Bandelettes, estimation géométrique des images et sélection de modèles

E. Le Pennec,

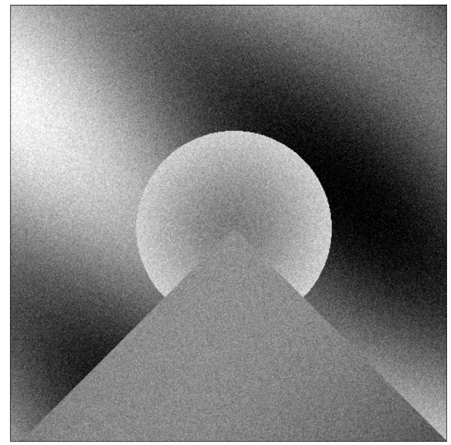
Ch. Dossal, S. Mallat

LPMA (Université Paris 7) – CMAP (École Polytechnique) – Let It Wave

Estimation géométrique des images

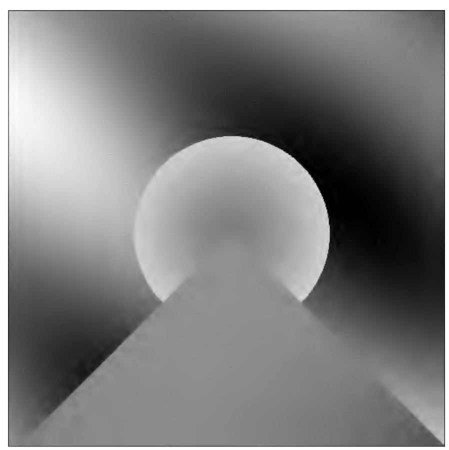
Estimation géométrique des images

Bruitée

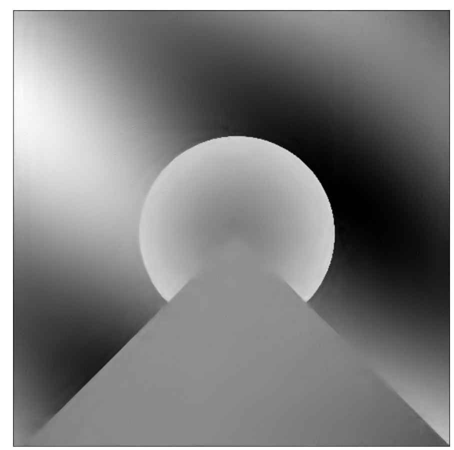


Restaurée

Sans géométrie

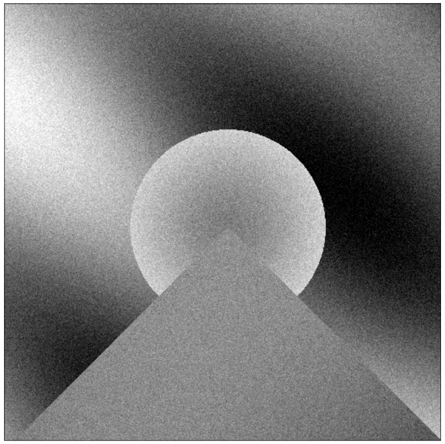


Avec géométrie

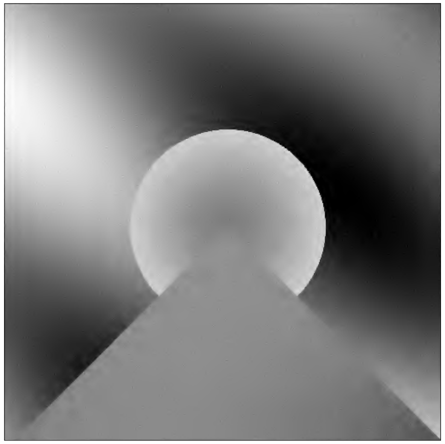


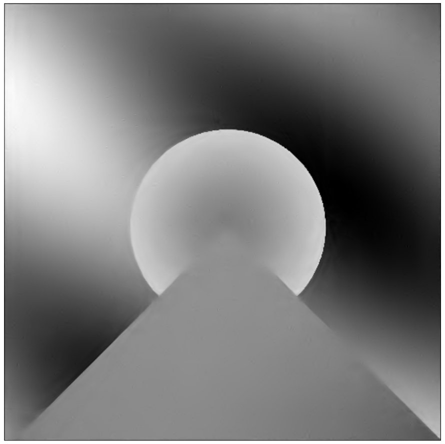
Estimation géométrique des images

Bruitée



Restaurée Sans géométrie Avec géométrie

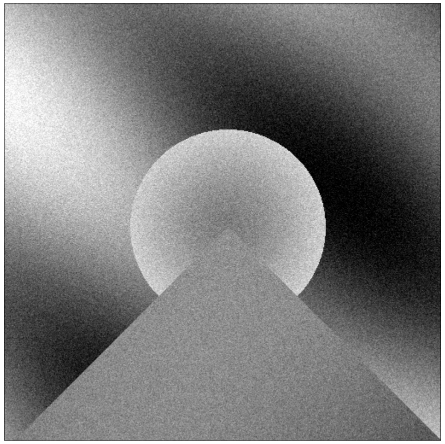




- Estimation par projection et sélection de modèles.

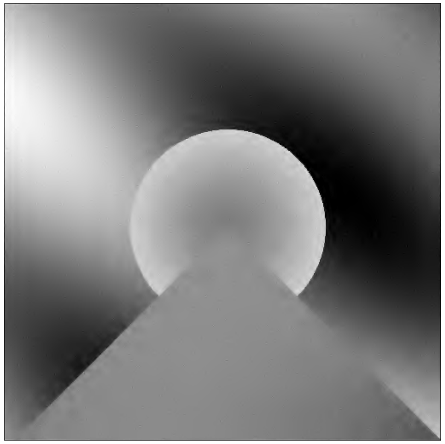
Estimation géométrique des images

Bruitée

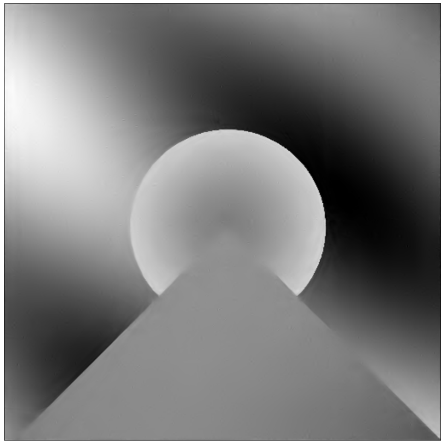


Restaurée

Sans géométrie



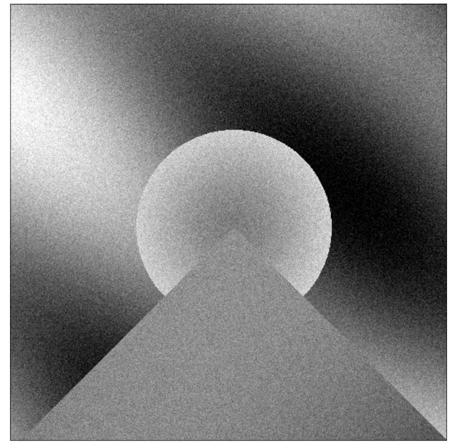
Avec géométrie



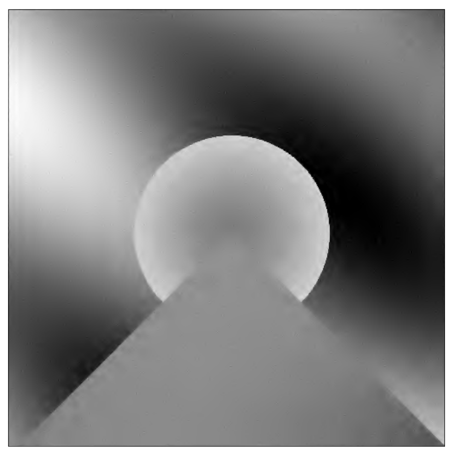
- Estimation par projection et sélection de modèles.
- Représentation géométrique des images (bandelettes).

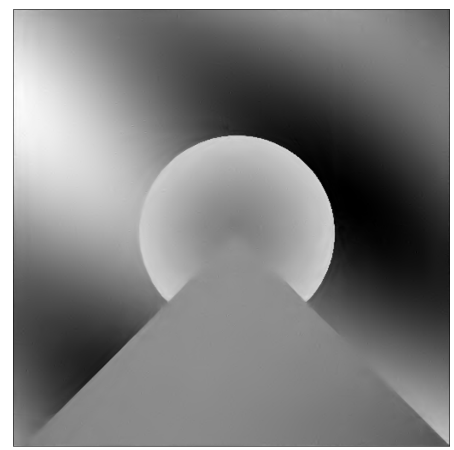
Estimation géométrique des images

Bruitée



Restaurée Sans géométrie Avec géométrie





- Estimation par projection et sélection de modèles.
- Représentation géométrique des images (bandelettes).
- Algorithmique.

Plan

Plan

- Modèle de bruit blanc et seuillage dans une base.

Plan

- Modèle de bruit blanc et seuillage dans une base.
- Lien avec l'approximation non linéaire.

Plan

- Modèle de bruit blanc et seuillage dans une base.
- Lien avec l'approximation non linéaire.
- Ondelettes 1D et 2D.

Plan

- Modèle de bruit blanc et seuillage dans une base.
- Lien avec l'approximation non linéaire.
- Ondelettes 1D et 2D.
- Représentations géométriques des images.

Plan

- Modèle de bruit blanc et seuillage dans une base.
- Lien avec l'approximation non linéaire.
- Ondelettes 1D et 2D.
- Représentations géométriques des images.
- Construction des bandelettes.

Plan

- Modèle de bruit blanc et seuillage dans une base.
- Lien avec l'approximation non linéaire.
- Ondelettes 1D et 2D.
- Représentations géométriques des images.
- Construction des bandelettes.
- Seuillage et sélection de modèles.

Plan

- Modèle de bruit blanc et seuillage dans une base.
- Lien avec l'approximation non linéaire.
- Ondelettes 1D et 2D.
- Représentations géométriques des images.
- Construction des bandelettes.
- Seuillage et sélection de modèles.
- Estimation géométrique des images.

Modèle de bruit blanc

Modèle de bruit blanc

- Modèle classique : $dY = f(t)dt + \epsilon dW(t)$.

Modèle de bruit blanc

- Modèle classique : $dY = f(t)dt + \epsilon dW(t)$.
- $f \in L^2$ fonction (image) à estimer et W processus de Wiener.

Modèle de bruit blanc

- Modèle classique : $dY = f(t)dt + \epsilon dW(t)$.
- $f \in L^2$ fonction (image) à estimer et W processus de Wiener.
- Observation corrompue par un bruit additif gaussien.

Modèle de bruit blanc

- Modèle classique : $dY = f(t)dt + \epsilon dW(t)$.
- $f \in L^2$ fonction (image) à estimer et W processus de Wiener.
- Observation corrompue par un bruit additif gaussien.
- Décomposition de dY dans une base orthonormée donne le modèle de bruit blanc pour les suites :

$$\langle dY, b_n \rangle = \langle f, b_n \rangle + \epsilon \langle dW, b_n \rangle \quad .$$

Modèle de bruit blanc

- Modèle classique : $dY = f(t)dt + \epsilon dW(t)$.
- $f \in L^2$ fonction (image) à estimer et W processus de Wiener.
- Observation corrompue par un bruit additif gaussien.
- Décomposition de dY dans une base orthonormée donne le modèle de bruit blanc pour les suites :

$$\langle dY, b_n \rangle = \langle f, b_n \rangle + \epsilon \langle dW, b_n \rangle \quad .$$

- En pratique, uniquement un nombre fini de composante (projection sur un sous espace de dimension fini).

Modèle de bruit blanc

- Modèle classique : $dY = f(t)dt + \epsilon dW(t)$.
- $f \in L^2$ fonction (image) à estimer et W processus de Wiener.
- Observation corrompue par un bruit additif gaussien.
- Décomposition de dY dans une base orthonormée donne le modèle de bruit blanc pour les suites :

$$\langle dY, b_n \rangle = \langle f, b_n \rangle + \epsilon \langle dW, b_n \rangle \quad .$$

- En pratique, uniquement un nombre fini de composante (projection sur un sous espace de dimension fini).
- Lien avec le modèle d'échantillonnage via $\epsilon = \frac{\sigma}{\sqrt{n}}$.

Seuillage dans une base

Seuillage dans une base

- Méthode simple d'estimation de f à partir de dY.

Seuillage dans une base

- Méthode simple d'estimation de f à partir de dY.
- Décomposition de dY dans une base orthonormée

$$dY = \sum_{b_n} \langle dY, b_n \rangle b_n(t) dt \quad .$$

Seuillage dans une base

- Méthode simple d'estimation de f à partir de dY.
- Décomposition de dY dans une base orthonormée

$$dY = \sum_{b_n} \langle dY, b_n \rangle b_n(t) dt \quad .$$

- Estimée F de f par seuillage avec un seuil T :

$$F = \sum_{\substack{b_n \\ |\langle dY, b_n \rangle| \geqslant T}} \langle dY, b_n \rangle b_n$$

Seuillage dans une base

- Méthode simple d'estimation de f à partir de dY.
- Décomposition de dY dans une base orthonormée

$$dY = \sum_{b_n} \langle dY, b_n \rangle b_n(t) dt \quad .$$

- Estimée F de f par seuillage avec un seuil T :

$$F = \sum_{\substack{b_n \\ |\langle dY, b_n \rangle| \geqslant T}} \langle dY, b_n \rangle b_n$$

- Performance ? : si $f \in \Theta$, convergence, vitesse, optimalité minimax ?

Estimateur par projection

Estimateur par projection

- $\{b_n\}_{n \leqslant N}$ base de dimension N.

Estimateur par projection

- $\{b_n\}_{n \leqslant N}$ base de dimension N.
- Estimateur par projection : $F_I = \sum_{n \in I} \langle dY, b_n \rangle b_n$.

