Problèmes inverses

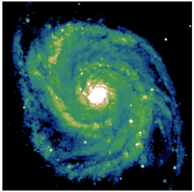
Ferréol Soulez

Centre de Recherche Astronomique de Lyon
Université Claude Bernard Lyon I
École Normale Supérieure de Lyon

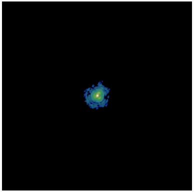
Master 2 AIMA / Télécom St Étienne / IOGS 2020

Introduction

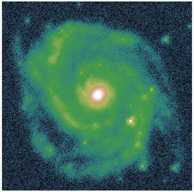
Exemple 1 : Convolution



$\star$

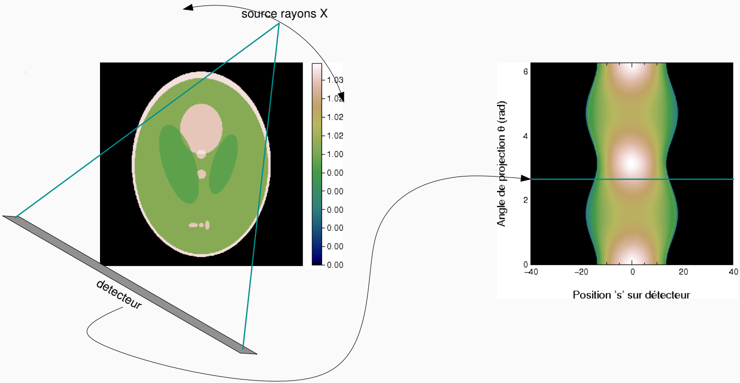


$+$ bruit $\longrightarrow$



- **observables**: image floue,
- **variables d'intérêt**: image nette,
- **modèle**: convolution par la réponse impulsionnelle (PSF).

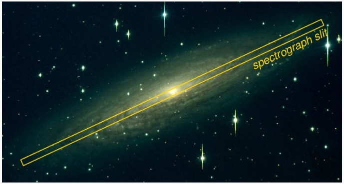
Exemple 2 : Tomographie



- **observables**: projections,
- **variables d'intérêt**: image de coupe,
- **modèle**: transformée de Radon.

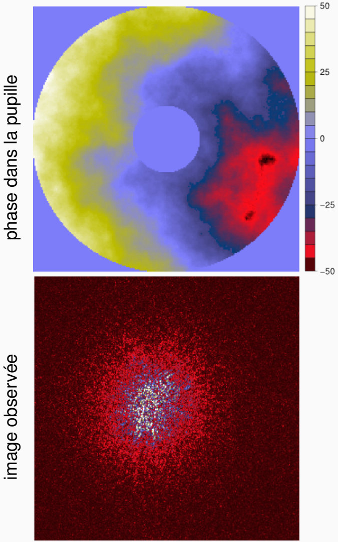
Exemple 3 : Dynamique galactique

- **observables :** distribution des vitesses



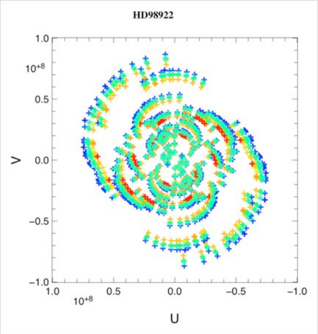
- **modèle physique :** dynamique disque galactique mince,
- **variables d'intérêt :** distribution des orbites et cinématique du gaz.

Exemple 4 : Reconstruction de phase

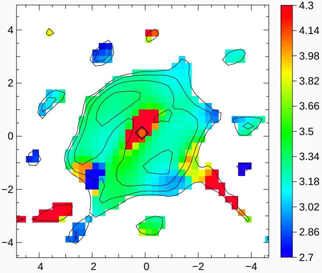


- **observables :** image dans le plan focal,
- **modèle physique :** optique de Fourier,
- **variables d'intérêt :** phase dans le plan pupille

plan (u,v)



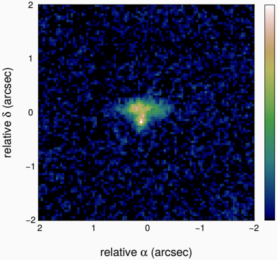
carte de température

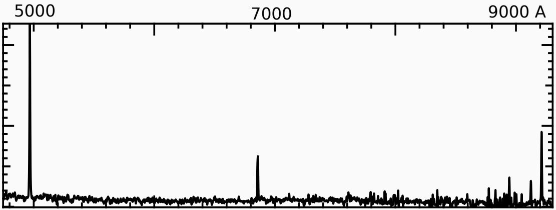


- **observables :** Mesures dans l'espace de Fourier (V2 + clotures),
- **modèle physique :** optique de Fourier + corps noir
- **variables d'intérêt :** température et flux du disque
- **résultats :** Herbig Be HD98922 (PIONIER/VLTI, 2012):
  - flux relatif du disque: 0.66 ,
  - température du disque: 1654 K .

Exemple 6 : Fusion de données

Image HST





Spectre mesuré avec MUSE

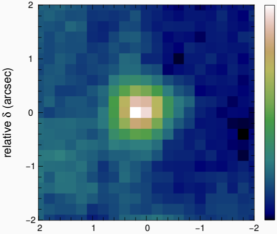
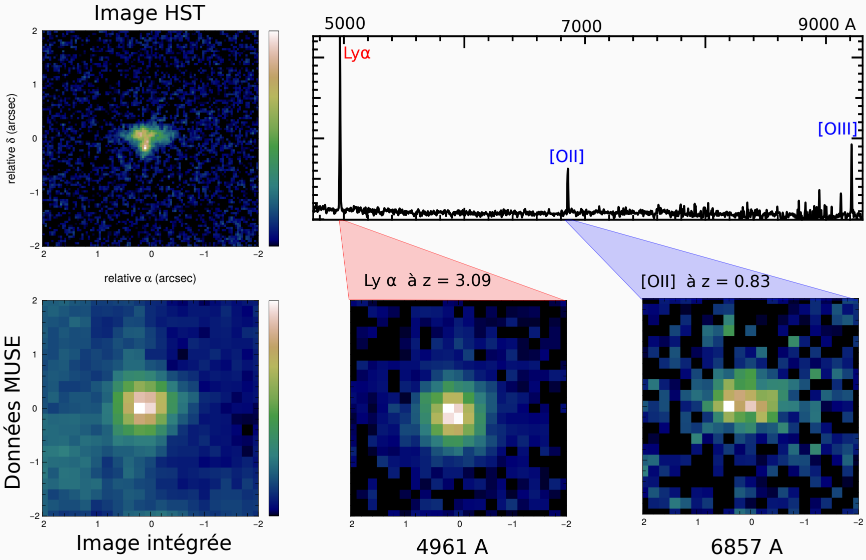
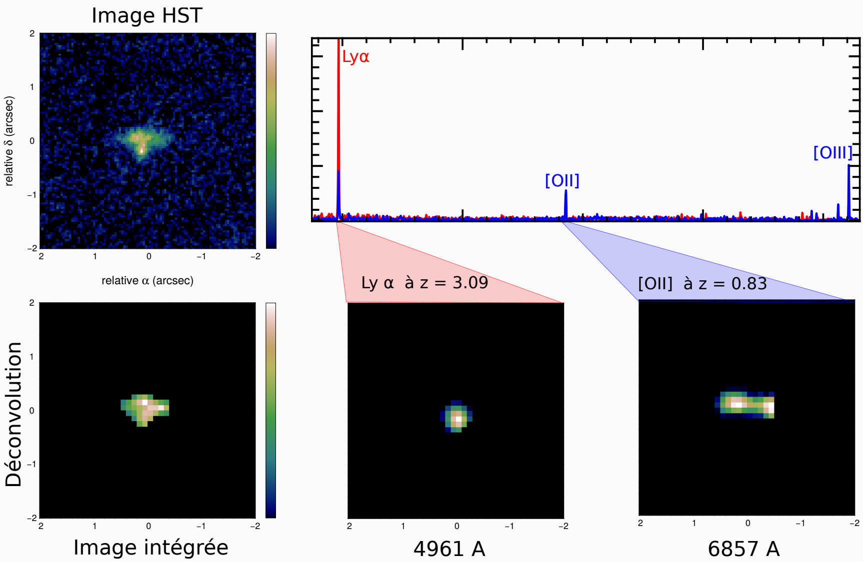


Image MUSE intégrée

Exemple 6 : Fusion de données



Exemple 6 : Fusion de données



Modèle d'observation

Object $f(u)$

Milieu d'observation

- diffusion,
- turbulence,
- ...

Instrument

- bougé,
- defocus,
- aberration,
- ...

Détecteur

- quantification,
- intégration,
- bruit ϵ,
- ...

Data y

Modèle direct de formation des données:
$y = \mathbf{H} \cdot x + \epsilon$

$\mathcal{H}()$: opérateur d'observation

Object $f(u)$

Milieu d’observation

- diffusion,
- turbulence,
- ...

Instrument

- bougé,
- defocus,
- aberration,
- ...

Détecteur

- quantification,
- intégration,
- bruit ϵ,
- ...

Data y

Modèle direct de formation des données:
$y = \mathbf{H} \cdot x + \epsilon$

$\mathcal{H}()$: opérateur d’observation

Qu'est-ce qu'un problème inverse ?

- Des observables disponibles et modélisables :

$$\boxed{y = \mathcal{H}(x) + e} \quad \Longleftarrow \quad \text{le } \textbf{\textit{modèle direct}}$$

  - les données $y \in \mathbb{R}^M$;
  - le modèle $\mathcal{H}$: $\mathbb{R}^N \mapsto \mathbb{R}^M$;
  - les variables d'intérêt $x \in \mathbb{R}^N$;
  - les erreurs $e \in \mathbb{R}^M$.
- **Objectif :** Retrouver les meilleurs paramètres x compte tenu des données y et du modèle $\mathcal{H}$.
- **Étapes :**
  - Contruire et analyser le modèle.
  - Établir un critère pour discriminer les meilleurs paramètres.
  - Élaborer une stratégie pour trouver la solution.

Qu'est-ce qu'un problème inverse ?

- Des observables disponibles et modélisables :

$$\boxed{y = \mathcal{H}(x) + e} \quad \Longleftarrow \quad \text{le } \textbf{\textit{modèle direct}}$$

  - les données $y \in \mathbb{R}^M$;
  - le modèle $\mathcal{H}$: $\mathbb{R}^N \mapsto \mathbb{R}^M$;
  - les variables d'intérêt $x \in \mathbb{R}^N$;
  - les erreurs $e \in \mathbb{R}^M$.
- **Objectif :** Retrouver les meilleurs paramètres x compte tenu des données y et du modèle $\mathcal{H}$.
- **Étapes :**
  - Contruire et analyser le modèle.
  - Établir un critère pour discriminer les meilleurs paramètres.
  - Élaborer une stratégie pour trouver la solution.

Qu'est-ce qu'un problème inverse ?

- Des observables disponibles et modélisables :

$$\boxed{\boldsymbol{y} = \mathcal{H}(\boldsymbol{x}) + \boldsymbol{e}} \quad \Longleftarrow \quad \text{le } \textbf{\textit{modèle direct}}$$

  - les données $\boldsymbol{y} \in \mathbb{R}^M$;
  - le modèle $\mathcal{H}\colon \mathbb{R}^N \mapsto \mathbb{R}^M$;
  - les variables d'intérêt $\boldsymbol{x} \in \mathbb{R}^N$;
  - les erreurs $\boldsymbol{e} \in \mathbb{R}^M$.
- **Objectif :** Retrouver les meilleurs paramètres $\boldsymbol{x}$ compte tenu des données $\boldsymbol{y}$ et du modèle $\mathcal{H}$.
- **Étapes :**
  - Contruire et analyser le modèle.
  - Établir un critère pour discriminer les meilleurs paramètres.
  - Élaborer une stratégie pour trouver la solution.

Modèle linéaire / non-linéaire

1. modèle linéaire :

$$\boxed{y = \mathbf{H} \cdot x + e}$$

- intégrale de Fredholm du 1[er] ordre (transformée de Radon, de Fourier, convolution, ...) ;
- base de fonctions ;

2. modèle non-linéaire :

$$\boxed{y = \mathcal{H}(x) + e}$$

- tout ou partie du noyau $h(r, s)$ à retrouver : déconvolution aveugle, synthèse spectrale, ...
- mesures de corrélation (à 2 ou 3 points)
- restoration de phase : cristallographie, clôtures de phases (interférométrie), ...

$\Rightarrow$ unicité ?

Modèle linéaire / non-linéaire

1. modèle linéaire :

$$\boxed{\boldsymbol{y} = \mathbf{H} \cdot \boldsymbol{x} + \boldsymbol{e}}$$

   - intégrale de Fredholm du 1er ordre (transformée de Radon, de Fourier, convolution, ...) ;
   - base de fonctions ;

2. modèle non-linéaire :

$$\boxed{\boldsymbol{y} = \mathcal{H}(\boldsymbol{x}) + \boldsymbol{e}}$$

   - tout ou partie du noyau $h(\boldsymbol{r}, \boldsymbol{s})$ à retrouver : déconvolution aveugle, synthèse spectrale, ...
   - mesures de corrélation (à 2 ou 3 points)
   - restoration de phase : cristallographie, clôtures de phases (interférométrie), ...

   $\Rightarrow$ unicité ?

Le modèle direct linéaire

Discrétisation de l'objet

$f(\boldsymbol{u})$ décrit par un nombre fini de paramètres $\boldsymbol{x}$ en utilisant la base $b_m(\boldsymbol{u})$:

$$f(\boldsymbol{u}) \approx \sum_n x_n\, b_n(\boldsymbol{u})\,.$$

Erreur de discrétisation : $\|f - \sum_n x_n\, b_n\|_2^2 > 0$.

Échantillonage par le détecteur

Distribution d'intensité $g(\boldsymbol{s})$ échantillonnée par le détecteur:

$$y_m = \int g(\boldsymbol{s})\, c_m(\boldsymbol{s})\, \mathrm{d}\boldsymbol{s} + e_m\,,$$

– $c_m(\boldsymbol{s})$: fonction d'intégration sur le pixel n
– e_m: bruit additif de variance σ_m.

Discrétisation

Discrétisation de l'objet

$f(\boldsymbol{u})$ décrit par un nombre fini de paramètres $\boldsymbol{x}$ en utilisant la base $b_m(\boldsymbol{u})$:

$$f(\boldsymbol{u}) \approx \sum_n x_n \, b_n(\boldsymbol{u}) \, .$$

Erreur de discrétisation : $\|f - \sum_n x_n \, b_n\|_2^2 > 0$.

Échantillonage par le détecteur

Distribution d'intensité $g(\boldsymbol{s})$ échantillonnée par le détecteur:

$$y_m = \int g(\boldsymbol{s}) \, c_m(\boldsymbol{s}) \, \mathrm{d}\boldsymbol{s} + e_m \, ,$$

– $c_m(\boldsymbol{s})$: fonction d'intégration sur le pixel n
– e_m: bruit additif de variance σ_m.

Discrétisation

Distribution d'intensité $g(\boldsymbol{s})$ échantillonnée par le détecteur:

$$\begin{aligned} y_m &= \int \left(\int \mathcal{H}(\boldsymbol{s},\boldsymbol{u}) \sum_n x_n\, b_n(\boldsymbol{u}) \mathrm{d}\boldsymbol{u} \right) c_m(\boldsymbol{s})\, \mathrm{d}\boldsymbol{s} + e_m \\ &= \sum_n x_n \iint \mathcal{H}(\boldsymbol{s},\boldsymbol{u})\, b_n(\boldsymbol{u})\, c_m(\boldsymbol{s})\, \mathrm{d}\boldsymbol{s}\, \mathrm{d}\boldsymbol{u} + e_m \\ \boldsymbol{y} &= \mathbf{H} \cdot \boldsymbol{x} + \boldsymbol{e}\,, \end{aligned}$$

avec l'opérateur (matrice):

$$H_{m,n} = \iint \mathcal{H}(\boldsymbol{s},\boldsymbol{u})\, b_n(\boldsymbol{u}) c_m(\boldsymbol{s})\, \mathrm{d}\boldsymbol{s} \mathrm{d}\boldsymbol{u}\,,$$

et le terme d'erreur $\boldsymbol{e}$ incluant les erreurs de discrétisation.

Pour une intégrale de Fredholm $g(s) = \int \mathcal{H}(s,u)\, f(u) \mathrm{d}u$, l'estimation de $f(u)$ étant données $g(s)$ et $\mathcal{H}(s,u)$ est un **problème mal posé**.

Conditions d'hadamard

Un problème est *bien posé* si et seulement si sa solution z:

- existe $z \in \mathrm{Im}(\mathcal{H})$,
- est unique $\mathrm{Ker}(\mathcal{H}) = \{0\}$,
- est stable $\|{*}\|\, z - z' \to 0 \Rightarrow \|\mathcal{H}(z) - \mathcal{H}(z')\| \to 0$.

Conditions d'Hadamard

Pour une intégrale de Fredholm $g(\boldsymbol{s}) = \int \mathcal{H}(\boldsymbol{s},\boldsymbol{u})\, f(\boldsymbol{u})\mathrm{d}\boldsymbol{u}$, l'estimation de $f(\boldsymbol{u})$ étant données $g(\boldsymbol{s})$ et $\mathcal{H}(\boldsymbol{s},\boldsymbol{u})$ est un **problème mal posé**.

Conditions d'hadamard

Un problème est *bien posé* si et seulement si sa solution z:

- existe $z \in \mathrm{Im}(\mathcal{H})$,
- est unique $\mathrm{Ker}(\mathcal{H}) = \{0\}$,
- est stable $\|*\|\, z - z' \to 0 \Rightarrow \|\mathcal{H}(z) - \mathcal{H}(z')\| \to 0$.

