qquad \phi_{ij}^{t} = \frac{h_{ij}^{(t-1)}}{(W^T W H^{(t-1)})_{ij}}$$

$$w_{ij}^{(t)} = w_{ij}^{(t-1)} \frac{(YH^T)_{ij}}{(W^{(t-1)} H H^T)_{ij}} \qquad h_{ij}^{(t)} = h_{ij}^{(t-1)} \frac{(W^T Y)_{ij}}{(W^T W H^{(t-1)})_{ij}}$$

- Constraint : Lagrange multiplier

$\mathrm{Min}_{W,H}\{\| Y - WH \|^2_F + \alpha J_1(W) + \beta J_2(H)\}$ where J_1 and J_2 are constraints

Spectral Analysis of small molecules in solution

D. Nuzillard, J.-M. Nuzillard, S. Bourg (LAM, Pharmacognosie, URCA)

- Nuclear Magnetic Resonance Spectroscopy
  - Interaction, excitation principe
  - Chemical displacement, response signal shape.
- Help to determine the structure of
  - molecules in the composition of one mixture,
  - a solely molecule,
  - molecules in the composition of many mixtures.
- Contributions
  - Take into account experimental constraints,
  - Developpement of *f*-SOBI.

Spectral Analysis of small molecules, NMR analysis

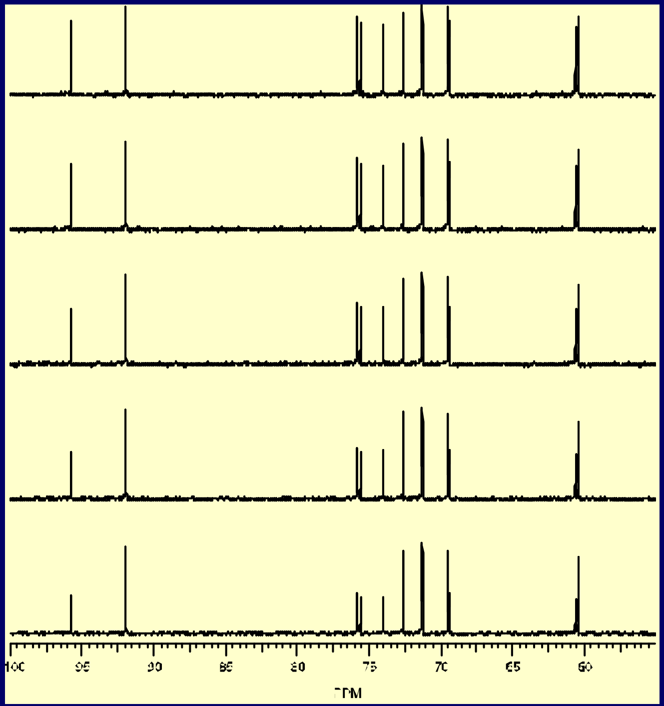
D. Nuzillard, J.-M. Nuzillard, S. Bourg (CReSTIC, Pharmacognosie, URCA)

- Raw data : 1-D, 2-D, 3-D signals
- Radio-frequency spectra produced by FT
  - 60 to 800 MHz for ^{1}H, 15 à 200 MHz for ^{13}C,
  - Frequency axis reflects the electronic environment,
    - The possible values of σ are very close to each others and they do not depend on B_0 ,
    - The values of ν depend on B_0.
    - Chemical displacement in ppm : we define a new variable of frequency from a reference substance , TMS (TetraMéthylSilane) : $Si(CH_3)_4$.

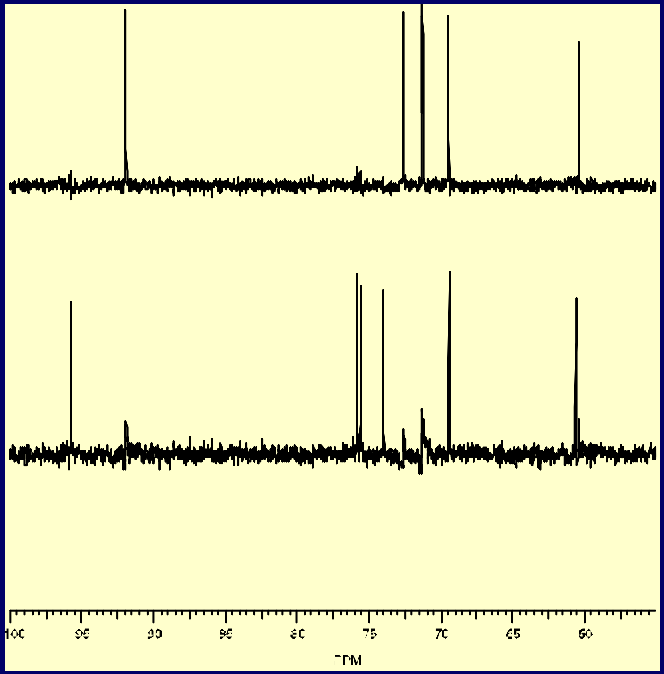
$$\delta = \frac{\nu - \nu_{ref}}{\nu_{ref}} \times 10^6$$

α-glucose to β-glucose Isomerisation

A solely mixture in variable concentration



Mixtures



Components

- RMN ^{13}C spectra are recorded every 45 minutes,
- Temporal data are separated using SOBI,
- FT gives separated spectra.

