HyperPCA

At the Interface of Random Matrix Theory and Laser-Induced Breakdown Spectroscopy

Riccardo Finotello *CEA Paris-Saclay (DIASI/SIALV/LVML)*

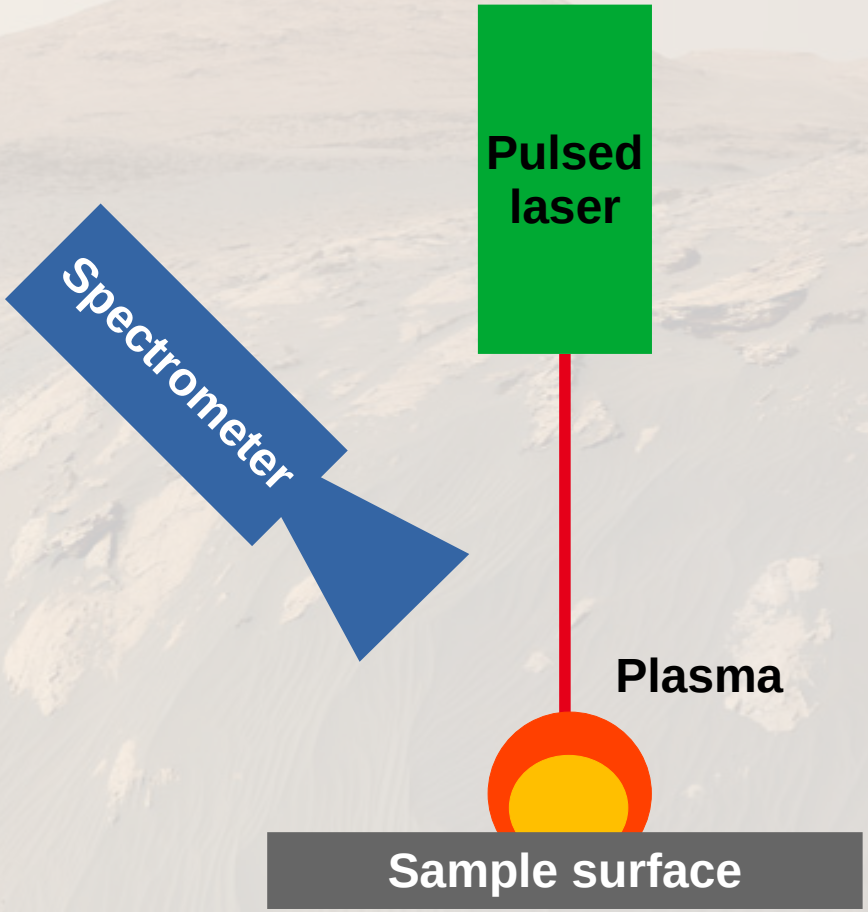
(loosely) based on Spectrochim. Acta B, 192 (2022), 106418

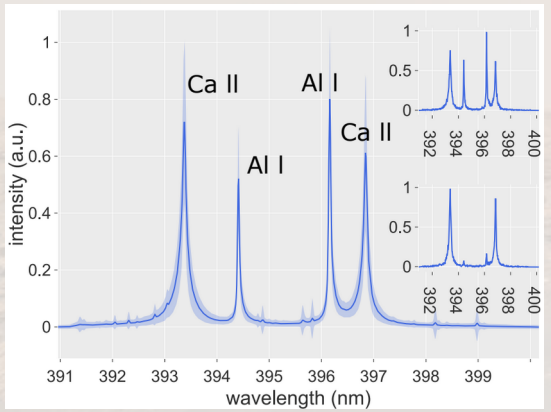
8e colloque de la SFPT-GH – Paris, 5-6 July 2023

1. Laser-Induced Breakdown Spectroscopy

The principle and a simple analysis

Laser-Induced Breakdown Spectroscopy

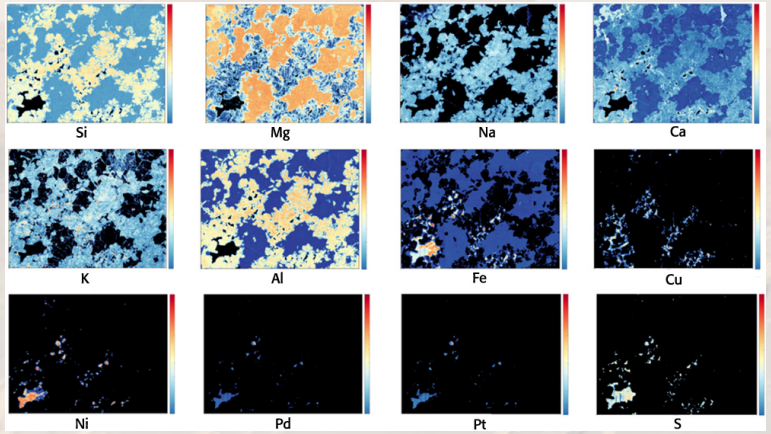




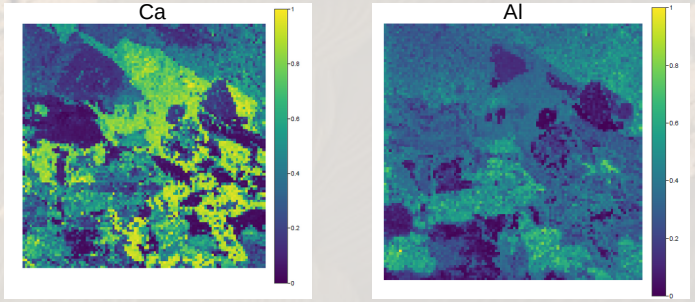
Advantages of **LIBS**:

- **Fast** measurements
- **Ambient** air
- **Light** elements (up to H)
- **High** resolution