Pseudo inverse:

Pour un problème discret $\boldsymbol{y} = \mathbf{H} \cdot \boldsymbol{x}$, le problème inverse possède toujours une inverse généralisée:

Moore-Penrose pseudo inverse

$$\mathbf{H}^{\dagger} = \left(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{H}\right)^{-1} \cdot \mathbf{H}^{\mathrm{T}}$$

$$\mathbf{H}^{\dagger} = \mathbf{H}^{\mathrm{T}} \cdot \left(\mathbf{H} \cdot \mathbf{H}^{\mathrm{T}}\right)^{-1}$$

si des valeurs propres λ_i de $\mathbf{H}^{\mathrm{T}} \cdot \mathbf{H}$ sont nulles on admet $\lambda_i^{-1} = 0$ (moindre norme).
La solution vérifie:

$$\boldsymbol{z} = \mathbf{H}^{\dagger} \cdot \boldsymbol{y} \Rightarrow \begin{cases} \|\boldsymbol{z}\|_2^2 \leq \|\boldsymbol{x}\|_2^2 \ , \ \forall \boldsymbol{x} \\ \|\mathbf{H} \cdot \boldsymbol{z} - \boldsymbol{y}\|_2^2 \leq \|\mathbf{H} \cdot \boldsymbol{x} - \boldsymbol{y}\|_2^2 \ , \ \forall \boldsymbol{x} \end{cases}$$

Problème mal conditionné

Valeurs singulières σ de $\mathbf{H}$:

$$\begin{aligned}\mathbf{H}\,\boldsymbol{u}_n &= \sigma_n\,\boldsymbol{v}_n\\ \mathbf{H}^*\,\boldsymbol{v}_n &= \sigma_n\,\boldsymbol{u}_n\,,\end{aligned}$$

Pour une petite perturbation additive $\delta\boldsymbol{y} = \varepsilon\boldsymbol{v}_n$ des données, la perturbation correspondante sur les paramètres $\delta\boldsymbol{x}$ est donnée par:

$$\begin{aligned}\mathbf{H}\cdot\delta\boldsymbol{x} &= \delta\boldsymbol{y}\\ &= \varepsilon\,\boldsymbol{v}_n\,,\\ &= \varepsilon\,\sigma_n^{-1}\,\sigma_n\,\boldsymbol{v}_n\,,\\ &= \varepsilon\,\sigma_n^{-1}\,\mathbf{H}\cdot\boldsymbol{u}_n\,,\\ \delta\boldsymbol{x} &= \varepsilon\,\sigma_n^{-1}\,\boldsymbol{u}_n\,.\end{aligned}$$

La perturbation est amplifiée par $\frac{||\delta\boldsymbol{x}||}{||\delta\boldsymbol{y}||} = \sigma_n^{-1} \Rightarrow$ **le système est instable**

La stabilité est quantifiée par le conditionnement $C(\mathbf{H}) = \frac{|\sigma_{\max}|}{|\sigma_{\min}|}$

Problème mal conditionné

Valeurs singulières σ de $\mathbf{H}$:

$$\begin{aligned} \mathbf{H}\,\boldsymbol{u}_n &= \sigma_n\,\boldsymbol{v}_n \\ \mathbf{H}^*\,\boldsymbol{v}_n &= \sigma_n\,\boldsymbol{u}_n\,, \end{aligned}$$

Pour une petite perturbation additive $\delta\boldsymbol{y} = \varepsilon\boldsymbol{v}_n$ des données, la perturbation correspondante sur les paramètres $\delta\boldsymbol{x}$ est donnée par:

$$\begin{aligned} \mathbf{H}\cdot\delta\boldsymbol{x} &= \delta\boldsymbol{y} \\ &= \varepsilon\,\boldsymbol{v}_n\,, \\ &= \varepsilon\,\sigma_n^{-1}\,\sigma_n\,\boldsymbol{v}_n\,, \\ &= \varepsilon\,\sigma_n^{-1}\,\mathbf{H}\cdot\boldsymbol{u}_n\,, \\ \delta\boldsymbol{x} &= \varepsilon\,\sigma_n^{-1}\,\boldsymbol{u}_n\,. \end{aligned}$$

La perturbation est amplifiée par $\frac{||\delta\boldsymbol{x}||}{||\delta\boldsymbol{y}||} = \sigma_n^{-1} \Rightarrow$ **le système est instable**

La stabilité est quantifiée par le conditionnement $C(\mathbf{H}) = \frac{|\sigma_{\max}|}{|\sigma_{\min}|}$

Problème mal conditionné

Valeurs singulières σ de $\mathbf{H}$:

$$\begin{aligned} \mathbf{H}\,\boldsymbol{u}_n &= \sigma_n\,\boldsymbol{v}_n \\ \mathbf{H}^*\,\boldsymbol{v}_n &= \sigma_n\,\boldsymbol{u}_n\,, \end{aligned}$$

Pour une petite perturbation additive $\delta\boldsymbol{y} = \varepsilon\boldsymbol{v}_n$ des données, la perturbation correspondante sur les paramètres $\delta\boldsymbol{x}$ est donnée par:

$$\begin{aligned} \mathbf{H}\cdot\delta\boldsymbol{x} &= \delta\boldsymbol{y} \\ &= \varepsilon\,\boldsymbol{v}_n\,, \\ &= \varepsilon\,\sigma_n^{-1}\,\sigma_n\,\boldsymbol{v}_n\,, \\ &= \varepsilon\,\sigma_n^{-1}\,\mathbf{H}\cdot\boldsymbol{u}_n\,, \\ \delta\boldsymbol{x} &= \varepsilon\,\sigma_n^{-1}\,\boldsymbol{u}_n\,. \end{aligned}$$

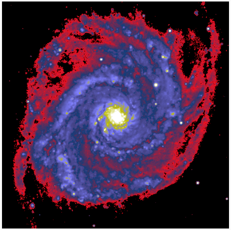
La perturbation est amplifiée par $\dfrac{\|\delta\boldsymbol{x}\|}{\|\delta\boldsymbol{y}\|} = \sigma_n^{-1} \Rightarrow$ **le système est instable**

La stabilité est quantifiée par le conditionnement $C(\mathbf{H}) = \frac{|\sigma_{\max}|}{|\sigma_{\min}|}$

Naive deconvolution

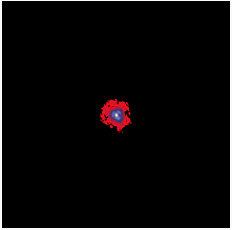
Simulation de M51

objet



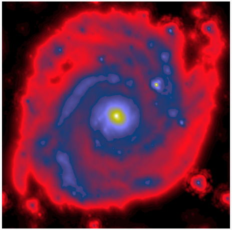
$\star$

PSF



$\mapsto$

objet flou



$+$

bruit

$\mapsfrom$

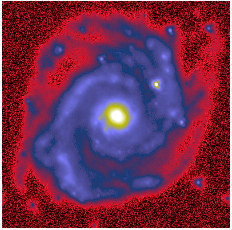


image observée

On suppose la PSF invariante par translation (isoplanétisme).

$$\boldsymbol{h} = \begin{bmatrix} h_0 \\ h_1 \\ h_2 \\ \vdots \\ h_{-2} \\ h_{-1} \end{bmatrix} \qquad \mathbf{H} = \begin{bmatrix} h_0 & h_{-1} & h_{-2} & \dots & h_{-N} \\ h_1 & h_0 & h_{-1} & \ddots & h_{-N+1} \\ h_2 & h_1 & h_0 & \ddots & h_{-N+2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ h_{N-2} & h_{N-3} & h_{N-2} & \ddots & h_{-1} \\ h_{N-1} & h_{N-2} & h_{N-3} & \dots & h_0 \end{bmatrix} \qquad \mathbf{H}_{\text{circ}} = \begin{bmatrix} h_0 & h_{-1} & h_{-2} & \dots & h_1 \\ h_1 & h_0 & h_{-1} & \ddots & h_2 \\ h_2 & h_1 & h_0 & \ddots & h_3 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ h_{-2} & h_{-3} & h_{-4} & \ddots & h_{-1} \\ h_{-1} & h_{-2} & h_{-3} & \dots & h_0 \end{bmatrix}$$

PSF

matrice de convolution

$\text{Card}(\boldsymbol{h}) = 2\,N$

approximation circulante

$\text{Card}(\boldsymbol{h}) = N$

approximation circulante

$\mathbf{H}_{\text{circ}}$ est diagonalisable dans l'espace de Fourier $\mathbf{H} = \mathbf{F}^{-1}\ \text{diag}(\widehat{\boldsymbol{h}})\,\mathbf{F}$ avec $\widehat{\boldsymbol{h}} = \mathbf{F}\,\boldsymbol{h}$.

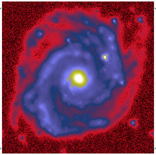
Si le bruit est négligeable, l’inversion directe est une division dans l’espace de Fourier:

$$\widehat{x}_k^{(\text{direct})} = \frac{\widehat{y}_k}{\widehat{h}_k} = \widehat{x}_k + \frac{\widehat{n}_k}{\widehat{h}_k}$$

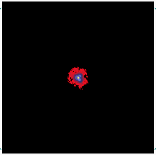
On a donc:

$\boldsymbol{x}^{(\text{direct})} = \mathrm{FFT}^{-1}$

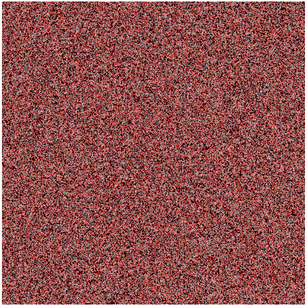
FFT



FFT



$=$



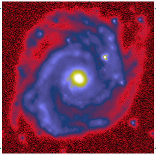
Inversion directe

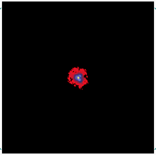
Si le bruit est négligeable, l'inversion directe est une division dans l'espace de Fourier:

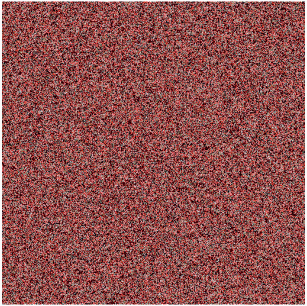
$$\widehat{x}_k^{\text{(direct)}} = \frac{\widehat{y}_k}{\widehat{h}_k} = \widehat{x}_k + \frac{\widehat{n}_k}{\widehat{h}_k}$$

On a donc:

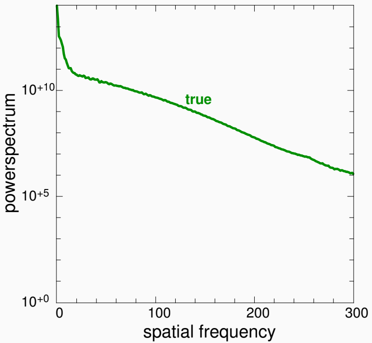
$$\boldsymbol{x}^{\text{(direct)}} = \mathrm{FFT}^{-1}\left(\frac{\mathrm{FFT}(\text{image 1})}{\mathrm{FFT}(\text{image 2})} \right) = \text{image 3}$$





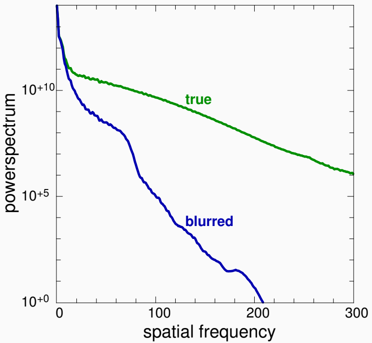


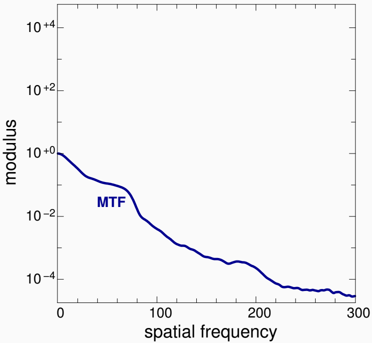
Explication



→ inversion directe: *amplification du bruit*

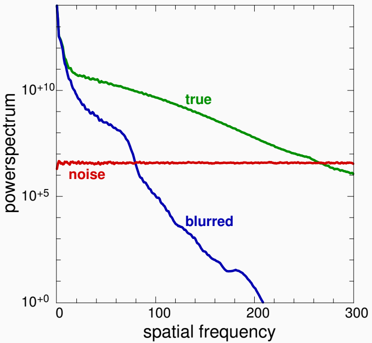
Explication

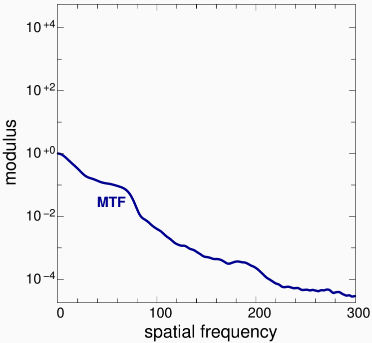




→ inversion directe: *amplification du bruit*

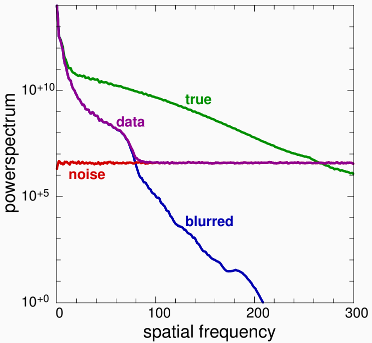
Explication

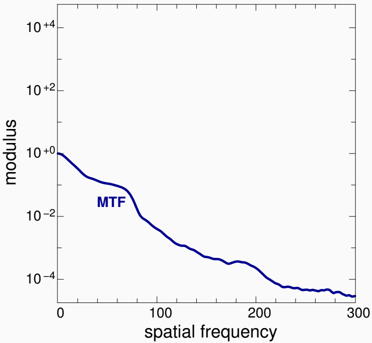




→ inversion directe: *amplification du bruit*

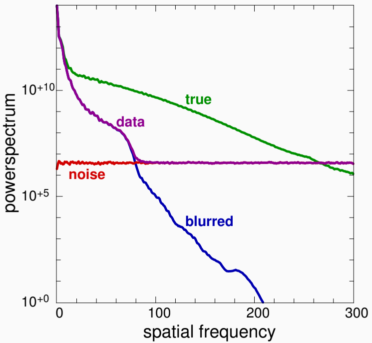
Explication

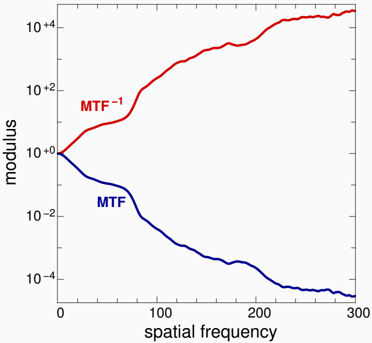




→ inversion directe: *amplification du bruit*

Explication



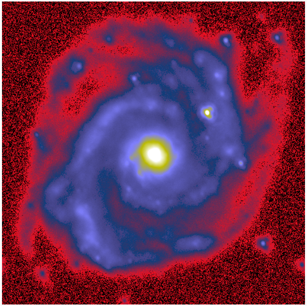


→ inversion directe: *amplification du bruit*

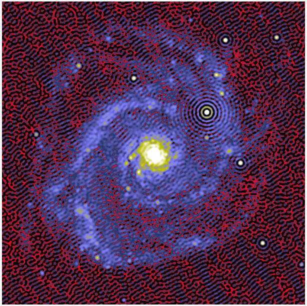
Troncature en fréquences

On peut essayer d'éviter l'amplification du bruit en introduisant une fréquence de coupure u_{cut}:

$$\widehat{x}_k^{(\text{cut})} = \begin{cases} \dfrac{\widehat{y}_k}{\widehat{h}_k} & \text{si } |\boldsymbol{u}_k| < u_{\text{cut}} \\ 0 & \text{sinon} \end{cases}$$



$\overset{\mapsto}{\text{(with } u_{\text{cut}} = 80 \text{ frequels)}}$



Filtre de Wiener

Filtre de Wiener: Principes

Le filtre de Wiener est dérivé des hypothèses suivantes:

- l'estimation $\boldsymbol{x}_{\mathrm{MMSE}}$ de l'objet vrai $\boldsymbol{x}$ est donné par une transformation linéaire des données $\boldsymbol{y}$:

  $$\boldsymbol{x}_{\mathrm{MMSE}} \equiv \widetilde{\mathbf{G}} \cdot \boldsymbol{y}$$

  Rationale: Cela peut être vu comme l'approximation au premier ordre du mapping $\boldsymbol{y} \mapsto \boldsymbol{x}_{\mathrm{MMSE}}$.
- les paramètres de cette transformation sont tel que l'erreur quadratique moyenne (MSE) par rapport à l'objet vrai $\boldsymbol{x}$ est minimale:

  $$\widetilde{\mathbf{G}} = \arg\min_{\mathbf{G}} \mathrm{E}\left\{ \frac{1}{2} \left\| \underbrace{\mathbf{G} \cdot \boldsymbol{y}}_{\text{estimator}} - \boldsymbol{x} \right\|^2 \right\}$$

→ *linear minimum mean squared error (LMMSE) estimator!*

seems magic!

Filtre de Wiener: Principes

Le filtre de Wiener est dérivé des hypothèses suivantes:

- l'estimation $\boldsymbol{x}_{\mathrm{MMSE}}$ de l'objet vrai $\boldsymbol{x}$ est donné par une transformation linéaire des données $\boldsymbol{y}$:

  $$\boldsymbol{x}_{\mathrm{MMSE}} \equiv \widetilde{\mathbf{G}} \cdot \boldsymbol{y}$$

  Rationale: Cela peut être vu comme l'approximation au premier ordre du mapping $\boldsymbol{y} \mapsto \boldsymbol{x}_{\mathrm{MMSE}}$.
- les paramètres de cette transformation sont tel que l'erreur quadratique moyenne (MSE) par rapport à l'objet vrai $\boldsymbol{x}$ est minimale:

  $$\widetilde{\mathbf{G}} = \arg\min_{\mathbf{G}} \mathrm{E}\left\{\frac{1}{2}\left\|\underbrace{\mathbf{G}\cdot\boldsymbol{y}}_{\text{estimator}} - \boldsymbol{x}\right\|^2\right\}$$

$\longrightarrow$ ***linear minimum mean squared error (LMMSE) estimator!***

seems magic!

Filtre de Wiener: Principes

Le filtre de Wiener est dérivé des hypothèses suivantes:

- l'estimation $\boldsymbol{x}_{\mathrm{MMSE}}$ de l'objet vrai $\boldsymbol{x}$ est donné par une transformation linéaire des données $\boldsymbol{y}$:

  $$\boldsymbol{x}_{\mathrm{MMSE}} \equiv \widetilde{\mathbf{G}} \cdot \boldsymbol{y}$$

  Rationale: Cela peut être vu comme l'approximation au premier ordre du mapping $\boldsymbol{y} \mapsto \boldsymbol{x}_{\mathrm{MMSE}}$.
- les paramètres de cette transformation sont tel que l'erreur quadratique moyenne (MSE) par rapport à l'objet vrai $\boldsymbol{x}$ est minimale:

  $$\widetilde{\mathbf{G}} = \arg\min_{\mathbf{G}} \mathrm{E}\left\{\frac{1}{2}\left\|\underbrace{\mathbf{G}\cdot\boldsymbol{y}}_{\text{estimator}} - \boldsymbol{x}\right\|^2\right\}$$

$\longrightarrow$ ***linear minimum mean squared error (LMMSE) estimator!***

seems magic!