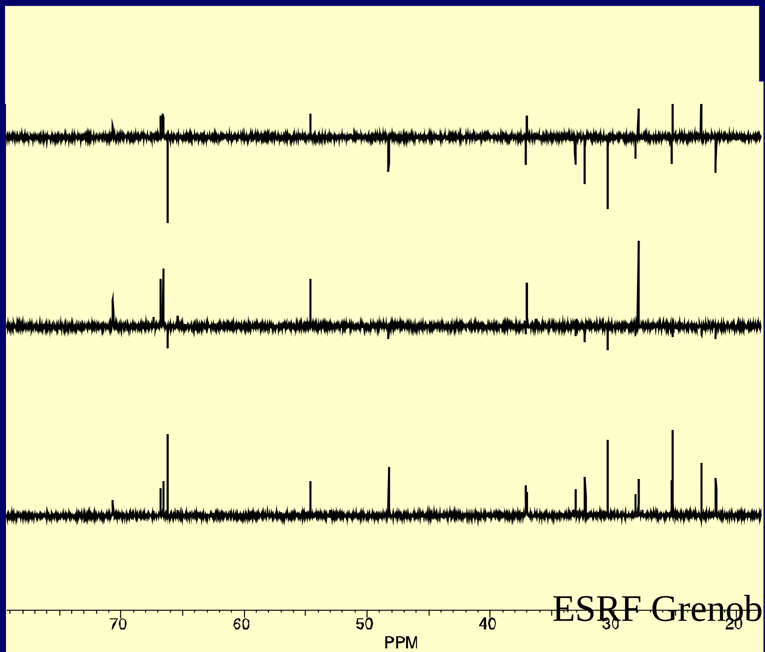
Molecule strcuture ^{13}C Spectra 1-D

Impulses parametered by their duration

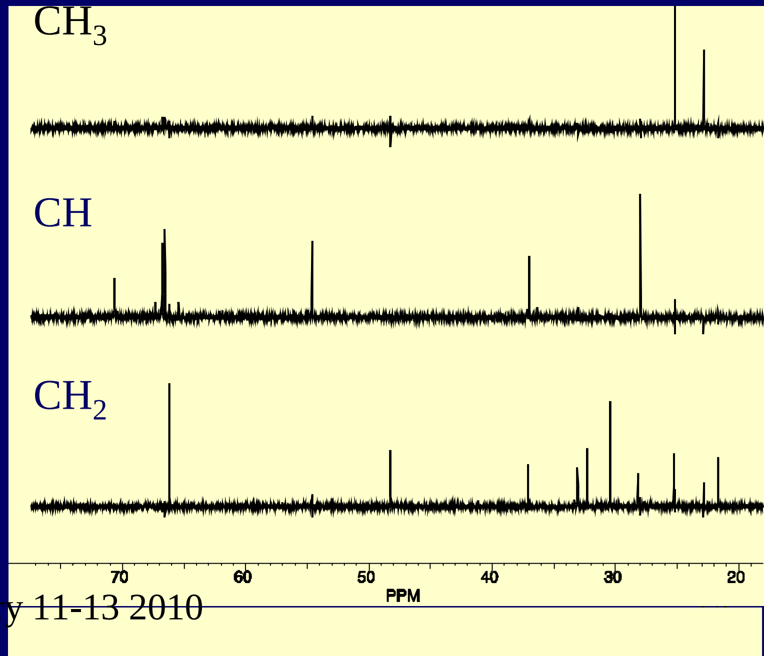
- Correspond to a rocking angle θ of the aimantation of the sample
- The intensity of rays depends on θ and on ‘n’ the number of H in each CH_n
- 3 angles θ not precisely known : about 45°, 90°, 135°

