Estimation régularisée d'attributs fractals par minimisation convexe pour la segmentation de textures.

Barbara Pascal

30 septembre 2020

Laboratoire de Physique à l'École Normale Supérieure de Lyon

Mme Laure Blanc-Féraud	Examinatrice
M. Bruno Torrésani	Rapporteur
M. Gabriel Peyré	Rapporteur
M. Rémi Bardenet	Examinateur
M. Jean-François Giovannelli	Examinateur
Mme Nelly Pustelnik	Co-directrice
M. Patrice Abry	Directeur

Estimation régularisée d'attributs fractals par minimisation convexe pour la segmentation de textures.

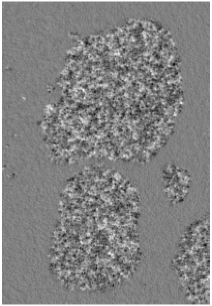
Barbara Pascal

30 septembre 2020

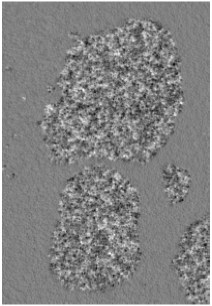
Laboratoire de Physique à l'École Normale Supérieure de Lyon

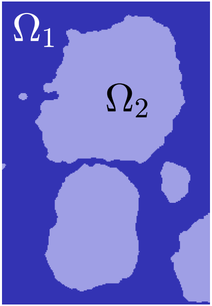
Mme Laure Blanc-Féraud	Examinatrice
M. Bruno Torrésani	Rapporteur
M. Gabriel Peyré	Rapporteur
M. Rémi Bardenet	Examinateur
M. Jean-François Giovannelli	Examinateur
Mme Nelly Pustelnik	Co-directrice
M. Patrice Abry	Directeur

Segmentation d'image



Segmentation d'image

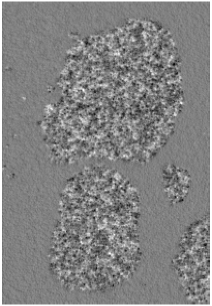


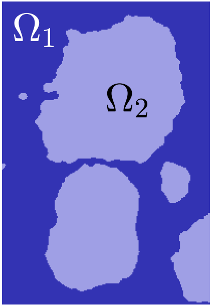


Objectif : obtenir une partition de l'image en K régions homogènes

$$\Omega = \Omega_1 \bigsqcup \ldots \bigsqcup \Omega_K$$

Segmentation d'image



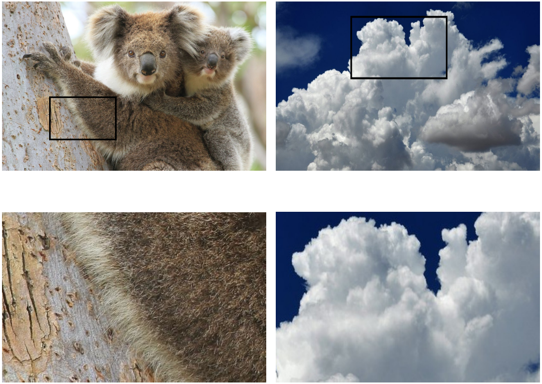


Objectif : obtenir une partition de l'image en K régions **homogènes**

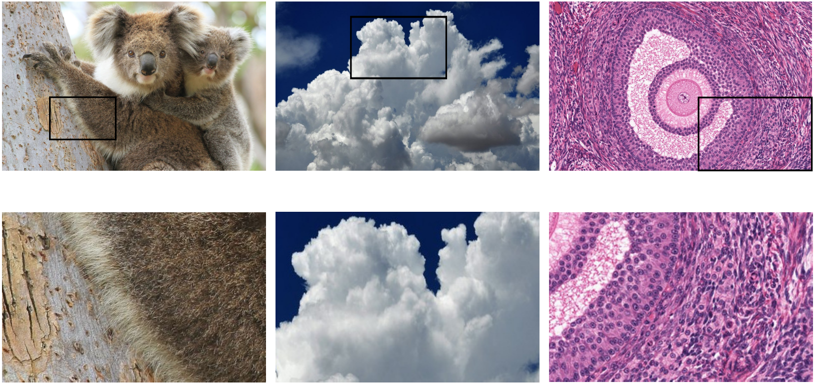
$$\Omega = \Omega_1 \bigsqcup \ldots \bigsqcup \Omega_K$$

Textures

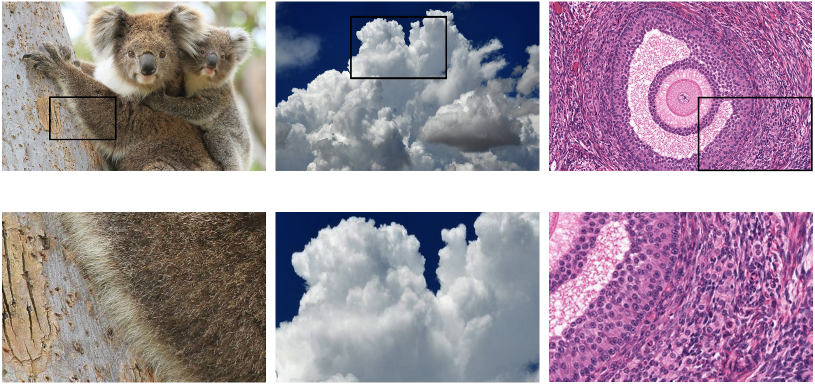
Textures



Textures



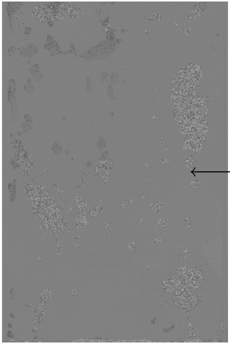
Textures



Crucial pour décrire les images réelles

Écoulement multiphasiques en milieu poreux

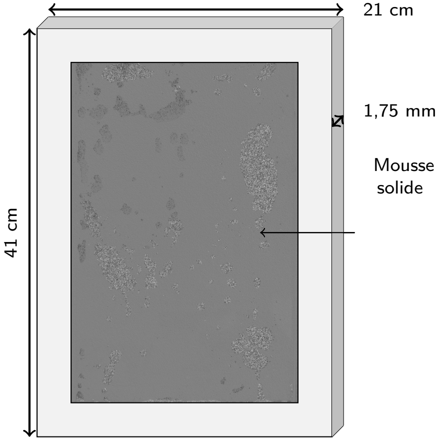
Laboratoire de Physique, ENS Lyon, V. Vidal, T. Busser, (M. Serres, IFPEN)



Mousse solide

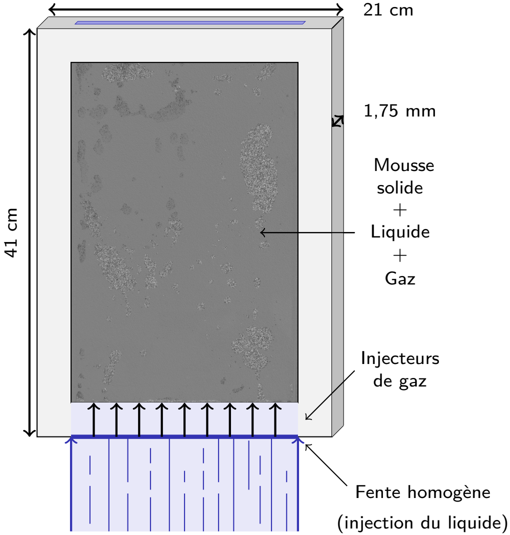
Écoulement multiphasiques en milieu poreux

Laboratoire de Physique, ENS Lyon, V. Vidal, T. Busser, (M. Serres, IFPEN)



Écoulement multiphasiques en milieu poreux

Laboratoire de Physique, ENS Lyon, V. Vidal, T. Busser, (M. Serres, IFPEN)



- → 1600 × 1100 pixels
- → vidéo : ∼ 1000 images
- → plan de phases : ∼ 10 débits

Plan de l'exposé

1. Caractérisation de textures

[Filtres de Gabor (Dunn, 1995)]
[Amplitude et fréquence locales (Havlicek, 1996)]
[Histogrammes spectraux (Yuan, 2015)]

Plan de l'exposé

1. Caractérisation de textures

$\rightarrow$ attributs fractals

- variance locale σ^2
- régularité locale h

[Filtres de Gabor (Dunn, 1995)]
[Amplitude et fréquence locales (Havlicek, 1996)]
[Histogrammes spectraux (Yuan, 2015)]

Plan de l'exposé

1. Caractérisation de textures

$\rightarrow$ attributs fractals

- variance locale σ^2
- régularité locale h

[Filtres de Gabor (Dunn, 1995)]
[Amplitude et fréquence locales (Havlicek, 1996)]
[Histogrammes spectraux (Yuan, 2015)]

2. Construction de fonctionnelles

[Champ de Markov (Geman, 1984)]
[Contours actifs (Chan, 2001)]
[Variation Totale et Seuillage (Cai, 2013)]

Plan de l'exposé

1. Caractérisation de textures

→ attributs fractals
- variance locale σ^2
- régularité locale h

[Filtres de Gabor (Dunn, 1995)]
[Amplitude et fréquence locales (Havlicek, 1996)]
[Histogrammes spectraux (Yuan, 2015)]

2. Construction de fonctionnelles

→ moindres carrés pénalisés
- contours libres
- contours co-localisés

[Champ de Markov (Geman, 1984)]
[Contours actifs (Chan, 2001)]
[Variation Totale et Seuillage (Cai, 2013)]

Plan de l'exposé

1. Caractérisation de textures

→ attributs fractals
- variance locale σ^2
- régularité locale h

[Filtres de Gabor (Dunn, 1995)]
[Amplitude et fréquence locales (Havlicek, 1996)]
[Histogrammes spectraux (Yuan, 2015)]

2. Construction de fonctionnelles

→ moindres carrés pénalisés
- contours libres
- contours co-localisés

[Champ de Markov (Geman, 1984)]
[Contours actifs (Chan, 2001)]
[Variation Totale et Seuillage (Cai, 2013)]

3. Algorithme de minimisation accéléré

[Forward-backward (Combettes, 2005)]
[FISTA (Beck, 2009)]
[Primal-dual (Chambolle, 2011)]

Plan de l'exposé

1. Caractérisation de textures

→ attributs fractals

- variance locale σ^2
- régularité locale h

[Filtres de Gabor (Dunn, 1995)]
[Amplitude et fréquence locales (Havlicek, 1996)]
[Histogrammes spectraux (Yuan, 2015)]

2. Construction de fonctionnelles

→ moindres carrés pénalisés

- contours libres
- contours co-localisés

[Champ de Markov (Geman, 1984)]
[Contours actifs (Chan, 2001)]
[Variation Totale et Seuillage (Cai, 2013)]

3. Algorithme de minimisation accéléré

→ algorithmes proximaux scindés

- calcul des opérateurs proximaux
- accélération par forte-convexité

[Forward-backward (Combettes, 2005)]
[FISTA (Beck, 2009)]
[Primal-dual (Chambolle, 2011)]

Plan de l'exposé

1. Caractérisation de textures

→ attributs fractals

- variance locale σ^2
- régularité locale h

[Filtres de Gabor (Dunn, 1995)]
[Amplitude et fréquence locales (Havlicek, 1996)]
[Histogrammes spectraux (Yuan, 2015)]

2. Construction de fonctionnelles

→ moindres carrés pénalisés

- contours libres
- contours co-localisés

[Champ de Markov (Geman, 1984)]
[Contours actifs (Chan, 2001)]
[Variation Totale et Seuillage (Cai, 2013)]

3. Algorithme de minimisation accéléré

→ algorithmes proximaux scindés

- calcul des opérateurs proximaux
- accélération par forte-convexité

[Forward-backward (Combettes, 2005)]
[FISTA (Beck, 2009)]
[Primal-dual (Chambolle, 2011)]

4. Réglage des hyperparamètres

[SURE (Stein, 1981)]
[SURE DFMC (Ramani, 2008)]
[GSURE (Eldar, 2008)]
[SUGAR (Deledalle, 2014)]

Plan de l'exposé

1. Caractérisation de textures

→ attributs fractals
- variance locale σ^2
- régularité locale h

[Filtres de Gabor (Dunn, 1995)]
[Amplitude et fréquence locales (Havlicek, 1996)]
[Histogrammes spectraux (Yuan, 2015)]

2. Construction de fonctionnelles

→ moindres carrés pénalisés
- contours libres
- contours co-localisés

[Champ de Markov (Geman, 1984)]
[Contours actifs (Chan, 2001)]
[Variation Totale et Seuillage (Cai, 2013)]

3. Algorithme de minimisation accéléré

→ algorithmes proximaux scindés
- calcul des opérateurs proximaux
- accélération par forte-convexité

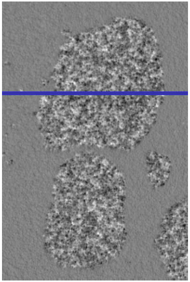
[Forward-backward (Combettes, 2005)]
[FISTA (Beck, 2009)]
[Primal-dual (Chambolle, 2011)]

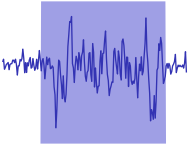
4. Réglage des hyperparamètres

→ SURE avec bruit gaussien corrélé
- erreur d'estimation projetée
- minimisation par quasi-Newton
  ↪ SUGAR généralisé

[SURE (Stein, 1981)]
[SURE DFMC (Ramani, 2008)]
[GSURE (Eldar, 2008)]
[SUGAR (Deledalle, 2014)]

Modèle monofractal par morceaux

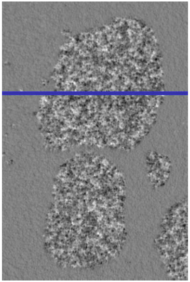


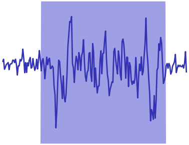


Modèle monofractal par morceaux

Attributs fractals

- variance σ^2 *amplitude des variations*

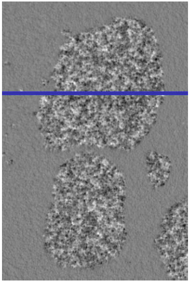


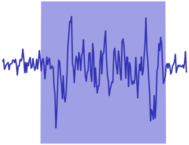


Modèle monofractal par morceaux

Attributs fractals

- variance σ^2 — *amplitude des variations*
- régularité locale h — *invariance d'échelle*



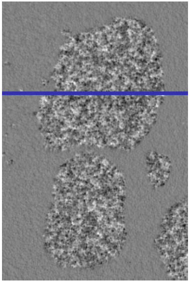


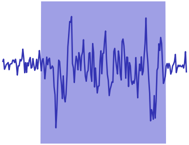
Modèle monofractal par morceaux

Attributs fractals

- variance σ^2 *amplitude des variations*
- régularité locale h *invariance d'échelle*

$$|f(x) - f(y)| \leq \sigma(x)|x - y|^{h(x)}$$



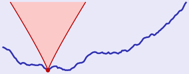


Modèle monofractal par morceaux

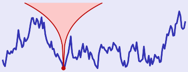
Attributs fractals

- variance σ^2 — *amplitude des variations*
- régularité locale h — *invariance d'échelle*

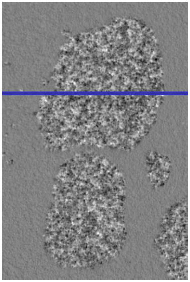
$$|f(x) - f(y)| \leq \sigma(x)|x - y|^{h(x)}$$

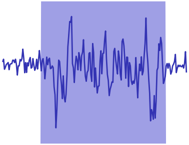


$h(x) \equiv h_1 = 0.9$



$h(x) \equiv h_2 = 0.3$



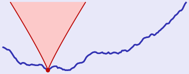


Modèle monofractal par morceaux

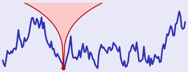
Attributs fractals

- variance σ^2 *amplitude des variations*
- régularité locale h *invariance d'échelle*

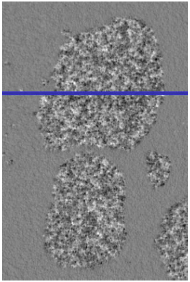
$$|f(x) - f(y)| \leq \sigma(x)|x - y|^{h(x)}$$

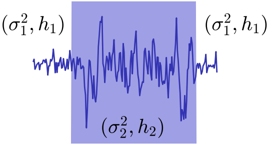


$h(x) \equiv h_1 = 0.9$



$h(x) \equiv h_2 = 0.3$





Segmentation

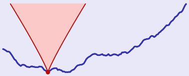
- h et σ^2 constants par morceaux

Modèle monofractal par morceaux

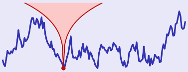
Attributs fractals

- variance σ^2 — *amplitude des variations*
- régularité locale h — *invariance d'échelle*

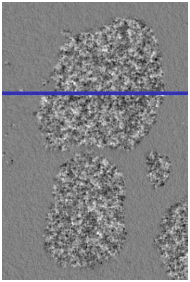
$$|f(x) - f(y)| \leq \sigma(x)|x-y|^{h(x)}$$

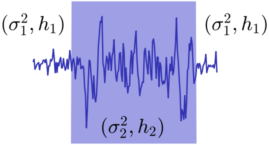


$h(x) \equiv h_1 = 0.9$



$h(x) \equiv h_2 = 0.3$



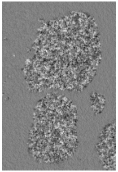


Segmentation

- h et σ^2 constants par morceaux
- région Ω_k caractérisée par (h_k, σ_k^2)

Analyse multi-échelle

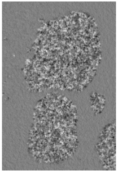
Image texturée

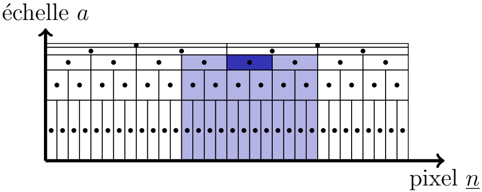


Analyse multi-échelle

Image texturée

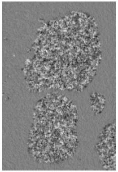
Maximum local des coefficients d'ondelettes : $\mathcal{L}_{a,\cdot}$





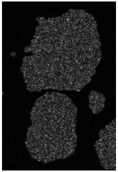
Analyse multi-échelle

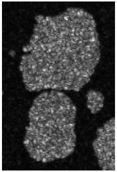
Image texturée

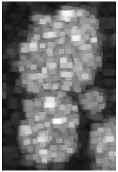


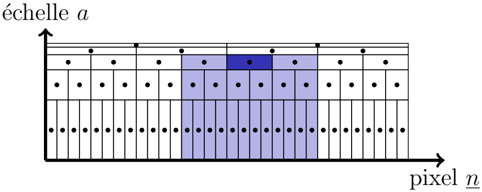
Maximum local des coefficients d'ondelettes : $\mathcal{L}_{a,\cdot}$

Échelle $a = 2^1$ $a = 2^2$ … $a = 2^5$



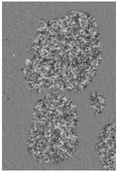






Analyse multi-échelle

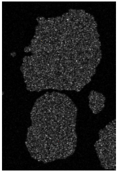
Image texturée



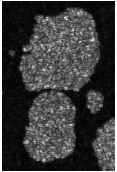
Maximum local des coefficients d'ondelettes : $\mathcal{L}_{a,\cdot}$

Échelle

$a = 2^1$

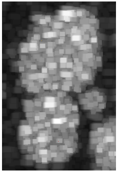


$a = 2^2$



$\ldots$

$a = 2^5$

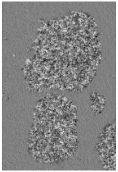


Proposition *(Jaffard, 2004), (Wendt, 2008)*

$$\log\left(\mathcal{L}_{a,\cdot}\right) \underset{a\to 0}{\simeq} \log(a) \underbrace{\boldsymbol{h}}_{\text{régularité}} + \underbrace{\boldsymbol{v}}_{\substack{\propto \log(\boldsymbol{\sigma}^2) \\ \text{(variance)}}}$$

Analyse multi-échelle

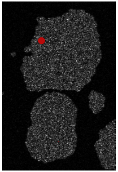
Image texturée



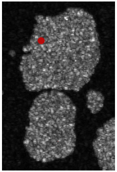
Maximum local des coefficients d'ondelettes : $\mathcal{L}_{a,\cdot}$

Échelle

$a = 2^1$

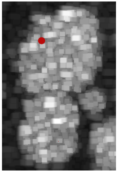


$a = 2^2$



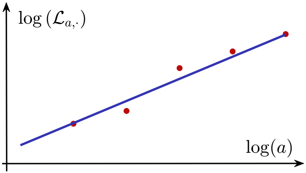
$\ldots$

$a = 2^5$



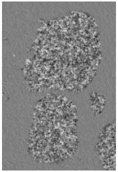
Proposition *(Jaffard, 2004), (Wendt, 2008)*

$$\log\left(\mathcal{L}_{a,\cdot}\right) \underset{a\to 0}{\simeq} \log(a) \underbrace{\boldsymbol{h}}_{\text{régularité}} + \underbrace{\boldsymbol{v}}_{\substack{\propto \log(\boldsymbol{\sigma}^2)\\ \text{(variance)}}}$$



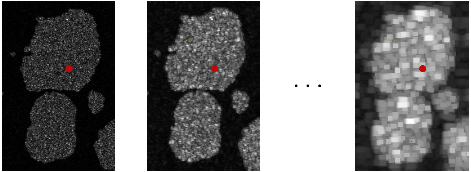
Analyse multi-échelle

Image texturée



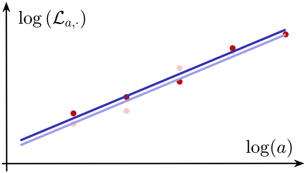
Maximum local des coefficients d'ondelettes : $\mathcal{L}_{a,\cdot}$

Échelle $a = 2^1$ $a = 2^2$ … $a = 2^5$



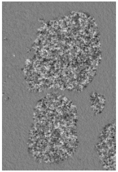
Proposition *(Jaffard, 2004), (Wendt, 2008)*

$$\log\left(\mathcal{L}_{a,\cdot}\right) \underset{a\to 0}{\simeq} \log(a) \underbrace{\boldsymbol{h}}_{\text{régularité}} + \underbrace{\boldsymbol{v}}_{\substack{\propto \log(\boldsymbol{\sigma}^2) \\ \text{(variance)}}}$$



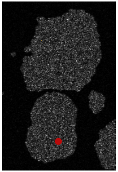
Analyse multi-échelle

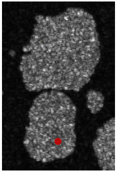
Image texturée

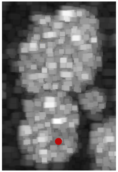


Maximum local des coefficients d'ondelettes : $\mathcal{L}_{a,\cdot}$

Échelle $a = 2^1$ $a = 2^2$ … $a = 2^5$

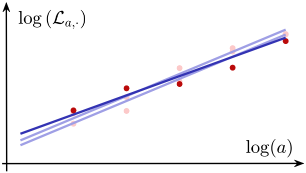






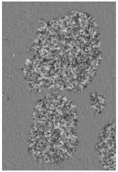
Proposition *(Jaffard, 2004), (Wendt, 2008)*

$$\log\left(\mathcal{L}_{a,\cdot}\right) \underset{a\to 0}{\simeq} \log(a) \underbrace{\boldsymbol{h}}_{\text{régularité}} + \underbrace{\boldsymbol{v}}_{\substack{\propto \log(\boldsymbol{\sigma}^2) \\ \text{(variance)}}}$$



Analyse multi-échelle

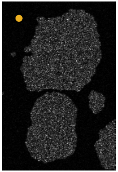
Image texturée



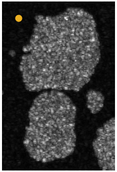
Maximum local des coefficients d'ondelettes : $\mathcal{L}_{a,\cdot}$

Échelle

$a = 2^1$

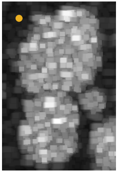


$a = 2^2$



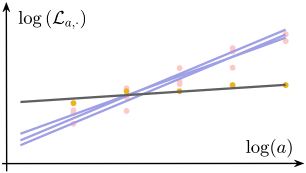
$\cdots$

$a = 2^5$



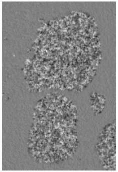
Proposition *(Jaffard, 2004), (Wendt, 2008)*

$$\log\left(\mathcal{L}_{a,\cdot}\right) \underset{a\to 0}{\simeq} \log(a) \underbrace{\boldsymbol{h}}_{\text{régularité}} + \underbrace{\boldsymbol{v}}_{\substack{\propto \log(\boldsymbol{\sigma}^2) \\ \text{(variance)}}}$$



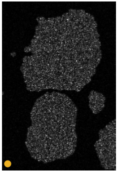
Analyse multi-échelle

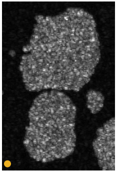
Image texturée

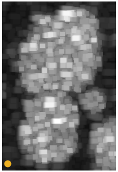


Maximum local des coefficients d'ondelettes : $\mathcal{L}_{a,\cdot}$

Échelle $a = 2^1$ $a = 2^2$ ... $a = 2^5$

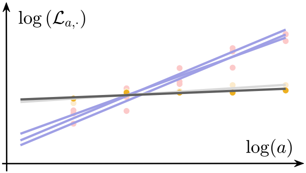






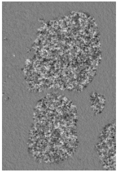
Proposition *(Jaffard, 2004), (Wendt, 2008)*

$$\log\left(\mathcal{L}_{a,\cdot}\right) \underset{a\to 0}{\simeq} \log(a) \underbrace{\boldsymbol{h}}_{\text{régularité}} + \underbrace{\boldsymbol{v}}_{\substack{\propto \log(\boldsymbol{\sigma}^2)\\ \text{(variance)}}}$$



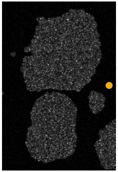
Analyse multi-échelle

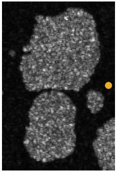
Image texturée

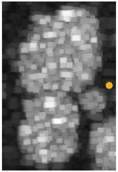


Maximum local des coefficients d'ondelettes : $\mathcal{L}_{a,\cdot}$

Échelle $a = 2^1$ $a = 2^2$ … $a = 2^5$

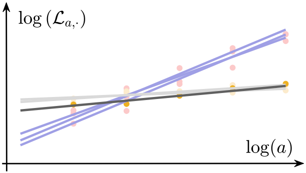






Proposition *(Jaffard, 2004), (Wendt, 2008)*

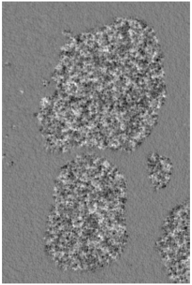
$$\log\left(\mathcal{L}_{a,\cdot}\right) \underset{a\to 0}{\simeq} \log(a) \underbrace{\boldsymbol{h}}_{\text{régularité}} + \underbrace{\boldsymbol{v}}_{\substack{\propto \log(\boldsymbol{\sigma}^2) \\ \text{(variance)}}}$$