Filtre de Wiener: dérivation

$$\widetilde{\mathbf{G}} = \arg\min_{\mathbf{G}} \mathrm{E}\left\{\frac{1}{2}\|\mathbf{G}\cdot\boldsymbol{y}-\boldsymbol{x}\|^2\right\} \Leftrightarrow \left.\frac{\partial \mathrm{E}\left\{\frac{1}{2}\|\mathbf{G}\cdot\boldsymbol{y}-\boldsymbol{x}\|^2\right\}}{\partial \mathbf{G}}\right|_{\mathbf{G}=\widetilde{\mathbf{G}}} = 0$$

$$\begin{aligned}
\frac{\partial \mathrm{E}\left\{\frac{1}{2}\|\mathbf{G}\cdot\boldsymbol{y}-\boldsymbol{x}\|^2\right\}}{\partial \mathbf{G}} &= \frac{1}{2}\mathrm{E}\left\{\frac{\partial\|\mathbf{G}\cdot\boldsymbol{y}-\boldsymbol{x}\|^2}{\partial \mathbf{G}}\right\} \\
&= \frac{1}{2}\mathrm{E}\left\{2\,(\mathbf{G}\cdot\boldsymbol{y}-\boldsymbol{x})\cdot\boldsymbol{y}^{\mathrm{T}}\right\} \\
&= \mathrm{E}\left\{\mathbf{G}\cdot\boldsymbol{y}\cdot\boldsymbol{y}^{\mathrm{T}}\right\} - \mathrm{E}\left\{\boldsymbol{x}\cdot\boldsymbol{y}^{\mathrm{T}}\right\} \\
&= \mathbf{G}\cdot\mathrm{E}\left\{\boldsymbol{y}\cdot\boldsymbol{y}^{\mathrm{T}}\right\} - \mathrm{E}\left\{\boldsymbol{x}\cdot\boldsymbol{y}^{\mathrm{T}}\right\}
\end{aligned}$$

$$\widetilde{\mathbf{G}} = \mathrm{E}\left\{\boldsymbol{x}\cdot\boldsymbol{y}^{\mathrm{T}}\right\}\cdot\mathrm{E}\left\{\boldsymbol{y}\cdot\boldsymbol{y}^{\mathrm{T}}\right\}^{-1}$$

Filtre de Wiener: dérivation

$$\widetilde{\mathbf{G}} = \arg\min_{\mathbf{G}} \mathrm{E}\left\{\frac{1}{2}\left\|\mathbf{G}\cdot\boldsymbol{y}-\boldsymbol{x}\right\|^2\right\} \Leftrightarrow \left.\frac{\partial \mathrm{E}\left\{\frac{1}{2}\left\|\mathbf{G}\cdot\boldsymbol{y}-\boldsymbol{x}\right\|^2\right\}}{\partial \mathbf{G}}\right|_{\mathbf{G}=\widetilde{\mathbf{G}}} = 0$$

$$\begin{aligned}
\frac{\partial \mathrm{E}\left\{\frac{1}{2}\left\|\mathbf{G}\cdot\boldsymbol{y}-\boldsymbol{x}\right\|^2\right\}}{\partial \mathbf{G}} &= \frac{1}{2}\,\mathrm{E}\left\{\frac{\partial\left\|\mathbf{G}\cdot\boldsymbol{y}-\boldsymbol{x}\right\|^2}{\partial \mathbf{G}}\right\} \\
&= \frac{1}{2}\,\mathrm{E}\left\{2\,(\mathbf{G}\cdot\boldsymbol{y}-\boldsymbol{x})\cdot\boldsymbol{y}^{\mathrm{T}}\right\} \\
&= \mathrm{E}\left\{\mathbf{G}\cdot\boldsymbol{y}\cdot\boldsymbol{y}^{\mathrm{T}}\right\} - \mathrm{E}\left\{\boldsymbol{x}\cdot\boldsymbol{y}^{\mathrm{T}}\right\} \\
&= \mathbf{G}\cdot\mathrm{E}\left\{\boldsymbol{y}\cdot\boldsymbol{y}^{\mathrm{T}}\right\} - \mathrm{E}\left\{\boldsymbol{x}\cdot\boldsymbol{y}^{\mathrm{T}}\right\}
\end{aligned}$$

$$\widetilde{\mathbf{G}} = \mathrm{E}\left\{\boldsymbol{x}\cdot\boldsymbol{y}^{\mathrm{T}}\right\}\cdot\mathrm{E}\left\{\boldsymbol{y}\cdot\boldsymbol{y}^{\mathrm{T}}\right\}^{-1}$$

Filtre de Wiener: dérivation

$$\widetilde{\mathbf{G}} = \arg\min_{\mathbf{G}} \mathrm{E}\left\{\frac{1}{2}\,\|\mathbf{G}\cdot\boldsymbol{y}-\boldsymbol{x}\|^2\right\} \Leftrightarrow \left.\frac{\partial \mathrm{E}\left\{\frac{1}{2}\,\|\mathbf{G}\cdot\boldsymbol{y}-\boldsymbol{x}\|^2\right\}}{\partial \mathbf{G}}\right|_{\mathbf{G}=\widetilde{\mathbf{G}}} = 0$$

$$\begin{aligned}
\frac{\partial \mathrm{E}\left\{\frac{1}{2}\,\|\mathbf{G}\cdot\boldsymbol{y}-\boldsymbol{x}\|^2\right\}}{\partial \mathbf{G}} &= \frac{1}{2}\,\mathrm{E}\left\{\frac{\partial\,\|\mathbf{G}\cdot\boldsymbol{y}-\boldsymbol{x}\|^2}{\partial \mathbf{G}}\right\} \\
&= \frac{1}{2}\,\mathrm{E}\left\{2\,(\mathbf{G}\cdot\boldsymbol{y}-\boldsymbol{x})\cdot\boldsymbol{y}^{\mathrm{T}}\right\} \\
&= \mathrm{E}\left\{\mathbf{G}\cdot\boldsymbol{y}\cdot\boldsymbol{y}^{\mathrm{T}}\right\} - \mathrm{E}\left\{\boldsymbol{x}\cdot\boldsymbol{y}^{\mathrm{T}}\right\} \\
&= \mathbf{G}\cdot\mathrm{E}\left\{\boldsymbol{y}\cdot\boldsymbol{y}^{\mathrm{T}}\right\} - \mathrm{E}\left\{\boldsymbol{x}\cdot\boldsymbol{y}^{\mathrm{T}}\right\}
\end{aligned}$$

$$\widetilde{\mathbf{G}} = \mathrm{E}\left\{\boldsymbol{x}\cdot\boldsymbol{y}^{\mathrm{T}}\right\}\cdot\mathrm{E}\left\{\boldsymbol{y}\cdot\boldsymbol{y}^{\mathrm{T}}\right\}^{-1}$$

Hypothèses

- Bruit centré: $\mathrm{E}\{\boldsymbol{e}\} = 0$,
- Objet centré: $\mathrm{E}\{\boldsymbol{x}\} = 0$,
- Le bruit et l'objet sont mutuellement indépendant: $\mathrm{Cov}(\boldsymbol{x}, \boldsymbol{e}) = 0$

- $\mathrm{E}\{\boldsymbol{y}\} = \mathrm{E}\{\mathbf{H} \cdot \boldsymbol{y} + \boldsymbol{e}\} = \mathbf{H} \cdot \mathrm{E}\{\boldsymbol{y}\} + \mathrm{E}\{\boldsymbol{e}\} = 0$
- $\mathrm{E}\left\{\boldsymbol{x} \cdot \boldsymbol{y}^{\mathrm{T}}\right\} = \underbrace{\mathrm{Cov}\left(\boldsymbol{x}, \boldsymbol{y}\right)}_{\mathbf{C}_{x,y}}$
- $\mathrm{E}\left\{\boldsymbol{y} \cdot \boldsymbol{y}^{\mathrm{T}}\right\} = \underbrace{\mathrm{Cov}\left(\boldsymbol{y}\right)}_{\mathbf{C}_{y}}$
-

$$\boxed{\widetilde{\mathbf{G}} = \mathbf{C}_{x,y} \cdot \mathbf{C}_{y}^{-1}}$$

Filtre de Wiener: dérivation

Hypothèses

- Bruit centré: $\mathrm{E}\{\boldsymbol{e}\} = 0$,
- Objet centré: $\mathrm{E}\{\boldsymbol{x}\} = 0$,
- Le bruit et l'objet sont mutuellement indépendant: $\mathrm{Cov}(\boldsymbol{x}, \boldsymbol{e}) = 0$

- $\mathrm{E}\{\boldsymbol{y}\} = \mathrm{E}\{\mathbf{H} \cdot \boldsymbol{y} + \boldsymbol{e}\} = \mathbf{H} \cdot \mathrm{E}\{\boldsymbol{y}\} + \mathrm{E}\{\boldsymbol{e}\} = 0$
- $\mathrm{E}\left\{\boldsymbol{x} \cdot \boldsymbol{y}^{\mathrm{T}}\right\} = \underbrace{\mathrm{Cov}\,(\boldsymbol{x}, \boldsymbol{y})}_{\mathbf{C}_{\boldsymbol{x},\boldsymbol{y}}}$
- $\mathrm{E}\left\{\boldsymbol{y} \cdot \boldsymbol{y}^{\mathrm{T}}\right\} = \underbrace{\mathrm{Cov}\,(\boldsymbol{y})}_{\mathbf{C}_{\boldsymbol{y}}}$
-

$$\boxed{\widetilde{\mathbf{G}} = \mathbf{C}_{\boldsymbol{x},\boldsymbol{y}} \cdot \mathbf{C}_{\boldsymbol{y}}^{-1}}$$

Filtre de Wiener: dérivation

$$
\begin{aligned}
\mathrm{E}\left\{\boldsymbol{y}\cdot\boldsymbol{y}^{\mathrm{T}}\right\} =& \mathrm{E}\left\{(\mathbf{H}\cdot\boldsymbol{x}+\boldsymbol{e})\cdot(\mathbf{H}\cdot\boldsymbol{x}+\boldsymbol{e})^{\mathrm{T}}\right\}\\
=& \mathrm{E}\left\{\mathbf{H}\cdot\boldsymbol{x}\cdot\boldsymbol{x}^{\mathrm{T}}\cdot\mathbf{H}^{\mathrm{T}}\right\}+\mathrm{E}\left\{\boldsymbol{e}\cdot\boldsymbol{e}^{\mathrm{T}}\right\}\\
&+\mathrm{E}\left\{\mathbf{H}\cdot\boldsymbol{x}\cdot\boldsymbol{e}^{\mathrm{T}}\right\}+\mathrm{E}\left\{\boldsymbol{e}\cdot\boldsymbol{x}^{\mathrm{T}}\cdot\mathbf{H}^{\mathrm{T}}\right\}\\
=& \mathbf{H}\cdot\mathrm{E}\left\{\boldsymbol{x}\cdot\boldsymbol{x}^{\mathrm{T}}\right\}\cdot\mathbf{H}^{\mathrm{T}}+\mathrm{E}\left\{\boldsymbol{e}\cdot\boldsymbol{e}^{\mathrm{T}}\right\}\\
=& \mathbf{H}\cdot\mathbf{C}_{\boldsymbol{x}}\cdot\mathbf{H}^{\mathrm{T}}+\mathbf{C}_{\boldsymbol{e}}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}\left\{\boldsymbol{x}\cdot\boldsymbol{y}^{\mathrm{T}}\right\} &= \mathrm{E}\left\{\boldsymbol{x}\cdot(\mathbf{H}\boldsymbol{x}+\boldsymbol{e})^{\mathrm{T}}\right\}\\
&= \mathrm{E}\left\{\boldsymbol{x}\cdot\boldsymbol{x}^{\mathrm{T}}\cdot\mathbf{H}^{\mathrm{T}}\right\}+\mathrm{E}\left\{\boldsymbol{x}\cdot\boldsymbol{e}^{\mathrm{T}}\right\}\\
&= \mathrm{E}\left\{\boldsymbol{x}\cdot\boldsymbol{x}^{\mathrm{T}}\right\}\cdot\mathbf{H}^{\mathrm{T}}+\mathrm{E}\left\{\boldsymbol{x}\right\}\cdot\mathrm{E}\left\{\boldsymbol{e}\right\}^{\mathrm{T}}\\
&= \mathbf{C}_{\boldsymbol{x}}\cdot\mathbf{H}^{\mathrm{T}}
\end{aligned}
$$

Filtre de Wiener (LMMSE)

$$
\begin{aligned}
\widetilde{\mathbf{G}} &= \mathbf{C}_{\boldsymbol{x},\boldsymbol{y}}\cdot\mathbf{C}_{\boldsymbol{y}}^{-1}\\
&= \mathbf{C}_{\boldsymbol{x}}\cdot\mathbf{H}^{\mathrm{T}}\cdot\left(\mathbf{H}\cdot\mathbf{C}_{\boldsymbol{x}}\cdot\mathbf{H}^{\mathrm{T}}+\mathbf{C}_{\boldsymbol{e}}\right)^{-1}\\
&= \left(\mathbf{H}^{\mathrm{T}}\cdot\mathbf{C}_{\boldsymbol{e}}^{-1}\cdot\mathbf{H}+\mathbf{C}_{\boldsymbol{x}}^{-1}\right)^{-1}\cdot\mathbf{H}^{\mathrm{T}}\cdot\mathbf{C}_{\boldsymbol{e}}^{-1}
\end{aligned}
$$

Pas besoin de connaître $\boldsymbol{x}$ et $\boldsymbol{e}$ uniquement leurs deux premiers moments statistiques: $\mathrm{E}\left\{\boldsymbol{x}\right\}$, $\mathbf{C}_{\boldsymbol{x}}$, $\mathrm{E}\left\{\boldsymbol{e}\right\}$ and $\mathbf{C}_{\boldsymbol{e}}$.

Filtre de Wiener: dérivation

$$\begin{aligned}
\mathrm{E}\left\{\boldsymbol{y}\cdot\boldsymbol{y}^{\mathrm{T}}\right\} =& \mathrm{E}\left\{(\mathbf{H}\cdot\boldsymbol{x}+\boldsymbol{e})\cdot(\mathbf{H}\cdot\boldsymbol{x}+\boldsymbol{e})^{\mathrm{T}}\right\}\\
=& \mathrm{E}\left\{\mathbf{H}\cdot\boldsymbol{x}\cdot\boldsymbol{x}^{\mathrm{T}}\cdot\mathbf{H}^{\mathrm{T}}\right\}+\mathrm{E}\left\{\boldsymbol{e}\cdot\boldsymbol{e}^{\mathrm{T}}\right\}\\
&+\mathrm{E}\left\{\mathbf{H}\cdot\boldsymbol{x}\cdot\boldsymbol{e}^{\mathrm{T}}\right\}+\mathrm{E}\left\{\boldsymbol{e}\cdot\boldsymbol{x}^{\mathrm{T}}\cdot\mathbf{H}^{\mathrm{T}}\right\}\\
=& \mathbf{H}\cdot\mathrm{E}\left\{\boldsymbol{x}\cdot\boldsymbol{x}^{\mathrm{T}}\right\}\cdot\mathbf{H}^{\mathrm{T}}+\mathrm{E}\left\{\boldsymbol{e}\cdot\boldsymbol{e}^{\mathrm{T}}\right\}\\
=& \mathbf{H}\cdot\mathbf{C}_{\boldsymbol{x}}\cdot\mathbf{H}^{\mathrm{T}}+\mathbf{C}_{\boldsymbol{e}}
\end{aligned}$$

$$\begin{aligned}
\mathrm{E}\left\{\boldsymbol{x}\cdot\boldsymbol{y}^{\mathrm{T}}\right\} &= \mathrm{E}\left\{\boldsymbol{x}\cdot(\mathbf{H}\boldsymbol{x}+\boldsymbol{e})^{\mathrm{T}}\right\}\\
&= \mathrm{E}\left\{\boldsymbol{x}\cdot\boldsymbol{x}^{\mathrm{T}}\cdot\mathbf{H}^{\mathrm{T}}\right\}+\mathrm{E}\left\{\boldsymbol{x}\cdot\boldsymbol{e}^{\mathrm{T}}\right\}\\
&= \mathrm{E}\left\{\boldsymbol{x}\cdot\boldsymbol{x}^{\mathrm{T}}\right\}\cdot\mathbf{H}^{\mathrm{T}}+\mathrm{E}\left\{\boldsymbol{x}\right\}\cdot\mathrm{E}\left\{\boldsymbol{e}\right\}^{\mathrm{T}}\\
&= \mathbf{C}_{\boldsymbol{x}}\cdot\mathbf{H}^{\mathrm{T}}
\end{aligned}$$

Filtre de Wiener (LMMSE)

$$\begin{aligned}
\widetilde{\mathbf{G}} &= \mathbf{C}_{x,y}\cdot\mathbf{C}_{y}^{-1}\\
&= \mathbf{C}_{x}\cdot\mathbf{H}^{\mathrm{T}}\cdot\left(\mathbf{H}\cdot\mathbf{C}_{x}\cdot\mathbf{H}^{\mathrm{T}}+\mathbf{C}_{e}\right)^{-1}\\
&= \left(\mathbf{H}^{\mathrm{T}}\cdot\mathbf{C}_{e}^{-1}\cdot\mathbf{H}+\mathbf{C}_{x}^{-1}\right)^{-1}\cdot\mathbf{H}^{\mathrm{T}}\cdot\mathbf{C}_{e}^{-1}
\end{aligned}$$

Pas besoin de connaître $\boldsymbol{x}$ et $\boldsymbol{e}$ uniquement leurs deux premiers moments statistiques: $\mathrm{E}\left\{\boldsymbol{x}\right\}$, $\mathbf{C}_{x}$, $\mathrm{E}\left\{\boldsymbol{e}\right\}$ and $\mathbf{C}_{e}$.

$$\begin{aligned}
\mathrm{E}\left\{\boldsymbol{y}\cdot\boldsymbol{y}^{\mathrm{T}}\right\} =& \mathrm{E}\left\{(\mathbf{H}\cdot\boldsymbol{x}+\boldsymbol{e})\cdot(\mathbf{H}\cdot\boldsymbol{x}+\boldsymbol{e})^{\mathrm{T}}\right\}\\
=& \mathrm{E}\left\{\mathbf{H}\cdot\boldsymbol{x}\cdot\boldsymbol{x}^{\mathrm{T}}\cdot\mathbf{H}^{\mathrm{T}}\right\}+\mathrm{E}\left\{\boldsymbol{e}\cdot\boldsymbol{e}^{\mathrm{T}}\right\}\\
&+\mathrm{E}\left\{\mathbf{H}\cdot\boldsymbol{x}\cdot\boldsymbol{e}^{\mathrm{T}}\right\}+\mathrm{E}\left\{\boldsymbol{e}\cdot\boldsymbol{x}^{\mathrm{T}}\cdot\mathbf{H}^{\mathrm{T}}\right\}\\
=& \mathbf{H}\cdot\mathrm{E}\left\{\boldsymbol{x}\cdot\boldsymbol{x}^{\mathrm{T}}\right\}\cdot\mathbf{H}^{\mathrm{T}}+\mathrm{E}\left\{\boldsymbol{e}\cdot\boldsymbol{e}^{\mathrm{T}}\right\}\\
=& \mathbf{H}\cdot\mathbf{C}_{\boldsymbol{x}}\cdot\mathbf{H}^{\mathrm{T}}+\mathbf{C}_{\boldsymbol{e}}
\end{aligned}$$

$$\begin{aligned}
\mathrm{E}\left\{\boldsymbol{x}\cdot\boldsymbol{y}^{\mathrm{T}}\right\} &= \mathrm{E}\left\{\boldsymbol{x}\cdot(\mathbf{H}\boldsymbol{x}+\boldsymbol{e})^{\mathrm{T}}\right\}\\
&= \mathrm{E}\left\{\boldsymbol{x}\cdot\boldsymbol{x}^{\mathrm{T}}\cdot\mathbf{H}^{\mathrm{T}}\right\}+\mathrm{E}\left\{\boldsymbol{x}\cdot\boldsymbol{e}^{\mathrm{T}}\right\}\\
&= \mathrm{E}\left\{\boldsymbol{x}\cdot\boldsymbol{x}^{\mathrm{T}}\right\}\cdot\mathbf{H}^{\mathrm{T}}+\mathrm{E}\left\{\boldsymbol{x}\right\}\cdot\mathrm{E}\left\{\boldsymbol{e}\right\}^{\mathrm{T}}\\
&= \mathbf{C}_{\boldsymbol{x}}\cdot\mathbf{H}^{\mathrm{T}}
\end{aligned}$$

Filtre de Wiener (LMMSE)

$$\begin{aligned}
\widetilde{\mathbf{G}} &= \mathbf{C}_{\boldsymbol{x},\boldsymbol{y}}\cdot\mathbf{C}_{\boldsymbol{y}}^{-1}\\
&= \mathbf{C}_{\boldsymbol{x}}\cdot\mathbf{H}^{\mathrm{T}}\cdot\left(\mathbf{H}\cdot\mathbf{C}_{\boldsymbol{x}}\cdot\mathbf{H}^{\mathrm{T}}+\mathbf{C}_{\boldsymbol{e}}\right)^{-1}\\
&= \left(\mathbf{H}^{\mathrm{T}}\cdot\mathbf{C}_{\boldsymbol{e}}^{-1}\cdot\mathbf{H}+\mathbf{C}_{\boldsymbol{x}}^{-1}\right)^{-1}\cdot\mathbf{H}^{\mathrm{T}}\cdot\mathbf{C}_{\boldsymbol{e}}^{-1}
\end{aligned}$$

Pas besoin de connaître $\boldsymbol{x}$ et $\boldsymbol{e}$ uniquement leurs deux premiers moments statistiques: $\mathrm{E}\left\{\boldsymbol{x}\right\}$, $\mathbf{C}_{\boldsymbol{x}}$, $\mathrm{E}\left\{\boldsymbol{e}\right\}$ and $\mathbf{C}_{\boldsymbol{e}}$.

