Robust Target Detection for Hyperspectral Imaging

Post-Doc Seminar

Joana Frontera-Pons
Laboratoire CosmoStat
Office 278

Working under the supervision of Jérôme Bobin and Florent Sureau


MALLORCA, SPAIN

Robust Target Detection for Hyperspectral Imaging

under the supervision of
J.-P. Ovarlez and F. Pascal

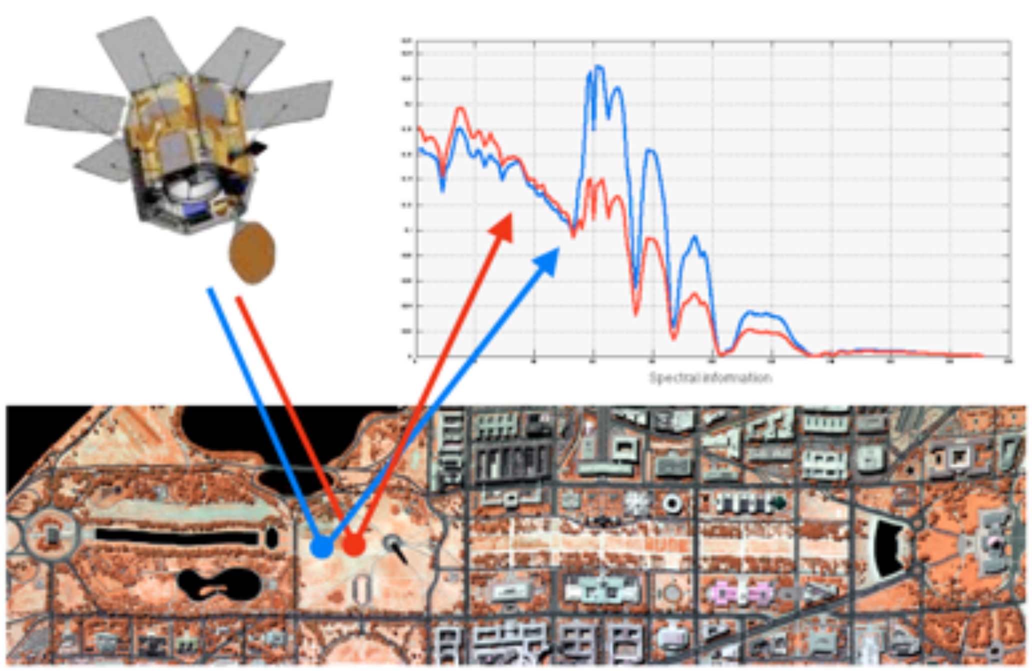
Problems Description

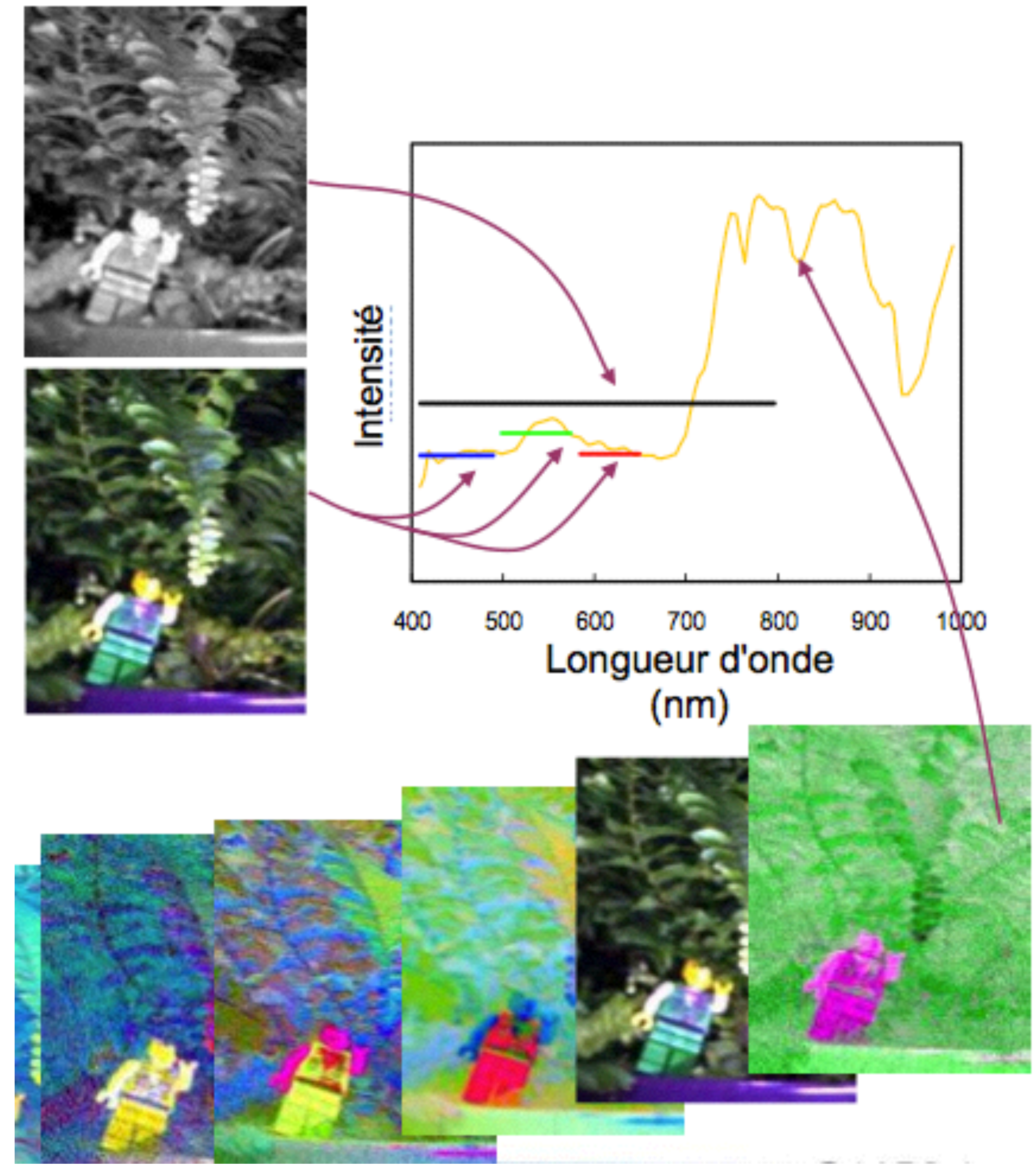
- Hyperspectral Imaging : Imaging Spectometer performing spectral measurements,
- Spectometer : Optical instrument (sensor) used for measuring the quantity of electromagenetic radiation over one wavelength or more,
- How? - Measurement over narrow contiguous bands in the optical domain creating a continous spectrum,
- Why? - Characterization of different materials with differents spectral signatures for:
  - Target detection and Change detection,
  - Determine the different types of materials which compose the image: Unmixing,
  - Classification.

Find the artificial plant and the hidden LEGO

Classes of Spectral Images

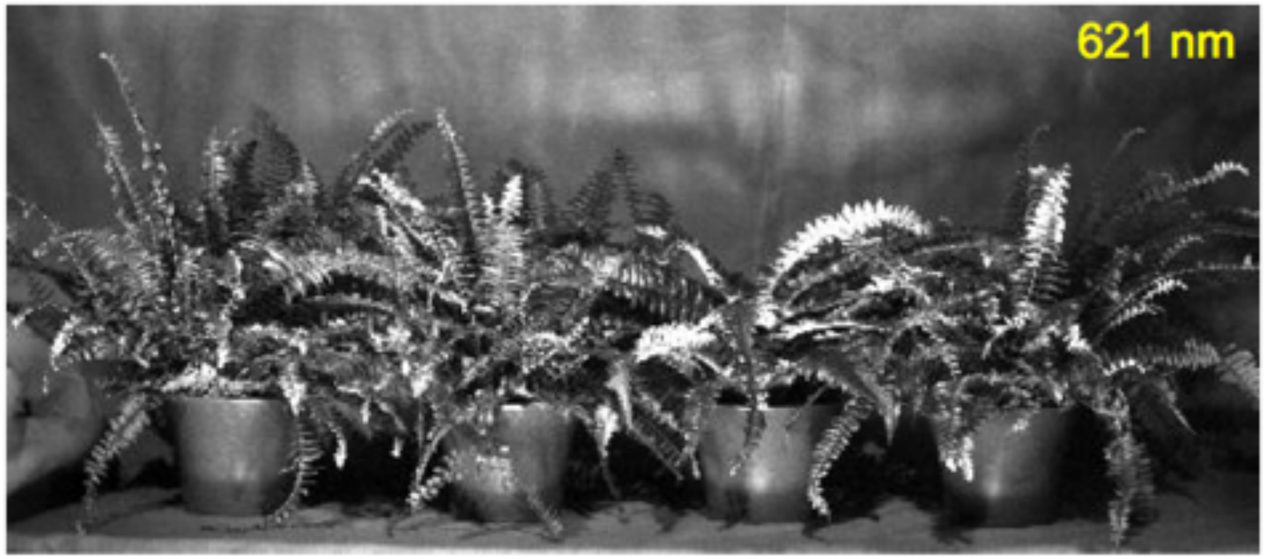
- <u>Monochromatic</u> : grayscale image without spectral information.
- <u>Multispectral</u> : frrom 2 to 10 bands, limited spectral information,
- <u>Hyperspectral</u> : from 10 to 100 narrow and contigous bands, detailed spectral information.

