



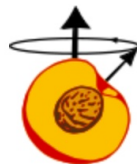
Optimized signal processing methods for NMR signal modeling and MRI data analysis for the estimation for relaxation time distributions

Saïd MOUSSAOUI

Ecole Centrale de Nantes, LS2N UMR CNRS 6004

said.moussaoui@ec-nantes.fr

Tuesday, September 18, 2018



MRFood2018

September 17 to 21, Rennes - France

Acknowledgments

- **Reconstruction algorithms**

Emilie CHOUZENOUX

Jérôme IDIER

- **NMR signal modeling**

François MARIETTE

Corinne RONDEAU-MOURO

- **MRI data analysis**

Guylaine COLLEWET

Maja MUSSE

Christian EL HAJJ

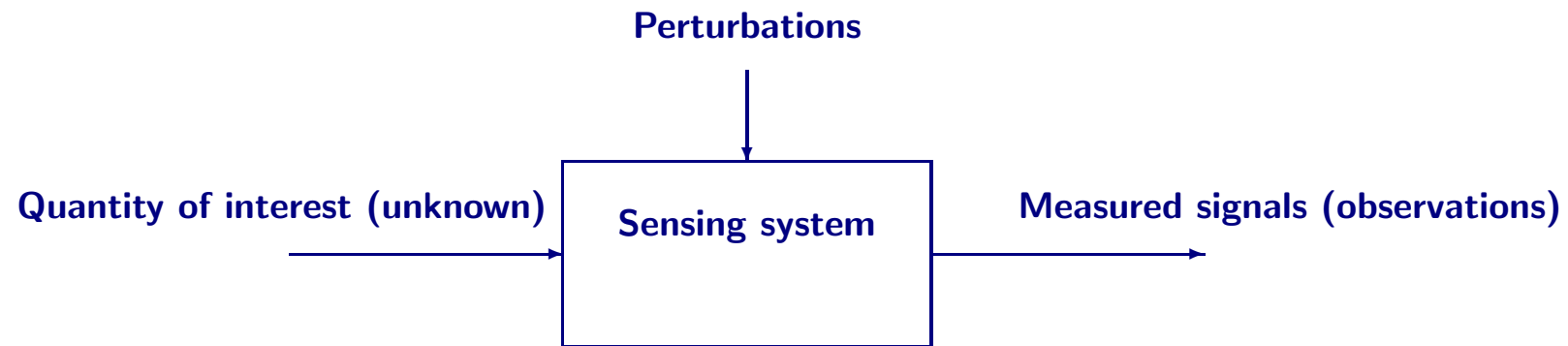
Qianyi JIANG

Outline

1. Overview of signal processing problems
 - Inverse problems
 - Resolution approaches
 - Examples
2. Relaxation times distribution estimation
 - 1D & 2D relaxation models
 - Optimized reconstruction approach
 - Dedicated computing tool
3. Summary

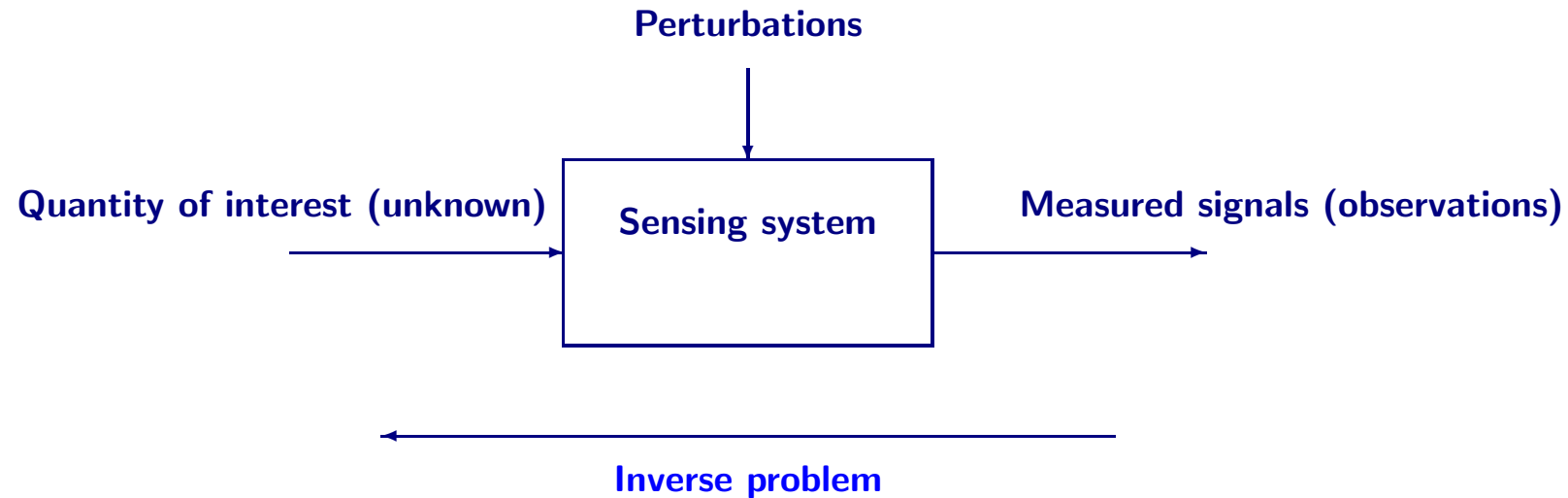
1. Overview of signal processing problems

1.1 Inverse problem



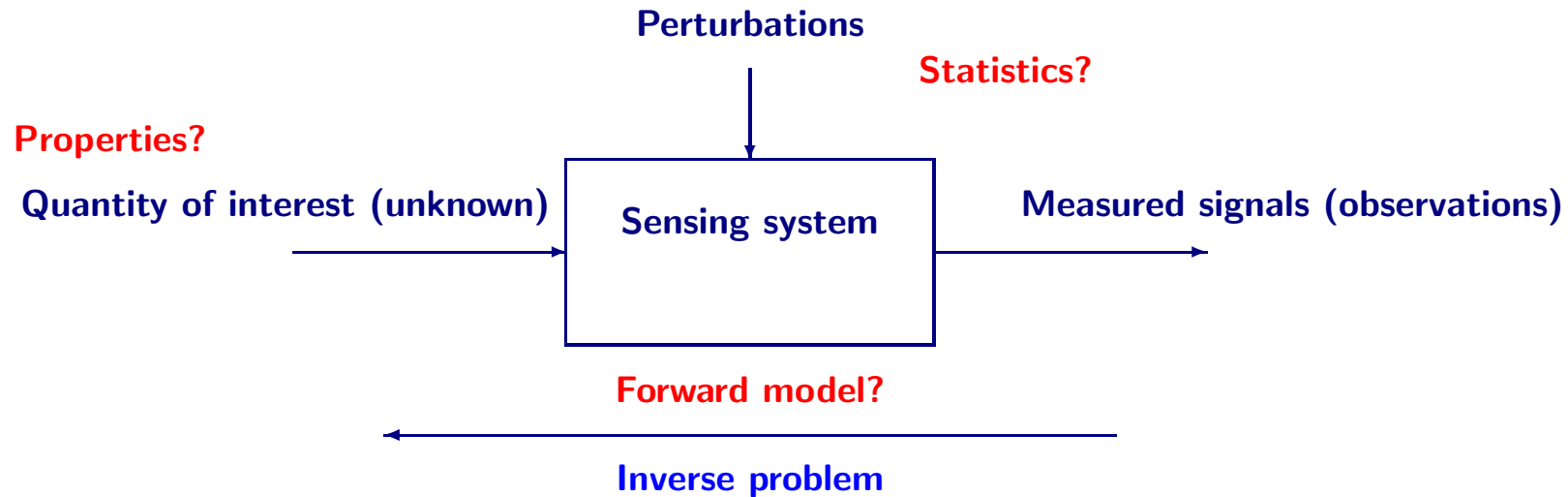
1. Overview of signal processing problems

1.1 Inverse problem



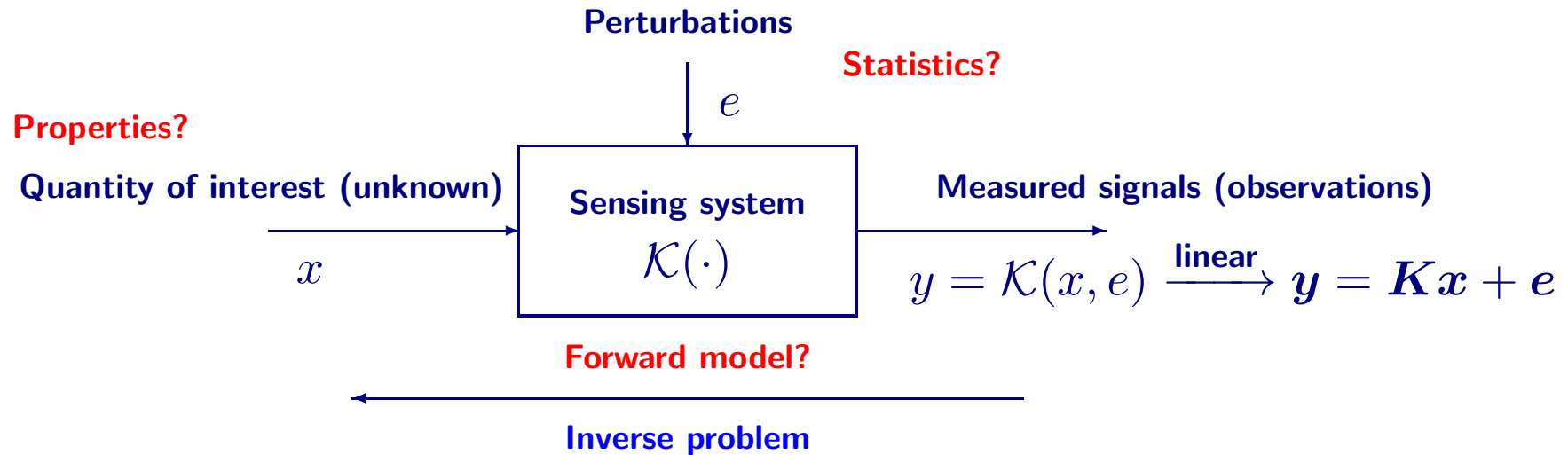
1. Overview of signal processing problems