$$I = I_0 . n . \cos^{n-1}(\theta) . \sin(\theta)$$

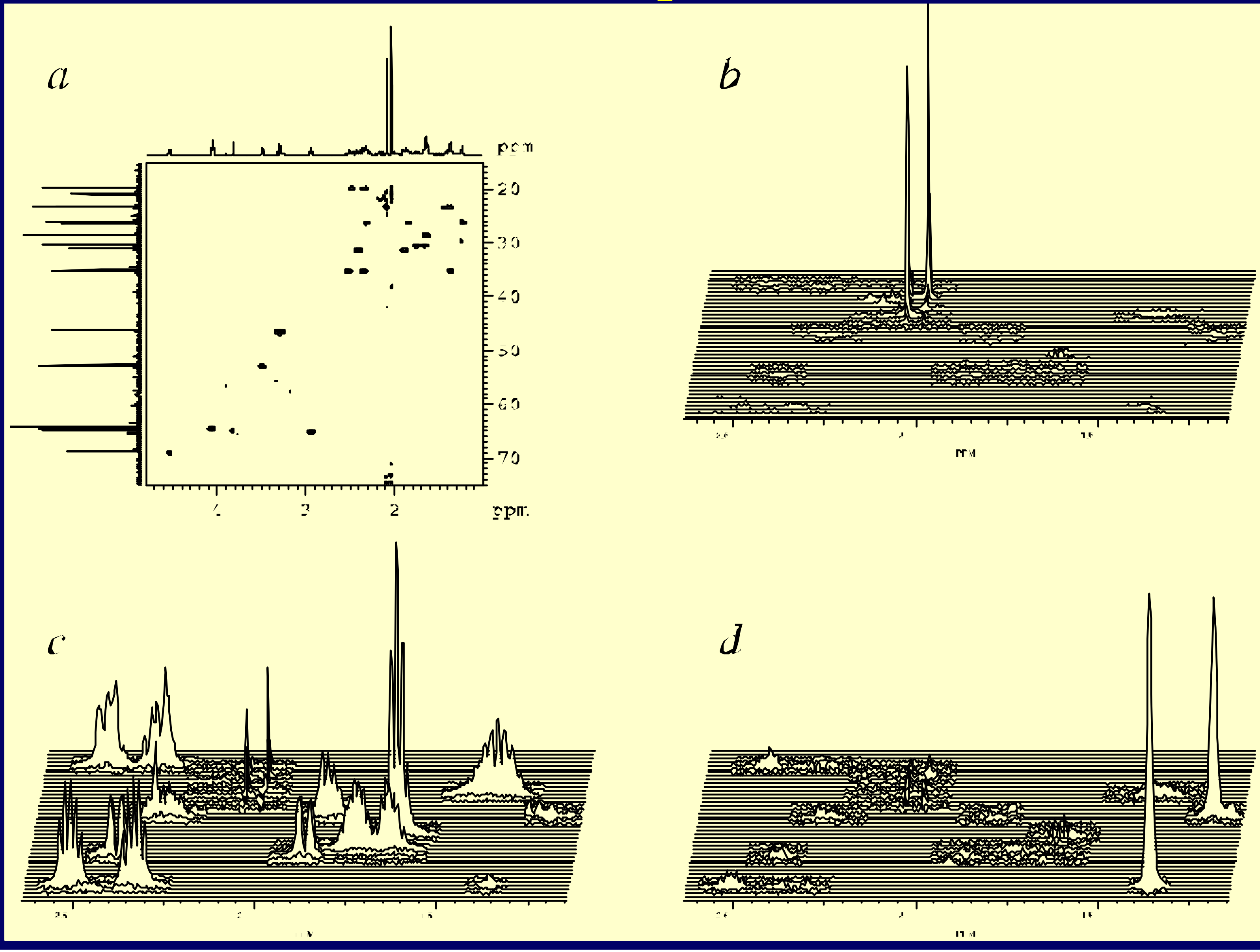
CH, CH_2, CH_3 Sub-spectra in mixture



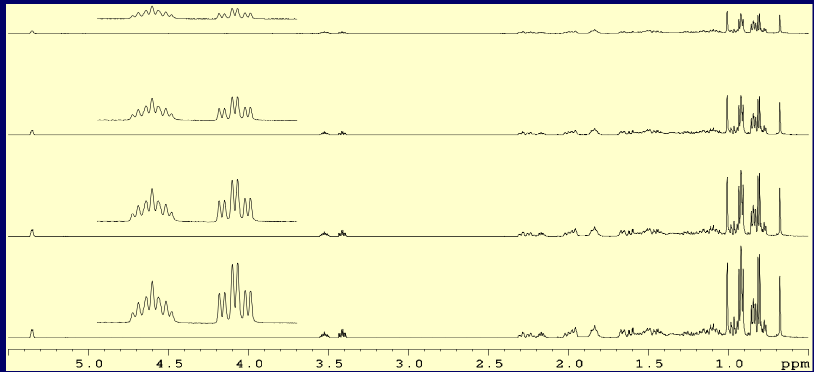
Separated components



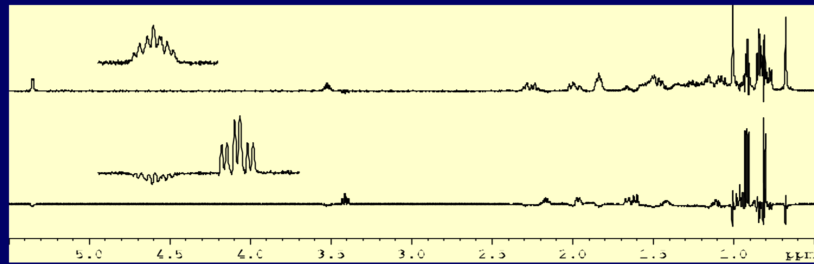
Molecule Structure
2-D correlation Spectra ^{1}H - ^{13}C