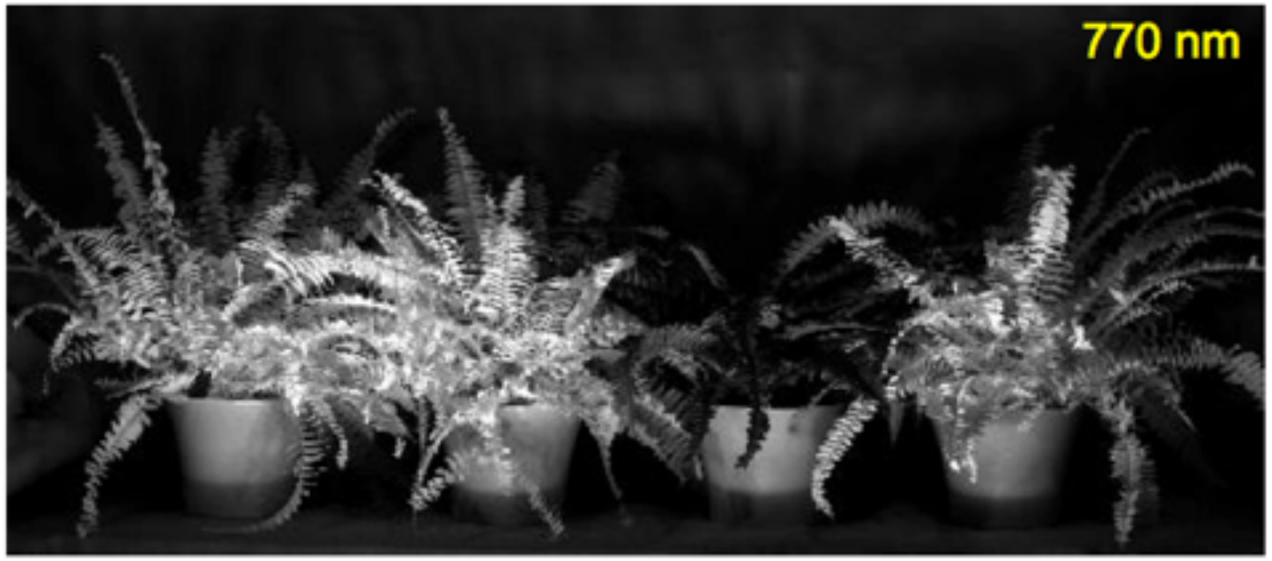




BANDS SELECTION

- How to analyze Hypspectral Images?
  - Visual Inspection : Slow process because the user sees only three bands at once,
  - Bands Selection : Using hyperspectral imagery to select a few spectral bands adapted to the problem to solve.

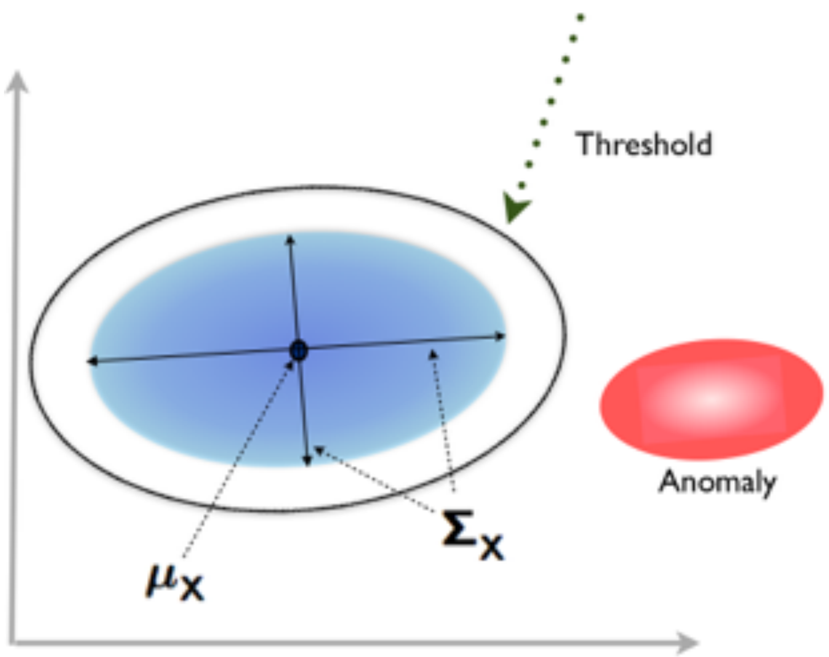


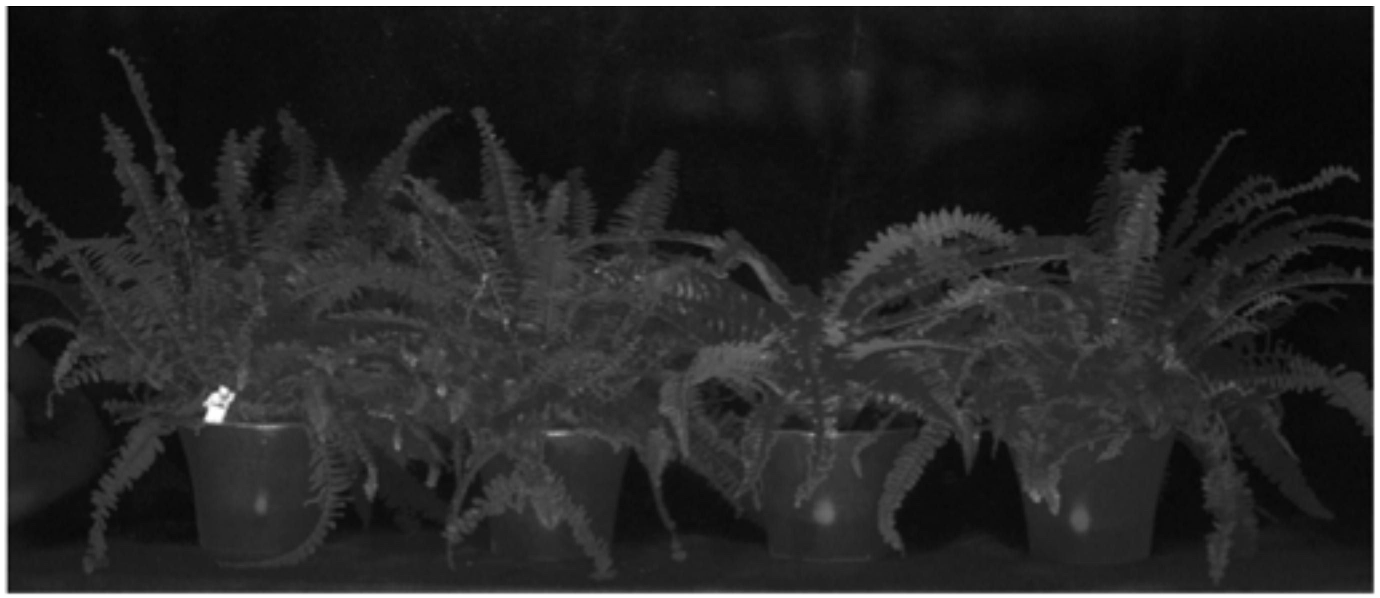


HYPERSPECTRAL TARGET DETECTION

- **ANOMALY DETECTION IN HYPERSPECTRAL IMAGES**

  To detect all that is « different » from the background (*Mahalanobis distance*) - Application to radiance images.





- **DETECTION OF TARGETS IN HYPERSPECTRAL IMAGES**

  To detect (GLRT) targets (characterized by a given spectral signature *p*) - Regulation of False Alarm. Application to reflectance images (after some atmospherical corrections or others).

GAUSSIAN DISTRIBUTION

A m-dimensional vector $\mathbf{x}$ has a complex Gaussian distribution denoted $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. If the probability density function exists, it is of the form:

$$f_{\mathbf{x}}(\mathbf{x}) = \pi^{-m} |\boldsymbol{\Sigma}|^{-1} \exp\{-(\mathbf{x} - \boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\}.$$

Maximum Likelihood Estimators:

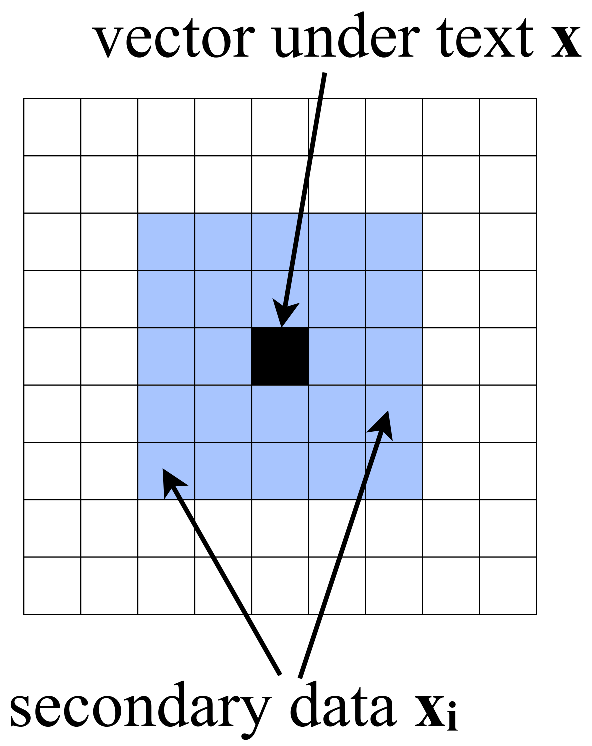
Let $\mathbf{x}_1, \ldots, \mathbf{x}_N$ be an IID N-sample, where $\mathbf{x}_i \sim \mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Thus, the SMV and the SCM can be written as:

$$\hat{\boldsymbol{\mu}}_{SMV} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{x}_i, \quad \hat{\boldsymbol{\Sigma}}_{SCM} = \frac{1}{N} \sum_{i=1}^{N} (\mathbf{x}_i - \hat{\boldsymbol{\mu}}_{SMV})(\mathbf{x}_i - \hat{\boldsymbol{\mu}}_{SMV})^H .$$

- Simplicity of analysis and well-known statistical properties: consistent, unbiased and efficient,
- $\hat{\boldsymbol{\Sigma}}_{SCM}$ is Wishart distributed and $\hat{\boldsymbol{\mu}}_{SMV}$ is Gaussian distributed.

Selection of the Secondary Data

- Rectangular Sliding window of size $n \times n$ moving all over the hyperspectral image,
- Size large enough to ensure the invertibility of the covariance matrix,
- and small enough to justify spatial homogeneity.

