

Optimisation non lisse pour la RdF et l'apprentissage

Jalal Fadili

GREYC, CNRS-ENSICAEN-Université de Caen,
<http://www.greyc.ensicaen.fr/~jfadili>

Réunion NormaSTIC Juin 2012



Vue d'ensemble

● **Thèmes de recherches:**

- Traitement du signal et des images.
- Mathématiques appliquées au signal et à l'image.

● **Socle théorique:**

- Mathématiques statistiques pour la modélisation et l'analyse.
- Analyse harmonique computationnelle, parcimonie.
- Optimisation, analyse convexe.
- Analyse non-lisse, analyse variationnelle.

● **Originalité:** couvrir toute la chaîne image de la physique d'acquisition à la modélisation, l'analyse, le traitement et le développement logiciel.

● **Applications:**

- Médicales.
- Astronomie/astrophysique.
- Contrôle non-destructif.
- Photographie computationnelle.

Optimisation non lisse

Optimisation non lisse

- **Score théorique:**
 - Optimisation convexe et non-convexe.
 - Optimisation non-lisse.
 - Optimisation sur les variétés.
 - Analyse de sensibilité et de perturbation.
 - Analyse variationnelle.

Optimisation non lisse

● **Score théorique:**

- Optimisation convexe et non-convexe.
- Optimisation non-lisse.
- Optimisation sur les variétés.
- Analyse de sensibilité et de perturbation.
- Analyse variationnelle.

● **Verrous:**

- Algorithmes accélérés.
- Problèmes non convexes et non lisses.
- Garanties (convergence, propriétés points d'accumulation, minimiseurs).
- Problèmes de grande taille.

Optimisation non lisse

● **Score théorique:**

- Optimisation convexe et non-convexe.
- Optimisation non-lisse.
- Optimisation sur les variétés.
- Analyse de sensibilité et de perturbation.
- Analyse variationnelle.

● **Verrous:**

- Algorithmes accélérés.
- Problèmes non convexes et non lisses.
- Garanties (convergence, propriétés points d'accumulation, minimiseurs).
- Problèmes de grande taille.

● **Contributions:**

- Nouveaux algorithmes avec garanties théoriques.
- Algorithmes à métrique variable en non lisse.
- Garanties sur les minimiseurs en régularisation parcimonieuse.
- Stabilité des minimiseurs et choix des paramètres.
- De très nombreuses instances en problèmes inverses, RdF et apprentissage.

Optimisation convexe non lisse

- Objectif : résoudre

$$\min_{x \in \mathcal{H}} \left\{ F(x) := \sum_{i=1}^n F_i(x) \right\},$$

$F_i : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$, $F_i \in \Gamma_0(\mathcal{H})$, et condition(s) de qualification.

- Apparaît dans de nombreux contextes : signal et image, problèmes inverses, RdF, apprentissage, classification, etc..
- Difficile car F convexe mais non nécessairement lisse.
- ● Exploiter la structure composite de la fonctionnelle primale, duale ou des deux \Rightarrow *décomposition d'opérateurs monotones maximaux*.
- Exploiter les propriétés individuelles des F_i : e.g. simplicité (au sens proximal), régularité, etc.
- S'applique facilement aux problèmes de grande taille.
- Pas de sous-itération (ou très peu coûteuse).

Exemples de problèmes

$$\min_{x \in \mathcal{H}} \quad \frac{1}{2} \|y - Ax\|_2^2 + \|x\|_1$$

$$\min_{x \in \mathcal{H}} \quad \frac{1}{2} \|y - Ax\|_2^2 + \|D^*x\|_1$$

$$\min_{x \in \mathcal{H}} \quad \frac{1}{2} \|y - Ax\|_2^2 + \lambda \|D^*x\|_1 + \iota_C(x)$$

$$\min_{x \in \mathcal{H}} \quad \|D^*x\|_1 \quad \text{t.q.} \quad \|y - Ax\|_p \leq \epsilon$$

$$\min_{x \in \mathcal{H}} \quad \frac{1}{m} \sum_{i=1}^m L(\langle x, z_i \rangle + b, y_i) + \lambda \|x\|_2^2$$

$$\min_{x \in \mathcal{H}} \quad \frac{1}{m} \sum_{i=1}^m L(\langle x, z_i \rangle + b, y_i) + \lambda \|x\|_1$$

⋮

L est le coût hinge, hinge carré ou logistique.

Focus: Generalized Forward-Backward

$$\min_{x \in \mathcal{H}} f(x) + \sum_{i=1}^n g_i(x)$$

- $f, g_i : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$, $f, g_i \in \Gamma_0(\mathcal{H})$;
- f C^1 with β -Lipschitz gradient, all g_i 's are simple.
- Domain qualification condition : $(0, \dots, 0) \in \text{sri}(\{(x - y, x - x_1, \dots, x - x_n) : x \in \mathcal{H} \text{ and } y \in \text{dom}(f), x_i \in \text{dom}(g_i)\})$;
- Set of minimizers $\mathcal{M}^* \neq \emptyset$.

For $w_i \in]0, 1]$ with $\sum_{i=1}^n w_i = 1$, let \mathcal{H} be the real Hilbert space obtained by endowing the Cartesian product \mathcal{H}^n with the scalar product $\sum_i w_i \langle x_i, y_i \rangle$.

$$\mathcal{S} = \{(x_1, \dots, x_n) \in \mathcal{H} : x_1 = x_2 = \dots = x_n\},$$

$$\Pi : \mathcal{H} \rightarrow \mathcal{S}, x \mapsto (x, \dots, x) \quad (\text{canonical isometry}).$$

Proximity operator

Definition (Proximity operator [J.-J. Moreau 1962]) Let $f \in \Gamma_0(\mathcal{H})$. Then, for every $x \in \mathcal{H}$, the function $z \mapsto \frac{1}{2} \|x - z\|^2 + f(z)$ achieves its infimum at a unique point denoted by $\text{prox}_f x$. The uniquely-valued operator $\text{prox}_f : \mathcal{H} \rightarrow \mathcal{H}$ thus defined is the proximity operator of f . $\text{rprox}_f = 2 \text{prox}_f - \text{I}$ is the reflection proximity operator.

- $\text{prox}_f(x)$ is single-valued.
- $\text{prox}_{\gamma f}$ is the resolvent of the subdifferential of f : $\text{prox}_{\gamma f} = (\text{I} + \gamma \partial f)^{-1}$, $\gamma > 0$.
- The proximity operator is firmly nonexpansive. Hence its is nonexpansive and so is its reflection operator.
- Many useful properties and a whole calculus framework.

GFB Algorithm

Initialization : Choose $(z_{i,0})_{1 \leq i \leq n} \in \mathcal{H}$, $\gamma \in]0, 2\beta[$, a sequence $(\lambda_k)_k$ in $]0, 1]$, weights $w_i \in]0, 1]$ (e.g. $1/n$). Let $x_0 = \sum_{i=1}^n w_i z_{i,0}$.

Main iteration :

repeat

1. Compute the resolvent points (in parallel if desired) :

for $i = 1$ **to** n **do**

$$z_{i,k+1} = z_{i,k} + \lambda_k \left(\text{prox}_{\gamma/w_i g_i} \left(2x_k - z_{i,k} - \gamma(\nabla f(x_k) + e_{2,k}) \right) + e_{1,i,k} - x_k \right) .$$

2. Update by averaging :

$$x_{k+1} = \sum_{i=1}^n w_i z_{i,k} .$$

3. $k \leftarrow k + 1$.

until *Convergence* ;

Output : x_k .