LIBS Mapping



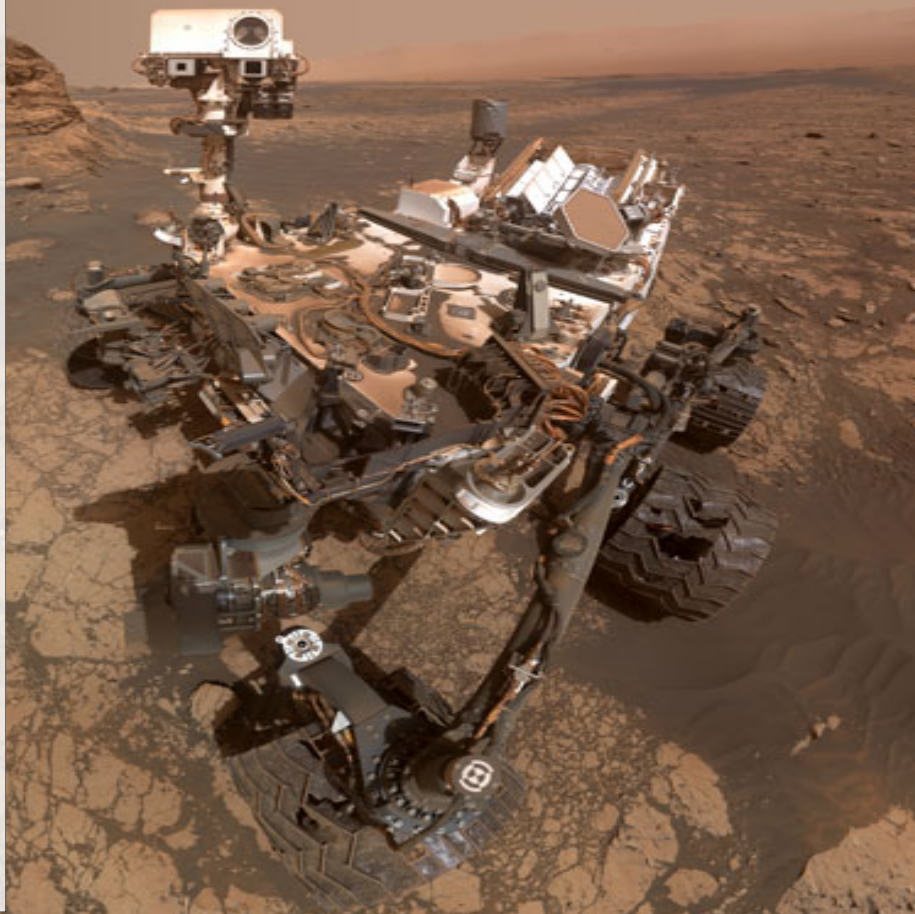
Senesi and Harmon. Spectroscopy Europe/World Vol. 33 Issue 4 (2021)



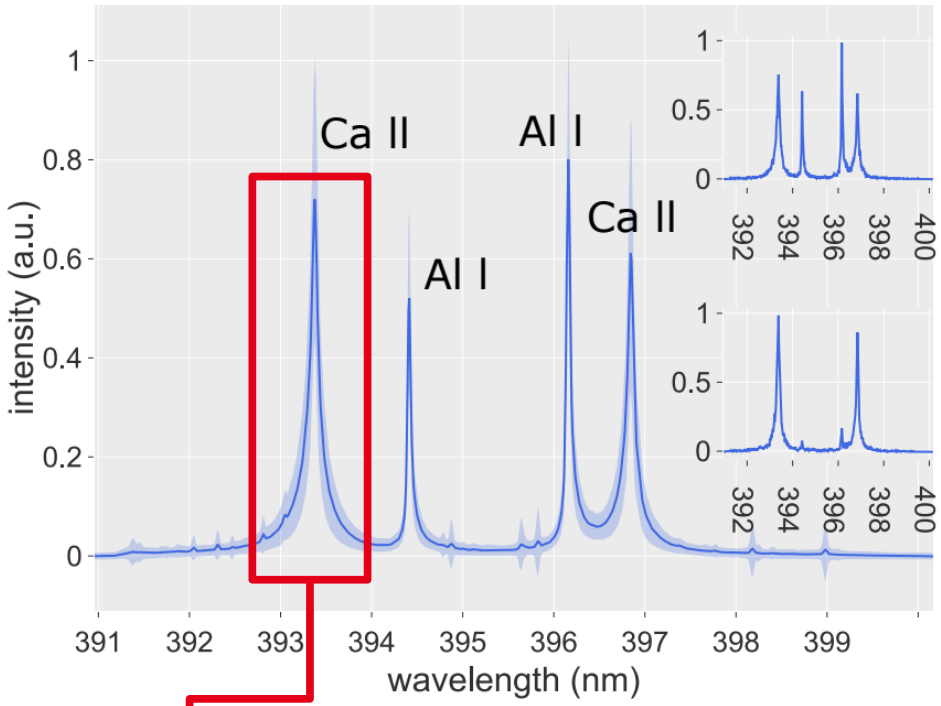
Finotello et al. Spectrochim. Acta B: At. Spectrosc., 192 (2022), 106418

LIBS **mapping**:

- **Resolution** <10 µm
- Ability to cover **large** surfaces *(very recent: 4K images with 1 kHz laser, see Harrel et al. (2023))*
- **Elemental** chemical maps
- **Minimal** sample preparation
- **Qualitative** and **quantitative** analyses
- Applications to geology, space sciences, biology, etc.

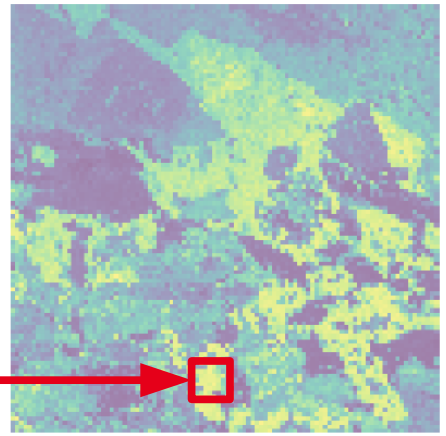


Standard Line Intensity Analysis



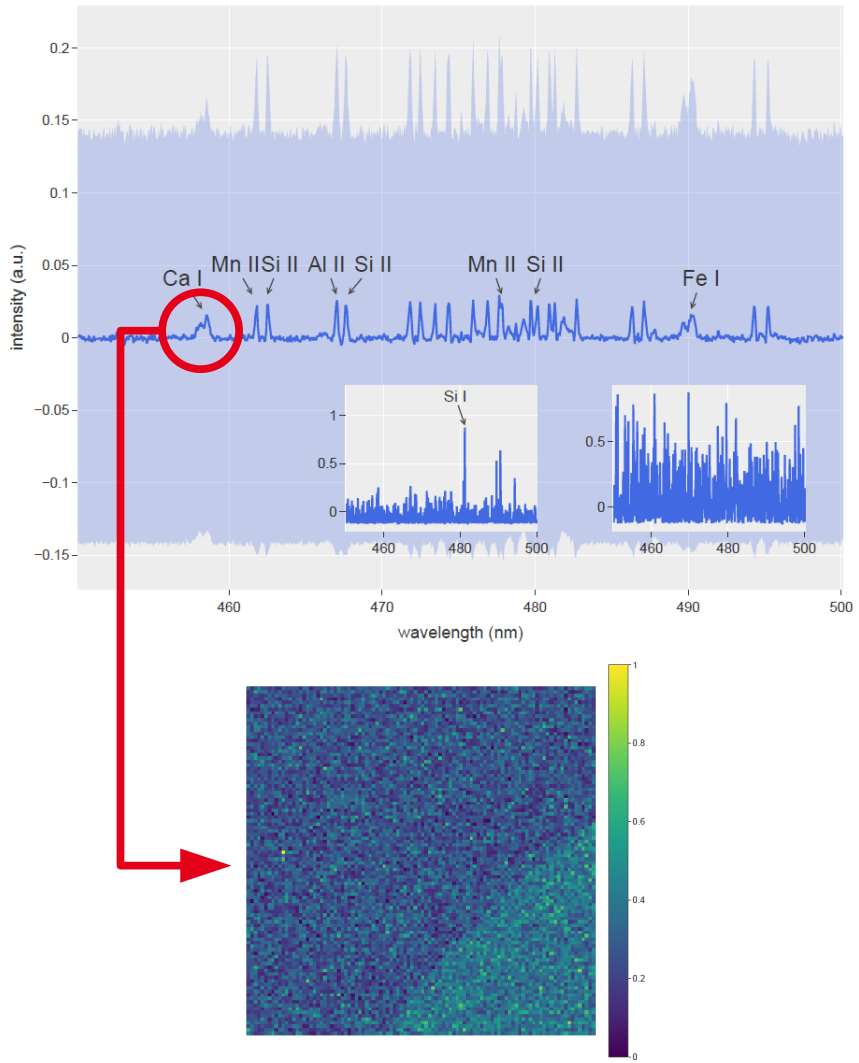
- Plasma radiation **characteristic** of **elements** in the sample
- Laser shot emission **spectrum** characterizing elements in a **pixel**
- Known **emission lines** (at given wavelengths) 2D **elemental maps**