Estimateur par projection

- $\{b_n\}_{n\leqslant N}$ base de dimension N.
- Estimateur par projection : $F_I = \sum_{n\in I}\langle dY, b_n\rangle b_n$.
- Comment choisir $I \subset 1,..N$ pour minimiser le risque quadratique

$$E(\|F_I - f\|^2) = \sum_{n>N} |\langle f, b_n\rangle|^2 + \sum_{n\notin I} |\langle f, b_n\rangle|^2 + \sum_{n\in I} \epsilon^2$$

Estimateur par projection

- $\{b_n\}_{n\leqslant N}$ base de dimension N.
- Estimateur par projection : $F_I = \sum_{n\in I}\langle dY, b_n\rangle b_n$.
- Comment choisir $I \subset 1,..N$ pour minimiser le risque quadratique

$$E(\|F_I - f\|^2) = \sum_{n>N} |\langle f, b_n\rangle|^2 + \sum_{n\notin I} |\langle f, b_n\rangle|^2 + \sum_{n\in I} \epsilon^2$$

- Solution : $I = \{n \in \{1,\ldots,N\},\ |\langle f, b_n\rangle| \geqslant \epsilon\}$ donne F_O.

Estimateur par projection

- $\{b_n\}_{n\leqslant N}$ base de dimension N.
- Estimateur par projection : $F_I = \sum_{n\in I}\langle dY, b_n\rangle b_n$.
- Comment choisir $I \subset 1, ..N$ pour minimiser le risque quadratique

$$E(\|F_I - f\|^2) = \sum_{n>N} |\langle f, b_n\rangle|^2 + \sum_{n\notin I} |\langle f, b_n\rangle|^2 + \sum_{n\in I} \epsilon^2$$

- Solution : $I = \{n \in \{1, \ldots, N\},\ |\langle f, b_n\rangle| \geqslant \epsilon\}$ donne F_O.
- Problème : nécessité d'un oracle (F_O dépend de f).

Estimateur par projection

- $\{b_n\}_{n\leqslant N}$ base de dimension N.
- Estimateur par projection : $F_I = \sum_{n\in I}\langle dY, b_n\rangle b_n$.
- Comment choisir $I \subset 1, ..N$ pour minimiser le risque quadratique

$$E(\|F_I - f\|^2) = \sum_{n>N} |\langle f, b_n\rangle|^2 + \sum_{n\notin I} |\langle f, b_n\rangle|^2 + \sum_{n\in I} \epsilon^2$$

- Solution : $I = \{n \in \{1, \ldots, N\},\ |\langle f, b_n\rangle| \geqslant \epsilon\}$ donne F_O.
- Problème : nécessité d'un oracle (F_O dépend de f).
- Lien avec le cadre déterministe de l'approximation non linéaire :

$$f_I = \sum_{n\in I}\langle f, b_n\rangle b_n$$

et minimisation à $|I| = M$ fixé de $\|f - f_I\|^2$ donne f_M.

Oracle et approximation non linéaire

Oracle et approximation non linéaire

- Risque :

$$E(\|F_I - f\|^2) = \sum_{n>N} |\langle f, b_n\rangle|^2 + \sum_{n\notin I} |\langle f, b_n\rangle|^2 + \sum_{n\in I} \epsilon^2$$

$$E(\|F_I - f\|^2) = \|f - f_{\{1,\dots,N\}}\|^2 + \|f_{\{1,\dots,N\}} - f_I\|^2 + |I|\epsilon^2$$

$$E(\|F_I - f\|^2) = \|f - f_I\|^2 + \epsilon^2|I| \quad .$$

Oracle et approximation non linéaire

- Risque :

$$E(\|F_I - f\|^2) = \sum_{n>N} |\langle f, b_n\rangle|^2 + \sum_{n\notin I} |\langle f, b_n\rangle|^2 + \sum_{n\in I} \epsilon^2$$

$$E(\|F_I - f\|^2) = \|f - f_{\{1,\dots,N\}}\|^2 + \|f_{\{1,\dots,N\}} - f_I\|^2 + |I|\epsilon^2$$

$$E(\|F_I - f\|^2) = \|f - f_I\|^2 + \epsilon^2|I| \quad .$$

- Formulation lagrangienne du problème d'approximation non linéaire ($f_I = f_M$ et $|I| = M$).

Oracle et approximation non linéaire

- Risque :

$$E(\|F_I - f\|^2) = \sum_{n>N} |\langle f, b_n\rangle|^2 + \sum_{n\notin I} |\langle f, b_n\rangle|^2 + \sum_{n\in I} \epsilon^2$$

$$E(\|F_I - f\|^2) = \|f - f_{\{1,\ldots,N\}}\|^2 + \|f_{\{1,\ldots,N\}} - f_I\|^2 + |I|\epsilon^2$$

$$E(\|F_I - f\|^2) = \|f - f_I\|^2 + \epsilon^2|I| \quad .$$

- Formulation lagrangienne du problème d'approximation non linéaire ($f_I = f_M$ et $|I| = M$).
- Décroissance du risque en fonction de ϵ liée à la classe Θ (régularité) et à la base utilisée.

Oracle et approximation non linéaire

- Risque :

$$E(\|F_I - f\|^2) = \sum_{n>N} |\langle f, b_n\rangle|^2 + \sum_{n\notin I} |\langle f, b_n\rangle|^2 + \sum_{n\in I} \epsilon^2$$

$$E(\|F_I - f\|^2) = \|f - f_{\{1,\dots,N\}}\|^2 + \|f_{\{1,\dots,N\}} - f_I\|^2 + |I|\epsilon^2$$

$$E(\|F_I - f\|^2) = \|f - f_I\|^2 + \epsilon^2|I| \quad .$$

- Formulation lagrangienne du problème d'approximation non linéaire ($f_I = f_M$ et $|I| = M$).
- Décroissance du risque en fonction de ϵ liée à la classe Θ (régularité) et à la base utilisée.
- Décroissance de l'erreur d'approximation $\|f - f_M\|^2$ en $M^{-\beta}$: convergence du risque quadratique oracle en $\epsilon^{2\beta/(\beta+1)}$.

Oracle et approximation non linéaire

- Risque :

$$E(\|F_I - f\|^2) = \sum_{n>N} |\langle f, b_n\rangle|^2 + \sum_{n\notin I} |\langle f, b_n\rangle|^2 + \sum_{n\in I} \epsilon^2$$
$$E(\|F_I - f\|^2) = \|f - f_{\{1,\ldots,N\}}\|^2 + \|f_{\{1,\ldots,N\}} - f_I\|^2 + |I|\epsilon^2$$
$$E(\|F_I - f\|^2) = \|f - f_I\|^2 + \epsilon^2|I| \quad .$$

- Formulation lagrangienne du problème d'approximation non linéaire ($f_I = f_M$ et $|I| = M$).
- Décroissance du risque en fonction de ϵ liée à la classe Θ (régularité) et à la base utilisée.
- Décroissance de l'erreur d'approximation $\|f - f_M\|^2$ en $M^{-\beta}$: convergence du risque quadratique oracle en $\epsilon^{2\beta/(\beta+1)}$.
- Optimisation : choix de la bonne base.

Oracle et seuillage

Oracle et seuillage

- Lien avec le seuillage : Donoho, Johnstone, Kerkyacharian, Picard...

Oracle et seuillage

- Lien avec le seuillage : Donoho, Johnstone, Kerkyacharian, Picard...
- Estimée F de f par seuillage avec un seuil T :

$$F = \sum_{\substack{b_n, n \leqslant N \\ |\langle dY, b_n \rangle| \geqslant T}} \langle dY, b_n \rangle b_n$$

Oracle et seuillage

- Lien avec le seuillage : Donoho, Johnstone, Kerkyacharian, Picard...
- Estimée F de f par seuillage avec un seuil T :

$$F = \sum_{\substack{b_n, n \leqslant N \\ |\langle dY, b_n \rangle| \geqslant T}} \langle dY, b_n \rangle b_n$$

- **Théorème :** Pour $T = \lambda \sqrt{\log N} \epsilon$, avec λ suffisament grand,

$$E(\|F - f\|^2) \leqslant C \log N (E(\|F_O - f\|^2) + \epsilon/N) \quad .$$

Oracle et seuillage

- Lien avec le seuillage : Donoho, Johnstone, Kerkyacharian, Picard...
- Estimée F de f par seuillage avec un seuil T :

$$F = \sum_{\substack{b_n, n \leqslant N \\ |\langle dY, b_n \rangle| \geqslant T}} \langle dY, b_n \rangle b_n$$

- **Théorème :** Pour $T = \lambda \sqrt{\log N} \epsilon$, avec λ suffisament grand,

$$E(\|F - f\|^2) \leqslant C \log N (E(\|F_O - f\|^2) + \epsilon/N) \quad .$$

- Estimateur par seuillage *aussi* efficace que l'estimateur oracle.

Oracle et seuillage

- Lien avec le seuillage : Donoho, Johnstone, Kerkyacharian, Picard...
- Estimée F de f par seuillage avec un seuil T :

$$F = \sum_{\substack{b_n, n \leqslant N \\ |\langle dY, b_n \rangle| \geqslant T}} \langle dY, b_n \rangle b_n$$

- **Théorème :** Pour $T = \lambda \sqrt{\log N} \epsilon$, avec λ suffisament grand,

$$E(\|F - f\|^2) \leqslant C \log N (E(\|F_O - f\|^2) + \epsilon/N) \quad .$$

- Estimateur par seuillage *aussi* efficace que l'estimateur oracle.
- Étude des performances de l'estimateur oracle.

Oracle et seuillage

- Lien avec le seuillage : Donoho, Johnstone, Kerkyacharian, Picard...
- Estimée F de f par seuillage avec un seuil T :

$$F = \sum_{\substack{b_n, n \leqslant N \\ |\langle dY, b_n \rangle| \geqslant T}} \langle dY, b_n \rangle b_n$$

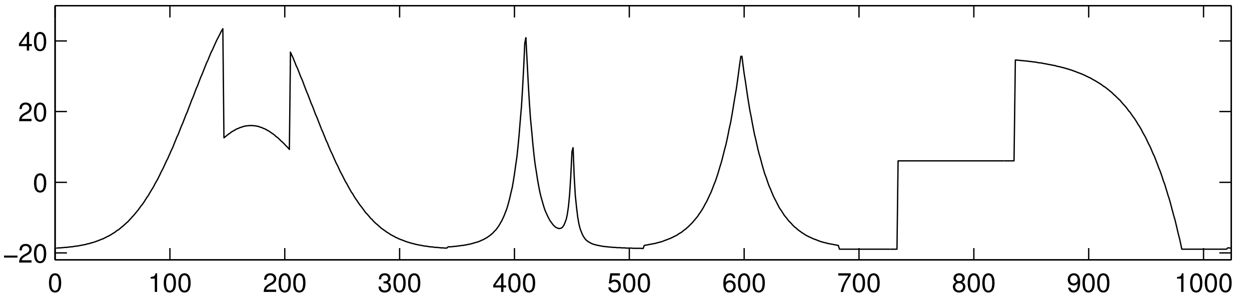
- **Théorème :** Pour $T = \lambda\sqrt{\log N}\epsilon$, avec λ suffisament grand,

$$E(\|F - f\|^2) \leqslant C \log N (E(\|F_O - f\|^2) + \epsilon/N) \quad .$$

- Estimateur par seuillage *aussi* efficace que l'estimateur oracle.
- Étude des performances de l'estimateur oracle.
- Adaptivité automatique.

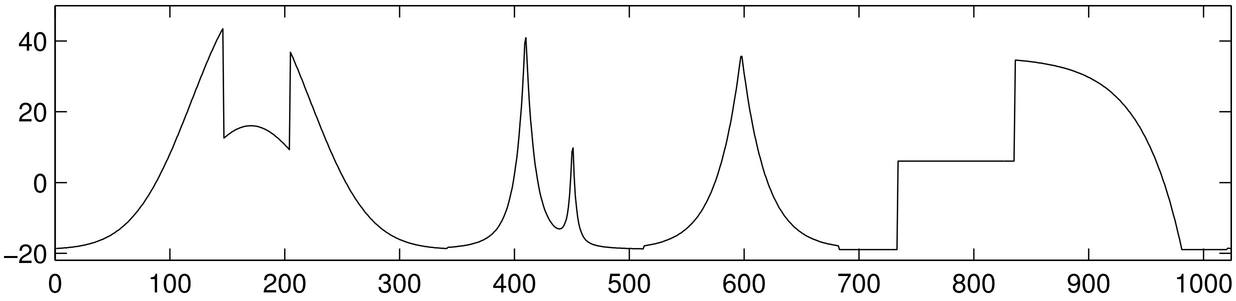
Régularité 1D

Régularité 1D



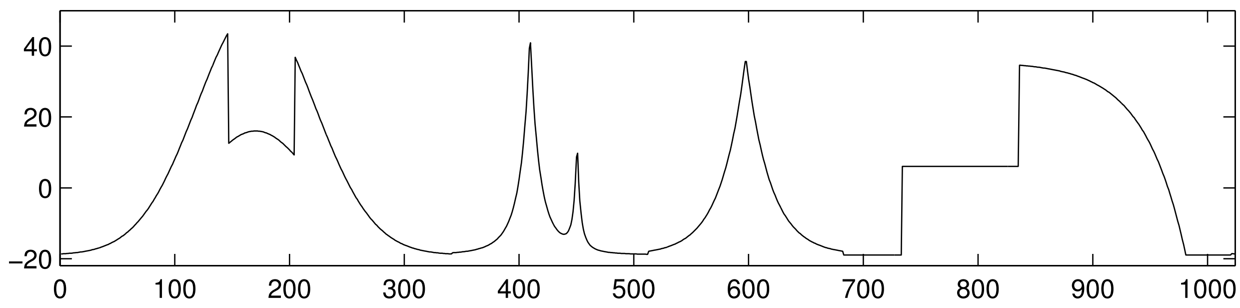
- $f \in L^2([0,1])$.