Convergence

Theorem (Fixed step-size) *Let $\gamma \in]0, 2\beta[$, let $(\lambda_k)_{k \in \mathbb{N}}$ be a sequence in $]0, 1]$. Let $z_0 \in \mathcal{H}$. Define the sequence of the above GFB algorithm. Assume that (i) $\liminf_t \lambda_k > 0$, and (ii) that the proximity operators $\text{prox}_{\gamma/w_i g_i}$ and ∇f are computed with errors that are summable in \mathcal{H} . Then,*

- $x_k \rightarrow x \in \mathcal{M}^*$.
- $\nabla f(x_k) \rightarrow \nabla f(x)$.
- Suppose that $\text{int}(\mathcal{M}^*) \neq \emptyset$. Then, $x_k \rightarrow x \in \mathcal{M}^*$.
- If f is uniformly convex then $x_k \rightarrow x$, the unique global minimizer.

Convergence

Theorem (Fixed step-size) *Let $\gamma \in]0, 2\beta[$, let $(\lambda_k)_{k \in \mathbb{N}}$ be a sequence in $]0, 1]$. Let $z_0 \in \mathcal{H}$. Define the sequence of the above GFB algorithm. Assume that (i) $\liminf_t \lambda_k > 0$, and (ii) that the proximity operators $\text{prox}_{\gamma/w_i g_i}$ and ∇f are computed with errors that are summable in \mathcal{H} . Then,*

- $x_k \rightarrow x \in \mathcal{M}^*$.
- $\nabla f(x_k) \rightarrow \nabla f(x)$.
- Suppose that $\text{int}(\mathcal{M}^*) \neq \emptyset$. Then, $x_k \rightarrow x \in \mathcal{M}^*$.
- If f is uniformly convex then $x_k \rightarrow x$, the unique global minimizer.

Corollary (Variable step-size) *In addition to the hypotheses of the previous theorem, let the sequence $(\gamma_k)_{k \in \mathbb{N}}$ such that $0 < \underline{\gamma} \leq \gamma_k \leq \bar{\gamma} < 2\beta$, $\gamma \in [\underline{\gamma}, \bar{\gamma}]$. Assume that $(\gamma_k - \gamma)_{k \in \mathbb{N}}$ is absolutely summable. Then, the above conclusions hold true.*

Monotone operator splitting

Find the zeros of a maximal monotone operator :

$$0 \in Ax + \sum_{i=1}^n B_i x$$

- $A, B_i : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ are maximal monotone ;
- A single-valued with $\beta A \in \mathcal{A}(\frac{1}{2})$, B_i simple $\forall i$;
- $\text{zer}(A + \sum_i B_i) \neq \emptyset$.

$$T \in \mathcal{A}(1/2) \iff T \text{ firmly expansive} \iff T = \frac{I_{\mathcal{H}} + C}{2}, C \in 1\text{-Lip}(\mathcal{H}).$$