1.1 Inverse problem



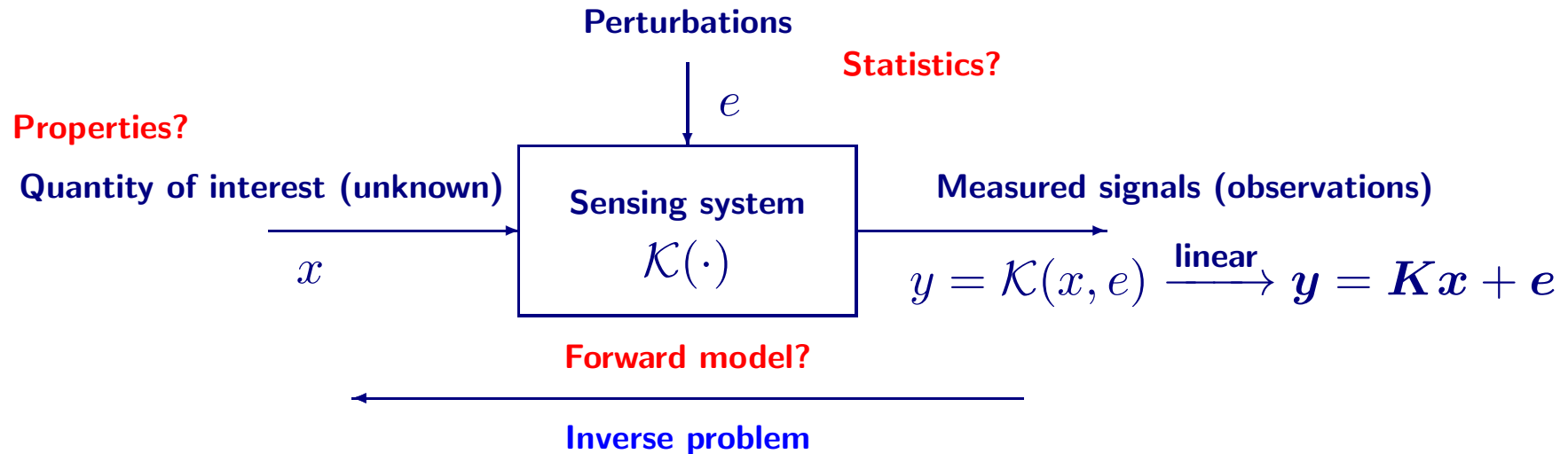
1. Overview of signal processing problems

1.1 Inverse problem



1. Overview of signal processing problems

1.1 Inverse problem

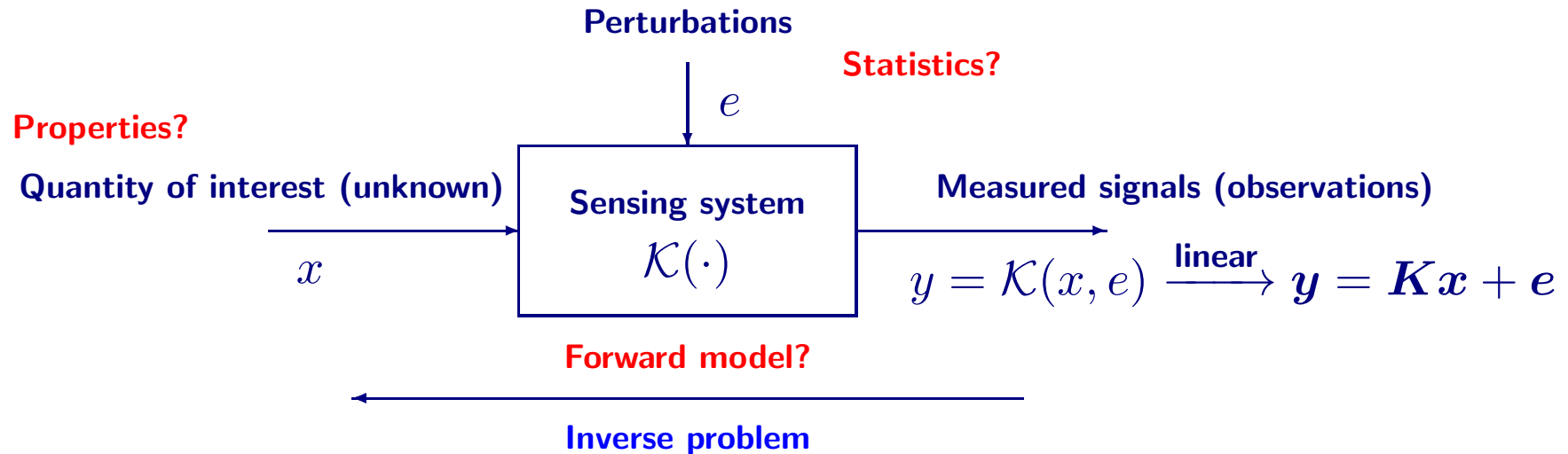


* Linear forward operators

- ① Convolution (*Image restoration, deblurring*)
- ② Projection, Radon transform (*Tomography*)
- ③ Mixing (*Partial volume effect*)
- ④ Laplace transform (*NMR Relaxation*)

1. Overview of signal processing problems

1.1 Inverse problem



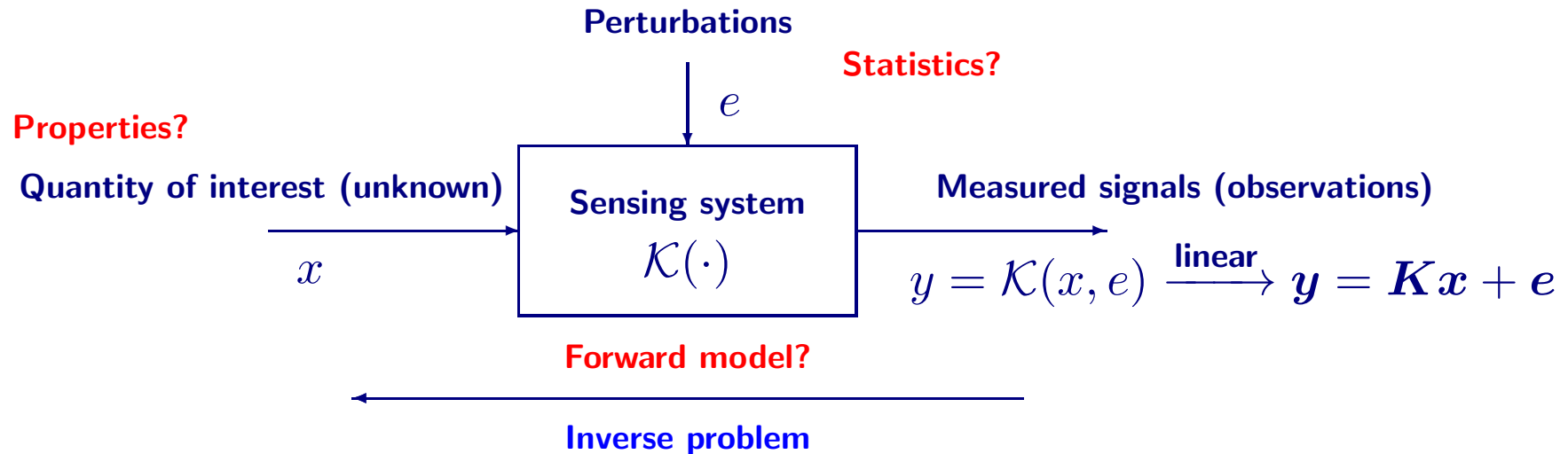
* Linear forward operators

- ① Convolution (*Image restoration, deblurring*)
- ② Projection, Radon transform (*Tomography*)
- ③ Mixing (*Partial volume effect*)
- ④ Laplace transform (*NMR Relaxation*)

* **Well-posedness conditions** [Hadamard, 1923]: a solution exists, it is unique and it is stable

1. Overview of signal processing problems

1.1 Inverse problem



* Linear forward operators

- ① Convolution (*Image restoration, deblurring*)
- ② Projection, Radon transform (*Tomography*)
- ③ Mixing (*Partial volume effect*)
- ④ Laplace transform (*NMR Relaxation*)

* **Well-posedness conditions** [Hadamard, 1923]: a solution exists, it is unique and it is stable

→ Otherwise the inverse problem is said *ill-posed*

⊛ Ill-posedness consequences

- Noise amplification \Rightarrow high uncertainty and inaccurate solution
- Non-meaningful solution (e.g. negative values)

→ Exploit prior knowledge, add hypotheses or impose constraints on the quantities of interest

⊛ Ill-posedness consequences

- Noise amplification \Rightarrow high uncertainty and inaccurate solution
- Non-meaningful solution (e.g. negative values)

\rightarrow Exploit prior knowledge, add hypotheses or impose constraints on the quantities of interest

⊛ Inverse problem regularization approach

- Compound criterion minimization: $F(\mathbf{x}) = \underbrace{Q(\mathbf{y}, \mathbf{x})}_{\text{data fitting}} + \beta \underbrace{R(\mathbf{x})}_{\text{penalization}}$

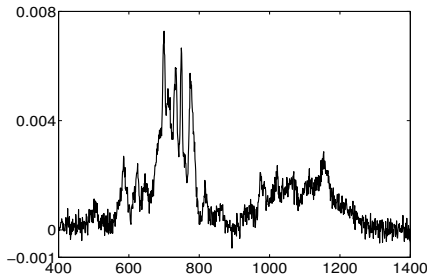
- Main challenges:

- ✓ What are the **relevant properties** of the sought quantities?
- ✓ How to quantify the data fitting $Q(\mathbf{y}, \mathbf{x})$?
- ✓ Which criterion $R(\mathbf{x})$ is more appropriate to encode the solution properties?
- ✓ How to set β , the weight of the regularization criterion?
- ✓ Is it possible to find an optimization algorithm having **fast convergence speed** and **low computing cost**?