Régularité 1D



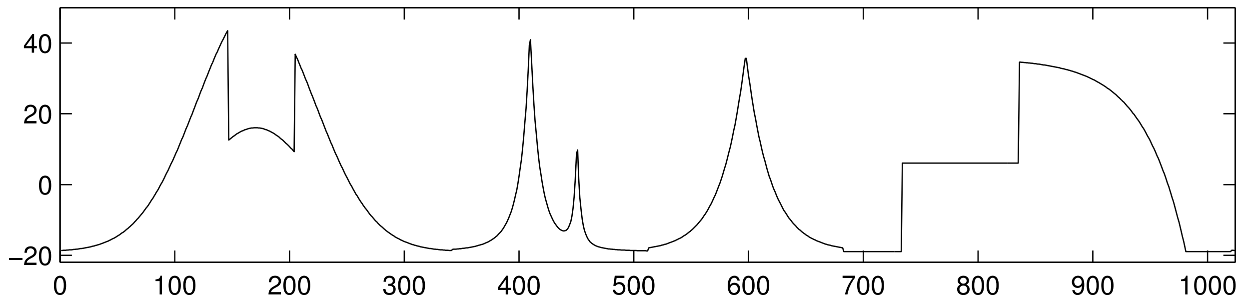
- $f \in L^2([0,1])$.
- Classe des fonctions régulières ($\mathbf{C}^\alpha$, classe de Hölder d'ordre α).

Régularité 1D



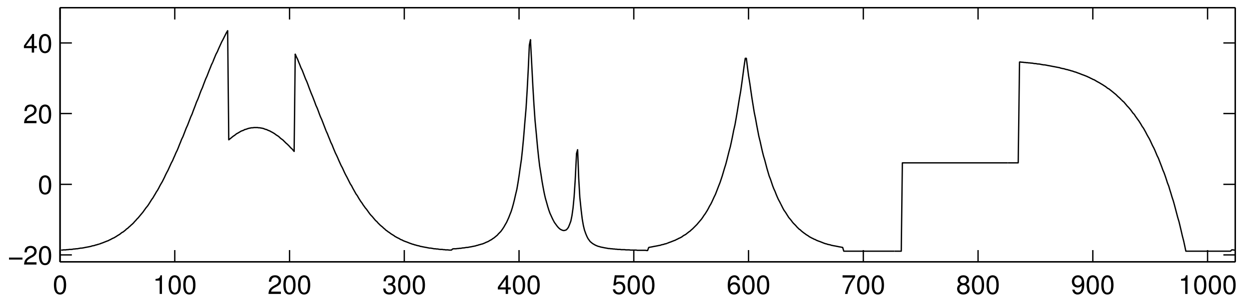
- $f \in L^2([0,1])$.
- Classe des fonctions régulières ($\mathbf{C}^\alpha$, classe de Hölder d'ordre α).
- Risque minimax $\propto \epsilon^{2\alpha/(\alpha+1/2)}$.

Régularité 1D



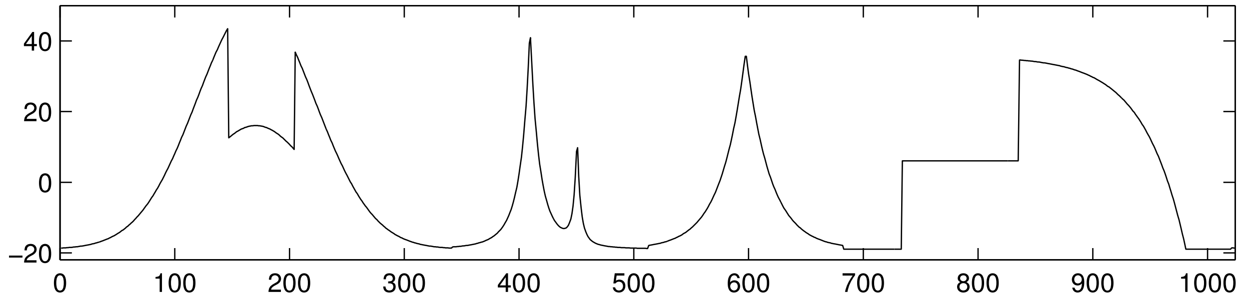
- $f \in L^2([0,1])$.
- Classe des fonctions régulières ($\mathbf{C}^\alpha$, classe de Hölder d'ordre α).
- Risque minimax $\propto \epsilon^{2\alpha/(\alpha+1/2)}$.
- Classe des fonctions régulières par morceaux ($\mathbf{C}^\alpha$ par morceaux).

Régularité 1D



- $f \in L^2([0,1])$.
- Classe des fonctions régulières ($\mathbf{C}^\alpha$, classe de Hölder d'ordre α).
- Risque minimax $\propto \epsilon^{2\alpha/(\alpha+1/2)}$.
- Classe des fonctions régulières par morceaux ($\mathbf{C}^\alpha$ par morceaux).
- Risque minimax : $\propto \epsilon^{2\alpha/(\alpha+1/2)}$.

Régularité 1D



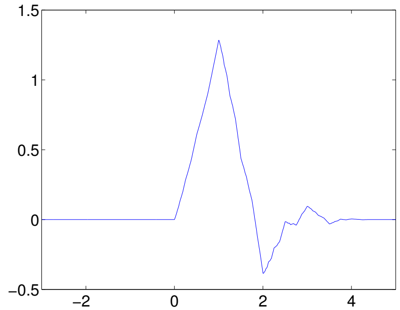
- $f \in L^2([0,1])$.
- Classe des fonctions régulières ($\mathbf{C}^\alpha$, classe de Hölder d'ordre α).
- Risque minimax $\propto \epsilon^{2\alpha/(\alpha+1/2)}$.
- Classe des fonctions régulières par morceaux ($\mathbf{C}^\alpha$ par morceaux).
- Risque minimax : $\propto \epsilon^{2\alpha/(\alpha+1/2)}$.
- Peut-on atteindre cette borne ?

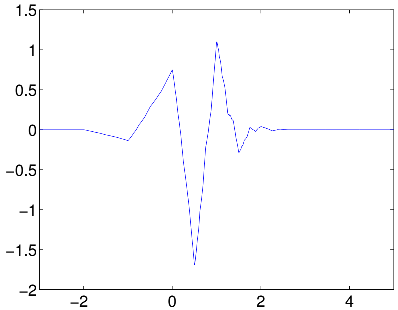
Base d'ondelettes 1D de $L^2[0,1]$

Base d'ondelettes 1D de $L^2[0,1]$

- Construite à partir d'une fonction d'échelle $\phi(x)$ et d'une ondelette mère $\psi(x)$

$\phi(x)$

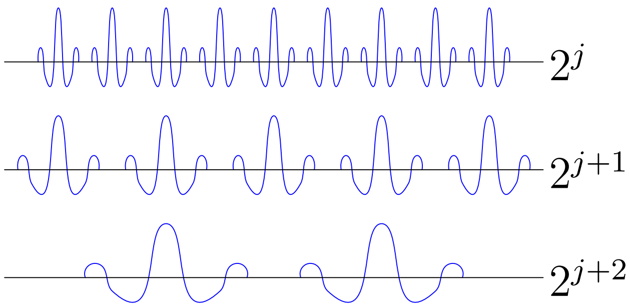




$\psi(x)$

qui sont dilatées par 2^j et translatées de $2^j n$

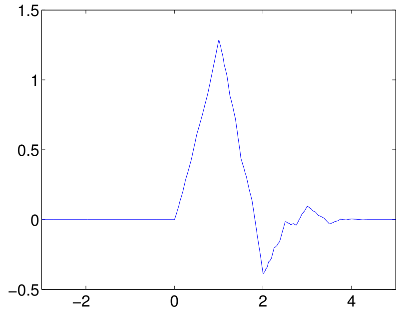
$$\phi_{j,n}(x) = \frac{1}{\sqrt{2^j}}\,\phi\Big(\frac{x - 2^j n}{2^j}\Big) \quad , \quad \psi_{j,n}(x) = \frac{1}{\sqrt{2^j}}\,\psi\Big(\frac{x - 2^j n}{2^j}\Big) \quad .$$

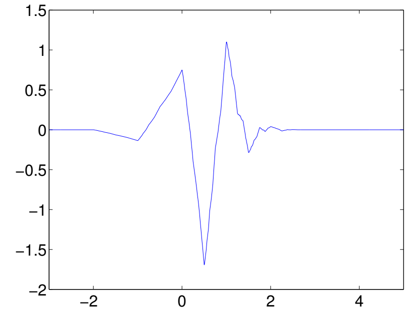


Base d'ondelettes 1D de $L^2[0,1]$

- Construite à partir d'une fonction d'échelle $\phi(x)$ et d'une ondelette mère $\psi(x)$

$\phi(x)$



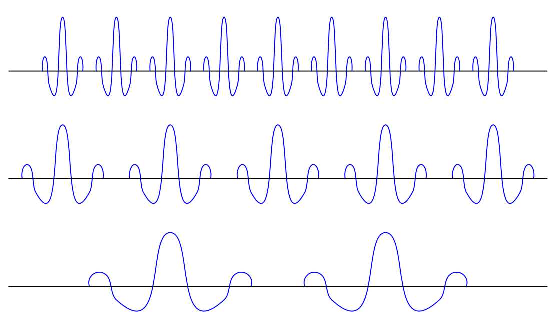


$\psi(x)$

qui sont dilatées par 2^j et translatées de $2^j n$

$$\phi_{j,n}(x) = \frac{1}{\sqrt{2^j}}\,\phi\Big(\frac{x-2^j n}{2^j}\Big) \quad , \quad \psi_{j,n}(x) = \frac{1}{\sqrt{2^j}}\,\psi\Big(\frac{x-2^j n}{2^j}\Big) \quad .$$

- $\mathbf{B} = \Big\{\psi_{j,n}\Big\}_{j\in\mathbb{N}\,,\,2^j n\in[0,1)}$ est une base orthonormale de $L^2[0,1]$.



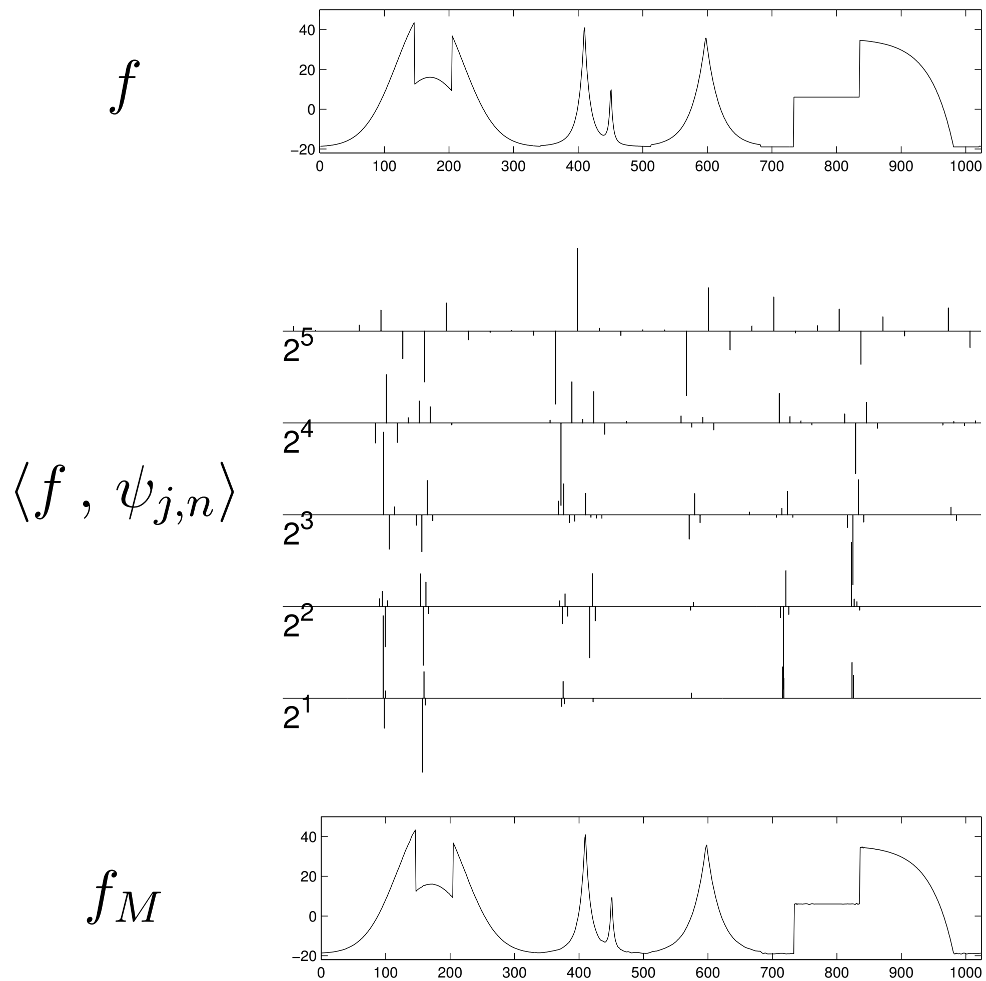
2^j

2^{j+1}

2^{j+2}

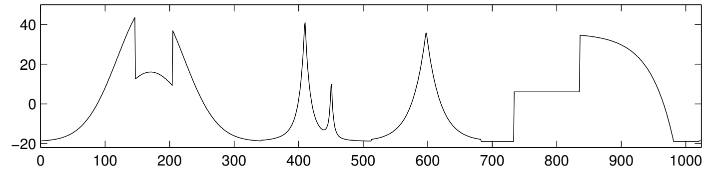
Approximation non linéaire en ondelettes

Approximation non linéaire en ondelettes

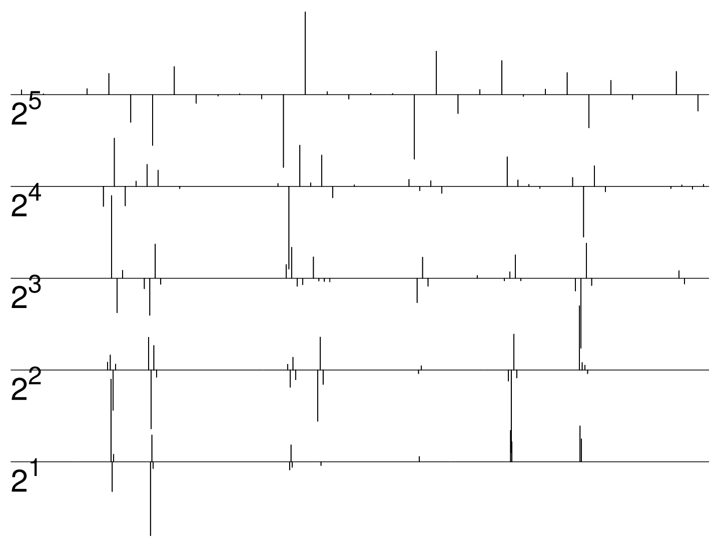


Approximation non linéaire en ondelettes

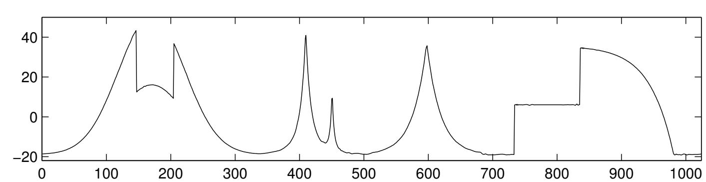
f



$\langle f\ ,\ \psi_{j,n}\rangle$



f_M



- Si f est $\mathbf{C}^\alpha$ par morceaux et ψ a $p > \alpha$ moments nuls alors

$$\|f - f_M\|^2 \leqslant C\, M^{-2\alpha} \quad .$$

Quasi optimalité

Quasi optimalité

- Si f est $\mathbf{C}^\alpha$ par morceaux, $\|f - f_M\|^2 \leqslant C\, M^{-2\alpha}$.