Débruitage

Original



Noisy

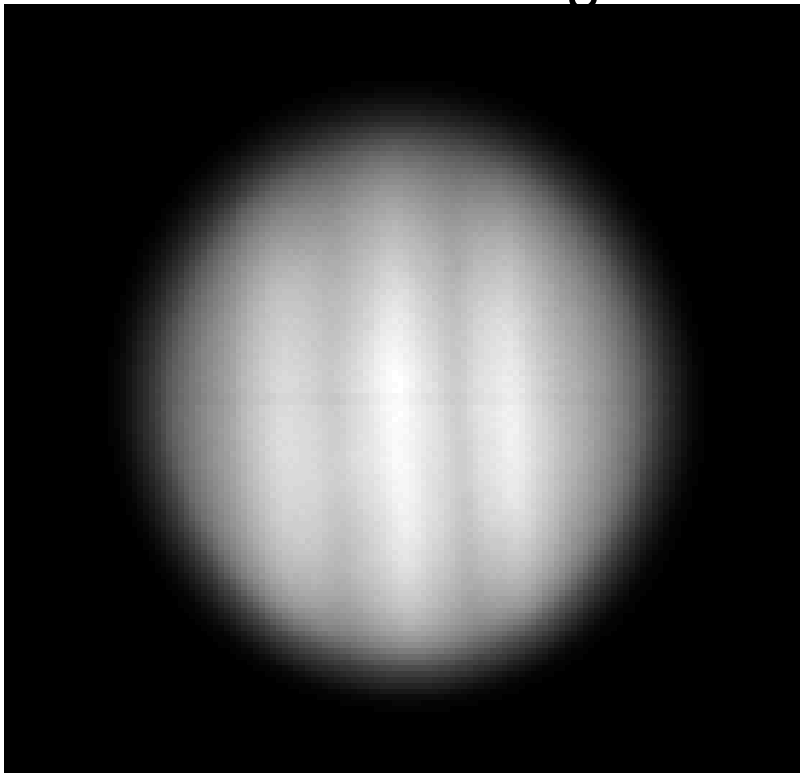


Denoised

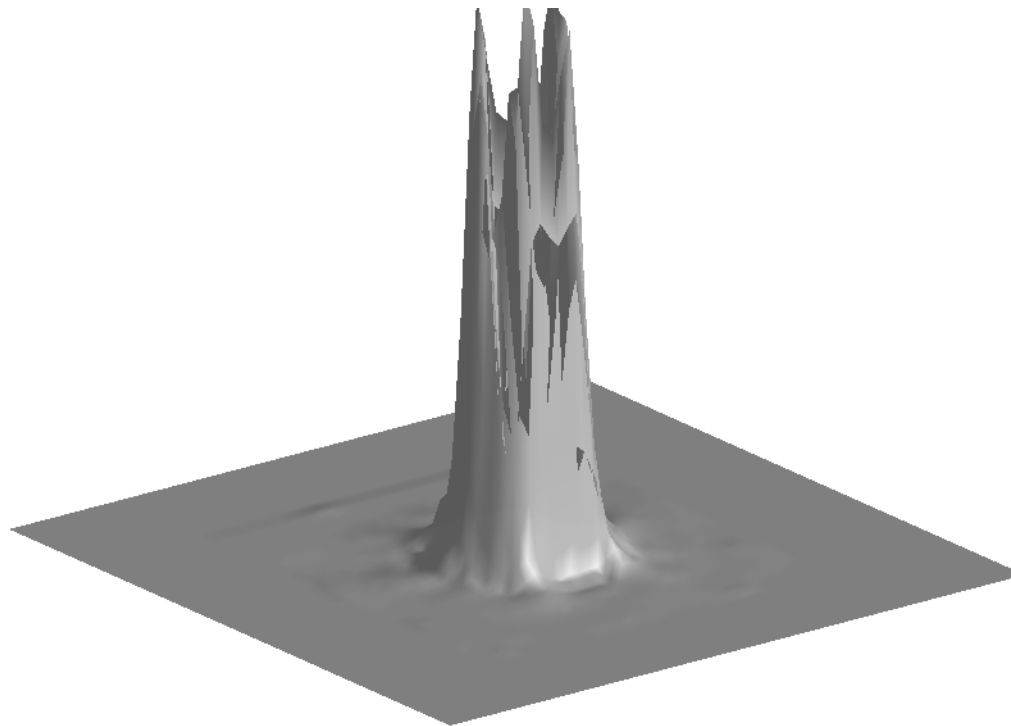


Déconvolution

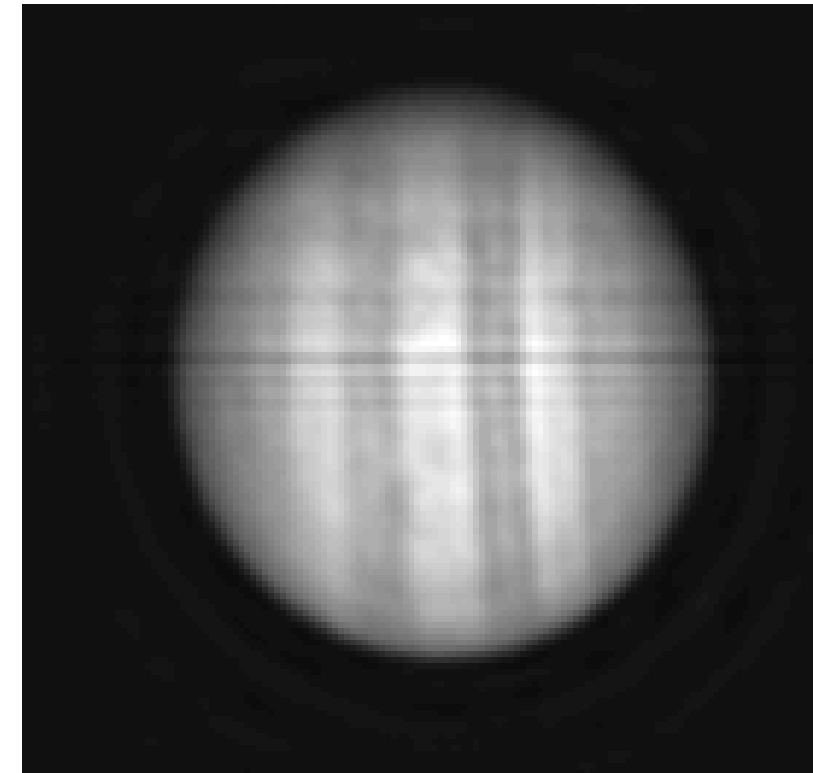
Observed image