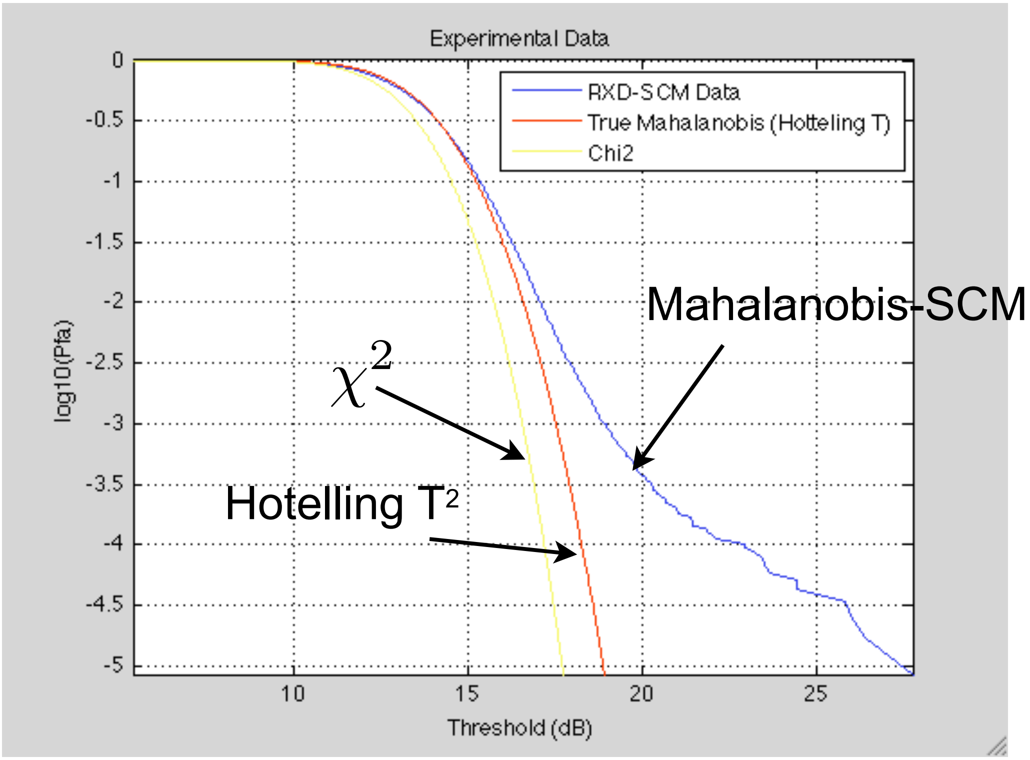


Assumptions

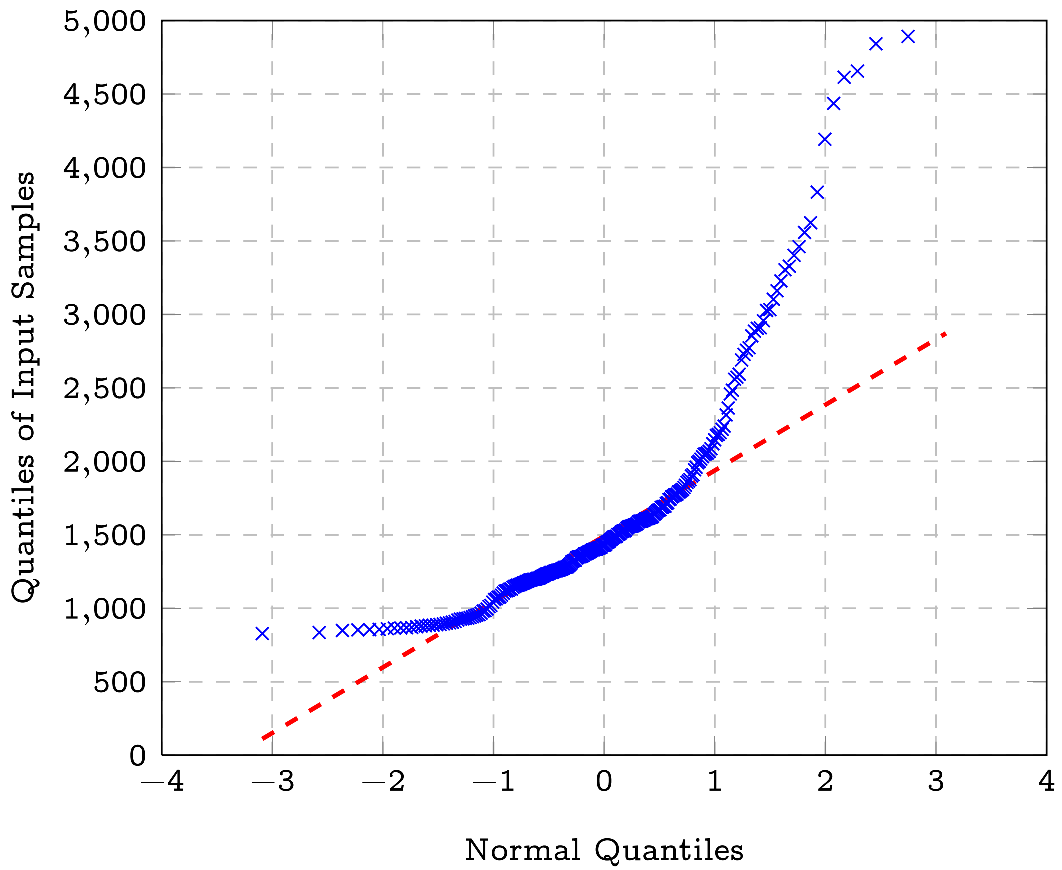
- Pixels of the mask are statistically independent, i.e. spatially independence.
- Pixels of the mask are identically distributed.

Motivation for Non-Gaussian Models

- Hyperspectral data are generally spatially heterogeneous in intensity and they cannot be only characterized by Gaussian distribution:

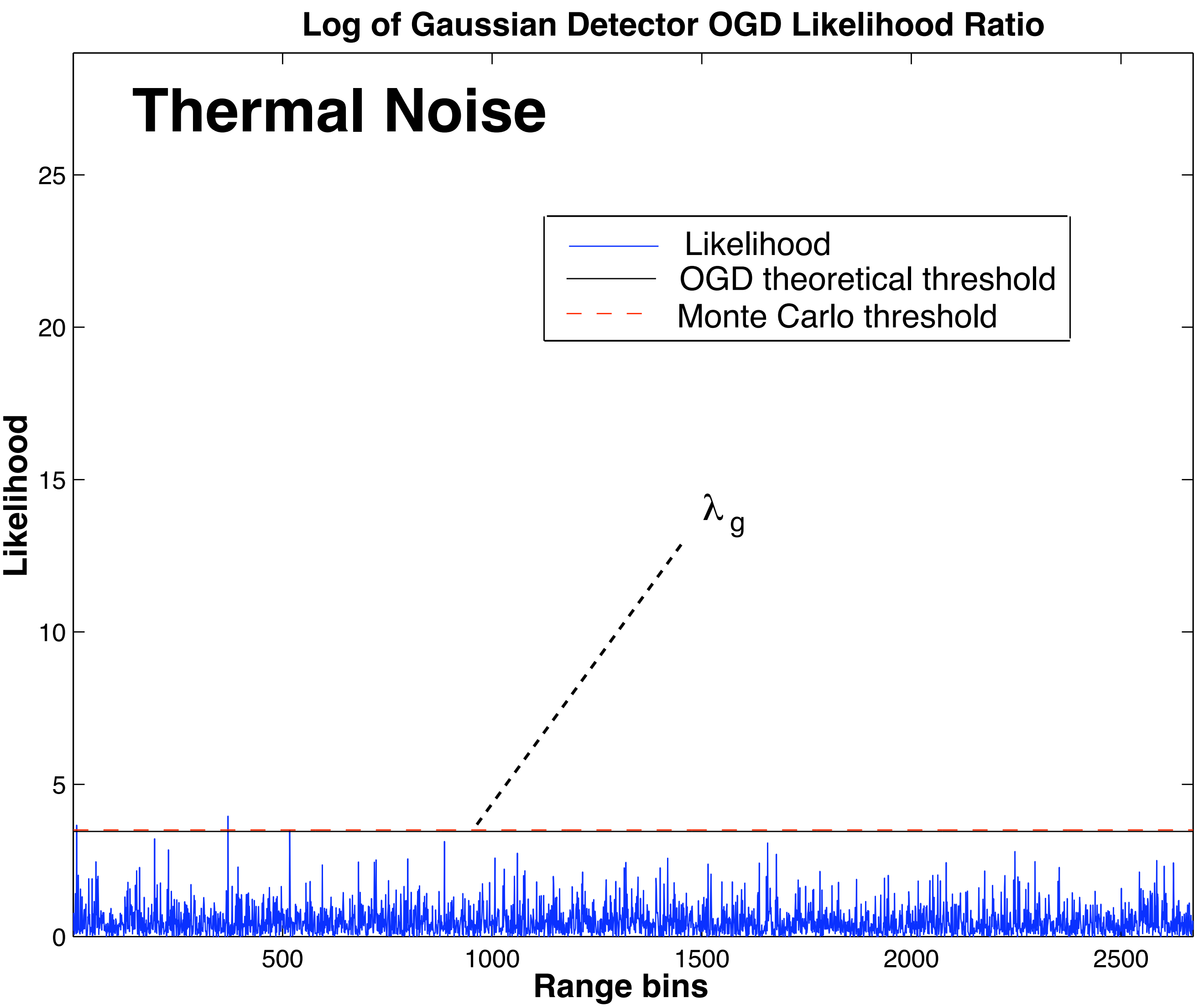


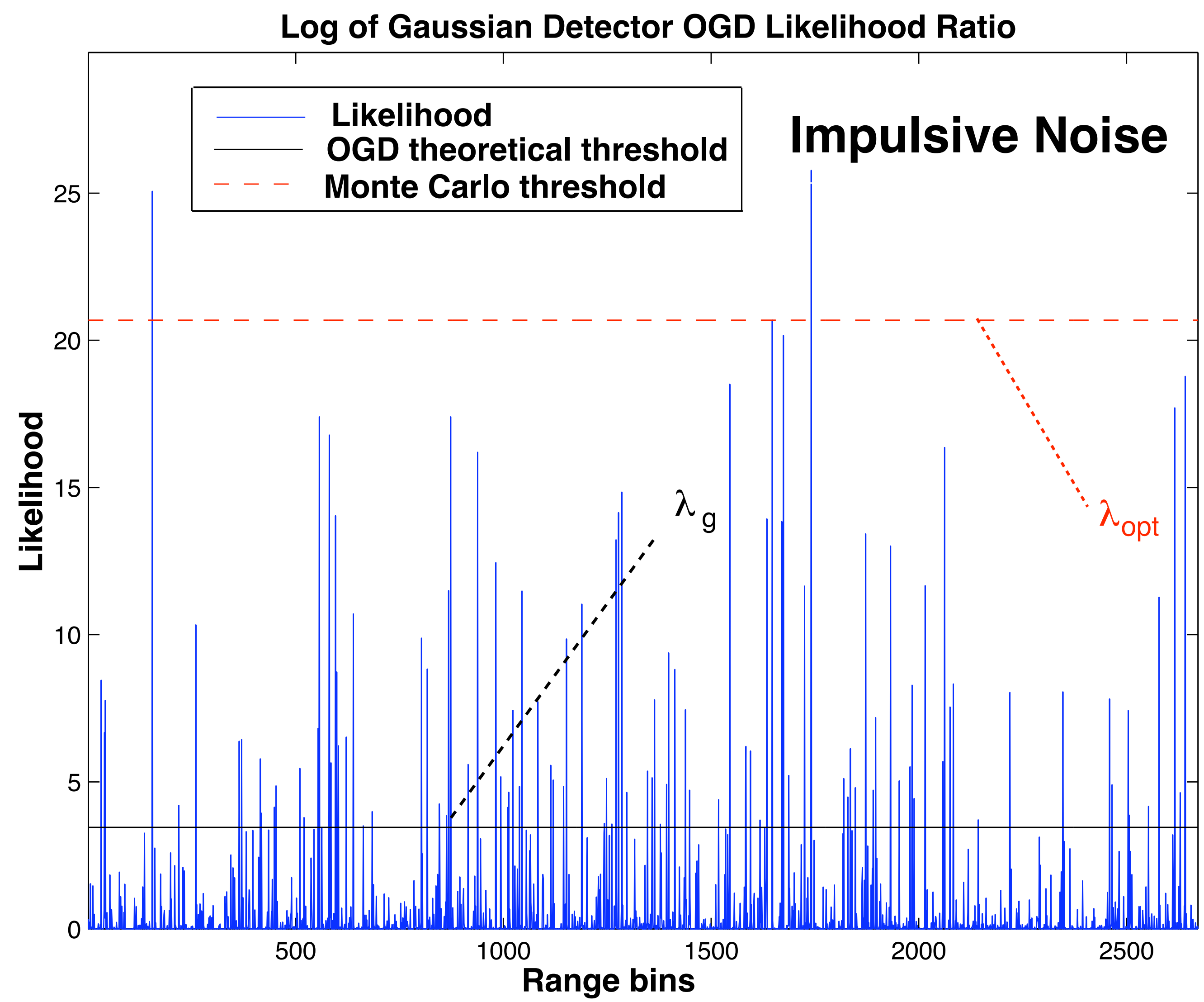
Mahalanobis on Experimental data



- Elliptical distribution models have started to be studied in the hyperspectral scientific community but one generally uses …. **Gaussian estimates !**

Motivation for Non-Gaussian Models

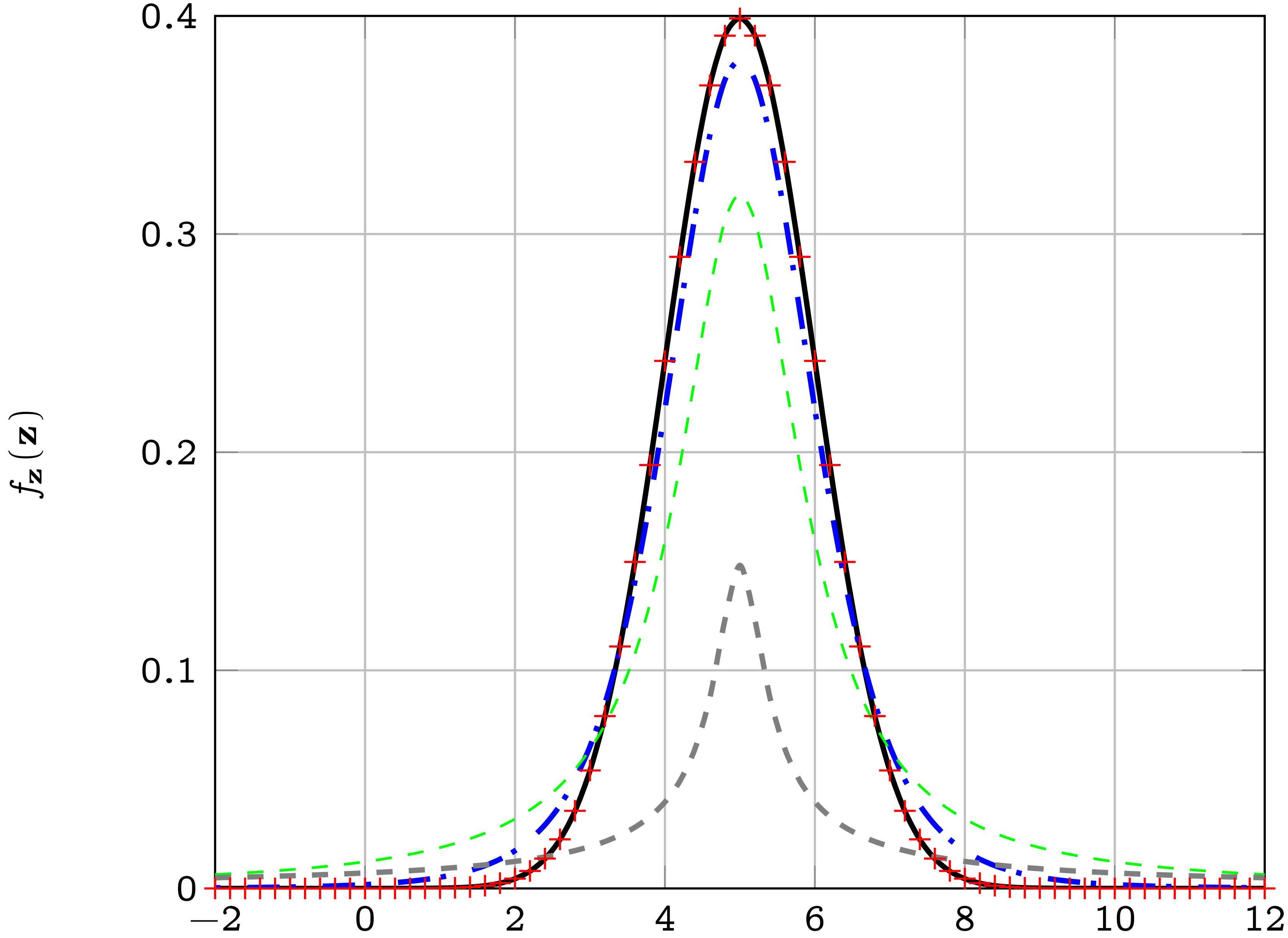




Elliptical Distributions

$$f_{\mathbf{x}}(\mathbf{x}) = \pi^{-m}|\boldsymbol{\Sigma}|^{-1}\exp\{-(\mathbf{x}-\boldsymbol{\mu})^H\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\}.$$

$$f_{\mathbf{z}}(\mathbf{z}) = |\boldsymbol{\Sigma}|^{-1} h_m\left((\mathbf{z}-\boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1}(\mathbf{z}-\boldsymbol{\mu})\right)$$



Robust *M*-estimators

M-estimators

The complex *M*-estimators of location and scatter are defined as the joint solutions of:

$$\hat{\boldsymbol{\mu}}_N = \frac{\sum_{i=1}^{N} u_1(t_i)\,\mathbf{z}_i}{\sum_{i=1}^{N} u_1(t_i)}, \quad \hat{\boldsymbol{\Sigma}}_N = \frac{1}{N}\sum_{i=1}^{N} u_2\left(t_i^2\right)\,(\mathbf{z}_i - \hat{\boldsymbol{\mu}}_N)(\mathbf{z}_i - \hat{\boldsymbol{\mu}}_N)^H,$$

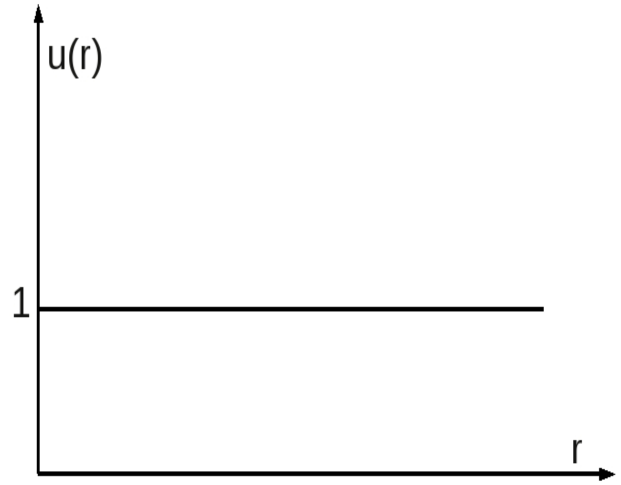
where $t_i = \left((\mathbf{z}_i - \hat{\boldsymbol{\mu}}_N)^H\,\hat{\boldsymbol{\Sigma}}_N^{-1}\,(\mathbf{z}_i - \hat{\boldsymbol{\mu}}_N)\right)^{1/2}$.