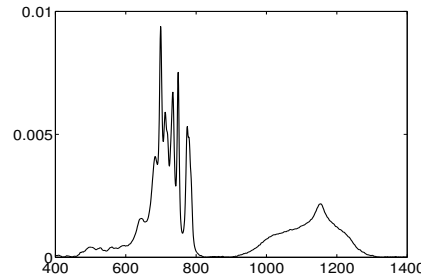
1.2 Signal denoising

$$y_t = x_t + b_t, \quad (\forall t = 1, 2, \dots, T) \Leftrightarrow \mathbf{y} = \mathbf{x} + \mathbf{b},$$

with $\mathbf{x} = [x_1, x_2, \dots, x_T]^t$ (same notation for \mathbf{y} and \mathbf{b}).



measured signal



useful signal

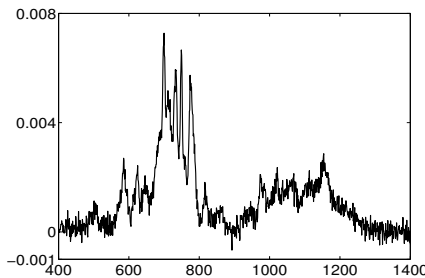
+ **perturbation (noise)**

◁ What are the noise statistics?

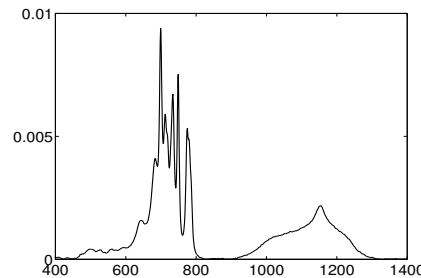
1.2 Signal denoising

$$y_t = x_t + b_t, \quad (\forall t = 1, 2, \dots, T) \Leftrightarrow \mathbf{y} = \mathbf{x} + \mathbf{b},$$

with $\mathbf{x} = [x_1, x_2, \dots, x_T]^t$ (same notation for \mathbf{y} and \mathbf{b}).



measured signal



useful signal

+ perturbation (noise)

◁ What are the noise statistics?

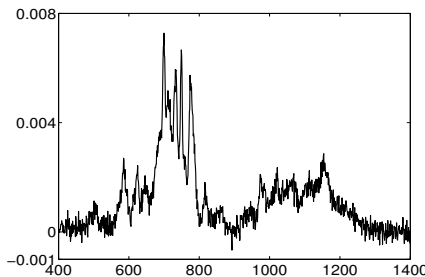
▷ **Noise** : independent samples distributed according to a **Gaussian probability density**

$$b_t \sim \mathcal{N}(0, \sigma_b^2) \implies p(\mathbf{y}|\mathbf{x}) = \prod_{t=1}^T \mathcal{N}(y_t - x_t, \sigma_b^2)$$

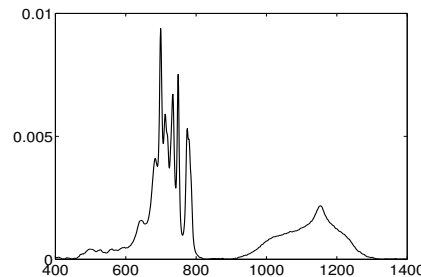
1.2 Signal denoising

$$y_t = x_t + b_t, \quad (\forall t = 1, 2, \dots, T) \Leftrightarrow \mathbf{y} = \mathbf{x} + \mathbf{b},$$

with $\mathbf{x} = [x_1, x_2, \dots, x_T]^t$ (same notation for \mathbf{y} and \mathbf{b}).



measured signal



useful signal

+ perturbation (noise)

◁ *What are the noise statistics?*

- ▷ **Noise** : independent samples distributed according to a **Gaussian probability density**

$$b_t \sim \mathcal{N}(0, \sigma_b^2) \implies p(\mathbf{y}|\mathbf{x}) = \prod_{t=1}^T \mathcal{N}(y_t - x_t, \sigma_b^2)$$

1) Maximum likelihood estimation

$$\hat{\mathbf{x}} = \arg \max_{\mathbf{x}} p(\mathbf{y}|\mathbf{x}) = \arg \min_{\mathbf{x}} (-\log p(\mathbf{y}|\mathbf{x}))$$

$$= \arg \min_{\mathbf{x}} \frac{1}{2\sigma_b^2} \sum_{t=1}^T (y_t - x_t)^2 \quad (\text{Least squares estimation})$$

- ▷ **Solution:** $\hat{\mathbf{x}} = \mathbf{y} \implies$ ◁ *no processing !!* ◁ *How to describe the sought signal properties?*

2) Maximum a posteriori estimation

- ▷ **Smooth signal:** k -th derivative distributed according to a **zero-mean Gaussian** probability density

$$\frac{d^k x_t}{dt^k} \sim \mathcal{N}(0, \sigma_d^2) \implies p(\mathbf{x}) = \prod_{t=1}^T \mathcal{N}(\mathbf{D}_k \mathbf{x}_t, \sigma_b^2)$$

Bayes' theorem: $p(\mathbf{x}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x}) \times p(\mathbf{x})$

2) Maximum a posteriori estimation

- ▷ **Smooth signal:** k -th derivative distributed according to a **zero-mean Gaussian** probability density

$$\frac{d^k x_t}{dt^k} \sim \mathcal{N}(0, \sigma_d^2) \implies p(\mathbf{x}) = \prod_{t=1}^T \mathcal{N}(\mathbf{D}_k \mathbf{x}_t, \sigma_b^2)$$

Bayes' theorem: $p(\mathbf{x}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x}) \times p(\mathbf{x})$

- ▷ **Solution:** minimization of $(-\log p(\mathbf{x}|\mathbf{y})) = \frac{1}{2\sigma_b^2} \sum_{t=1}^T (y_t - x_t)^2 + \frac{1}{2\sigma_d^2} \sum_{t=1}^T (\mathbf{D}_k \mathbf{x}_t)^2$

$$J(\mathbf{x}) = \|\mathbf{y} - \mathbf{x}\|_2^2 + \beta \|\mathbf{D}_k \mathbf{x}\|_2^2 \quad \text{with } \beta = \frac{\sigma_b^2}{\sigma_d^2}$$

2) Maximum a posteriori estimation

- ▷ **Smooth signal:** k -th derivative distributed according to a **zero-mean Gaussian** probability density

$$\frac{d^k x_t}{dt^k} \sim \mathcal{N}(0, \sigma_d^2) \implies p(\mathbf{x}) = \prod_{t=1}^T \mathcal{N}(\mathbf{D}_k x_t, \sigma_b^2)$$

Bayes' theorem: $p(\mathbf{x}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x}) \times p(\mathbf{x})$

- ▷ **Solution:** minimization of $(-\log p(\mathbf{x}|\mathbf{y})) = \frac{1}{2\sigma_b^2} \sum_{t=1}^T (y_t - x_t)^2 + \frac{1}{2\sigma_d^2} \sum_{t=1}^T (\mathbf{D}_k x_t)^2$

$$J(\mathbf{x}) = \|\mathbf{y} - \mathbf{x}\|_2^2 + \beta \|\mathbf{D}_k \mathbf{x}\|_2^2 \quad \text{with } \beta = \frac{\sigma_b^2}{\sigma_d^2}$$

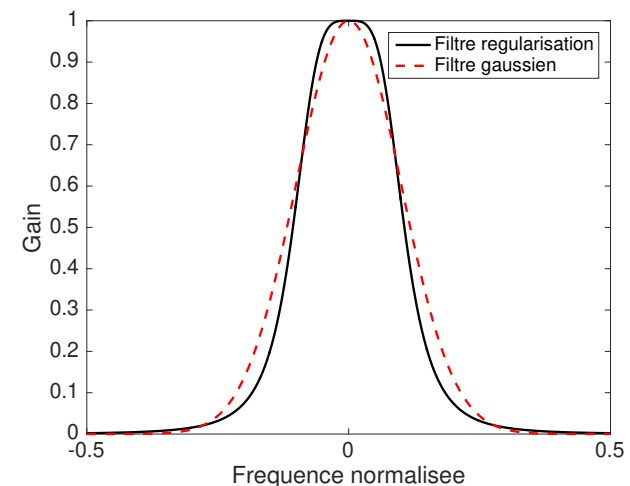
$$\implies \hat{\mathbf{x}} = [\mathbf{I} + \beta \mathbf{D}_k^t \mathbf{D}_k]^{-1} \mathbf{y}$$

Frequency-domain formulation :

$$\hat{X}(w) = \frac{1}{1 + \beta w^{2k}} Y(w)$$

◁ *Low-pass filter with a band-pass controlled by β*

◁ *Similar shape a Gaussien filter*



1.3 Parametric decomposition/modeling of a signal

$$x_t = \mathcal{M}(t, \boldsymbol{\theta}) = \sum_{k=1}^K \alpha_k f_k(t; \mu_k, \sigma_k),$$

with functions f_k of known forms (Exponential, Gaussian, Lorentzian, Voigt, etc.).

The parameters $\boldsymbol{\theta} = [(\alpha_1, \mu_1, \sigma_1), \dots, (\alpha_K, \mu_K, \sigma_K)]^t$ are estimated by minimizing a least squares criterion

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \sum_{t=1}^T (y_t - \mathcal{M}(t, \boldsymbol{\theta}))^2$$

1.3 Parametric decomposition/modeling of a signal

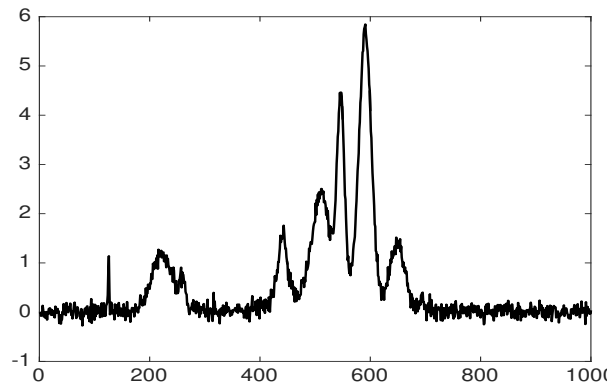
$$x_t = \mathcal{M}(t, \boldsymbol{\theta}) = \sum_{k=1}^K \alpha_k f_k(t; \mu_k, \sigma_k),$$

with functions f_k of known forms (Exponential, Gaussian, Lorentzian, Voigt, etc.).