Deconvolution
PSF



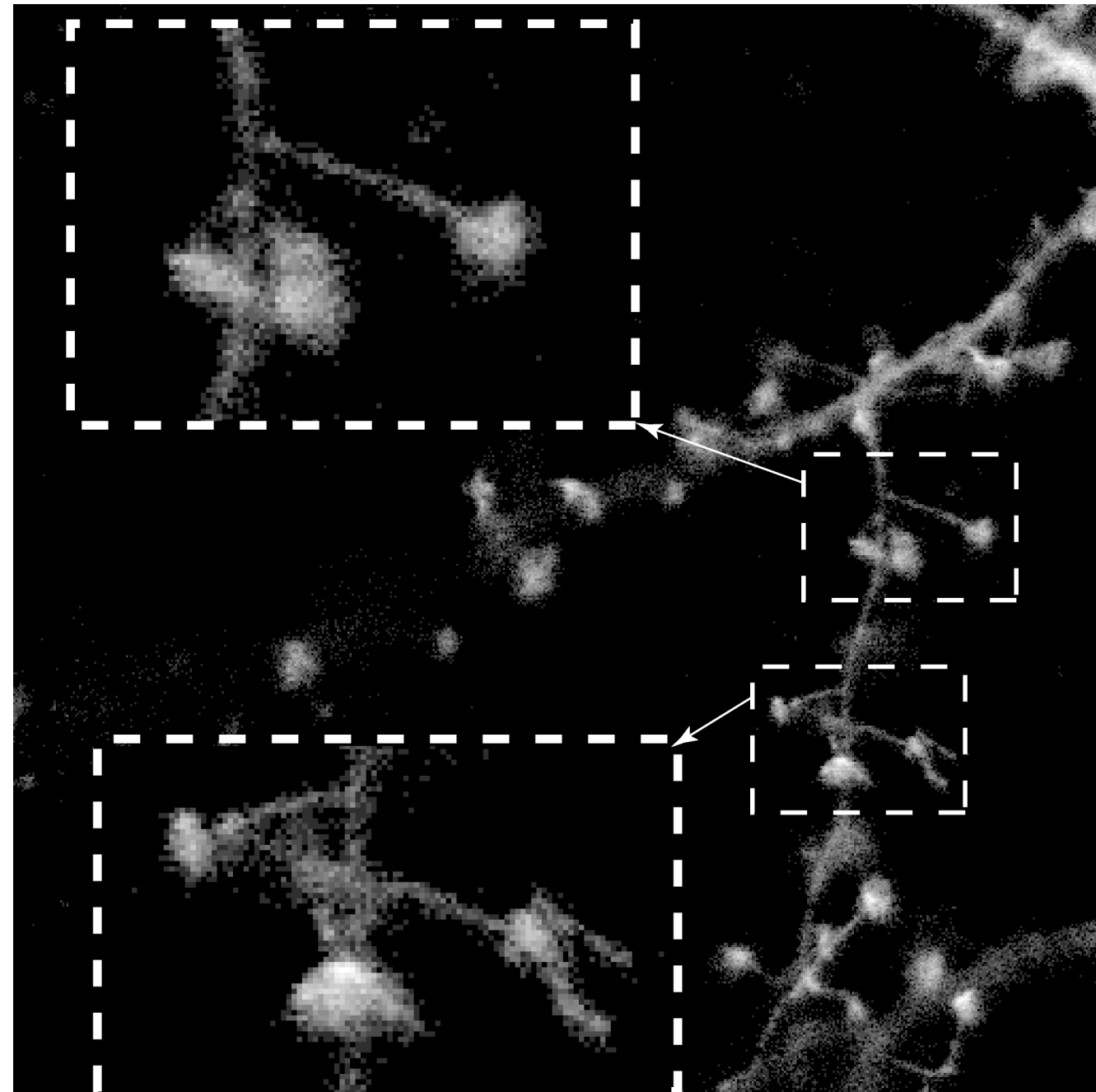
Restored



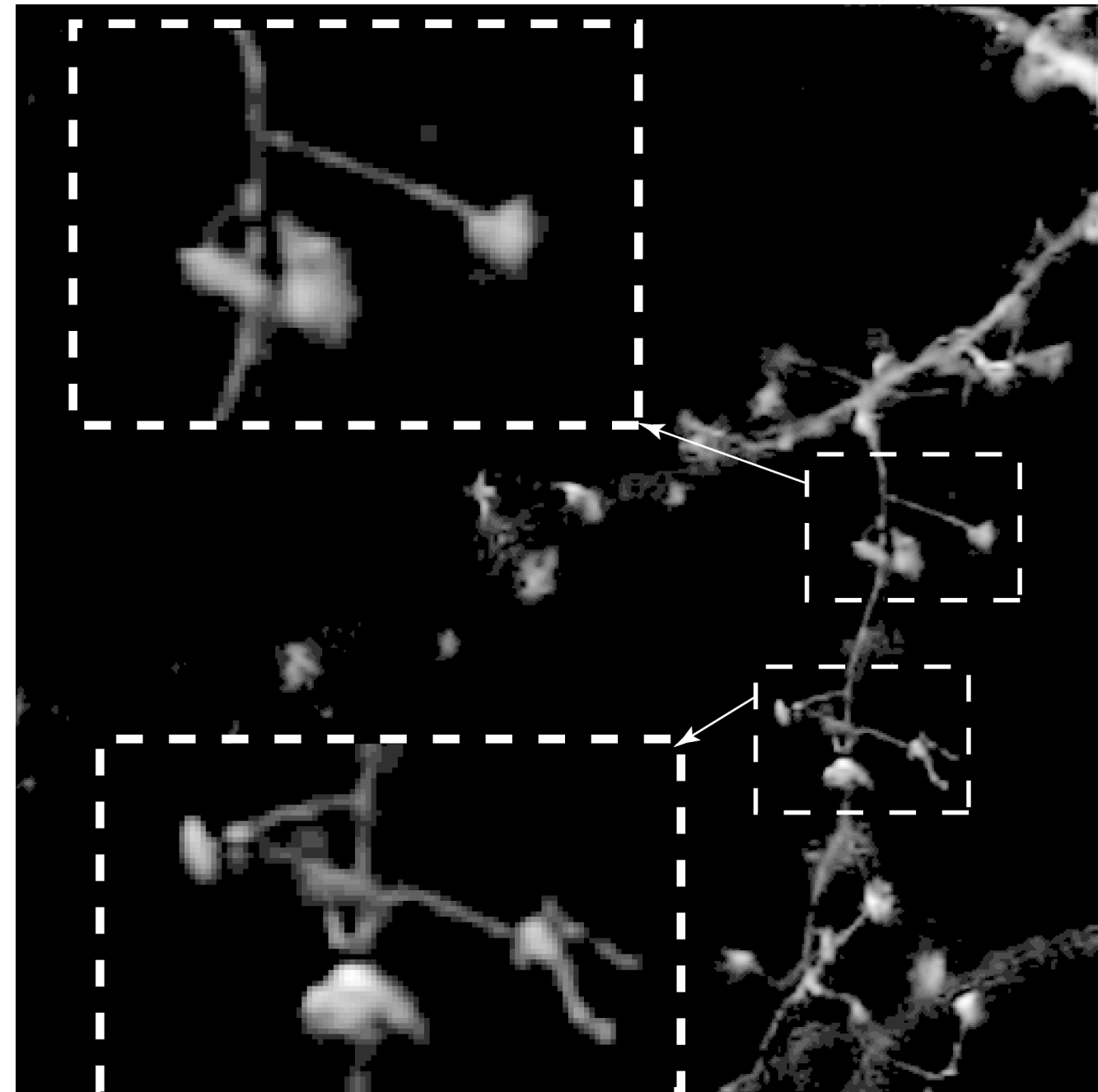
Déconvolution

Déconvolution

Observed image



Restored image



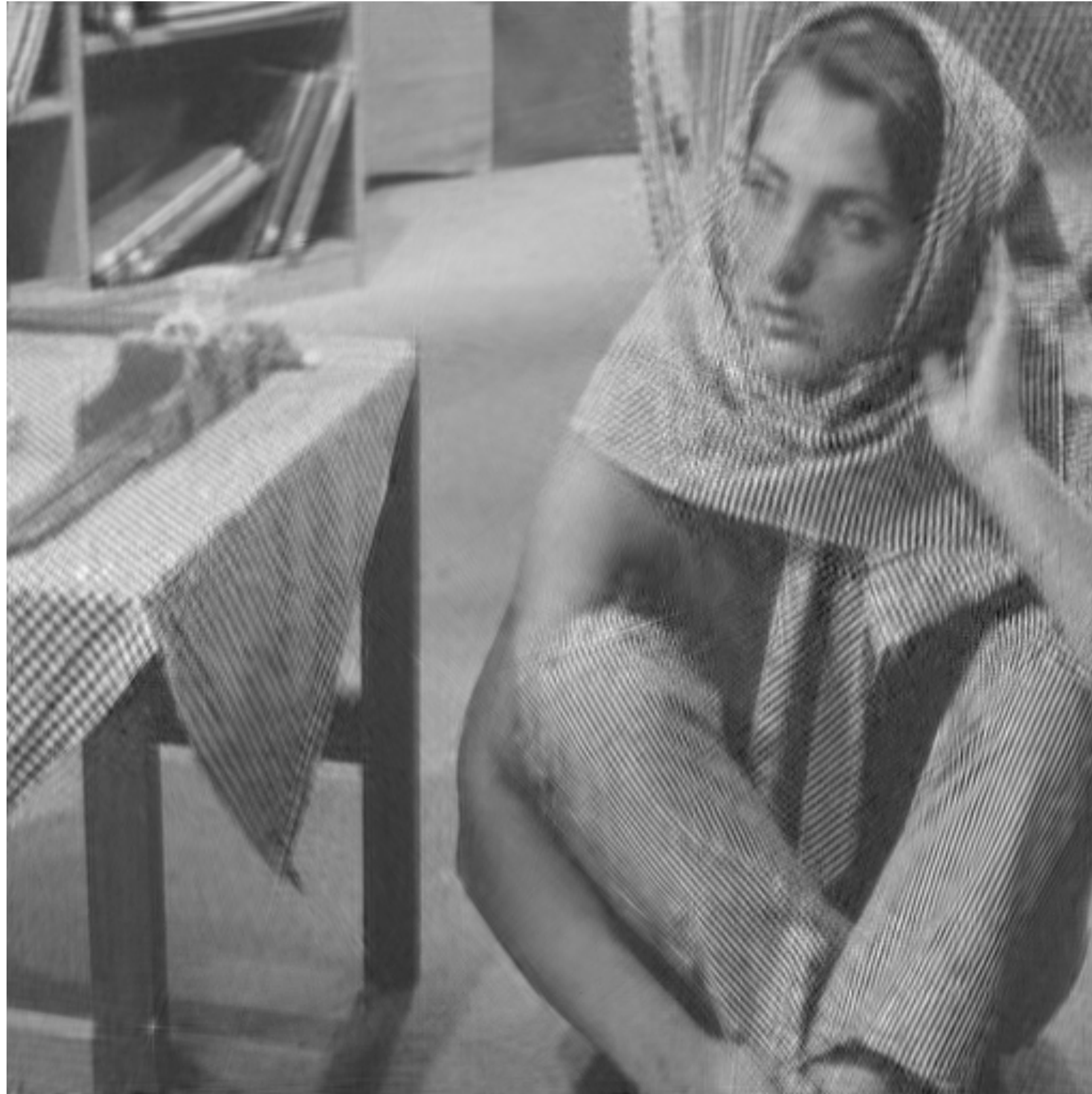