Estimation directe ponctuelle

Régression linéaire $\log\left(\mathcal{L}_{a,\cdot}\right) \simeq \log(a) \underbrace{\boldsymbol{h}}_{\text{régularité}} + \underbrace{\boldsymbol{v}}_{\propto \log(\boldsymbol{\sigma}^2)}$

Image texturée

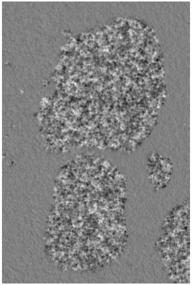


Estimation directe ponctuelle

Régression linéaire $\log\left(\mathcal{L}_{a,\cdot}\right) \simeq \log(a) \underbrace{\boldsymbol{h}}_{\text{régularité}} + \underbrace{\boldsymbol{v}}_{\propto \log(\boldsymbol{\sigma}^2)}$

$$\left(\widehat{\boldsymbol{h}}^{\mathrm{RL}}, \widehat{\boldsymbol{v}}^{\mathrm{RL}}\right) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a=a_{\min}}^{a_{\max}} \left\|\log\left(\mathcal{L}_{a,\cdot}\right) - \log(a)\boldsymbol{h} - \boldsymbol{v}\right\|^2$$

Image texturée

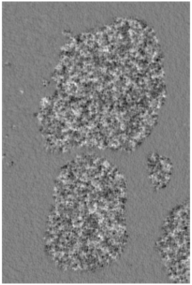


Estimation directe ponctuelle

Régression linéaire $\log(\boldsymbol{\mathcal{L}}_{a,\cdot}) \simeq \log(a) \underbrace{\boldsymbol{h}}_{\text{régularité}} + \underbrace{\boldsymbol{v}}_{\propto \log(\boldsymbol{\sigma}^2)}$

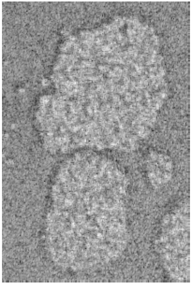
$$\left(\widehat{\boldsymbol{h}}^{\mathrm{RL}}, \widehat{\boldsymbol{v}}^{\mathrm{RL}}\right) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a=a_{\min}}^{a_{\max}} \left\| \log(\boldsymbol{\mathcal{L}}_{a,\cdot}) - \log(a)\boldsymbol{h} - \boldsymbol{v} \right\|^2$$

Image texturée



Régularité locale $\widehat{\boldsymbol{h}}^{\mathrm{RL}}$

Puissance locale $\widehat{\boldsymbol{v}}^{\mathrm{RL}}$

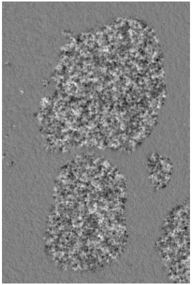


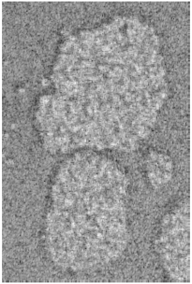
Estimation directe ponctuelle

Régression linéaire $\underbrace{\mathbb{E}\log\left(\mathcal{L}_{a,\cdot}\right)}_{\text{espérance}} = \log(a)\underbrace{\bar{\boldsymbol{h}}}_{\text{régularité}} + \underbrace{\bar{\boldsymbol{v}}}_{\propto\log(\boldsymbol{\sigma}^2)}$

$$\left(\widehat{\boldsymbol{h}}^{\mathrm{RL}},\widehat{\boldsymbol{v}}^{\mathrm{RL}}\right) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a=a_{\min}}^{a_{\max}} \left\|\log\left(\mathcal{L}_{a,\cdot}\right) - \log(a)\boldsymbol{h} - \boldsymbol{v}\right\|^2$$

Image texturée | Régularité locale $\widehat{\boldsymbol{h}}^{\mathrm{RL}}$ | Puissance locale $\widehat{\boldsymbol{v}}^{\mathrm{RL}}$





$\longrightarrow$ variance d'estimation élevée

Régularisation *a posteriori*

Régression linéaire $\widehat{\boldsymbol{h}}^{\mathrm{RL}}$

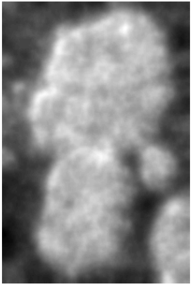
Régularisation *a posteriori*

Lissage par filtrage (linéaire)

$$\left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \widehat{\boldsymbol{h}}^{\mathrm{RL}}$$

Régression linéaire $\widehat{\boldsymbol{h}}^{\mathrm{RL}}$

Lissage



Régularisation *a posteriori*

Lissage par filtrage (linéaire)

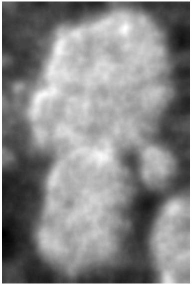
$$\left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \widehat{\boldsymbol{h}}^{\mathrm{RL}}$$

Débruitage ROF (non linéaire)

$$\underset{\boldsymbol{h}}{\operatorname{argmin}} \; \|\boldsymbol{h} - \widehat{\boldsymbol{h}}^{\mathrm{RL}}\|^2 + \lambda \|\mathbf{D}\boldsymbol{h}\|_{2,1}$$

Régression linéaire $\widehat{\boldsymbol{h}}^{\mathrm{RL}}$

Lissage



ROF

Régularisation *a posteriori*

Lissage par filtrage (linéaire)

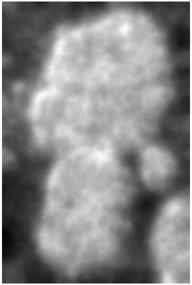
$$\left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \widehat{\boldsymbol{h}}^{\mathrm{RL}}$$

Débruitage ROF (non linéaire)

$$\underset{\boldsymbol{h}}{\operatorname{argmin}} \ \|\boldsymbol{h} - \widehat{\boldsymbol{h}}^{\mathrm{RL}}\|^2 + \lambda \|\mathbf{D}\boldsymbol{h}\|_{2,1}$$

Régression linéaire $\widehat{\boldsymbol{h}}^{\mathrm{RL}}$

Lissage



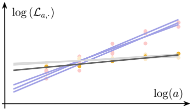
ROF

$\longrightarrow$ cumul de la variance d'estimation et du biais de régularisation

Fonctionnelles à contours libres et à contours co-localisés

$$\sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,.} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}}$$

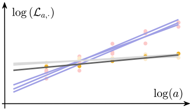
$\rightarrow$ fidélité au modèle log-linéaire



Fonctionnelles à contours libres et à contours co-localisés

$$\sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)}_{\textbf{Variation Totale}}$$

$\rightarrow$ fidélité au modèle log-linéaire



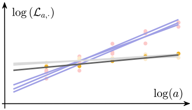
$\rightarrow$ favorise la constance par morceaux

Fonctionnelles à contours libres et à contours co-localisés

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \quad \underbrace{\sum_a \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)}_{\textbf{Variation Totale}}$$

$\rightarrow$ fidélité au modèle log-linéaire

$\rightarrow$ favorise la constance par morceaux

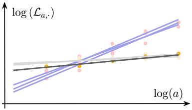


Fonctionnelles à contours libres et à contours co-localisés

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \quad \underbrace{\sum_a \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \quad + \quad \lambda \, \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Variation Totale}}$$

$\rightarrow$ fidélité au modèle log-linéaire

$\rightarrow$ favorise la constance par morceaux

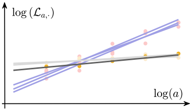


Différences finies $\mathbf{D}_1\boldsymbol{x}$ (horizontales), $\mathbf{D}_2\boldsymbol{x}$ (verticales) en chaque pixel

Fonctionnelles à contours libres et à contours co-localisés

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} + \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)}_{\textbf{Variation Totale}}$$

$\rightarrow$ fidélité au modèle log-linéaire



$\rightarrow$ favorise la constance par morceaux

Différences finies $\mathbf{D}\boldsymbol{x} = [\mathbf{D}_1\boldsymbol{x},\mathbf{D}_2\boldsymbol{x}]$

libres : $\boldsymbol{h}$, $\boldsymbol{v}$ sont **indépendamment** constantes par morceaux

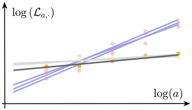
$$\mathcal{Q}_{\mathrm{L}}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha) = \alpha\|\mathbf{D}\boldsymbol{h}\|_{2,1} + \|\mathbf{D}\boldsymbol{v}\|_{2,1}$$

Fonctionnelles à contours libres et à contours co-localisés

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \quad \underbrace{\sum_a \|\log \mathcal{L}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)}_{\textbf{Variation Totale}}$$

→ fidélité au modèle log-linéaire

→ favorise la constance par morceaux



Différences finies $\mathbf{D}\boldsymbol{x} = [\mathbf{D}_1\boldsymbol{x},\mathbf{D}_2\boldsymbol{x}]$

<u>libres :</u> $\boldsymbol{h}$, $\boldsymbol{v}$ sont **indépendamment** constantes par morceaux

$$\mathcal{Q}_{\mathrm{L}}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha) = \alpha\|\mathbf{D}\boldsymbol{h}\|_{2,1} + \|\mathbf{D}\boldsymbol{v}\|_{2,1}$$

<u>co-localisés :</u> $\boldsymbol{h}$, $\boldsymbol{v}$ sont **concomitamment** constantes par morceaux

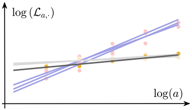
$$\mathcal{Q}_{\mathrm{C}}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha) = \|[\alpha\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v}]\|_{2,1}$$

Fonctionnelles à contours libres et à contours co-localisés

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \underbrace{\sum_a \|\log \mathcal{L}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} + \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)}_{\textbf{Variation Totale}}$$

→ fidélité au modèle log-linéaire

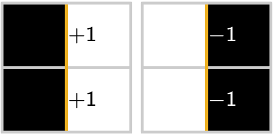
→ favorise la constance par morceaux



Contours disjoints

$\boldsymbol{h} \in \mathbb{R}^{2\times 2}$ $\boldsymbol{v} \in \mathbb{R}^{2\times 2}$

Contours communs



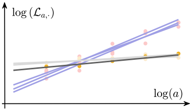
$\boldsymbol{h} \in \mathbb{R}^{2\times 2}$ $\boldsymbol{v} \in \mathbb{R}^{2\times 2}$

Fonctionnelles à contours libres et à contours co-localisés

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} + \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)}_{\textbf{Variation Totale}}$$

$\rightarrow$ fidélité au modèle log-linéaire

$\rightarrow$ favorise la constance par morceaux



Contours disjoints

$\boldsymbol{h} \in \mathbb{R}^{2\times 2}$ $\quad$ $\boldsymbol{v} \in \mathbb{R}^{2\times 2}$

$\mathcal{Q}_L(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};1) = 4$

Contours communs

$\boldsymbol{h} \in \mathbb{R}^{2\times 2}$ $\quad$ $\boldsymbol{v} \in \mathbb{R}^{2\times 2}$

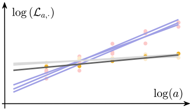
$\mathcal{Q}_L(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};1) = 4$

Fonctionnelles à contours libres et à contours co-localisés

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \quad \underbrace{\sum_a \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Variation Totale}}$$

→ fidélité au modèle log-linéaire

→ favorise la constance par morceaux



Contours disjoints

$\boldsymbol{h} \in \mathbb{R}^{2\times 2}$ $\quad$ $\boldsymbol{v} \in \mathbb{R}^{2\times 2}$

$\mathcal{Q}_L(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; 1) = 4$

$\mathcal{Q}_C(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; 1) = 2 + \sqrt{2} \simeq 3{,}4$

Contours communs

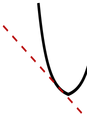
$\boldsymbol{h} \in \mathbb{R}^{2\times 2}$ $\quad$ $\boldsymbol{v} \in \mathbb{R}^{2\times 2}$

$\mathcal{Q}_L(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; 1) = 4$

$\mathcal{Q}_C(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; 1) = 2\sqrt{2} \simeq 2{,}8$

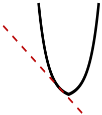
Minimisation de fonctionnelle

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \; \underbrace{\sum_a \|\log \boldsymbol{\mathcal{L}}_{a,.} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \; + \; \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)}_{\textbf{Variation Totale}}$$



Minimisation de fonctionnelle

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \ \sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \ + \ \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)}_{\textbf{Variation Totale}}$$

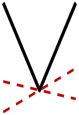


- descente de gradient $\boldsymbol{x}^{n+1} = \boldsymbol{x}^n - \tau\nabla\varphi(\boldsymbol{x}^n)$

Minimisation de fonctionnelle

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \; \underbrace{\sum_a \|\log \mathcal{L}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \; + \; \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Variation Totale}}$$

non lisse



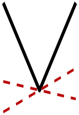
- descente de gradient $\boldsymbol{x}^{n+1} = \boldsymbol{x}^n - \tau \nabla \varphi(\boldsymbol{x}^n)$
- descente de sous-gradient implicite : algorithme du point proximal

$$\boldsymbol{x}^{n+1} = \boldsymbol{x}^n - \boldsymbol{u}^n,\; \boldsymbol{u}^n \in \partial\varphi(\boldsymbol{x}^{n+1}) \;\Leftrightarrow\; \boldsymbol{x}^{n+1} = \mathrm{prox}_{\tau\varphi}(\boldsymbol{x}^n)$$

Minimisation de fonctionnelle

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \; \underbrace{\sum_a \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \; + \; \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Variation Totale}}$$

non lisse



- descente de gradient $\boldsymbol{x}^{n+1} = \boldsymbol{x}^n - \tau \nabla \varphi(\boldsymbol{x}^n)$
- descente de sous-gradient implicite : algorithme du point proximal

$$\boldsymbol{x}^{n+1} = \boldsymbol{x}^n - \boldsymbol{u}^n, \; \boldsymbol{u}^n \in \partial\varphi(\boldsymbol{x}^{n+1}) \;\Leftrightarrow\; \boldsymbol{x}^{n+1} = \operatorname{prox}_{\tau\varphi}(\boldsymbol{x}^n)$$

- algorithme proximal scindé

$$\boldsymbol{y}^{n+1} = \operatorname{prox}_{\sigma(\lambda\mathcal{Q})^*}\left(\boldsymbol{y}^n + \sigma \mathbf{D}\bar{\boldsymbol{x}}^n\right)$$

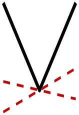
$$\boldsymbol{x}^{n+1} = \operatorname{prox}_{\tau\|\boldsymbol{\mathcal{L}} - \boldsymbol{\Phi}\cdot\|_2^2}\left(\boldsymbol{x}^n - \tau \mathbf{D}^\top \boldsymbol{y}^{n+1}\right), \qquad \boldsymbol{\Phi} : (\boldsymbol{h}, \boldsymbol{v}) \mapsto \{\log(a)\boldsymbol{h} + \boldsymbol{v}\}_a$$

$$\bar{\boldsymbol{x}}^{n+1} = 2\boldsymbol{x}^{n+1} - \boldsymbol{x}^n$$

Minimisation de fonctionnelle

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \quad \underbrace{\sum_a \|\log \mathcal{L}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)}_{\textbf{Variation Totale}}$$

non lisse



- descente de gradient $\boldsymbol{x}^{n+1} = \boldsymbol{x}^n - \tau\nabla\varphi(\boldsymbol{x}^n)$
- descente de sous-gradient implicite : algorithme du point proximal

$$\boldsymbol{x}^{n+1} = \boldsymbol{x}^n - \boldsymbol{u}^n,\ \boldsymbol{u}^n \in \partial\varphi(\boldsymbol{x}^{n+1}) \ \Leftrightarrow\ \boldsymbol{x}^{n+1} = \operatorname{prox}_{\tau\varphi}(\boldsymbol{x}^n)$$

- algorithme proximal scindé $\quad \operatorname{prox}_{\tau\varphi}(\boldsymbol{x}) = \underset{\boldsymbol{u}}{\operatorname{argmin}} \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{u}\|^2 + \tau\varphi(\boldsymbol{u})$

$$\boldsymbol{y}^{n+1} = \operatorname{prox}_{\sigma(\lambda\mathcal{Q})^*}\left(\boldsymbol{y}^n + \sigma\mathbf{D}\bar{\boldsymbol{x}}^n\right)$$

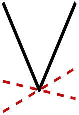
$$\boldsymbol{x}^{n+1} = \operatorname{prox}_{\tau\|\mathcal{L}-\boldsymbol{\Phi}\cdot\|_2^2}\left(\boldsymbol{x}^n - \tau\mathbf{D}^\top\boldsymbol{y}^{n+1}\right), \qquad \boldsymbol{\Phi} : (\boldsymbol{h},\boldsymbol{v}) \mapsto \{\log(a)\boldsymbol{h} + \boldsymbol{v}\}_a$$

$$\bar{\boldsymbol{x}}^{n+1} = 2\boldsymbol{x}^{n+1} - \boldsymbol{x}^n$$

Calcul des opérateurs proximaux

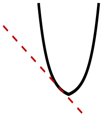
$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \sum_a \underset{\textbf{Moindres Carrés}}{\underline{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}} \quad + \quad \lambda \underset{\textbf{Variation Totale}}{\underline{\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)}}$$

non lisse

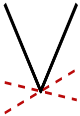


Calcul des opérateurs proximaux

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \; \sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \; + \; \lambda \; \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)}_{\textbf{Variation Totale}}$$



non lisse



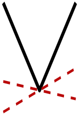
Ex. Norme mixte : pour $\boldsymbol{z} = [\boldsymbol{z}_1; \ldots,; \boldsymbol{z}_I]$

$$\mathcal{Q}(\boldsymbol{z}) = \|\boldsymbol{z}\|_{2,1} = \sum_{\underline{n}\in\Omega} \sqrt{\sum_{i=1}^{I} z_i^2(\underline{n})} = \sum_{\underline{n}\in\Omega} \|\boldsymbol{z}(\underline{n})\|_2$$

Calcul des opérateurs proximaux

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \quad \underbrace{\sum_a \|\log \boldsymbol{\mathcal{L}}_{a,.} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)}_{\textbf{Variation Totale}}$$

non lisse



Ex. Norme mixte : pour $\boldsymbol{z} = [\boldsymbol{z}_1; \ldots,; \boldsymbol{z}_I]$

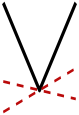
$$\mathcal{Q}(\boldsymbol{z}) = \|\boldsymbol{z}\|_{2,1} = \sum_{\underline{n}\in\Omega} \sqrt{\sum_{i=1}^{I} z_i^2(\underline{n})} = \sum_{\underline{n}\in\Omega} \|\boldsymbol{z}(\underline{n})\|_2$$

$$\boldsymbol{p} = \mathrm{prox}_{\lambda\|\cdot\|_{2,1}}(\boldsymbol{z}) \quad \Leftrightarrow \quad p_i(\underline{n}) = \max\left(0,1 - \frac{\lambda}{\|\boldsymbol{z}(\underline{n})\|_2}\right) z_i(\underline{n})$$

Calcul des opérateurs proximaux

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \quad + \quad \lambda \, \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)}_{\textbf{Variation Totale}}$$

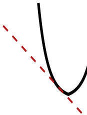
non lisse



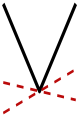
Moindres carrés : $\|\log \boldsymbol{\mathcal{L}} - \Phi(\boldsymbol{h},\boldsymbol{v})\|^2, \quad \Phi : (\boldsymbol{h},\boldsymbol{v}) \mapsto \{\log(a)\boldsymbol{h} + \boldsymbol{v}\}_a$

Calcul des opérateurs proximaux

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \quad \sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Variation Totale}}$$



non lisse



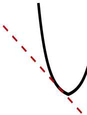
Moindres carrés : $\|\log \boldsymbol{\mathcal{L}} - \boldsymbol{\Phi}(\boldsymbol{h},\boldsymbol{v})\|^2, \quad \boldsymbol{\Phi} : (\boldsymbol{h},\boldsymbol{v}) \mapsto \{\log(a)\boldsymbol{h} + \boldsymbol{v}\}_a$

Proposition *(Pascal, 2019)*

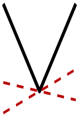
$$(\widetilde{\boldsymbol{h}},\widetilde{\boldsymbol{v}}) = \mathrm{prox}_{\tau\|\boldsymbol{\mathcal{L}}-\boldsymbol{\Phi}\cdot\|^2}(\boldsymbol{h},\boldsymbol{v}) \Longleftrightarrow (\widetilde{\boldsymbol{h}},\widetilde{\boldsymbol{v}}) = \left(\mathbf{I} + \tau\boldsymbol{\Phi}^\top\boldsymbol{\Phi}\right)^{-1}\left((\boldsymbol{h},\boldsymbol{v}) + \tau\boldsymbol{\Phi}^\top \log \boldsymbol{\mathcal{L}}\right)$$

Calcul des opérateurs proximaux

$$\operatorname*{minimiser}_{\boldsymbol{h},\boldsymbol{v}} \sum_a \underbrace{\|\log \mathcal{L}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)}_{\textbf{Variation Totale}}$$



non lisse



Moindres carrés : $\|\log \mathcal{L} - \Phi(\boldsymbol{h},\boldsymbol{v})\|^2, \quad \Phi : (\boldsymbol{h},\boldsymbol{v}) \mapsto \{\log(a)\boldsymbol{h} + \boldsymbol{v}\}_a$

Proposition *(Pascal, 2019)*

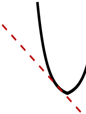
Soit $S_m = \sum_a \log^m(a), \;\; \mathcal{D} = (1+\tau S_2)(1+\tau S_0) - \tau^2 S_1^2,$
$\boldsymbol{\mathcal{T}} = \sum_a \log \mathcal{L}_a$ et $\boldsymbol{\mathcal{G}} = \sum_a \log(a) \log \mathcal{L}_a$, alors

$$(\widetilde{\boldsymbol{h}},\widetilde{\boldsymbol{v}}) = \mathrm{prox}_{\tau\|\mathcal{L}-\Phi\cdot\|^2}(\boldsymbol{h},\boldsymbol{v}) \Longleftrightarrow (\widetilde{\boldsymbol{h}},\widetilde{\boldsymbol{v}}) = \left(\mathbf{I} + \tau \Phi^\top \Phi\right)^{-1} \left((\boldsymbol{h},\boldsymbol{v}) + \tau \Phi^\top \log \mathcal{L}\right)$$

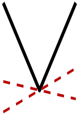
$$\Longleftrightarrow \begin{cases} \widetilde{\boldsymbol{h}} = \mathcal{D}^{-1}\left((1+\tau S_0)(\tau \boldsymbol{\mathcal{G}} + \boldsymbol{h}) - \tau S_1(\tau \boldsymbol{\mathcal{T}} + \boldsymbol{v})\right) \\ \widetilde{\boldsymbol{v}} = \mathcal{D}^{-1}\left((1+\tau S_2)(\tau \boldsymbol{\mathcal{T}} + \boldsymbol{v}) - \tau S_1(\tau \boldsymbol{\mathcal{G}} + \boldsymbol{h})\right) \end{cases}$$

Algorithme accéléré par forte-convexité

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} + \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)}_{\textbf{Variation Totale}}$$



non lisse

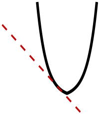


Algorithme primal-dual *(Chambolle, 2011)*

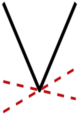
$$\delta \text{ : gap de dualité, } \delta(\boldsymbol{x}^n,\boldsymbol{y}^n) \underset{n\to\infty}{\longrightarrow} 0$$

Algorithme accéléré par forte-convexité

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \ \underbrace{\sum_a \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} + \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)}_{\textbf{Variation Totale}}$$

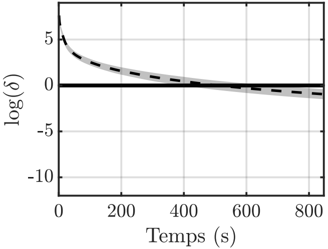


non lisse



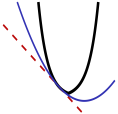
Algorithme primal-dual *(Chambolle, 2011)*

δ : gap de dualité, $\delta(\boldsymbol{x}^n,\boldsymbol{y}^n) \underset{n\to\infty}{\longrightarrow} 0$



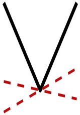
Propriétés de convexité

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \sum_a \underbrace{\left\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\right\|^2}_{\textbf{Moindres Carrés}} + \lambda\, \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)}_{\textbf{Variation Totale}}$$



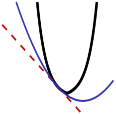
$\boldsymbol{\mu}$-fortement convexe

non lisse



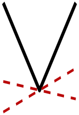
Propriétés de convexité

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \quad \underbrace{\sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \quad + \quad \lambda \, \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Variation Totale}}$$



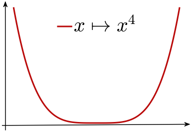
$\boldsymbol{\mu}$-fortement convexe

non lisse

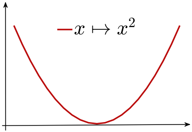


Forte convexité

- φ $\boldsymbol{\mu}$-fortement convexe ssi $\varphi - \dfrac{\boldsymbol{\mu}}{2}\|\cdot\|^2$ convexe



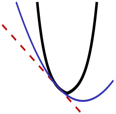
✓ strictement convexe
✗ non fortement convexe



✓ strictement convexe
✓ 1-fortement convexe

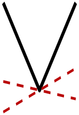
Propriétés de convexité

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \; \underbrace{\sum_a \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \; + \; \underbrace{\lambda\, \mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Variation Totale}}$$



μ-fortement convexe

non lisse

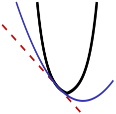


Forte convexité

- φ μ-fortement convexe ssi $\varphi - \frac{\mu}{2}\|\cdot\|^2$ convexe
- φ $\mathcal{C}^2$ de hessienne $\boldsymbol{H}\varphi \succeq 0 \implies \mu = \min \mathrm{Sp}(\boldsymbol{H}\varphi)$

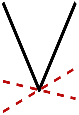
Propriétés de convexité

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \ \underbrace{\sum_a \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \ + \ \lambda \ \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)}_{\textbf{Variation Totale}}$$



μ-fortement convexe

non lisse



Forte convexité

- φ μ-fortement convexe ssi $\varphi - \dfrac{\mu}{2}\|\cdot\|^2$ convexe
- φ $\mathcal{C}^2$ de hessienne $\boldsymbol{H}\varphi \succeq 0 \implies \mu = \min \mathrm{Sp}(\boldsymbol{H}\varphi)$

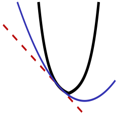
Proposition *(Pascal, 2019)*

$\sum_a \|\log \boldsymbol{\mathcal{L}} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2$ est μ-fortement convexe.