- $u_1(\cdot)$, $u_2(\cdot)$ are two weighting functions acting on the quadratic form, i.e. Mahalanobis distance,
- The choice of $u_1(\cdot)$, $u_2(\cdot)$ results in different estimates for the covariance matrix and the mean vector,
- Existence and uniqueness of the solution have been proven provided $u_1(\cdot)$, $u_2(\cdot)$ satisfy given conditions [Maronna 1976],

Examples of *M*-estimators

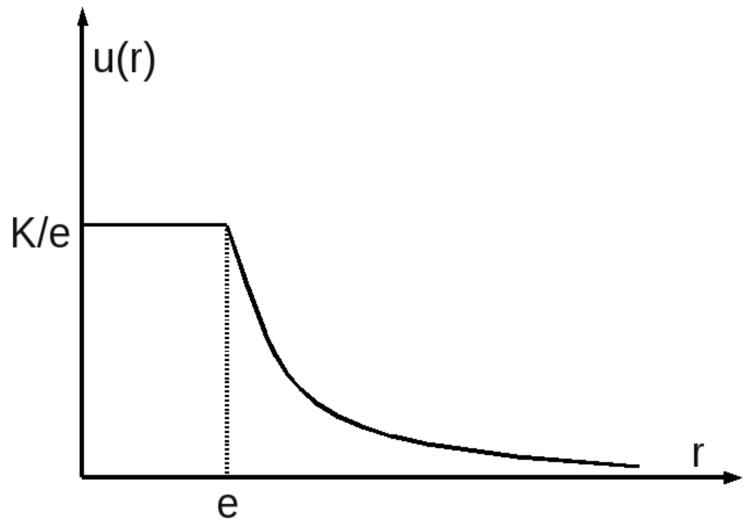
SCM:

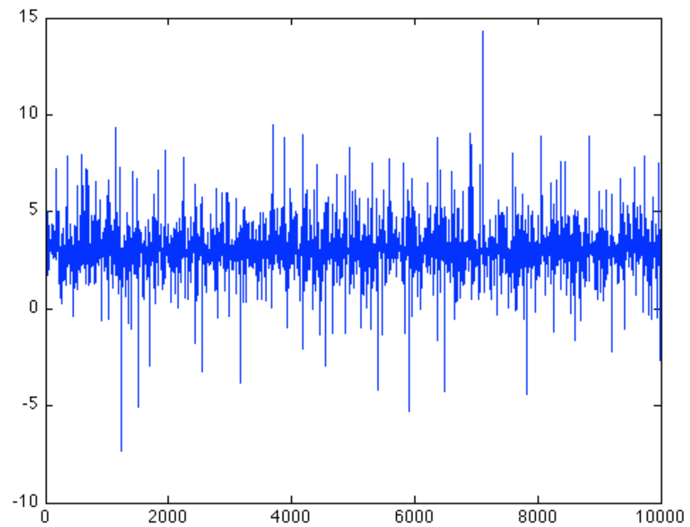
$$u(t) = 1$$



Huber's M-estimator:

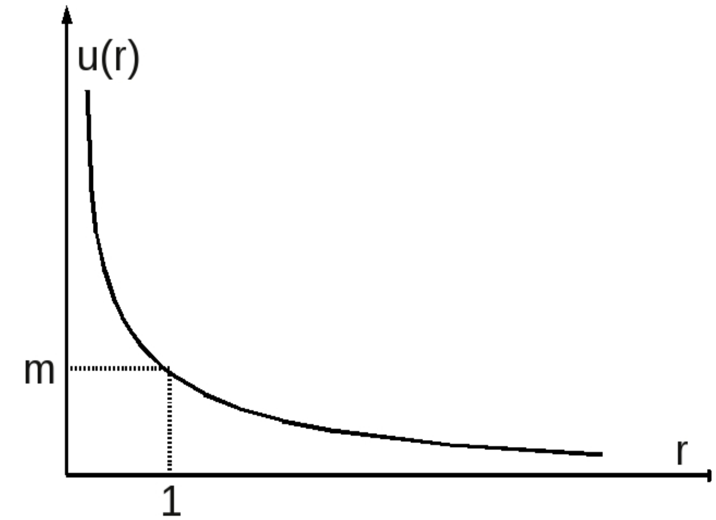
$$u(t) = \begin{cases} 1/k^2 \text{ if } t <= k^2 \\ 1/t \text{ if } t > k^2 \end{cases}$$



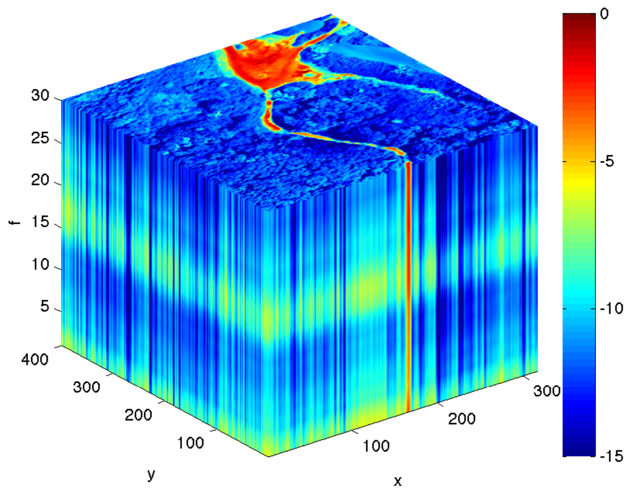


FPE (Tyler):

$$u(t) = \frac{m}{t}$$



ADAPTIVE DETECTION



ELLIPTICAL-CFAR TEST

$$\Lambda(\hat{\mathbf{M}}_{FP}, \widehat{\boldsymbol{\mu}}) = \frac{\left| \mathbf{p}^H \, \hat{\mathbf{M}}_{FP}^{-1} \, (\mathbf{c} - \widehat{\boldsymbol{\mu}}) \right|^2}{\left(\mathbf{p}^H \, \hat{\mathbf{M}}_{FP}^{-1} \, \mathbf{p} \right) \left((\mathbf{c} - \widehat{\boldsymbol{\mu}})^H \, \hat{\mathbf{M}}_{FP}^{-1} \, (\mathbf{c} - \widehat{\boldsymbol{\mu}}) \right)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda$$

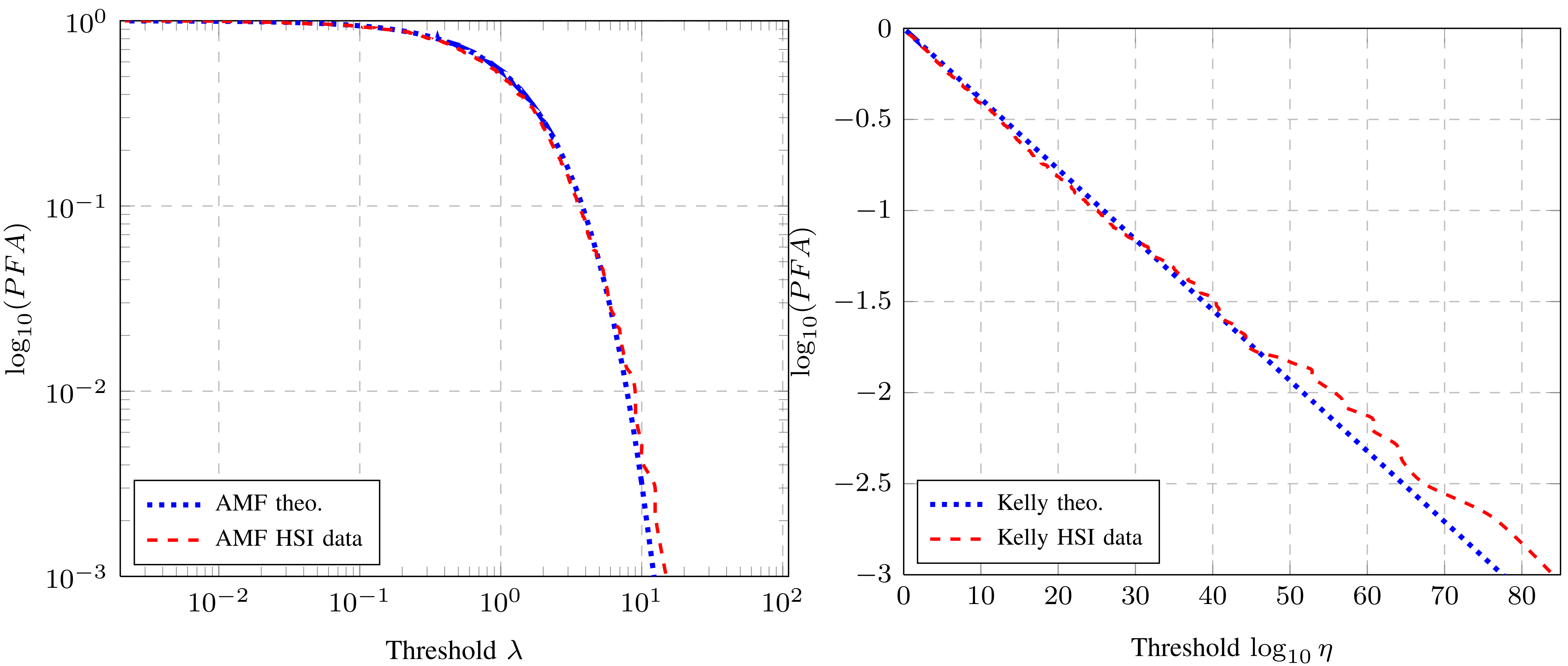
c: cell under test
p: spectral steering vector of the target

The hyperspectral data are real and positive as they represent radiance or reflectance.