Inpainting



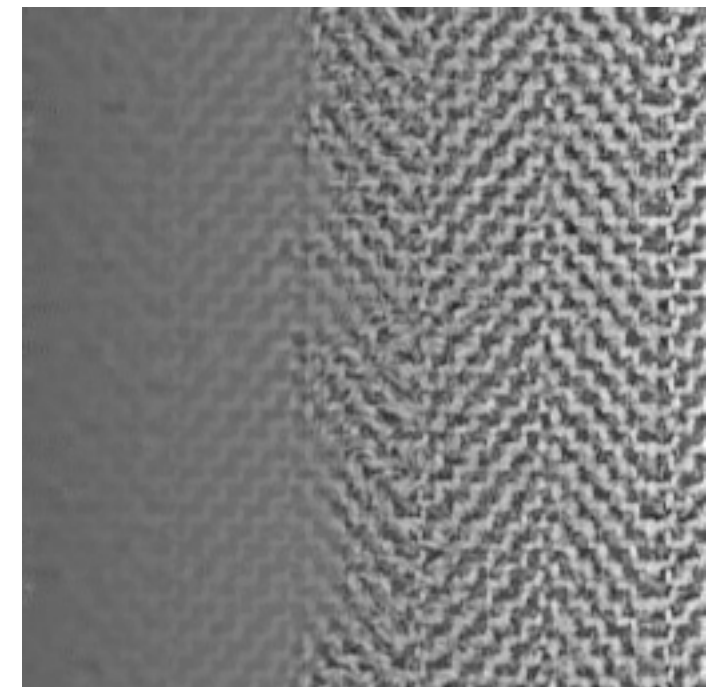
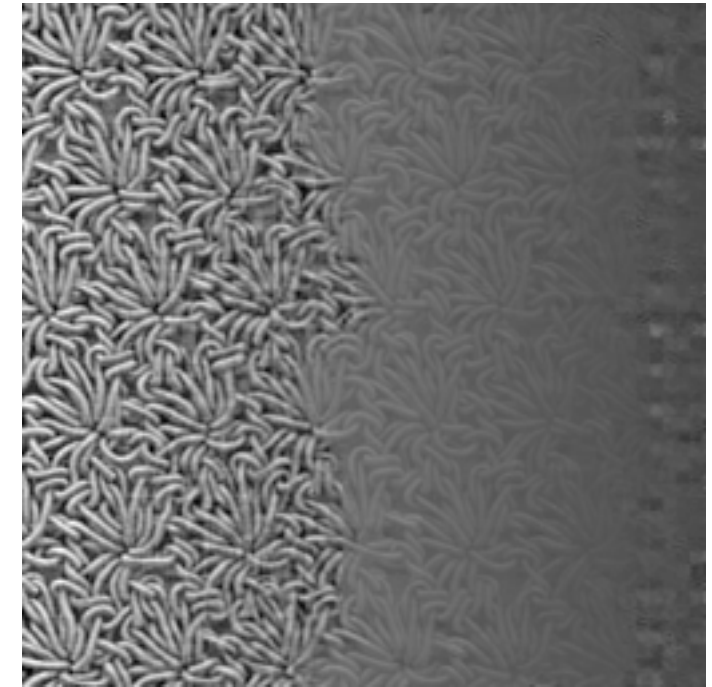
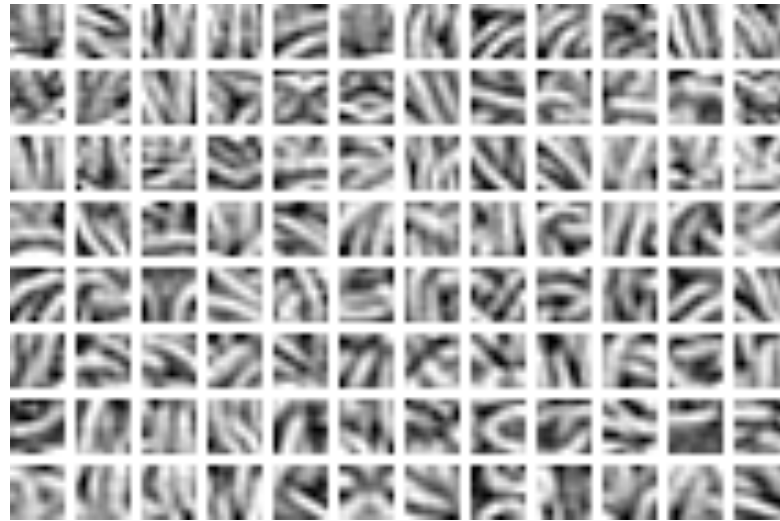
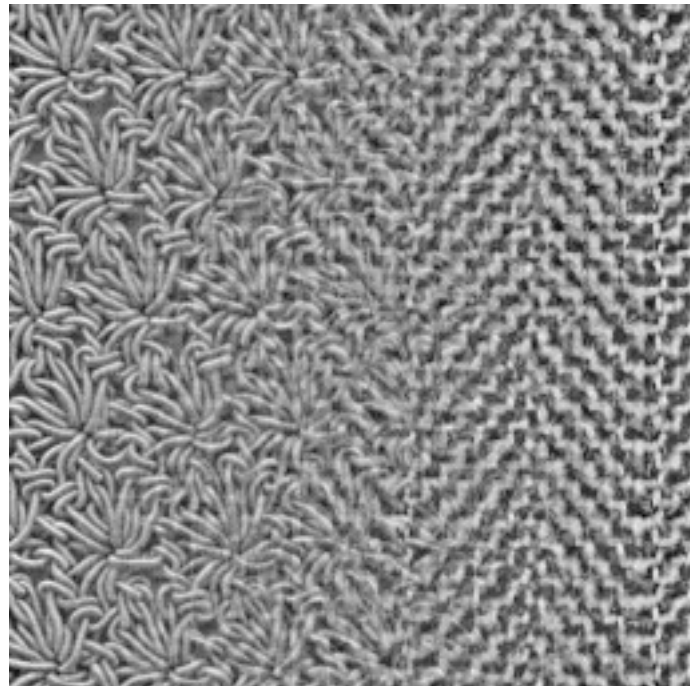
Inpainting