Quasi optimalité

- Si f est $\mathbf{C}^\alpha$ par morceaux, $\|f - f_M\|^2 \leqslant C\, M^{-2\alpha}$.
- Risque oracle correspondant : $\propto \epsilon^{2\alpha/(\alpha+1/2)}$ (risque minimax).

Quasi optimalité

- Si f est $\mathbf{C}^\alpha$ par morceaux, $\|f - f_M\|^2 \leqslant C\,M^{-2\alpha}$.
- Risque oracle correspondant : $\propto \epsilon^{2\alpha/(\alpha+1/2)}$ (risque minimax).
- Risque de l'estimateur par seuillage : $\propto (\log \epsilon)\,\epsilon^{2\alpha/(\alpha+1/2)}$ si $N \leqslant \epsilon^{-1}$.

Quasi optimalité

- Si f est $\mathbf{C}^\alpha$ par morceaux, $\|f - f_M\|^2 \leqslant C\, M^{-2\alpha}$.
- Risque oracle correspondant : $\propto \epsilon^{2\alpha/(\alpha+1/2)}$ (risque minimax).
- Risque de l'estimateur par seuillage : $\propto (\log \epsilon)\, \epsilon^{2\alpha/(\alpha+1/2)}$ si $N \leqslant \epsilon^{-1}$.
- Condition non restrictive (théorie de l'approximation linéaire) :

$$\|f - f_{1\ldots N}\|^2 \leqslant \|f - f_M\|^2 \quad .$$

Quasi optimalité

- Si f est $\mathbf{C}^\alpha$ par morceaux, $\|f - f_M\|^2 \leqslant C\, M^{-2\alpha}$.
- Risque oracle correspondant : $\propto \epsilon^{2\alpha/(\alpha+1/2)}$ (risque minimax).
- Risque de l'estimateur par seuillage : $\propto (\log \epsilon)\, \epsilon^{2\alpha/(\alpha+1/2)}$ si $N \leqslant \epsilon^{-1}$.
- Condition non restrictive (théorie de l'approximation linéaire) :

$$\|f - f_{1\dots N}\|^2 \leqslant \|f - f_M\|^2 \quad .$$

- Quasi optimalité de l'estimateur par seuillage en ondelettes pour les fonctions 1D.

Base d'ondelettes 2D séparables

Base d'ondelettes 2D séparables

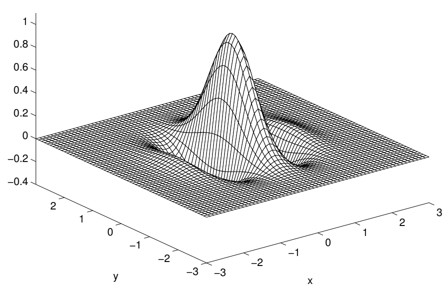
- La famille
$$\left\{ \begin{array}{ccc} \phi_{j,n_1}(x_1)\,\psi_{j,n_2}(x_2) & , & \psi_{j,n_1}(x_1)\,\phi_{j,n_2}(x_2) \\ & , & \psi_{j,n_1}(x_1)\,\psi_{j,n_2}(x_2) \end{array} \right\}_{(j,n_1,n_2)\in\mathbb{Z}^3}$$
est une base orthonormée de $L^2[0,1]^2$.

Base d'ondelettes 2D séparables

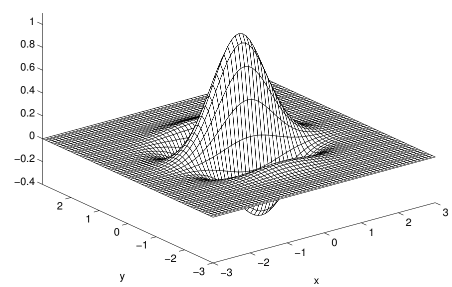
- La famille

$$\left\{ \begin{array}{ccc} \phi_{j,n_1}(x_1)\,\psi_{j,n_2}(x_2) & , & \psi_{j,n_1}(x_1)\,\phi_{j,n_2}(x_2) \\ & , & \psi_{j,n_1}(x_1)\,\psi_{j,n_2}(x_2) \end{array} \right\}_{(j,n_1,n_2)\in\mathbb{Z}^3}$$

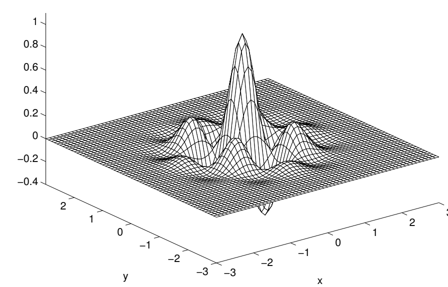
est une base orthonormée de $L^2[0,1]^2$.



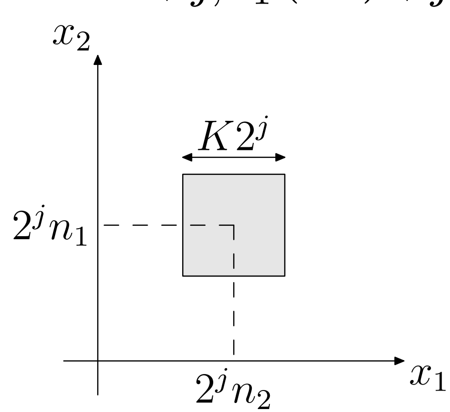
$\phi_{j,n_1}(x_1)\,\psi_{j,n_2}(x_2)$



$\psi_{j,n_1}(x_1)\,\phi_{j,n_2}(x_2)$



$\psi_{j,n_1}(x_1)\,\psi_{j,n_2}(x_2)$



Support
des ondelettes
isotropes

Seuillage en ondelettes 2D

Seuillage en ondelettes 2D

- $f \in L^2([0,1]^2)$.

Seuillage en ondelettes 2D

- $f \in L^2([0,1]^2)$.
- Modèle de bruit blanc : $dY = f(x)dx + \epsilon dW(x)$

Seuillage en ondelettes 2D

- $f \in L^2([0,1]^2)$.
- Modèle de bruit blanc : $dY = f(x)dx + \epsilon dW(x)$
- Estimateur : $F = \sum_{\substack{j \geqslant j_0 \\ |\langle dY, \psi_{j,k} \rangle| \geqslant T}} \langle dY, \psi_{j,k} \rangle \psi_{j,k}$ avec $N = 2^{-j_0}$ et $T = \lambda \sqrt{\log N} \epsilon$.

Seuillage en ondelettes 2D

- $f \in L^2([0,1]^2)$.
- Modèle de bruit blanc : $dY = f(x)dx + \epsilon dW(x)$
- Estimateur : $F = \sum_{\substack{j \geqslant j_0 \\ |\langle dY, \psi_{j,k}\rangle| \geqslant T}} \langle dY, \psi_{j,k}\rangle \psi_{j,k}$ avec $N = 2^{-j_0}$ et $T = \lambda\sqrt{\log N}\epsilon$.
- Quasi optimal pour les fonctions régulières $\mathbf{C}^\alpha$ et autres espaces classiques.

Seuillage en ondelettes 2D

- $f \in L^2([0,1]^2)$.
- Modèle de bruit blanc : $dY = f(x)dx + \epsilon dW(x)$
- Estimateur : $F = \sum_{\substack{j \geqslant j_0 \\ |\langle dY, \psi_{j,k} \rangle| \geqslant T}} \langle dY, \psi_{j,k} \rangle \psi_{j,k}$ avec $N = 2^{-j_0}$ et $T = \lambda \sqrt{\log N} \epsilon$.
- Quasi optimal pour les fonctions régulières $\mathbf{C}^\alpha$ et autres espaces classiques.
- Pour les images, modèle classique : $\mathbf{C}^\alpha - \mathbf{C}^\alpha$.

Seuillage en ondelettes 2D

- $f \in L^2([0,1]^2)$.
- Modèle de bruit blanc : $dY = f(x)dx + \epsilon dW(x)$
- Estimateur : $F = \sum_{\substack{j \geqslant j_0 \\ |\langle dY, \psi_{j,k} \rangle| \geqslant T}} \langle dY, \psi_{j,k} \rangle \psi_{j,k}$ avec $N = 2^{-j_0}$ et $T = \lambda \sqrt{\log N} \epsilon$.
- Quasi optimal pour les fonctions régulières $\mathbf{C}^\alpha$ et autres espaces classiques.
- Pour les images, modèle classique : $\mathbf{C}^\alpha - \mathbf{C}^\alpha$.
- Ici, seul modèle possible : $f \in BV$ et $\|f - f_M\|^2 \leqslant C\, M^{-1}$.

Seuillage en ondelettes 2D

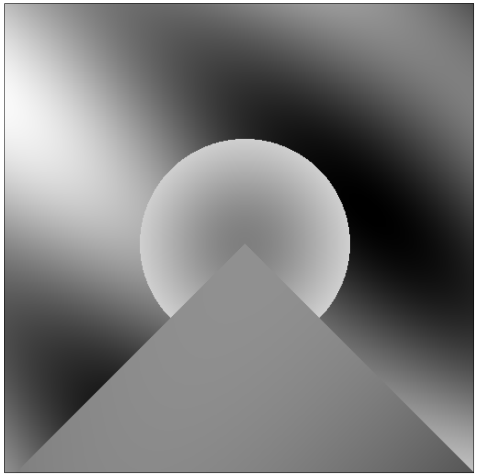
- $f \in L^2([0,1]^2)$.
- Modèle de bruit blanc : $dY = f(x)dx + \epsilon dW(x)$
- Estimateur : $F = \sum_{\substack{j \geqslant j_0 \\ |\langle dY, \psi_{j,k} \rangle| \geqslant T}} \langle dY, \psi_{j,k} \rangle \psi_{j,k}$ avec $N = 2^{-j_0}$ et $T = \lambda \sqrt{\log N} \epsilon$.
- Quasi optimal pour les fonctions régulières $\mathbf{C}^\alpha$ et autres espaces classiques.
- Pour les images, modèle classique : $\mathbf{C}^\alpha - \mathbf{C}^\alpha$.
- Ici, seul modèle possible : $f \in BV$ et $\|f - f_M\|^2 \leqslant C\, M^{-1}$.
- Risque quadratique de l'ordre de ϵ^2 qui est sous optimal si $\alpha \geqslant 1$.

Seuillage en ondelettes 2D

- $f \in L^2([0,1]^2)$.
- Modèle de bruit blanc : $dY = f(x)dx + \epsilon dW(x)$
- Estimateur : $F = \sum_{\substack{j \geqslant j_0 \\ |\langle dY, \psi_{j,k} \rangle| \geqslant T}} \langle dY, \psi_{j,k} \rangle \psi_{j,k}$ avec $N = 2^{-j_0}$ et $T = \lambda \sqrt{\log N} \epsilon$.
- Quasi optimal pour les fonctions régulières $\mathbf{C}^\alpha$ et autres espaces classiques.
- Pour les images, modèle classique : $\mathbf{C}^\alpha - \mathbf{C}^\alpha$.
- Ici, seul modèle possible : $f \in BV$ et $\|f - f_M\|^2 \leqslant C\, M^{-1}$.
- Risque quadratique de l'ordre de ϵ^2 qui est sous optimal si $\alpha \geqslant 1$.
- Problème d'approximation non linéaire.

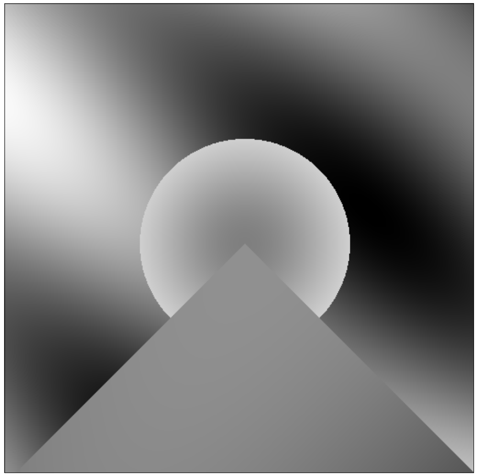
Modèle de régularité géométrique

Modèle de régularité géométrique



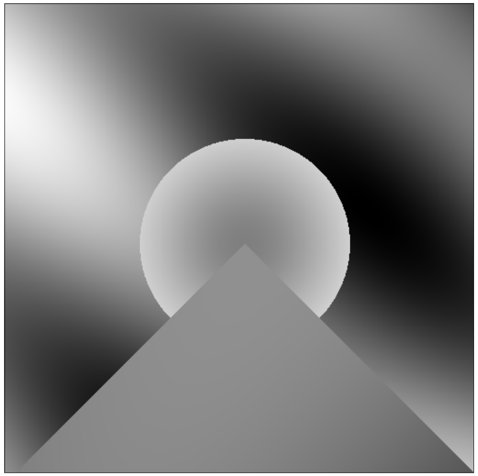
- Fonctions régulières par morceaux avec des discontinuités le long de courbes régulières.