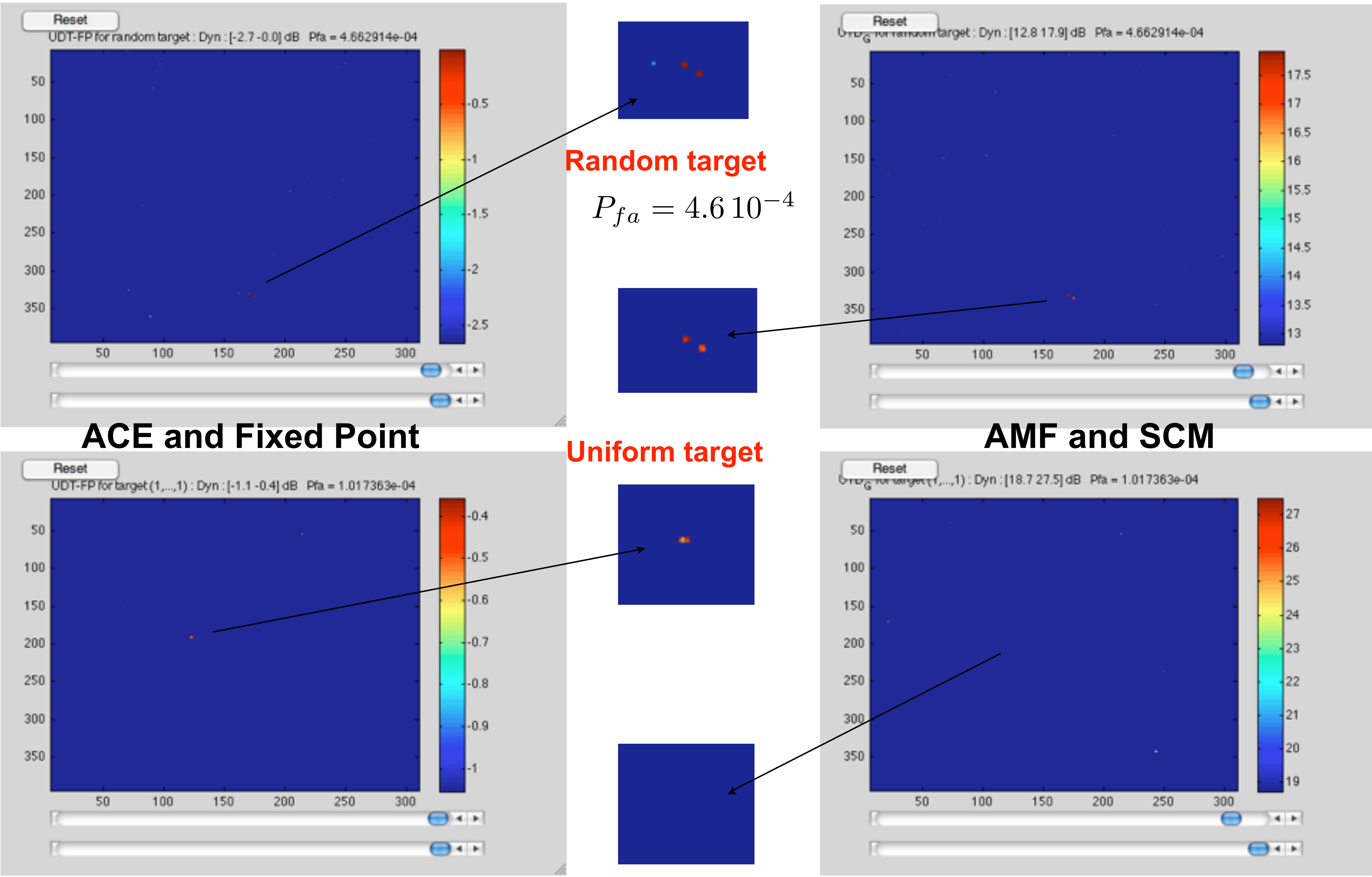
- A mean vector has to be included in the model and estimated jointly with the covariance matrix,
- The real data can be transformed into complex ones by a linear Hilbert filter.

FALSE ALARM REGULATION

- Decide Target presence and in fact only noise and background found - False Alarm
- Goal : Build detectors that keep the False Alarm Rate Constant,
- Derivation of "FA-threshold" relationships to automatically set the threshold.