Inpainting

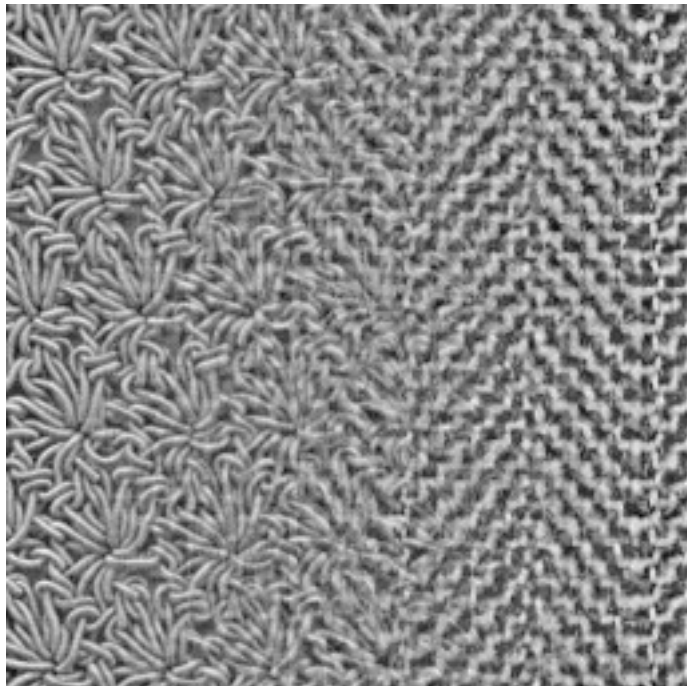


Décomposition parcimonieuse

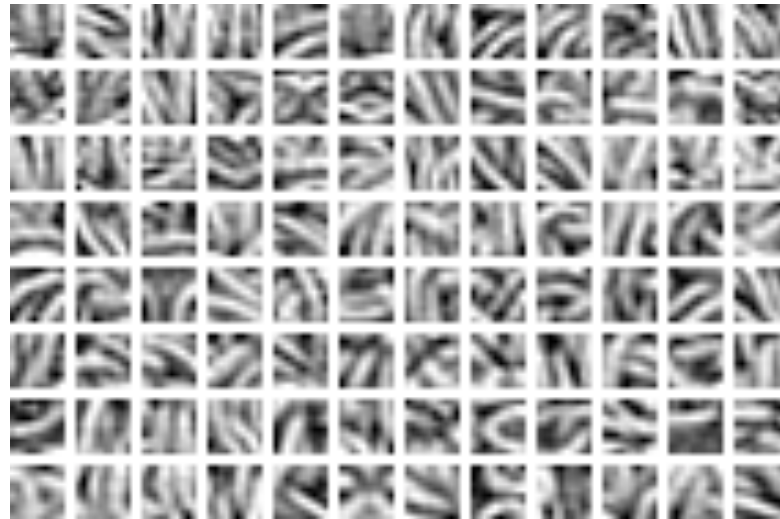


Décomposition parcimonieuse

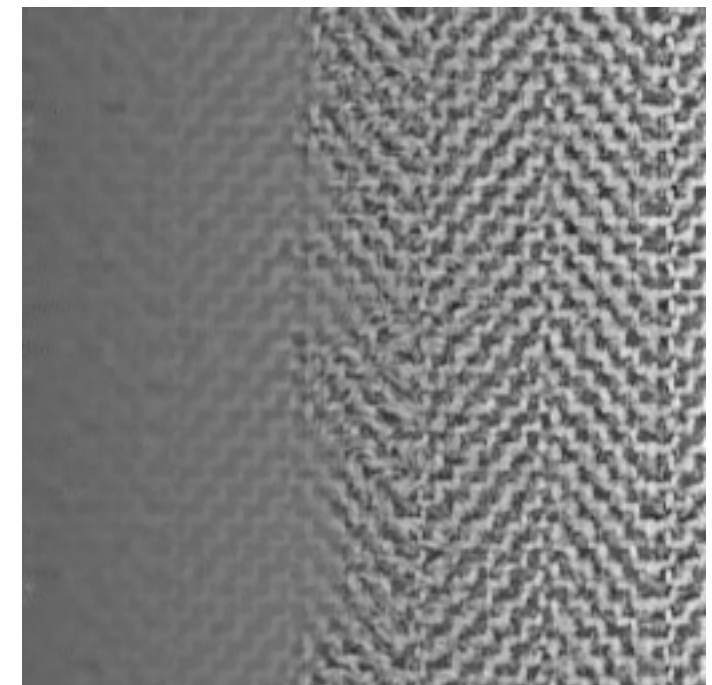
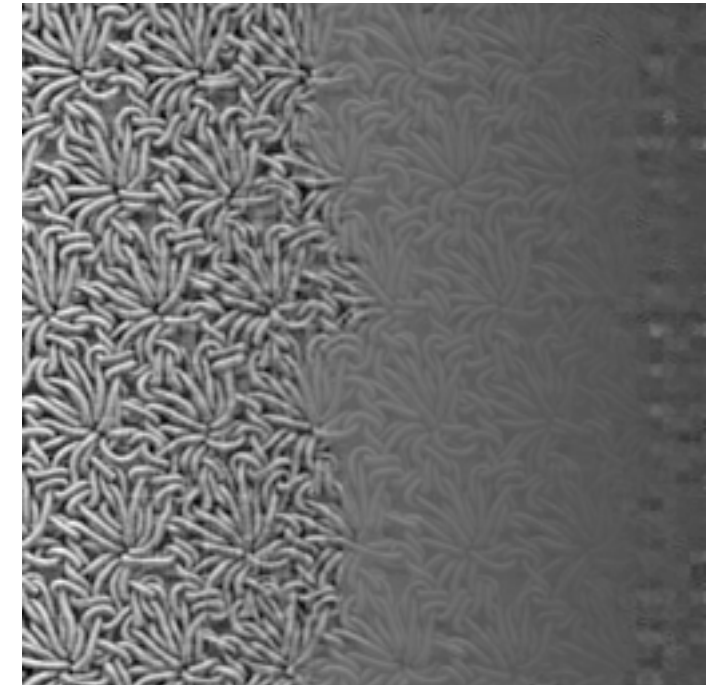
Observed mixture



Learned dictionaries
(interactive)