$a_{\min} = 2^1$, $a_{\max}$	2^2	2^3	2^4	2^5	2^6
$\mu = \min \mathrm{Sp}\left(2\boldsymbol{\Phi}^\top\boldsymbol{\Phi}\right)$	0.29	**0.72**	1.20	1.69	2.20

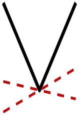
Algorithme accéléré par forte-convexité

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \ \underbrace{\sum_a \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \quad + \quad \lambda \ \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)}_{\textbf{Variation Totale}}$$



μ-fortement convexe

non lisse



Algorithme primal-dual **accéléré** *(Chambolle, 2011)*

for $n = 0,1,\ldots$ $\qquad \boldsymbol{x} = (\boldsymbol{h},\boldsymbol{v})$

$$\boldsymbol{y}^{n+1} = \operatorname{prox}_{\sigma_n(\lambda\mathcal{Q})^*}\left(\boldsymbol{y}^n + \sigma_n \mathbf{D}\bar{\boldsymbol{x}}^n\right)$$

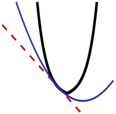
$$\boldsymbol{x}^{n+1} = \operatorname{prox}_{\tau_n\|\boldsymbol{\mathcal{L}}-\boldsymbol{\Phi}\cdot\|_2^2}\left(\boldsymbol{x}^n - \tau_n \mathbf{D}^\top \boldsymbol{y}^{n+1}\right)$$

$$\theta_n = \sqrt{1+2\mu\tau_n}, \quad \tau_{n+1} = \tau_n/\theta_n, \quad \sigma_{n+1} = \theta_n\sigma_n$$

$$\bar{\boldsymbol{x}}^{n+1} = \boldsymbol{x}^{n+1} + \theta_n^{-1}\left(\boldsymbol{x}^{n+1} - \boldsymbol{x}^n\right)$$

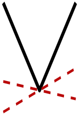
Algorithme accéléré par forte-convexité

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} + \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)}_{\textbf{Variation Totale}}$$



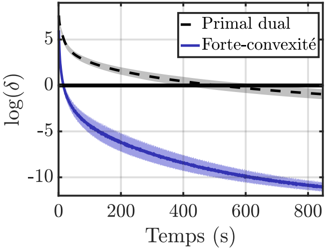
$\boldsymbol{\mu}$-fortement convexe

non lisse



Algorithme primal-dual **accéléré** *(Chambolle, 2011)*

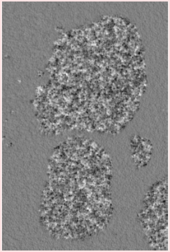
$$\delta \text{ : gap de dualité, } \delta(\boldsymbol{x}^n,\boldsymbol{y}^n) \underset{n\to\infty}{\longrightarrow} 0$$



Segmentation par seuillage itéré

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \quad \sum_{a} \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,.} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)}_{\textbf{Variation Totale}}$$

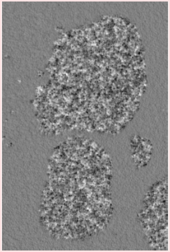
Image texturée



Segmentation par seuillage itéré

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \; \underbrace{\sum_a \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \; + \; \underbrace{\lambda \, \mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Variation Totale}}$$

Image texturée

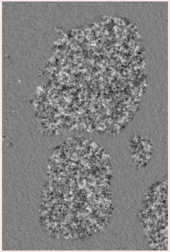


Rég. lin. $\widehat{\boldsymbol{h}}^{\mathrm{RL}}$

Segmentation par seuillage itéré

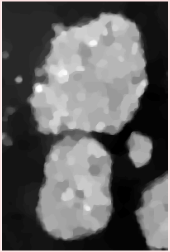
$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \quad \underbrace{\sum_a \|\log \mathcal{L}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \quad + \quad \underbrace{\lambda\, \mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Variation Totale}}$$

Image texturée



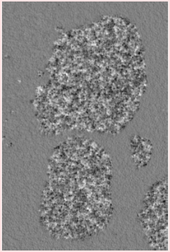
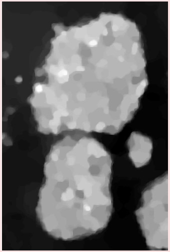
Rég. lin. $\widehat{\boldsymbol{h}}^{\mathrm{RL}}$

Contours co-localisés $\widehat{\boldsymbol{h}}^{\mathrm{C}}$



Segmentation par seuillage itéré

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \quad \underbrace{\sum_a \|\log \mathcal{L}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\text{Moindres Carrés}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)}_{\text{Variation Totale}}$$

Image texturée	Rég. lin. $\widehat{\boldsymbol{h}}^{\mathrm{RL}}$	Contours co-localisés $\widehat{\boldsymbol{h}}^{\mathrm{C}}$	Estimée seuillée[†] $S\widehat{\boldsymbol{h}}^{\mathrm{C}}$
			

[†] *(Cai, 2013)*

Méthodes de l'état-de-l'art en segmentation de texture

ROF-Seuillé sur $\widehat{\boldsymbol{h}}^{\mathrm{RL}}$

(Nafornita, 2014), (Pustelnik, 2016)

$$\underset{\boldsymbol{h}}{\operatorname{argmin}} \|\boldsymbol{h} - \widehat{\boldsymbol{h}}^{\mathrm{RL}}\|^2 + \lambda \|\mathbf{D}\boldsymbol{h}\|_{2,1}$$

S'appuie uniquement sur la régularité $\boldsymbol{h}$.

† https://sites.google.com/site/factorizationsegmentation/

Méthodes de l'état-de-l'art en segmentation de texture

ROF-Seuillé sur $\widehat{\boldsymbol{h}}^{\mathrm{RL}}$

(Nafornita, 2014), (Pustelnik, 2016)

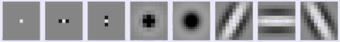
$$\underset{\boldsymbol{h}}{\operatorname{argmin}} \|\boldsymbol{h} - \widehat{\boldsymbol{h}}^{\mathrm{RL}}\|^2 + \lambda \|\mathbf{D}\boldsymbol{h}\|_{2,1}$$

Rég. lin. | ROF | Seuillage

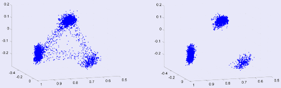
S'appuie uniquement sur la régularité $\boldsymbol{h}$.

Segmentation par factorisation matricielle[†] *(Yuan, 2015)*

(i) histogrammes locaux



(ii) factorisation matricielle



(a) (b)

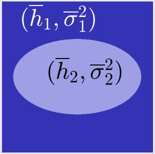
Fig. 2. Scatterplot of features in subspace. (a) Scatterplot of features projected onto the 3-d subspace. (b) Scatterplot after removing features with high edgeness.

[†] `https://sites.google.com/site/factorizationsegmentation/`

Performances comparées sur des textures synthétiques

Synthèse de texture monofractale par morceaux *(Pascal, 2019)*

- masque : $\Omega = \Omega_1 \sqcup \Omega_2$,
- attributs : $(\overline{h}_k, \overline{\sigma}_k^2)_{k=1,2}$

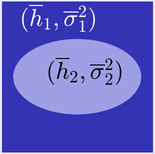


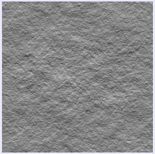
Performances comparées sur des textures synthétiques

Synthèse de texture monofractale par morceaux *(Pascal, 2019)*

- masque : $\Omega = \Omega_1 \sqcup \Omega_2$,
- attributs : $(\bar{h}_k, \bar{\sigma}_k^2)_{k=1,2}$

Ex. $\bar{h}_1 = 0{,}5$, $\bar{\sigma}_1^2 = 0{,}6$
$\bar{h}_2 = 0{,}6$, $\bar{\sigma}_2^2 = 0{,}7$



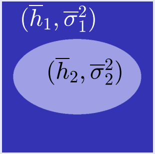


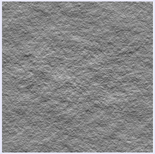
Performances comparées sur des textures synthétiques

Synthèse de texture monofractale par morceaux *(Pascal, 2019)*

- masque : $\Omega = \Omega_1 \sqcup \Omega_2$,
- attributs : $(\overline{h}_k, \overline{\sigma}_k^2)_{k=1,2}$

Ex. $\overline{h}_1 = 0{,}5,\ \overline{\sigma}_1^2 = 0{,}6$
$\overline{h}_2 = 0{,}6,\ \overline{\sigma}_2^2 = 0{,}7$

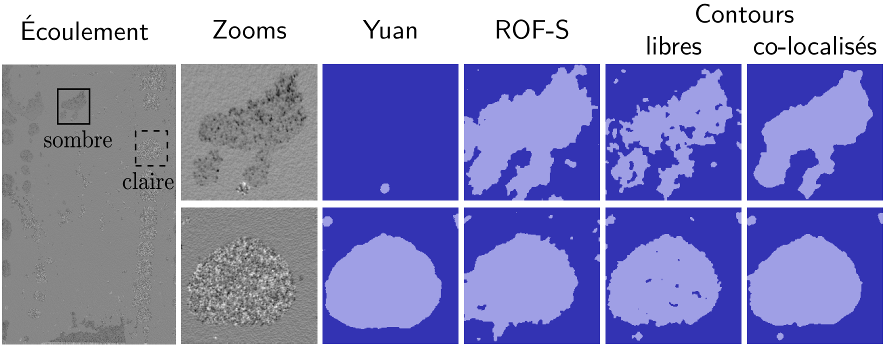




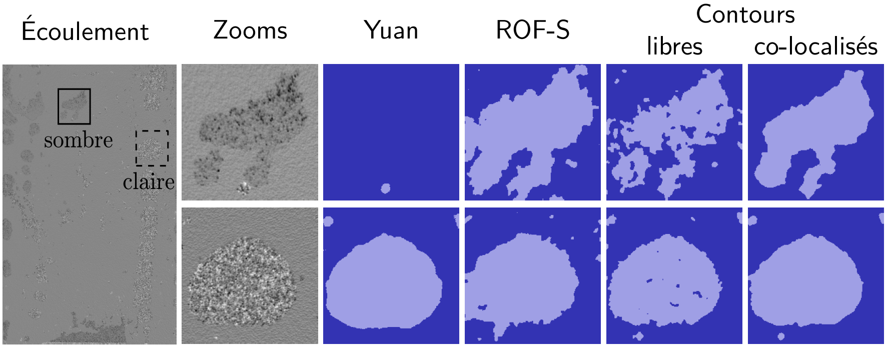
Performances de segmentation moyennées sur 5 réalisations

Yuan	ROF-S	Contours libres	Contours co-localisés
			
$71{,}1 \pm 1{,}3\%$	$78{,}5 \pm 1{,}1\%$	$90{,}2 \pm 1{,}9\%$	$91{,}1 \pm 1{,}5\%$

Faible activité : $Q_{\mathrm{G}} = 300\text{mL}/\text{min}$ - $Q_{\mathrm{L}} = 300\text{mL}/\text{min}$



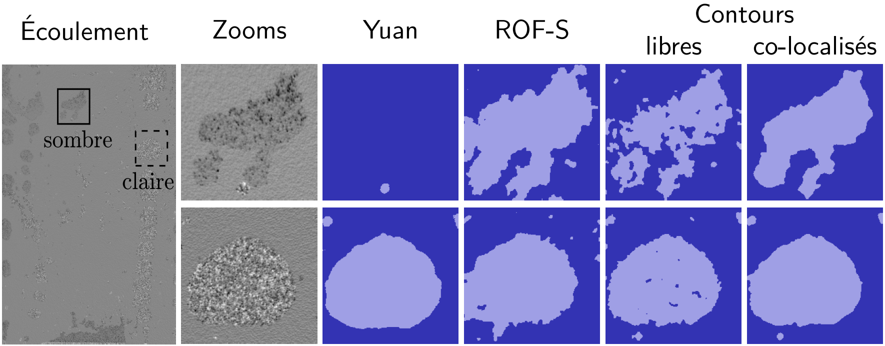
Faible activité : $Q_\mathrm{G} = 300\mathrm{mL/min}$ - $Q_\mathrm{L} = 300\mathrm{mL/min}$



Liquide : $h_\mathrm{L} = 0{,}4$

Gaz : $h_\mathrm{G} = 0{,}9$

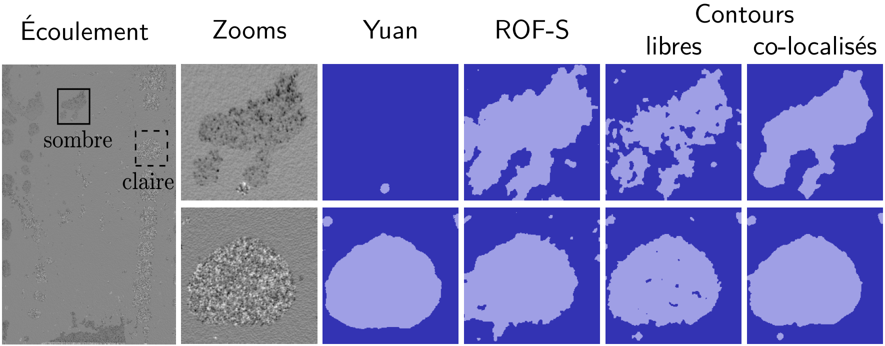
Faible activité : $Q_{\mathrm{G}} = 300\text{mL}/\text{min}$ - $Q_{\mathrm{L}} = 300\text{mL}/\text{min}$



Liquide : $h_{\mathrm{L}} = 0{,}4$ $\sigma^2_{\mathrm{sombre}} = 10^{-2}$

Gaz : $h_{\mathrm{G}} = 0{,}9$

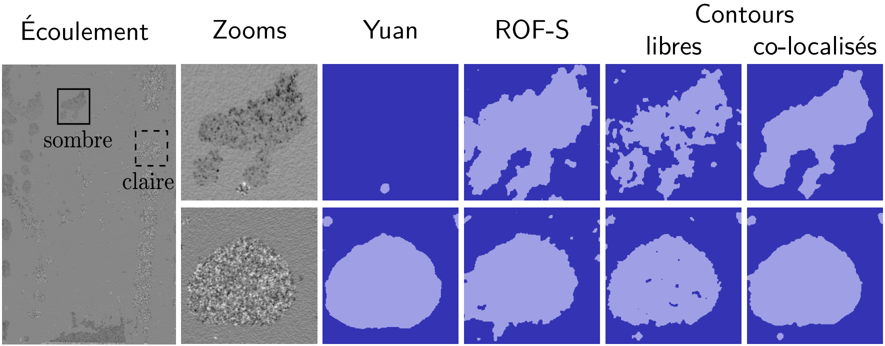
Faible activité : $Q_{\mathrm{G}} = 300\mathrm{mL}/\mathrm{min}$ - $Q_{\mathrm{L}} = 300\mathrm{mL}/\mathrm{min}$



Liquide : $h_{\mathrm{L}} = 0{,}4$ $\sigma^2_{\mathrm{sombre}} = 10^{-2}$

Gaz : $h_{\mathrm{G}} = 0{,}9$

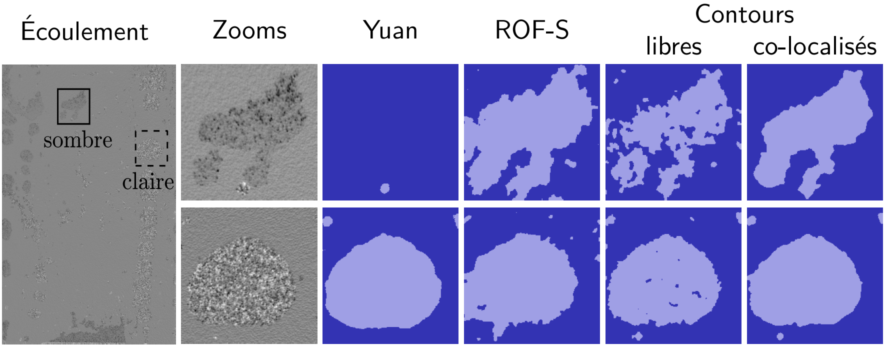
Faible activité : $Q_{\mathrm{G}} = 300\mathrm{mL}/\mathrm{min}$ - $Q_{\mathrm{L}} = 300\mathrm{mL}/\mathrm{min}$



Liquide : $h_{\mathrm{L}} = 0{,}4$ $\quad \sigma^2_{\mathrm{sombre}} = 10^{-2}$

Gaz : $h_{\mathrm{G}} = 0{,}9$ $\quad \sigma^2_{\mathrm{sombre}} = 10^{-2}$ (bulles sombres)

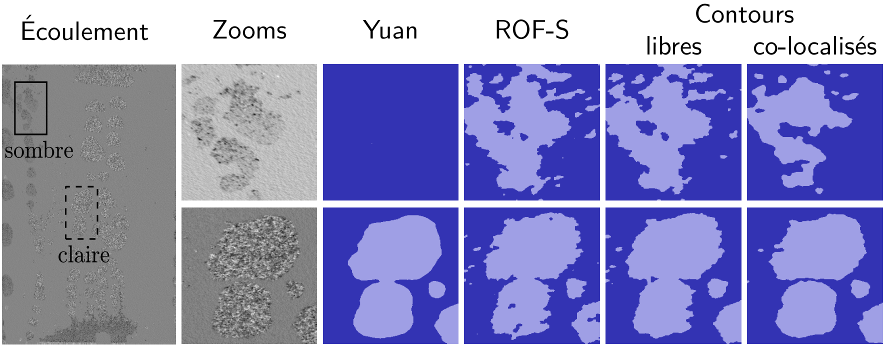
Faible activité : $Q_{\mathrm{G}} = 300\mathrm{mL}/\mathrm{min}$ - $Q_{\mathrm{L}} = 300\mathrm{mL}/\mathrm{min}$



Liquide : $h_{\mathrm{L}} = 0{,}4 \qquad \sigma^2_{\mathrm{sombre}} = 10^{-2}$

Gaz : $h_{\mathrm{G}} = 0{,}9$ $\left| \begin{array}{ll} \sigma^2_{\mathrm{sombre}} = 10^{-2} & \text{(bulles sombres)} \\ \sigma^2_{\mathrm{claire}} = 10^{-1} & \text{(bulles claires)} \end{array} \right.$

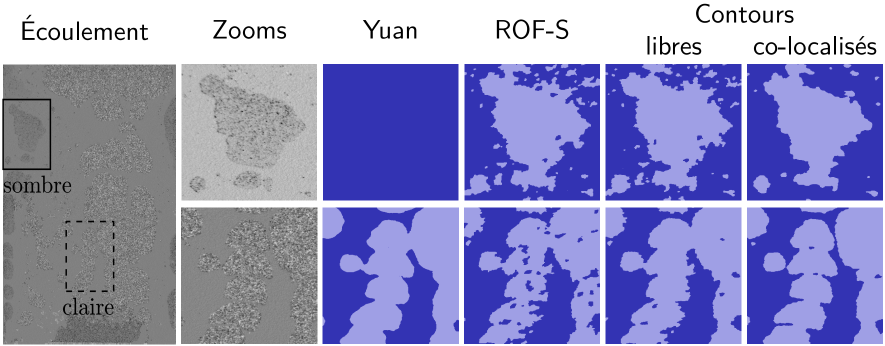
Transition : $Q_{\mathrm{G}} = 400\mathrm{mL}/\mathrm{min}$ - $Q_{\mathrm{L}} = 700\mathrm{mL}/\mathrm{min}$



Liquide : $h_{\mathrm{L}} = 0{,}4 \qquad \sigma^2_{\mathrm{sombre}} = 10^{-2}$

Gaz : $h_{\mathrm{G}} = 0{,}9$ $\left| \begin{array}{ll} \sigma^2_{\mathrm{sombre}} = 10^{-2} & \text{(bulles sombres)} \\ \sigma^2_{\mathrm{claire}} = 10^{-1} & \text{(bulles claires).} \end{array} \right.$

Forte activité : $Q_{\mathrm{G}} = 1200$mL/min - $Q_{\mathrm{L}} = 300$mL/min

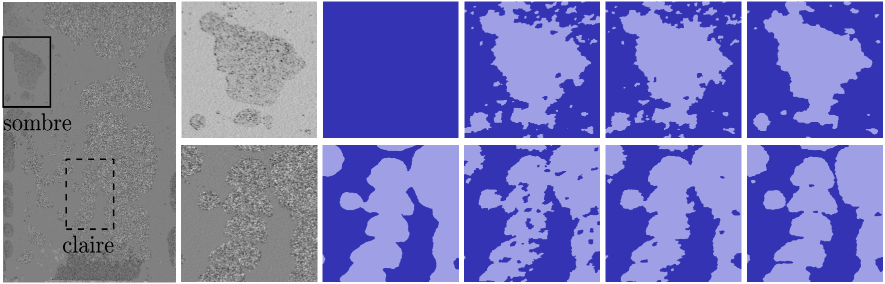


Liquide : $h_{\mathrm{L}} = 0{,}4 \qquad \sigma^2_{\mathrm{sombre}} = 10^{-2}$

Gaz : $h_{\mathrm{G}} = 0{,}9 \quad \left| \begin{array}{ll} \sigma^2_{\mathrm{sombre}} = 10^{-2} & \text{(bulles sombres)} \\ \sigma^2_{\mathrm{claire}} = 10^{-1} & \text{(bulles claires).} \end{array} \right.$

Forte activité : $Q_{\mathrm{G}} = 1200\text{mL}/\text{min}$ - $Q_{\mathrm{L}} = 300\text{mL}/\text{min}$

Écoulement	Zooms	Yuan	ROF-S	Contours libres	Contours co-localisés



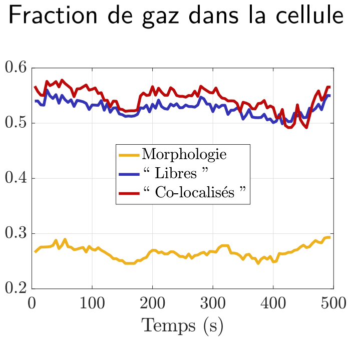
Temps de calcul		1s	12s	700s	**2100s**

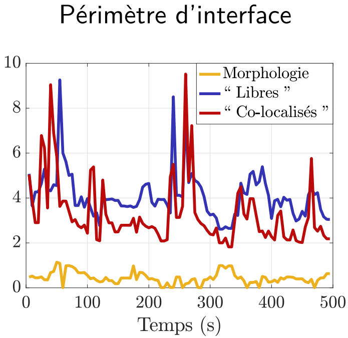
Liquide : $h_{\mathrm{L}} = 0{,}4 \qquad \sigma^2_{\mathrm{sombre}} = 10^{-2}$

Gaz : $h_{\mathrm{G}} = 0{,}9$ $\Big|$ $\sigma^2_{\mathrm{sombre}} = 10^{-2}$ (bulles sombres)
$\sigma^2_{\mathrm{claire}} = 10^{-1}$ (bulles claires).

Écoulement multiphasiques en milieu poreux

Laboratoire de Physique, ENS Lyon, V. Vidal, T. Busser, (M. Serres, IFPEN)





Réglage des paramètres de régularisation

$$\left(\widehat{\boldsymbol{h}},\widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}};\lambda,\alpha) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

Réglage des paramètres de régularisation

$$\left(\widehat{\boldsymbol{h}},\widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}};\lambda,\alpha) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,.} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

Rég. lin. $\widehat{\boldsymbol{h}}^{\mathrm{RL}}$

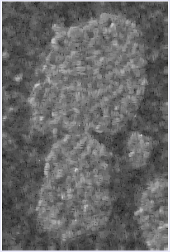
$(\lambda;\alpha) = (0;0)$

Réglage des paramètres de régularisation

$$\left(\widehat{\boldsymbol{h}},\widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}};\lambda,\alpha) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

Rég. lin. $\widehat{\boldsymbol{h}}^{\mathrm{RL}}$ Estimée $\widehat{\boldsymbol{h}}^{\mathrm{C}}$ à contours co-localisés

$(\lambda;\alpha) = (0;0)$ $(\lambda,\alpha) = (0{,}5;0{,}5)$



trop faible

Réglage des paramètres de régularisation

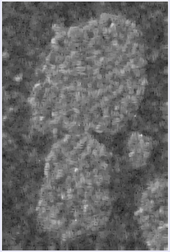
$$\left(\widehat{\boldsymbol{h}},\widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}};\lambda,\alpha) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

Rég. lin. $\widehat{\boldsymbol{h}}^{\mathrm{RL}}$

Estimée $\widehat{\boldsymbol{h}}^{\mathrm{C}}$ à contours co-localisés

$(\lambda;\alpha) = (0;0)$

$(\lambda,\alpha) = (0{,}5;0{,}5)$



trop faible

$(\lambda;\alpha) = (500;500)$

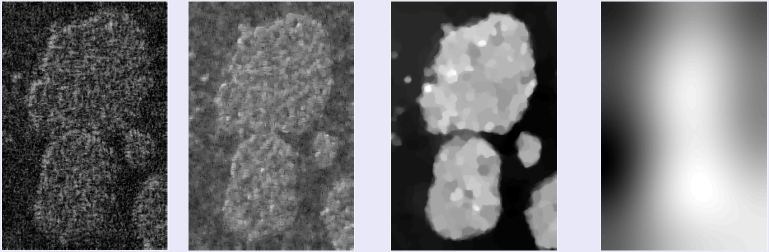
trop grand

Réglage des paramètres de régularisation

$$\left(\widehat{\boldsymbol{h}},\widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}};\lambda,\alpha) = \underset{\boldsymbol{h},\boldsymbol{v}}{\mathrm{argmin}} \sum_a \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

Rég. lin. $\widehat{\boldsymbol{h}}^{\mathrm{RL}}$ — Estimée $\widehat{\boldsymbol{h}}^{\mathrm{C}}$ à contours co-localisés

$(\lambda;\alpha) = (0;0)$ | $(\lambda,\alpha) = (0{,}5;0{,}5)$ | $(\lambda^\dagger,\alpha^\dagger) = (11{,}5;0{,}8)$ | $(\lambda;\alpha) = (500;500)$



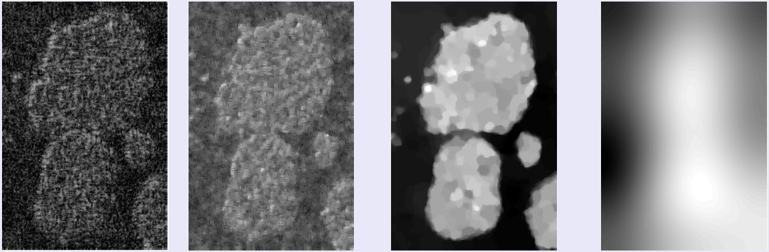
trop faible | optimal | trop grand

Réglage des paramètres de régularisation

$$\left(\widehat{\boldsymbol{h}},\widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}};\lambda,\alpha) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

Rég. lin. $\widehat{\boldsymbol{h}}^{\mathrm{RL}}$ | Estimée $\widehat{\boldsymbol{h}}^{\mathrm{C}}$ à contours co-localisés

$(\lambda;\alpha) = (0;0)$ | $(\lambda,\alpha) = (0{,}5;0{,}5)$ | $(\lambda^\dagger,\alpha^\dagger) = (11{,}5;0{,}8)$ | $(\lambda;\alpha) = (500;500)$



trop faible | optimal | trop grand

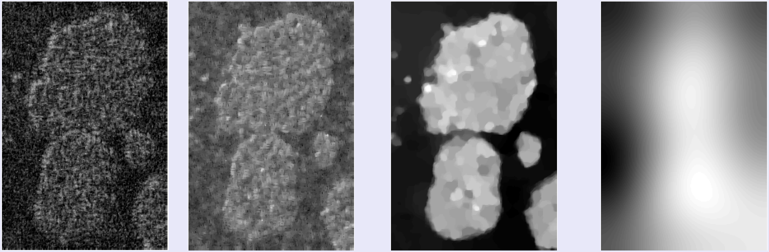
Que signifie *optimal* ?