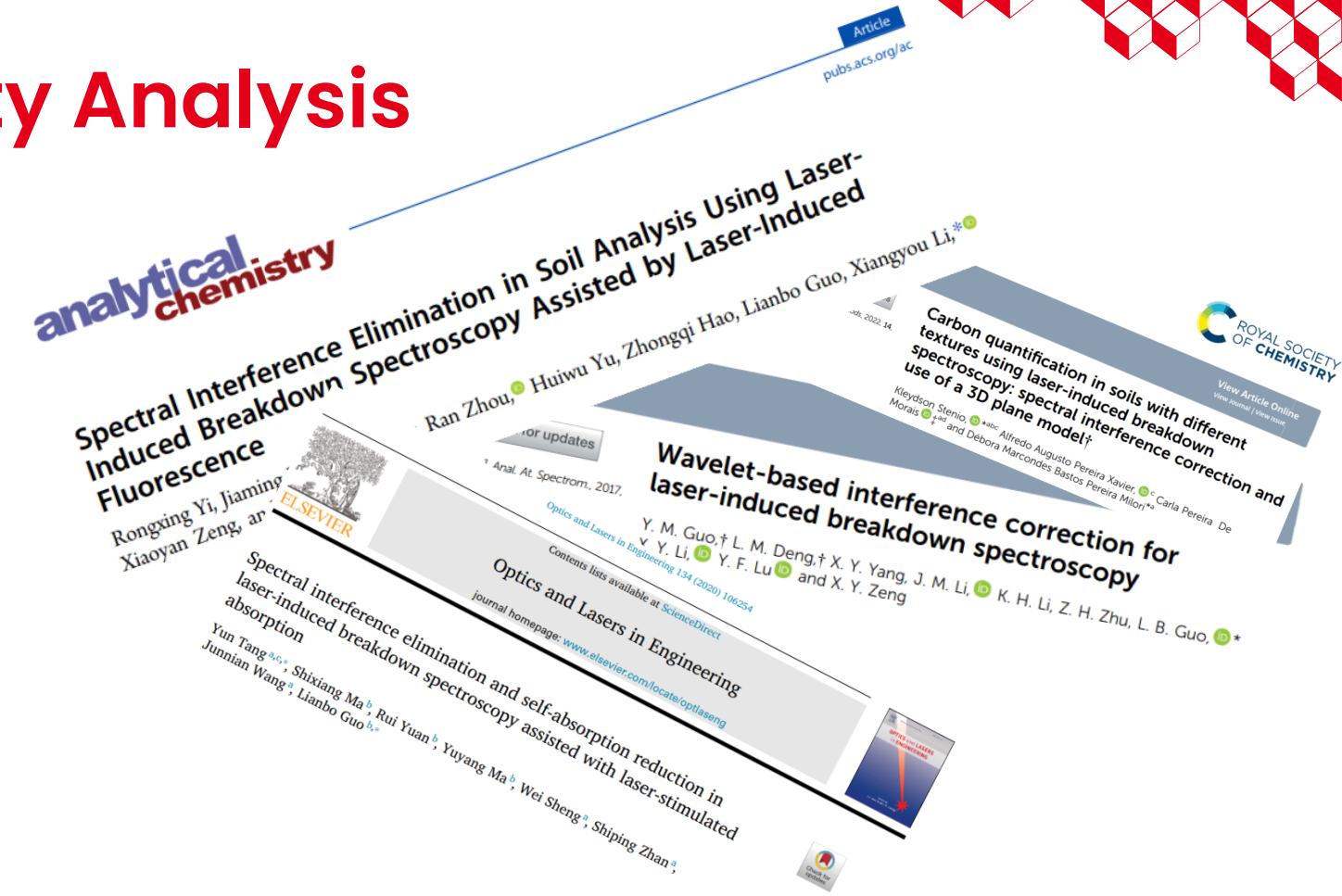
Intensity



MAPPINGS ARE ASSOCIATED TO WAVELENGTHS!

Standard Line Intensity Analysis



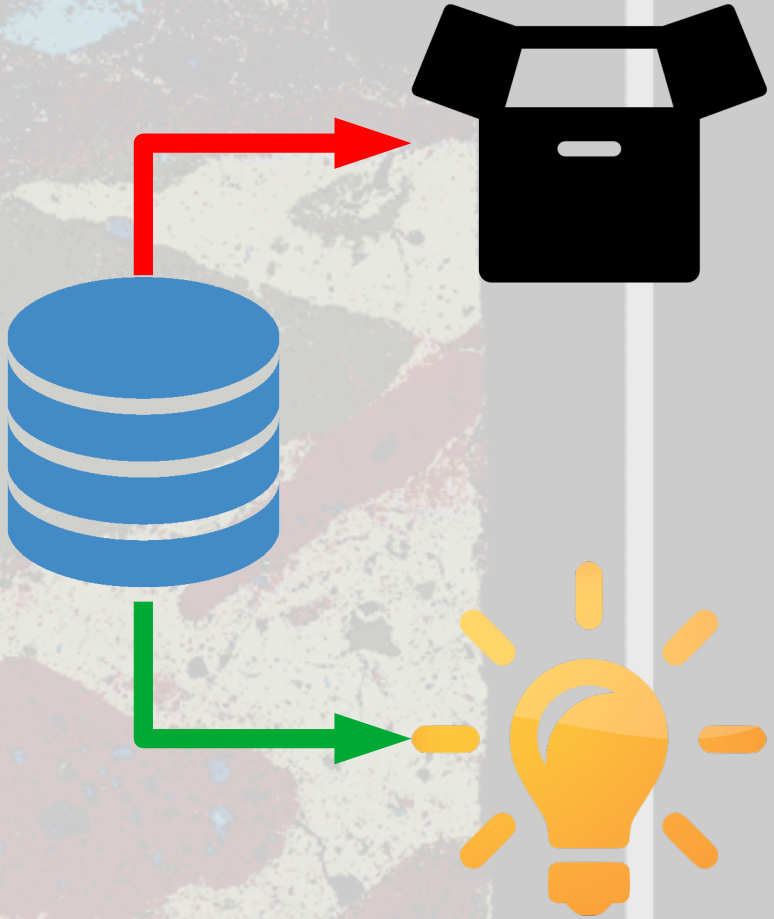


- **Spectral interference** is a longstanding problem in elemental spectral
- Element discovery can be difficult due to the specific **expertise required**

2. Multivariate Analysis and Random Matrices

Towards an advanced Principal Components Analysis

Extracting LIBS Information



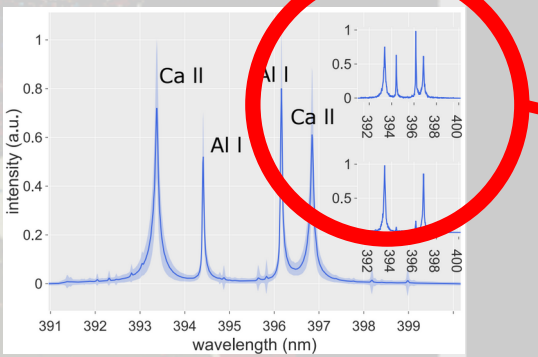
Black box approaches (CNNs, ViTs, etc.)