Opérateur de résolution

- Ce qui nous intéresse c'est de trouver les paramètres $\boldsymbol{x}$ correspondant aux mesures donc il faut comparer $\boldsymbol{x}_{\text{MMSE}}$ et $\boldsymbol{x}$, en moyennant au sens des erreurs $\boldsymbol{e}$ seulement :

$$\langle \boldsymbol{x}_{\text{MMSE}} \rangle_{\boldsymbol{e}} = \widetilde{\mathbf{G}} \cdot \mathbf{H} \cdot \boldsymbol{x}$$

soit

$$\begin{aligned} \langle \boldsymbol{x}_{\text{MMSE}} \rangle_{\boldsymbol{e}} &= \widetilde{\mathbf{G}} \cdot \mathbf{H} \cdot \boldsymbol{x} \\ &= \mathbf{R}_{\text{MMSE}} \cdot \boldsymbol{x} \end{aligned}$$

où

$$\boxed{\mathbf{R}_{\text{MMSE}} \equiv \widetilde{\mathbf{G}} \cdot \mathbf{H}}$$

est l'**opérateur de résolution** (Backus & Gilbert, 1968) ;

Opérateur de résolution

nous avons vu que :

$$\langle \boldsymbol{x}_{\text{MMSE}} \rangle_{\boldsymbol{e}} = \mathbf{R}_{\text{MMSE}} \cdot \boldsymbol{x}$$

avec

$$\boxed{\begin{aligned} \mathbf{R}_{\text{MMSE}} &= \widetilde{\mathbf{G}} \cdot \mathbf{H} \\ &= \mathbf{C}_{\boldsymbol{x}} \cdot \mathbf{H}^{\mathrm{T}} \cdot \left(\mathbf{H} \cdot \mathbf{C}_{\boldsymbol{x}} \cdot \mathbf{H}^{\mathrm{T}} + \mathbf{C}_{\boldsymbol{e}}\right)^{-1} \cdot \mathbf{H} \\ &= \left(\mathbf{C}_{\boldsymbol{x}}^{-1} + \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \mathbf{H}\right)^{-1} \cdot \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \mathbf{H} \\ &= \mathbf{I} - \left(\mathbf{C}_{\boldsymbol{x}}^{-1} + \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \mathbf{H}\right)^{-1} \cdot \mathbf{C}_{\boldsymbol{x}}^{-1} \end{aligned}}$$

on remarque que :

1. $\mathbf{R}_{\text{MMSE}} \neq \mathbf{I}$ donc $\boldsymbol{x}_{\text{MMSE}}$ est ***biaisé*** au sens où $\langle \boldsymbol{x}_{\text{MMSE}} \rangle_{\boldsymbol{e}} \neq \boldsymbol{x}$
2. le filtrage par $\mathbf{R}_{\text{MMSE}}$ est responsable d'une ***atténuation*** (*i.e.* $\mathbf{R}_{\text{MMSE}} \leq \mathbf{I}$) ;

Hypothèses:

- $\mathbf{H}$ est une convolution circulante: $\mathbf{H} = \mathbf{F}^{-1} \cdot \operatorname{diag}\left(\widehat{\boldsymbol{h}}\right) \cdot \mathbf{F}$,
- les variables aléatoires $\boldsymbol{x}$ et $\boldsymbol{e}$ sont centrées: $\mathrm{E}\{x\} = 0$ et $\mathrm{E}\{e\} = 0$,
- le bruit $\boldsymbol{e}$ est stationnaire et de densité spectrale de puissance $\mathrm{E}\left\{|\widehat{\boldsymbol{e}}|^2\right\}$:
  $\mathbf{C}_e = \mathbf{F}^{-1} \cdot \operatorname{diag}\left(\mathrm{E}\left\{|\widehat{\boldsymbol{e}}|^2\right\}\right) \cdot \mathbf{F}$,
- on connaît *a priori* la densité spectrale de puissance de l'objet $\mathrm{E}\left\{|\widehat{\boldsymbol{x}}|^2\right\}$:
  $\mathbf{C}_x = \mathbf{F}^{-1} \cdot \operatorname{diag}\left(\mathrm{E}\left\{|\widehat{\boldsymbol{x}}|^2\right\}\right) \cdot \mathbf{F}$.

Filtre de Wiener: Fourier

Le filtre de Wiener est $\widetilde{\mathbf{G}}$ est une convolution et prend une forme très simple dans l'espace de Fourier:

$$\widehat{x}_k^{(\text{Wiener})} = \widehat{g}_k\,\widehat{y}_k = \frac{\widehat{h}_k^\star}{|\widehat{h}_k|^2 + \dfrac{\mathrm{E}\left\{|\widehat{e}_k|^2\right\}}{\mathrm{E}\left\{|\widehat{x}_k|^2\right\}}}\,\widehat{y}_k$$

Filtre de Wiener: déconvolution

Hypothèses:

- $\mathbf{H}$ est une convolution circulante: $\mathbf{H} = \mathbf{F}^{-1} \cdot \mathrm{diag}\left(\widehat{\boldsymbol{h}}\right) \cdot \mathbf{F}$,
- les variables aléatoires $\boldsymbol{x}$ et $\boldsymbol{e}$ sont centrées: $\mathrm{E}\{x\} = 0$ et $\mathrm{E}\{e\} = 0$,
- le bruit $\boldsymbol{e}$ est stationnaire et de densité spectrale de puissance $\mathrm{E}\left\{|\widehat{\boldsymbol{e}}|^2\right\}$:
  $\mathbf{C}_e = \mathbf{F}^{-1} \cdot \mathrm{diag}\left(\mathrm{E}\left\{|\widehat{\boldsymbol{e}}|^2\right\}\right) \cdot \mathbf{F}$,
- on connaît *a priori* la densité spectrale de puissance de l'objet $\mathrm{E}\left\{|\widehat{\boldsymbol{x}}|^2\right\}$:
  $\mathbf{C}_x = \mathbf{F}^{-1} \cdot \mathrm{diag}\left(\mathrm{E}\left\{|\widehat{\boldsymbol{x}}|^2\right\}\right) \cdot \mathbf{F}$.

Filtre de Wiener: Fourier

Le filtre de Wiener est $\widetilde{\mathbf{G}}$ est une convolution et prend une forme très simple dans l'espace de Fourier:

$$\widehat{x}_k^{\text{(Wiener)}} = \widehat{g}_k\,\widehat{y}_k = \frac{\widehat{h}_k^\star}{|\widehat{h}_k|^2 + \dfrac{\mathrm{E}\left\{|\widehat{e}_k|^2\right\}}{\mathrm{E}\left\{|\widehat{x}_k|^2\right\}}}\,\widehat{y}_k$$

Filtre de Wiener: déconvolution

Hypothèses:

- $\mathbf{H}$ est une convolution circulante: $\mathbf{H} = \mathbf{F}^{-1} \cdot \operatorname{diag}\left(\widehat{\boldsymbol{h}}\right) \cdot \mathbf{F}$,
- les variables aléatoires $\boldsymbol{x}$ et $\boldsymbol{e}$ sont centrées: $\mathrm{E}\{x\} = 0$ et $\mathrm{E}\{e\} = 0$,
- le bruit $\boldsymbol{e}$ est stationnaire et de densité spectrale de puissance $\mathrm{E}\left\{|\widehat{\boldsymbol{e}}|^2\right\}$:
  $\mathbf{C}_{\boldsymbol{e}} = \mathbf{F}^{-1} \cdot \operatorname{diag}\left(\mathrm{E}\left\{|\widehat{\boldsymbol{e}}|^2\right\}\right) \cdot \mathbf{F}$,
- on connaît *a priori* la densité spectrale de puissance de l'objet $\mathrm{E}\left\{|\widehat{\boldsymbol{x}}|^2\right\}$:
  $\mathbf{C}_{\boldsymbol{x}} = \mathbf{F}^{-1} \cdot \operatorname{diag}\left(\mathrm{E}\left\{|\widehat{\boldsymbol{x}}|^2\right\}\right) \cdot \mathbf{F}$.

Filtre de Wiener: Fourier

Le filtre de Wiener est $\widetilde{\mathbf{G}}$ est une convolution et prend une forme très simple dans l'espace de Fourier:

$$\widehat{x}_k^{\text{(Wiener)}} = \widehat{g}_k\,\widehat{y}_k = \frac{\widehat{h}_k^\star}{|\widehat{h}_k|^2 + \dfrac{\mathrm{E}\left\{|\widehat{e}_k|^2\right\}}{\mathrm{E}\left\{|\widehat{x}_k|^2\right\}}}\,\widehat{y}_k$$

Filtre de Wiener: déconvolution

Hypothèses:

- $\mathbf{H}$ est une convolution circulante: $\mathbf{H} = \mathbf{F}^{-1} \cdot \operatorname{diag}\left(\widehat{\boldsymbol{h}}\right) \cdot \mathbf{F}$,
- les variables aléatoires $\boldsymbol{x}$ et $\boldsymbol{e}$ sont centrées: $\mathrm{E}\{x\} = 0$ et $\mathrm{E}\{e\} = 0$,
- le bruit $\boldsymbol{e}$ est stationnaire et de densité spectrale de puissance $\mathrm{E}\left\{|\widehat{\boldsymbol{e}}|^2\right\}$:
  $\mathbf{C}_{\boldsymbol{e}} = \mathbf{F}^{-1} \cdot \operatorname{diag}\left(\mathrm{E}\left\{|\widehat{\boldsymbol{e}}|^2\right\}\right) \cdot \mathbf{F}$,
- on connaît *a priori* la densité spectrale de puissance de l'objet $\mathrm{E}\left\{|\widehat{\boldsymbol{x}}|^2\right\}$:
  $\mathbf{C}_{\boldsymbol{x}} = \mathbf{F}^{-1} \cdot \operatorname{diag}\left(\mathrm{E}\left\{|\widehat{\boldsymbol{x}}|^2\right\}\right) \cdot \mathbf{F}$.

Filtre de Wiener: Fourier

Le filtre de Wiener est $\widetilde{\mathrm{G}}$ est une convolution et prend une forme très simple dans l'espace de Fourier:

$$\widehat{x}_k^{\text{(Wiener)}} = \widehat{g}_k\, \widehat{y}_k = \frac{\widehat{h}_k^\star}{|\widehat{h}_k|^2 + \dfrac{\mathrm{E}\left\{|\widehat{e}_k|^2\right\}}{\mathrm{E}\left\{|\widehat{x}_k|^2\right\}}}\, \widehat{y}_k$$

Filtre de Wiener: déconvolution

Hypothèses:

- $\mathbf{H}$ est une convolution circulante: $\mathbf{H} = \mathbf{F}^{-1} \cdot \operatorname{diag}\left(\widehat{\boldsymbol{h}}\right) \cdot \mathbf{F}$,
- les variables aléatoires $\boldsymbol{x}$ et $\boldsymbol{e}$ sont centrées: $\mathrm{E}\{x\} = 0$ et $\mathrm{E}\{e\} = 0$,
- le bruit $\boldsymbol{e}$ est stationnaire et de densité spectrale de puissance $\mathrm{E}\left\{|\widehat{\boldsymbol{e}}|^2\right\}$:
  $\mathbf{C}_{\boldsymbol{e}} = \mathbf{F}^{-1} \cdot \operatorname{diag}\left(\mathrm{E}\left\{|\widehat{\boldsymbol{e}}|^2\right\}\right) \cdot \mathbf{F}$,
- on connaît *a priori* la densité spectrale de puissance de l'objet $\mathrm{E}\left\{|\widehat{\boldsymbol{x}}|^2\right\}$:
  $\mathbf{C}_{\boldsymbol{x}} = \mathbf{F}^{-1} \cdot \operatorname{diag}\left(\mathrm{E}\left\{|\widehat{\boldsymbol{x}}|^2\right\}\right) \cdot \mathbf{F}$.

Filtre de Wiener: Fourier

Le filtre de Wiener est $\widetilde{\mathbf{G}}$ est une convolution et prend une forme très simple dans l'espace de Fourier:

$$\widehat{x}_k^{\text{(Wiener)}} = \widehat{g}_k \, \widehat{y}_k = \frac{\widehat{h}_k^\star}{|\widehat{h}_k|^2 + \dfrac{\mathrm{E}\left\{|\widehat{e}_k|^2\right\}}{\mathrm{E}\left\{|\widehat{x}_k|^2\right\}}} \, \widehat{y}_k$$

Filtre de Wiener: déconvolution

Le filtre de Wiener est:

$$\widehat{g}_k = \frac{\widehat{h}_k^\star}{|\widehat{h}_k|^2 + \widehat{q}_k} \quad \text{avec} \quad \widehat{q}_k = \frac{\mathrm{E}\left\{|\widehat{e}_k|^2\right\}}{\mathrm{E}\left\{|\widehat{x}_k|^2\right\}} > 0$$

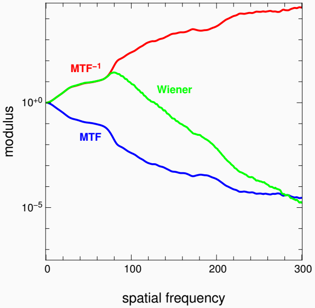
Comportement:

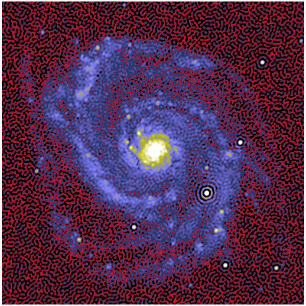
$$\widehat{g}_u \simeq \begin{cases} 1/\widehat{h}_u & \text{si } \mathrm{SNR}_u \gg 1 \\ 0 & \text{si } \mathrm{SNR}_u \ll 1 \end{cases}$$

avec $\mathrm{SNR}_u \equiv \sqrt{\frac{|\widehat{h}_u|^2\,\mathrm{E}\{|\widehat{x}_u|^2\}}{\mathrm{E}\{|\widehat{e}_u|^2\}}}$.

Mais:

- Les densités spectrales sont rarement connue et doivent être devinée d'après les données,
- Solution possiblement non physique (valeur négative, rebonds,...).

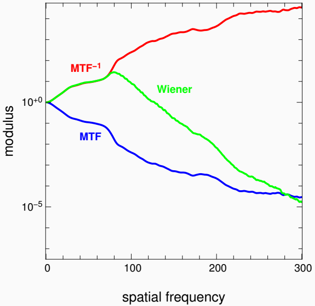




Filtre de Wiener: déconvolution

Le filtre de Wiener est:

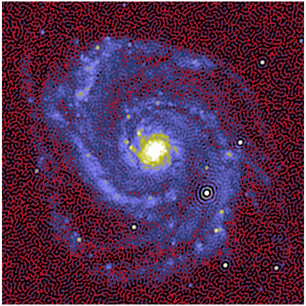
$$\widehat{g}_k = \frac{\widehat{h}_k^\star}{|\widehat{h}_k|^2 + \widehat{q}_k} \quad \text{avec} \quad \widehat{q}_k = \frac{\mathrm{E}\left\{|\widehat{e}_k|^2\right\}}{\mathrm{E}\left\{|\widehat{x}_k|^2\right\}} > 0$$



Comportement:

$$\widehat{g}_u \simeq \begin{cases} 1/\widehat{h}_u & \text{si } \mathrm{SNR}_u \gg 1 \\ 0 & \text{si } \mathrm{SNR}_u \ll 1 \end{cases}$$

avec $\mathrm{SNR}_u \equiv \sqrt{\frac{|\widehat{h}_u|^2\,\mathrm{E}\{|\widehat{\bar{x}}_u|^2\}}{\mathrm{E}\{|\widehat{e}_u|^2\}}}$.



Mais:

- Les densités spectrales sont rarement connue et doivent être devinée d'après les données,
- Solution possiblement non physique (valeur négative, rebonds,. . .).

Filtre de Wiener: déconvolution

Le filtre de Wiener est:

$$\widehat{g}_k = \frac{\widehat{h}_k^\star}{|\widehat{h}_k|^2 + \widehat{q}_k} \quad \text{avec} \quad \widehat{q}_k = \frac{\mathrm{E}\left\{|\widehat{e}_k|^2\right\}}{\mathrm{E}\left\{|\widehat{x}_k|^2\right\}} > 0$$

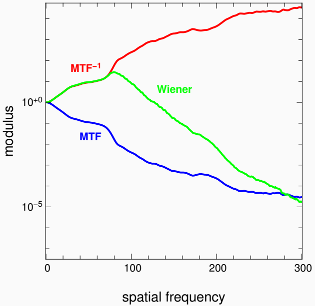
Comportement:

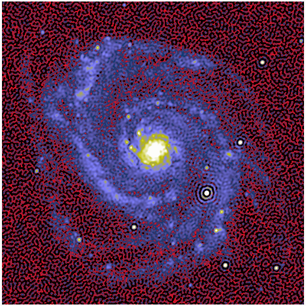
$$\widehat{g}_u \simeq \begin{cases} 1/\widehat{h}_u & \text{si } \mathrm{SNR}_u \gg 1 \\ 0 & \text{si } \mathrm{SNR}_u \ll 1 \end{cases}$$

avec $\mathrm{SNR}_u \equiv \sqrt{\frac{|\widehat{h}_u|^2\,\mathrm{E}\{|\widehat{\bar{x}}_u|^2\}}{\mathrm{E}\{|\widehat{e}_u|^2\}}}$.

Mais:

- Les densités spectrales sont rarement connue et doivent être devinée d'après les données,
- Solution possiblement non physique (valeur négative, rebonds,...).





En pratique, contruire le filtre de Wiener

On suppose:

- Le bruit indépendant et stationnaire: $\mathrm{E}\left\{|\widehat{e_k}|^2\right\} = \mathrm{cte}$
- La distribution d'intensité de l'objet suit une loi simple (loi de puissance): $\mathrm{E}\left\{|\widehat{x_k}|^2\right\} \propto \|\boldsymbol{\omega}_k\|^{-\beta}$; avec $\beta \geq 0$:

$$\widehat{g_k} = \frac{\widehat{h_k^\star}}{|\widehat{h_k}|^2 + \alpha\ \|\boldsymbol{u}_k\|^\beta}\,.$$

Il reste plus qu'a déterminer les valeurs de α et β visuellement ou automatiquement (via GCV, GML, SURE,. . .).

Maximum de vraisemblance

Maximum de vraisemblance : Objectif

- Quel est le meilleur modèle ?

⇒ **réponse :** c'est celui qui maximise la probabilité d'avoir observé les données :

$$\boxed{\boldsymbol{x}_{\mathrm{ML}} = \arg\max_{\boldsymbol{x}} \Pr(\boldsymbol{y}|\boldsymbol{x})}$$

où **ML = *Maximum Likelihood***

▲ C'est la meilleure solution au sens de la statistiques des erreurs...

Maximum de vraisemblance : Objectif

- Quel est le meilleur modèle ?

⇒ **réponse :** c'est celui qui maximise la probabilité d'avoir observé les données :

$$\boxed{\boldsymbol{x}_{\mathrm{ML}} = \arg\max_{\boldsymbol{x}} \Pr(\boldsymbol{y}|\boldsymbol{x})}$$

où **ML = *Maximum Likelihood***

▲ C'est la meilleure solution au sens de la statistiques des erreurs...

Maximum de vraisemblance : Objectif

- Quel est le meilleur modèle ?

⇒ **réponse :** c'est celui qui maximise la probabilité d'avoir observé les données :

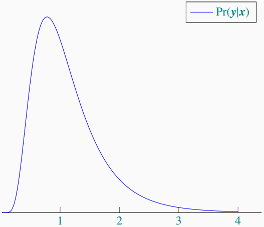
$$\boxed{\boldsymbol{x}_{\mathrm{ML}} = \arg\max_{\boldsymbol{x}} \Pr(\boldsymbol{y}|\boldsymbol{x})}$$

où **ML = *Maximum Likelihood***

▲ C'est la meilleure solution au sens de la statistiques des erreurs...

Maximum de vraisemblance : Fonction de pénalisation

- Maximum de vraisemblance: on cherche la solution qui maximise $\Pr(\boldsymbol{y}|\boldsymbol{x})$:



- solution au sens du maximum de vraisemblance :

$$\begin{aligned} \boldsymbol{x}_{\mathrm{ML}} &= \arg\max_{\boldsymbol{x}} \Pr(\boldsymbol{y}|\boldsymbol{x}) \\ &= \arg\min_{\boldsymbol{x}} f_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{x}) \end{aligned}$$

- fonction de pénalisation :

$$\boxed{f_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{x}) = -\log \Pr(\boldsymbol{y}|\boldsymbol{x}) + \text{const}}$$

Approximation gaussienne pour les erreurs

- modèle direct :

$$\boldsymbol{y} = m(\boldsymbol{x}) + \boldsymbol{e}$$

  où $\boldsymbol{e}$ incorpore les erreurs de modélisation et le bruit ;
- la statistique des erreurs est ***normale centrée*** :

$$\boxed{\boldsymbol{e} \,|\, \boldsymbol{x} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_{\boldsymbol{e}|\boldsymbol{x}})}$$

  c'est-à-dire que la fonction de distribution des erreurs sachant $\boldsymbol{x}$ est :

$$\mathrm{PDF}(\boldsymbol{e} \,|\, \boldsymbol{x}) = \frac{\exp\left\{-\frac{1}{2}\,\boldsymbol{e}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}|\boldsymbol{x}}^{-1} \cdot \boldsymbol{e}\right\}}{(2\,\pi)^{M/2}\,\left|\mathbf{C}_{\boldsymbol{e}|\boldsymbol{x}}\right|^{\frac{1}{2}}}$$

  où M est le nombre de mesures.