Réglage des paramètres de régularisation

$$\left(\widehat{\boldsymbol{h}},\widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}};\lambda,\alpha) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_a \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

Rég. lin. $\widehat{\boldsymbol{h}}^{\mathrm{RL}}$ — Estimée $\widehat{\boldsymbol{h}}^{\mathrm{C}}$ à contours co-localisés

$(\lambda;\alpha) = (0;0)$ — $(\lambda,\alpha) = (0{,}5;0{,}5)$ — $(\lambda^\dagger,\alpha^\dagger) = (11{,}5;0{,}8)$ — $(\lambda;\alpha) = (500;500)$



trop faible — optimal — trop grand

Que signifie *optimal* ? Comment déterminer $\lambda^\dagger$ et $\alpha^\dagger$?

Réglage des paramètres (Recherche systématique)

$$\left(\widehat{\boldsymbol{h}},\widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}};\lambda,\alpha) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_a \|\log \boldsymbol{\mathcal{L}}_{a,.} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

$\boldsymbol{h}$: *discriminant*, $\boldsymbol{v}$: *auxiliaire*

Réglage des paramètres (Recherche systématique)

$$\left(\widehat{\boldsymbol{h}},\widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}};\lambda,\alpha) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,.} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

$\boldsymbol{h}$: *discriminant*, $\boldsymbol{v}$: *auxiliaire*

$\bar{\boldsymbol{h}}$: *vraie* régularité

$$\mathcal{R}(\lambda,\alpha) = \left\|\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}};\lambda,\alpha) - \bar{\boldsymbol{h}}\right\|^2$$

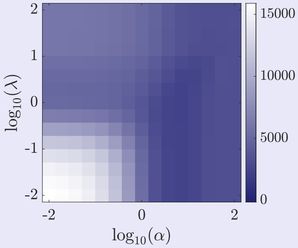
Réglage des paramètres (Recherche systématique)

$$\left(\widehat{\boldsymbol{h}},\widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}};\lambda,\alpha) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_a \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

$\boldsymbol{h}$: *discriminant*, $\boldsymbol{v}$: *auxiliaire*

$\bar{\boldsymbol{h}}$: *vraie* régularité

$$\mathcal{R}(\lambda,\alpha) = \left\|\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}};\lambda,\alpha) - \bar{\boldsymbol{h}}\right\|^2$$



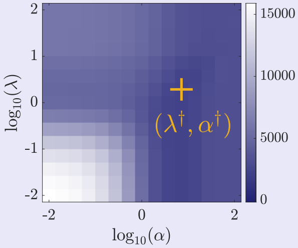
Réglage des paramètres (Recherche systématique)

$$\left(\widehat{\boldsymbol{h}},\widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}};\lambda,\alpha) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,.} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

$\boldsymbol{h}$: *discriminant*, $\boldsymbol{v}$: *auxiliaire*

$\bar{\boldsymbol{h}}$: *vraie* régularité

$$\mathcal{R}(\lambda,\alpha) = \left\|\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}};\lambda,\alpha) - \bar{\boldsymbol{h}}\right\|^2$$



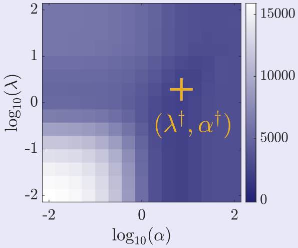
Réglage des paramètres (Recherche systématique)

$$\left(\widehat{\boldsymbol{h}},\widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}};\lambda,\alpha) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_a \|\log \boldsymbol{\mathcal{L}}_{a,.} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

$\boldsymbol{h}$: *discriminant*, $\boldsymbol{v}$: *auxiliaire*

$\bar{\boldsymbol{h}}$: *vraie* régularité

$$\mathcal{R}(\lambda,\alpha) = \left\|\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}};\lambda,\alpha) - \bar{\boldsymbol{h}}\right\|^2$$



$\bar{\boldsymbol{h}}$: inconnue !

?

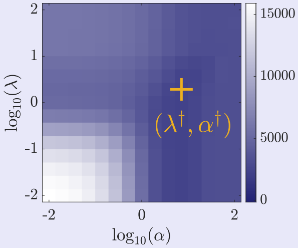
Réglage des paramètres (Recherche systématique)

$$\left(\widehat{\boldsymbol{h}},\widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}};\lambda,\alpha) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

$\boldsymbol{h}$: *discriminant*, $\boldsymbol{v}$: *auxiliaire*

$\bar{\boldsymbol{h}}$: *vraie* régularité

$$\mathcal{R}(\lambda,\alpha) = \left\|\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}};\lambda,\alpha) - \bar{\boldsymbol{h}}\right\|^2$$



$\bar{\boldsymbol{h}}$: inconnue !

?

Stein Unbiased Risk Estimate
(SURE)

Stein Unbiased Risk Estimate (Principe)

Observations $\boldsymbol{y} = \bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P$, $\bar{\boldsymbol{x}}$: vérité et $\zeta \sim \mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$

Stein Unbiased Risk Estimate (Principe)

Observations $\boldsymbol{y} = \bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P$, $\bar{\boldsymbol{x}}$: vérité et $\zeta \sim \mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$

Estimateur paramétrique $(\boldsymbol{y}; \lambda) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)$

Ex. $$\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) = \begin{cases} \left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \boldsymbol{y} & \text{(linéaire)} \\ \underset{\boldsymbol{x}}{\operatorname{argmin}} \, \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{x}) & \text{(non linéaire)} \end{cases}$$

Stein Unbiased Risk Estimate (Principe)

Observations $\boldsymbol{y} = \bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P$, $\bar{\boldsymbol{x}}$: vérité et $\boldsymbol{\zeta} \sim \mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$

Estimateur paramétrique $(\boldsymbol{y}; \lambda) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)$

Ex. $$\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) = \begin{cases} \left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \boldsymbol{y} & \text{(linéaire)} \\ \underset{\boldsymbol{x}}{\operatorname{argmin}} \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{x}) & \text{(non linéaire)} \end{cases}$$

Erreur quadratique $R(\lambda) \triangleq \mathbb{E}_{\boldsymbol{\zeta}} \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \bar{\boldsymbol{x}}\|^2 \overset{?}{=} \mathbb{E}_{\boldsymbol{\zeta}} \widehat{R}(\boldsymbol{y}; \lambda)$ $\bar{\boldsymbol{x}}$ inconnue

Stein Unbiased Risk Estimate (Principe)

Observations $\boldsymbol{y} = \bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P$, $\bar{\boldsymbol{x}}$: vérité et $\zeta \sim \mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$

Estimateur paramétrique $(\boldsymbol{y}; \lambda) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)$

Ex. $$\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) = \begin{cases} \left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \boldsymbol{y} & \text{(linéaire)} \\ \underset{\boldsymbol{x}}{\operatorname{argmin}} \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{x}) & \text{(non linéaire)} \end{cases}$$

Erreur quadratique $R(\lambda) \triangleq \mathbb{E}_\zeta \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \bar{\boldsymbol{x}}\|^2 \overset{?}{=} \mathbb{E}_\zeta \widehat{R}(\boldsymbol{y}; \lambda)$ $\bar{\boldsymbol{x}}$ inconnue

Théorème *(Stein, 1981)*

Soit $(\boldsymbol{y}; \lambda) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)$ un estimateur de $\bar{\boldsymbol{x}}$

- différentiable au sens faible par rapport à $\boldsymbol{y}$,
- tel que $\zeta \mapsto \langle \widehat{\boldsymbol{x}}(\bar{\boldsymbol{x}} + \zeta; \lambda), \zeta \rangle$ est intégrable par rapport à $\mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$.

$$\widehat{R}(\boldsymbol{y}; \lambda) \triangleq \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}\|^2 + 2\rho^2 \operatorname{tr}\left(\partial_{\boldsymbol{y}} \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)\right) - \rho^2 P$$
$$\Longrightarrow R(\lambda) = \mathbb{E}_\zeta[\widehat{R}(\boldsymbol{y}; \lambda)].$$

Stein Unbiased Risk Estimate généralisé

Observations $\boldsymbol{y} = \Phi \bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N, \ \Phi : \mathbb{R}^{P \times N}$ et $\zeta \sim \mathcal{N}(\mathbf{0}, \mathcal{S})$

Stein Unbiased Risk Estimate généralisé

Observations $\boldsymbol{y} = \Phi\bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N$, $\Phi : \mathbb{R}^{P\times N}$ et $\zeta \sim \mathcal{N}(\mathbf{0},\mathcal{S})$

Ex. des estimateurs $\widehat{h}(\mathcal{L}; \lambda,\alpha)$ à contours libres ou co-localisés

$\log \boldsymbol{\mathcal{L}} = \Phi(\bar{\boldsymbol{h}},\bar{\boldsymbol{v}}) + \zeta$

$\Phi : (\boldsymbol{h},\boldsymbol{v}) \mapsto \{\log(a)\boldsymbol{h} + \boldsymbol{v}\}_a$

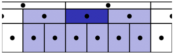
Stein Unbiased Risk Estimate généralisé

Observations $\boldsymbol{y} = \Phi\bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N,\ \Phi : \mathbb{R}^{P\times N}$ et $\zeta \sim \mathcal{N}(\mathbf{0},\mathcal{S})$

Ex. des estimateurs $\widehat{h}(\mathcal{L};\lambda,\alpha)$ à contours libres ou co-localisés

$\log \boldsymbol{\mathcal{L}} = \Phi(\bar{\boldsymbol{h}},\bar{\boldsymbol{v}}) + \zeta \qquad \zeta \sim \mathcal{N}(\mathbf{0},\mathcal{S})$

$\Phi : (\boldsymbol{h},\boldsymbol{v}) \mapsto \{\log(a)\boldsymbol{h} + \boldsymbol{v}\}_a$



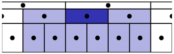
Stein Unbiased Risk Estimate généralisé

Observations $\boldsymbol{y} = \Phi\bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N,\ \Phi : \mathbb{R}^{P\times N}$ et $\zeta \sim \mathcal{N}(\boldsymbol{0},\mathcal{S})$

Ex. des estimateurs $\widehat{h}(\mathcal{L};\lambda,\alpha)$ à contours libres ou co-localisés

$\log \boldsymbol{\mathcal{L}} = \Phi(\bar{\boldsymbol{h}},\bar{\boldsymbol{v}}) + \zeta \qquad \zeta \sim \mathcal{N}(\boldsymbol{0},\mathcal{S}) \qquad \mathcal{R} = \|\widehat{\boldsymbol{h}} - \bar{\boldsymbol{h}}\|^2$

$\Phi : (\boldsymbol{h},\boldsymbol{v}) \mapsto \{\log(a)\boldsymbol{h} + \boldsymbol{v}\}_a$



$\Pi : (\boldsymbol{h},\boldsymbol{v}) \mapsto (\boldsymbol{h},\boldsymbol{0})$

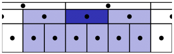
Stein Unbiased Risk Estimate généralisé

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N, \ \boldsymbol{\Phi} : \mathbb{R}^{P\times N}$ et $\zeta \sim \mathcal{N}(\mathbf{0},\mathcal{S})$

Ex. des estimateurs $\widehat{h}(\mathcal{L}; \lambda,\alpha)$ à contours libres ou co-localisés

$\log \boldsymbol{\mathcal{L}} = \boldsymbol{\Phi}(\bar{\boldsymbol{h}},\bar{\boldsymbol{v}}) + \zeta$ $\qquad \zeta \sim \mathcal{N}(\mathbf{0},\mathcal{S})$ $\qquad \mathcal{R} = \|\widehat{\boldsymbol{h}} - \bar{\boldsymbol{h}}\|^2$

$\boldsymbol{\Phi} : (\boldsymbol{h},\boldsymbol{v}) \mapsto \{\log(a)\boldsymbol{h} + \boldsymbol{v}\}_a$



$\boldsymbol{\Pi} : (\boldsymbol{h},\boldsymbol{v}) \mapsto (\boldsymbol{h},\mathbf{0})$

Erreur d'estimation projetée $\quad R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) \triangleq \mathbb{E}_\zeta \|\boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda}) - \boldsymbol{\Pi}\bar{\boldsymbol{x}}\|^2$

Stein Unbiased Risk Estimate généralisé (Calcul)

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N,\ \boldsymbol{\Phi} : \mathbb{R}^{P\times N}$ et $\boldsymbol{\zeta} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\mathcal{S}})$

$$R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) \triangleq \mathbb{E}_{\boldsymbol{\zeta}} \| \boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda}) - \boldsymbol{\Pi}\bar{\boldsymbol{x}} \|^2$$

Stein Unbiased Risk Estimate généralisé (Calcul)

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N, \ \boldsymbol{\Phi} : \mathbb{R}^{P\times N}$ et $\boldsymbol{\zeta} \sim \mathcal{N}(\boldsymbol{0},\mathcal{S})$

$$R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) \triangleq \mathbb{E}_{\boldsymbol{\zeta}} \|\boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{\Pi}\bar{\boldsymbol{x}}\|^2$$

$$= \mathbb{E}_{\boldsymbol{\zeta}} \left\|\boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top\boldsymbol{\Phi}(\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \bar{\boldsymbol{x}})\right\|^2 \qquad \mathbf{A} \triangleq \boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top$$

Stein Unbiased Risk Estimate généralisé (Calcul)

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N, \ \boldsymbol{\Phi} : \mathbb{R}^{P\times N}$ et $\boldsymbol{\zeta} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\mathcal{S}})$

$$
\begin{aligned}
R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) &\triangleq \mathbb{E}_{\boldsymbol{\zeta}} \|\boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{\Pi}\bar{\boldsymbol{x}}\|^2 \\
&= \mathbb{E}_{\boldsymbol{\zeta}} \left\| \boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top\boldsymbol{\Phi}(\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \bar{\boldsymbol{x}}) \right\|^2 \qquad \mathbf{A} \triangleq \boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top \\
&= \mathbb{E}_{\boldsymbol{\zeta}} \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \boldsymbol{y} - \boldsymbol{\Phi}\bar{\boldsymbol{x}})\|^2
\end{aligned}
$$

Stein Unbiased Risk Estimate généralisé (Calcul)

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N,\ \boldsymbol{\Phi} : \mathbb{R}^{P\times N}$ et $\boldsymbol{\zeta} \sim \mathcal{N}(\boldsymbol{0}, \mathcal{S})$

$$\begin{aligned}
R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) &\triangleq \mathbb{E}_{\zeta} \|\boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{\Pi}\bar{\boldsymbol{x}}\|^2 \\
&= \mathbb{E}_{\zeta} \left\|\boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top\boldsymbol{\Phi}(\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \bar{\boldsymbol{x}})\right\|^2 \qquad \mathbf{A} \triangleq \boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top \\
&= \mathbb{E}_{\zeta} \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \boldsymbol{y} - \boldsymbol{\Phi}\bar{\boldsymbol{x}})\|^2 \\
&= \mathbb{E}_{\zeta} \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \boldsymbol{\zeta})\|^2
\end{aligned}$$

Stein Unbiased Risk Estimate généralisé (Calcul)

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N,\ \boldsymbol{\Phi} : \mathbb{R}^{P\times N}$ et $\zeta \sim \mathcal{N}(\mathbf{0}, \mathcal{S})$

$$
\begin{aligned}
R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) &\triangleq \mathbb{E}_\zeta \|\boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{\Pi}\bar{\boldsymbol{x}}\|^2 \\
&= \mathbb{E}_\zeta \left\|\boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top\boldsymbol{\Phi}(\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \bar{\boldsymbol{x}})\right\|^2 \qquad \mathbf{A} \triangleq \boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top \\
&= \mathbb{E}_\zeta \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \boldsymbol{y} - \boldsymbol{\Phi}\bar{\boldsymbol{x}})\|^2 \\
&= \mathbb{E}_\zeta \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \zeta)\|^2 \\
&= \mathbb{E}_\zeta \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\left\langle \mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y}), \mathbf{A}\zeta\right\rangle + \|\mathbf{A}\zeta\|^2\right]
\end{aligned}
$$

Stein Unbiased Risk Estimate généralisé (Calcul)

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N, \; \boldsymbol{\Phi} : \mathbb{R}^{P \times N}$ et $\zeta \sim \mathcal{N}(\boldsymbol{0}, \mathcal{S})$

$$
\begin{aligned}
R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) &\triangleq \mathbb{E}_\zeta \|\boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{\Pi}\bar{\boldsymbol{x}}\|^2 \\
&= \mathbb{E}_\zeta \left\|\boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top\boldsymbol{\Phi}(\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \bar{\boldsymbol{x}})\right\|^2 \qquad \mathbf{A} \triangleq \boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top \\
&= \mathbb{E}_\zeta \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \boldsymbol{y} - \boldsymbol{\Phi}\bar{\boldsymbol{x}})\|^2 \\
&= \mathbb{E}_\zeta \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \zeta)\|^2 \\
&= \mathbb{E}_\zeta \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\langle \mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y}), \mathbf{A}\zeta\rangle + \|\mathbf{A}\zeta\|^2\right] \\
&= \mathbb{E}_\zeta \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\zeta\rangle - 2\langle \mathbf{A}\boldsymbol{y}, \mathbf{A}\zeta\rangle + \|\mathbf{A}\zeta\|^2\right]
\end{aligned}
$$

Stein Unbiased Risk Estimate généralisé (Calcul)

Observations $\quad \boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N, \ \boldsymbol{\Phi} : \mathbb{R}^{P \times N}$ et $\zeta \sim \mathcal{N}(\boldsymbol{0}, \mathcal{S})$

$$
\begin{aligned}
R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) &\triangleq \mathbb{E}_\zeta \|\boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{\Pi}\bar{\boldsymbol{x}}\|^2 \\
&= \mathbb{E}_\zeta \left\|\boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top\boldsymbol{\Phi}(\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \bar{\boldsymbol{x}})\right\|^2 \qquad \mathbf{A} \triangleq \boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top \\
&= \mathbb{E}_\zeta \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \boldsymbol{y} - \boldsymbol{\Phi}\bar{\boldsymbol{x}})\|^2 \\
&= \mathbb{E}_\zeta \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \zeta)\|^2 \\
&= \mathbb{E}_\zeta \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\langle \mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y}), \mathbf{A}\zeta\rangle + \|\mathbf{A}\zeta\|^2\right] \\
&= \mathbb{E}_\zeta \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\zeta\rangle - 2\langle \mathbf{A}\boldsymbol{y}, \mathbf{A}\zeta\rangle + \|\mathbf{A}\zeta\|^2\right] \\
&= \mathbb{E}_\zeta \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\zeta\rangle - 2\langle \mathbf{A}(\boldsymbol{\Phi}\bar{\boldsymbol{x}} + \zeta), \mathbf{A}\zeta\rangle + \|\mathbf{A}\zeta\|^2\right]
\end{aligned}
$$

Stein Unbiased Risk Estimate généralisé (Calcul)

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N, \ \boldsymbol{\Phi} : \mathbb{R}^{P\times N}$ et $\zeta \sim \mathcal{N}(\mathbf{0},\mathcal{S})$

$$\begin{aligned}
R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) &\triangleq \mathbb{E}_\zeta \|\boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{\Pi}\bar{\boldsymbol{x}}\|^2 \\
&= \mathbb{E}_\zeta \left\|\boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top\boldsymbol{\Phi}(\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \bar{\boldsymbol{x}})\right\|^2 \qquad \mathbf{A} \triangleq \boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top \\
&= \mathbb{E}_\zeta \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \boldsymbol{y} - \boldsymbol{\Phi}\bar{\boldsymbol{x}})\|^2 \\
&= \mathbb{E}_\zeta \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \zeta)\|^2 \\
&= \mathbb{E}_\zeta \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\langle \mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y}), \mathbf{A}\zeta\rangle + \|\mathbf{A}\zeta\|^2\right] \\
&= \mathbb{E}_\zeta \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\zeta\rangle - 2\langle \mathbf{A}\boldsymbol{y}, \mathbf{A}\zeta\rangle + \|\mathbf{A}\zeta\|^2\right] \\
&= \mathbb{E}_\zeta \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\zeta\rangle - 2\langle \mathbf{A} \qquad \zeta, \mathbf{A}\zeta\rangle + \|\mathbf{A}\zeta\|^2\right]
\end{aligned}$$

Stein Unbiased Risk Estimate généralisé (Calcul)

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N, \ \boldsymbol{\Phi} : \mathbb{R}^{P\times N}$ et $\zeta \sim \mathcal{N}(\mathbf{0}, \mathcal{S})$

$$
\begin{aligned}
R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) &\triangleq \mathbb{E}_\zeta \|\boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{\Pi}\bar{\boldsymbol{x}}\|^2 \\
&= \mathbb{E}_\zeta \left\|\boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top\boldsymbol{\Phi}(\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \bar{\boldsymbol{x}})\right\|^2 \qquad \mathbf{A} \triangleq \boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top \\
&= \mathbb{E}_\zeta \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \boldsymbol{y} - \boldsymbol{\Phi}\bar{\boldsymbol{x}})\|^2 \\
&= \mathbb{E}_\zeta \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \zeta)\|^2 \\
&= \mathbb{E}_\zeta \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\langle \mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y}), \mathbf{A}\zeta\rangle + \|\mathbf{A}\zeta\|^2\right] \\
&= \mathbb{E}_\zeta \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\zeta\rangle - 2\langle \mathbf{A}\boldsymbol{y}, \mathbf{A}\zeta\rangle + \|\mathbf{A}\zeta\|^2\right] \\
&= \mathbb{E}_\zeta \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\zeta\rangle - 2\langle \mathbf{A} \qquad \zeta, \mathbf{A}\zeta\rangle + \|\mathbf{A}\zeta\|^2\right] \\
&= \mathbb{E}_\zeta \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\mathbb{E}_\zeta \langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\zeta\rangle - \mathbb{E}_\zeta \|\mathbf{A}\zeta\|^2
\end{aligned}
$$

Stein Unbiased Risk Estimate généralisé (Calcul)

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N, \ \boldsymbol{\Phi} : \mathbb{R}^{P\times N}$ et $\boldsymbol{\zeta} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\mathcal{S}})$

$$\begin{aligned}
R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) &\triangleq \mathbb{E}_{\boldsymbol{\zeta}} \|\boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{\Pi}\bar{\boldsymbol{x}}\|^2 \\
&= \mathbb{E}_{\boldsymbol{\zeta}} \left\|\boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top\boldsymbol{\Phi}(\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \bar{\boldsymbol{x}})\right\|^2 \qquad \mathbf{A} \triangleq \boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top \\
&= \mathbb{E}_{\boldsymbol{\zeta}} \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \boldsymbol{y} - \boldsymbol{\Phi}\bar{\boldsymbol{x}})\|^2 \\
&= \mathbb{E}_{\boldsymbol{\zeta}} \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \boldsymbol{\zeta})\|^2 \\
&= \mathbb{E}_{\boldsymbol{\zeta}} \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\left\langle \mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y}), \mathbf{A}\boldsymbol{\zeta}\right\rangle + \|\mathbf{A}\boldsymbol{\zeta}\|^2\right] \\
&= \mathbb{E}_{\boldsymbol{\zeta}} \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\left\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\boldsymbol{\zeta}\right\rangle - 2\left\langle \mathbf{A}\boldsymbol{y}, \mathbf{A}\boldsymbol{\zeta}\right\rangle + \|\mathbf{A}\boldsymbol{\zeta}\|^2\right] \\
&= \mathbb{E}_{\boldsymbol{\zeta}} \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\left\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\boldsymbol{\zeta}\right\rangle - 2\left\langle \mathbf{A} \qquad \boldsymbol{\zeta}, \mathbf{A}\boldsymbol{\zeta}\right\rangle + \|\mathbf{A}\boldsymbol{\zeta}\|^2\right] \\
&= \mathbb{E}_{\boldsymbol{\zeta}} \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\mathbb{E}_{\boldsymbol{\zeta}} \left\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\boldsymbol{\zeta}\right\rangle - \operatorname{tr}\left(\mathbf{A}\boldsymbol{\mathcal{S}}\mathbf{A}^\top\right)
\end{aligned}$$

Stein Unbiased Risk Estimate généralisé (Calcul)

Observations $\quad \boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N, \ \boldsymbol{\Phi} : \mathbb{R}^{P\times N}$ et $\boldsymbol{\zeta} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\mathcal{S}})$

$$\begin{aligned}
R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) &\triangleq \mathbb{E}_{\boldsymbol{\zeta}} \|\boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{\Pi}\bar{\boldsymbol{x}}\|^2 \\
&= \mathbb{E}_{\boldsymbol{\zeta}} \left\|\boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top\boldsymbol{\Phi}(\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \bar{\boldsymbol{x}})\right\|^2 \qquad \mathbf{A} \triangleq \boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top \\
&= \mathbb{E}_{\boldsymbol{\zeta}} \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \boldsymbol{y} - \boldsymbol{\Phi}\bar{\boldsymbol{x}})\|^2 \\
&= \mathbb{E}_{\boldsymbol{\zeta}} \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \boldsymbol{\zeta})\|^2 \\
&= \mathbb{E}_{\boldsymbol{\zeta}} \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\left\langle \mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y}), \mathbf{A}\boldsymbol{\zeta}\right\rangle + \|\mathbf{A}\boldsymbol{\zeta}\|^2\right] \\
&= \mathbb{E}_{\boldsymbol{\zeta}} \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\left\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\boldsymbol{\zeta}\right\rangle - 2\left\langle \mathbf{A}\boldsymbol{y}, \mathbf{A}\boldsymbol{\zeta}\right\rangle + \|\mathbf{A}\boldsymbol{\zeta}\|^2\right] \\
&= \mathbb{E}_{\boldsymbol{\zeta}} \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\left\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\boldsymbol{\zeta}\right\rangle - 2\left\langle \mathbf{A} \qquad \boldsymbol{\zeta}, \mathbf{A}\boldsymbol{\zeta}\right\rangle + \|\mathbf{A}\boldsymbol{\zeta}\|^2\right] \\
&= \mathbb{E}_{\boldsymbol{\zeta}} \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\mathbb{E}_{\boldsymbol{\zeta}} \left\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\boldsymbol{\zeta}\right\rangle \underbrace{- \operatorname{tr}\left(\mathbf{A}\boldsymbol{\mathcal{S}}\mathbf{A}^\top\right)}_{\text{accessible}}
\end{aligned}$$

Stein Unbiased Risk Estimate généralisé (Calcul)