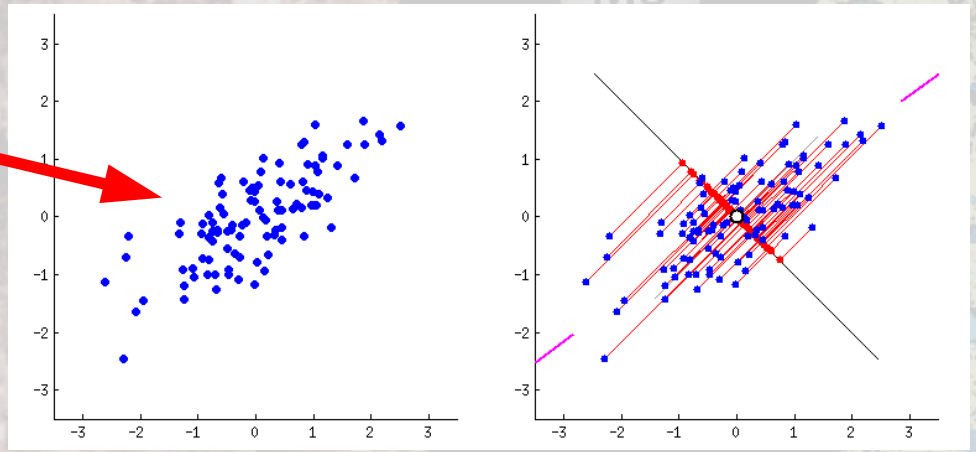
Interpretable statistical model

- Use **more information** contained in spectra
- Consider data as **compact structure** (image) rather than single pixel
- Avoid complicated models: **streamline** complexity and computation
- **Reduce expertise** required to interpret the results

Background: Herreyre et al. J. Anal. At. Spectrom., 2023, 38, 730-741

Principal Components Analysis





- Find **new representation**
- **Minimize** reconstruction error

$$W_{PCA} = \arg \min_{W^T W = \mathbb{I}} \left\{ ||X' - X\,W||_2^2 \right\}$$

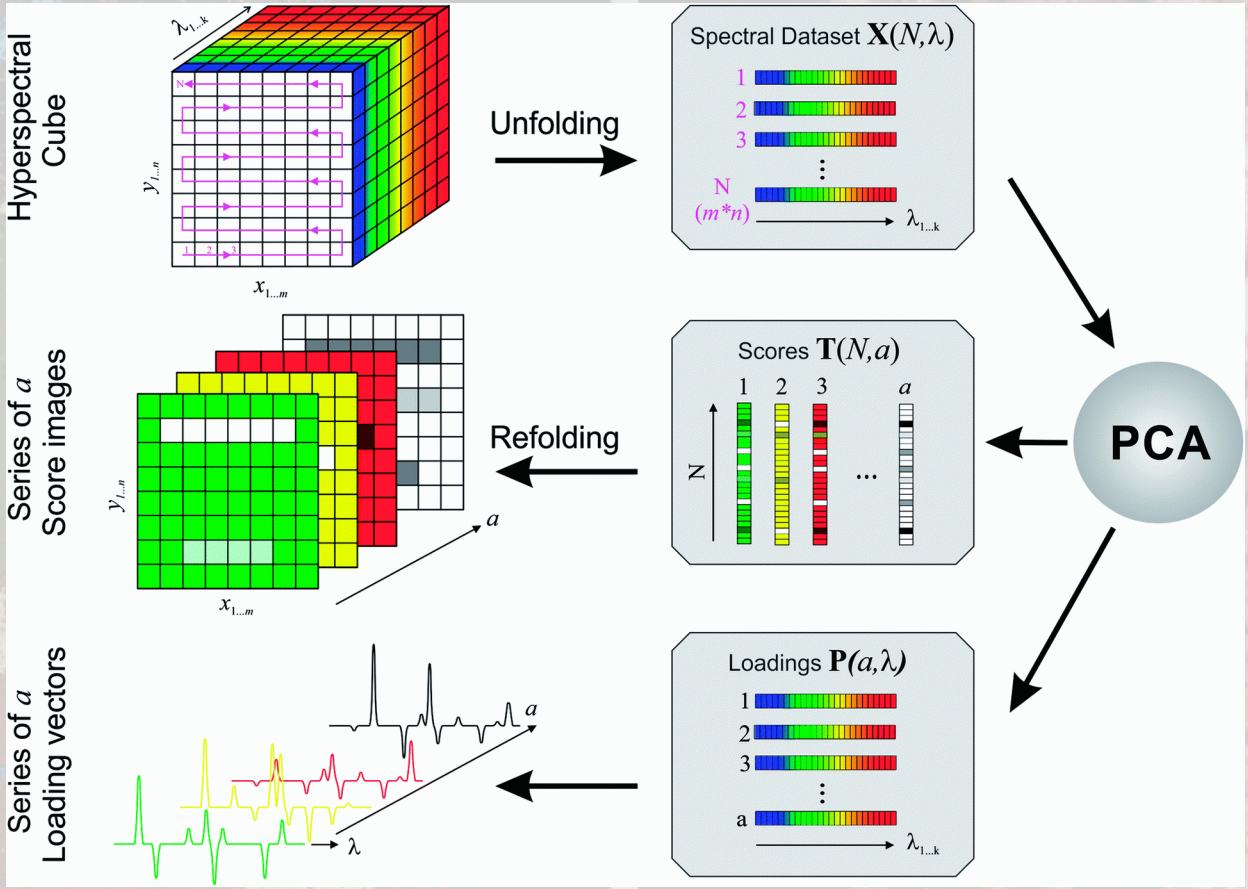
- **Maximize** variance → find signal

$$w_{(1)} = \arg \max_{w^T w = 1} \left\{ \sum_{i=1}^{n} (x_{(i)} \cdot w)^2 \right\}$$

$$\text{PCA} : \mathbb{R}^{n \times p} \to \mathbb{R}^{n \times p}$$
$$X \mapsto X' = X\,W$$