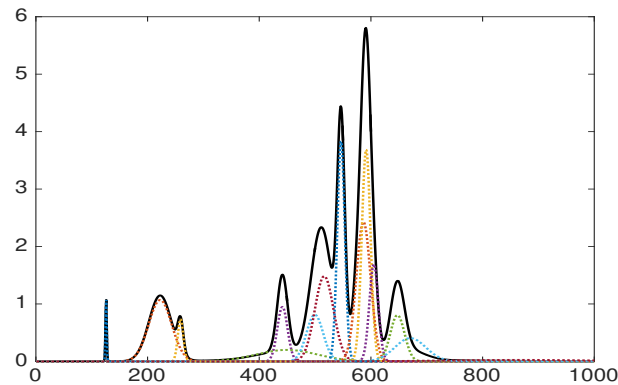
The parameters $\boldsymbol{\theta} = [(\alpha_1, \mu_1, \sigma_1), \dots, (\alpha_K, \mu_K, \sigma_K)]^t$ are estimated by minimizing a least squares criterion

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \sum_{t=1}^T (y_t - \mathcal{M}(t, \boldsymbol{\theta}))^2$$

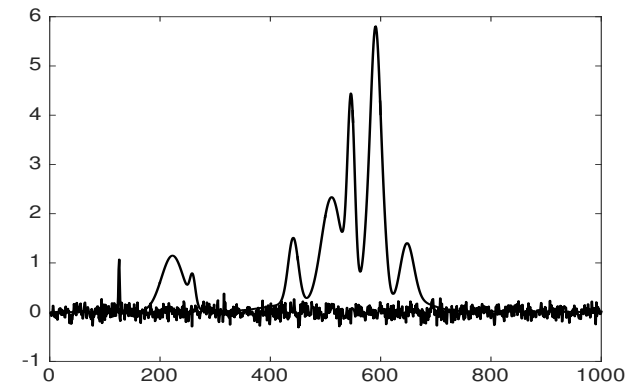
Iterative descent algorithms (Gauss-Newton, Levenberg-Marquardt) or Monte-Carlo methods are used.



raw spectrum



decomposition

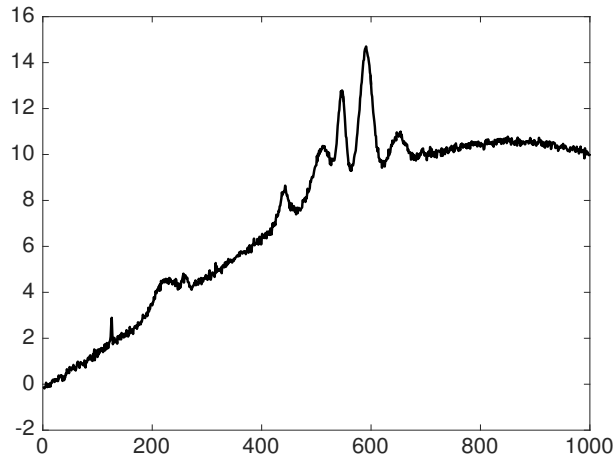


residual

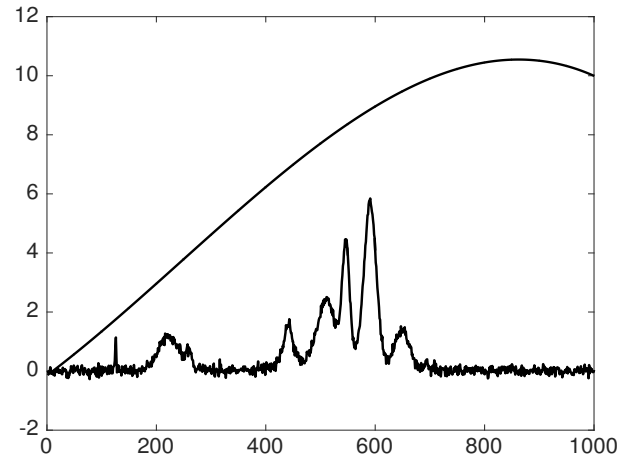
1.4 Background signal removal

$$y_t = x_t + p_t + b_t$$

with $p_t = a_0 + a_1t + \dots + a_nt^n$, n -th order polynomial.



signal with baseline

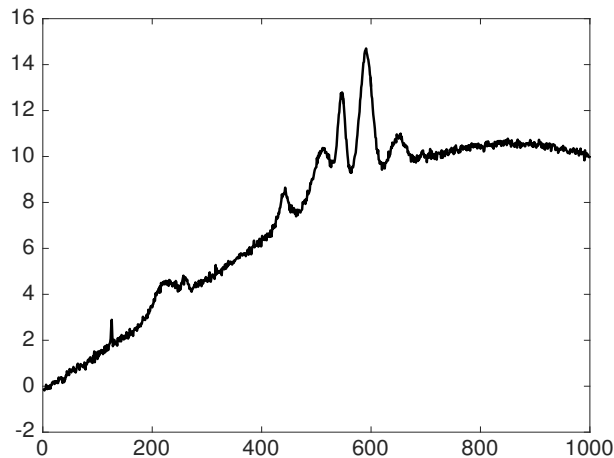


decomposition

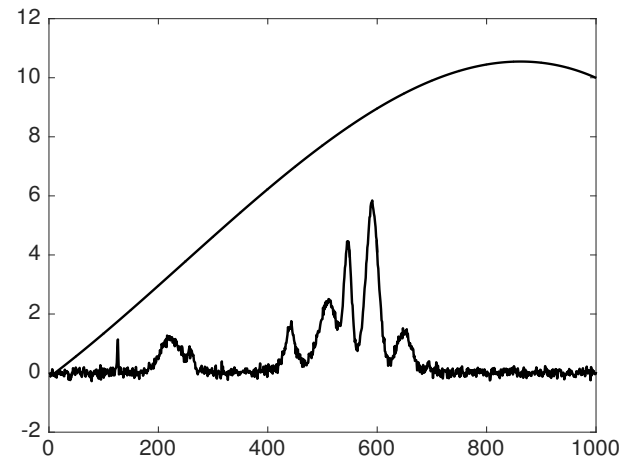
1.4 Background signal removal

$$y_t = x_t + p_t + b_t$$

with $p_t = a_0 + a_1t + \dots + a_nt^n$, n -th order polynomial.



signal with baseline



decomposition

Estimation of the parameters $\mathbf{a} = [a_0, \dots, a_n]$ of the polynomial by **weighted least squares** minimization

$$J(\mathbf{a}) = \sum_{t=1}^T \frac{1}{2\sigma_t^2} (y_t - p_t)^2 \quad \text{such that } \sigma_t^2 \text{ takes high values for large residuals.}$$

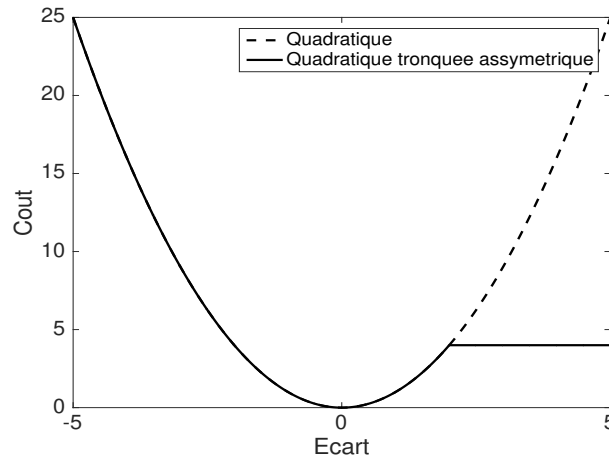
It requires the specification of the values of σ_t^2 .

Alternative approach: robust regression using non-convex cost functions

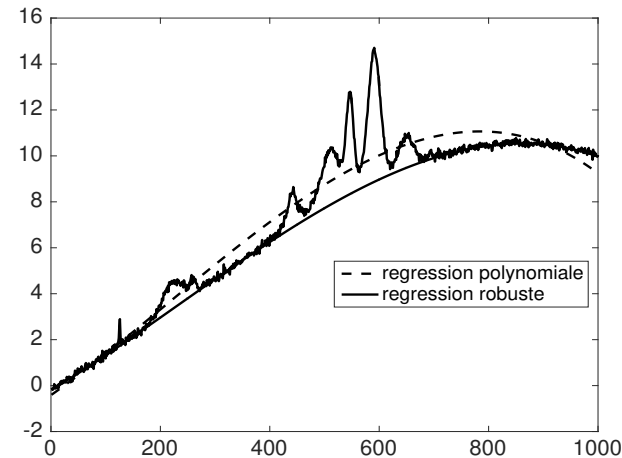
$$J(\mathbf{a}) = \sum_{t=1}^T \psi(y_t - p_t).$$

Example: asymmetric truncated quadratic, adapted for the estimation of spectra [Mazet, 2005].

⊗ Illustration



robust cost function



estimation result

⋄ *Need to specify the polynomial order, the threshold level and the noise variance.*

⋄ *trial and error test for the selection of these tuning parameters*

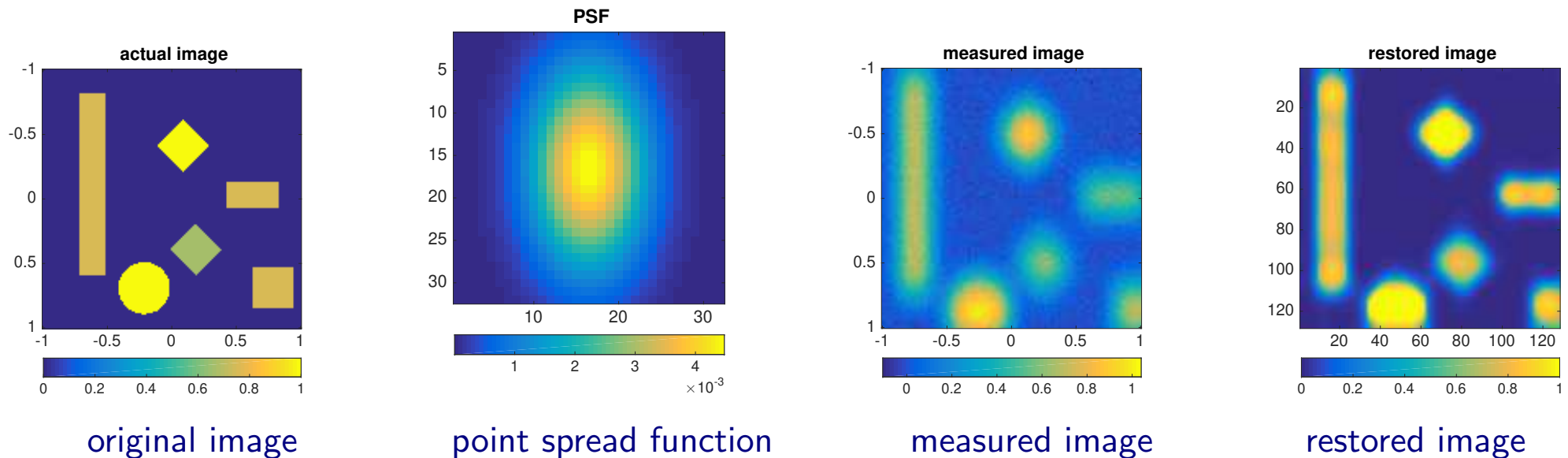
1.5 Image deconvolution

The optical response of an image acquisition system can introduce a blur that can be represented by a convolution

$$y_{ij} = \sum_{k=0}^K \sum_{l=0}^L h_{kl} x_{(i-k,j-l)} + b_{ij}$$

H_{kl} are the coefficients of the point spread function (PSF) of the imaging system.