Modèle de régularité géométrique



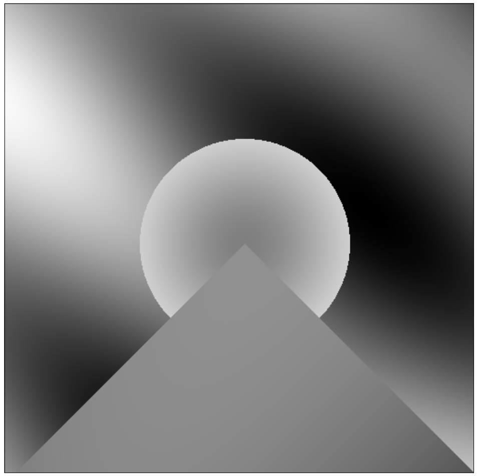
- Fonctions régulières par morceaux avec des discontinuités le long de courbes régulières.
- Cf modèle horizon ou fonctions étoilées (*Tsybakov*, *Donoho*).

Modèle de régularité géométrique



- Fonctions régulières par morceaux avec des discontinuités le long de courbes régulières.
- Cf modèle horizon ou fonctions étoilées (*Tsybakov, Donoho*).
- Fonctions de régularité géométrique $\mathbf{C}^\alpha$:
  - $f = \tilde{f}$ ou $f = \tilde{f} \star h$,
  - $\tilde{f}$: $\mathbf{C}^\alpha(\Lambda)$ avec $\Lambda = [0,1]^2 - \{\mathcal{C}_\gamma\}_{1 \leqslant \gamma \leqslant G}$,
  - $\mathcal{C}_\gamma$: $\mathbf{C}^\alpha$ + conditions géométriques de non tangence,
  - h : $\mathbf{C}^\alpha$ à support $\subset [-s, s]^2$ et $\|h\|_{\mathbf{C}^\alpha} \leqslant s^{-(2+\alpha)}$.

Modèle de régularité géométrique

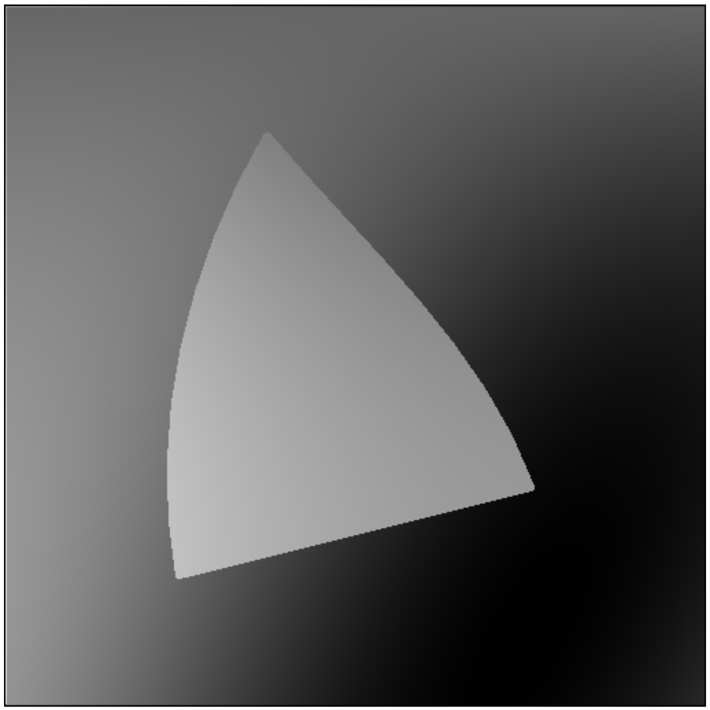


- Fonctions régulières par morceaux avec des discontinuités le long de courbes régulières.
- Cf modèle horizon ou fonctions étoilées (*Tsybakov, Donoho*).
- Fonctions de régularité géométrique $\mathbf{C}^\alpha$:
  - $f = \tilde{f}$ ou $f = \tilde{f} \star h$,
  - $\tilde{f}$: $\mathbf{C}^\alpha(\Lambda)$ avec $\Lambda = [0,1]^2 - \{\mathcal{C}_\gamma\}_{1 \leqslant \gamma \leqslant G}$,
  - $\mathcal{C}_\gamma$: $\mathbf{C}^\alpha$ + conditions géométriques de non tangence,
  - h : $\mathbf{C}^\alpha$ à support $\subset [-s,s]^2$ et $\|h\|_{\mathbf{C}^\alpha} \leqslant s^{-(2+\alpha)}$.
- Risque minimax : $\propto \epsilon^{2\alpha/(\alpha+1)}$.

Approximation en ondelettes des contours

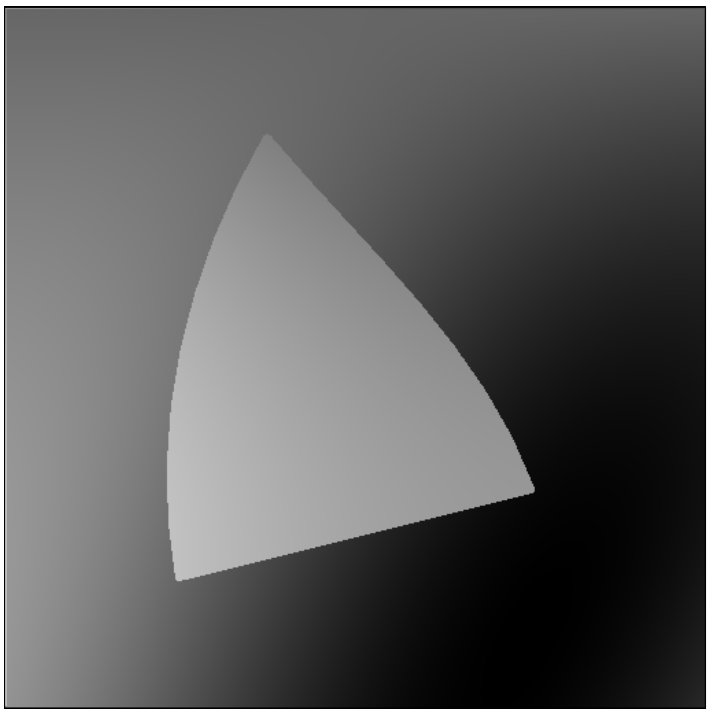
Approximation en ondelettes des contours

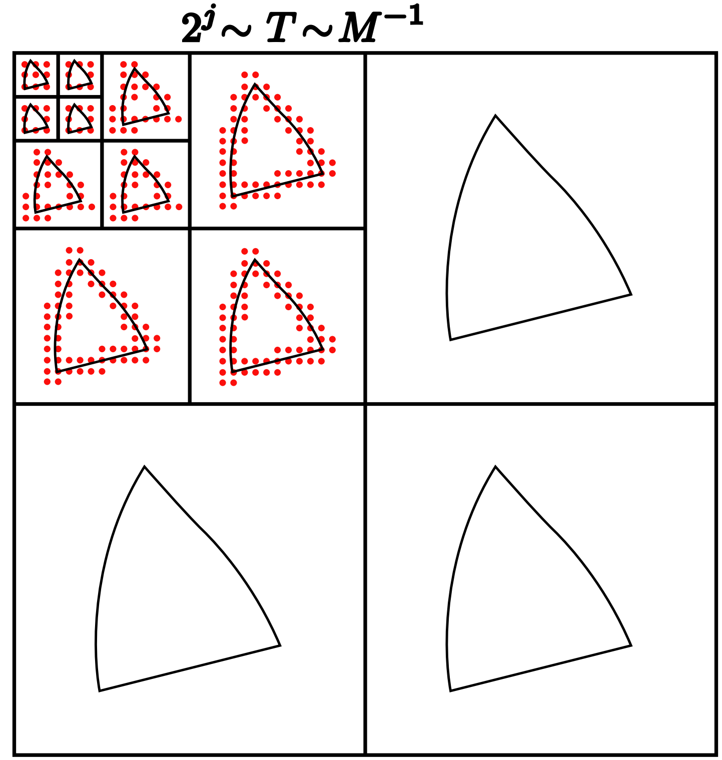
- Approximation de f qui est $\mathbf{C}^\alpha$ en dehors de contours $\mathbf{C}^\alpha$:



Approximation en ondelettes des contours

- Approximation de f qui est $\mathbf{C}^\alpha$ en dehors de contours $\mathbf{C}^\alpha$:

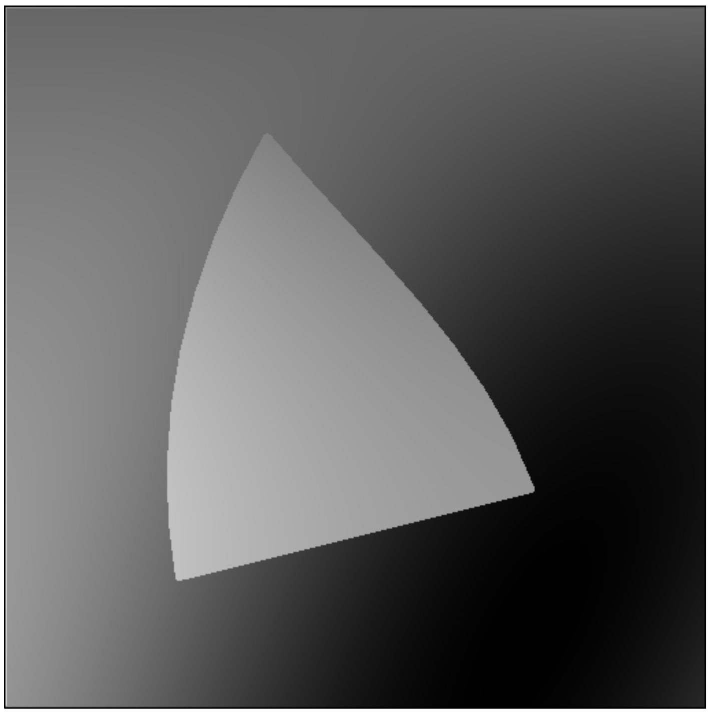


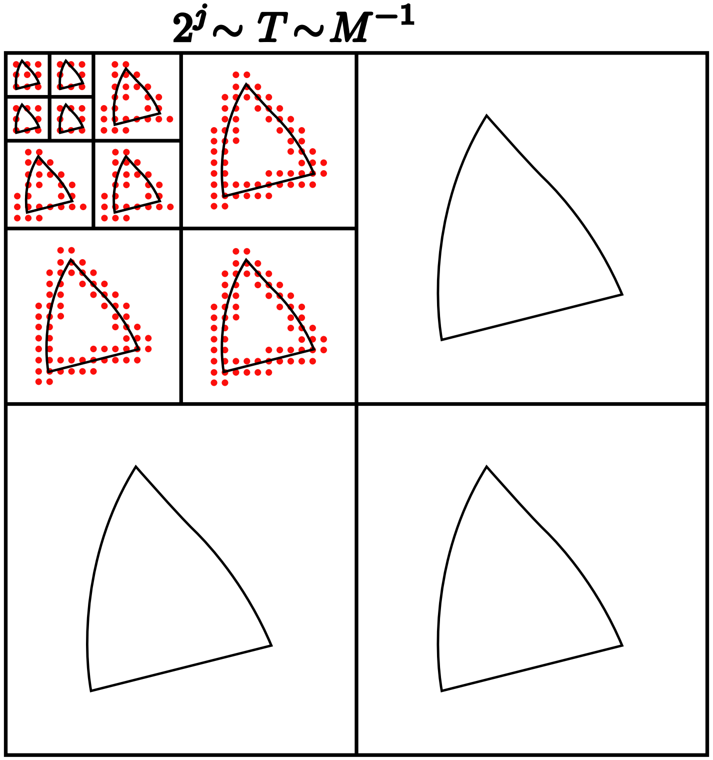


- Avec M ondelettes : $\|f - f_M\|^2 \leqslant C\, M^{-1}$.

Approximation en ondelettes des contours

- Approximation de f qui est $\mathbf{C}^\alpha$ en dehors de contours $\mathbf{C}^\alpha$:



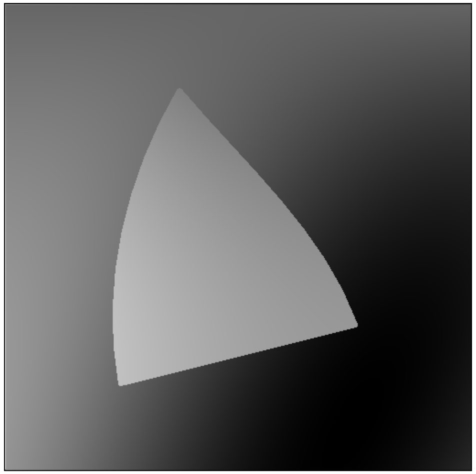


- Avec M ondelettes : $\|f - f_M\|^2 \leqslant C\, M^{-1}$.
- Pour obtenir le risque minimax, besoin de $\|f - f_M\|^2 \leqslant C\, M^{-\alpha}$.