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N, \ \boldsymbol{\Phi} : \mathbb{R}^{P\times N}$ et $\zeta \sim \mathcal{N}(\mathbf{0}, \mathcal{S})$

$$
\begin{aligned}
R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) &\triangleq \mathbb{E}_\zeta \|\boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{\Pi}\bar{\boldsymbol{x}}\|^2 \\
&= \mathbb{E}_\zeta \left\|\boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top\boldsymbol{\Phi}(\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \bar{\boldsymbol{x}})\right\|^2 \qquad \mathbf{A} \triangleq \boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top \\
&= \mathbb{E}_\zeta \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \boldsymbol{y} - \boldsymbol{\Phi}\bar{\boldsymbol{x}})\|^2 \\
&= \mathbb{E}_\zeta \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \zeta)\|^2 \\
&= \mathbb{E}_\zeta \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\langle \mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y}), \mathbf{A}\zeta\rangle + \|\mathbf{A}\zeta\|^2\right] \\
&= \mathbb{E}_\zeta \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\zeta\rangle - 2\langle \mathbf{A}\boldsymbol{y}, \mathbf{A}\zeta\rangle + \|\mathbf{A}\zeta\|^2\right] \\
&= \mathbb{E}_\zeta \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\zeta\rangle - 2\langle \mathbf{A} \qquad \zeta, \mathbf{A}\zeta\rangle + \|\mathbf{A}\zeta\|^2\right] \\
&= \mathbb{E}_\zeta \underbrace{\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2}_{\text{accessible}} + 2\mathbb{E}_\zeta \langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\zeta\rangle - \underbrace{\operatorname{tr}\left(\mathbf{A}\mathcal{S}\mathbf{A}^\top\right)}_{\text{accessible}}
\end{aligned}
$$

Stein Unbiased Risk Estimate généralisé (Calcul)

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N, \ \boldsymbol{\Phi} : \mathbb{R}^{P\times N}$ et $\zeta \sim \mathcal{N}(\mathbf{0}, \mathcal{S})$

$$\mathbf{A} \triangleq \boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top$$

$$\begin{aligned}
R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) &\triangleq \mathbb{E}_\zeta \|\boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{\Pi}\bar{\boldsymbol{x}}\|^2 \\
&= \mathbb{E}_\zeta \left\|\boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top\boldsymbol{\Phi}(\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \bar{\boldsymbol{x}})\right\|^2 \\
&= \mathbb{E}_\zeta \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \boldsymbol{y} - \boldsymbol{\Phi}\bar{\boldsymbol{x}})\|^2 \\
&= \mathbb{E}_\zeta \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \zeta)\|^2 \\
&= \mathbb{E}_\zeta \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\langle \mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y}), \mathbf{A}\zeta\rangle + \|\mathbf{A}\zeta\|^2\right] \\
&= \mathbb{E}_\zeta \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\zeta\rangle - 2\langle \mathbf{A}\boldsymbol{y}, \mathbf{A}\zeta\rangle + \|\mathbf{A}\zeta\|^2\right] \\
&= \mathbb{E}_\zeta \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\zeta\rangle - 2\langle \mathbf{A} \qquad \zeta, \mathbf{A}\zeta\rangle + \|\mathbf{A}\zeta\|^2\right] \\
&= \mathbb{E}_\zeta \underbrace{\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2}_{\text{accessible}} + 2\mathbb{E}_\zeta \langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\zeta\rangle - \underbrace{\mathrm{tr}\left(\mathbf{A}\mathcal{S}\mathbf{A}^\top\right)}_{\text{accessible}}
\end{aligned}$$

Stein Unbiased Risk Estimate généralisé (Calcul)

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N, \ \boldsymbol{\Phi} : \mathbb{R}^{P\times N}$ et $\boldsymbol{\zeta} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\mathcal{S}})$

$$
\begin{aligned}
R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) &\triangleq \mathbb{E}_{\boldsymbol{\zeta}} \|\boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{\Pi}\bar{\boldsymbol{x}}\|^2 \\
&= \mathbb{E}_{\boldsymbol{\zeta}} \left\|\boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top\boldsymbol{\Phi}(\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \bar{\boldsymbol{x}})\right\|^2 \qquad \mathbf{A} \triangleq \boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top \\
&= \mathbb{E}_{\boldsymbol{\zeta}} \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \boldsymbol{y} - \boldsymbol{\Phi}\bar{\boldsymbol{x}})\|^2 \\
&= \mathbb{E}_{\boldsymbol{\zeta}} \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \boldsymbol{\zeta})\|^2 \\
&= \mathbb{E}_{\boldsymbol{\zeta}} \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\left\langle \mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y}), \mathbf{A}\boldsymbol{\zeta}\right\rangle + \|\mathbf{A}\boldsymbol{\zeta}\|^2\right] \\
&= \mathbb{E}_{\boldsymbol{\zeta}} \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\left\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\boldsymbol{\zeta}\right\rangle - 2\left\langle \mathbf{A}\boldsymbol{y}, \mathbf{A}\boldsymbol{\zeta}\right\rangle + \|\mathbf{A}\boldsymbol{\zeta}\|^2\right] \\
&= \mathbb{E}_{\boldsymbol{\zeta}} \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\left\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\boldsymbol{\zeta}\right\rangle - 2\left\langle \mathbf{A} \quad \boldsymbol{\zeta}, \mathbf{A}\boldsymbol{\zeta}\right\rangle + \|\mathbf{A}\boldsymbol{\zeta}\|^2\right] \\
&= \mathbb{E}_{\boldsymbol{\zeta}} \underbrace{\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2}_{\text{accessible}} + 2\mathbb{E}_{\boldsymbol{\zeta}} \left\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\boldsymbol{\zeta}\right\rangle - \underbrace{\mathrm{tr}\left(\mathbf{A}\boldsymbol{\mathcal{S}}\mathbf{A}^\top\right)}_{\text{accessible}}
\end{aligned}
$$

$$
\mathbb{E}_{\boldsymbol{\zeta}} \left\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\boldsymbol{\zeta}\right\rangle = \mathbb{E}_{\boldsymbol{\zeta}} \left\langle \boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\boldsymbol{\zeta}\right\rangle
$$

Stein Unbiased Risk Estimate généralisé (Calcul)

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N, \ \boldsymbol{\Phi} : \mathbb{R}^{P\times N}$ et $\zeta \sim \mathcal{N}(\mathbf{0}, \mathcal{S})$

$$
\begin{aligned}
R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) &\triangleq \mathbb{E}_\zeta \|\boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{\Pi}\bar{\boldsymbol{x}}\|^2 \\
&= \mathbb{E}_\zeta \left\|\boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top\boldsymbol{\Phi}(\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \bar{\boldsymbol{x}})\right\|^2 \qquad \mathbf{A} \triangleq \boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top \\
&= \mathbb{E}_\zeta \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \boldsymbol{y} - \boldsymbol{\Phi}\bar{\boldsymbol{x}})\|^2 \\
&= \mathbb{E}_\zeta \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \zeta)\|^2 \\
&= \mathbb{E}_\zeta \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\langle \mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y}), \mathbf{A}\zeta\rangle + \|\mathbf{A}\zeta\|^2\right] \\
&= \mathbb{E}_\zeta \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\zeta\rangle - 2\langle \mathbf{A}\boldsymbol{y}, \mathbf{A}\zeta\rangle + \|\mathbf{A}\zeta\|^2\right] \\
&= \mathbb{E}_\zeta \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\zeta\rangle - 2\langle \mathbf{A} \qquad \zeta, \mathbf{A}\zeta\rangle + \|\mathbf{A}\zeta\|^2\right] \\
&= \mathbb{E}_\zeta \underbrace{\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2}_{\text{accessible}} + 2\mathbb{E}_\zeta \langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\zeta\rangle - \underbrace{\operatorname{tr}\left(\mathbf{A}\mathcal{S}\mathbf{A}^\top\right)}_{\text{accessible}}
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}_\zeta \langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\zeta\rangle &= \mathbb{E}_\zeta \langle \boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}), \mathbf{A}\zeta\rangle \\
&= \int \langle \boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{\Phi}\bar{\boldsymbol{x}} + \zeta;\lambda), \mathbf{A}\zeta\rangle \exp\left(-\frac{\zeta^\top\mathcal{S}^{-1}\zeta}{2}\right) \mathrm{d}\zeta
\end{aligned}
$$

Stein Unbiased Risk Estimate généralisé (Calcul)

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N, \ \boldsymbol{\Phi} : \mathbb{R}^{P\times N}$ et $\zeta \sim \mathcal{N}(\boldsymbol{0},\mathcal{S})$

$$
\begin{aligned}
R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) &\triangleq \mathbb{E}_\zeta \|\boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{\Pi}\bar{\boldsymbol{x}}\|^2 \\
&= \mathbb{E}_\zeta \left\|\boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top\boldsymbol{\Phi}(\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \bar{\boldsymbol{x}})\right\|^2 \qquad \mathbf{A} \triangleq \boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top \\
&= \mathbb{E}_\zeta \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \boldsymbol{y} - \boldsymbol{\Phi}\bar{\boldsymbol{x}})\|^2 \\
&= \mathbb{E}_\zeta \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y} + \zeta)\|^2 \\
&= \mathbb{E}_\zeta \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\left\langle \mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y}),\mathbf{A}\zeta\right\rangle + \|\mathbf{A}\zeta\|^2\right] \\
&= \mathbb{E}_\zeta \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\left\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}),\mathbf{A}\zeta\right\rangle - 2\left\langle \mathbf{A}\boldsymbol{y},\mathbf{A}\zeta\right\rangle + \|\mathbf{A}\zeta\|^2\right] \\
&= \mathbb{E}_\zeta \left[\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\left\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}),\mathbf{A}\zeta\right\rangle - 2\left\langle \mathbf{A} \qquad \zeta\ ,\mathbf{A}\zeta\right\rangle + \|\mathbf{A}\zeta\|^2\right] \\
&= \mathbb{E}_\zeta \underbrace{\|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y})\|^2}_{\text{accessible}} + 2\mathbb{E}_\zeta \left\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}),\mathbf{A}\zeta\right\rangle - \underbrace{\operatorname{tr}\left(\mathbf{A}\mathcal{S}\mathbf{A}^\top\right)}_{\text{accessible}}
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}_\zeta \left\langle \mathbf{A}\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}),\mathbf{A}\zeta\right\rangle &= \mathbb{E}_\zeta \left\langle \boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}),\mathbf{A}\zeta\right\rangle \\
&= \int \left\langle \boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{\Phi}\bar{\boldsymbol{x}} + \zeta;\lambda),\mathbf{A}\zeta\right\rangle \exp\left(-\frac{\zeta^\top\mathcal{S}^{-1}\zeta}{2}\right) \mathrm{d}\zeta \\
\text{(I.P.P. gén.)} &= \mathbb{E}_\zeta \operatorname{tr}\left(\mathcal{S}\mathbf{A}^\top\boldsymbol{\Pi}\partial_{\boldsymbol{y}}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda})\right)
\end{aligned}
$$

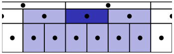
Stein Unbiased Risk Estimate généralisé

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N, \ \boldsymbol{\Phi} : \mathbb{R}^{P \times N}$ et $\boldsymbol{\zeta} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\mathcal{S}})$

Ex. des estimateurs $\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha)$ à contours libres ou co-localisés

$\log \boldsymbol{\mathcal{L}} = \boldsymbol{\Phi}(\bar{\boldsymbol{h}}, \bar{\boldsymbol{v}}) + \boldsymbol{\zeta}$ $\quad \boldsymbol{\zeta} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\mathcal{S}})$ $\quad \mathcal{R} = \|\widehat{\boldsymbol{h}} - \bar{\boldsymbol{h}}\|^2$

$\boldsymbol{\Phi} : (\boldsymbol{h}, \boldsymbol{v}) \mapsto \{\log(a)\boldsymbol{h} + \boldsymbol{v}\}_a$



$\boldsymbol{\Pi} : (\boldsymbol{h}, \boldsymbol{v}) \mapsto (\boldsymbol{h}, \mathbf{0})$

Erreur d'estimation projetée $\quad R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) \triangleq \mathbb{E}_{\boldsymbol{\zeta}} \|\boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda}) - \boldsymbol{\Pi}\bar{\boldsymbol{x}}\|^2$

Théorème *(Pascal, 2020)*

Soit $(\boldsymbol{y}; \boldsymbol{\Lambda}) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda})$ un estimateur de $\bar{\boldsymbol{x}}$

- différentiable au sens faible par rapport à $\boldsymbol{y}$,
- tel que $\boldsymbol{\zeta} \mapsto \langle \boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\bar{\boldsymbol{x}} + \boldsymbol{\zeta}; \lambda), \mathbf{A}\boldsymbol{\zeta} \rangle$ est intégrable par rapport à $\mathcal{N}(\mathbf{0}, \boldsymbol{\mathcal{S}})$.

$$\widehat{R}(\boldsymbol{\Lambda}) \triangleq \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\mathrm{tr}\left(\boldsymbol{\mathcal{S}}\mathbf{A}^\top \boldsymbol{\Pi} \partial_{\boldsymbol{y}} \widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda})\right) - \mathrm{tr}\left(\mathbf{A}\boldsymbol{\mathcal{S}}\mathbf{A}^\top\right)$$

$$\Longrightarrow R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) = \mathbb{E}_{\boldsymbol{\zeta}}[\widehat{R}(\boldsymbol{\Lambda})].$$

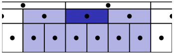
Stein Unbiased Risk Estimate généralisé

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N, \ \boldsymbol{\Phi} : \mathbb{R}^{P\times N}$ et $\boldsymbol{\zeta} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\mathcal{S}})$

Ex. des estimateurs $\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha)$ à contours libres ou co-localisés

$\log \boldsymbol{\mathcal{L}} = \boldsymbol{\Phi}(\bar{\boldsymbol{h}}, \bar{\boldsymbol{v}}) + \boldsymbol{\zeta}$ $\qquad$ $\boldsymbol{\zeta} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\mathcal{S}})$ $\qquad$ $\mathcal{R} = \|\widehat{\boldsymbol{h}} - \bar{\boldsymbol{h}}\|^2$

$\boldsymbol{\Phi} : (\boldsymbol{h}, \boldsymbol{v}) \mapsto \{\log(a)\boldsymbol{h} + \boldsymbol{v}\}_a$



$\boldsymbol{\Pi} : (\boldsymbol{h}, \boldsymbol{v}) \mapsto (\boldsymbol{h}, \boldsymbol{0})$

Erreur d'estimation projetée $\quad R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) \triangleq \mathbb{E}_{\boldsymbol{\zeta}} \|\boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda}) - \boldsymbol{\Pi}\bar{\boldsymbol{x}}\|^2$

Théorème *(Pascal, 2020)*

Soit $(\boldsymbol{y}; \boldsymbol{\Lambda}) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda})$ un estimateur de $\bar{\boldsymbol{x}}$

- différentiable au sens faible par rapport à $\boldsymbol{y}$,
- tel que $\boldsymbol{\zeta} \mapsto \langle \boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\bar{\boldsymbol{x}} + \boldsymbol{\zeta}; \lambda), \mathbf{A}\boldsymbol{\zeta}\rangle$ est intégrable par rapport à $\mathcal{N}(\boldsymbol{0}, \boldsymbol{\mathcal{S}})$.

$$\widehat{R}(\boldsymbol{\Lambda}) \triangleq \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\mathrm{tr}\left(\boldsymbol{\mathcal{S}}\mathbf{A}^\top \boldsymbol{\Pi}\partial_{\boldsymbol{y}}\widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda})\right) - \mathrm{tr}\left(\mathbf{A}\boldsymbol{\mathcal{S}}\mathbf{A}^\top\right)$$

$$\Longrightarrow R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) = \mathbb{E}_{\boldsymbol{\zeta}}[\widehat{R}(\boldsymbol{\Lambda})].$$

Calcul des degrés de liberté

Degrés de liberté $\quad \mathrm{dof} \triangleq \mathrm{tr}\left(\boldsymbol{\mathcal{S}}\mathbf{A}^\top \boldsymbol{\Pi} \partial_{\boldsymbol{y}} \widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda})\right)$

Calcul des degrés de liberté

Degrés de liberté $\quad \mathrm{dof} \triangleq \mathrm{tr}\left(\boldsymbol{\mathcal{S}}\mathbf{A}^\top \boldsymbol{\Pi} \partial_{\boldsymbol{y}} \widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda})\right)$

- **Stratégie de Monte Carlo** (MC) $\quad$ **$\mathbf{M} \in \mathbb{R}^{P\times P}$ de grande taille**

$$\mathrm{tr}(\mathbf{M}) = \mathbb{E}_{\boldsymbol{\varepsilon}} \langle \mathbf{M}\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon} \rangle, \quad \boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I}_P)$$

Calcul des degrés de liberté

Degrés de liberté $\quad \mathrm{dof} \triangleq \mathrm{tr}\left(\boldsymbol{\mathcal{S}}\mathbf{A}^\top \boldsymbol{\Pi} \partial_{\boldsymbol{y}} \widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda})\right)$

- **Stratégie de Monte Carlo** (MC) $\quad$ **$\mathbf{M} \in \mathbb{R}^{P\times P}$ de grande taille**

$$\mathrm{tr}(\mathbf{M}) = \mathbb{E}_{\boldsymbol{\varepsilon}} \langle \mathbf{M}\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}\rangle, \quad \boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I}_P)$$

- **Différences Finies** (DF) $\quad$ **Jacobienne inaccessible**

$$\partial_{\boldsymbol{y}} \widehat{\boldsymbol{x}}[\boldsymbol{\varepsilon}] \underset{\nu\to 0}{\simeq} \frac{1}{\nu}\left(\widehat{\boldsymbol{x}}(\boldsymbol{y}+\nu\boldsymbol{\varepsilon};\boldsymbol{\Lambda}) - \widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda})\right)$$

Calcul des degrés de liberté

Degrés de liberté $\quad \mathrm{dof} \triangleq \mathrm{tr}\left(\boldsymbol{\mathcal{S}}\mathbf{A}^\top \boldsymbol{\Pi} \partial_{\boldsymbol{y}} \widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda})\right)$

- **Stratégie de Monte Carlo** (MC) — **$\mathbf{M} \in \mathbb{R}^{P \times P}$ de grande taille**

$$\mathrm{tr}(\mathbf{M}) = \mathbb{E}_{\boldsymbol{\varepsilon}} \left\langle \mathbf{M}\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon} \right\rangle, \quad \boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I}_P)$$

- **Différences Finies** (DF) — **Jacobienne inaccessible**

$$\partial_{\boldsymbol{y}} \widehat{\boldsymbol{x}}\left[\boldsymbol{\varepsilon}\right] \underset{\nu \to 0}{\simeq} \frac{1}{\nu}\left(\widehat{\boldsymbol{x}}(\boldsymbol{y} + \nu\boldsymbol{\varepsilon}; \boldsymbol{\Lambda}) - \widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda})\right)$$

Proposition *(Pascal, 2020)*

Soit $(\boldsymbol{y}; \boldsymbol{\Lambda}) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda})$ un estimateur de $\bar{\boldsymbol{x}}$

- uniformément lipschitzien par rapport à $\boldsymbol{y}$,
- tel que $\forall \boldsymbol{\Lambda} \in \mathbb{R}^L$, $\widehat{\boldsymbol{x}}(\mathbf{0}_P; \boldsymbol{\Lambda}) = \mathbf{0}_N$. Alors

$$\mathbb{E}_{\boldsymbol{\zeta}}\left[\mathrm{dof}\right] = \lim_{\nu \to 0} \mathbb{E}_{\boldsymbol{\zeta}, \boldsymbol{\varepsilon}}\left[\frac{1}{\nu}\left\langle \boldsymbol{\mathcal{S}}\mathbf{A}^\top \boldsymbol{\Pi}\left(\widehat{\boldsymbol{x}}(\boldsymbol{y} + \nu\boldsymbol{\varepsilon}; \boldsymbol{\Lambda}) - \widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda})\right), \boldsymbol{\varepsilon}\right\rangle\right]$$

Stein Unbiased Risk Estimate (Calcul)

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N, \ \boldsymbol{\Phi} : \mathbb{R}^{P \times N}$ et $\boldsymbol{\zeta} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\mathcal{S}})$

Erreur d'estimation projetée $R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) \triangleq \mathbb{E}_{\boldsymbol{\zeta}} \|\boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{\Pi}\bar{\boldsymbol{x}}\|^2$

SURE généralisé Différences Finies Monte Carlo

$$\widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{y};\boldsymbol{\Lambda} \,|\, \boldsymbol{\mathcal{S}}) \triangleq \left\|\mathbf{A}\left(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y}\right)\right\|^2 + \frac{2}{\nu}\left\langle \boldsymbol{\mathcal{S}}\mathbf{A}^\top\boldsymbol{\Pi}\left(\widehat{\boldsymbol{x}}(\boldsymbol{y}+\nu\boldsymbol{\varepsilon};\boldsymbol{\Lambda}) - \widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda})\right), \boldsymbol{\varepsilon}\right\rangle - \operatorname{tr}\left(\mathbf{A}\boldsymbol{\mathcal{S}}\mathbf{A}^\top\right)$$

Stein Unbiased Risk Estimate (Calcul)

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N, \ \boldsymbol{\Phi} : \mathbb{R}^{P\times N}$ et $\zeta \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\mathcal{S}})$

Erreur d'estimation projetée $R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) \triangleq \mathbb{E}_\zeta \|\boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{\Pi}\bar{\boldsymbol{x}}\|^2$

SURE généralisé Différences Finies Monte Carlo

$$\widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{y};\boldsymbol{\Lambda}\,|\,\boldsymbol{\mathcal{S}}) \triangleq \|\mathbf{A}\left(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y}\right)\|^2 + \frac{2}{\nu}\left\langle \boldsymbol{\mathcal{S}}\mathbf{A}^\top\boldsymbol{\Pi}\left(\widehat{\boldsymbol{x}}(\boldsymbol{y}+\nu\boldsymbol{\varepsilon};\boldsymbol{\Lambda}) - \widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda})\right), \boldsymbol{\varepsilon}\right\rangle - \operatorname{tr}\left(\mathbf{A}\boldsymbol{\mathcal{S}}\mathbf{A}^\top\right)$$

Théorème *(Pascal, 2020)*

Soit $(\boldsymbol{y};\boldsymbol{\Lambda}) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda})$ un estimateur de $\bar{\boldsymbol{x}}$

- uniformément lipschitzien par rapport à $\boldsymbol{y}$,
- tel que $\forall \boldsymbol{\Lambda} \in \mathbb{R}^L$, $\widehat{\boldsymbol{x}}(\mathbf{0}_P;\boldsymbol{\Lambda}) = \mathbf{0}_N$. Alors

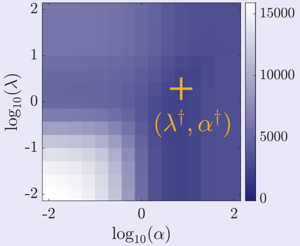
$$R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) = \lim_{\nu\to 0} \mathbb{E}_{\zeta,\boldsymbol{\varepsilon}}\left[\widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{y};\boldsymbol{\Lambda}\,|\,\boldsymbol{\mathcal{S}})\right]$$

Réglage des paramètres (Recherche systématique)

$$\left(\widehat{\boldsymbol{h}},\widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}};\lambda,\alpha) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_a \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

$\bar{\boldsymbol{h}}$: *vraie* régularité

$$\mathcal{R}(\lambda,\alpha) = \left\|\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}};\lambda,\alpha) - \bar{\boldsymbol{h}}\right\|^2$$



$\bar{\boldsymbol{h}}$: inconnue !

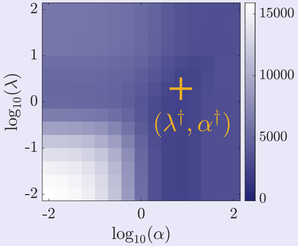
$$\widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{\mathcal{L}};\lambda,\alpha|\,\boldsymbol{\mathcal{S}})$$

Réglage des paramètres (Recherche systématique)

$$\left(\widehat{\boldsymbol{h}},\widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}};\lambda,\alpha) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,.} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

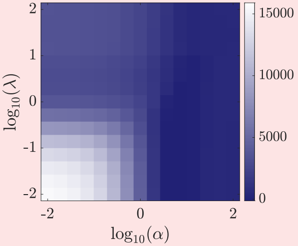
$\bar{\boldsymbol{h}}$: *vraie* régularité

$$\mathcal{R}(\lambda,\alpha) = \left\|\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}};\lambda,\alpha) - \bar{\boldsymbol{h}}\right\|^2$$



$\bar{\boldsymbol{h}}$: inconnue !

$$\widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{\mathcal{L}};\lambda,\alpha|\,\boldsymbol{\mathcal{S}})$$

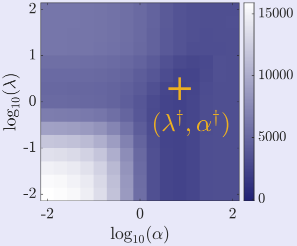


Réglage des paramètres (Recherche systématique)

$$\left(\widehat{\boldsymbol{h}},\widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}};\lambda,\alpha) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_a \|\log \boldsymbol{\mathcal{L}}_{a,.} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

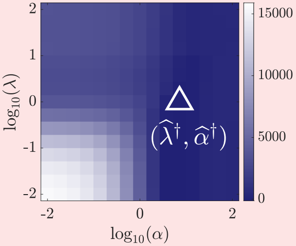
$\bar{\boldsymbol{h}}$: *vraie* régularité

$$\mathcal{R}(\lambda,\alpha) = \left\|\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}};\lambda,\alpha) - \bar{\boldsymbol{h}}\right\|^2$$



$\bar{\boldsymbol{h}}$: inconnue !

$$\widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{\mathcal{L}};\lambda,\alpha|\,\boldsymbol{\mathcal{S}})$$

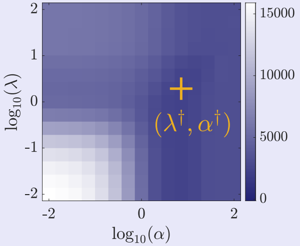


Réglage des paramètres (Recherche systématique)

$$\left(\widehat{\boldsymbol{h}},\widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}};\lambda,\alpha)=\underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}}\sum_{a}\left\|\log\boldsymbol{\mathcal{L}}_{a,\cdot}-\log(a)\boldsymbol{h}-\boldsymbol{v}\right\|^2+\lambda\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

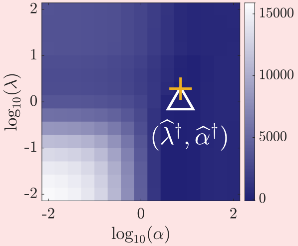
$\bar{\boldsymbol{h}}$: *vraie* régularité

$$\mathcal{R}(\lambda,\alpha)=\left\|\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}};\lambda,\alpha)-\bar{\boldsymbol{h}}\right\|^2$$



$\bar{\boldsymbol{h}}$: inconnue !