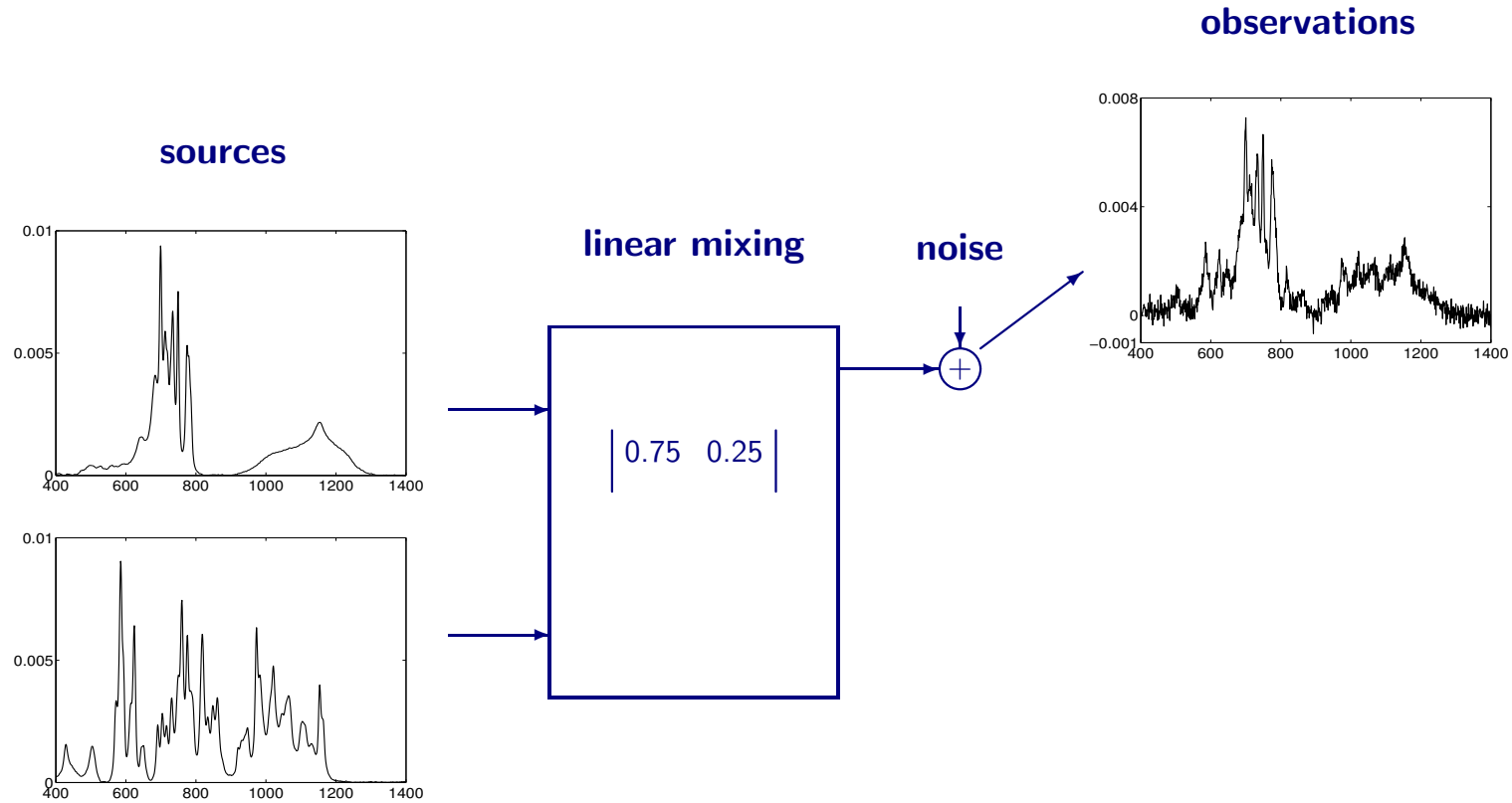
⊗ Example:



⊲ Iterative optimization required to manage the computation cost and the memory usage

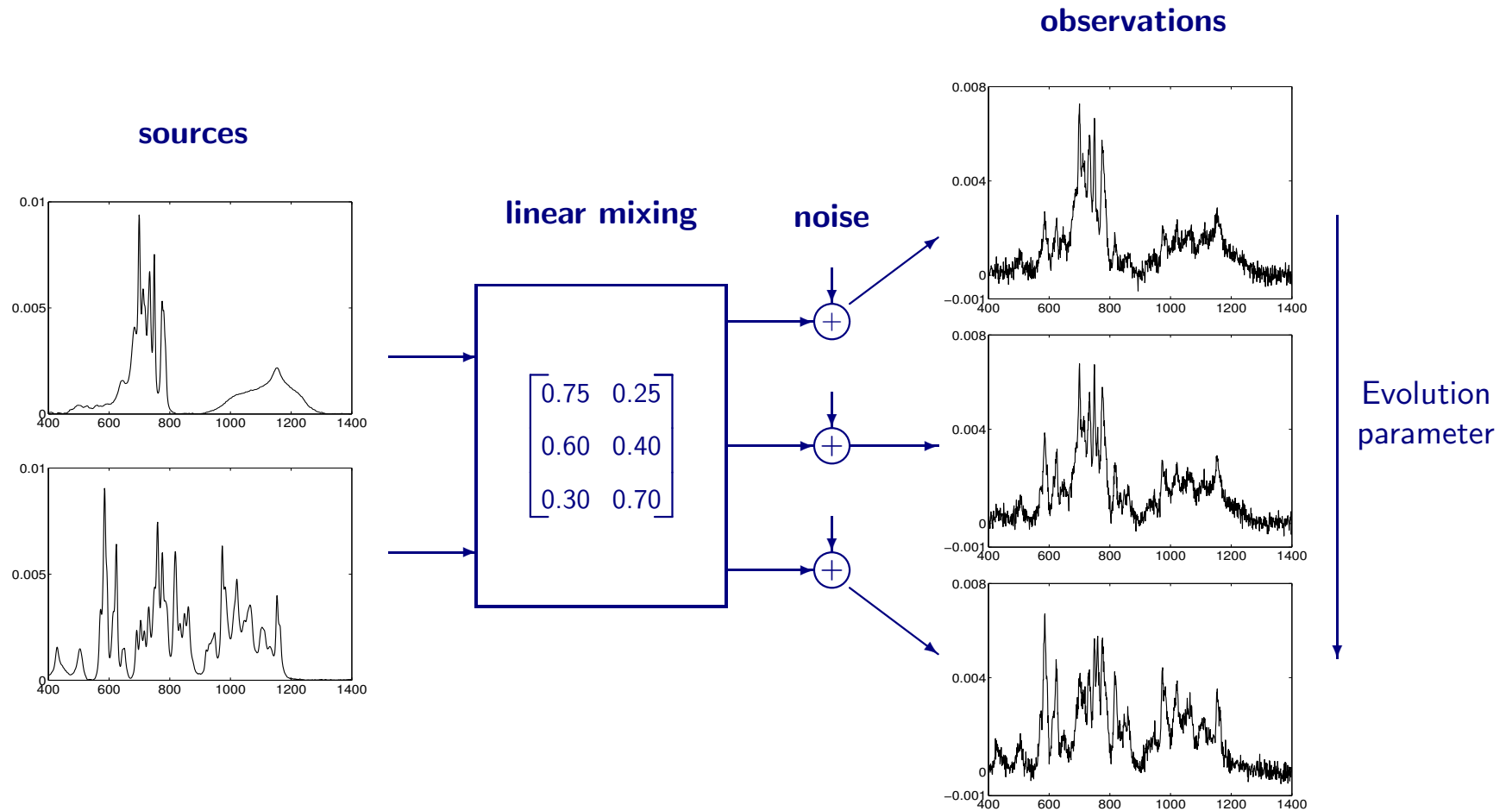
1.6 Signal separation (Mixture analysis)

⊛ Partial volume effect due to low spatial image resolution.