Éléments géométriques pour les contours

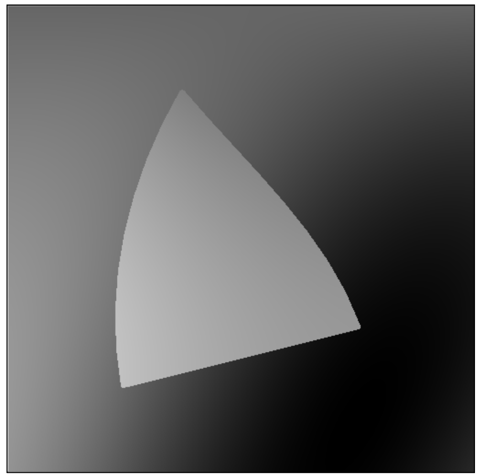
Éléments géométriques pour les contours

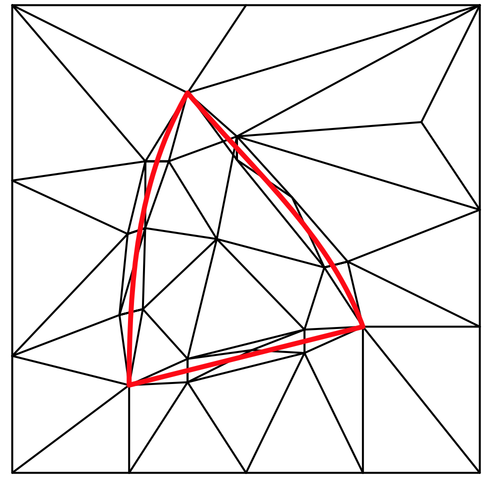
- Approximation de f qui est $\mathbf{C}^\alpha$ en dehors de contours $\mathbf{C}^\alpha$:



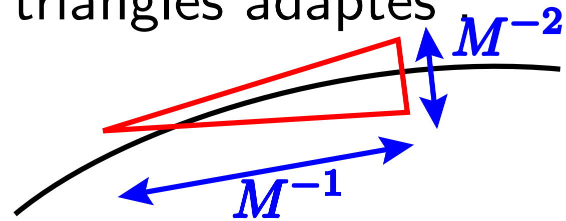
Éléments géométriques pour les contours

- Approximation de f qui est $\mathbf{C}^\alpha$ en dehors de contours $\mathbf{C}^\alpha$:



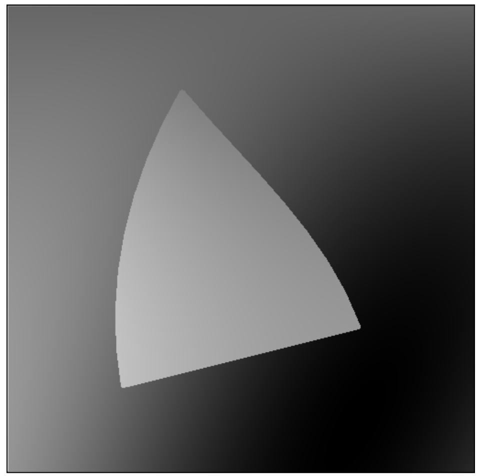
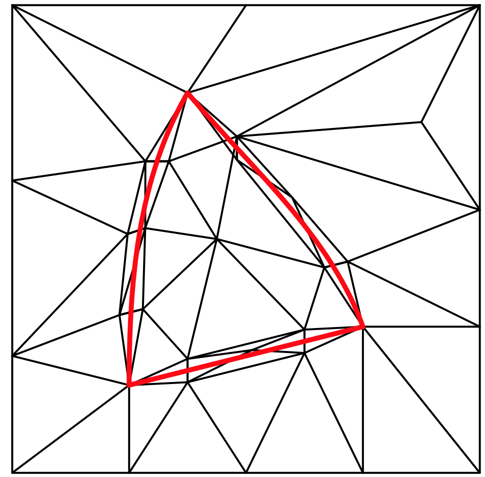


- Approximation linéaire par morceau sur M triangles adaptés :
  si $\alpha \geqslant 2$ alors $\|f - f_M\|^2 \leqslant C\, M^{-2}$,

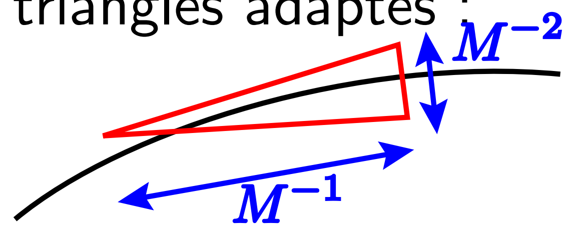


Éléments géométriques pour les contours

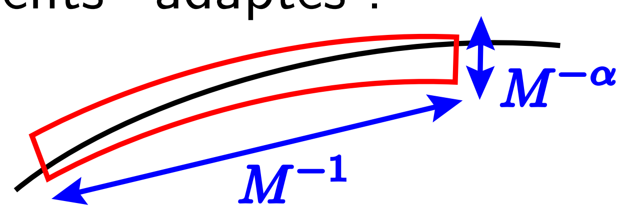
- Approximation de f qui est $\mathbf{C}^\alpha$ en dehors de contours $\mathbf{C}^\alpha$:

- Approximation linéaire par morceau sur M triangles adaptés :
si $\alpha \geqslant 2$ alors $\|f - f_M\|^2 \leqslant C\, M^{-2}$,

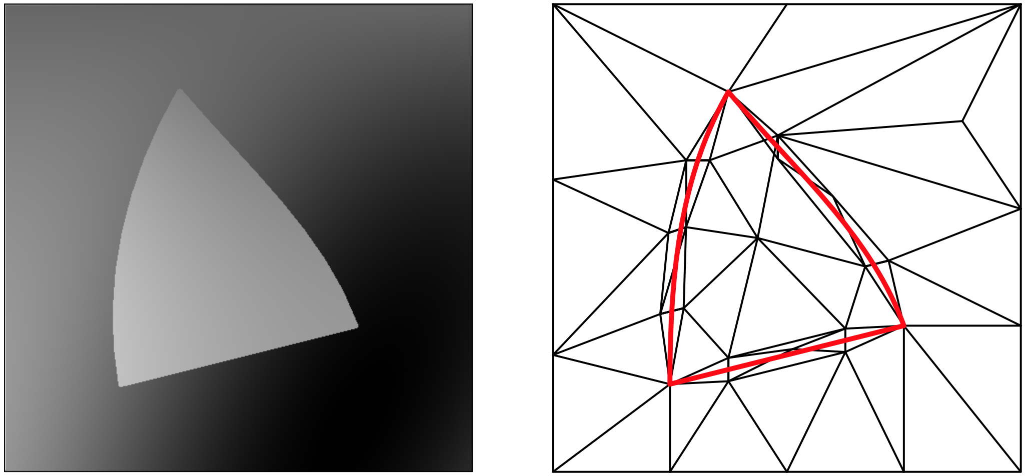


- Approximation d'ordre élevé avec M "éléments" adaptés :
$\|f - f_M\|^2 \leqslant C\, M^{-\alpha}$.

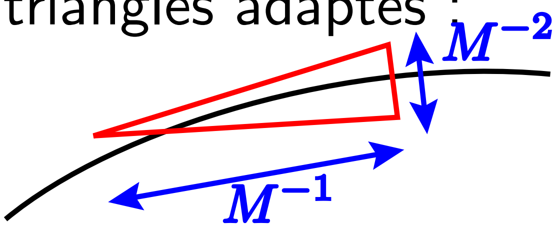


Éléments géométriques pour les contours

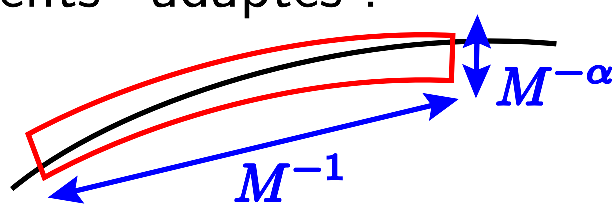
- Approximation de f qui est $\mathbf{C}^\alpha$ en dehors de contours $\mathbf{C}^\alpha$:



- Approximation linéaire par morceau sur M triangles adaptés :
  si $\alpha \geqslant 2$ alors $\|f - f_M\|^2 \leqslant C\, M^{-2}$,



- Approximation d'ordre élevé avec M "éléments" adaptés :
  $\|f - f_M\|^2 \leqslant C\, M^{-\alpha}$.

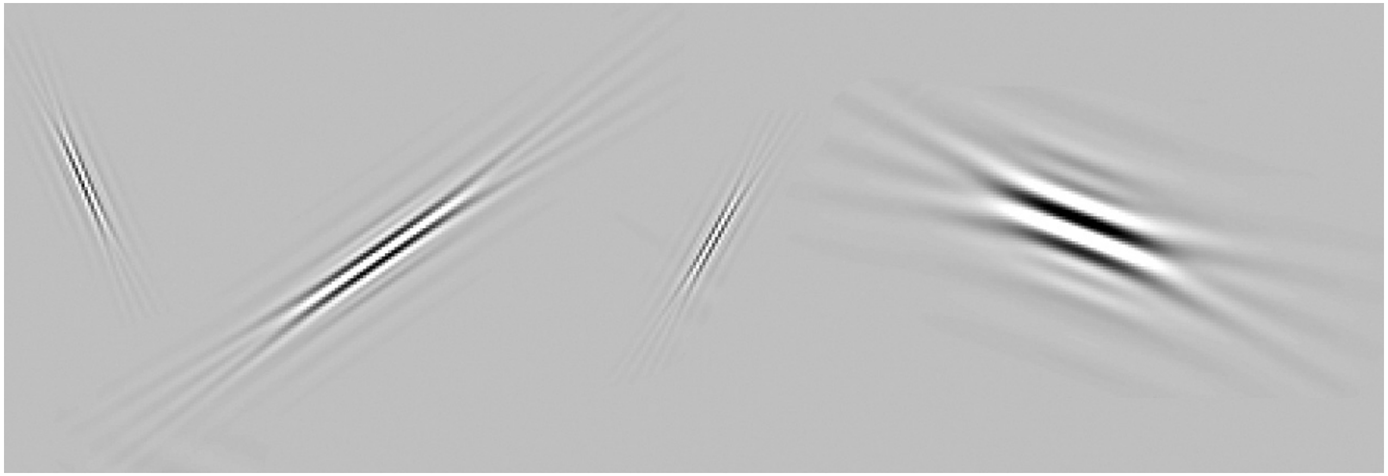


- Difficile de trouver une solution optimale mais bonnes solutions "gloutonnes" (*Dekel,Demaret, Dyn, Iske*).

Approximation en curvelets

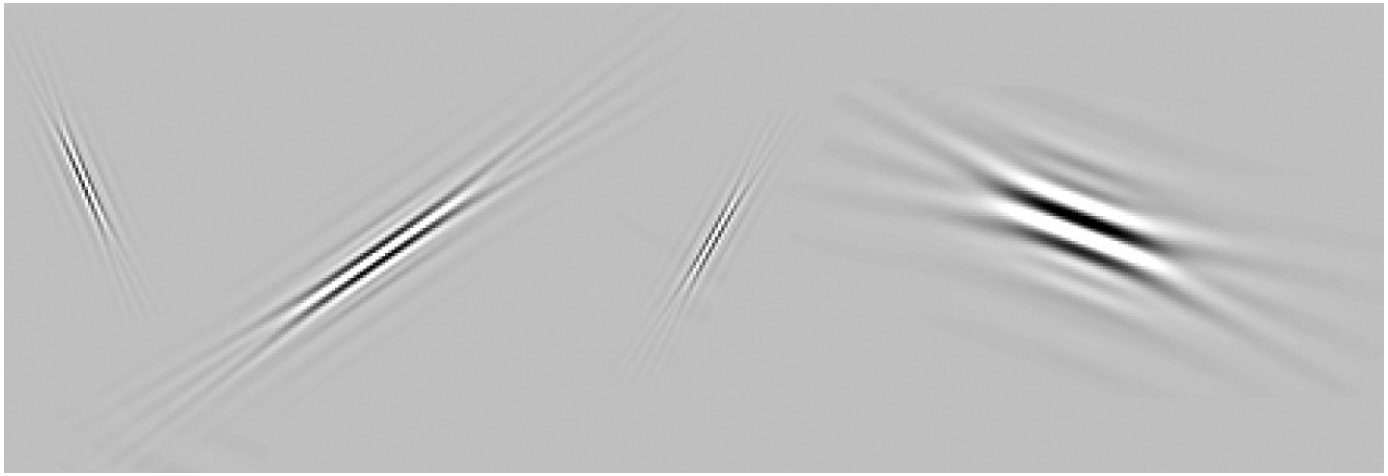
Approximation en curvelets

- Les curvelets définissent un "tight frame" de $L^2[0,1]^2$ avec des éléments allongés et orientés (*Candes, Donoho*) : $\{c_j(R_\theta x - \eta)\}_{j,\theta,\eta}$



Approximation en curvelets

- Les curvelets définissent un "tight frame" de $L^2[0,1]^2$ avec des éléments allongés et orientés (*Candes, Donoho*) : $\{c_j(R_\theta x - \eta)\}_{j,\theta,\eta}$

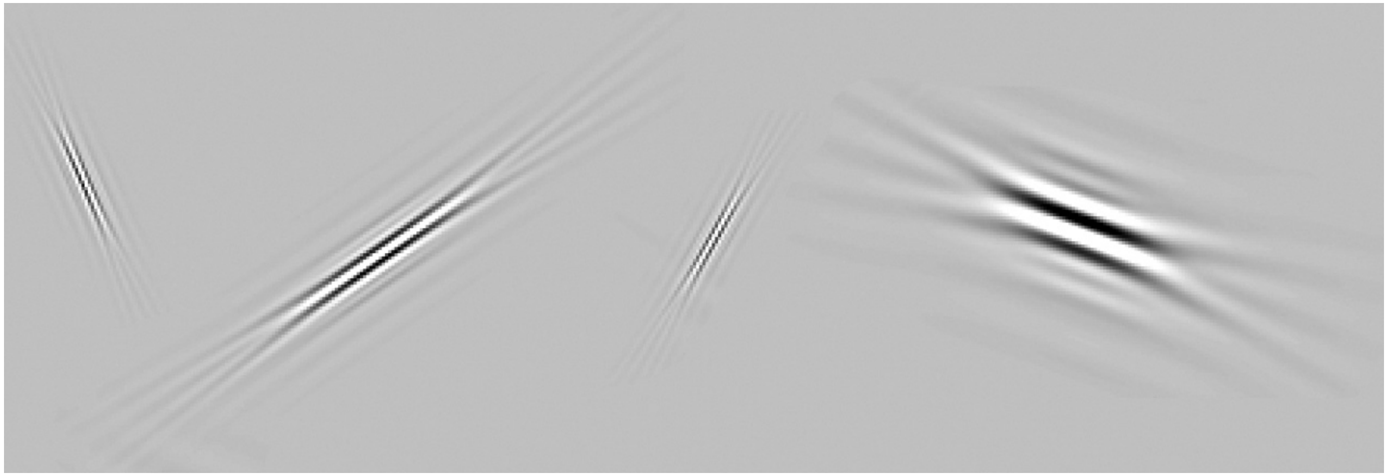


- Si f est de régularité géométrique $\mathbf{C}^\alpha$ alors avec M curvelets :

$$\|f - f_M\|^2 \leqslant C\,(\log M)^3\, M^{-2} \quad \text{si } \alpha \geqslant 2.$$