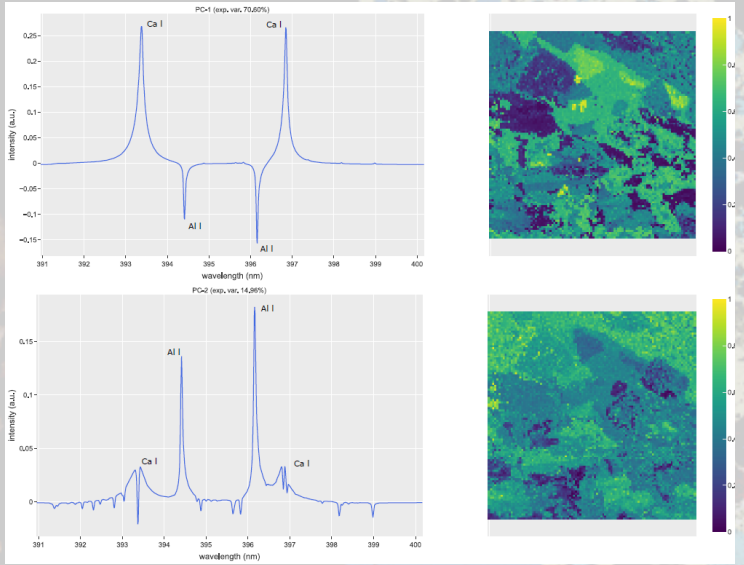
Background: Herreyre et al. J. Anal. At. Spectrom., 2023, 38, 730-741

Principal Components Analysis in LIBS



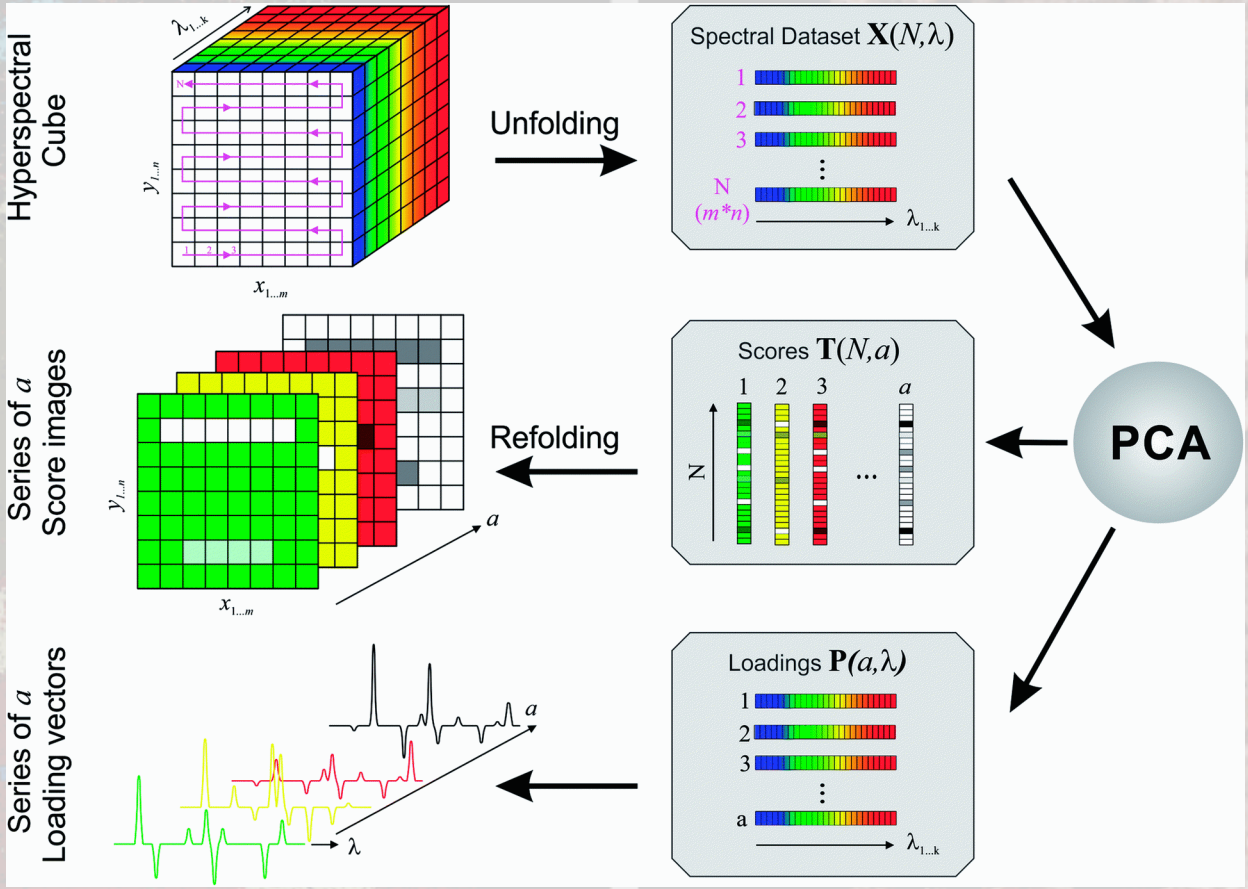
Moncayo et al. J. Anal. At. Spectrom., 2018,33, 210-220

- **Scores** can be refolded to **images** after projection on the same subspace of all subpixels
- **Loadings** are **fingerprints** of the associated 2D map, representing the coefficients of the "rotation" of the components



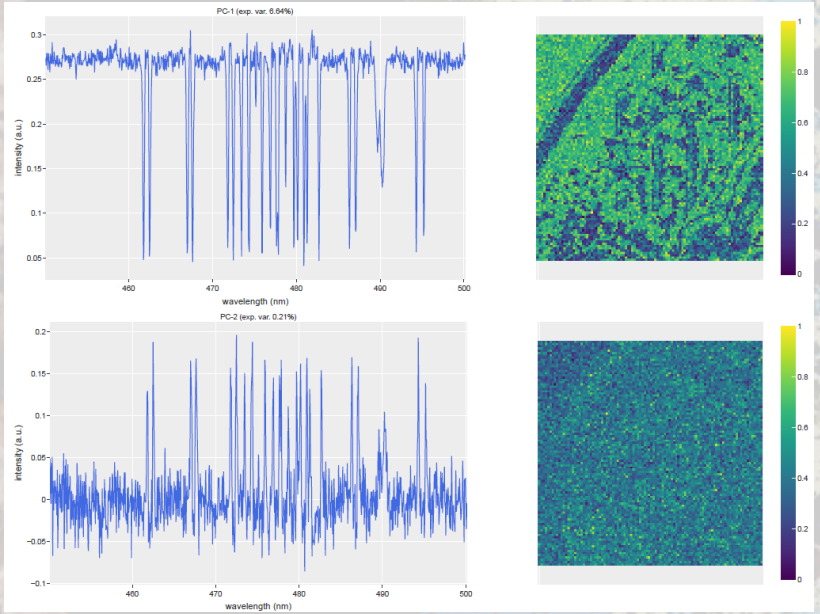
Background: Herreyre et al. J. Anal. At. Spectrom., 2023, 38, 730-741

Principal Components Analysis in LIBS



Moncayo et al. J. Anal. At. Spectrom., 2018,33, 210-220

- Highly sensitive to **noise** and spectral **interference**
- Loadings may have a **difficult interpretation**



Background: Herreyre et al. J. Anal. At. Spectrom., 2023, 38, 730-741

A Random Matrix Approach

$$Y = \Sigma^{1/2} Z = (\mathbb{I} + X)^{1/2} Z = \overbrace{\left(\mathbb{I} + \sum_{i=1}^{k} \beta_k u_i u_i^T \right)^{1/2}}^{\text{signal}} \underbrace{Z}_{\text{noise}}$$

PCA: $Y^T Y \longrightarrow f(Y^T Y) \simeq f(\Sigma) + R(\Sigma, Z)$

A particular case → **Σ is sparse** 

$$R(\Sigma, Z) = \mathcal{O}(n^{-\alpha})$$
$$f(0) = f'(0) = f''(0) = 0$$

We can **asymptotically** work with the **sparse** signal!