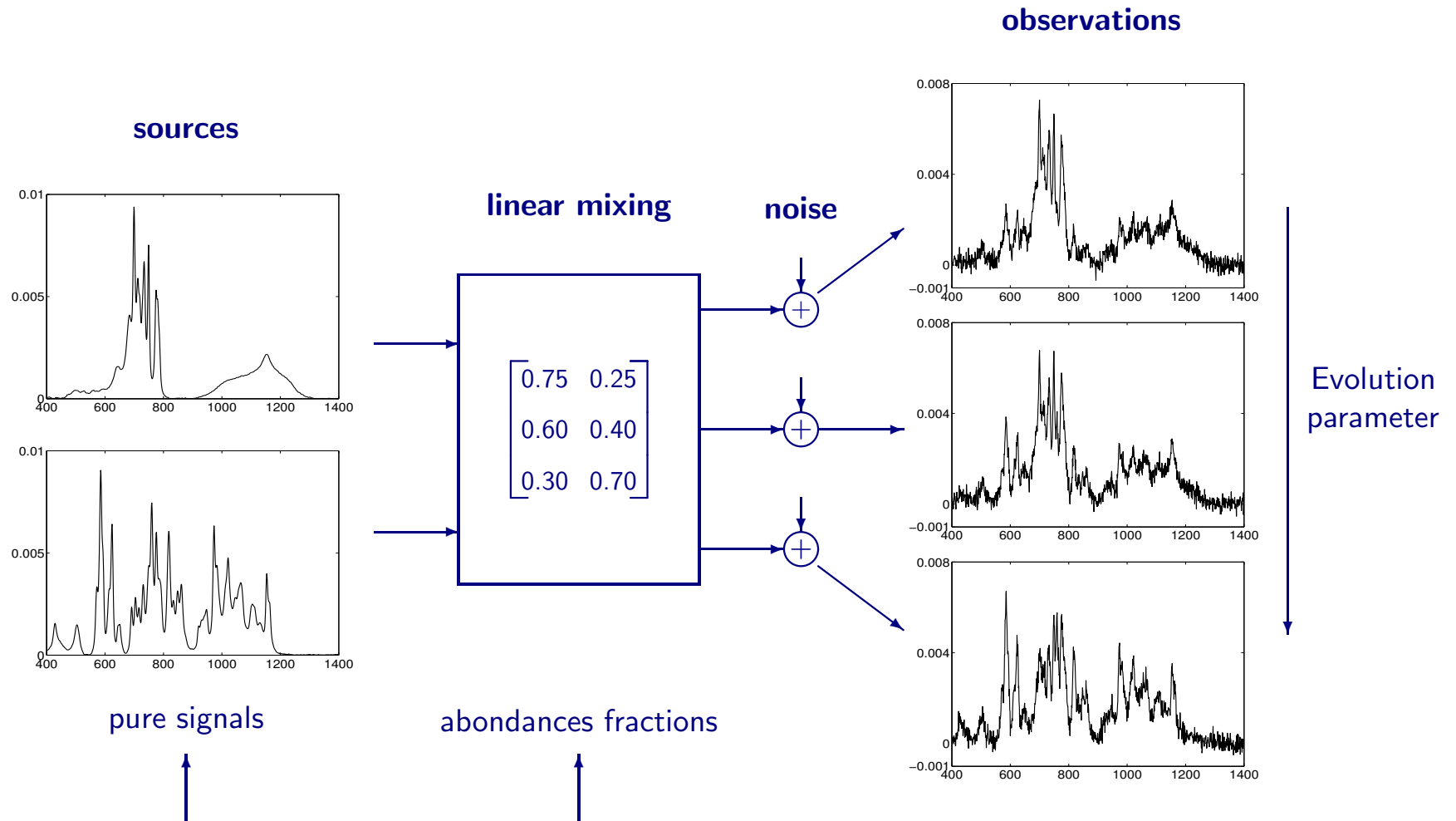
1.6 Signal separation (Mixture analysis)

⊗ Partial volume effect due to low spatial image resolution.



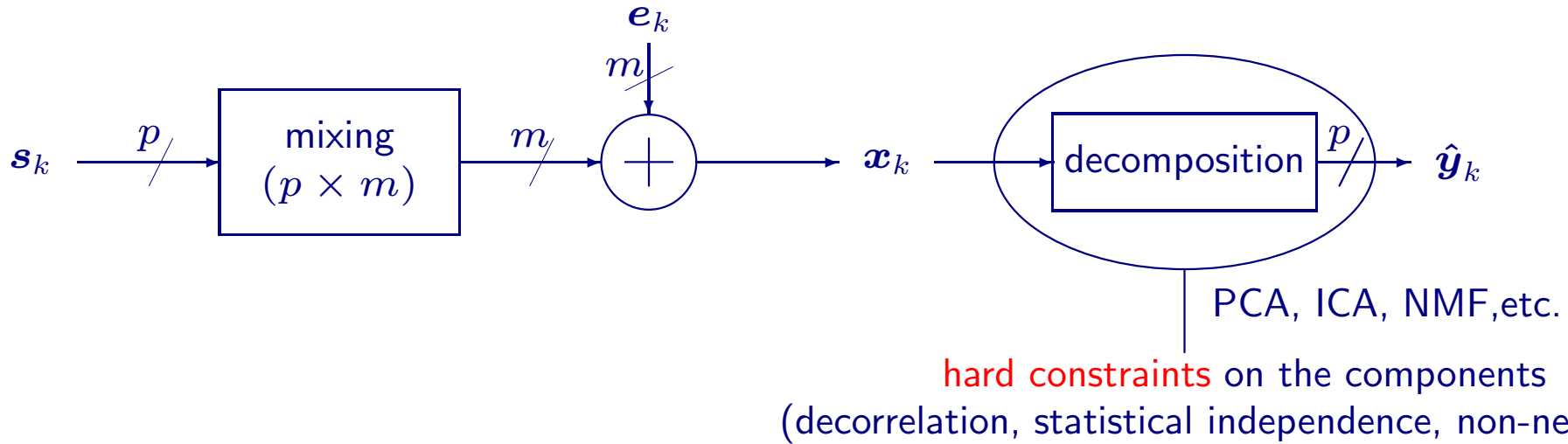
1.6 Signal separation (Mixture analysis)

⊗ Partial volume effect due to low spatial image resolution.

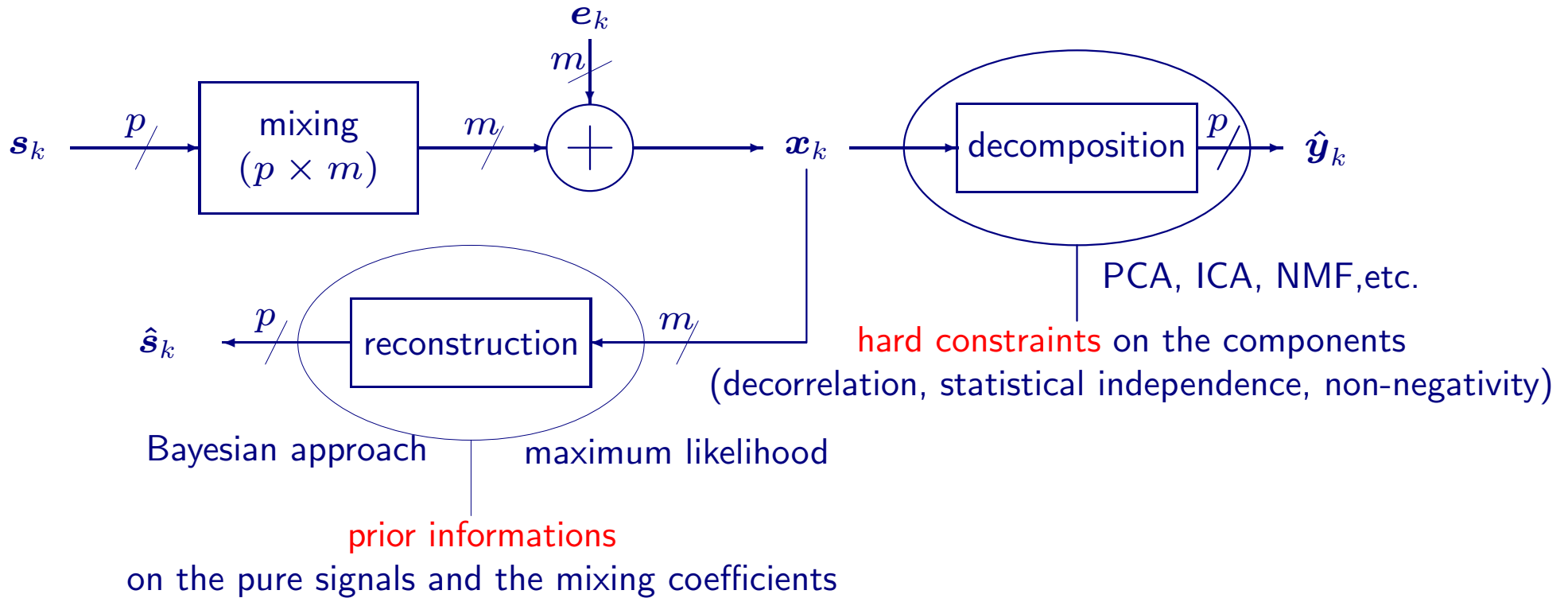


Inverse problem: determine the pure signals and their fractions.

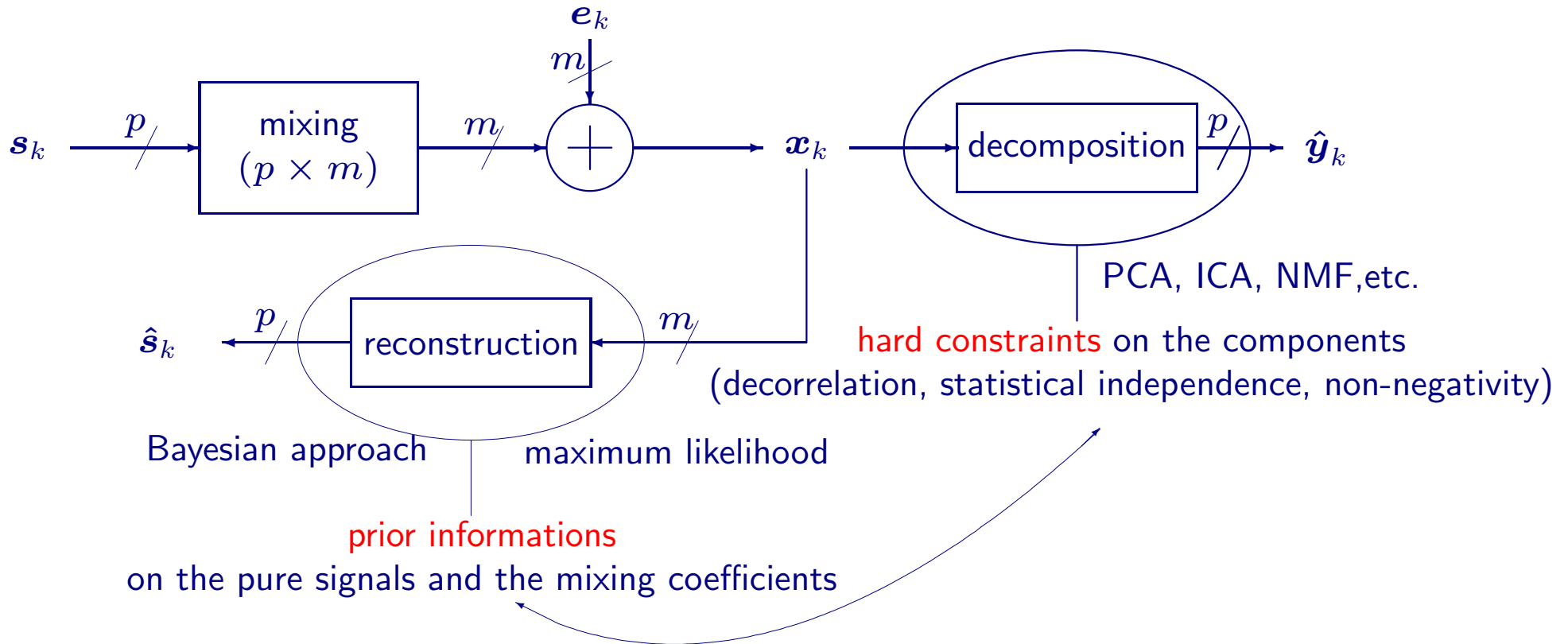
⊛ Separation approaches:



⊛ Separation approaches:



⊛ Separation approaches:



Similar results when the constraints match prior information

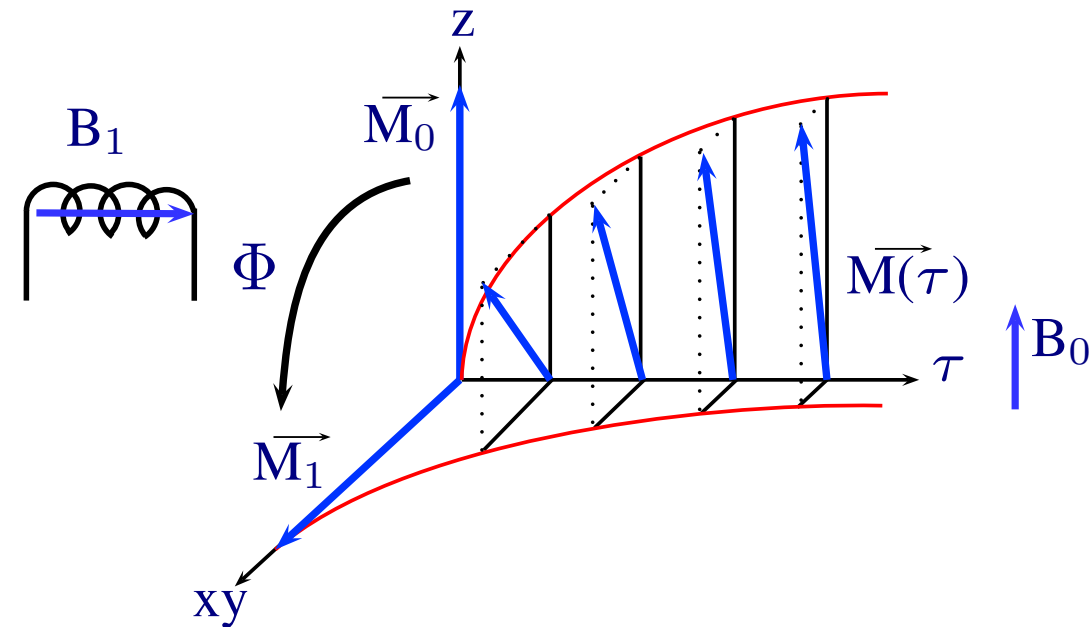
$$\hat{y}_k \equiv \hat{s}_k$$

Outline

1. Overview of signal processing problems
 - Inverse problems
 - Resolution approaches
 - Examples
2. Relaxation times distribution estimation
 - 1D & 2D relaxation models
 - Optimized reconstruction approach
 - Dedicated computing tool
3. Summary

2. Relaxation times distribution estimation

2.1 Data acquisition principle



- Static field $B_0 \Rightarrow$ nuclear spin alignment (z axis)
- Short magnetic pulse $B_1 \Rightarrow$ initial flip angle Φ
- Relaxation: return to the equilibrium state
 1. Longitudinal dynamics (z axis)
 $\Rightarrow T_1$ relaxation: $x_1(\tau) = M_z(\tau)$
 2. Transverse dynamics (xy plane)
 $\Rightarrow T_2$ relaxation: $x_2(\tau) = M_{xz}(\tau)$

⊗ **Quantities of interest:** values of longitudinal and transverse relaxation times T_1 and T_2 .

- ✓ Modeling of the relaxation signal decay
- ✓ Recover the relaxation times **distribution** from the measured data

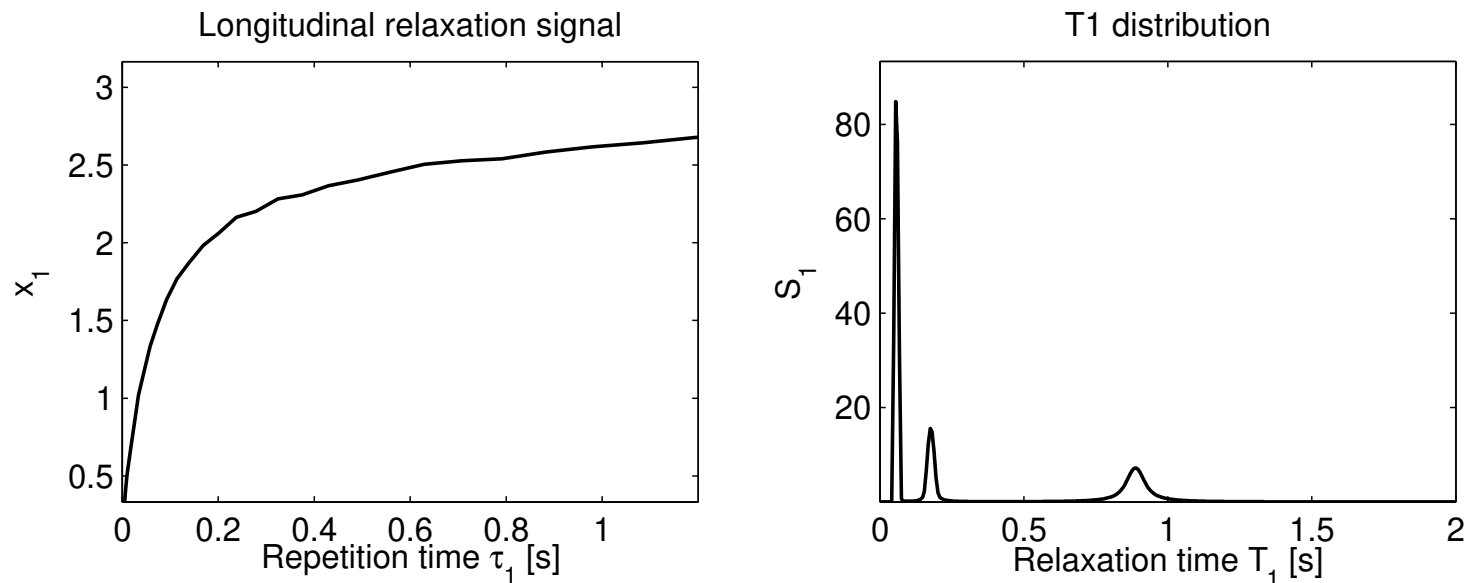
1) One-dimensionnal analysis

- ✓ Measured signal: Laplace transform of relaxation times distribution

$$x_i(\tau_i) = \int k_i(\tau_i, T_i) S(T_i) dT_i \longrightarrow \mathbf{y} = \mathbf{K}\mathbf{s} + \mathbf{e}$$

with $k_1(\tau_1) = 1 - \gamma e^{-\tau_1/T_1}$ in T_1 relaxation and $k_2(\tau_2) = e^{-\tau_2/T_2}$ for T_2 relaxation.

The parameter γ is related to the flip angle ($\gamma = 1 - \cos \Phi$).



∠ Find T_1 or T_2 distributions \implies Numerical inversion of a Laplace transform

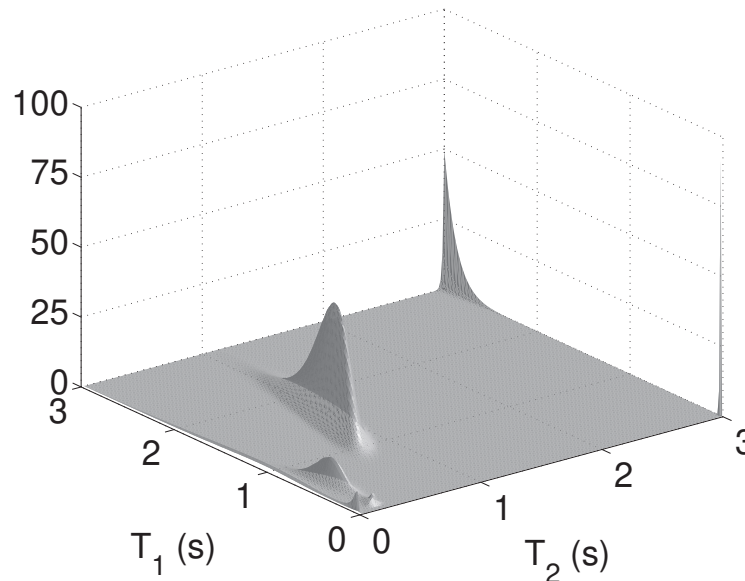
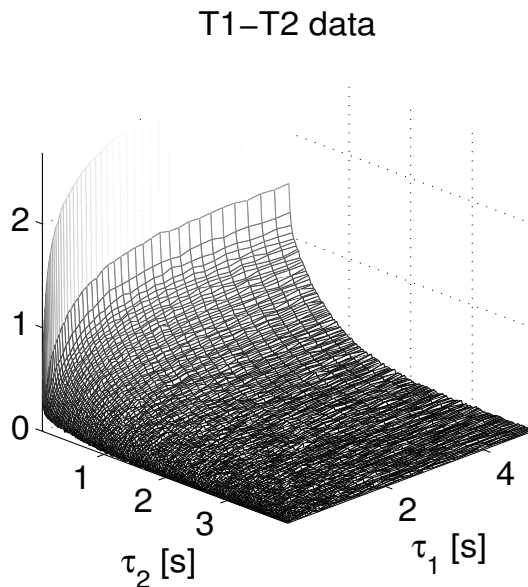
2) Two-dimensional analysis [English 1991]