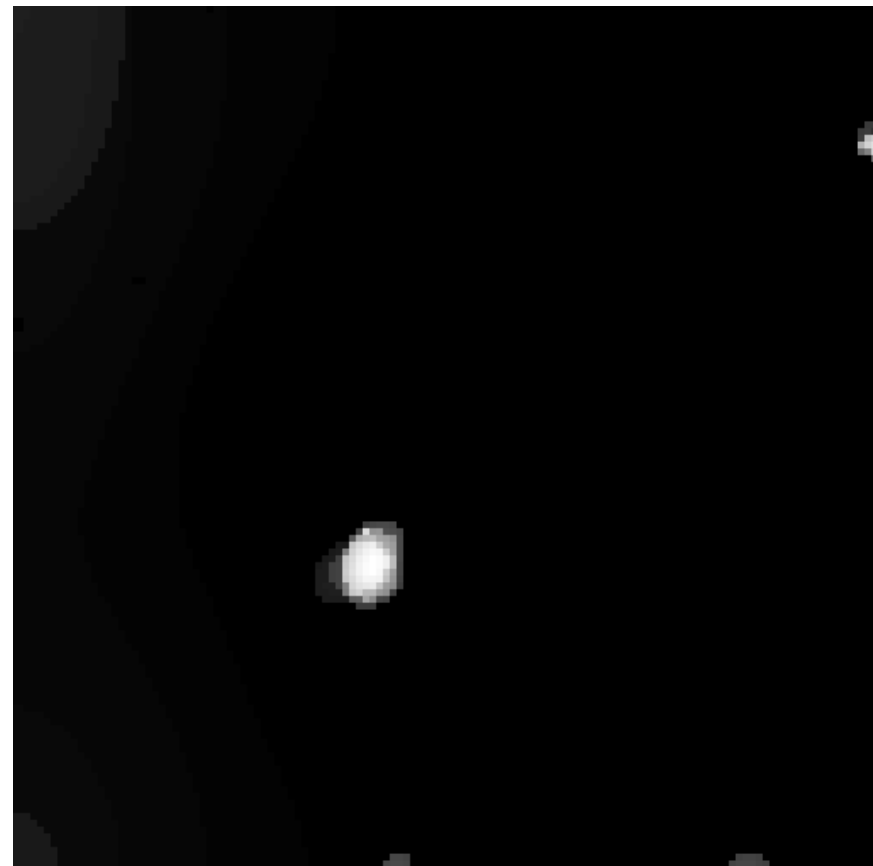
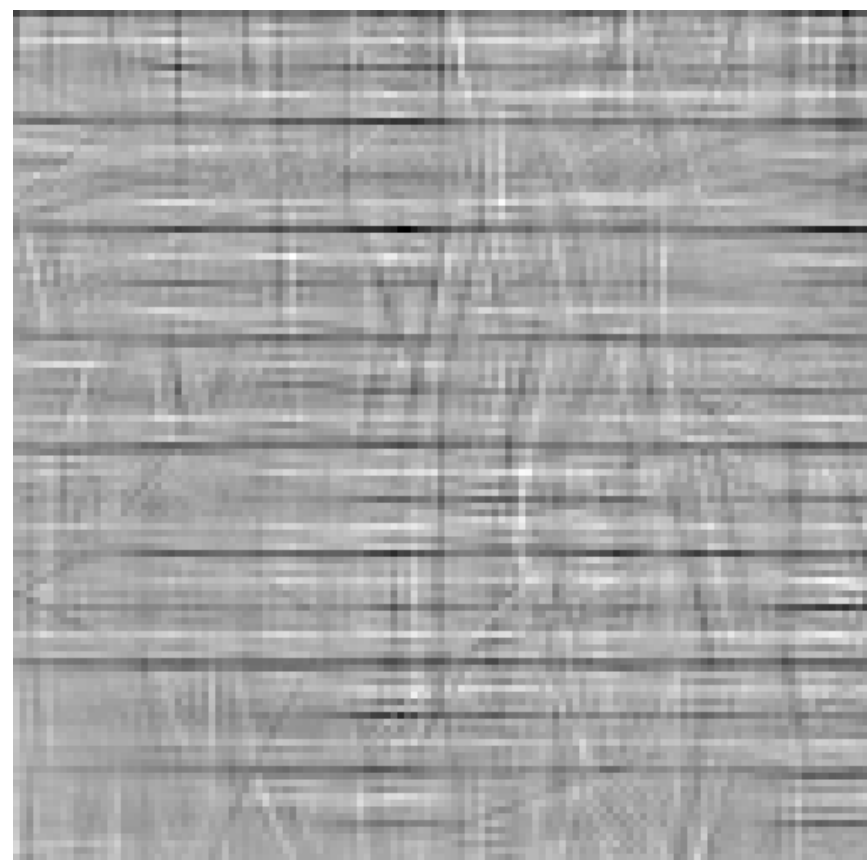
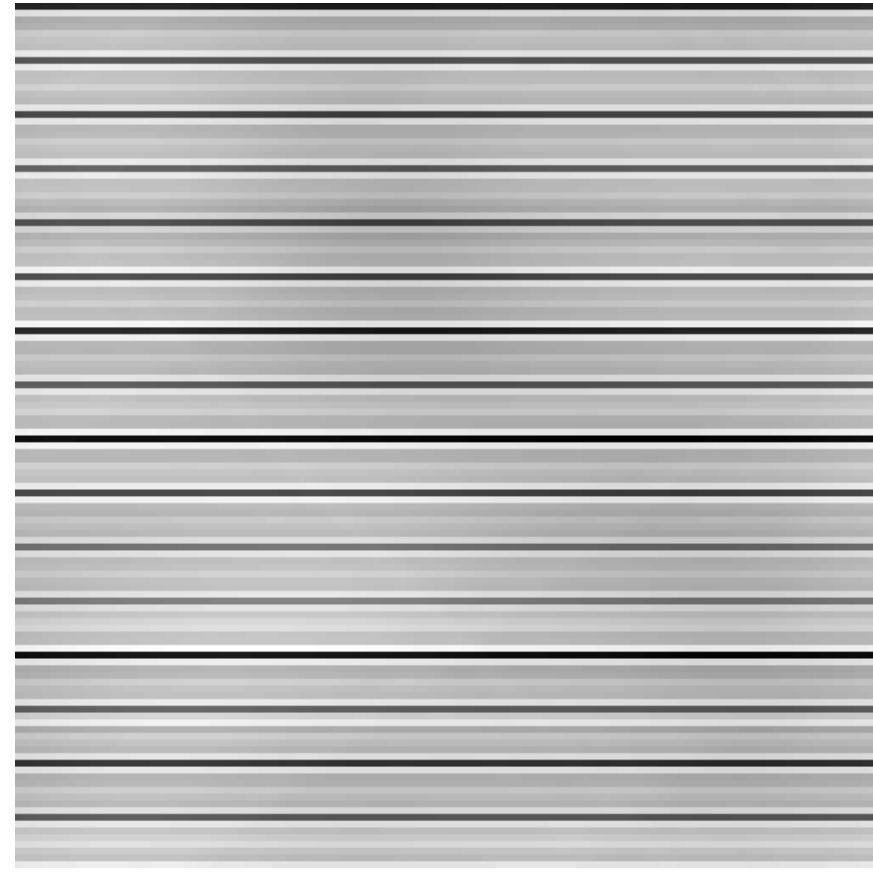
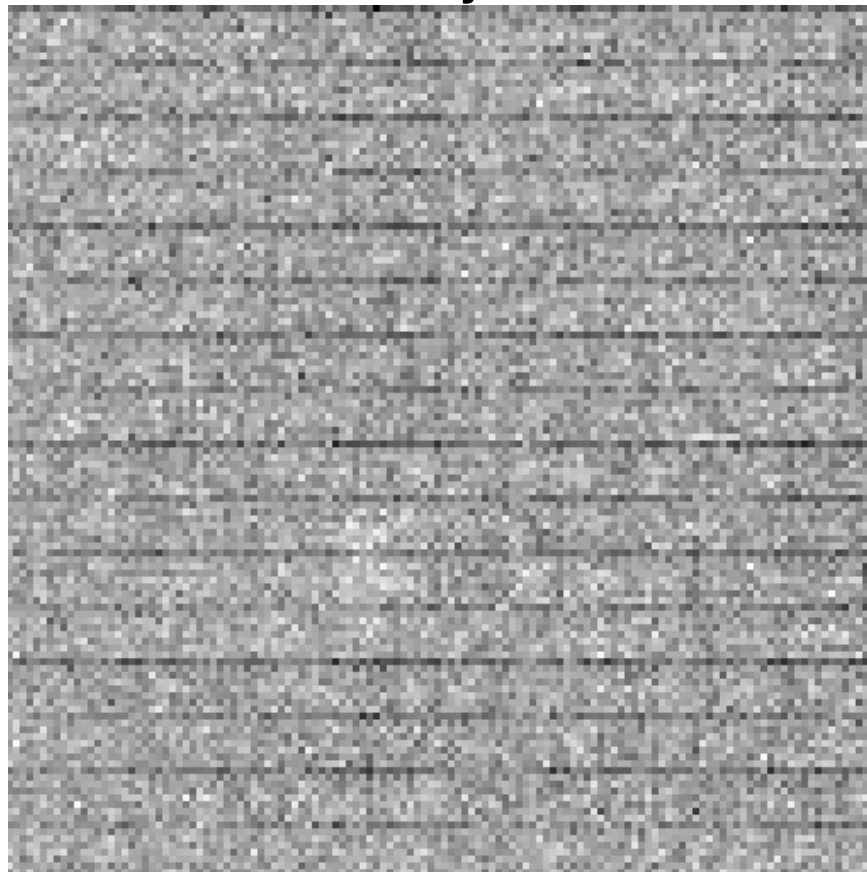