Maximum de vraisemblance : Approximation gaussienne, modèle non linéaire

- modèle :

$$\boldsymbol{y} = m(\boldsymbol{x}) + \boldsymbol{e}$$

- fonction de distribution des erreurs :

$$\mathrm{PDF}(\boldsymbol{y}|\boldsymbol{x}) = \frac{\exp\left\{-\frac{1}{2}\left[\boldsymbol{y} - \boldsymbol{m}(\boldsymbol{x})\right]^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \left[\boldsymbol{y} - \boldsymbol{m}(\boldsymbol{x})\right]\right\}}{(2\,\pi)^{M/2}\,|\mathbf{C}_{\boldsymbol{e}}|^{\frac{1}{2}}}$$

- fonction de pénalisation :

$$f_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{y} - m(\boldsymbol{x}))^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot (\boldsymbol{y} - m(\boldsymbol{x}))$$

Maximum de vraisemblance: détection unique

Amas des Hyades

On recherche la position $\boldsymbol{p}$ et l'amplitude α d'un motif $\boldsymbol{h}$ dans l'image

$$\boldsymbol{m} = m(\alpha, \boldsymbol{p}) = \alpha\, \boldsymbol{h}(\boldsymbol{x} - \boldsymbol{p})$$

avec la PSF $\boldsymbol{h}(\boldsymbol{x})$ et $\boldsymbol{g} = \boldsymbol{h}(\boldsymbol{x} - \boldsymbol{p})$

$$\begin{aligned}
f_{\boldsymbol{y}|\boldsymbol{x}}(\alpha, \boldsymbol{p}) &= \frac{1}{2}(\boldsymbol{y} - \boldsymbol{m})^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot (\boldsymbol{y} - \boldsymbol{m}) \\
&= \frac{1}{2}\boldsymbol{m}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{m} - \boldsymbol{m}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y} + \frac{1}{2}\boldsymbol{y}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y} \\
&= \frac{1}{2}\alpha^2\, \boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{g} + \frac{1}{2}\boldsymbol{y}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y} \\
&\quad - \alpha\, \boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y}
\end{aligned}$$

Maximum de vraisemblance: détection unique

Amas des Hyades

On recherche la position $\boldsymbol{p}$ et l'amplitude α d'un motif $\boldsymbol{h}$ dans l'image

$$\boldsymbol{m} = m(\alpha, \boldsymbol{p}) = \alpha\, \boldsymbol{h}(\boldsymbol{x} - \boldsymbol{p})$$

avec la PSF $\boldsymbol{h}(\boldsymbol{x})$ et $\boldsymbol{g} = \boldsymbol{h}(\boldsymbol{x} - \boldsymbol{p})$

$$\begin{aligned} f_{\boldsymbol{y}|\boldsymbol{x}}(\alpha, \boldsymbol{p}) &= \frac{1}{2}(\boldsymbol{y} - \boldsymbol{m})^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot (\boldsymbol{y} - \boldsymbol{m}) \\ &= \frac{1}{2}\boldsymbol{m}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{m} - \boldsymbol{m}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y} + \frac{1}{2}\boldsymbol{y}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y} \\ &= \frac{1}{2}\alpha^2\, \boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{g} + \frac{1}{2}\boldsymbol{y}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y} \\ &\quad - \alpha\, \boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y} \end{aligned}$$

Détection unique: estimation de l'amplitude

$$f_{\boldsymbol{y}|\boldsymbol{x}}(\alpha) = \frac{1}{2}\alpha^2\, \boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g} - \alpha\, \boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y} + \frac{1}{2}\boldsymbol{y}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}\,.$$

Estimation de l'amplitude α

$$\left.\frac{\partial f_{\boldsymbol{y}|\boldsymbol{x}}(\alpha)}{\partial \alpha}\right|_{\alpha=\alpha^+} = 0 \Longleftrightarrow \alpha^+ \boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g} - \boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y} = 0\,,$$

$$\Longleftrightarrow \alpha^+ = \frac{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}}{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}}\,.$$

$$\begin{aligned} \left. f_{\boldsymbol{y}|\boldsymbol{x}}(\alpha, \boldsymbol{p})\right|_{\alpha=\alpha^+} &= \frac{1}{2}\left(\frac{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}}{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}}\right)^2 \boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g} - \frac{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}}{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}}\, \boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y} \\ &= -\frac{1}{2}\frac{\left(\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}\right)^2}{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}} \end{aligned}$$

Détection unique: estimation de l'amplitude

$$f_{\boldsymbol{y}|\boldsymbol{x}}(\alpha) = \frac{1}{2}\alpha^2\, \boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g} - \alpha\, \boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y} + \frac{1}{2}\boldsymbol{y}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}\,.$$

Estimation de l'amplitude α

$$\left.\frac{\partial f_{\boldsymbol{y}|\boldsymbol{x}}(\alpha)}{\partial \alpha}\right|_{\alpha=\alpha^+} = 0 \iff \alpha^+ \boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g} - \boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y} = 0\,,$$

$$\iff \alpha^+ = \frac{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}}{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}}\,.$$

$$\left. f_{\boldsymbol{y}|\boldsymbol{x}}(\alpha, \boldsymbol{p})\right|_{\alpha=\alpha^+} = \frac{1}{2}\left(\frac{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}}{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}}\right)^2 \boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g} - \frac{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}}{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}}\, \boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}$$

$$= -\frac{1}{2}\frac{\left(\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}\right)^2}{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}}$$

Détection unique: estimation de l'amplitude

$$f_{\boldsymbol{y}|\boldsymbol{x}}(\alpha) = \frac{1}{2}\alpha^2\, \boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g} - \alpha\, \boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y} + \frac{1}{2}\boldsymbol{y}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}\,.$$

Estimation de l'amplitude α

$$\left.\frac{\partial f_{\boldsymbol{y}|\boldsymbol{x}}(\alpha)}{\partial \alpha}\right|_{\alpha=\alpha^+} = 0 \Longleftrightarrow \alpha^+ \boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g} - \boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y} = 0\,,$$

$$\Longleftrightarrow \alpha^+ = \frac{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}}{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}}\,.$$

$$\begin{aligned} \left. f_{\boldsymbol{y}|\boldsymbol{x}}(\alpha, \boldsymbol{p})\right|_{\alpha=\alpha^+} &= \frac{1}{2}\left(\frac{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}}{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}}\right)^2 \boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g} - \frac{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}}{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}}\, \boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y} \\ &= -\frac{1}{2}\frac{\left(\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}\right)^2}{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}} \end{aligned}$$

Détection unique: estimation rapide de la position

Approximation: bruit blanc non stationnaire

$\mathbf{C}_{\boldsymbol{e}}^{-1} = \mathrm{diag}(\boldsymbol{w})$ avec $w_k = \dfrac{1}{\sigma_k^2} = \dfrac{1}{\mathrm{Var}(y_k)}$.

Estimation de la position au pixel près

En 1D : $g[k] = h[k-p]$

$$
\begin{aligned}
f_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{\alpha}, \boldsymbol{p})\Big|_{\alpha=\alpha^+} &= -\frac{1}{2}\frac{\left(\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}\right)^2}{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}}, \\
&= -\frac{1}{2}\frac{\left(\sum_k h[k-p]\, w[k]\, y[k]\right)^2}{\sum_k h^2[k-p]\, w[k]} \\
&= -\frac{1}{2}\frac{(\text{Inter-corrélation})^2}{\text{Normalisation}}
\end{aligned}
$$

Détection

Pour p qui maximise $\frac{(\text{Inter-corrélation})^2}{\text{Normalisation}}$ ou $\frac{\text{Inter-corrélation}}{\text{Normalisation}}$ si $\alpha > 0$. Calcul rapide par FFT.

Détection unique: estimation rapide de la position

Approximation: bruit blanc non stationnaire

$\mathbf{C}_{\boldsymbol{e}}^{-1} = \mathrm{diag}(\boldsymbol{w})$ avec $w_k = \dfrac{1}{\sigma_k^2} = \dfrac{1}{\mathrm{Var}(y_k)}$.

Estimation de la position au pixel près

En 1D : $g[k] = h[k-p]$

$$
\begin{aligned}
f_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{\alpha}, \boldsymbol{p})\Big|_{\alpha=\alpha^+} &= -\frac{1}{2}\,\frac{\left(\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}\right)^2}{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}}\,, \\
&= -\frac{1}{2}\,\frac{\left(\sum_k h[k-p]\, w[k]\, y[k]\right)^2}{\sum_k h^2[k-p]\, w[k]} \\
&= -\frac{1}{2}\,\frac{(\text{Inter-corrélation})^2}{\text{Normalisation}}
\end{aligned}
$$

Détection

Pour p qui maximise $\frac{(\text{Inter-corrélation})^2}{\text{Normalisation}}$ ou $\frac{\text{Inter-corrélation}}{\text{Normalisation}}$ si $\alpha > 0$. Calcul rapide par FFT.

Approximation: bruit blanc non stationnaire

$\mathbf{C}_{\boldsymbol{e}}^{-1} = \mathrm{diag}(\boldsymbol{w})$ avec $w_k = \dfrac{1}{\sigma_k^2} = \dfrac{1}{\mathrm{Var}(y_k)}$.

Estimation de la position au pixel près

En 1D : $g[k] = h[k-p]$

$$
\begin{aligned}
f_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{\alpha}, \boldsymbol{p})\Big|_{\alpha=\alpha^+} &= -\frac{1}{2}\,\frac{\left(\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}\right)^2}{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}}\,, \\
&= -\frac{1}{2}\,\frac{\left(\sum_k h[k-p]\,w[k]\,y[k]\right)^2}{\sum_k h^2[k-p]\,w[k]} \\
&= -\frac{1}{2}\,\frac{(\text{Inter-corrélation})^2}{\text{Normalisation}}
\end{aligned}
$$

Détection

Pour p qui maximise $\frac{(\text{Inter-corrélation})^2}{\text{Normalisation}}$ ou $\frac{\text{Inter-corrélation}}{\text{Normalisation}}$ si $\alpha > 0$. Calcul rapide par FFT.

Détection unique: estimation rapide de la position

Approximation: bruit blanc non stationnaire

$\mathbf{C}_{\boldsymbol{e}}^{-1} = \mathrm{diag}(\boldsymbol{w})$ avec $w_k = \dfrac{1}{\sigma_k^2} = \dfrac{1}{\mathrm{Var}(y_k)}$.

Estimation de la position au pixel près

En 1D : $g[k] = h[k-p]$

$$
\begin{aligned}
f_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{\alpha}, \boldsymbol{p})\Big|_{\alpha=\alpha^+} &= -\frac{1}{2}\frac{\left(\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}\right)^2}{\boldsymbol{g}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}}\,, \\
&= -\frac{1}{2}\frac{\left(\sum_k h[k-p]\, w[k]\, y[k]\right)^2}{\sum_k h^2[k-p]\, w[k]} \\
&= -\frac{1}{2}\frac{(\text{Inter-corrélation})^2}{\text{Normalisation}}
\end{aligned}
$$

Détection

Pour p qui maximise $\frac{(\text{Inter-corrélation})^2}{\text{Normalisation}}$ ou $\frac{\text{Inter-corrélation}}{\text{Normalisation}}$ si $\alpha > 0$. Calcul rapide par FFT.

Maximum de vraisemblance: détection multiple

Amas des Hyades

On recherche la position $\boldsymbol{p}$ et l'amplitude α de N étoiles

$$\boldsymbol{m} = m(\alpha, \boldsymbol{p}) = \sum_{n}^{N} \alpha_n \, \boldsymbol{h}(\boldsymbol{x} - \boldsymbol{p}_n)$$

avec la PSF $\boldsymbol{h}(\boldsymbol{x})$ et $\boldsymbol{g}_n = \boldsymbol{h}(\boldsymbol{x} - \boldsymbol{p}_n)$

$$\begin{aligned}
f_{\boldsymbol{y}|\boldsymbol{x}}(\alpha, \boldsymbol{p}) &= \frac{1}{2}(\boldsymbol{y} - \boldsymbol{m})^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot (\boldsymbol{y} - \boldsymbol{m}) \\
&= \frac{1}{2}\boldsymbol{m}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{m} - \boldsymbol{m}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y} + \frac{1}{2}\boldsymbol{y}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y} \\
&= \frac{1}{2}\sum_{n} \alpha_n^2 \, \boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}_n + \frac{1}{2}\boldsymbol{y}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y} \\
&\quad \frac{1}{2}\sum_{n}\sum_{m \neq n} \alpha_n \, \alpha_m \, \boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}_m \\
&\quad - \sum_{n} \alpha_n \, \boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}
\end{aligned}$$

Un problème combinatoire!!!

Amas des Hyades

On recherche la position $\boldsymbol{p}$ et l'amplitude α de N étoiles

$$\boldsymbol{m} = m(\alpha, \boldsymbol{p}) = \sum_{n}^{N} \alpha_n \, \boldsymbol{h}(\boldsymbol{x} - \boldsymbol{p}_n)$$

avec la PSF $\boldsymbol{h}(\boldsymbol{x})$ et $\boldsymbol{g}_n = \boldsymbol{h}(\boldsymbol{x} - \boldsymbol{p}_n)$

$$\begin{aligned}
f_{\boldsymbol{y}|\boldsymbol{x}}(\alpha, \boldsymbol{p}) =& \frac{1}{2}(\boldsymbol{y} - \boldsymbol{m})^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot (\boldsymbol{y} - \boldsymbol{m}) \\
=& \frac{1}{2}\boldsymbol{m}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{m} - \boldsymbol{m}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y} + \frac{1}{2}\boldsymbol{y}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y} \\
=& \frac{1}{2} \sum_{n} \alpha_n^2 \, \boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}_n + \frac{1}{2}\boldsymbol{y}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y} \\
& \frac{1}{2} \sum_{n} \sum_{m \neq n} \alpha_n \, \alpha_m \, \boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}_m \\
& - \sum_{n} \alpha_n \, \boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}
\end{aligned}$$

Un problème combinatoire!!!

Maximum de vraisemblance: détection multiple

Amas des Hyades

On recherche la position $\boldsymbol{p}$ et l'amplitude α de N étoiles

$$\boldsymbol{m} = m(\alpha, \boldsymbol{p}) = \sum_{n}^{N} \alpha_n \, \boldsymbol{h}(\boldsymbol{x} - \boldsymbol{p}_n)$$

avec la PSF $\boldsymbol{h}(\boldsymbol{x})$ et $\boldsymbol{g}_n = \boldsymbol{h}(\boldsymbol{x} - \boldsymbol{p}_n)$

$$\begin{aligned}
f_{\boldsymbol{y}|\boldsymbol{x}}(\alpha, \boldsymbol{p}) &= \frac{1}{2}(\boldsymbol{y} - \boldsymbol{m})^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot (\boldsymbol{y} - \boldsymbol{m}) \\
&= \frac{1}{2}\boldsymbol{m}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{m} - \boldsymbol{m}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y} + \frac{1}{2}\boldsymbol{y}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y} \\
&= \frac{1}{2} \sum_{n} \alpha_n^2 \, \boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}_n + \frac{1}{2}\boldsymbol{y}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y} \\
&\quad \frac{1}{2} \sum_{n} \sum_{m \neq n} \alpha_n \, \alpha_m \, \boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}_m \\
&\quad - \sum_{n} \alpha_n \, \boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}
\end{aligned}$$

Un problème combinatoire!!!

Maximum de vraisemblance: détection

Méthode séparable

Négliger les termes croisés,

$$
\begin{aligned}
\boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}_m &\approx 0 \\
f_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{\alpha}, \boldsymbol{p}) &\approx \frac{1}{2} \sum_n \alpha_n^2 \, \boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}_n - \sum_n \alpha_n \, \boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}
\end{aligned}
$$

Méthode gloutone

$$
\begin{aligned}
f_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{\alpha}, \boldsymbol{p}) &\approx \frac{1}{2} \sum_n \alpha_n^2 \, \boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}_n - \sum_n \alpha_n \, \boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y} + \frac{1}{2} \sum_n \sum_{m<n} \alpha_n \, \alpha_m \, \boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}_m \\
&\approx \frac{1}{2} \sum_n \alpha_n^2 \, \boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}_n - \sum_n \alpha_n \, \boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \underbrace{\left(\boldsymbol{y} - \sum_{m<n} \alpha_m \, \boldsymbol{r}_m \right)}_{\boldsymbol{r}_r} .
\end{aligned}
$$

détection dans les résidus $\boldsymbol{r}_n$ (CLEAN):

$$
r_n = \boldsymbol{y} - \sum_{m<n} \alpha_m \, \boldsymbol{r}_r
$$

Maximum de vraisemblance: détection

Méthode séparable

Négliger les termes croisés,

$$\boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}_m \approx 0$$

$$f_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{\alpha}, \boldsymbol{p}) \approx \frac{1}{2} \sum_n \alpha_n^2 \, \boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}_n - \sum_n \alpha_n \, \boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y}$$

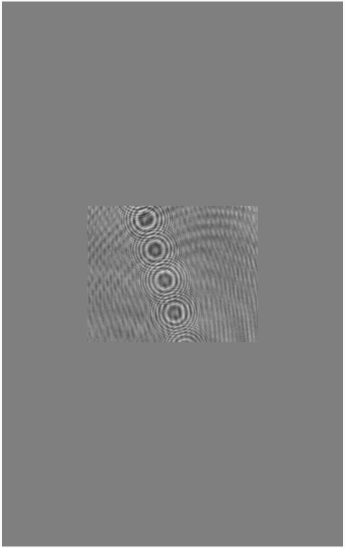
Méthode gloutone

$$\begin{aligned} f_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{\alpha}, \boldsymbol{p}) &\approx \frac{1}{2} \sum_n \alpha_n^2 \, \boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}_n - \sum_n \alpha_n \, \boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{y} + \frac{1}{2} \sum_n \sum_{m<n} \alpha_n \, \alpha_m \, \boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}_m \\ &\approx \frac{1}{2} \sum_n \alpha_n^2 \, \boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \boldsymbol{g}_n - \sum_n \alpha_n \, \boldsymbol{g}_n^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \underbrace{\left(\boldsymbol{y} - \sum_{m<n} \alpha_m \, \boldsymbol{r}_m \right)}_{\boldsymbol{r}_r} . \end{aligned}$$

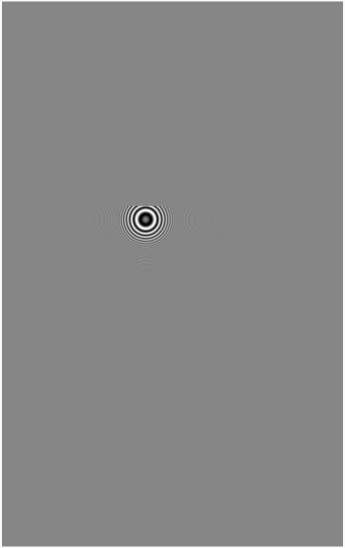
détection dans les résidus $\boldsymbol{r}_n$ (CLEAN):