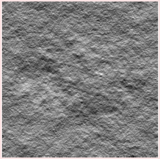
$$\widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{\mathcal{L}};\lambda,\alpha|\,\boldsymbol{\mathcal{S}})$$

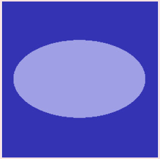


Recherche systématique des paramètres de régularisation

$$\left(\widehat{\boldsymbol{h}}^{\mathsf{L}}, \widehat{\boldsymbol{v}}^{\mathsf{L}}\right)(\mathcal{L};\boldsymbol{\Lambda}) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \mathcal{L}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}_{\mathsf{L}}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

Exemple

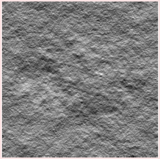


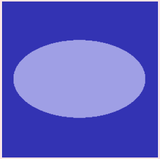


Recherche systématique des paramètres de régularisation

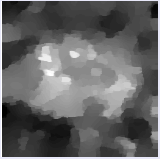
$$\left(\widehat{\boldsymbol{h}}^{\mathsf{L}},\widehat{\boldsymbol{v}}^{\mathsf{L}}\right)(\boldsymbol{\mathcal{L}};\boldsymbol{\Lambda}) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_a \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}_{\mathsf{L}}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

Exemple





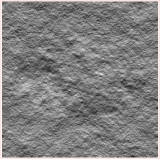
$\widehat{\boldsymbol{h}}^{\mathsf{L}}(\boldsymbol{\mathcal{L}};\lambda^\dagger,\alpha^\dagger)$ (grille)

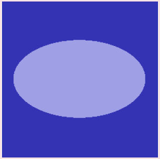


Recherche systématique des paramètres de régularisation

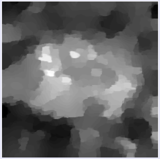
$$\left(\widehat{\boldsymbol{h}}^{\mathsf{L}},\widehat{\boldsymbol{v}}^{\mathsf{L}}\right)(\boldsymbol{\mathcal{L}};\boldsymbol{\Lambda}) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}_{\mathsf{L}}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

Exemple

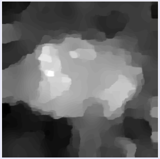




$\widehat{\boldsymbol{h}}^{\mathsf{L}}(\boldsymbol{\mathcal{L}};\lambda^\dagger,\alpha^\dagger)$ (grille)



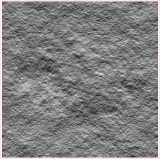
$\widehat{\boldsymbol{h}}^{\mathsf{L}}(\boldsymbol{\mathcal{L}};\widehat{\lambda}^\dagger,\widehat{\alpha}^\dagger)$ (grille)

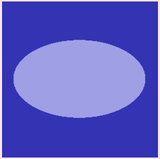


Recherche systématique des paramètres de régularisation

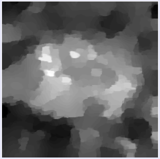
$$\left(\widehat{\boldsymbol{h}}^{\mathrm{L}},\widehat{\boldsymbol{v}}^{\mathrm{L}}\right)(\boldsymbol{\mathcal{L}};\boldsymbol{\Lambda}) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}_{\mathrm{L}}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

Exemple

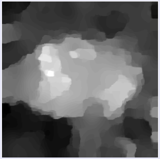




$\widehat{\boldsymbol{h}}^{\mathrm{L}}(\boldsymbol{\mathcal{L}};\lambda^{\dagger},\alpha^{\dagger})$ (grille)



$\widehat{\boldsymbol{h}}^{\mathrm{L}}(\boldsymbol{\mathcal{L}};\widehat{\lambda}^{\dagger},\widehat{\alpha}^{\dagger})$ (grille)



$15 \times 15 =$ **225** paramètres $\longrightarrow$ recherche sur grille très coûteuse

Minimisation automatique de l'estimateur SURE

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N$, $\boldsymbol{\Phi} : \mathbb{R}^{P \times N}$ et $\boldsymbol{\zeta} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\mathcal{S}})$

SURE généralisé DFMC $\lim_{\nu \to 0} \mathbb{E}_{\boldsymbol{\zeta},\boldsymbol{\varepsilon}} \widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{y}; \boldsymbol{\Lambda} \,|\, \boldsymbol{\mathcal{S}}) = R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda})$

Minimisation automatique de l'estimateur SURE

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N, \ \boldsymbol{\Phi} : \mathbb{R}^{P\times N}$ et $\boldsymbol{\zeta} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\mathcal{S}})$

SURE généralisé DFMC $\lim_{\nu\to 0} \mathbb{E}_{\boldsymbol{\zeta},\boldsymbol{\varepsilon}} \widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{y}; \boldsymbol{\Lambda} \,|\, \boldsymbol{\mathcal{S}}) = R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda})$

But : $\underset{\boldsymbol{\Lambda}}{\operatorname{minimiser}} \ \widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{y}; \boldsymbol{\Lambda} \,|\, \boldsymbol{\mathcal{S}})$ pour $\boldsymbol{y}$, $\boldsymbol{\mathcal{S}}$ donnés

Minimisation automatique de l'estimateur SURE

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N,\ \boldsymbol{\Phi} : \mathbb{R}^{P\times N}$ et $\boldsymbol{\zeta} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\mathcal{S}})$

SURE généralisé DFMC $\lim\limits_{\nu\to 0} \mathbb{E}_{\boldsymbol{\zeta},\boldsymbol{\varepsilon}} \widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{y}; \boldsymbol{\Lambda} \,|\, \boldsymbol{\mathcal{S}}) = R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda})$

But : $\underset{\boldsymbol{\Lambda}}{\mathrm{minimiser}}\ \widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{y}; \boldsymbol{\Lambda} \,|\, \boldsymbol{\mathcal{S}}) \equiv \widehat{R}(\boldsymbol{\Lambda})$ pour $\boldsymbol{y}$, $\boldsymbol{\mathcal{S}}$ donnés

Minimisation automatique de l'estimateur SURE

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P$, $\bar{\boldsymbol{x}} \in \mathbb{R}^N$, $\boldsymbol{\Phi} : \mathbb{R}^{P\times N}$ et $\boldsymbol{\zeta} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\mathcal{S}})$

SURE généralisé DFMC $\lim_{\nu\to 0} \mathbb{E}_{\boldsymbol{\zeta},\boldsymbol{\varepsilon}} \widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{y}; \boldsymbol{\Lambda} \,|\, \boldsymbol{\mathcal{S}}) = R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda})$

But : $\underset{\boldsymbol{\Lambda}}{\text{minimiser}}\ \widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{y}; \boldsymbol{\Lambda} \,|\, \boldsymbol{\mathcal{S}}) \equiv \widehat{R}(\boldsymbol{\Lambda})$ pour $\boldsymbol{y}$, $\boldsymbol{\mathcal{S}}$ donnés

Quasi-Newton de Broyden-Fletcher-Goldfarb-Shanno *(Nocedal, 2006)*

for $t = 0, 1, \ldots$

$\boldsymbol{d}^{[t]} = -\boldsymbol{H}^{[t]} \partial_{\boldsymbol{\Lambda}} \widehat{R}(\boldsymbol{\Lambda}^{[t]})$ *direction de descente*

$\alpha^{[t]} \in \underset{\alpha\in\mathbb{R}}{\text{Argmin}}\ \widehat{R}(\boldsymbol{\Lambda}^{[t]} + \alpha \boldsymbol{d}^{[t]})$ *recherche sur une ligne*

$\boldsymbol{\Lambda}^{[t+1]} = \boldsymbol{\Lambda}^{[t]} + \alpha^{[t]} \boldsymbol{d}^{[t]}$

$\boldsymbol{u}^{[t]} = \partial_{\boldsymbol{\Lambda}} \widehat{R}(\boldsymbol{\Lambda}^{[t+1]}) - \partial_{\boldsymbol{\Lambda}} \widehat{R}(\boldsymbol{\Lambda}^{[t]})$ *variation du gradient*

$\boldsymbol{H}^{[t+1]} = \text{BFGS}(\boldsymbol{H}^{[t]}, \boldsymbol{d}^{[t]}, \boldsymbol{u}^{[t]})$ *mise à jour "hessienne inverse"*

Minimisation automatique de l'estimateur SURE

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}} \in \mathbb{R}^N,\ \boldsymbol{\Phi} : \mathbb{R}^{P\times N}$ et $\boldsymbol{\zeta} \sim \mathcal{N}(\mathbf{0},\boldsymbol{\mathcal{S}})$

SURE généralisé DFMC $\lim_{\nu\to 0} \mathbb{E}_{\boldsymbol{\zeta},\boldsymbol{\varepsilon}} \widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{y}; \boldsymbol{\Lambda} \,|\, \boldsymbol{\mathcal{S}}) = R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda})$

But : $\operatorname*{minimiser}_{\boldsymbol{\Lambda}} \widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{y}; \boldsymbol{\Lambda} \,|\, \boldsymbol{\mathcal{S}}) \equiv \widehat{R}(\boldsymbol{\Lambda})$ pour $\boldsymbol{y}$, $\boldsymbol{\mathcal{S}}$ donnés

Quasi-Newton de Broyden-Fletcher-Goldfarb-Shanno *(Nocedal, 2006)*

for $t = 0,1,\dots$

$\boldsymbol{d}^{[t]} = -\boldsymbol{H}^{[t]}\partial_{\boldsymbol{\Lambda}}\widehat{R}(\boldsymbol{\Lambda}^{[t]})$ *direction de descente*

$\alpha^{[t]} \in \operatorname*{Argmin}_{\alpha\in\mathbb{R}} \widehat{R}(\boldsymbol{\Lambda}^{[t]} + \alpha \boldsymbol{d}^{[t]})$ *recherche sur une ligne*

$\boldsymbol{\Lambda}^{[t+1]} = \boldsymbol{\Lambda}^{[t]} + \alpha^{[t]}\boldsymbol{d}^{[t]}$

$\boldsymbol{u}^{[t]} = \partial_{\boldsymbol{\Lambda}}\widehat{R}(\boldsymbol{\Lambda}^{[t+1]}) - \partial_{\boldsymbol{\Lambda}}\widehat{R}(\boldsymbol{\Lambda}^{[t]})$ *variation du gradient*

$\boldsymbol{H}^{[t+1]} = \mathrm{BFGS}(\boldsymbol{H}^{[t]},\boldsymbol{d}^{[t]},\boldsymbol{u}^{[t]})$ *mise à jour "hessienne inverse"*

Stein Unbiased GrAdient Risk estimate

SURE généralisé DFMC

$$\widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{y};\boldsymbol{\Lambda}\,|\,\boldsymbol{\mathcal{S}}) = \left\|\mathbf{A}\left(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y}\right)\right\|^2 + \frac{2}{\nu}\left\langle \boldsymbol{\mathcal{S}}\mathbf{A}^\top\boldsymbol{\Pi}\left(\widehat{\boldsymbol{x}}(\boldsymbol{y}+\nu\boldsymbol{\varepsilon};\boldsymbol{\Lambda}) - \widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda})\right), \boldsymbol{\varepsilon}\right\rangle - \operatorname{tr}\left(\mathbf{A}\boldsymbol{\mathcal{S}}\mathbf{A}^\top\right)$$

Stein Unbiased GrAdient Risk estimate

SURE généralisé DFMC

$$\widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{y};\boldsymbol{\Lambda}\,|\,\boldsymbol{\mathcal{S}}) = \left\|\mathbf{A}\left(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y}\right)\right\|^2 + \frac{2}{\nu}\left\langle \boldsymbol{\mathcal{S}}\mathbf{A}^\top\boldsymbol{\Pi}\left(\widehat{\boldsymbol{x}}(\boldsymbol{y}+\nu\boldsymbol{\varepsilon};\boldsymbol{\Lambda}) - \widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda})\right), \boldsymbol{\varepsilon}\right\rangle - \operatorname{tr}\left(\mathbf{A}\boldsymbol{\mathcal{S}}\mathbf{A}^\top\right)$$

SUGAR généralisé Différences Finies Monte Carlo

$$\begin{aligned}\partial_{\boldsymbol{\Lambda}}\widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{y};\boldsymbol{\Lambda}\,|\,\boldsymbol{\mathcal{S}}) &= 2\left(\mathbf{A}\boldsymbol{\Phi}\partial_{\boldsymbol{\Lambda}}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda})\right)^\top \mathbf{A}\left(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y}\right)\\ &\quad + \frac{2}{\nu}\left\langle \boldsymbol{\mathcal{S}}\mathbf{A}^\top\boldsymbol{\Pi}\left(\partial_{\boldsymbol{\Lambda}}\widehat{\boldsymbol{x}}(\boldsymbol{y}+\nu\boldsymbol{\varepsilon};\boldsymbol{\Lambda}) - \partial_{\boldsymbol{\Lambda}}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda})\right), \boldsymbol{\varepsilon}\right\rangle\end{aligned}$$

Stein Unbiased GrAdient Risk estimate

SURE généralisé DFMC

$$\widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{y};\boldsymbol{\Lambda}\,|\,\boldsymbol{\mathcal{S}}) = \left\|\mathbf{A}\left(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y}\right)\right\|^2 + \frac{2}{\nu}\left\langle \boldsymbol{\mathcal{S}}\mathbf{A}^\top\boldsymbol{\Pi}\left(\widehat{\boldsymbol{x}}(\boldsymbol{y}+\nu\boldsymbol{\varepsilon};\boldsymbol{\Lambda}) - \widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda})\right), \boldsymbol{\varepsilon}\right\rangle - \operatorname{tr}\left(\mathbf{A}\boldsymbol{\mathcal{S}}\mathbf{A}^\top\right)$$

SUGAR généralisé Différences Finies Monte Carlo

$$\begin{aligned}\partial_{\boldsymbol{\Lambda}}\widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{y};\boldsymbol{\Lambda}\,|\,\boldsymbol{\mathcal{S}}) &= 2\left(\mathbf{A}\boldsymbol{\Phi}\partial_{\boldsymbol{\Lambda}}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda})\right)^\top\mathbf{A}\left(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda}) - \boldsymbol{y}\right)\\ &\quad + \frac{2}{\nu}\left\langle \boldsymbol{\mathcal{S}}\mathbf{A}^\top\boldsymbol{\Pi}\left(\partial_{\boldsymbol{\Lambda}}\widehat{\boldsymbol{x}}(\boldsymbol{y}+\nu\boldsymbol{\varepsilon};\boldsymbol{\Lambda}) - \partial_{\boldsymbol{\Lambda}}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda})\right), \boldsymbol{\varepsilon}\right\rangle\end{aligned}$$

Théorème *(Pascal, 2020)*

Soit $(\boldsymbol{y};\boldsymbol{\Lambda}) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda})$ un estimateur de $\bar{\boldsymbol{x}}$

- uniformément lipschitzien par rapport à $\boldsymbol{y}$
- tel que $\forall\boldsymbol{\Lambda}\in\mathbb{R}^L$, $\widehat{\boldsymbol{x}}(\mathbf{0}_P;\boldsymbol{\Lambda}) = \mathbf{0}_N$,
- uniformément L-lipschitzien par rapport à $\boldsymbol{\Lambda}$, L indép. de $\boldsymbol{y}$. Alors

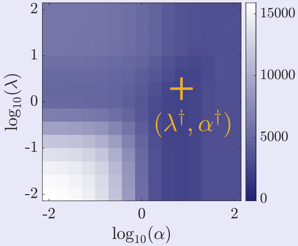
$$\partial_{\boldsymbol{\Lambda}}R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) = \lim_{\nu\to 0}\mathbb{E}_{\boldsymbol{\zeta},\boldsymbol{\varepsilon}}\left[\partial_{\boldsymbol{\Lambda}}\widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{y};\boldsymbol{\Lambda}\,|\,\boldsymbol{\mathcal{S}})\right]$$

Réglage des paramètres (Recherche automatique)

$$\left(\widehat{\boldsymbol{h}},\widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}};\lambda,\alpha) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_a \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

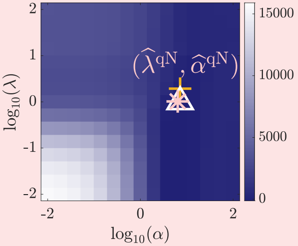
$\bar{\boldsymbol{h}}$: *vraie* régularité

$$\mathcal{R}(\lambda,\alpha) = \left\|\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}};\lambda,\alpha) - \bar{\boldsymbol{h}}\right\|^2$$



$\bar{\boldsymbol{h}}$: inconnue !

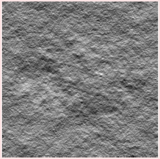
$$\widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{\mathcal{L}};\lambda,\alpha|\,\boldsymbol{\mathcal{S}})$$

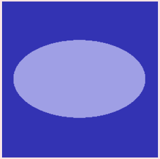


Recherche automatique des paramètres de régularisation

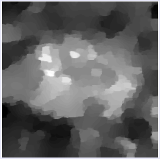
$$\left(\widehat{\boldsymbol{h}}^{\mathsf{L}},\widehat{\boldsymbol{v}}^{\mathsf{L}}\right)(\boldsymbol{\mathcal{L}};\boldsymbol{\Lambda}) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}_{\mathsf{L}}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

Exemple

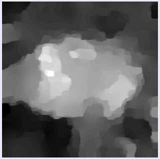




$\widehat{\boldsymbol{h}}^{\mathsf{L}}(\boldsymbol{\mathcal{L}};\lambda^\dagger,\alpha^\dagger)$ (grille)



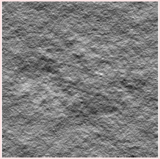
$\widehat{\boldsymbol{h}}^{\mathsf{L}}(\boldsymbol{\mathcal{L}};\widehat{\lambda}^\dagger,\widehat{\alpha}^\dagger)$ (grille)

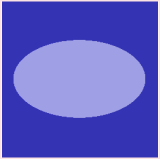


Recherche automatique des paramètres de régularisation

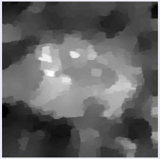
$$\left(\widehat{\boldsymbol{h}}^{\mathrm{L}},\widehat{\boldsymbol{v}}^{\mathrm{L}}\right)(\boldsymbol{\mathcal{L}};\boldsymbol{\Lambda}) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}_{\mathrm{L}}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

Exemple

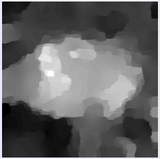




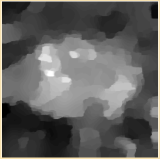
$\widehat{\boldsymbol{h}}^{\mathrm{L}}(\boldsymbol{\mathcal{L}};\lambda^{\dagger},\alpha^{\dagger})$ (grille)



$\widehat{\boldsymbol{h}}^{\mathrm{L}}(\boldsymbol{\mathcal{L}};\widehat{\lambda}^{\dagger},\widehat{\alpha}^{\dagger})$ (grille)



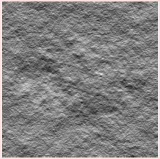
$\widehat{\boldsymbol{h}}^{\mathrm{L}}(\boldsymbol{\mathcal{L}};\widehat{\lambda}^{\mathrm{qN}},\widehat{\alpha}^{\mathrm{qN}})$ (quasi-Newton)

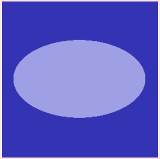


Recherche automatique des paramètres de régularisation

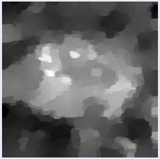
$$\left(\widehat{\boldsymbol{h}}^{\mathrm{L}},\widehat{\boldsymbol{v}}^{\mathrm{L}}\right)(\boldsymbol{\mathcal{L}};\boldsymbol{\Lambda}) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}_{\mathrm{L}}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)$$

Exemple

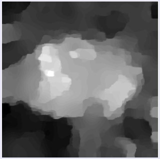




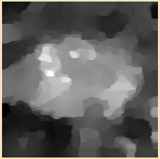
$\widehat{\boldsymbol{h}}^{\mathrm{L}}(\boldsymbol{\mathcal{L}};\lambda^{\dagger},\alpha^{\dagger})$ (grille)



$\widehat{\boldsymbol{h}}^{\mathrm{L}}(\boldsymbol{\mathcal{L}};\widehat{\lambda}^{\dagger},\widehat{\alpha}^{\dagger})$ (grille)



$\widehat{\boldsymbol{h}}^{\mathrm{L}}(\boldsymbol{\mathcal{L}};\widehat{\lambda}^{\mathrm{qN}},\widehat{\alpha}^{\mathrm{qN}})$ (quasi-Newton)



40 appels de l'estimateurs v.s. **225** sur une grille

Bilan de la présentation

- Régularité et variance locale [ICIP, 2018]

Bilan de la présentation

- Régularité et variance locale [ICIP, 2018]
  - aptes à caractériser des textures réelles

Bilan de la présentation

- Régularité et variance locale [ICIP, 2018]
  - aptes à caractériser des textures réelles
  - attributs complémentaires → capacité à discriminer finement

Bilan de la présentation

- Régularité et variance locale [ICIP, 2018]
  - aptes à caractériser des textures réelles
  - attributs complémentaires → capacité à discriminer finement
- Estimation et régularisation simultanées [ACHA, 2019]

Bilan de la présentation

- Régularité et variance locale [ICIP, 2018]
  - aptes à caractériser des textures réelles
  - attributs complémentaires $\rightarrow$ capacité à discriminer finement
- Estimation et régularisation simultanées [ACHA, 2019]
  - forte diminution de l'erreur d'estimation

Bilan de la présentation

- Régularité et variance locale [ICIP, 2018]
  - aptes à caractériser des textures réelles
  - attributs complémentaires → capacité à discriminer finement
- Estimation et régularisation simultanées [ACHA, 2019]
  - forte diminution de l'erreur d'estimation
  - contours précis et réguliers grâce à la pénalisation co-localisée

Bilan de la présentation

- Régularité et variance locale [ICIP, 2018]
  - aptes à caractériser des textures réelles
  - attributs complémentaires $\rightarrow$ capacité à discriminer finement
- Estimation et régularisation simultanées [ACHA, 2019]
  - forte diminution de l'erreur d'estimation
  - contours précis et réguliers grâce à la pénalisation co-localisée
- Algorithmes rapides et réglage automatique des paramètres [JMIV, 2020]

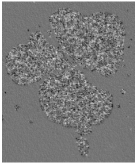
Bilan de la présentation

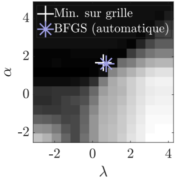
- Régularité et variance locale [ICIP, 2018]
  - aptes à caractériser des textures réelles
  - attributs complémentaires → capacité à discriminer finement
- Estimation et régularisation simultanées [ACHA, 2019]
  - forte diminution de l'erreur d'estimation
  - contours précis et réguliers grâce à la pénalisation co-localisée
- Algorithmes rapides et réglage automatique des paramètres [JMIV, 2020]
  - possibilité de traiter de gros volumes de données

Bilan de la présentation

- Régularité et variance locale [ICIP, 2018]
  - aptes à caractériser des textures réelles
  - attributs complémentaires $\rightarrow$ capacité à discriminer finement
- Estimation et régularisation simultanées [ACHA, 2019]
  - forte diminution de l'erreur d'estimation
  - contours précis et réguliers grâce à la pénalisation co-localisée
- Algorithmes rapides et réglage automatique des paramètres [JMIV, 2020]
  - possibilité de traiter de gros volumes de données
  - objectivité et reproductibilité

$\longrightarrow$ **En cours** : traitement automatisé de séries temporelles issues de l'étude des écoulements multiphasiques [Ann. Telecom, 2020]





Autres contributions

- Synthèse de textures monofractales par morceaux [ACHA, 2019]

Autres contributions

- Synthèse de textures monofractales par morceaux [ACHA, 2019]
  - segmentation et attributs prescrits $\rightarrow$ performances

Autres contributions

- Synthèse de textures monofractales par morceaux [ACHA, 2019]
  - segmentation et attributs prescrits → performances
  - possibilité de générer de grosses bases de données

Autres contributions

- Synthèse de textures monofractales par morceaux [ACHA, 2019]
  - segmentation et attributs prescrits → performances
  - possibilité de générer de grosses bases de données
- Comparaison d'algorithmes proximaux [ACHA, 2019]

Autres contributions

- Synthèse de textures monofractales par morceaux [ACHA, 2019]
  - segmentation et attributs prescrits → performances
  - possibilité de générer de grosses bases de données
- Comparaison d'algorithmes proximaux [ACHA, 2019]
  - stratégie FISTA pour accélérer l'algorithme forward-backward

Autres contributions

- Synthèse de textures monofractales par morceaux [ACHA, 2019]
  - segmentation et attributs prescrits → performances
  - possibilité de générer de grosses bases de données
- Comparaison d'algorithmes proximaux [ACHA, 2019]
  - stratégie FISTA pour accélérer l'algorithme forward-backward
  - calcul du gap de dualité et choix d'un critère d'arrêt

Autres contributions

- Synthèse de textures monofractales par morceaux [ACHA, 2019]
  - segmentation et attributs prescrits → performances
  - possibilité de générer de grosses bases de données
- Comparaison d'algorithmes proximaux [ACHA, 2019]
  - stratégie FISTA pour accélérer l'algorithme forward-backward
  - calcul du gap de dualité et choix d'un critère d'arrêt
- Réseaux convolutionnels pour la segmentation de textures [EUSIPCO, 2020]

Autres contributions

- Synthèse de textures monofractales par morceaux [ACHA, 2019]
  - segmentation et attributs prescrits → performances
  - possibilité de générer de grosses bases de données
- Comparaison d'algorithmes proximaux [ACHA, 2019]
  - stratégie FISTA pour accélérer l'algorithme forward-backward
  - calcul du gap de dualité et choix d'un critère d'arrêt
- Réseaux convolutionnels pour la segmentation de textures [EUSIPCO, 2020]
  - performance et robustesse

Autres contributions

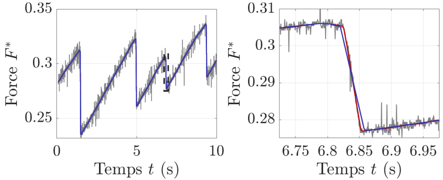
- Synthèse de textures monofractales par morceaux [ACHA, 2019]
  - segmentation et attributs prescrits → performances
  - possibilité de générer de grosses bases de données
- Comparaison d'algorithmes proximaux [ACHA, 2019]
  - stratégie FISTA pour accélérer l'algorithme forward-backward
  - calcul du gap de dualité et choix d'un critère d'arrêt
- Réseaux convolutionnels pour la segmentation de textures [EUSIPCO, 2020]
  - performance et robustesse
  - comparaison coûts de calcul et mémoire

Autres contributions

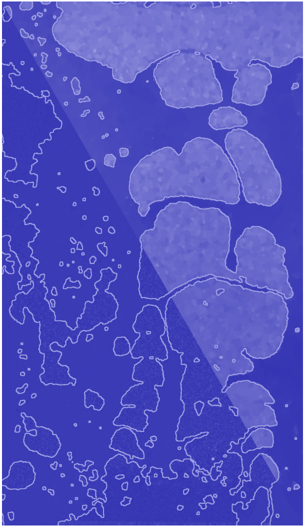
- Synthèse de textures monofractales par morceaux [ACHA, 2019]
  - segmentation et attributs prescrits → performances
  - possibilité de générer de grosses bases de données
- Comparaison d'algorithmes proximaux [ACHA, 2019]
  - stratégie FISTA pour accélérer l'algorithme forward-backward
  - calcul du gap de dualité et choix d'un critère d'arrêt
- Réseaux convolutionnels pour la segmentation de textures [EUSIPCO, 2020]
  - performance et robustesse
  - comparaison coûts de calcul et mémoire
- Application à la physique du formalisme de Stein généralisé [Ann. Telecom, 2020]
  - segmentation de texture → écoulements multiphasiques