Separated textures



Décomposition parcimonieuse

Galaxy SBS 0335-052 from Gemini instrument



Séparation de sources

Original sources f_i



2 out of 4 channels y_j



Séparation de sources

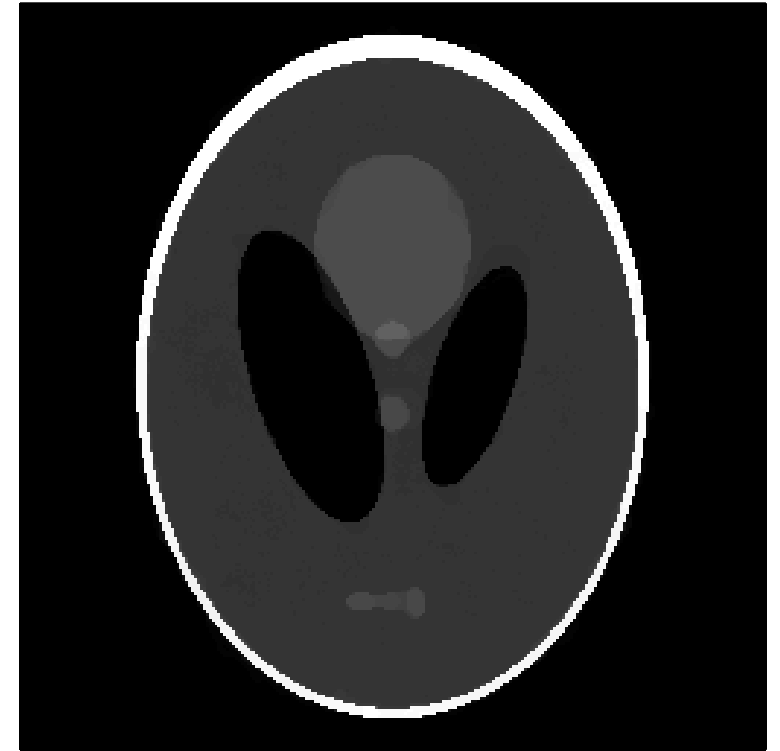
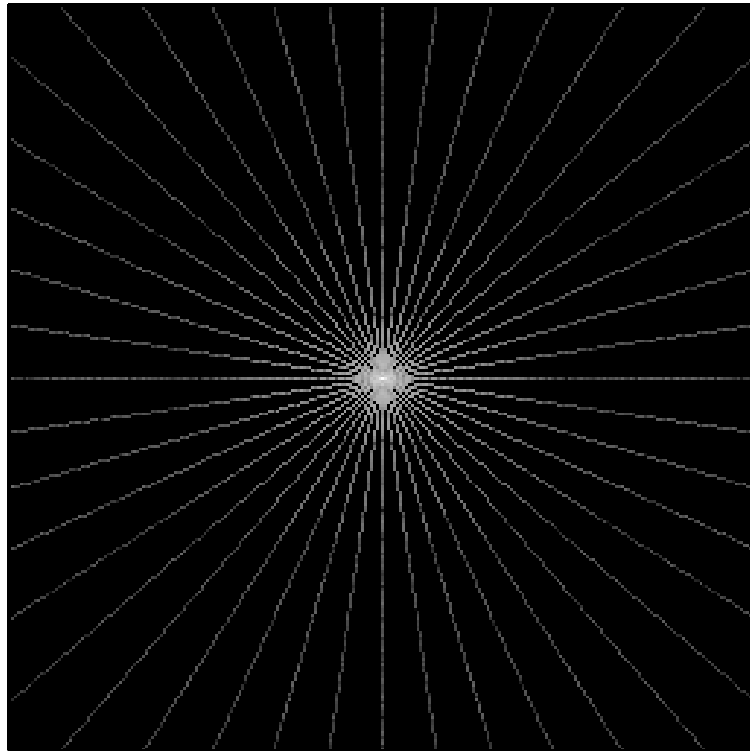
Estimated sources



Original sources

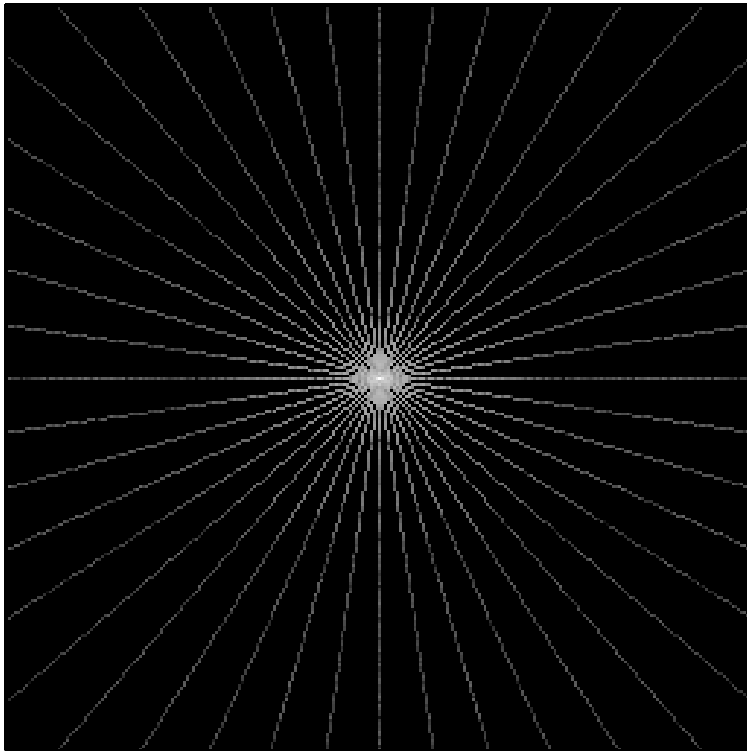


Echantillonnage compressé



Echantillonnage compressé

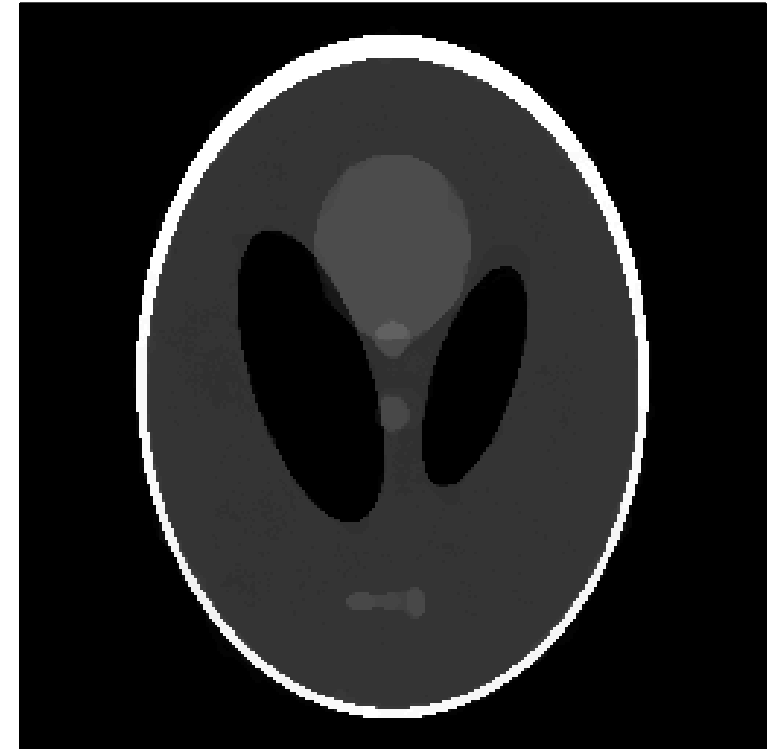
Measurements



Original image



Reconstructed image



Merci
Vos questions ?