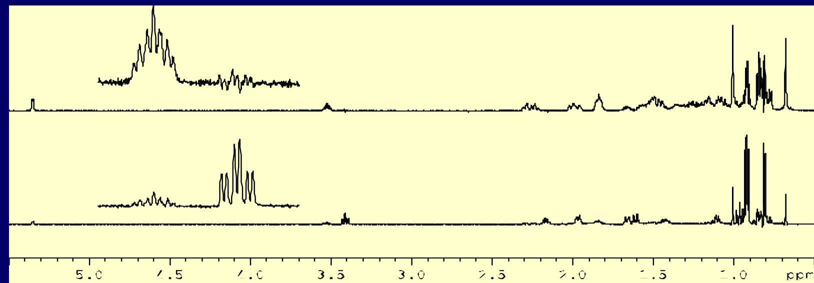
Analysis of a Mixture



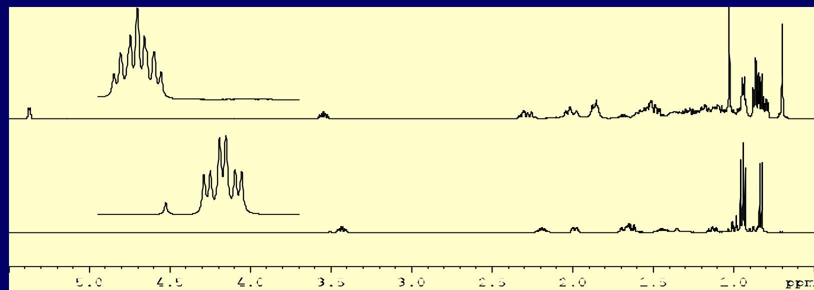
Mixtures : β-sitosterol and Menthol



Semi-blind



Blind + ALS



References

Non orthogonal Spectra

- Physical Constraint

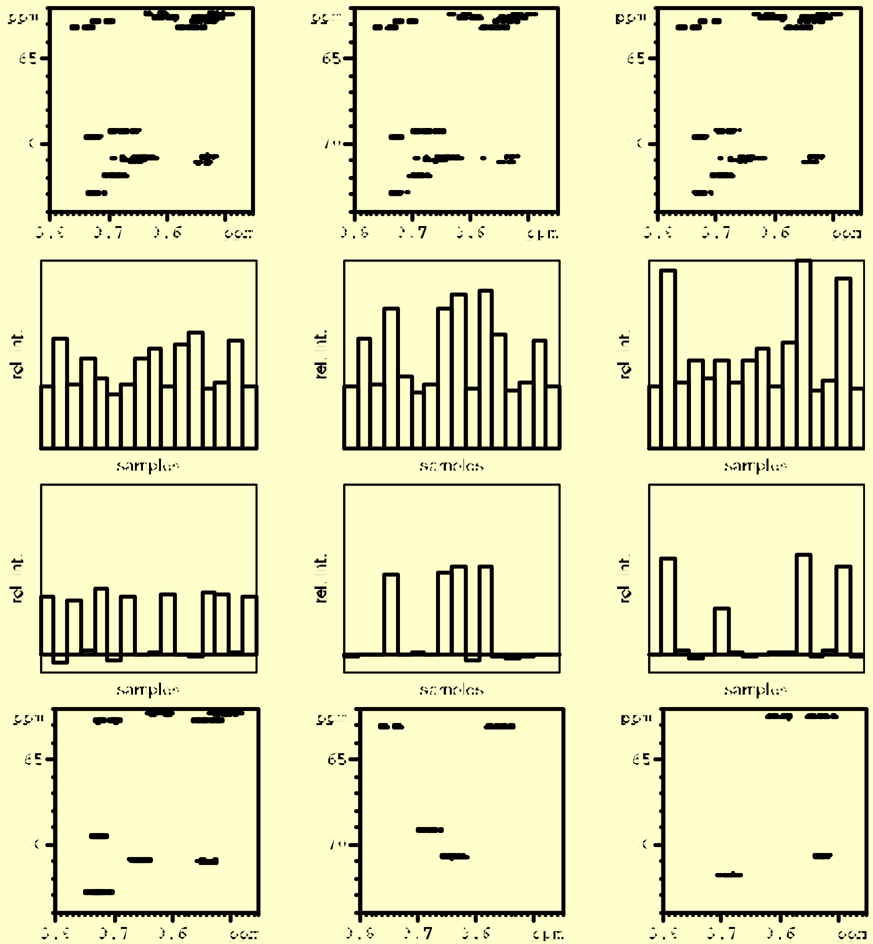
Strategy 1

- Separation on a orthogonal part
- Application of the separation matrix on the set of spectra

Strategy 2

- Blind Separation
- Applications of the positivity constraints of sources and mixtures

Many mixtures of many molecules, frequential variation due to the concentration

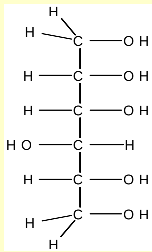


Integration, positivity constraint

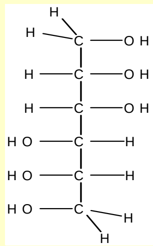
Spectres

Pseudo spectra

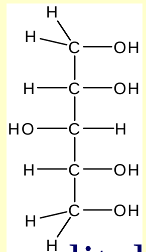
Separated pseudo-spectra



sorbitol



mannitol



xylitol

Reconstructed sources-spectra

Separation in the spectral domain
f-SOBI

Presented to GRETI in 1999, Signal Processing 2003

- Spectral data have not the property of cross-correlation required by SOBI
- Their inverse TF possesses this property
- The separation of spectral data can be done by the following operations IFT-SOBI-TF
- The algorithm *f*-SOBI, based on the FT of the correlation coefficients, allow to group these 3 steps in only one operation

Conclusion on RMN spectroscopy

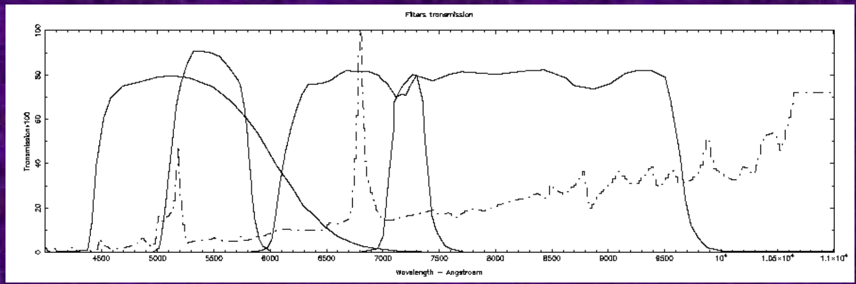
- Varied examples : Molecule structure, analysis of a solely mixture and of many maixtures,
- Take into account of constraints
  - Frequential instabilities due to
    - The instrumentation,
    - The variation of the concentration of the chemicals,
  - Positivity of the sources and of the mixtures.
- Separation in the Fourier space of data
  - Development of *f*-SOBI
    - A few information in the direct space
    - Correlation more important in the Fourier space
    - This is an alternative to SOBI for the choice of the correlation space
    - Prove its efficientcy with others applications

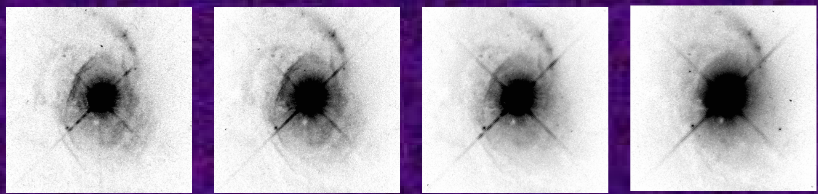
3C120 Radio-source composition

D. Nuzillard, A. Bijoui (CReSTIC, Observatoire de Nice)

- Multispectral anlysis in astronomy
  - Different wavelengths
    - ↳ Different phenomena
- Independent physical phenomena
  - Independent sources
    - ↳ Extraction of physical parameters through models.