Seddik et al. ICLR 2019

3. Back to LIBS

Working with (hyper)spectral images

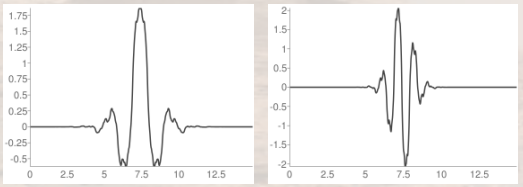
HyperPCA for LIBS

HYPERPCA

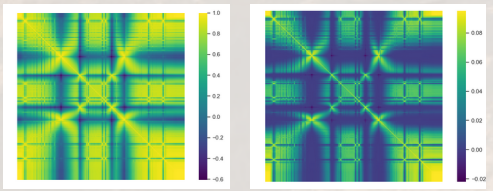
1 **SIGNAL SPARSIFICATION**

2 **Kernel Function**

3 **Principal Components**

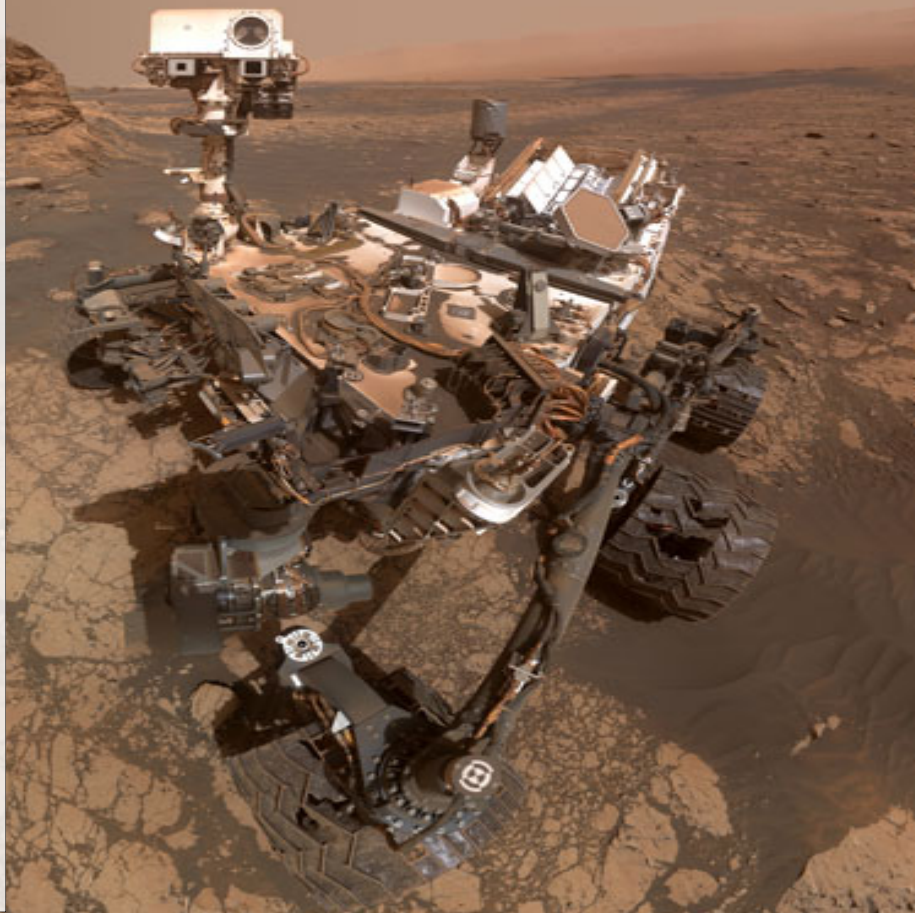


Wavelet decomposition



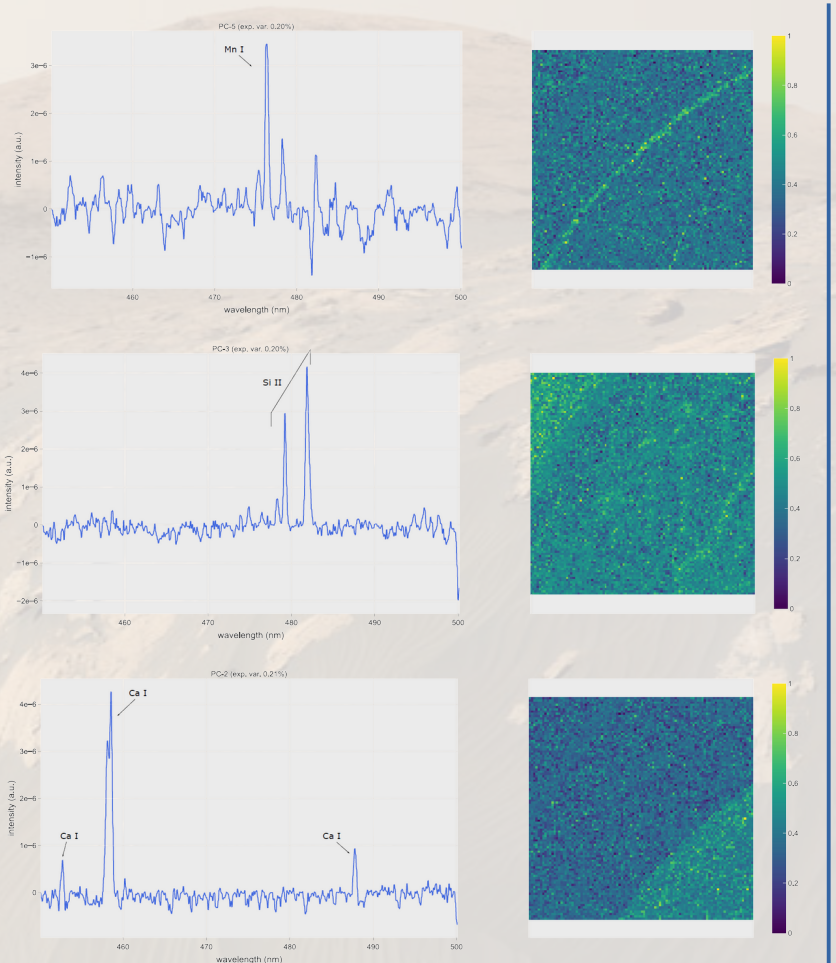
$$f_a(t) = t\left(1 - e^{-at^2}\right)$$