Approximation en curvelets

- Les curvelets définissent un "tight frame" de $L^2[0,1]^2$ avec des éléments allongés et orientés (*Candes, Donoho*) : $\{c_j(R_\theta x - \eta)\}_{j,\theta,\eta}$



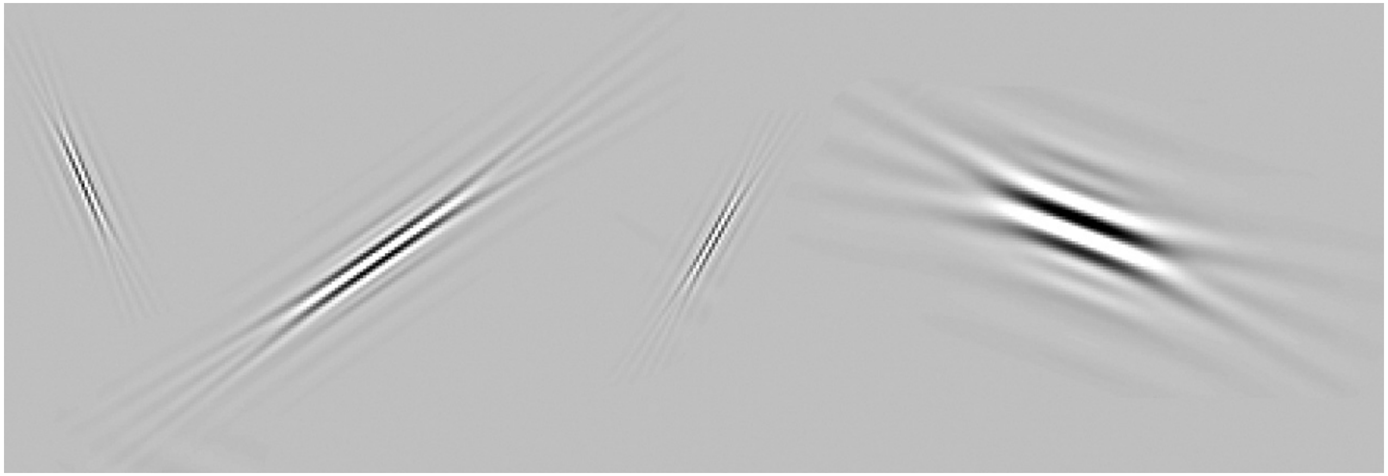
- Si f est de régularité géométrique $\mathbf{C}^\alpha$ alors avec M curvelets :

$$\|f - f_M\|^2 \leqslant C\,(\log M)^3\,M^{-2} \quad \text{si } \alpha \geqslant 2.$$

- Quasi optimal pour $\alpha = 2$.

Approximation en curvelets

- Les curvelets définissent un "tight frame" de $L^2[0,1]^2$ avec des éléments allongés et orientés (*Candes, Donoho*) : $\{c_j(R_\theta x - \eta)\}_{j,\theta,\eta}$



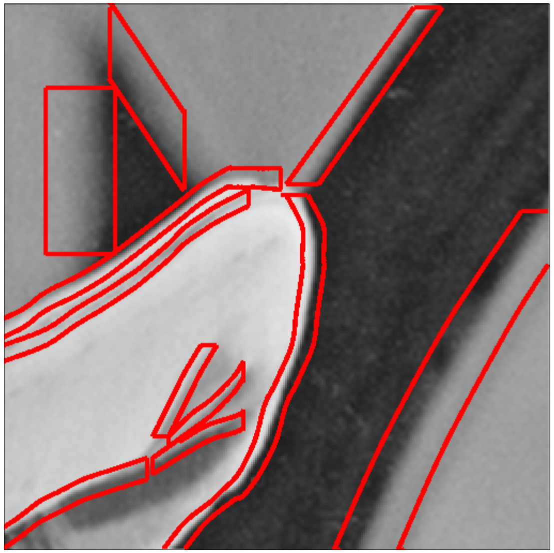
- Si f est de régularité géométrique $\mathbf{C}^\alpha$ alors avec M curvelets :

$$\|f - f_M\|^2 \leqslant C\,(\log M)^3\, M^{-2} \quad \text{si } \alpha \geqslant 2.$$

- Quasi optimal pour $\alpha = 2$.
- Difficile d'obtenir des bases orthogonales ou des bases de Riesz : (*Vetterli & Minh Do*).

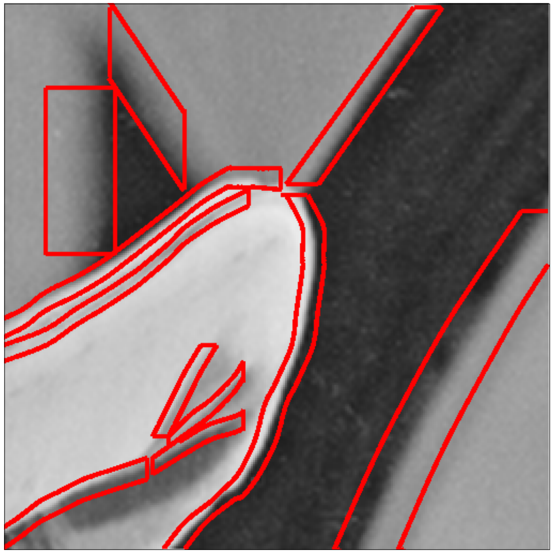
Bandelettes

Bandelettes

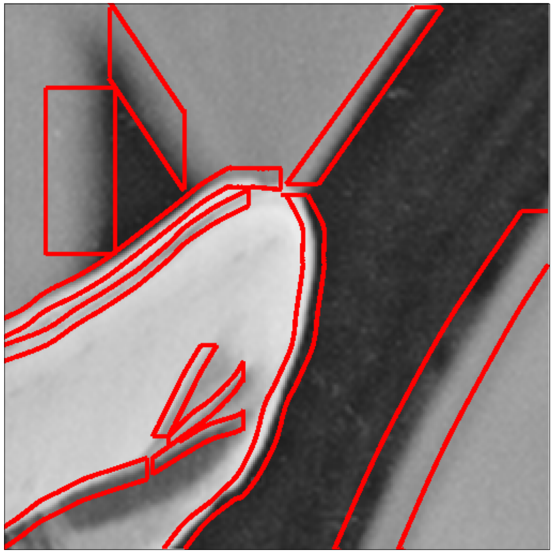
- Famille de base de bandelettes : bases adaptées à la géométrie des images.

Bandelettes

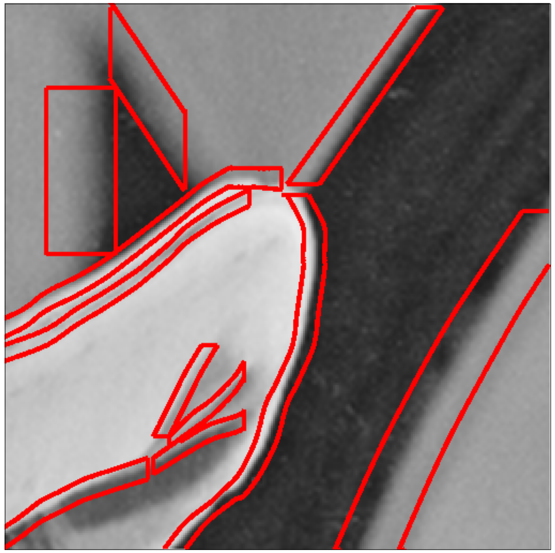
- Famille de base de bandelettes : bases adaptées à la géométrie des images.
- Introduites pour un problème de compression (frame) et modifiées pour obtenir des bases orthonormées.

Bandelettes

- Famille de base de bandelettes : bases adaptées à la géométrie des images.
- Introduites pour un problème de compression (frame) et modifiées pour obtenir des bases orthonormées.
- 3 ingrédients :
  - segmentation hiérarchique (strucrure d'arbre),
  - déformation géométrique locale,
  - ondelettes hyperboliques.

Bandelettes

- Famille de base de bandelettes : bases adaptées à la géométrie des images.
- Introduites pour un problème de compression (frame) et modifiées pour obtenir des bases orthonormées.
- 3 ingrédients :
  - segmentation hiérarchique (strucrure d'arbre),
  - déformation géométrique locale,
  - ondelettes hyperboliques.
- Représentation optimale et algorithme de recherche de la meilleure représentation.

Flot géométrique segmenté

Flot géométrique segmenté

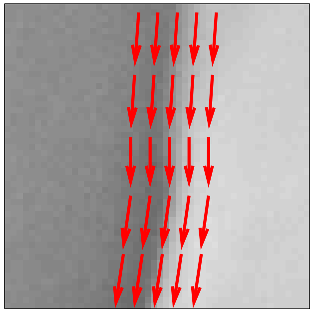
- Spécification de la géométrie.

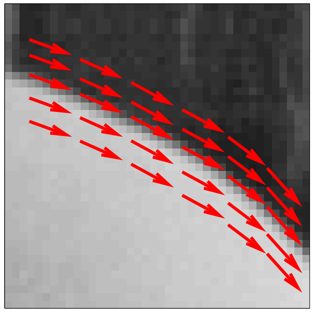
Flot géométrique segmenté

- Spécification de la géométrie.
- Flot géométrique : champ de vecteurs $\vec{\tau}(x_1, x_2)$ donnant des directions dans lesquelles l'image est localement régulière.

Flot géométrique segmenté

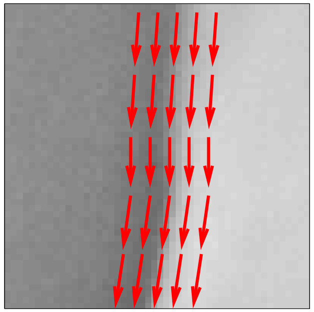
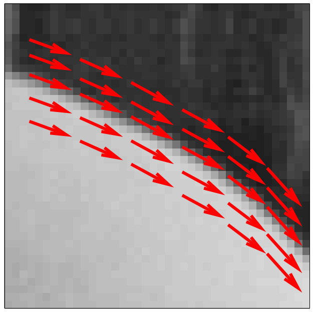
- Spécification de la géométrie.
- Flot géométrique : champ de vecteurs $\vec{\tau}(x_1, x_2)$ donnant des directions dans lesquelles l'image est localement régulière.
- Dans une région, le flot est constant horizontalement ou verticalement et induit par un polynôme.



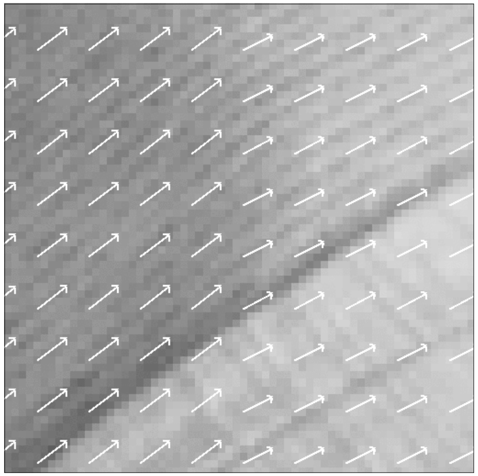
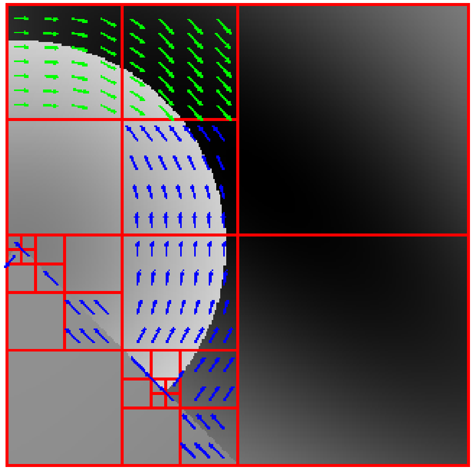


Flot géométrique segmenté

- Spécification de la géométrie.
- Flot géométrique : champ de vecteurs $\vec{\tau}(x_1, x_2)$ donnant des directions dans lesquelles l'image est localement régulière.
- Dans une région, le flot est constant horizontalement ou verticalement et induit par un polynôme.

- Nécessité d'une segmentation de l'image.