- 3C120 Radiosource
- Observations HST, WFPC2
  - F547M (V1): OIII + Continuum
  - F555W (V): F547M + Continuum
  - F675W (R): Hα + Continuum
  - F814W (I): weak rays + Continuum



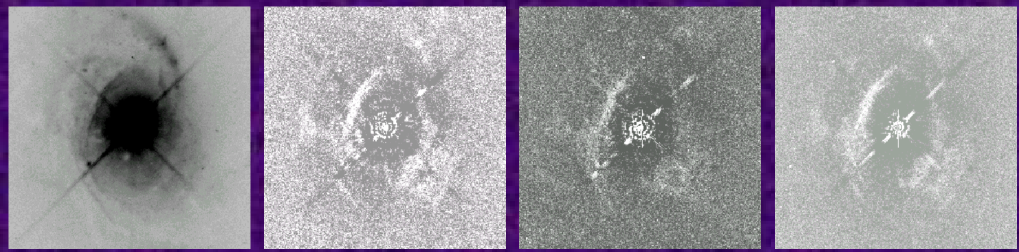


Noisy images ↳ photon noise

3C120 images Separations

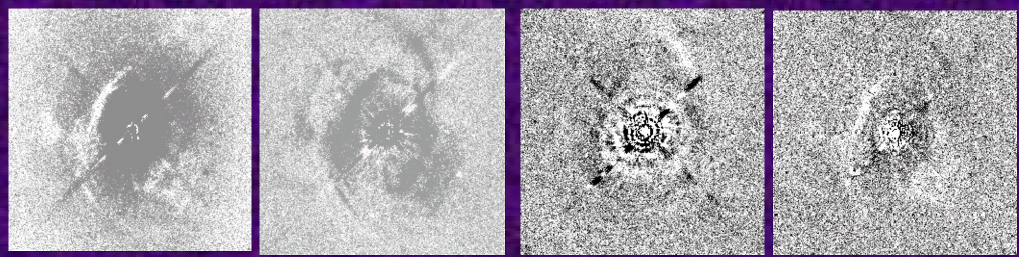
- Karhunen-Loève Transformation

Energy concentration in the first eigen vectors



- Best separation with FastICA [th(y)]