- ✓ Apply two successive magnetic pulses with a predefined time spacing τ_1

$$x(\tau_1, \tau_2) = \int \int k_1(\tau_1, T_1) S(T_1, T_2) k_2(\tau_2, T_2) dT_1 dT_2$$

$$Y = K_1 S K_2^t + E \iff y = (K_1 \otimes K_2) s + e$$

- ✓ Find the joint distribution $S(T_1, T_2)$ of the relaxation time constants



2.2 Relaxation times estimation ... Problem statement

- $k_1(\tau_1, T_1)$ and $k_2(\tau_2, T_2)$ depend on the used NMR sequence (CPMG, FID) [Rondeau-Mouro, 2016]

$$k_2(\tau_2, T_2) = \begin{cases} e^{-\tau_2/T_2} & \text{for CPMG} \\ (e^{-\tau_2/T_2} + e^{-\tau_2/T_2^2}) \operatorname{sinc}\left(w \frac{\tau_2}{T_2}\right) & \text{for FID} \end{cases}$$

- Ill-conditioned matrices $\mathbf{K}_1, \mathbf{K}_2$. The singular values of \mathbf{K}_1 and \mathbf{K}_2 decay exponentially

◁ *Direct inversion yields unstable results and negative intensities*

- Large-size problem in the case of a two-dimensional analysis

1. $m_1 = 50$ repetition time values τ_1
2. $m_2 = 5000$ echo time instants τ_2
3. $N_1 = N_2 = 300$ values of T_1 and T_2

◁ *Matrix $\mathbf{K} := \mathbf{K}_1 \otimes \mathbf{K}_2$ of size $m_1 m_2 \times N_1 N_2$ contains over 10^{10} elements !!*

2.3 Relaxation times estimation ... Regularization framework

⊛ The relaxation times distribution is a solution of

$$\min_{\mathbf{s} \in \mathbb{R}^{N+}} \left(F(\mathbf{s}) = \frac{1}{2} \|\mathbf{y} - \mathbf{K}\mathbf{s}\|_2^2 + \beta R(\mathbf{s}) \right)$$

where $R(\mathbf{s})$ is a *convex and differentiable* regularization criterion

→ Solve a *non-negativity constrained* optimization problem

- ▷ Classical approach [Venkataramanan, 2002]
 - Data compression and Tikhonov regularization
 - Non-negativity constrained least squares
- ▷ Low complexity iterative resolution [Mariette, 1996] [Chouzenoux, 2010]
- ✓ Regularization criterion based on **Maximum entropy**

$$R(\mathbf{s}) = \mathbf{s}^t \log \mathbf{s} = \sum_{i=1}^N s_i \log s_i$$

⇒ ensure **non-negativity** and **sparsity** of the distribution

⊛ Optimized reconstruction algorithm

Solve

$$\min_{\mathbf{s} \in \mathbb{R}^{N+}} \left(F(\mathbf{s}) = \frac{1}{2} \|\mathbf{y} - \mathbf{K}\mathbf{s}\|_2^2 + \beta R(\mathbf{s}) \right)$$

Starting from an initial value \mathbf{s}_0 and for each iteration k , repeat steps (1) and (2) until convergence,

▷ (1) Truncated Newton direction

- Evaluate $(\mathbf{g}_k, \mathbf{H}_k)$, **first** and **second order derivatives** of $F(\mathbf{s})$ at \mathbf{s}_k ,
- Calculate the **descent direction** \mathbf{d}_k , approximate solution of a linear system $\mathbf{H}_k \mathbf{d} = -\mathbf{g}_k$

▷ (2) Descent step length

- Find a **step length** α_k ensuring a sufficient decrease of $\phi(\alpha) = F(\mathbf{s}_k + \alpha \mathbf{d}_k)$ (*dedicated search method*)
- Make the update of the distribution value : $\mathbf{s}_{k+1} = \mathbf{s}_k + \alpha_k \mathbf{d}_k$

▷ Implementation issues

- Solving the linear system $\mathbf{H}_k \mathbf{d} = -\mathbf{g}_k$ using preconditioned conjugate gradient
- Avoid the storing of matrix \mathbf{K} in the 2D case by exploiting the factorized form $\mathbf{Y} = \mathbf{K}_1 \mathbf{S} \mathbf{K}_2^t$
- On the fly tuning of the regularization parameter (Khi-2 test: *residuals at the same level as noise*)

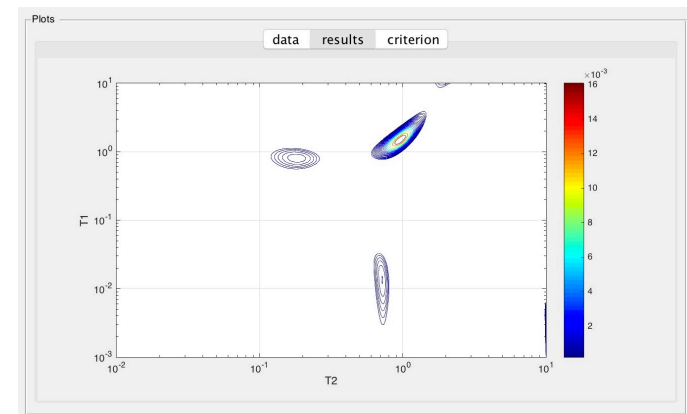
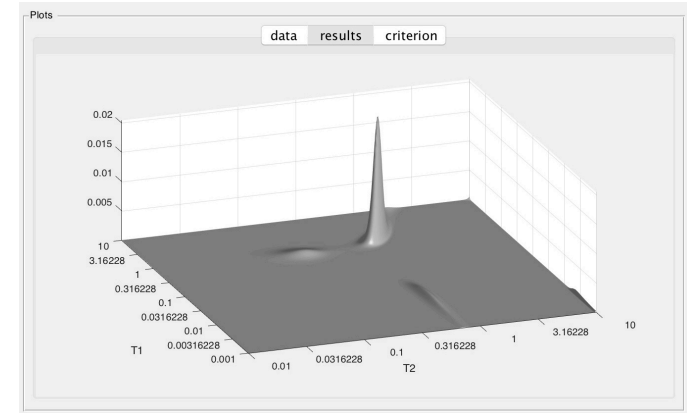
2.4 Relaxation times estimation ... Dedicated tool (Emilio)

* Entropy Maximization for Iterative Laplace Inversion by Optimization

The screenshot displays the 'Entropy Maximization for Iterative Laplace Inversion by Optimization' software interface. The window has tabs for 'Data', 'Results', and 'Advanced'. The 'Data' tab is active, showing 'T1-T2 data' with file name 'eau-ableS-55-001', first point [0.03; 0.0003], last point [1.5; 2.0001], and number of points [50; 10000]. The 'Model' section shows 'T1-T2' selected, with checkboxes for 'FID' and 'CPMG', and input fields for 'FID duration (mus)', 'Flip angle (deg)', 'FID frequency', 'min', 'max', 'length', and 'grid'. The 'Reconstruction' section includes 'Noise variance' (0.00010211), 'Regularization' (1), 'Algorithm' (T. Newton), and 'Advanced settings' (Set). The 'Results' section shows 'Iteration: 673', 'Khi2: 335.0199', 'RMSE (%): 63.7915', and checkboxes for 'T1', 'T2', and 'Contour plot'. The 'Plots' section shows a 3D surface plot of 'T1-T2 data' with axes for relaxation times and signal intensity.

* Main features of Emilio tool

1. Load **measurement data** file, (T_1 , T_2 and T_1 - T_2 data) including CPMG and/or FID data
2. Define the **forward model** parameters
 - Model on T_1 and T_2 relaxations
 - Flip angle value
 - FID sequence duration and frequency
3. Specify the **resolution algorithm** parameters
 - Noise variance estimation
 - Automatic setting of the regularization parameter
 - Algorithm stopping criteria
4. Exploit the **estimated results**
 - Calculate and visualize 1D/2D distributions
 - Quantify peak area
 - Export in various file formats

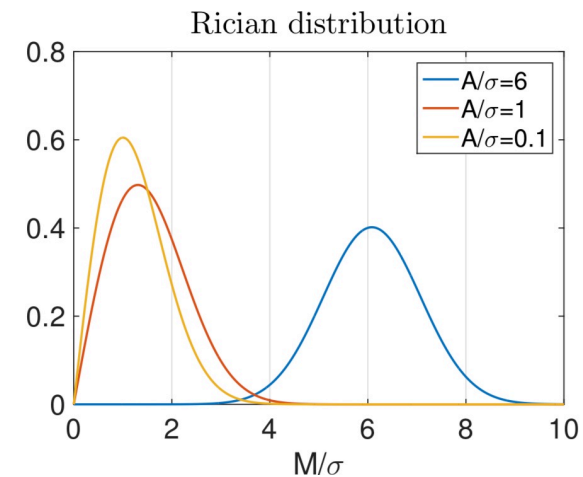


2.5 Relaxation times estimation ... Challenges

1) Processing magnitude signals [Jiang, 2017]

- Methods based on least squares assumes Gaussian noise
- Magnitude data follow a Rician distribution

$$p(M|A) = \frac{M}{\sigma^2} e^{-\frac{M^2+A^2}{2\sigma^2}} I_0\left(\frac{AM}{\sigma^2}\right)$$



⋄ Gaussian approximation valid only for high SNR

- ✓ Adequate optimization methods for Maximum likelihood estimation

2) Relaxation times from MRI data [El Hajj, 2018]

- Model the partial volume effect
- Account for spatial information
- Handle the large-scale optimization problem

3. Summary

- Signal processing methods
 - Need adequate forward models
 - Relevant information related to the sought quantities
- Relaxation times estimation
 - Exponential and/or Gaussian decay curves ... depending on the applied MR sequences
 - Maximum entropy regularization adapted for distribution estimation
 - Handle large-size optimization problems ... dedicated optimization tools
- Challenging issues
 - Accounting for the Rician noise model
 - Applications to MR images