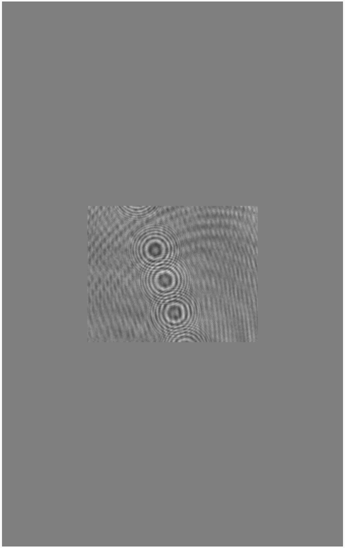
$$r_n = \boldsymbol{y} - \sum_{m<n} \alpha_m \, \boldsymbol{r}_r$$

Détection hors-champ (0)

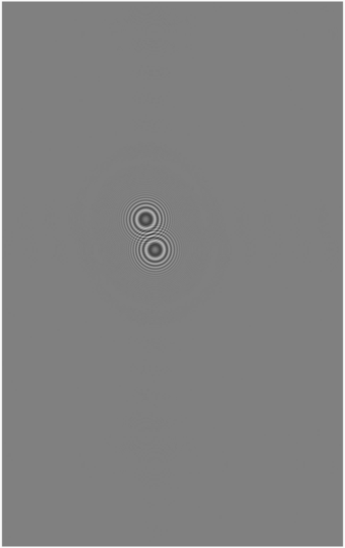


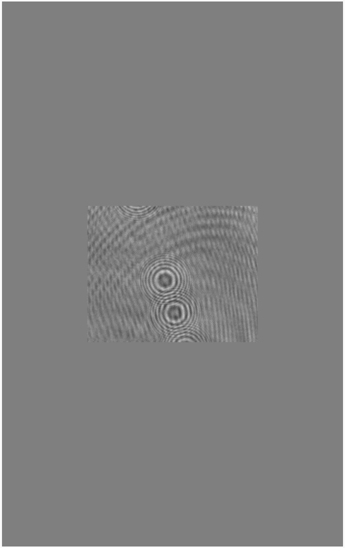
Détection hors-champ (1)



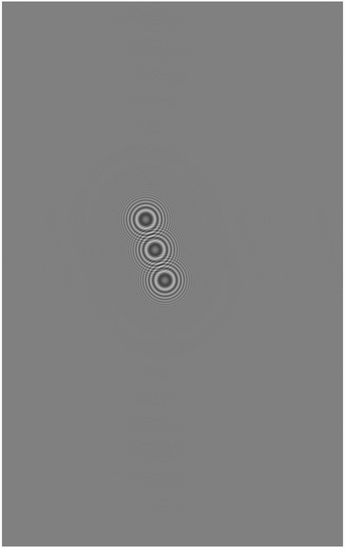


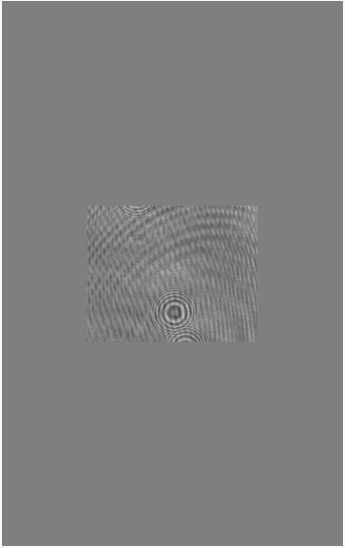
Détection hors-champ (2)



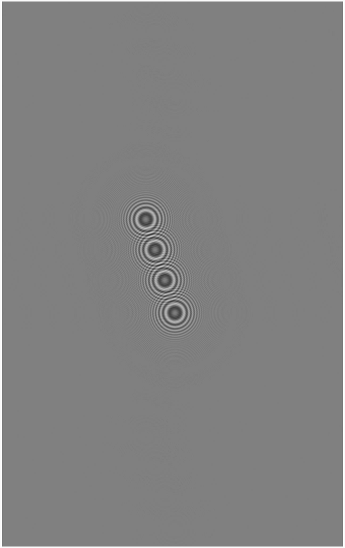


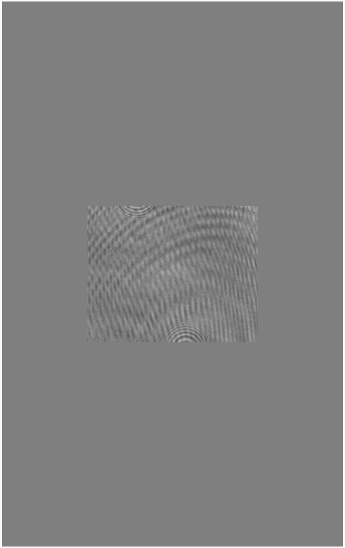
Détection hors-champ (3)



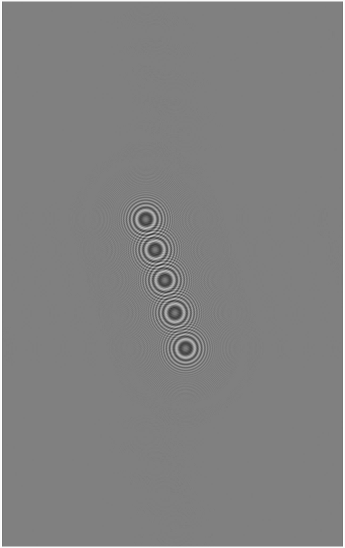


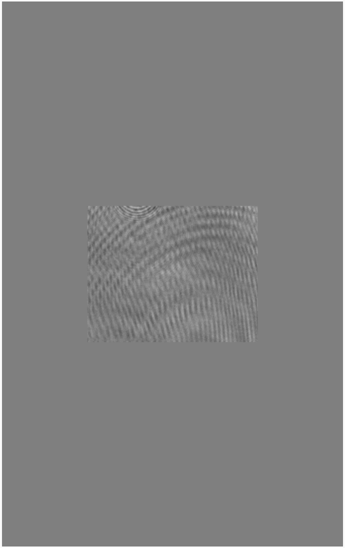
Détection hors-champ (4)



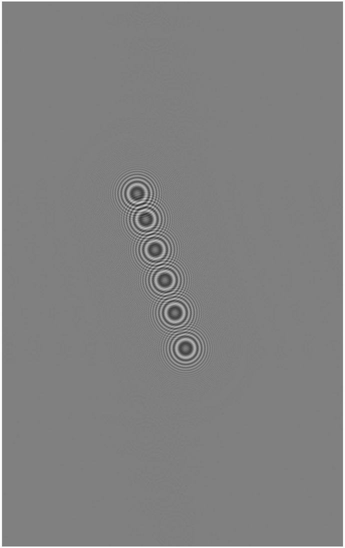


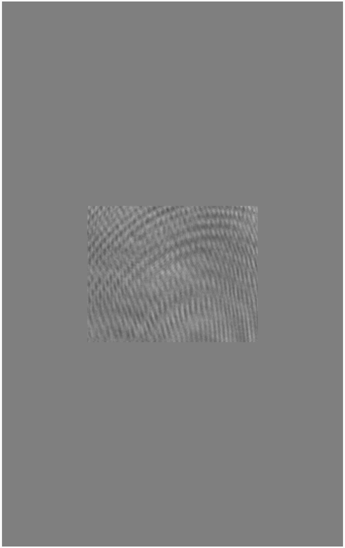
Détection hors-champ (5)



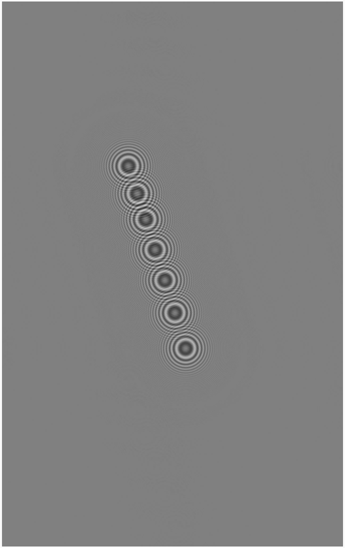


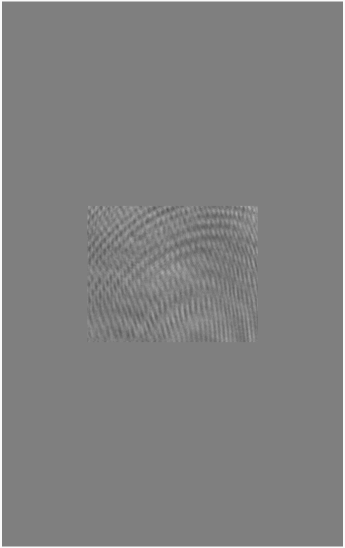
Détection hors-champ (6)



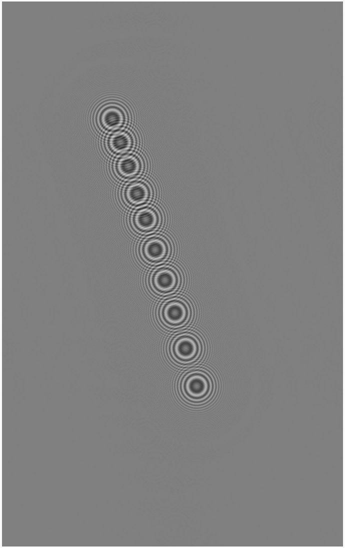


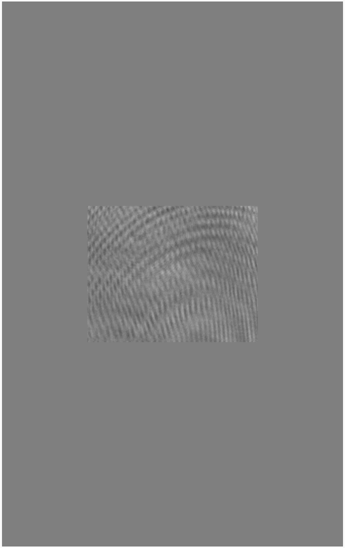
Détection hors-champ(7)



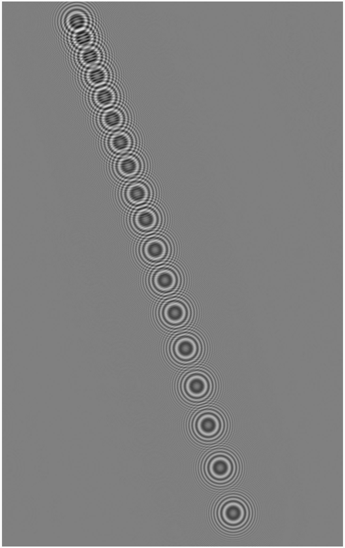


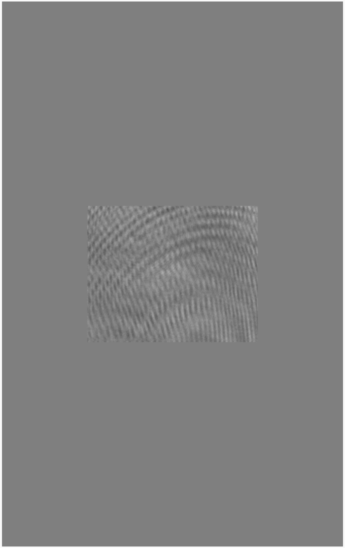
Détection hors-champ(10)





Détection hors-champ(18)





Maximum de vraisemblance: déconvolution

Maximum de vraisemblance : Approximation gaussienne, modèle linéaire

- modèle :

$$\boldsymbol{y} = \mathbf{H} \cdot \boldsymbol{x} + \boldsymbol{e}$$

- fonction de distribution des erreurs :

$$\mathrm{PDF}(\boldsymbol{y}|\boldsymbol{x}) = \frac{\exp\left\{-\frac{1}{2}\,[\boldsymbol{y} - \mathbf{H}\cdot\boldsymbol{x}]^{\mathrm{T}} \cdot \mathbf{C}_{e}^{-1} \cdot [\boldsymbol{y} - \mathbf{H}\cdot\boldsymbol{x}]\right\}}{(2\,\pi)^{M/2}\,|\mathbf{C}_{e}|^{\frac{1}{2}}}$$

- fonction de pénalisation :

$$f_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{x}) = (\boldsymbol{y} - \mathbf{H}\cdot\boldsymbol{x})^{\mathrm{T}} \cdot \mathbf{C}_{e}^{-1} \cdot (\boldsymbol{y} - \mathbf{H}\cdot\boldsymbol{x})$$

- condition d'optimalité du 1er ordre (***équations normales***) :

$$\left.\frac{\partial f_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{x})}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}=\boldsymbol{x}_{\mathrm{ML}}} = \mathbf{0}$$

$$\iff \quad \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{e}^{-1} \cdot \mathbf{H} \cdot \boldsymbol{x}_{\mathrm{ML}} = \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{e}^{-1} \cdot \boldsymbol{y}$$

$$\iff \quad \boxed{\boldsymbol{x}_{\mathrm{ML}} = \left(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{e}^{-1} \cdot \mathbf{H}\right)^{-1} \cdot \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{e}^{-1} \cdot \boldsymbol{y}}$$

Maximum de vraisemblance: déconvolution

- Approximation circulante: $\mathbf{H} = \mathbf{F}^{-1} \cdot \mathrm{diag}(\hat{\boldsymbol{h}}) \cdot \mathbf{F}$.
- Bruit diagonal dans Fourier: $\mathbf{C}_{\boldsymbol{e}} = \mathbf{F}^{-1} \cdot \mathrm{diag}(\hat{\boldsymbol{e}}) \cdot \mathbf{F}$.
- $$\begin{aligned} x^{\mathrm{ML}} &= \left(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{e} \cdot \mathbf{H}\right)^{-1} \cdot \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{e} \cdot \boldsymbol{y}\,, \\ \hat{x}_u^{(\mathrm{ML})} &= \frac{\hat{h}_u^* \, \hat{y}_u}{|\hat{h}_u|^2} = \frac{\hat{y}_u}{\hat{h}_u}\,, \end{aligned}$$
- C'est l'inversion directe

Maximum de vraisemblance: déconvolution

- Approximation circulante: $\mathbf{H} = \mathbf{F}^{-1} \cdot \mathrm{diag}(\hat{\boldsymbol{h}}) \cdot \mathbf{F}$.
- Bruit diagonal dans Fourier: $\mathbf{C}_{\boldsymbol{e}} = \mathbf{F}^{-1} \cdot \mathrm{diag}(\hat{\boldsymbol{e}}) \cdot \mathbf{F}$.
-

$$
\begin{aligned}
x^{\mathrm{ML}} &= \left(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}} \cdot \mathbf{H}\right)^{-1} \cdot \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}} \cdot \boldsymbol{y}\,, \\
\hat{x}_u^{(\mathrm{ML})} &= \frac{\hat{h}_u^* \, \hat{y}_u}{|\hat{h}_u|^2} = \frac{\hat{y}_u}{\hat{h}_u}\,,
\end{aligned}
$$

- C'est l'inversion directe

Maximum de vraisemblance: déconvolution

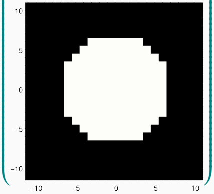
- Approximation circulante: $\mathbf{H} = \mathbf{F}^{-1} \cdot \mathrm{diag}(\hat{\boldsymbol{h}}) \cdot \mathbf{F}$.
- Bruit diagonal dans Fourier: $\mathbf{C}_{\boldsymbol{e}} = \mathbf{F}^{-1} \cdot \mathrm{diag}(\hat{\boldsymbol{e}}) \cdot \mathbf{F}$.
-
$$
\begin{aligned}
x^{\mathrm{ML}} &= \left(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}} \cdot \mathbf{H}\right)^{-1} \cdot \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}} \cdot \boldsymbol{y}\,, \\
\hat{x}_u^{(\mathrm{ML})} &= \frac{\hat{h}_u^* \, \hat{y}_u}{|\hat{h}_u|^2} = \frac{\hat{y}_u}{\hat{h}_u}\,,
\end{aligned}
$$
- C'est l'inversion directe

Inversion

$\boldsymbol{x} = \mathrm{FFT}^{-1}$

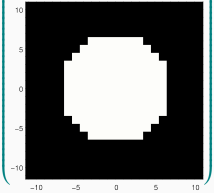
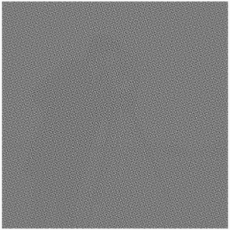
FFT

FFT



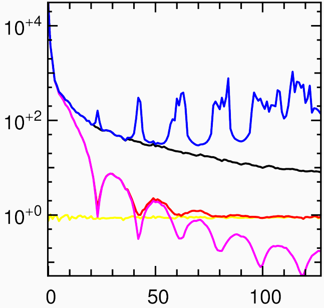
$$\widehat{x}(\boldsymbol{\nu}) = \frac{\widehat{y}(\boldsymbol{\nu})}{\widehat{h}(\boldsymbol{\nu})}$$

Inversion

$x = \mathrm{FFT}^{-1}$ (FFT () / FFT ()) = 

$$\widehat{x}(\nu) = \frac{\widehat{y}(\nu)}{\widehat{h}(\nu)}$$

Quel est le problème?



Profils radiaux des spectres de puissance

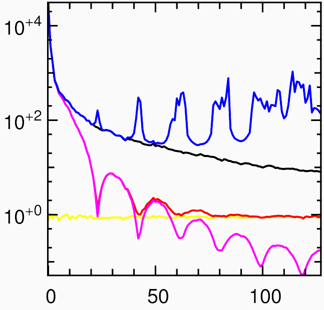
- — object,
- — blurred object,
- — noise,
- — blurred object + noise,
- — direct inversion.

- inversion directe :

$$\widehat{x}_{\mathrm{ML}}(\nu) = \frac{\widehat{y}(\nu)}{\widehat{h}(\nu)} = \widehat{x}(\nu) + \frac{\widehat{e}(\nu)}{\widehat{h}(\nu)}$$

▶ Amplification du bruit C'est bien un problème mal conditionné.

Quel est le problème?



Profils radiaux des spectres de puissance

- — object,
- — blurred object,
- — noise,
- — blurred object + noise,
- — direct inversion.

- inversion directe :

$$\widehat{x}_{\mathrm{ML}}(\nu) = \frac{\widehat{y}(\nu)}{\widehat{h}(\nu)} = \widehat{x}(\nu) + \frac{\widehat{e}(\nu)}{\widehat{h}(\nu)}$$

► Amplification du bruit C'est bien un problème mal conditionné.