Optimize a local criterion of independance



SAS et spatial Correlations

Cross-correlations space

SOBI : Data space

f-SOBI : Fourier space

f-SOBI 2-D : Fourier space, signals 2-D

SOBI 2-D : Data space, signals 2-D

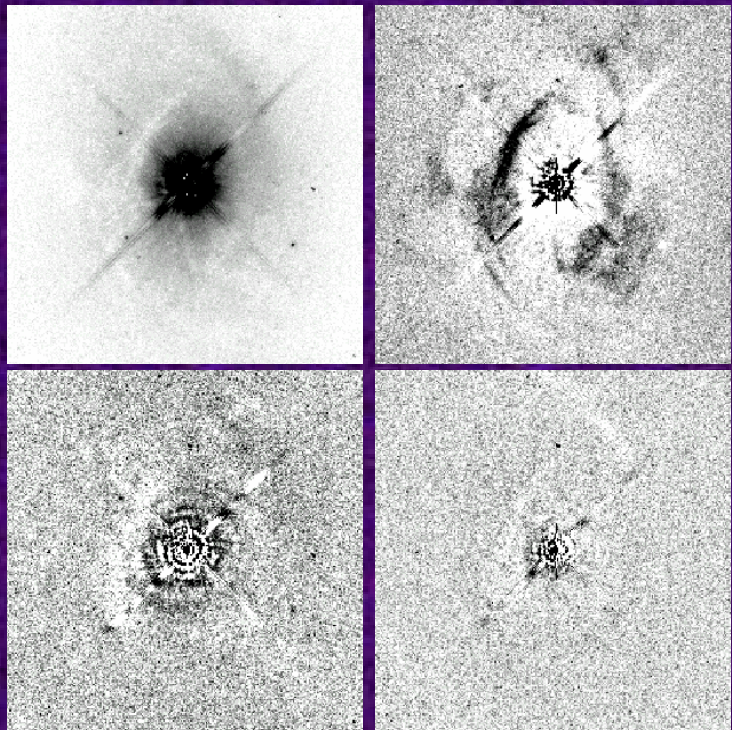
IM Selection : *f*-SOBI-16

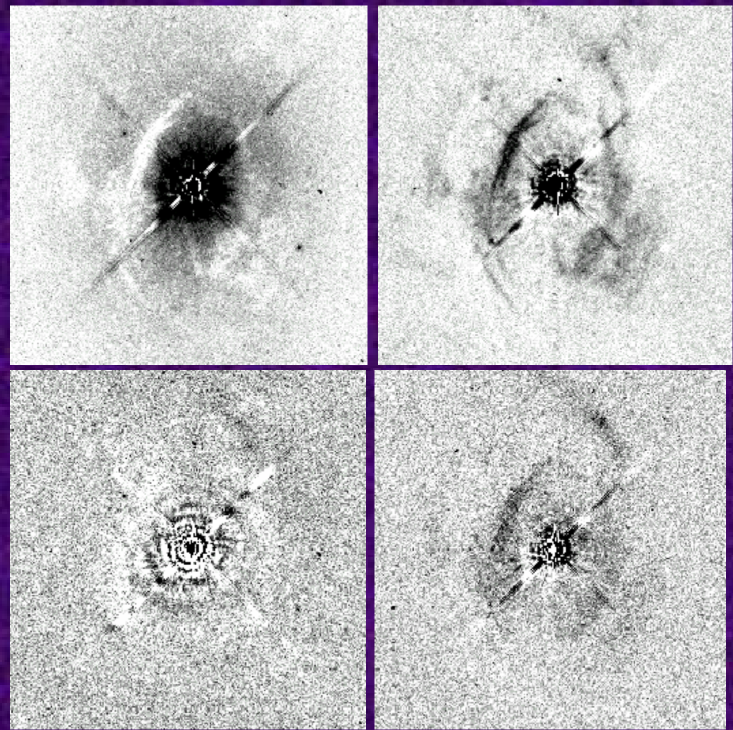
- Observed Mutual Information

$$I(S_1, \ldots, S_I) = \sum_I E(S_I) - E(X_1, \ldots, X_n) - log_2 |\det B|$$

- Empirical estimation by a couple of sources, the weakest

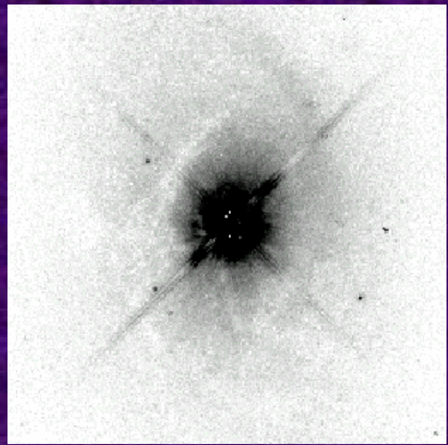
Visual selection : SOBI 2D-8

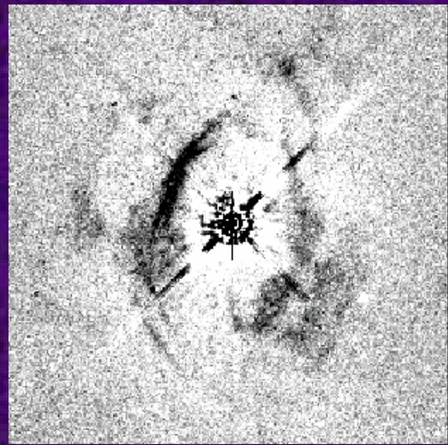


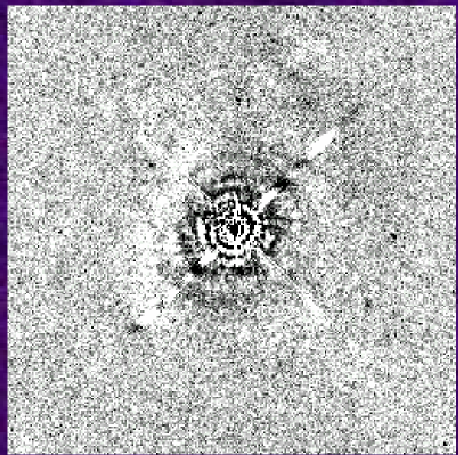


Sources Interpretation

- Source 1 :
  - Central region
- Source 2 :
  - Ionised areas around the kernel
  - OIII Rays play the most essential role
- Source 3 :
  - Rings due to the spread fnction in the center in the Hα ray







Model with 2 components:

- A kernel very brillant
- A gazeous area

Explanatory tools adapted to the very big data bases of images (*Data Mining)* *hyperspectral* observations

Improvement of the vision in diffusant media

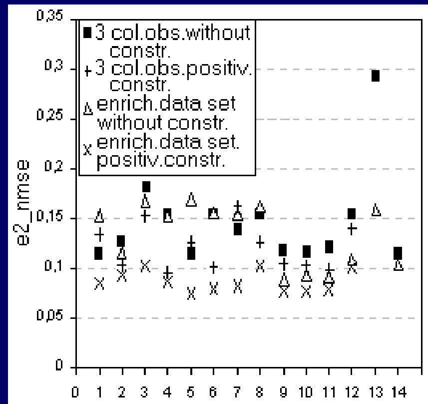
S. Curila, A. Elhafid, M. Curila, D. Nuzillard, J. Padet

Non uniform perturbation

Original image

Extracted Source

Source obtained with additive information

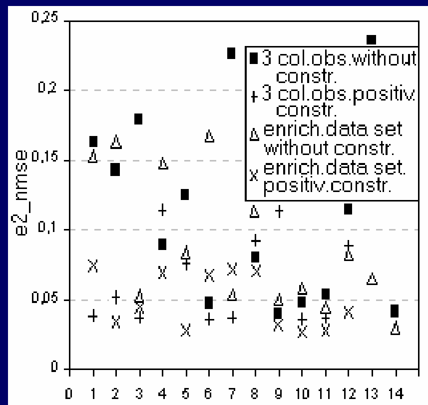


Uniform Perturbation

Original Image

Extrated Source

Source obtained with additive information



Legend

FastICA 1-8

1.defl,pow3

2.defl,tanh

3.defl,gauss

4.defl,skew

5.symm,pow3

6.symm,tanh

7.symm,gauss

8.symm,skew

9.SOBI

10.*f-SOBI*

11.*f-SOBI*2-D

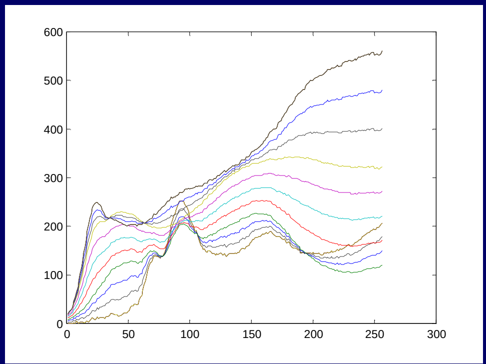
12.SOBI2-D

13.NNSC

14.NMF

A new possibility for analysing series of electron energy loss spectra.