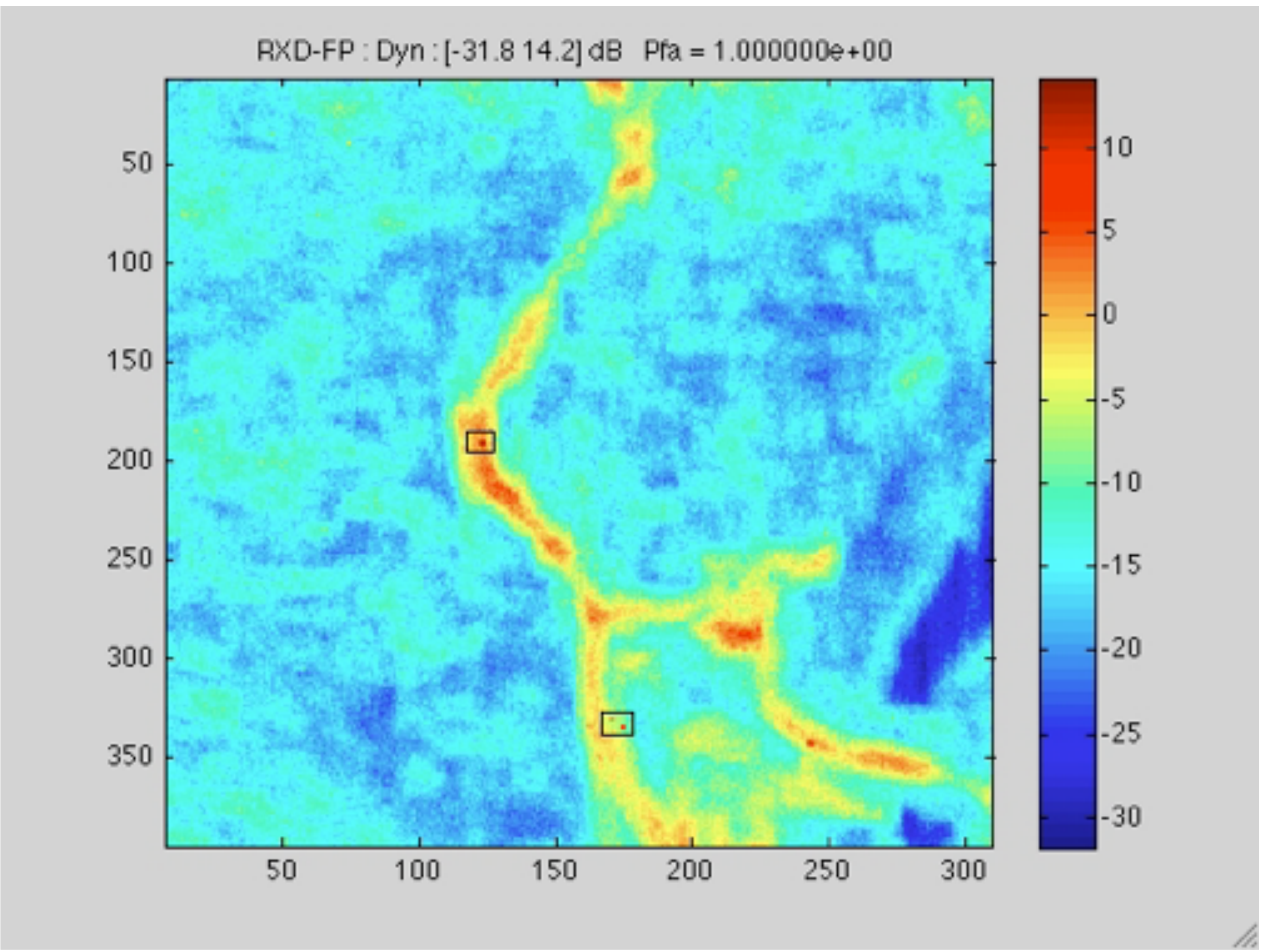


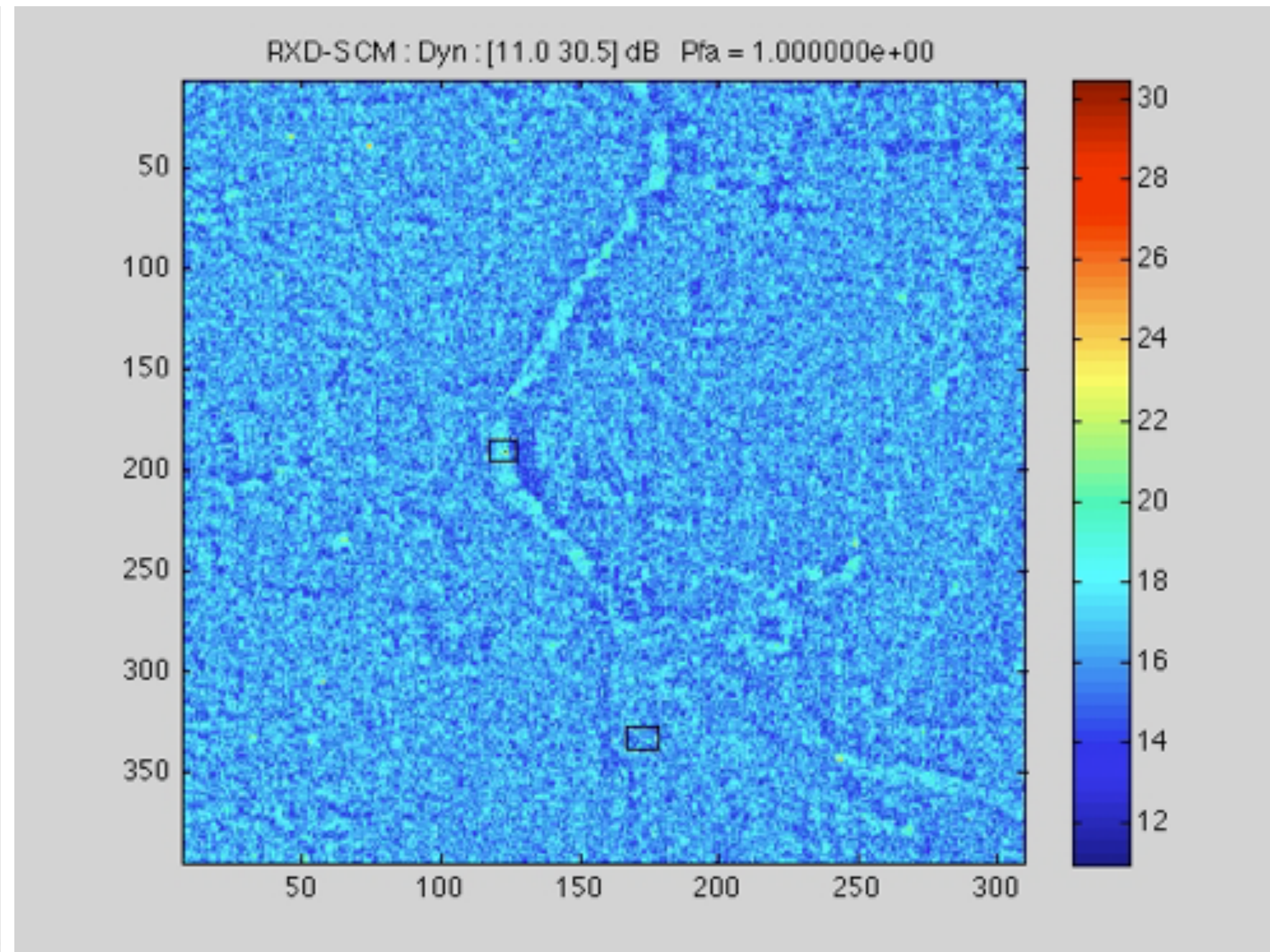
Target Detection on Real Data



Anomaly Detection on Real Data

Local Covariance Matrix estimate approach





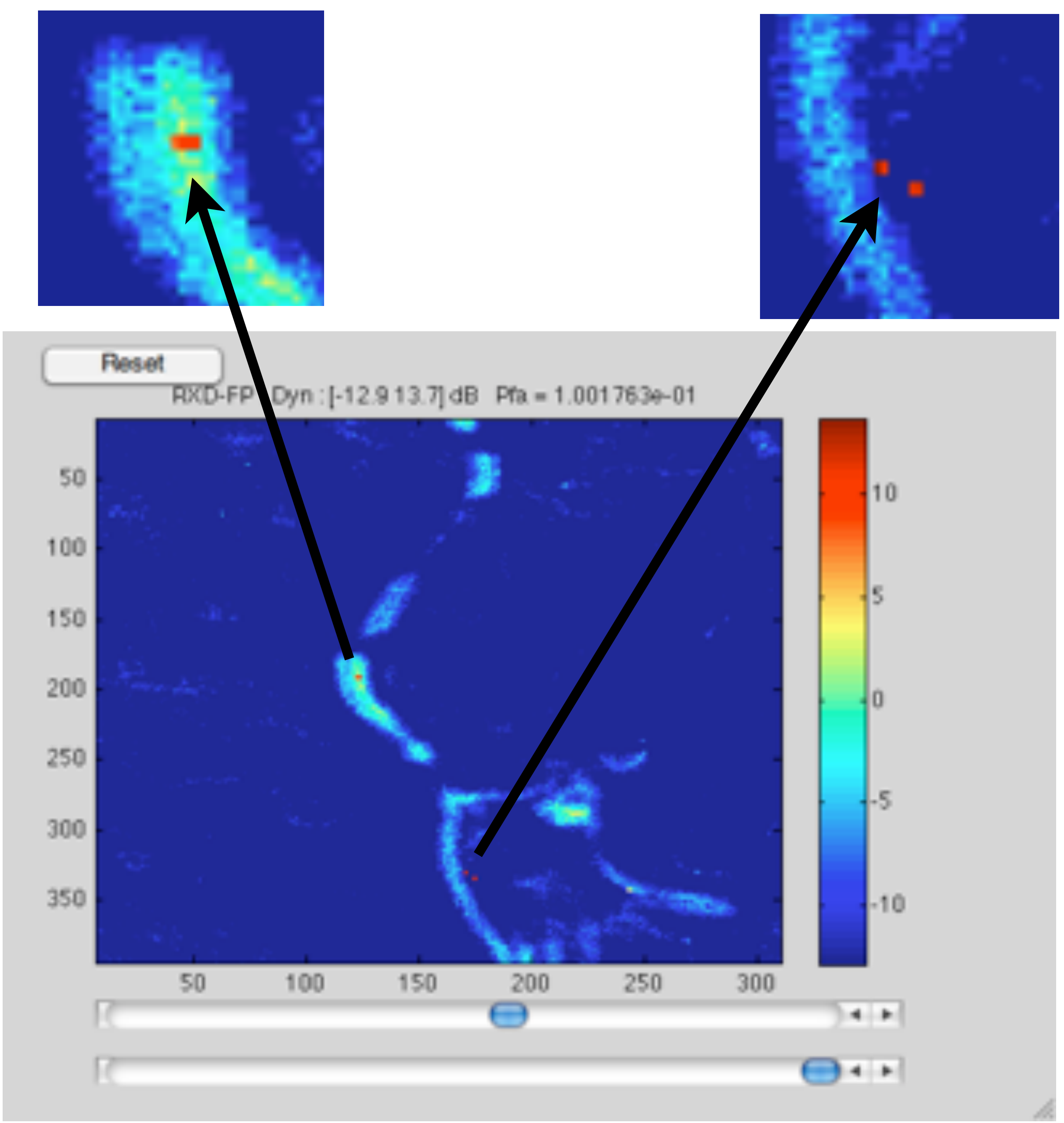
$$RXD_{FP} = (\mathbf{c}_k - \widehat{\boldsymbol{\mu}})^H \, \hat{\mathbf{M}}_{FP}^{-1} \, (\mathbf{c}_k - \widehat{\boldsymbol{\mu}})$$

$$RXD_{SCM} = (\mathbf{c}_k - \widehat{\boldsymbol{\mu}})^H \, \hat{\mathbf{M}}_{SCM}^{-1} \, (\mathbf{c}_k - \widehat{\boldsymbol{\mu}})$$

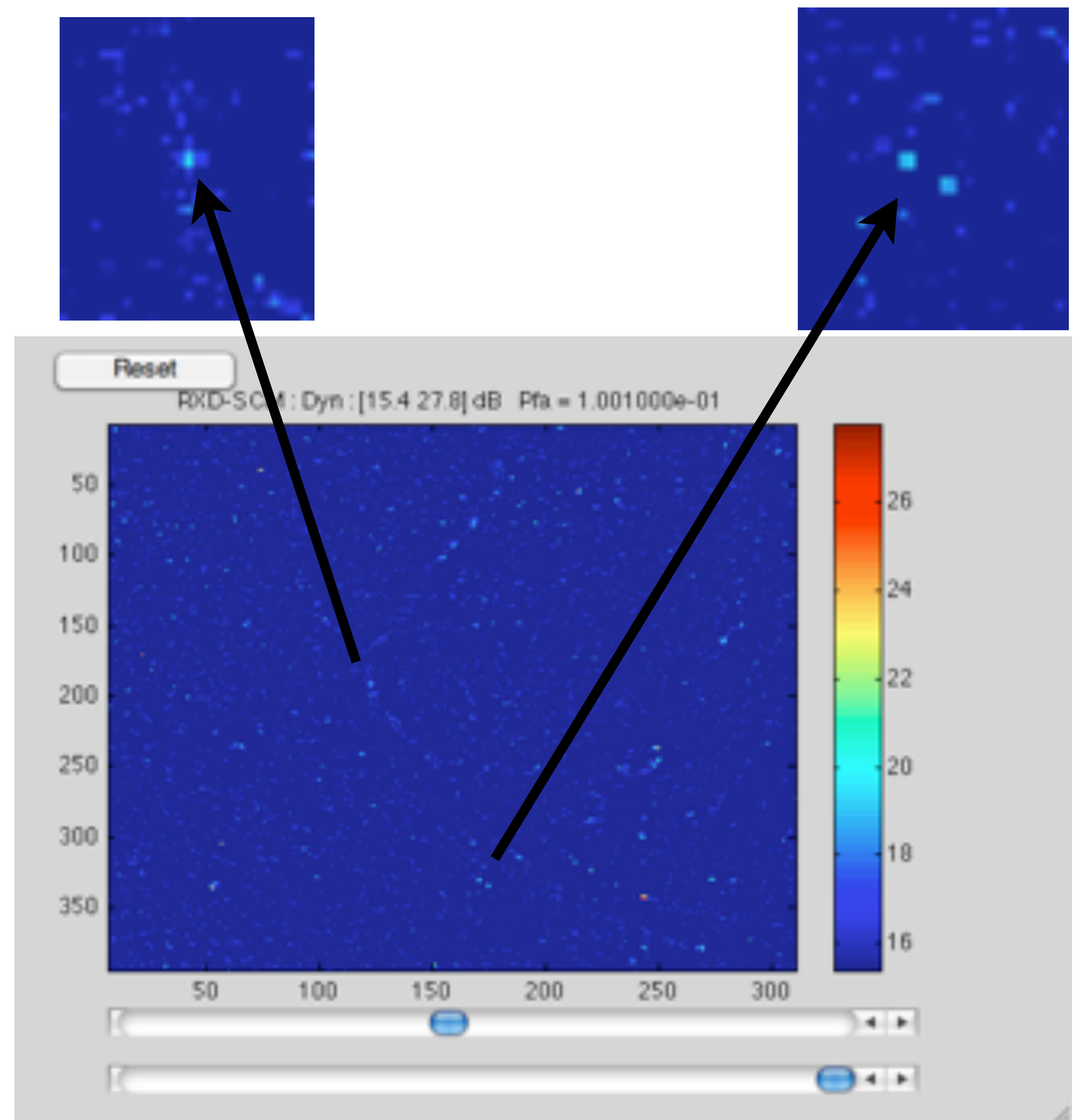
[*Reed and Yu, 1990*]

Anomaly Detection on Real Data

Mahanalobis Distance with the Classical SCM and the Fixed Point (Pfa = 0.1)



$$RXD_{FP} = (\mathbf{c}_k - \boldsymbol{\mu})^H \, \hat{\mathbf{M}}_{FP}^{-1} \, (\mathbf{c}_k - \boldsymbol{\mu})$$



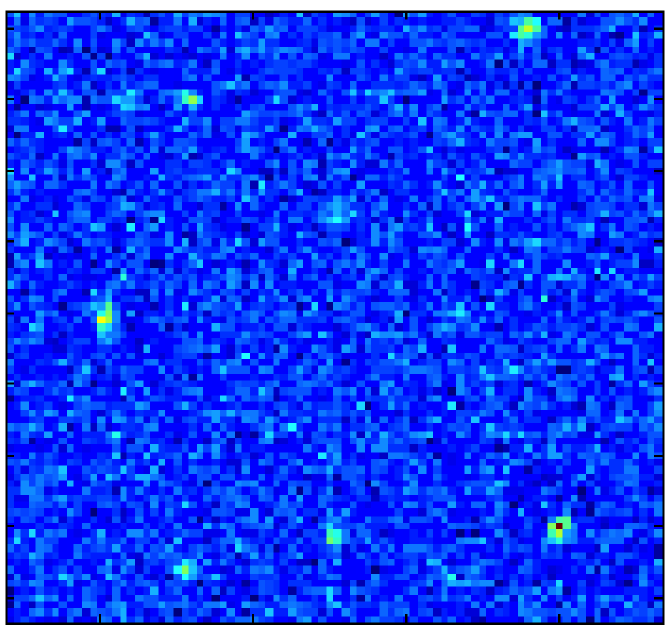
$$RXD_{SCM} = (\mathbf{c}_k - \boldsymbol{\mu})^H \, \hat{\mathbf{M}}_{SCM}^{-1} \, (\mathbf{c}_k - \boldsymbol{\mu})$$

Galaxy Detection

Problem of detecting galaxies in HS MUSE (Multi Unit Spectroscopic Explorer) data (465- 930 nm)