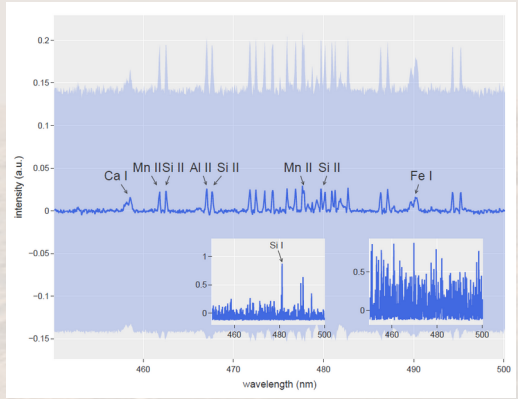
A traditional PCA with the newly defined covariance matrix



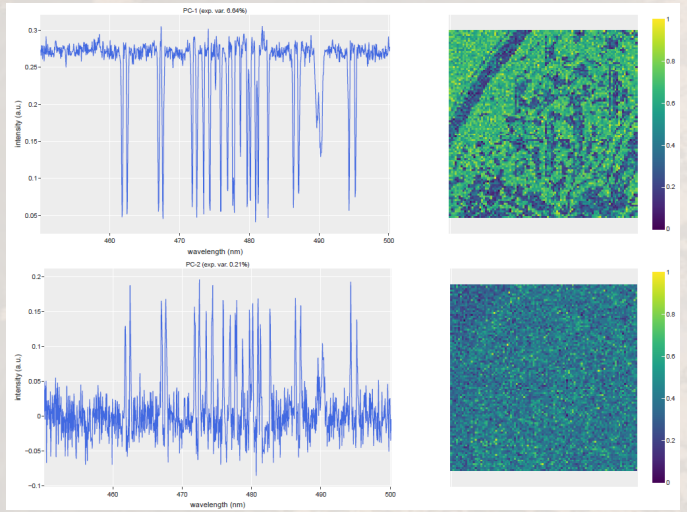
Finotello et al. Spectrochim. Acta B: At. Spectrosc., 192 (2022), 106418

HyperPCA for LIBS

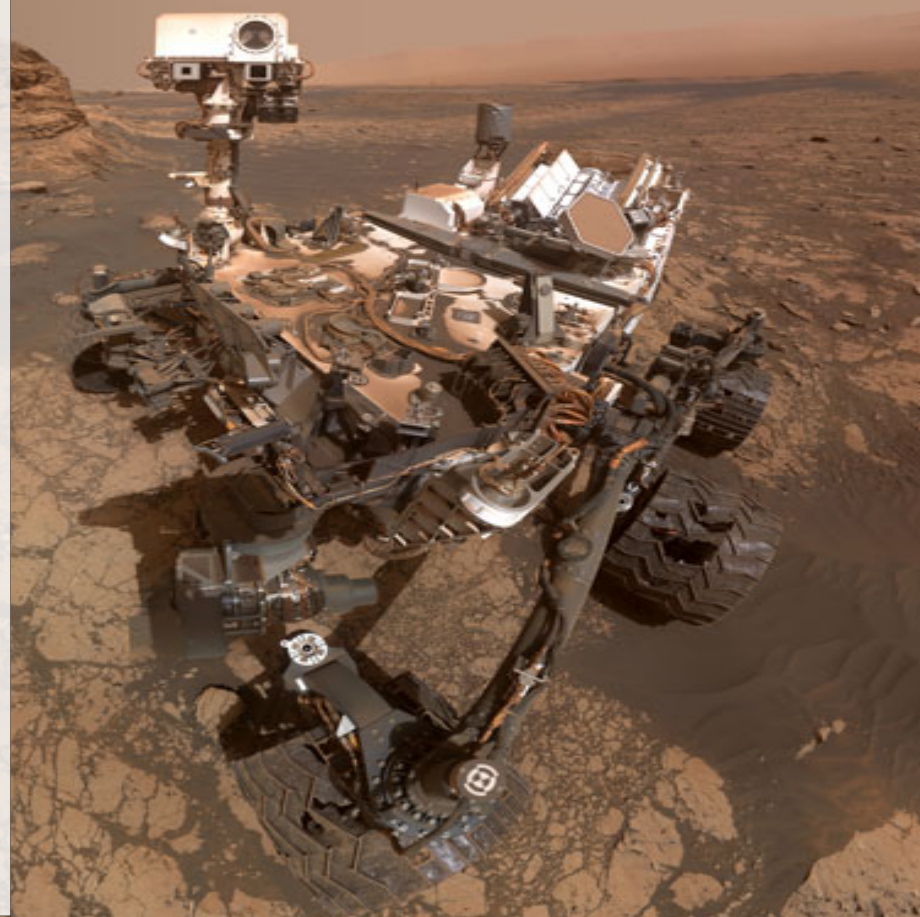




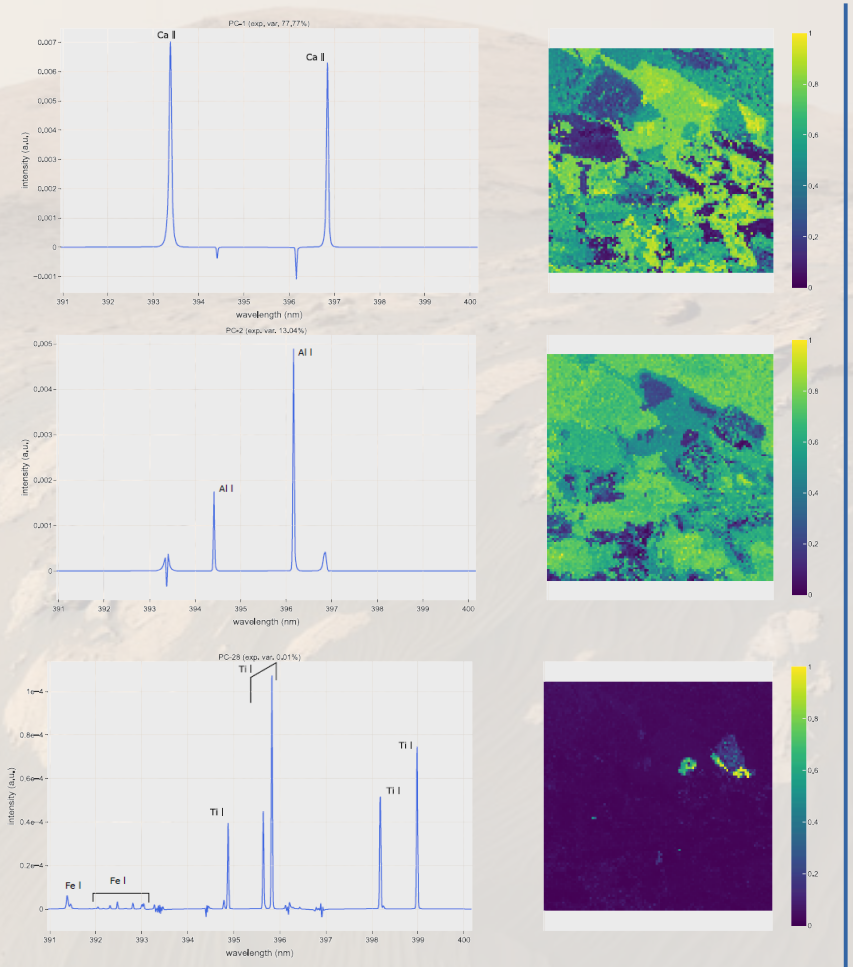
Standard PCA



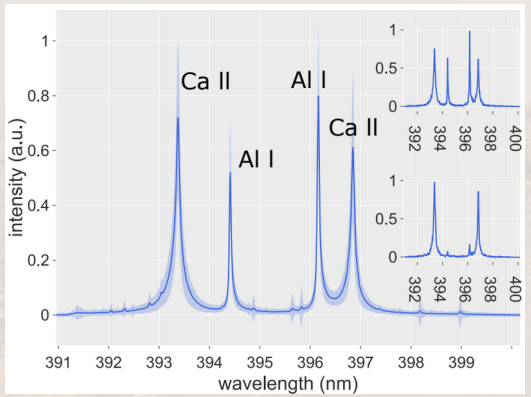
Finotello et al. Spectrochim. Acta B: At. Spectrosc., 192 (2022), 106418