Autres contributions

- Synthèse de textures monofractales par morceaux [ACHA, 2019]
  - segmentation et attributs prescrits → performances
  - possibilité de générer de grosses bases de données
- Comparaison d'algorithmes proximaux [ACHA, 2019]
  - stratégie FISTA pour accélérer l'algorithme forward-backward
  - calcul du gap de dualité et choix d'un critère d'arrêt
- Réseaux convolutionnels pour la segmentation de textures [EUSIPCO, 2020]
  - performance et robustesse
  - comparaison coûts de calcul et mémoire
- Application à la physique du formalisme de Stein généralisé [Ann. Telecom, 2020]
  - segmentation de texture → écoulements multiphasiques
  - débruitage linéaire par morceaux → frottement solide



Merci



Grazie

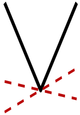
Gracias

Thank you

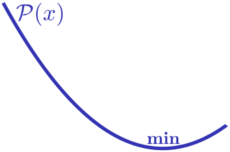
Définition du gap de dualité

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \sum_{a} \underbrace{\left\|\log \boldsymbol{\mathcal{L}}_{a,.} - \log(a)\boldsymbol{h} - \boldsymbol{v}\right\|^2}_{\textbf{Moindres Carrés}} \quad + \quad \lambda\, \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)}_{\textbf{Variation Totale}}$$

non lisse



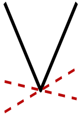
Primal $\quad \min_{\boldsymbol{x}} \; \mathrm{MC}(\boldsymbol{x}) + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{x})$



Définition du gap de dualité

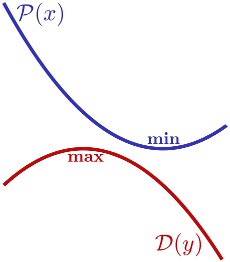
$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \quad \sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \quad + \quad \lambda \; \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h},\mathbf{D}\boldsymbol{v};\alpha)}_{\textbf{Variation Totale}}$$

non lisse



Primal $\quad \min_{\boldsymbol{x}} \; \mathrm{MC}(\boldsymbol{x}) + \lambda\mathcal{Q}(\mathbf{D}\boldsymbol{x})$

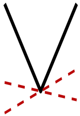
Dual $\quad \max_{\boldsymbol{y}} \; -\mathrm{MC}^*(-\mathbf{D}^\top \boldsymbol{y}) - (\lambda\mathcal{Q})^*(\boldsymbol{y})$



Définition du gap de dualité

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \quad \underbrace{\sum_a \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Variation Totale}}$$

non lisse

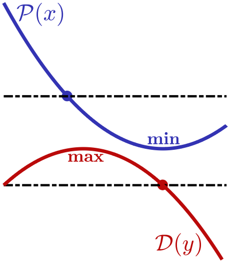


Primal $\min_{\boldsymbol{x}} \ \mathrm{MC}(\boldsymbol{x}) + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{x})$

Dual $\max_{\boldsymbol{y}} \ -\mathrm{MC}^*(-\mathbf{D}^\top \boldsymbol{y}) - (\lambda\mathcal{Q})^*(\boldsymbol{y})$

Proposition *(Bauschke, 2011)*

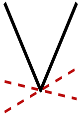
Soit $\delta(\boldsymbol{x};\boldsymbol{y}) \triangleq \mathcal{P}(\boldsymbol{x}) - \mathcal{D}(\boldsymbol{y})$ le gap de dualité,



Définition du gap de dualité

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimiser}} \ \sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Moindres Carrés}} \ + \ \lambda \ \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Variation Totale}}$$

non lisse



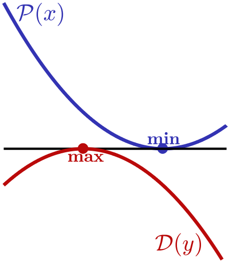
Primal $\min_{\boldsymbol{x}} \ \mathrm{MC}(\boldsymbol{x}) + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{x})$

Dual $\max_{\boldsymbol{y}} \ -\mathrm{MC}^*(-\mathbf{D}^\top \boldsymbol{y}) - (\lambda\mathcal{Q})^*(\boldsymbol{y})$

Proposition *(Bauschke, 2011)*

Soit $\delta(\boldsymbol{x}; \boldsymbol{y}) \triangleq \mathcal{P}(\boldsymbol{x}) - \mathcal{D}(\boldsymbol{y})$ le gap de dualité,

$$\delta(\widehat{\boldsymbol{x}}; \widehat{\boldsymbol{y}}) = \mathcal{P}(\widehat{\boldsymbol{x}}) - \mathcal{D}(\widehat{\boldsymbol{y}}) = 0$$



Conjuguée convexe des moindres carrés

$$\mathrm{MC}^*(\boldsymbol{h},\boldsymbol{v}) = \sup_{\widetilde{\boldsymbol{h}},\widetilde{\boldsymbol{v}}} \langle\widetilde{\boldsymbol{h}},\boldsymbol{h}\rangle + \langle\widetilde{\boldsymbol{v}},\boldsymbol{v}\rangle - \mathrm{MC}(\widetilde{\boldsymbol{h}},\widetilde{\boldsymbol{v}}) \underset{\text{(si le sup est atteint)}}{=} \langle\bar{\boldsymbol{h}},\boldsymbol{h}\rangle + \langle\bar{\boldsymbol{v}},\boldsymbol{v}\rangle - \mathrm{MC}(\bar{\boldsymbol{h}},\bar{\boldsymbol{v}}).$$

Conjuguée convexe des moindres carrés

$$\mathrm{MC}^*(\boldsymbol{h},\boldsymbol{v}) = \sup_{\widetilde{\boldsymbol{h}},\widetilde{\boldsymbol{v}}} \langle \widetilde{\boldsymbol{h}},\boldsymbol{h}\rangle + \langle \widetilde{\boldsymbol{v}},\boldsymbol{v}\rangle - \mathrm{MC}(\widetilde{\boldsymbol{h}},\widetilde{\boldsymbol{v}}) \underset{\text{(si le sup est atteint)}}{=} \langle \bar{\boldsymbol{h}},\boldsymbol{h}\rangle + \langle \bar{\boldsymbol{v}},\boldsymbol{v}\rangle - \mathrm{MC}(\bar{\boldsymbol{h}},\bar{\boldsymbol{v}}).$$

Condition d'optimalité

$$\begin{cases} \boldsymbol{h} - 2\sum_a \log(a)\left(\bar{\boldsymbol{v}} + \log(a)\bar{\boldsymbol{h}} - \log \boldsymbol{\mathcal{L}}_{a,.}\right) = 0 \\ \boldsymbol{v} - 2\sum_a \left(\bar{\boldsymbol{v}} + \log(a)\bar{\boldsymbol{h}} - \log \boldsymbol{\mathcal{L}}_{a,.}\right) = 0 \end{cases}$$

Conjuguée convexe des moindres carrés

$$\mathrm{MC}^*(\boldsymbol{h},\boldsymbol{v}) = \sup_{\widetilde{\boldsymbol{h}},\widetilde{\boldsymbol{v}}} \langle \widetilde{\boldsymbol{h}},\boldsymbol{h}\rangle + \langle \widetilde{\boldsymbol{v}},\boldsymbol{v}\rangle - \mathrm{MC}(\widetilde{\boldsymbol{h}},\widetilde{\boldsymbol{v}}) \underset{\text{(si le sup est atteint)}}{=} \langle \bar{\boldsymbol{h}},\boldsymbol{h}\rangle + \langle \bar{\boldsymbol{v}},\boldsymbol{v}\rangle - \mathrm{MC}(\bar{\boldsymbol{h}},\bar{\boldsymbol{v}}).$$

Condition d'optimalité

$$\begin{cases} \boldsymbol{h} - 2\sum_a \log(a)\left(\bar{\boldsymbol{v}} + \log(a)\bar{\boldsymbol{h}} - \log \mathcal{L}_{a,.}\right) = 0 \\ \boldsymbol{v} - 2\sum_a \left(\bar{\boldsymbol{v}} + \log(a)\bar{\boldsymbol{h}} - \log \mathcal{L}_{a,.}\right) = 0 \end{cases} \iff \boldsymbol{\Phi}^*\boldsymbol{\Phi}\begin{pmatrix} \bar{\boldsymbol{h}} \\ \bar{\boldsymbol{v}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{h}/2 + \mathcal{G} \\ \boldsymbol{v}/2 + \mathcal{T} \end{pmatrix}$$

$$\mathcal{T} = \sum_a \log \mathcal{L}_{a,.} \quad \text{and} \quad \mathcal{G} = \sum_a \log(a) \log \mathcal{L}_{a,.},$$

Conjuguée convexe des moindres carrés

$$\mathrm{MC}^*(\boldsymbol{h},\boldsymbol{v}) = \sup_{\widetilde{\boldsymbol{h}},\widetilde{\boldsymbol{v}}} \langle \widetilde{\boldsymbol{h}},\boldsymbol{h}\rangle + \langle \widetilde{\boldsymbol{v}},\boldsymbol{v}\rangle - \mathrm{MC}(\widetilde{\boldsymbol{h}},\widetilde{\boldsymbol{v}}) \underset{\text{(si le sup est atteint)}}{=} \langle \bar{\boldsymbol{h}},\boldsymbol{h}\rangle + \langle \bar{\boldsymbol{v}},\boldsymbol{v}\rangle - \mathrm{MC}(\bar{\boldsymbol{h}},\bar{\boldsymbol{v}}).$$

Condition d'optimalité

$$\begin{cases} \boldsymbol{h} - 2\sum_a \log(a)\left(\bar{\boldsymbol{v}} + \log(a)\bar{\boldsymbol{h}} - \log \boldsymbol{\mathcal{L}}_{a,.}\right) = 0 \\ \boldsymbol{v} - 2\sum_a \left(\bar{\boldsymbol{v}} + \log(a)\bar{\boldsymbol{h}} - \log \boldsymbol{\mathcal{L}}_{a,.}\right) = 0 \end{cases} \iff \boldsymbol{\Phi}^*\boldsymbol{\Phi}\begin{pmatrix}\bar{\boldsymbol{h}} \\ \bar{\boldsymbol{v}}\end{pmatrix} = \begin{pmatrix}\boldsymbol{h}/2 + \boldsymbol{\mathcal{G}} \\ \boldsymbol{v}/2 + \boldsymbol{\mathcal{T}}\end{pmatrix}$$

$$\boldsymbol{\mathcal{T}} = \sum_a \log \boldsymbol{\mathcal{L}}_{a,.} \quad \text{and} \quad \boldsymbol{\mathcal{G}} = \sum_a \log(a) \log \boldsymbol{\mathcal{L}}_{a,.},$$

$$\forall m = \{0,1,2\}, S_m = \sum_a (\log a)^m, \quad \boldsymbol{\Phi}^*\boldsymbol{\Phi} = \begin{pmatrix} S_2\boldsymbol{I} & S_1\boldsymbol{I} \\ S_1\boldsymbol{I} & S_0\boldsymbol{I}\end{pmatrix}$$

Conjuguée convexe des moindres carrés

$$\mathrm{MC}^*(\boldsymbol{h},\boldsymbol{v}) = \sup_{\widetilde{\boldsymbol{h}},\widetilde{\boldsymbol{v}}} \langle \widetilde{\boldsymbol{h}},\boldsymbol{h}\rangle + \langle \widetilde{\boldsymbol{v}},\boldsymbol{v}\rangle - \mathrm{MC}(\widetilde{\boldsymbol{h}},\widetilde{\boldsymbol{v}}) \underset{\text{(si le sup est atteint)}}{=} \langle \bar{\boldsymbol{h}},\boldsymbol{h}\rangle + \langle \bar{\boldsymbol{v}},\boldsymbol{v}\rangle - \mathrm{MC}(\bar{\boldsymbol{h}},\bar{\boldsymbol{v}}).$$

Condition d'optimalité

$$\left\{ \begin{array}{l} \boldsymbol{h} - 2\sum_a \log(a)\left(\bar{\boldsymbol{v}} + \log(a)\bar{\boldsymbol{h}} - \log \boldsymbol{\mathcal{L}}_{a,.}\right) = 0 \\ \boldsymbol{v} - 2\sum_a \left(\bar{\boldsymbol{v}} + \log(a)\bar{\boldsymbol{h}} - \log \boldsymbol{\mathcal{L}}_{a,.}\right) = 0 \end{array} \right. \iff \boldsymbol{\Phi}^*\boldsymbol{\Phi} \begin{pmatrix} \bar{\boldsymbol{h}} \\ \bar{\boldsymbol{v}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{h}/2 + \boldsymbol{\mathcal{G}} \\ \boldsymbol{v}/2 + \boldsymbol{\mathcal{T}} \end{pmatrix}$$

$$\boldsymbol{\mathcal{T}} = \sum_a \log \boldsymbol{\mathcal{L}}_{a,.} \quad \text{and} \quad \boldsymbol{\mathcal{G}} = \sum_a \log(a) \log \boldsymbol{\mathcal{L}}_{a,.},$$

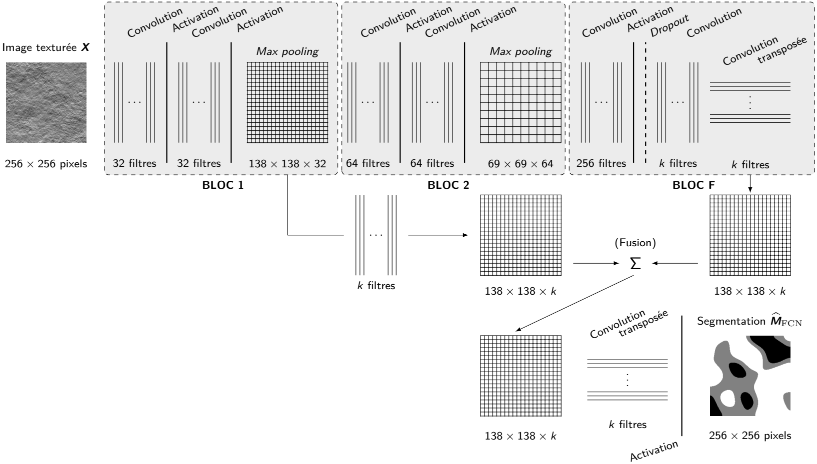
$$\forall m = \{0,1,2\}, S_m = \sum_a (\log a)^m, \quad \boldsymbol{\Phi}^*\boldsymbol{\Phi} = \begin{pmatrix} S_2 \boldsymbol{I} & S_1 \boldsymbol{I} \\ S_1 \boldsymbol{I} & S_0 \boldsymbol{I} \end{pmatrix}$$

$$\mathrm{MC}^*(\boldsymbol{h},\boldsymbol{v}) = \frac{1}{4}\langle(\boldsymbol{h},\boldsymbol{v}),(\boldsymbol{\Phi}^*\boldsymbol{\Phi})^{-1}(\boldsymbol{h},\boldsymbol{v})\rangle + \langle(\boldsymbol{\mathcal{G}},\boldsymbol{\mathcal{T}}),(\boldsymbol{\Phi}^*\boldsymbol{\Phi})^{-1}(\boldsymbol{h},\boldsymbol{v})\rangle + \mathcal{C}$$

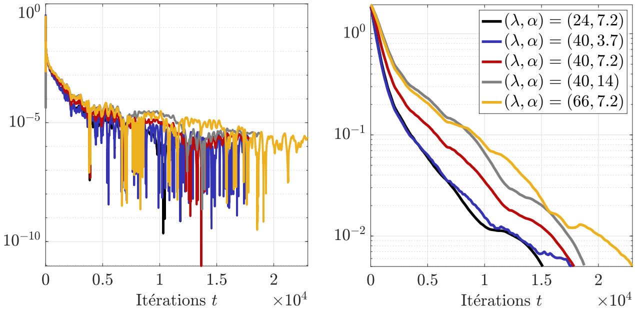
où $\mathcal{C}$ est une constante dépendant uniquement de $\boldsymbol{\mathcal{L}}$.

Architecture pour la segmentation de texture

Avec connexions *résiduelles*



Critère d'arrêt



Performances de segmentation
Configuration I

	2 classes	3 classes	4 classes
Entraîné sur la Config. I, testé sur la Config. I			
Segmentation à contours « libres »	$93{,}2 \pm 0{,}8\%$	$69{,}3 \pm 2{,}8\%$	$58{,}6 \pm 1{,}5\%$
Entraîné sur 2000 images			
Réseau à $8 \cdot 10^7$ poids / $P = 2000$	$97.3 \pm 0{,}6\%$	$97{,}8 \pm 0{,}3\%$	$97{,}1 \pm 0{,}4\%$
Réseau à $2 \cdot 10^6$ poids / $P = 2000$	$97{,}4 \pm 0{,}6\%$	$98.1 \pm 0{,}3\%$	$96{,}8 \pm 0{,}5\%$
Réseau à $4 \cdot 10^5$ poids / $P = 2000$	$96{,}9 \pm 0{,}7\%$	$98{,}0 \pm 0{,}3\%$	$96.5 \pm 0{,}5\%$
Entraîné sur 20 images			
Réseau à $8 \cdot 10^7$ poids / $P = 20$	$95{,}5 \pm 0{,}9\%$	$97{,}5 \pm 0{,}4\%$	$95{,}4 \pm 0{,}8\%$
Réseau à $2 \cdot 10^6$ poids / $P = 20$	$95{,}4 \pm 1{,}1\%$	$97{,}4 \pm 0{,}5\%$	$95{,}9 \pm 0{,}7\%$
Réseau à $4 \cdot 10^5$ poids / $P = 20$	$96{,}6 \pm 0{,}7\%$	$98{,}0 \pm 0{,}4\%$	$96{,}5 \pm 0{,}5\%$

Performances de segmentation
Configuration II

	2 classes	3 classes	4 classes
Entraîné sur la Config. II, testé sur la Config. II			
Segmentation à contours « libres »	$97{,}8 \pm 0{,}2\%$	$95{,}2 \pm 3{,}1\%$	$64{,}9 \pm 1{,}4\%$
Entraîné sur 2000 images			
Réseau à $8 \cdot 10^7$ poids / $P = 2000$	$99{,}1 \pm 0{,}2\%$	$98{,}3 \pm 0{,}3\%$	$95{,}7 \pm 0{,}5\%$
Réseau à $2 \cdot 10^6$ poids / $P = 2000$	$99{,}0 \pm 0{,}2\%$	$98{,}5 \pm 0{,}3\%$	$95{,}6 \pm 0{,}5\%$
Réseau à $4 \cdot 10^5$ poids / $P = 2000$	$99{,}1 \pm 0{,}2\%$	$98{,}4 \pm 0{,}3\%$	$95{,}2 \pm 0{,}6\%$
Entraîné sur 20 images			
Réseau à $8 \cdot 10^7$ poids / $P = 20$	$98{,}8 \pm 0{,}2\%$	$97{,}9 \pm 0{,}3\%$	$94{,}5 \pm 0{,}7\%$
Réseau à $2 \cdot 10^6$ poids / $P = 20$	$98{,}6 \pm 0{,}3\%$	$97{,}4 \pm 0{,}4\%$	$93{,}0 \pm 0{,}9\%$
Réseau à $4 \cdot 10^5$ poids / $P = 20$	$98{,}8 \pm 0{,}3\%$	$98{,}3 \pm 0{,}3\%$	$94{,}8 \pm 0{,}6\%$

Robustesse

Entraîné sur la Config. I, testé sur la Config. II

	2 classes	3 classes	4 classes
Entraîné sur la Config. I, testé sur la Config. II			
Segmentation à contours « libres »	$79{,}2 \pm 2{,}9\%$	$95{,}2 \pm 1{,}2\%$	$66{,}3 \pm 1{,}1\%$
Entraîné sur 2000 images			
Réseau à $8 \cdot 10^7$ poids / $P = 2000$	$91{,}2 \pm 2{,}1\%$	$65{,}7 \pm 7{,}2\%$	$55{,}6 \pm 3{,}4\%$
Réseau à $2 \cdot 10^6$ poids / $P = 2000$	$87{,}9 \pm 2{,}5\%$	$69{,}0 \pm 7{,}6\%$	$50{,}8 \pm 4{,}0\%$
Réseau à $4 \cdot 10^5$ poids / $P = 2000$	$81{,}8 \pm 3{,}8\%$	$65{,}2 \pm 7{,}2\%$	$46{,}4 \pm 3{,}7\%$
Entraîné sur 20 images			
Réseau à $8 \cdot 10^7$ poids / $P = 20$	$91{,}4 \pm 1{,}6\%$	$63{,}3 \pm 7{,}1\%$	$54{,}7 \pm 3{,}3\%$
Réseau à $2 \cdot 10^6$ poids / $P = 20$	$92{,}4 \pm 1{,}6\%$	$65{,}6 \pm 7{,}4\%$	$44{,}4 \pm 3{,}4\%$
Réseau à $4 \cdot 10^5$ poids / $P = 20$	$86{,}3 \pm 2{,}6\%$	$64{,}9 \pm 7{,}2\%$	$48{,}4 \pm 3{,}8\%$

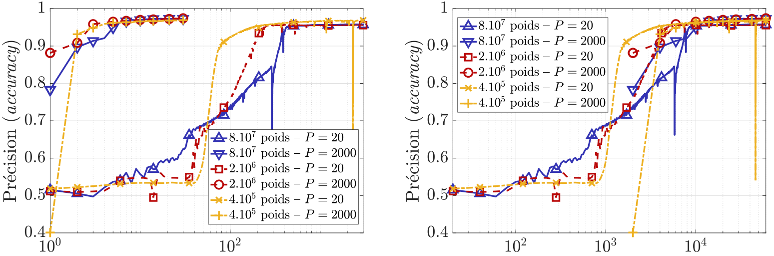
Robustesse

Entraîné sur la Config. II, testé sur la Config. I

	2 classes	3 classes	4 classes
Entraîné sur la Config. II, testé sur la Config. I			
Segmentation à contours « libres »	$90{,}9 \pm 2{,}8\%$	$66{,}7 \pm 2{,}5\%$	$52{,}0 \pm 1{,}5\%$
Entraîné sur 2000 images			
Réseau à $8 \cdot 10^7$ poids / $P = 2000$	$56{,}2 \pm 13{,}5\%$	$73{,}5 \pm 8{,}2\%$	$50{,}9 \pm 3{,}9\%$
Réseau à $2 \cdot 10^6$ poids / $P = 2000$	$55{,}1 \pm 14.0\%$	$74{,}9 \pm 8{,}2\%$	$51{,}3 \pm 4{,}3\%$
Réseau à $5\,10^5$ poids / $P = 2000$	$55{,}5 \pm 13{,}8\%$	$72{,}6 \pm 8{,}1\%$	$50{,}2 \pm 3.8\%$
Entraîné sur 20 images			
Réseau à $8 \cdot 10^7$ poids / $P = 20$	$57{,}1 \pm 13{,}3\%$	$71{,}1 \pm 8{,}2\%$	$52{,}6 \pm 3{,}8\%$
Réseau à $2 \cdot 10^6$ poids / $P = 20$	$55{,}3 \pm 14{,}0\%$	$71{,}7 \pm 8{,}4\%$	$49{,}6 \pm 4{,}2\%$
Réseau à $5\,10^5$ poids / $P = 20$	$62{,}3 \pm 11{,}5\%$	$71{,}0 \pm 8{,}2\%$	$54{,}1 \pm 3{,}7\%$

Convergence de la phase d'entraînement

Évolution du score de segmentation des trois réseaux au cours de l'entraînement sur la Config. I, avec deux classes $k = 2$

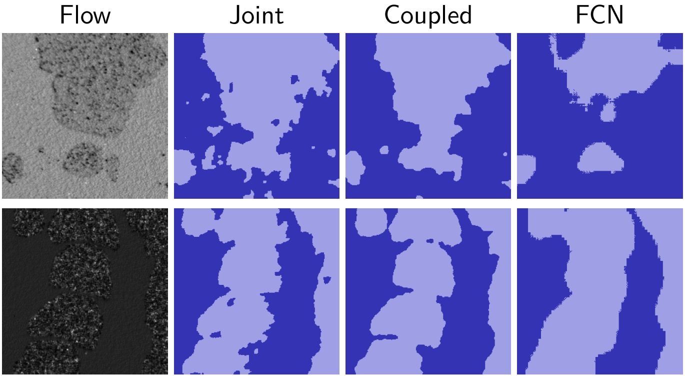


Comparaison effort de calcul $\mathcal{C}$

	$\mathcal{W}$	P	$\mathcal{I}$	$\mathcal{C}$
Segmentation à contours ≪ libres ≫	2	1	10^7	$2\,10^7$
Réseau à $8 \cdot 10^7$ poids / $P = \{20{,}2000\}$	$8 \cdot 10^7$	$\{20{,}2000\}$	$\{3000{,}30\}$	$4{,}8 \cdot 10^{12}$
Réseau à $2 \cdot 10^6$ poids / $P = \{20{,}2000\}$	$2 \cdot 10^6$	$\{20{,}2000\}$	$\{3000{,}30\}$	$1{,}2 \cdot 10^{11}$
Réseau à $4 \cdot 10^5$ poids / $P = \{20{,}2000\}$	$4 \cdot 10^5$	$\{20{,}2000\}$	$\{3000{,}30\}$	$2{,}4 \cdot 10^{10}$

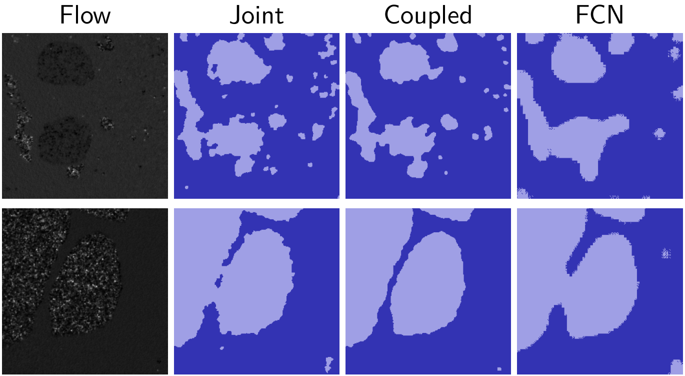
- $\mathcal{W}$: nombre de poids
- P : taille base d'entraînement
- $\mathcal{I}$: nombre d'*epochs*

Réseaux de neurones convolutionnels†



† V. Andrearczyk, https://arxiv.org/abs/1703.05230

Réseaux de neurones convolutionnels†



† V. Andrearczyk, `https://arxiv.org/abs/1703.05230`

Stein Unbiased Risk Estimate (Principe)

Observations $\boldsymbol{y} = \bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P$, $\bar{\boldsymbol{x}}$: vérité et $\boldsymbol{\zeta} \sim \mathcal{N}(\mathbf{0},\rho^2\mathbf{I})$

Stein Unbiased Risk Estimate (Principe)

Observations $\boldsymbol{y} = \bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P$, $\bar{\boldsymbol{x}}$: vérité et $\boldsymbol{\zeta} \sim \mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$