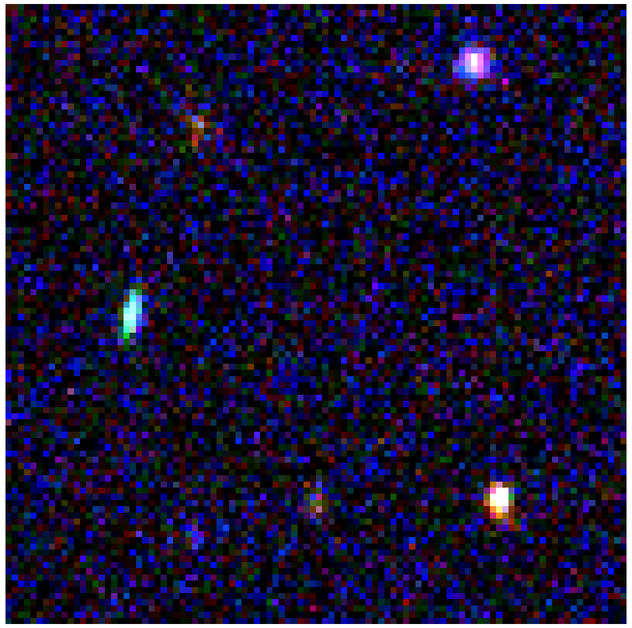
)

Classical RX Detector

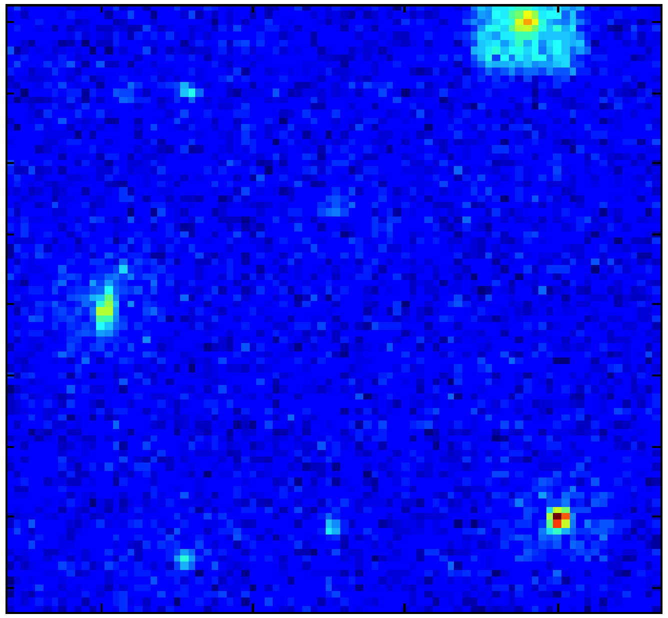


$$RXD_{SCM} = (\mathbf{c} - \hat{\boldsymbol{\mu}}_{SCM})^H \, \hat{\mathbf{M}}_{SCM}^{-1} \, (\mathbf{c} - \hat{\boldsymbol{\mu}}_{SCM})$$

MUSE images: 300 x 300 pixels, 3578 spectral bands

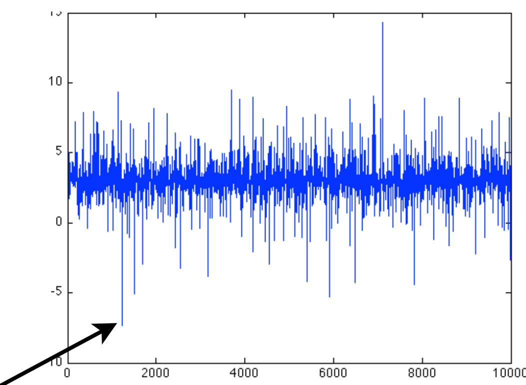


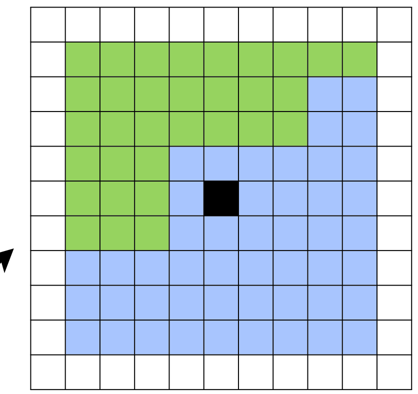
Enhanced RX Detector

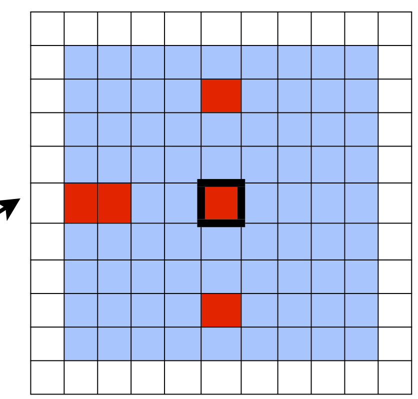


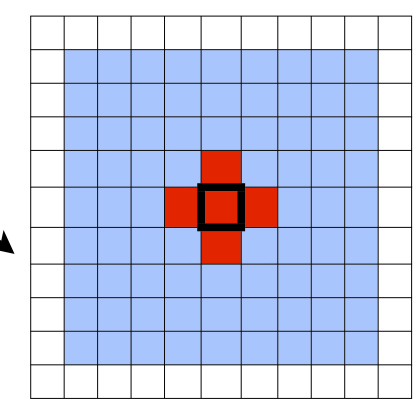
$$RXD_{FP} = (\mathbf{c} - \hat{\boldsymbol{\mu}}_{FP})^H \, \hat{\mathbf{M}}_{FP}^{-1} \, (\mathbf{c} - \hat{\boldsymbol{\mu}}_{FP})$$

Why it Works Better with M-estimators









- Impulsive samples in the secondary data (non-Gaussian distribution assumption),
- Heterogeneity due to the size of the sliding window (problems at the edges),
- The intended targets are closer than the window size,
- The targets are larger than one pixel cell resolution.

Guard pixels on the sliding window do not solve these problems!

***M*-estimators are ROBUST to the presence of strong scatterers and impulsive samples in the secondary data**

CONCLUSIONS

- Hyperspectral Imaging provides an image array that enable material discrimination and targets location that could not be spatially detected,

- Robust Target Detection improves classical methods in twofold. It is robust to corrupted information and to changes in distribution assumptions. Additionally, it allows a theoretical False Alarm Regulation which is critical in real-life detector design,

- The main obstacles in the development of effective detection algorithms are the inherent variability in target and background spectra, and the computational problems in high dimensional domains. The use of critical band selection deals quite effectively with both difficulties in target and anomaly detection,

- When extended to the Hyperspectral context, all the methodologies (Random Matrix Theory, non-Gaussian modelling, robust estimation, ...) developed for radar applications can enhance performance in detection and classification problems: source localization, linear spectral unmixing, sub-spaces techniques, detection, estimation, classification, ...

WORK AT CEA

Goal : Solve inverse problems

Signal to be retrieved

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}$$

Data

Observation operator

Noise

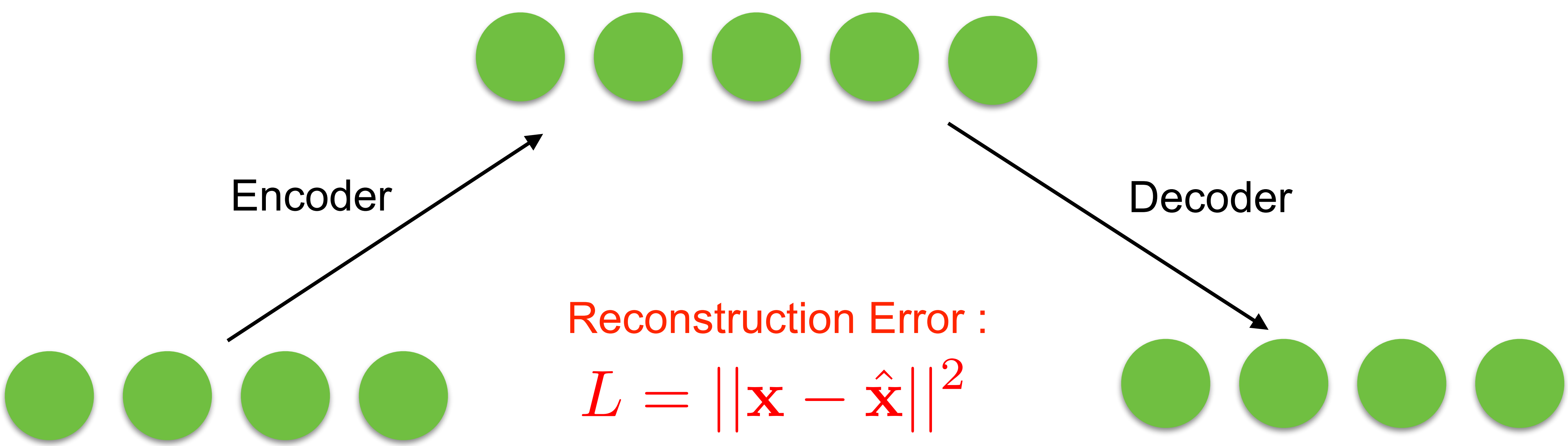
- Sparse representations provide some of the state of the art methods for solving inverse problems such as denoising, inpainting, component separation, etc.
- Classical linear models suffer from some limitations for real data modeling,
- Generalization of sparse representations to the non-linear case.

Linear models ⇒ Non-Linear models

Work at CEA

- Deep Learning : Unsupervised Learning methods that can learn invariant features hierarchies,
- Non-linear representations obtained with deep layer structures allow to bring out complex relationships and disentangle the variation factors of the inputs,
- How? - AutoEncoders, ConvNets, Deep Belief Networks,…
- Extending deep neural networks to learn sparse representations.

Hidden representation : $\mathbf{h} = \sigma(\mathbf{W}_e^T \mathbf{x} + \mathbf{b}_e)$



$$L = ||\mathbf{x} - \hat{\mathbf{x}}||^2$$

Input : $\mathbf{x} \in \mathbb{R}^d$

$$\hat{\mathbf{x}} = \sigma(\mathbf{W}_d^T \mathbf{h} + \mathbf{b}_d)$$

Applications in Astrophysics : PSF interpolation, Galaxy classification,…