Noël Bonnet and Danielle Nuzillard

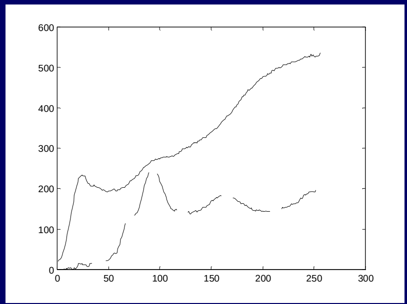


- The two spectra far from the interface Si-SiO2 are known and a third unknown component is suspected close to the interface
- The aim is to infer the shape of the unknown spectrum and to deduce the variation of composition across the interface, i.e. the weights of the different spectra in the mixture.

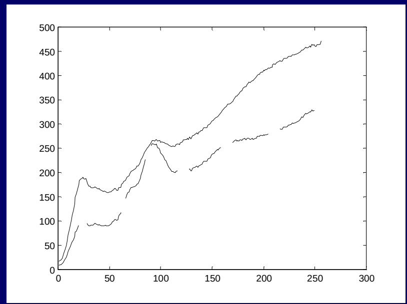
A new possibility for analysing series of electron energy loss spectra.

Noël Bonnet and Danielle Nuzillard

Simple demonstration

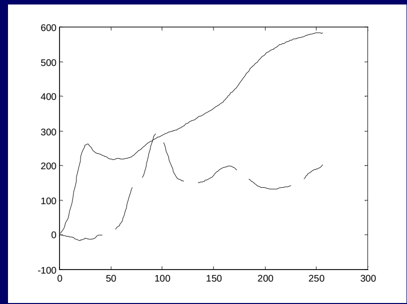


Two electron energy-loss spectra



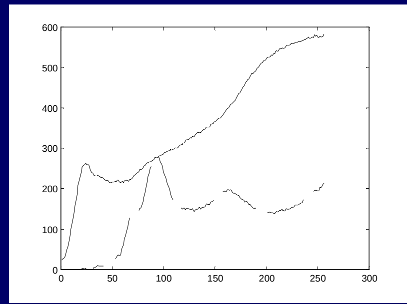
Two mixtures of the spectra

$$\begin{pmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{pmatrix}$$



Recovery of the two spectra, without applying constraints

$$\begin{pmatrix} 0.78 & 0.22 \\ 0.45 & 0.55 \end{pmatrix}$$



Recovery of the two spectra, with application of a positivity constraint.

$$\begin{pmatrix} 0.75 & 0.25 \\ 0.41 & 0.59 \end{pmatrix}$$

- The result could not be obtained from the raw spectra, but only from the derivative spectra
- EEL spectra are not composed of separated peaks. Raw EEL spectra do not fulfil the conditions for blind separation. They are not even uncorrelated. On the other hand, derivative spectra fulfil these conditions.

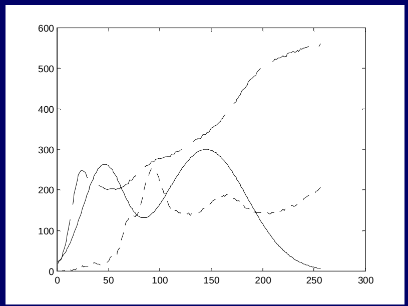
•Blind separation was thus performed with the derivative spectra and the resulting spectra were then integrated. This is not a problem for subsequent quantification since derivation is a linear process.

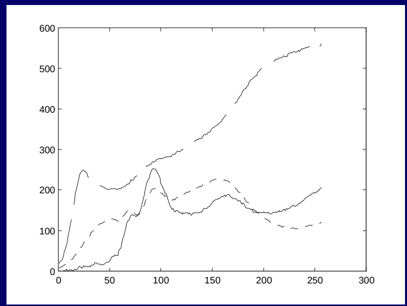
$$\begin{pmatrix} 0.78 & 0.22 \\ 0.45 & 0.55 \end{pmatrix}$$

A new possibility for analysing series of electron energy loss spectra.

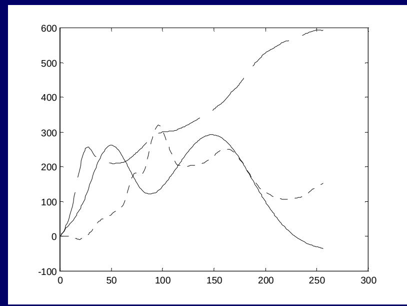
Noël Bonnet and Danielle Nuzillard

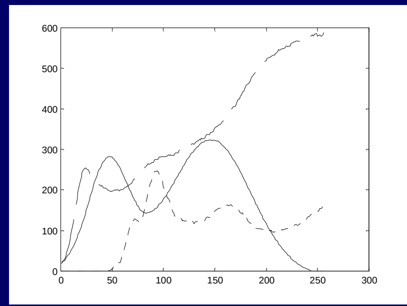
A three-component mixture





$$\begin{pmatrix} 1.0 & 0.0 & 0.0 \\ 0.58 & 0.42 & 0.0 \\ 0.0 & 0.0 & 1.0 \end{pmatrix}$$



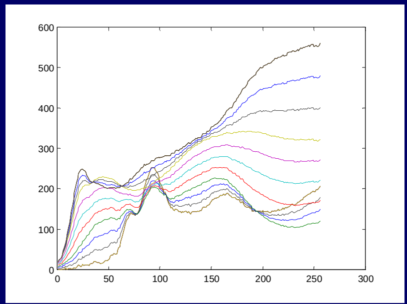


$$\begin{pmatrix} 0.90 & 0.0 & 0.10 \\ 0.58 & 0.34 & 0.08 \\ 0.0 & 0.04 & 0.96 \end{pmatrix}$$

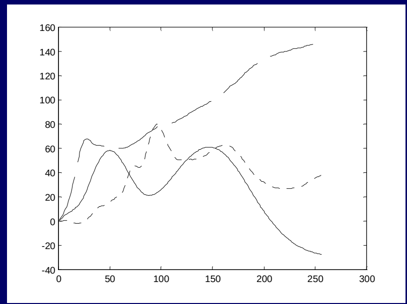
- The two spectra far from the interface are known and a third unknown component is suspected close to the interface
- The aim is to infer the shape of the unknown spectrum and to deduce the variation of composition across the interface, i.e. the weights of the different spectra in the mixture.