Estimateur paramétrique $(\boldsymbol{y}; \lambda) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)$

Stein Unbiased Risk Estimate (Principe)

Observations $\boldsymbol{y} = \bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P$, $\bar{\boldsymbol{x}}$: vérité et $\boldsymbol{\zeta} \sim \mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$

Estimateur paramétrique $(\boldsymbol{y}; \lambda) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)$

Ex. $$\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) = \begin{cases} \left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \boldsymbol{y} & \text{(linéaire)} \\ \underset{\boldsymbol{x}}{\operatorname{argmin}} \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{x}) & \text{(non linéaire)} \end{cases}$$

Stein Unbiased Risk Estimate (Principe)

Observations $\boldsymbol{y} = \bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P$, $\bar{\boldsymbol{x}}$: vérité et $\zeta \sim \mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$

Estimateur paramétrique $(\boldsymbol{y}; \lambda) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)$

Ex. $$\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) = \begin{cases} \left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \boldsymbol{y} & \text{(linéaire)} \\ \underset{\boldsymbol{x}}{\operatorname{argmin}} \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{x}) & \text{(non linéaire)} \end{cases}$$

Erreur quadratique $R(\lambda) \triangleq \mathbb{E}_\zeta \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \bar{\boldsymbol{x}}\|^2 = \mathbb{E}_\zeta \widehat{R}(\boldsymbol{y}; \lambda)$ $\bar{\boldsymbol{x}}$ inconnue

Stein Unbiased Risk Estimate (Principe)

Observations $\boldsymbol{y} = \bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P$, $\bar{\boldsymbol{x}}$: vérité et $\zeta \sim \mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$

Estimateur paramétrique $(\boldsymbol{y}; \lambda) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)$

Ex. $$\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) = \begin{cases} \left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \boldsymbol{y} & \text{(linéaire)} \\ \underset{\boldsymbol{x}}{\operatorname{argmin}} \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{x}) & \text{(non linéaire)} \end{cases}$$

Erreur quadratique $R(\lambda) \triangleq \mathbb{E}_\zeta \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \bar{\boldsymbol{x}}\|^2 = \mathbb{E}_\zeta \widehat{R}(\boldsymbol{y}; \lambda)$ $\bar{\boldsymbol{x}}$ inconnue

$$R(\lambda) = \mathbb{E}_\zeta \left\|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y} + \boldsymbol{y} - \bar{\boldsymbol{x}}\right\|^2$$

Stein Unbiased Risk Estimate (Principe)

Observations $\boldsymbol{y} = \bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P$, $\bar{\boldsymbol{x}}$: vérité et $\zeta \sim \mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$

Estimateur paramétrique $(\boldsymbol{y}; \lambda) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)$

Ex. $\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) = \begin{cases} \left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \boldsymbol{y} & \text{(linéaire)} \\ \operatorname{argmin}_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{x}) & \text{(non linéaire)} \end{cases}$

Erreur quadratique $R(\lambda) \triangleq \mathbb{E}_\zeta \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \bar{\boldsymbol{x}}\|^2 = \mathbb{E}_\zeta \widehat{R}(\boldsymbol{y}; \lambda)$ $\bar{\boldsymbol{x}}$ inconnue

$$\begin{aligned} R(\lambda) &= \mathbb{E}_\zeta \left\|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y} + \boldsymbol{y} - \bar{\boldsymbol{x}}\right\|^2 \\ &= \mathbb{E}_\zeta \left\|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}\right\|^2 + 2\mathbb{E}_\zeta \left\langle \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}, \boldsymbol{y} - \bar{\boldsymbol{x}} \right\rangle + \mathbb{E}_\zeta \left\|\boldsymbol{y} - \bar{\boldsymbol{x}}\right\|^2 \end{aligned}$$

Stein Unbiased Risk Estimate (Principe)

Observations $\boldsymbol{y} = \bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P$, $\bar{\boldsymbol{x}}$: vérité et $\zeta \sim \mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$

Estimateur paramétrique $(\boldsymbol{y}; \lambda) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)$

$$\text{Ex.} \quad \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) = \begin{cases} \left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \boldsymbol{y} & \text{(linéaire)} \\ \operatorname{argmin}_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{x}) & \text{(non linéaire)} \end{cases}$$

Erreur quadratique $R(\lambda) \triangleq \mathbb{E}_\zeta \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \bar{\boldsymbol{x}}\|^2 = \mathbb{E}_\zeta \widehat{R}(\boldsymbol{y}; \lambda)$ $\bar{\boldsymbol{x}}$ inconnue

$$\begin{aligned} R(\lambda) &= \mathbb{E}_\zeta \left\|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y} + \boldsymbol{y} - \bar{\boldsymbol{x}}\right\|^2 \\ &= \mathbb{E}_\zeta \left\|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}\right\|^2 + 2\mathbb{E}_\zeta \left\langle \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}, \boldsymbol{y} - \bar{\boldsymbol{x}} \right\rangle + \mathbb{E}_\zeta \left\|\boldsymbol{y} - \bar{\boldsymbol{x}}\right\|^2 \\ &= \mathbb{E}_\zeta \left\|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}\right\|^2 + 2\mathbb{E}_\zeta \left\langle \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda), \zeta \right\rangle - 2\mathbb{E}_\zeta \left\langle \bar{\boldsymbol{x}} + \zeta, \zeta \right\rangle + \mathbb{E}_\zeta \left\|\zeta\right\|^2 \end{aligned}$$

Stein Unbiased Risk Estimate (Principe)

Observations $\boldsymbol{y} = \bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P$, $\bar{\boldsymbol{x}}$: vérité et $\zeta \sim \mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$

Estimateur paramétrique $(\boldsymbol{y}; \lambda) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)$

Ex. $\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) = \begin{cases} \left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \boldsymbol{y} & \text{(linéaire)} \\ \underset{\boldsymbol{x}}{\operatorname{argmin}} \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{x}) & \text{(non linéaire)} \end{cases}$

Erreur quadratique $R(\lambda) \triangleq \mathbb{E}_\zeta \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \bar{\boldsymbol{x}}\|^2 = \mathbb{E}_\zeta \widehat{R}(\boldsymbol{y}; \lambda)$ $\bar{\boldsymbol{x}}$ inconnue

$$
\begin{aligned}
R(\lambda) &= \mathbb{E}_\zeta \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y} + \boldsymbol{y} - \bar{\boldsymbol{x}}\|^2 \\
&= \mathbb{E}_\zeta \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}\|^2 + 2\mathbb{E}_\zeta \langle \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}, \boldsymbol{y} - \bar{\boldsymbol{x}} \rangle + \mathbb{E}_\zeta \|\boldsymbol{y} - \bar{\boldsymbol{x}}\|^2 \\
&= \mathbb{E}_\zeta \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}\|^2 + 2\mathbb{E}_\zeta \langle \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda), \zeta \rangle - 2\mathbb{E}_\zeta \langle \zeta, \zeta \rangle + \mathbb{E}_\zeta \|\zeta\|^2
\end{aligned}
$$

Stein Unbiased Risk Estimate (Principe)

Observations $\boldsymbol{y} = \bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P$, $\bar{\boldsymbol{x}}$: vérité et $\zeta \sim \mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$

Estimateur paramétrique $(\boldsymbol{y}; \lambda) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)$

$$\text{Ex.} \quad \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) = \begin{cases} \left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \boldsymbol{y} & \text{(linéaire)} \\ \underset{\boldsymbol{x}}{\operatorname{argmin}} \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{x}) & \text{(non linéaire)} \end{cases}$$

Erreur quadratique $R(\lambda) \triangleq \mathbb{E}_\zeta \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \bar{\boldsymbol{x}}\|^2 = \mathbb{E}_\zeta \widehat{R}(\boldsymbol{y}; \lambda)$ $\bar{\boldsymbol{x}}$ inconnue

$$\begin{aligned} R(\lambda) &= \mathbb{E}_\zeta \left\|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y} + \boldsymbol{y} - \bar{\boldsymbol{x}}\right\|^2 \\ &= \mathbb{E}_\zeta \left\|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}\right\|^2 + 2\mathbb{E}_\zeta \left\langle \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}, \boldsymbol{y} - \bar{\boldsymbol{x}} \right\rangle + \mathbb{E}_\zeta \left\|\boldsymbol{y} - \bar{\boldsymbol{x}}\right\|^2 \\ &= \mathbb{E}_\zeta \left\|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}\right\|^2 + 2\mathbb{E}_\zeta \left\langle \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda), \zeta \right\rangle - 2\mathbb{E}_\zeta \left\langle \zeta, \zeta \right\rangle + \mathbb{E}_\zeta \left\|\zeta\right\|^2 \\ &= \mathbb{E}_\zeta \underline{\left\|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}\right\|^2} + 2\mathbb{E}_\zeta \left\langle \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda), \zeta \right\rangle - \underline{\mathbb{E}_\zeta \left\|\zeta\right\|^2} \end{aligned}$$

Stein Unbiased Risk Estimate (Principe)

Observations $\boldsymbol{y} = \bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P$, $\bar{\boldsymbol{x}}$: vérité et $\zeta \sim \mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$

Estimateur paramétrique $(\boldsymbol{y}; \lambda) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)$

$$\text{Ex.} \quad \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) = \begin{cases} \left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \boldsymbol{y} & \text{(linéaire)} \\ \underset{\boldsymbol{x}}{\operatorname{argmin}} \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{x}) & \text{(non linéaire)} \end{cases}$$

Erreur quadratique $R(\lambda) \triangleq \mathbb{E}_\zeta \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \bar{\boldsymbol{x}}\|^2 = \mathbb{E}_\zeta \widehat{R}(\boldsymbol{y}; \lambda)$ $\bar{\boldsymbol{x}}$ inconnue

$$\begin{aligned} R(\lambda) &= \mathbb{E}_\zeta \left\|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y} + \boldsymbol{y} - \bar{\boldsymbol{x}}\right\|^2 \\ &= \mathbb{E}_\zeta \left\|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}\right\|^2 + 2\mathbb{E}_\zeta \left\langle \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}, \boldsymbol{y} - \bar{\boldsymbol{x}} \right\rangle + \mathbb{E}_\zeta \left\|\boldsymbol{y} - \bar{\boldsymbol{x}}\right\|^2 \\ &= \mathbb{E}_\zeta \left\|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}\right\|^2 + 2\mathbb{E}_\zeta \left\langle \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda), \zeta \right\rangle - 2\mathbb{E}_\zeta \left\langle \zeta, \zeta \right\rangle + \mathbb{E}_\zeta \left\|\zeta\right\|^2 \\ &= \underbrace{\mathbb{E}_\zeta \left\|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}\right\|^2}_{\text{accessible}} + 2\mathbb{E}_\zeta \left\langle \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda), \zeta \right\rangle - \underline{\mathbb{E}_\zeta \left\|\zeta\right\|^2} \end{aligned}$$

Stein Unbiased Risk Estimate (Principe)

Observations $\boldsymbol{y} = \bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P$, $\bar{\boldsymbol{x}}$: vérité et $\zeta \sim \mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$

Estimateur paramétrique $(\boldsymbol{y}; \lambda) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)$

Ex. $$\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) = \begin{cases} \left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \boldsymbol{y} & \text{(linéaire)} \\ \operatorname*{argmin}_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{x}) & \text{(non linéaire)} \end{cases}$$

Erreur quadratique $R(\lambda) \triangleq \mathbb{E}_\zeta \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \bar{\boldsymbol{x}}\|^2 = \mathbb{E}_\zeta \widehat{R}(\boldsymbol{y}; \lambda)$ $\bar{\boldsymbol{x}}$ inconnue

$$\begin{aligned} R(\lambda) &= \mathbb{E}_\zeta \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y} + \boldsymbol{y} - \bar{\boldsymbol{x}}\|^2 \\ &= \mathbb{E}_\zeta \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}\|^2 + 2\mathbb{E}_\zeta \langle \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}, \boldsymbol{y} - \bar{\boldsymbol{x}} \rangle + \mathbb{E}_\zeta \|\boldsymbol{y} - \bar{\boldsymbol{x}}\|^2 \\ &= \mathbb{E}_\zeta \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}\|^2 + 2\mathbb{E}_\zeta \langle \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda), \zeta \rangle - 2\mathbb{E}_\zeta \langle \zeta, \zeta \rangle + \mathbb{E}_\zeta \|\zeta\|^2 \\ &= \underbrace{\mathbb{E}_\zeta \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}\|^2}_{\text{accessible}} + 2\mathbb{E}_\zeta \langle \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda), \zeta \rangle - \underbrace{\mathbb{E}_\zeta \|\zeta\|^2}_{\rho^2 P} \end{aligned}$$

Stein Unbiased Risk Estimate (Principe)

Observations $\boldsymbol{y} = \bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P$, $\bar{\boldsymbol{x}}$: vérité et $\zeta \sim \mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$

Estimateur paramétrique $(\boldsymbol{y}; \lambda) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)$

$$\text{Ex.} \quad \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) = \begin{cases} \left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \boldsymbol{y} & \text{(linéaire)} \\ \operatorname{argmin}_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{x}) & \text{(non linéaire)} \end{cases}$$

Erreur quadratique $R(\lambda) \triangleq \mathbb{E}_\zeta \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \bar{\boldsymbol{x}}\|^2 = \mathbb{E}_\zeta \widehat{R}(\boldsymbol{y}; \lambda)$ $\bar{\boldsymbol{x}}$ inconnue

$$\begin{aligned} R(\lambda) &= \mathbb{E}_\zeta \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y} + \boldsymbol{y} - \bar{\boldsymbol{x}}\|^2 \\ &= \mathbb{E}_\zeta \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}\|^2 + 2\mathbb{E}_\zeta \langle \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}, \boldsymbol{y} - \bar{\boldsymbol{x}} \rangle + \mathbb{E}_\zeta \|\boldsymbol{y} - \bar{\boldsymbol{x}}\|^2 \\ &= \mathbb{E}_\zeta \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}\|^2 + 2\mathbb{E}_\zeta \langle \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda), \zeta \rangle - 2\mathbb{E}_\zeta \langle \zeta, \zeta \rangle + \mathbb{E}_\zeta \|\zeta\|^2 \\ &= \mathbb{E}_\zeta \underbrace{\|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}\|^2}_{\text{accessible}} + 2\mathbb{E}_\zeta \langle \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda), \zeta \rangle - \underbrace{\mathbb{E}_\zeta \|\zeta\|^2}_{\rho^2 P} \end{aligned}$$

$$\mathbb{E}_\zeta \langle \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda), \zeta \rangle = \int \langle \widehat{\boldsymbol{x}}(\bar{\boldsymbol{x}} + \zeta; \lambda), \zeta \rangle \exp\left(-\frac{\|\zeta\|^2}{2\rho^2}\right) \mathrm{d}\zeta$$

Stein Unbiased Risk Estimate (Principe)

Observations $\quad \boldsymbol{y} = \bar{\boldsymbol{x}} + \zeta \in \mathbb{R}^P, \quad \bar{\boldsymbol{x}}$: vérité et $\zeta \sim \mathcal{N}(\mathbf{0},\rho^2\mathbf{I})$

Estimateur paramétrique $\quad (\boldsymbol{y};\lambda) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y};\lambda)$

Ex. $$\widehat{\boldsymbol{x}}(\boldsymbol{y};\lambda) = \begin{cases} \left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \boldsymbol{y} & \text{(linéaire)} \\ \operatorname*{argmin}_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{x}) & \text{(non linéaire)} \end{cases}$$

Erreur quadratique $R(\lambda) \triangleq \mathbb{E}_\zeta \|\widehat{\boldsymbol{x}}(\boldsymbol{y};\lambda) - \bar{\boldsymbol{x}}\|^2 = \mathbb{E}_\zeta \widehat{R}(\boldsymbol{y};\lambda)$ $\quad \bar{\boldsymbol{x}}$ inconnue

$$\begin{aligned} R(\lambda) &= \mathbb{E}_\zeta \|\widehat{\boldsymbol{x}}(\boldsymbol{y};\lambda) - \boldsymbol{y} + \boldsymbol{y} - \bar{\boldsymbol{x}}\|^2 \\ &= \mathbb{E}_\zeta \|\widehat{\boldsymbol{x}}(\boldsymbol{y};\lambda) - \boldsymbol{y}\|^2 + 2\mathbb{E}_\zeta \langle \widehat{\boldsymbol{x}}(\boldsymbol{y};\lambda) - \boldsymbol{y}, \boldsymbol{y} - \bar{\boldsymbol{x}} \rangle + \mathbb{E}_\zeta \|\boldsymbol{y} - \bar{\boldsymbol{x}}\|^2 \\ &= \mathbb{E}_\zeta \|\widehat{\boldsymbol{x}}(\boldsymbol{y};\lambda) - \boldsymbol{y}\|^2 + 2\mathbb{E}_\zeta \langle \widehat{\boldsymbol{x}}(\boldsymbol{y};\lambda), \zeta \rangle - 2\mathbb{E}_\zeta \langle \zeta, \zeta \rangle + \mathbb{E}_\zeta \|\zeta\|^2 \\ &= \mathbb{E}_\zeta \underbrace{\|\widehat{\boldsymbol{x}}(\boldsymbol{y};\lambda) - \boldsymbol{y}\|^2}_{\text{accessible}} + 2\mathbb{E}_\zeta \langle \widehat{\boldsymbol{x}}(\boldsymbol{y};\lambda), \zeta \rangle - \underbrace{\mathbb{E}_\zeta \|\zeta\|^2}_{\rho^2 P} \end{aligned}$$

$$\mathbb{E}_\zeta \langle \widehat{\boldsymbol{x}}(\boldsymbol{y};\lambda), \zeta \rangle = \int \langle \widehat{\boldsymbol{x}}(\bar{\boldsymbol{x}} + \zeta;\lambda), \zeta \rangle \exp\left(-\frac{\|\zeta\|^2}{2\rho^2}\right) \mathrm{d}\zeta \stackrel{\text{I.P.P.}}{=} \rho^2 \mathbb{E}_\zeta \operatorname{tr}\left(\partial_{\boldsymbol{y}} \widehat{\boldsymbol{x}}(\boldsymbol{y};\lambda)\right)$$

Estimateur séquentiel et différentiation récursive

$$\boldsymbol{\Lambda} \triangleq (\lambda,\alpha), \quad \mathbf{U}_{\boldsymbol{\Lambda}} : (\boldsymbol{h},\boldsymbol{v}) \mapsto \lambda[\alpha \mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}]$$

Primal-dual accéléré

$$\widetilde{\boldsymbol{z}}^n = \boldsymbol{z}^n + \tau_n \mathbf{U}_{\boldsymbol{\Lambda}} \boldsymbol{w}^n$$
$$\boldsymbol{z}^{n+1} = \mathrm{prox}_{\tau_n(\|\cdot\|_{2,1})^*}\left(\widetilde{\boldsymbol{z}}^n\right)$$
$$\widetilde{\boldsymbol{x}}^n = \boldsymbol{x}^n - \sigma_n \mathbf{U}_{\boldsymbol{\Lambda}}^* \boldsymbol{z}^{n+1}$$
$$\boldsymbol{x}^{n+1} = \mathrm{prox}_{\sigma_n\|\mathbf{D}\mathcal{L}-\boldsymbol{\Phi}\cdot\|_2^2}\left(\widetilde{\boldsymbol{x}}^n\right)$$
$$\theta_n = (1+2\mu\sigma_n)^{-\frac{1}{2}},$$
$$\tau_{n+1} = \tau_n/\theta_n,\ \sigma_{n+1} = \theta_n\sigma_n$$
$$\boldsymbol{w}^{n+1} = \boldsymbol{x}^n + \theta^n\left(\boldsymbol{x}^{n+1} - \boldsymbol{x}^n\right)$$

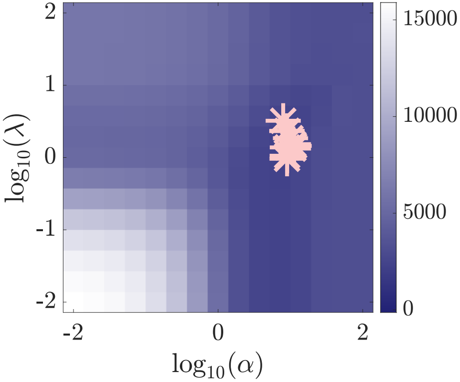
Primal-dual accéléré différentié

$$\partial_{\boldsymbol{\Lambda}}\widetilde{\boldsymbol{z}}^n = \partial_{\boldsymbol{\Lambda}}\boldsymbol{z}^n + \tau_n \mathbf{U}_{\boldsymbol{\Lambda}}\partial_{\boldsymbol{\Lambda}}\boldsymbol{w}^n + \tau_n\partial_{\boldsymbol{\Lambda}}\mathbf{U}_{\boldsymbol{\Lambda}}\boldsymbol{w}^n$$
$$\partial_{\boldsymbol{\Lambda}}\boldsymbol{z}^{n+1} = \partial_{\widetilde{\boldsymbol{z}}}\mathrm{prox}_{\tau_n(\|\cdot\|_{2,1})^*}\left(\widetilde{\boldsymbol{z}}^n\right)[\partial_{\boldsymbol{\Lambda}}\widetilde{\boldsymbol{z}}^n]$$
$$\partial_{\boldsymbol{\Lambda}}\widetilde{\boldsymbol{x}}^n = \partial_{\boldsymbol{\Lambda}}\boldsymbol{x}^n - \sigma_n \mathbf{U}_{\boldsymbol{\Lambda}}^*\partial_{\boldsymbol{\Lambda}}\boldsymbol{z}^{n+1} - \sigma_n\partial_{\boldsymbol{\Lambda}}\mathbf{U}_{\boldsymbol{\Lambda}}\boldsymbol{z}^{n+1}$$
$$\partial_{\boldsymbol{\Lambda}}\boldsymbol{x}^{n+1} = \partial_{\widetilde{\boldsymbol{x}}}\mathrm{prox}_{\sigma_n\|\mathbf{D}\mathcal{L}-\boldsymbol{\Phi}\cdot\|_2^2}\left(\widetilde{\boldsymbol{x}}^n\right)[\partial_{\boldsymbol{\Lambda}}\widetilde{\boldsymbol{x}}^n]$$
$$\partial_{\boldsymbol{\Lambda}}\boldsymbol{w}^{n+1} = \partial_{\boldsymbol{\Lambda}}\boldsymbol{x}^n + \theta^n\left(\partial_{\boldsymbol{\Lambda}}\boldsymbol{x}^{n+1} - \partial_{\boldsymbol{\Lambda}}\boldsymbol{x}^n\right)$$

Recherche automatique des paramètres de régularisation

Moyenne sur dix réalisations de texture

Matrice de covariance estimée $\boldsymbol{\mathcal{S}}$

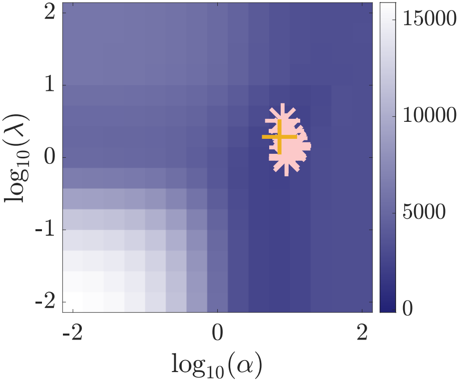


$\widehat{\lambda}^{\mathrm{qN}} = 1{,}74 \pm 0{,}43 \qquad \widehat{\alpha}^{\mathrm{qN}} = 9{,}62 \pm 0{,}70$

Recherche automatique des paramètres de régularisation

Moyenne sur dix réalisations de texture

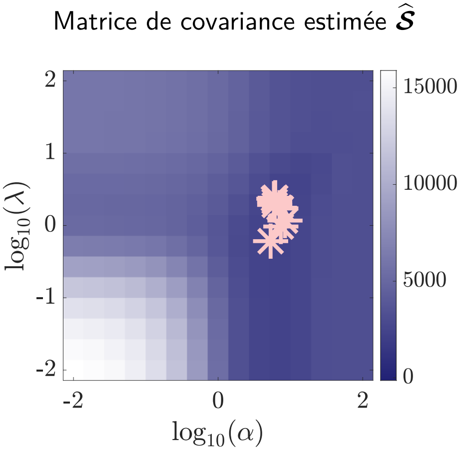
Matrice de covariance estimée $\boldsymbol{\mathcal{S}}$



$\widehat{\lambda}^{\mathrm{qN}} = 1{,}74 \pm 0{,}43 \qquad \widehat{\alpha}^{\mathrm{qN}} = 9{,}62 \pm 0{,}70$

Recherche automatique des paramètres de régularisation

Moyenne sur dix réalisations de texture

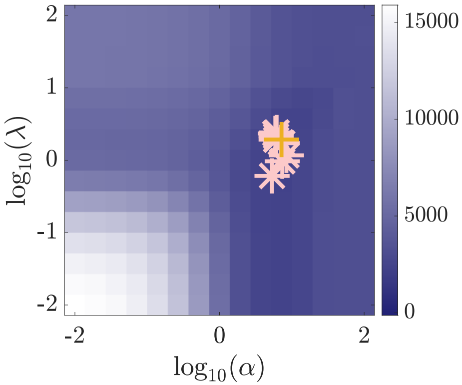


$\widehat{\lambda}^{\mathsf{qN}} = 1{,}68 \pm 0{,}37 \qquad \widehat{\alpha}^{\mathsf{qN}} = 6{,}70 \pm 0{,}58$

Recherche automatique des paramètres de régularisation

Moyenne sur dix réalisations de texture

Matrice de covariance estimée $\widehat{\mathcal{S}}$



$\widehat{\lambda}^{\mathrm{qN}} = 1{,}68 \pm 0{,}37 \qquad \widehat{\alpha}^{\mathrm{qN}} = 6{,}70 \pm 0{,}58$

Initialisation quasi-Newton

- Hyperparamètres $\mathbf{\Lambda}^{[0]} = \big(\lambda^{[0]}, \alpha^{[0]}\big)$, avec

$$\lambda^{[0]} = \frac{\mathrm{tr}(\boldsymbol{\mathcal{S}})}{2\,\mathrm{TV}(\widehat{\boldsymbol{v}}^{\mathrm{RL}}(\boldsymbol{\mathcal{L}}))}, \quad \text{et} \quad \alpha^{[0]} = \frac{\mathrm{TV}(\widehat{\boldsymbol{v}}^{\mathrm{RL}}(\boldsymbol{\mathcal{L}}))}{\mathrm{TV}(\widehat{\boldsymbol{h}}^{\mathrm{RL}}(\boldsymbol{\mathcal{L}}))}.$$

- Approximation de l'inverse de la hessienne

$$\boldsymbol{H}^{[0]} = \mathrm{diag}\left(\left|\frac{\kappa\lambda^{[0]}}{\partial_\lambda \widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{\mathcal{L}}; \mathbf{\Lambda}^{[0]} | \boldsymbol{\mathcal{S}})}\right|, \left|\frac{\kappa\alpha^{[0]}}{\partial_\alpha \widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{\mathcal{L}}; \mathbf{\Lambda}^{[0]} | \boldsymbol{\mathcal{S}})}\right|\right).$$