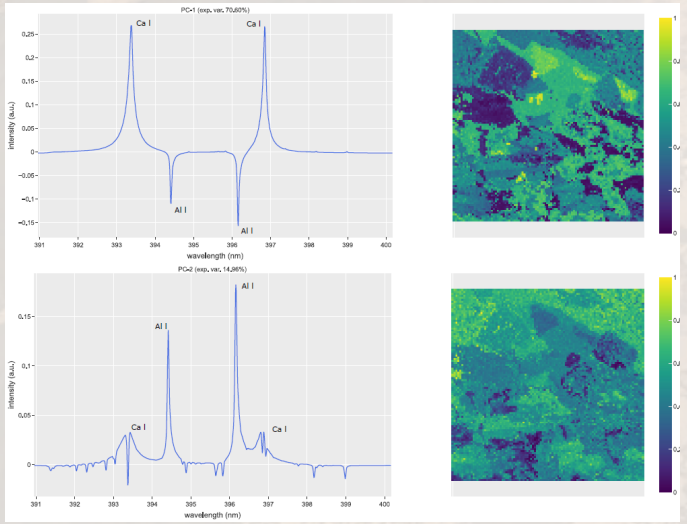
HyperPCA for LIBS



Finotello et al. Spectrochim. Acta B: At. Spectrosc., 192 (2022), 106418

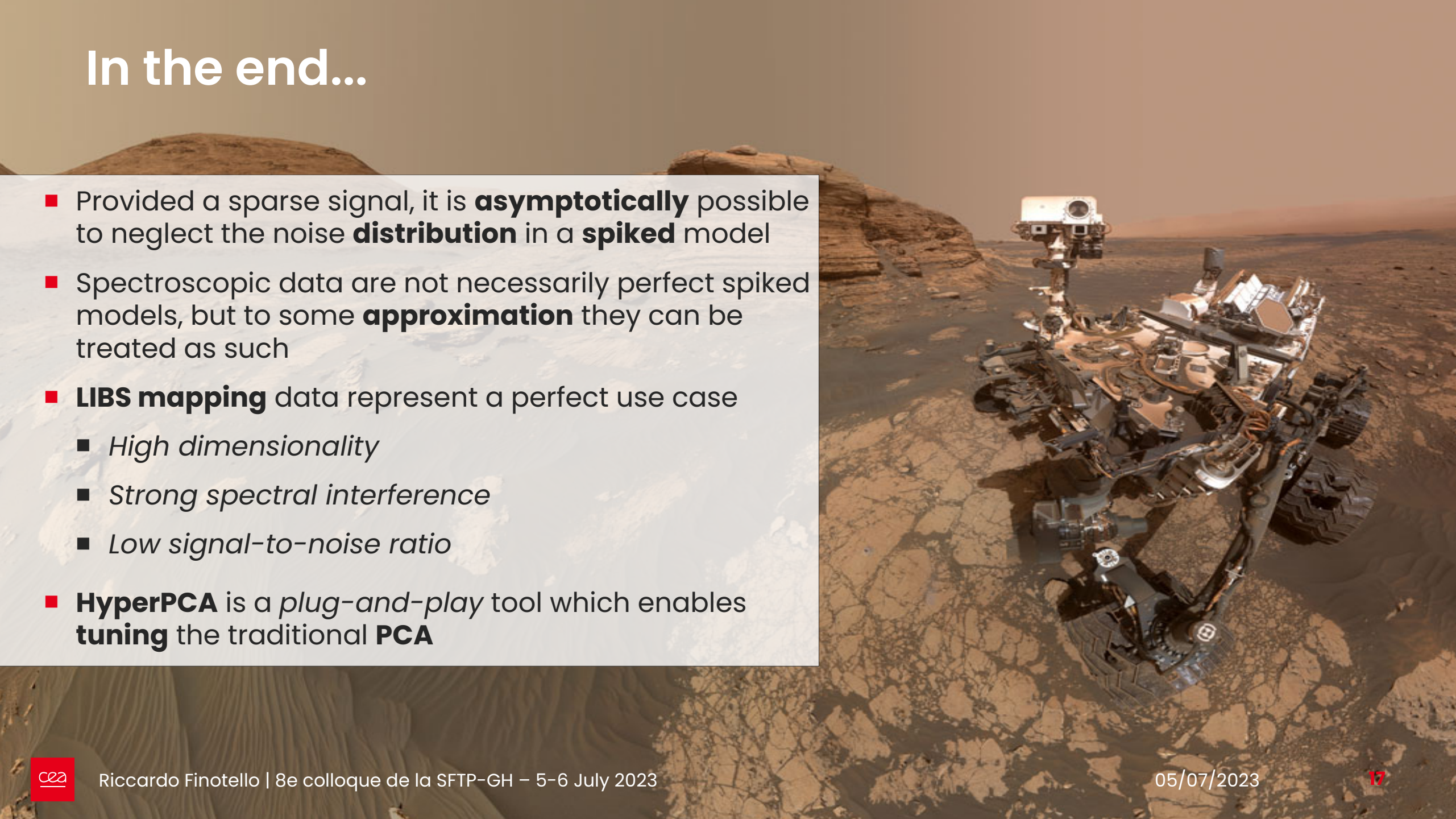


Standard PCA



In the end...

- Provided a sparse signal, it is **asymptotically** possible to neglect the noise **distribution** in a **spiked** model
- Spectroscopic data are not necessarily perfect spiked models, but to some **approximation** they can be treated as such
- **LIBS mapping** data represent a perfect use case
  - *High dimensionality*
  - *Strong spectral interference*
  - *Low signal-to-noise ratio*
- **HyperPCA** is a *plug-and-play* tool which enables **tuning** the traditional **PCA**



HyperPCA

At the Interface of Random Matrix Theory and Laser-Induced Breakdown Spectroscopy

Riccardo FINOTELLO (DRT/LIST/DIASI/SIALV/LVML | riccardo.finotello@cea.fr)

In collaboration with:

M. Tamaazousti (DRT/LIST/DIASI/SIALV/LVML) and J.-B. Sirven (DES/ISAS/DPC/SPC/LANIE)