Méthodes itératives

$$f_{y|x}(\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{y} - \mathbf{H}\cdot\boldsymbol{x})^{\mathrm{T}}\cdot\mathbf{C}_e^{-1}\cdot(\boldsymbol{y} - \mathbf{H}\cdot\boldsymbol{x})\,,$$
$$\nabla f_{y|x}(\boldsymbol{x}) = \mathbf{H}^{\mathrm{T}}\cdot\mathbf{C}_e^{-1}\cdot\mathbf{H}\cdot\boldsymbol{x} - \mathbf{H}^{\mathrm{T}}\cdot\mathbf{C}_e^{-1}\cdot\boldsymbol{y}$$

- **descente de gradient**:
  $$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \gamma\,\nabla f_{y|x}(\boldsymbol{x})$$
  taille du pas: $\gamma < 1/L$ avec L la constante de Lipschitz de $f_{y|x}(\boldsymbol{x})$.
- **méthodes du second ordre** (Newton):
  $$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \mathbf{B}\cdot\nabla f_{y|x}(\boldsymbol{x})$$
  où $\mathbf{B} = \left(\mathbf{H}^{\mathrm{T}}\cdot\mathbf{C}_e^{-1}\cdot\mathbf{H}\right)^{-1}$ est l'inverse de la matrice Hessienne. Approximation de $\mathbf{B}$ possible (BFGS,. . .).
- **gradients conjugués** Le plus efficace pour les fonctions quadratiques.
- méthode de points fixes

Méthodes itératives

$$f_{y|x}(\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{y} - \mathbf{H}\cdot\boldsymbol{x})^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot (\boldsymbol{y} - \mathbf{H}\cdot\boldsymbol{x}),$$

$$\nabla f_{y|x}(\boldsymbol{x}) = \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \mathbf{H} \cdot \boldsymbol{x} - \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y}$$

- **descente de gradient**:

  $$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \gamma\, \nabla f_{y|x}(\boldsymbol{x})$$

  taille du pas: $\gamma < 1/L$ avec L la constante de Lipschitz de $f_{y|x}(\boldsymbol{x})$.
- **méthodes du second ordre** (Newton):

  $$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \mathbf{B} \cdot \nabla f_{y|x}(\boldsymbol{x})$$

  où $\mathbf{B} = \left(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \mathbf{H}\right)^{-1}$ est l'inverse de la matrice Hessienne. Approximation de $\mathbf{B}$ possible (BFGS,. . .).
- **gradients conjugués** Le plus efficace pour les fonctions quadratiques.
- méthode de points fixes

Méthodes itératives

$$f_{y|x}(\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{y} - \mathbf{H}\cdot\boldsymbol{x})^{\mathrm{T}}\cdot\mathbf{C}_e^{-1}\cdot(\boldsymbol{y} - \mathbf{H}\cdot\boldsymbol{x}),$$

$$\nabla f_{y|x}(\boldsymbol{x}) = \mathbf{H}^{\mathrm{T}}\cdot\mathbf{C}_e^{-1}\cdot\mathbf{H}\cdot\boldsymbol{x} - \mathbf{H}^{\mathrm{T}}\cdot\mathbf{C}_e^{-1}\cdot\boldsymbol{y}$$

- **descente de gradient**:
  $$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \gamma\,\nabla f_{y|x}(\boldsymbol{x})$$
  taille du pas: $\gamma < 1/L$ avec L la constante de Lipschitz de $f_{y|x}(\boldsymbol{x})$.
- **méthodes du second ordre** (Newton):
  $$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \mathbf{B}\cdot\nabla f_{y|x}(\boldsymbol{x})$$
  où $\mathbf{B} = \left(\mathbf{H}^{\mathrm{T}}\cdot\mathbf{C}_e^{-1}\cdot\mathbf{H}\right)^{-1}$ est l'inverse de la matrice Hessienne. Approximation de $\mathbf{B}$ possible (BFGS,. . .).
- **gradients conjugués** Le plus efficace pour les fonctions quadratiques.
- méthode de points fixes

Méthodes itératives

$$f_{y|x}(\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{y} - \mathbf{H}\cdot\boldsymbol{x})^{\mathrm{T}}\cdot\mathbf{C}_e^{-1}\cdot(\boldsymbol{y} - \mathbf{H}\cdot\boldsymbol{x})\,,$$
$$\nabla f_{y|x}(\boldsymbol{x}) = \mathbf{H}^{\mathrm{T}}\cdot\mathbf{C}_e^{-1}\cdot\mathbf{H}\cdot\boldsymbol{x} - \mathbf{H}^{\mathrm{T}}\cdot\mathbf{C}_e^{-1}\cdot\boldsymbol{y}$$

- **descente de gradient**:
$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \gamma\,\nabla f_{y|x}(\boldsymbol{x})$$
taille du pas: $\gamma < 1/L$ avec L la constante de Lipschitz de $f_{y|x}(\boldsymbol{x})$.
- **méthodes du second ordre** (Newton):
$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \mathbf{B}\cdot\nabla f_{y|x}(\boldsymbol{x})$$
où $\mathbf{B} = \left(\mathbf{H}^{\mathrm{T}}\cdot\mathbf{C}_e^{-1}\cdot\mathbf{H}\right)^{-1}$ est l'inverse de la matrice Hessienne. Approximation de $\mathbf{B}$ possible (BFGS,. . .).
- **gradients conjugués** Le plus efficace pour les fonctions quadratiques.
- méthode de points fixes

Méthodes itératives

$$f_{y|x}(\mathbf{x}) = \frac{1}{2}(\mathbf{y} - \mathbf{H}\cdot\mathbf{x})^{\mathrm{T}}\cdot\mathbf{C}_e^{-1}\cdot(\mathbf{y} - \mathbf{H}\cdot\mathbf{x}),$$
$$\nabla f_{y|x}(\mathbf{x}) = \mathbf{H}^{\mathrm{T}}\cdot\mathbf{C}_e^{-1}\cdot\mathbf{H}\cdot\mathbf{x} - \mathbf{H}^{\mathrm{T}}\cdot\mathbf{C}_e^{-1}\cdot\mathbf{y}$$

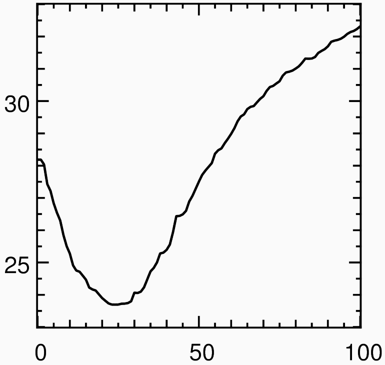
- **descente de gradient**:
$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \gamma\,\nabla f_{y|x}(\mathbf{x})$$
taille du pas: $\gamma < 1/L$ avec L la constante de Lipschitz de $f_{y|x}(\mathbf{x})$.
- **méthodes du second ordre** (Newton):
$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \mathbf{B}\cdot\nabla f_{y|x}(\mathbf{x})$$
où $\mathbf{B} = \left(\mathbf{H}^{\mathrm{T}}\cdot\mathbf{C}_e^{-1}\cdot\mathbf{H}\right)^{-1}$ est l'inverse de la matrice Hessienne. Approximation de $\mathbf{B}$ possible (BFGS,. . .).
- **gradients conjugués** Le plus efficace pour les fonctions quadratiques.
- méthode de points fixes

Si on laisse converger les algorithmes itératifs, la solution sera très bruitées (solution MV).

Erreur en fonction des itérations

Solution à l'itération 24.

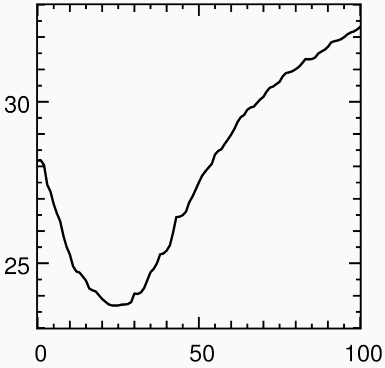
Si on laisse converger les algorithmes itératifs, la solution sera très bruitées (solution MV).



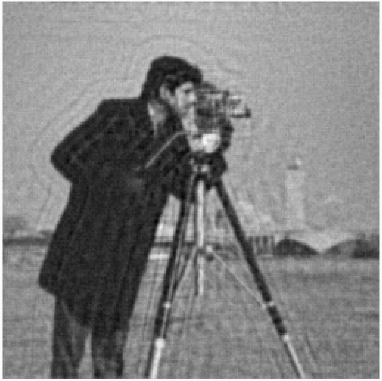
Erreur en fonction des itérations

Solution à l'itération 24.

Si on laisse converger les algorithmes itératifs, la solution sera très bruitées (solution MV).



Erreur en fonction des itérations


Solution à l'itération 24.

Maximum de vraisemblance avec contraintes

Restriction de l’espace des paramètres

Solution naive:

$$\boldsymbol{x}_{\mathrm{ML}} = \mathbf{H}^{-1} \cdot \boldsymbol{y} = \boldsymbol{x} + \mathbf{H}^{-1} \cdot \boldsymbol{e}$$

Recherche de la solution dans un sous-espace où $\mathbf{H}^{-1} \cdot \boldsymbol{e}$ est petit:

- Espace décrit par les vecteurs propres corespondant aux grandes valeurs propres de $\mathbf{H}$:

En coupant les fréquences trop élevées (ici $u_c = 40$ frequels).

Restriction de l'espace des paramètres

Solution naive:

$$\boldsymbol{x}_{\mathrm{ML}} = \mathbf{H}^{-1} \cdot \boldsymbol{y} = \boldsymbol{x} + \mathbf{H}^{-1} \cdot \boldsymbol{e}$$

Recherche de la solution dans un sous-espace où $\mathbf{H}^{-1} \cdot \boldsymbol{e}$ est petit:

- Espace décrit par les vecteurs propres corespondant aux grandes valeurs propres de $\mathbf{H}$:

En coupant les fréquences trop élevées (ici $u_c = 40\,\text{frequels}$).

Débruitage de la solution naïve

Solution naive:

$$\boldsymbol{x}_{\mathrm{ML}} = \mathbf{H}^{-1} \cdot \boldsymbol{y} = \boldsymbol{x} + \mathbf{H}^{-1} \cdot \boldsymbol{e}$$

Recherche de la solution dans un sous-espace où $\mathbf{H}^{-1} \cdot \boldsymbol{e}$ est bien séparée de $\boldsymbol{x}$:

- Espace décrit par une base d'ondelettes $\mathbf{M}$ (ForWarD [Neelamani,2004])

$$\boldsymbol{x} = \mathbf{M}^{-1}\, S\,(\mathbf{M}\, \boldsymbol{x}_{\mathrm{ML}})$$

avec $S\,()$ un seuillage.

Restriction aux solution positive

- **Gradient projeté**

$$\boldsymbol{x}^{(k+1)} = \mathcal{P}_{>0}\left(\boldsymbol{x}^{(k)} - \gamma \, \nabla f_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{x})\right)$$

avec la projection

$$\mathcal{P}_{>0}(x) = \begin{cases} x & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$$

- **Methodes multiplicatives**
ISRA (points fixe)

$$\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \mathbf{H} \cdot \boldsymbol{x} = \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y}$$

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} \frac{\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y}}{\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \mathbf{H} \cdot \boldsymbol{x}^{(k)}}$$

Solution positive

Maximum a posteriori

Approche Bayesienne : Maximum a posteriori

- maximiser la probabilité du model étant donnée les mesures (MAP = *maximum a posteriori*) :

$$
\begin{aligned}
\boldsymbol{x}_{\mathrm{MAP}} &= \arg\max_{\boldsymbol{x}} \Pr(\boldsymbol{x} \mid \boldsymbol{y}) \\
&= \arg\max_{\boldsymbol{x}} \frac{\Pr(\boldsymbol{y}|\boldsymbol{x})\ \Pr(\boldsymbol{x})}{\Pr(\boldsymbol{y})} \qquad \text{(Bayes)} \\
&= \arg\min_{\boldsymbol{x}} \Big\{ \underbrace{-\log \Pr(\boldsymbol{y}|\boldsymbol{x})}_{f_{y|x}(\boldsymbol{x})} \underbrace{-\log \Pr(\boldsymbol{x})}_{f_{x}(\boldsymbol{x})} \Big\} \\
&= \arg\min_{\boldsymbol{x}} f_{\mathrm{post}}(\boldsymbol{x})
\end{aligned}
$$

- $f_{\mathrm{post}}(\boldsymbol{x})$ = fonction de coût ***a posteriori*** :

$$
f_{\mathrm{post}}(\boldsymbol{x}) = f_{y|x}(\boldsymbol{x}) + f_{x}(\boldsymbol{x})
$$

- $f_{y|x}(\boldsymbol{x})$ = fonction de ***vraisemblance***
- $f_{x}(\boldsymbol{x})$ = fonction de***régularisation*** (a priori)

Approximation Gaussienne

- PDF des mesures :

$$\mathrm{PDF}(\boldsymbol{y}|\boldsymbol{x}) = \frac{\exp\{-\frac{1}{2}\,[\boldsymbol{y} - \boldsymbol{m}(\boldsymbol{x})]^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot [\boldsymbol{y} - \boldsymbol{m}(\boldsymbol{x})]\}}{(2\,\pi)^{N_y/2}\,|\mathbf{C}_{\boldsymbol{e}}|^{\frac{1}{2}}}$$

- ► terme de vraisemblance :

$$f_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{x}) = (\boldsymbol{y} - \mathbf{H}\cdot\boldsymbol{x})^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot (\boldsymbol{y} - \mathbf{H}\cdot\boldsymbol{x})$$

- PDF a priori:

$$\mathrm{PDF}(\boldsymbol{x}) = \frac{\exp\{-\frac{1}{2}\,(\boldsymbol{x} - \overline{\boldsymbol{x}})^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{x}}^{-1} \cdot (\boldsymbol{x} - \overline{\boldsymbol{x}})\}}{(2\,\pi)^{N_x/2}\,|\mathbf{C}_{\boldsymbol{x}}|^{\frac{1}{2}}}$$

- ► terme de regularisation :

$$f_{\boldsymbol{x}}(\boldsymbol{x}) = (\boldsymbol{x} - \overline{\boldsymbol{x}})^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{x}}^{-1} \cdot (\boldsymbol{x} - \overline{\boldsymbol{x}})$$

Approximation gaussienne, modèle linéaire

- fonction de coût a posteriori :

$$\begin{aligned} f_{\text{post}}(\boldsymbol{x}) &= f_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{x}) + f_{\boldsymbol{x}}(\boldsymbol{x}) \\ &= (\boldsymbol{y} - \mathbf{H}\cdot\boldsymbol{x})^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot (\boldsymbol{y} - \mathbf{H}\cdot\boldsymbol{x}) + (\boldsymbol{x} - \overline{\boldsymbol{x}})^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{x}}^{-1} \cdot (\boldsymbol{x} - \overline{\boldsymbol{x}}) \end{aligned}$$

- gradient :

$$\frac{\partial f_{\text{post}}(\boldsymbol{x})}{\partial \boldsymbol{x}} = 2\,\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot (\mathbf{H}\cdot\boldsymbol{x} - \boldsymbol{y}) + 2\,\mathbf{C}_{\boldsymbol{x}}^{-1} \cdot (\boldsymbol{x} - \overline{\boldsymbol{x}})$$

- solution MAP :

$$\boxed{\begin{aligned} \boldsymbol{x}_{\text{MAP}} &= \arg\min_{\boldsymbol{x}} f_{\text{post}}(\boldsymbol{x}) \\ &= (\mathbf{C}_{\boldsymbol{x}}^{-1} + \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \mathbf{H})^{-1} \cdot (\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y} + \mathbf{C}_{\boldsymbol{x}}^{-1} \cdot \overline{\boldsymbol{x}}) \end{aligned}}$$

solution MAP

$$
\begin{aligned}
\boldsymbol{x}_{\mathrm{MAP}} &= \arg\max_{\boldsymbol{x}} \Pr(\boldsymbol{x} \mid \boldsymbol{y}) \\
&= (\mathbf{C}_{x}^{-1} + \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{e}^{-1} \cdot \mathbf{H})^{-1} \cdot (\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{e}^{-1} \cdot \boldsymbol{y} + \mathbf{C}_{x}^{-1} \cdot \overline{\boldsymbol{x}}) \\
&= \overline{\boldsymbol{x}} + (\mathbf{C}_{x}^{-1} + \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{e}^{-1} \cdot \mathbf{H})^{-1} \cdot \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{e}^{-1} \cdot (\boldsymbol{y} - \mathbf{H} \cdot \overline{\boldsymbol{x}}) \\
&= \overline{\boldsymbol{x}} + \mathbf{C}_{x} \cdot \mathbf{H}^{\mathrm{T}}(\mathbf{H} \cdot \mathbf{C}_{x} \cdot \mathbf{H}^{\mathrm{T}} + \mathbf{C}_{e})^{-1} \cdot (\boldsymbol{y} - \mathbf{H} \cdot \overline{\boldsymbol{x}}) \\
&= \boldsymbol{x}_{\mathrm{MMSE}} \\
&= \arg\min_{\widetilde{\boldsymbol{x}}} \langle \|\boldsymbol{x} - \widetilde{\boldsymbol{x}}\|^{2} \rangle
\end{aligned}
$$

- Dans le cas Gaussien le MAP a les mêmes propriétés que le MMSE (PDF, expectations,. . .)

En pratique

Calcul direct (filtre de Wiener)

Toutes les matrices ($\mathbf{H}$, $\mathbf{C}_x$ and $\mathbf{C}_e$) doivent être diagonales dans le même espace.
Attention: $\mathbf{C}_e$ est en général diagonale que dans l'espace des mesures.

Optimisation de $f_{\text{post}}(\boldsymbol{x})$

$(\mathbf{C}_x^{-1} + \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \mathbf{H}) \cdot \boldsymbol{x} = \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y}$ résolu via les gradients conjugués ou des méthode de Newton ou quasi-Newton (Hessian: $\mathbf{B} = \mathbf{C}_x^{-1} + \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \mathbf{H}$)

Méthodes du point fixe (ISRA / RLA)

$$\frac{(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y})}{(\mathbf{C}_x^{-1} + \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \mathbf{H}) \cdot \boldsymbol{x}} = 1 \quad \Rightarrow \quad \boldsymbol{x}^{t+1} = \boldsymbol{x}^{t} \times \frac{\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y}}{(\mathbf{C}_x^{-1} + \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \mathbf{H}) \cdot \boldsymbol{x}} \,.$$

En pratique

Calcul direct (filtre de Wiener)

Toutes les matrices ($\mathbf{H}$, $\mathbf{C}_x$ and $\mathbf{C}_e$) doivent être diagonales dans le même espace.
Attention: $\mathbf{C}_e$ est en général diagonale que dans l'espace des mesures.

Optimisation de $f_{\text{post}}(\boldsymbol{x})$

$(\mathbf{C}_x^{-1} + \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \mathbf{H}) \cdot \boldsymbol{x} = \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y}$ résolu via les gradients conjugués ou des méthode de Newton ou quasi-Newton (Hessian: $\mathbf{B} = \mathbf{C}_x^{-1} + \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \mathbf{H}$)

Méthodes du point fixe (ISRA / RLA)

$$\frac{(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y})}{(\mathbf{C}_x^{-1} + \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \mathbf{H}) \cdot \boldsymbol{x}} = 1 \quad \Rightarrow \quad \boldsymbol{x}^{t+1} = \boldsymbol{x}^{t} \times \frac{\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y}}{(\mathbf{C}_x^{-1} + \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \mathbf{H}) \cdot \boldsymbol{x}}\,.$$

En pratique

Calcul direct (filtre de Wiener)

Toutes les matrices ($\mathbf{H}$, $\mathbf{C}_x$ and $\mathbf{C}_e$) doivent être diagonales dans le même espace.
Attention: $\mathbf{C}_e$ est en général diagonale que dans l'espace des mesures.

Optimisation de $f_{\text{post}}(\boldsymbol{x})$

$(\mathbf{C}_x^{-1} + \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \mathbf{H}) \cdot \boldsymbol{x} = \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y}$ résolu via les gradients conjugués ou des méthode de Newton ou quasi-Newton (Hessian: $\mathbf{B} = \mathbf{C}_x^{-1} + \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \mathbf{H}$)

Méthodes du point fixe (ISRA / RLA)

$$\frac{(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y})}{(\mathbf{C}_x^{-1} + \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \mathbf{H}) \cdot \boldsymbol{x}} = 1 \quad \Rightarrow \quad \boldsymbol{x}^{t+1} = \boldsymbol{x}^{t} \times \frac{\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y}}{(\mathbf{C}_x^{-1} + \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \mathbf{H}) \cdot \boldsymbol{x}}\,.$$

En pratique

Calcul direct (filtre de Wiener)

Toutes les matrices ($\mathbf{H}$, $\mathbf{C}_x$ and $\mathbf{C}_e$) doivent être diagonales dans le même espace.
Attention: $\mathbf{C}_e$ est en général diagonale que dans l'espace des mesures.