A new possibility for analysing series of electron energy loss spectra.

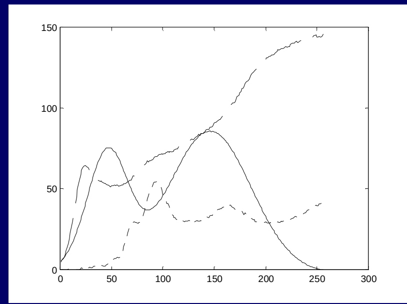
Simulation



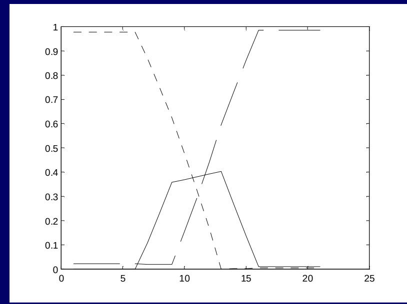
21 spectra (mixtures) used to recover the unknown spectrum



three spectra recovered without the positivity constraint

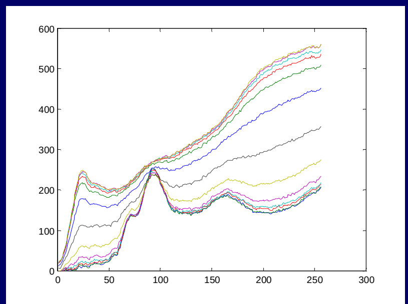


spectra recovered when using the positivity constraint

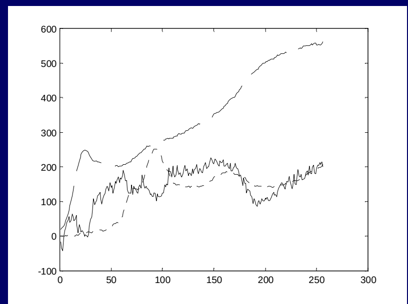


the weight of the three spectra across the interface

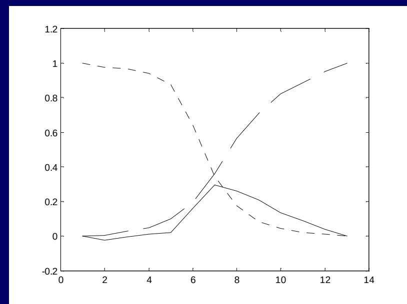
A series of 13 electron energy-loss spectra recorded across a Si-SiO2 interface, for an energy loss between 99 and 124 eV recorded at Orsay [Brun *et al.*, 1996].



13 electron energy-loss spectra



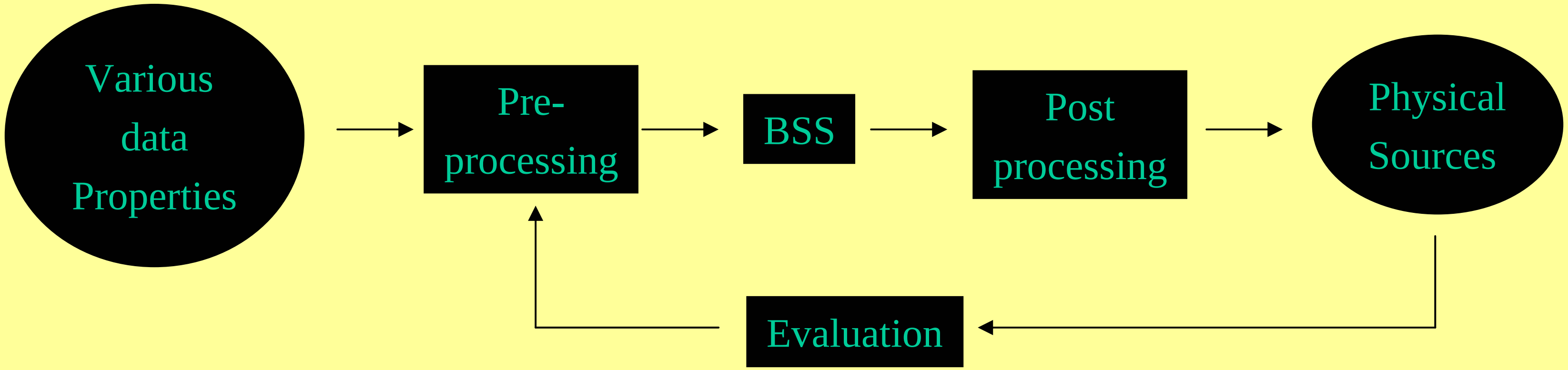
three expected components



weight of these three components across the interface

Conclusion

Scheme



- Information diversity
- Choice of space representation
  - Augmentation
  - Selection
    - Reduction / Compression
    - Denoising
- Evaluation

- Tools
  - Multi-resolution Analysis
  - Time-frequency Analysis
  - Classification / fusion
  - Others …