Optimisation de $f_{\mathrm{post}}(\boldsymbol{x})$

$(\mathbf{C}_x^{-1} + \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \mathbf{H}) \cdot \boldsymbol{x} = \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y}$ résolu via les gradients conjugués ou des méthode de Newton ou quasi-Newton (Hessian: $\mathbf{B} = \mathbf{C}_x^{-1} + \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \mathbf{H}$)

Méthodes du point fixe (ISRA / RLA)

$$\frac{(\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y})}{(\mathbf{C}_x^{-1} + \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \mathbf{H}) \cdot \boldsymbol{x}} = 1 \quad \Rightarrow \quad \boldsymbol{x}^{t+1} = \boldsymbol{x}^{t} \times \frac{\mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \boldsymbol{y}}{(\mathbf{C}_x^{-1} + \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_e^{-1} \cdot \mathbf{H}) \cdot \boldsymbol{x}} \,.$$

Probabilité a posteriori

$$\begin{aligned}
\mathrm{PDF}(\boldsymbol{x}|\boldsymbol{y}) &= \frac{\mathrm{PDF}(\boldsymbol{y}|\boldsymbol{x})\ \mathrm{PDF}(\boldsymbol{x})}{\int \mathrm{PDF}(\boldsymbol{y}|\boldsymbol{x}')\ \mathrm{PDF}(\boldsymbol{x}')\,\mathrm{d}\boldsymbol{x}'} \\
&= \frac{\exp\{-\frac{1}{2}\,(\boldsymbol{x} - \boldsymbol{x}_{\mathrm{MAP}}(\boldsymbol{y}))^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{x},\boldsymbol{y}}^{-1} \cdot (\boldsymbol{x} - \boldsymbol{x}_{\mathrm{MAP}}(\boldsymbol{y}))\}}{(2\,\pi)^{N_x/2}\,|\mathbf{C}_{\boldsymbol{x},\boldsymbol{y}}|^{\frac{1}{2}}}
\end{aligned}$$

avec la covariance *a posteriori*:

$$\begin{aligned}
\mathbf{C}_{\boldsymbol{x},\boldsymbol{y}} &= (\mathbf{C}_{\boldsymbol{x}}^{-1} + \mathbf{H}^{\mathrm{T}} \cdot \mathbf{C}_{\boldsymbol{e}}^{-1} \cdot \mathbf{H})^{-1} \\
&\leq \mathbf{C}_{\boldsymbol{x}}
\end{aligned}$$

MAP non linéaire

MAP généralisation

$$\boldsymbol{x}_{\text{MAP}} = \arg\min_{\boldsymbol{x}} f_{\text{ML}}(\boldsymbol{x}) + f_{\text{prior}}(\boldsymbol{x})$$

Cas Gaussien centré

$$f_{\text{prior}}(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}} \mathbf{C}_{\boldsymbol{x}}^{-1} \boldsymbol{x}$$

Cas Général

$$f_{\text{prior}}(\boldsymbol{x}) = \mu \, \varphi(\mathbf{D} \, \boldsymbol{x})$$

- μ: hyper-paramètre
- φ: fonction de coût,
- $\mathbf{D}$: opérateur linéaire.

MAP: régularisation non linéraire:

Une littérature importante sur les régularisation non linéaires:

- maximum d'entropie $f_{\text{prior}}(\boldsymbol{x}) = \mu \sum_k \left[p_k - x_k + x_k \log\left(\frac{x_k}{p_k}\right)\right]$
- variation totale $f_{\text{prior}}(\boldsymbol{x}) = \mu \; \|\mathbf{D}\,\boldsymbol{x}\|_1$
- $\ell_2 - \ell_1$: $\varphi(\boldsymbol{x}) = \sqrt{\|\boldsymbol{x}\|_2^2 + \epsilon^2} - \epsilon$
- ...

Bonne propriétés:

- convexité,
- dérivabilité,
- rapide à calculer.

Restoration with TV $\text{MSE} = 21\,\text{dB}$

Approche analyse, approche synthèse

approche analyse

$$x_{\mathrm{MAP}} = \arg\min_{x} \|\mathbf{H}\, x - y\|_{\mathbf{W}}^2 + \mu\, \varphi(\mathbf{D}\, x)$$

approche synthèse

Si $\mathbf{D}$ est inversible (base d'ondelettes)

$$x_{\mathrm{MAP}} = \arg\min_{x} \left\|\mathbf{H}\,\mathbf{D}^{-1}\, x - y\right\|_{\mathbf{W}}^2 + \mu\, \varphi(x)$$

Solutions identiques si $\varphi(x) = \|x\|_2^2$ (Parseval).

Approche analyse, approche synthèse

approche analyse

$$\boldsymbol{x}_{\text{MAP}} = \arg\min_{\boldsymbol{x}} \|\mathbf{H}\,\boldsymbol{x} - \boldsymbol{y}\|^2_{\mathbf{W}} + \mu\,\varphi(\mathbf{D}\,\boldsymbol{x})$$

approche synthèse

Si $\mathbf{D}$ est inversible (base d'ondelettes)

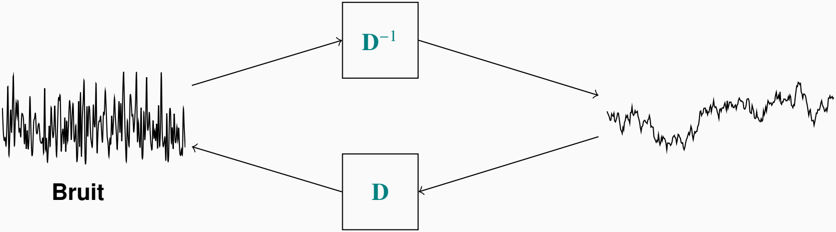
$$\boldsymbol{x}_{\text{MAP}} = \arg\min_{\boldsymbol{x}} \left\|\mathbf{H}\,\mathbf{D}^{-1}\,\boldsymbol{x} - \boldsymbol{y}\right\|^2_{\mathbf{W}} + \mu\,\varphi(\boldsymbol{x})$$

Solutions identiques si $\varphi(\boldsymbol{x}) = \|\boldsymbol{x}\|_2^2$ (Parseval).

Parcimonie

$$f_{\text{prior}}(\boldsymbol{x}) = \mu\, \varphi(\mathbf{D}\, \boldsymbol{x})$$

- $\mathbf{D}$ “blanchit” le signal: $\mathrm{Cov}(\mathbf{D}\, \boldsymbol{x}) = \mathbf{I}$
- $\varphi(z) = -\log\left(-\log \Pr(z)\right)$

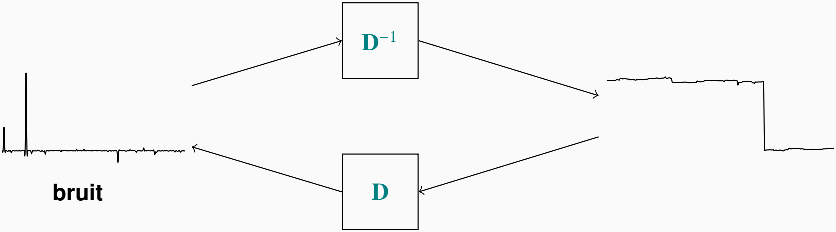


Bruit Gaussien $\Longrightarrow$ Mouvement Brownien

régularisation: interprétation

$$f_{\text{prior}}(\boldsymbol{x}) = \mu\,\varphi(\mathbf{D}\,\boldsymbol{x})$$

- $\mathbf{D}$ “blanchit” le signal: $\mathrm{Cov}(\mathbf{D}\,\boldsymbol{x}) = \mathbf{I}$
- $\varphi(z) = -\log\left(-\log \Pr(z)\right)$



Bruit Cauchy $\Longrightarrow$ vol de Lévy

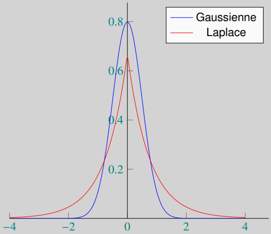
Sparse stochastic processes

Beaucoup de signaux semblent mieux modélisés par des processus parcimonieux.

Laplace

$$\Pr(z) = \frac{\lambda}{2} \mathrm{e}^{-\lambda |z|}$$

$$\varphi(z) = -\log\left(\Pr(z)\right) \propto \lambda\,|z|$$



MAP: Parcimonie

Il existe une base ψ où le signal peut être décrit par peu de coefficients (compressibilité)

MAP a priori de parcimonie

$$\boldsymbol{x}_{\mathrm{MAP}} = \arg\min_{\boldsymbol{x}} \|\mathbf{H}\,\boldsymbol{x} - \boldsymbol{y}\|^2_{\mathbf{W}} + \mu\,|\psi\,\boldsymbol{x}|$$

φ est la norme ℓ_1 (norme ℓ_p: $\|\boldsymbol{x}\|_p = \left(\sum_k x_k^p\right)^{1/p}$)

Compressive sensing

Extrêmement efficace lorsque l'espace des mesures est incohérent avec la base ψ.
Cohérence:

$$\mu(\psi, \mathbf{H}) = \sqrt{n} \max_{1\leq k, j\leq n} |\langle \psi_{\cdot,k}, H_{j,\cdot}\rangle| \in [1, \sqrt{n}]$$

$\mu(\psi, \mathbf{H}) = 1 \Rightarrow$ bases totalement incohérentes.
Nombre de mesures nécessaires $m \propto \mu^2(\psi, \mathbf{H})\, S \; \log(n)$ pour un signal S-parcimonieux.

Il existe une base ψ où le signal peut être décrit par peu de coefficients (compressibilité)

MAP a priori de parcimonie

$$\boldsymbol{x}_{\mathrm{MAP}} = \arg\min_{\boldsymbol{x}} \|\mathbf{H}\,\boldsymbol{x} - \boldsymbol{y}\|^2_{\mathbf{W}} + \mu\,|\psi\,\boldsymbol{x}|$$

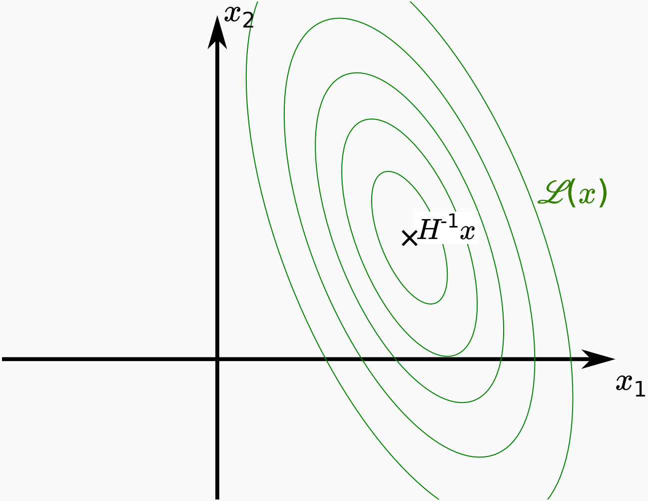
φ est la norme ℓ_1 (norme ℓ_p: $\|\boldsymbol{x}\|_p = \left(\sum_k x_k^p\right)^{1/p}$)

Compressive sensing

Extrêmement efficace lorsque l'espace des mesures est incohérent avec la base ψ.
Cohérence:

$$\mu(\psi, \mathbf{H}) = \sqrt{n}\max_{1\leq k, j\leq n} |\langle \psi_{.,k}, H_{j,.}\rangle| \in [1, \sqrt{n}]$$

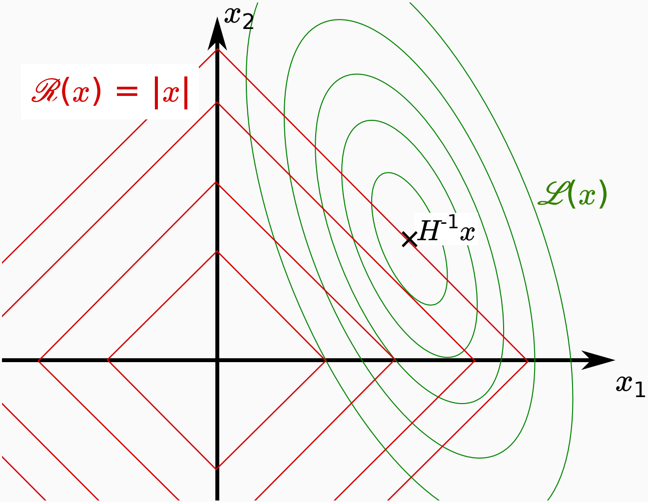
$\mu(\psi, \mathbf{H}) = 1 \Rightarrow$ bases totalement incohérentes.
Nombre de mesures nécessaires $m \propto \mu^2(\psi, \mathbf{H})\,S\,\log(n)$ pour un signal S-parcimonieux.



Level set of the likelihood
and the regularization
At convergence level sets
are tangent:
$\nabla\mathcal{L}(\boldsymbol{x}) = -\mu\nabla\mathcal{R}(\boldsymbol{x})$
Solution path $\boldsymbol{x}^{+}(\mu)$

$$\boldsymbol{x}^{+} = \arg\min_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}) + \mu\,\mathcal{R}(\boldsymbol{x})$$

Effect of the ℓ_1 norm



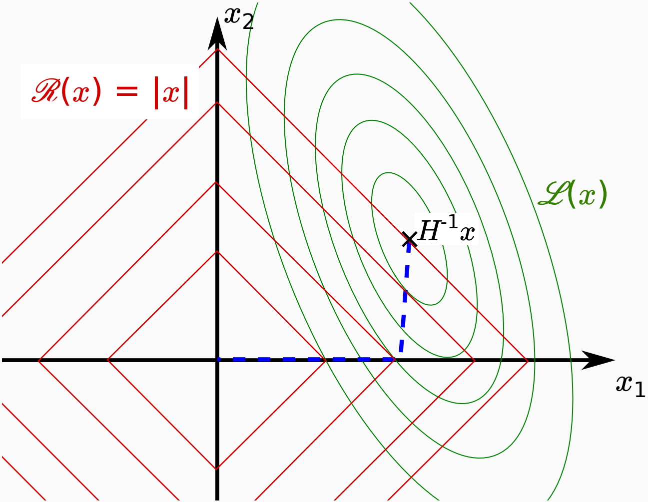
Level set of the likelihood and the regularization

At convergence level sets are tangent:

$\nabla \mathcal{L}(x) = -\mu \nabla \mathcal{R}(x)$

Solution path $x^{+}(\mu)$

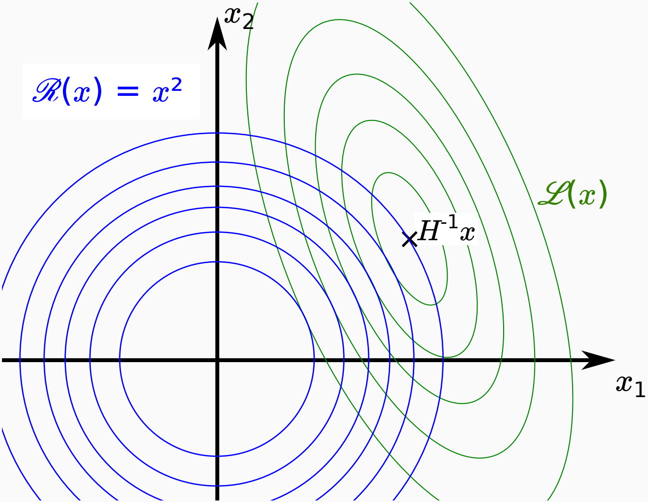
$$\boldsymbol{x}^{+} = \arg\min_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}) + \mu\, \mathcal{R}(\boldsymbol{x})$$



Level set of the likelihood and the regularization
At convergence level sets are tangent:
$\nabla \mathcal{L}(\boldsymbol{x}) = -\mu \nabla \mathcal{R}(\boldsymbol{x})$
Solution path $\boldsymbol{x}^+(\mu)$

$$\boldsymbol{x}^+ = \arg\min_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}) + \mu \mathcal{R}(\boldsymbol{x})$$

Effect of the ℓ_1 norm



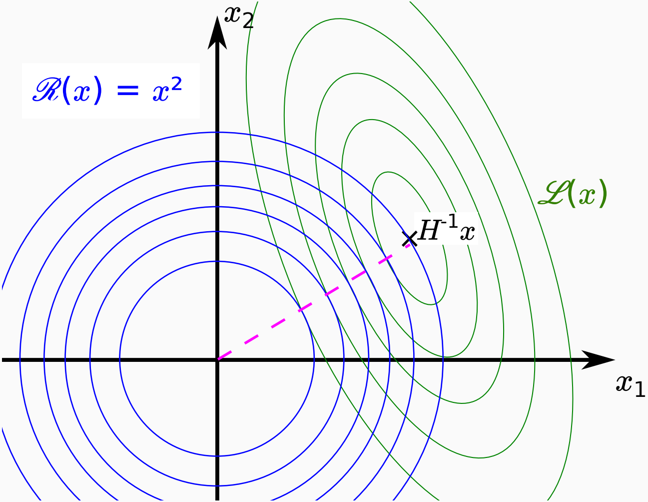
Level set of the likelihood and the regularization

At convergence level sets are tangent:

$\nabla \mathcal{L}(\boldsymbol{x}) = -\mu \nabla \mathcal{R}(\boldsymbol{x})$

Solution path $\boldsymbol{x}^+(\mu)$

$$\boldsymbol{x}^+ = \arg\min_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}) + \mu \mathcal{R}(\boldsymbol{x})$$



Level set of the likelihood and the regularization
At convergence level sets are tangent:
$\nabla \mathcal{L}(\boldsymbol{x}) = -\mu \nabla \mathcal{R}(\boldsymbol{x})$
Solution path $\boldsymbol{x}^+(\mu)$

$$\boldsymbol{x}^+ = \arg\min_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}) + \mu\, \mathcal{R}(\boldsymbol{x})$$

Optimisation non différentiable

$|\boldsymbol{x}|$ n'est pas dérivable en $x = 0$

Rappel: contraintes convexes (e.g. positivité)

Estimer le maximum de vraisemblance sous contraintes de positivité revient à:

$$\boldsymbol{x}_{\mathrm{ML}} = \arg\min_{\boldsymbol{x}} f_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{x}) + c_{>0}(\boldsymbol{x})$$

$$\text{avec l'indicatrice } c_{\geq 0}(\boldsymbol{x}) = \begin{cases} 0 & \text{si } x \geq 0 \\ +\infty & \text{sinon} \end{cases}$$

L'indicatrice $c_{\geq 0}(\boldsymbol{x})$ n'est pas dérivable en 0.
Solution par descente de gradient projeté:

$$\boldsymbol{x}^{(k+1)} = \mathcal{P}_{\geq 0}\left(\boldsymbol{x}^{(k)} - \gamma \nabla f_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{x})\right)$$

avec la projection (séparable):

$$\begin{aligned} \mathcal{P}_{\geq 0}(y) &= \arg\min_{\boldsymbol{x}} \frac{1}{2} \|\boldsymbol{x} - \boldsymbol{y}\|_2^2 + c_{\geq 0}(\boldsymbol{x}) \\ &= \begin{cases} x & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases} \end{aligned}$$

Optimisation non différentiable

$|\boldsymbol{x}|$ n'est pas dérivable en $x = 0$

Problème

$$\boldsymbol{x}_{\mathrm{ML}} = \arg\min_{\boldsymbol{x}} f_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{x}) + \lambda\, g(\boldsymbol{x})$$

$f_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{x})$ est dérivable et $g(\boldsymbol{x})$ est semi-continue par dessous.

Algorithme de gradient proximaux

Similaire à la descente de gradient projeté:

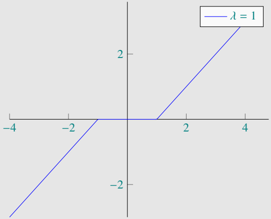
$$\boldsymbol{x}^{(k+1)} = \mathrm{prox}_{\lambda g}\left(\boldsymbol{x}^{(k)} - \gamma\, \nabla f_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{x})\right)$$

avec l'opérateur proximal:

$$\mathrm{prox}_{\lambda g}(\boldsymbol{y}) = \arg\min_{\boldsymbol{x}} \frac{1}{2}\,\|\boldsymbol{x} - \boldsymbol{y}\|_2^2 + \lambda\, g(\boldsymbol{x})$$

Opérateur proximal de ℓ_1

$$\begin{aligned}\mathrm{prox}_{\lambda \ell_1}(y) &= \arg\min_{\boldsymbol{x}} \frac{1}{2} \|\boldsymbol{x} - \boldsymbol{y}\|_2^2 + \lambda \,|\boldsymbol{x}| \,. \\ &= \mathrm{sign}(y) \max(|\boldsymbol{y}| - \lambda, 0) \,.\end{aligned}$$



Seuillage doux