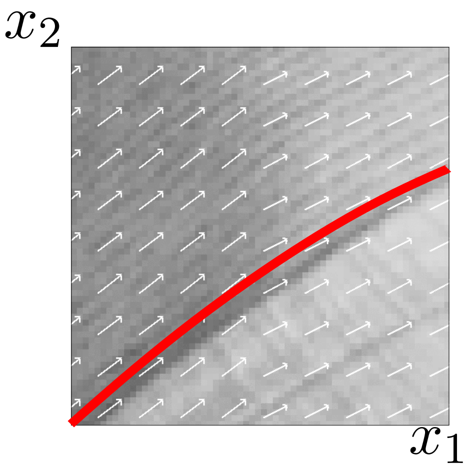



Base d'ondelettes déformées

Base d'ondelettes déformées

- Supposons le flot constant verticalement :

$$\vec{\tau}(x_1, x_2) = (1\,,\, c'(x_1))\ .$$

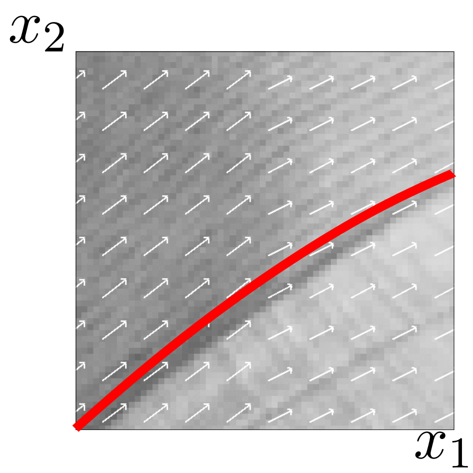


$c(x_1) = \int_{x_{1,\min}}^{x_1} c'(u)\,\mathrm{d}u$

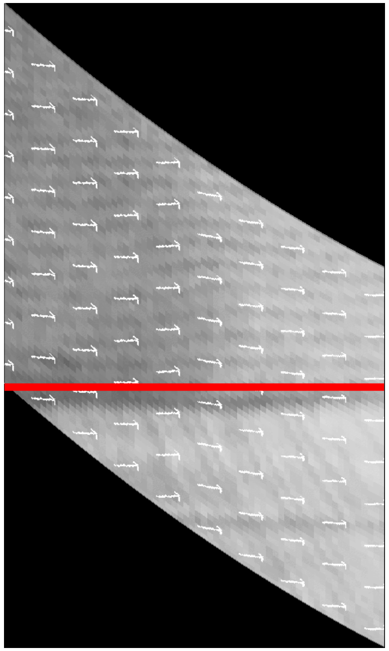
Base d'ondelettes déformées

- Supposons le flot constant verticalement :

$$\vec{\tau}(x_1, x_2) = (1\,,\, c'(x_1))\ .$$



$c(x_1) = \int_{x_{1,\min}}^{x_1} c'(u)\,\mathrm{d}u$

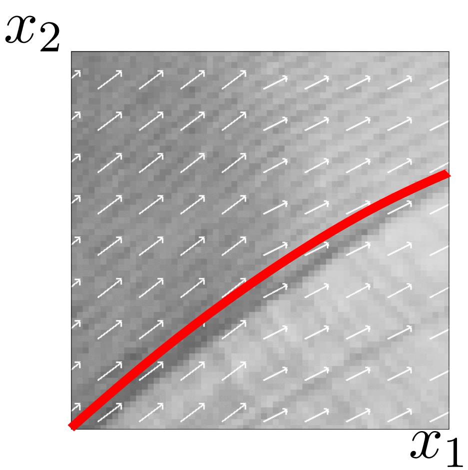


- À x_2 fixé, $f(x_1, x_2 + c(x_1))$ est une fonction régulière de x_1.

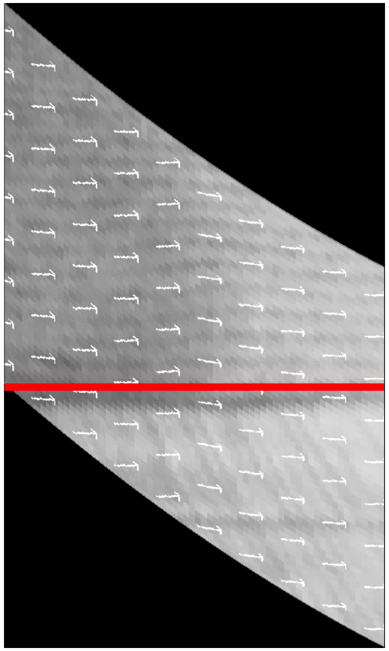
Base d'ondelettes déformées

- Supposons le flot constant verticalement :

$$\vec{\tau}(x_1, x_2) = (1\,,\, c'(x_1))\ .$$



$c(x_1) = \int_{x_{1,\min}}^{x_1} c'(u)\, \mathrm{d}u$

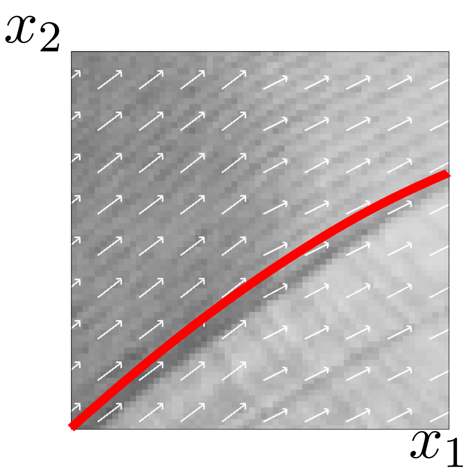


- À x_2 fixé, $f(x_1, x_2 + c(x_1))$ est une fonction régulière de x_1.
- $\langle f(x_1, x_2 + c(x_1))\,,\, \Psi(x_1, x_2)\rangle = \langle f(x_1, x_2)\,,\, \Psi(x_1, x_2 - c(x_1))\rangle$.

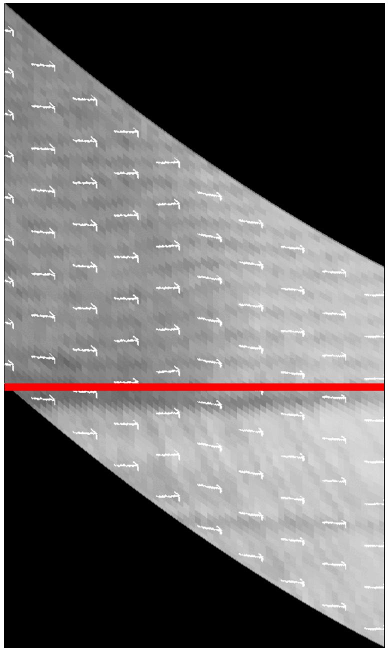
Base d'ondelettes déformées

- Supposons le flot constant verticalement :

$$\vec{\tau}(x_1, x_2) = (1\,,\, c'(x_1))\ .$$



$c(x_1) = \int_{x_{1,\min}}^{x_1} c'(u)\,\mathrm{d}u$



- À x_2 fixé, $f(x_1, x_2 + c(x_1))$ est une fonction régulière de x_1.
- $\langle f(x_1, x_2 + c(x_1))\,,\, \Psi(x_1, x_2)\rangle = \langle f(x_1, x_2)\,,\, \Psi(x_1, x_2 - c(x_1))\rangle\ .$

- On décompose donc f dans une *base d'ondelettes déformées* de $L^2(\Omega)$:

$$\left\{ \begin{array}{ccc} \phi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2 - c(x_1)) & , & \psi_{j,m_1}(x_1)\,\phi_{j,m_2}(x_2 - c(x_1)) \\ & , & \psi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2 - c(x_1)) \end{array} \right\}_{j,m_1,m_2}\ .$$

Bandelettisation

Bandelettisation

- Les fonctions de la base doivent avoir des moments nuls selon x_1 (direction du flot).

Bandelettisation

- Les fonctions de la base doivent avoir des moments nuls selon x_1 (direction du flot).

- Base d'ondelettes déformées de $L^2(\Omega)$:

$$\left\{ \begin{array}{ccc} \underbrace{\phi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2 - c(x_1))}_{\text{Isotrope}} & , & \psi_{j,m_1}(x_1)\,\phi_{j,m_2}(x_2 - c(x_1)) \\ & , & \underbrace{\psi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2 - c(x_1))}_{\text{Isotrope}} \end{array} \right\}_{\substack{j \\ m_1, m_2}} .$$

Bandelettisation

- Les fonctions de la base doivent avoir des moments nuls selon x_1 (direction du flot).
- Bandelettisation : remplace $\{\phi_{j,m_1}(x_1)\}_{m_1}$ par une famille d'ondelettes $\{\psi_{l,m_1}(x_1)\}_{l>j\,,\,m_1}$ qui génère le même espace.
- Base d'ondelettes déformées de $L^2(\Omega)$:

$$\left\{ \begin{array}{ccc} \phi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2-c(x_1)) & , & \psi_{j,m_1}(x_1)\,\phi_{j,m_2}(x_2-c(x_1)) \\ & , & \psi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2-c(x_1)) \end{array} \right\}_{\substack{j \\ m_1,m_2}} .$$

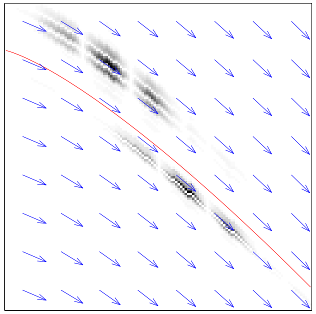
Isotrope Isotrope

Bandelettisation

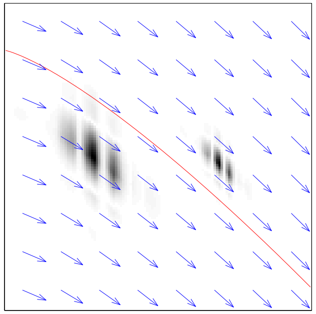
- Les fonctions de la base doivent avoir des moments nuls selon x_1 (direction du flot).
- Bandelettisation : remplace $\{\phi_{j,m_1}(x_1)\}_{m_1}$ par une famille d'ondelettes $\{\psi_{l,m_1}(x_1)\}_{l>j\,,\,m_1}$ qui génère le même espace.
- Obtention d'une *base de bandelettes* de $L^2(\Omega)$:

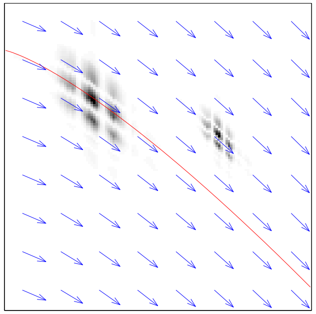
$$\left\{ \begin{array}{lcl} \psi_{l,m_1}(x_1)\,\psi_{j,m_2}(x_2 - c(x_1)) & , & \psi_{j,m_1}(x_1)\,\phi_{j,m_2}(x_2 - c(x_1)) \\ & , & \psi_{j,m_1}(x_1)\,\psi_{j,m_2}(x_2 - c(x_1)) \end{array} \right\}_{\substack{j,l>j \\ m_1,m_2}} .$$

Anisotrope



Isotrope





Base segmentée de bandelettes

Base segmentée de bandelettes

- Le support de l'image est segmenté en régions munies soit
  - d'une base de bandelettes à flot constant verticalement,
  - d'une base de bandelettes à flot constant horizontalement,
  - d'une base d'ondelettes sans flot (régularité isotrope).

Base segmentée de bandelettes

- Le support de l'image est segmenté en régions munies soit
  - d'une base de bandelettes à flot constant verticalement,
  - d'une base de bandelettes à flot constant horizontalement,
  - d'une base d'ondelettes sans flot (régularité isotrope).

- Transformée rapide en bandelettes ($O(N^2)$) :
  - rééchantillonnage, transformée en ondelettes déformées, bandelettisation.

Base segmentée de bandelettes

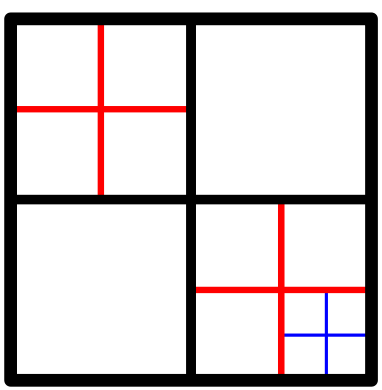
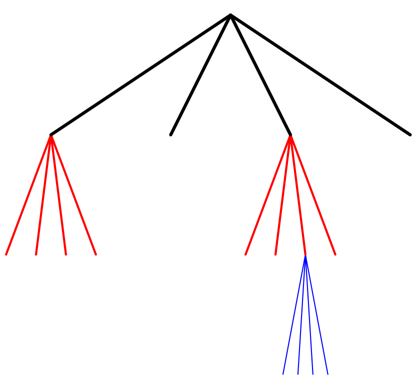
- Le support de l'image est segmenté en régions munies soit
  - d'une base de bandelettes à flot constant verticalement,
  - d'une base de bandelettes à flot constant horizontalement,
  - d'une base d'ondelettes sans flot (régularité isotrope).

- Transformée rapide en bandelettes $(O(N^2))$:
  - rééchantillonnage, transformée en ondelettes déformées, bandelettisation.
- Pas de discontinuités aux frontières grâce à un schéma de lifting adapté.

Approximation M termes

Approximation M termes

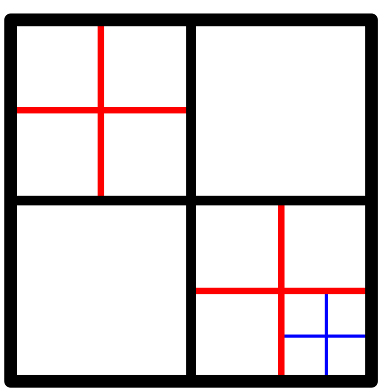
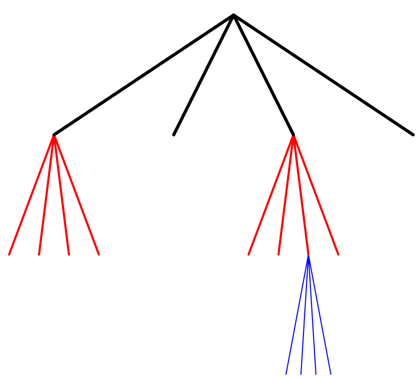
- Une approximation en bandelettes est donnée par :
  - une segmentation en rectangle, représentée par les M_s nœuds intérieurs de l'arbre de la segmentation,
  - à l'intérieur de chaque carré Ω_i de la segmentation par :
    - $M_{g,i}$ coefficients du flot géométrique,
    - $M_{b,i}$ coefficients de bandelettes au dessus d'un seuil T.


Approximation M termes

- Une approximation en bandelettes est donnée par :
  - une segmentation en rectangle, représentée par les M_s nœuds intérieurs de l'arbre de la segmentation,
  - à l'intérieur de chaque carré Ω_i de la segmentation par :
    - $M_{g,i}$ coefficients du flot géométrique,
    - $M_{b,i}$ coefficients de bandelettes au dessus d'un seuil T.
- Nombre total de paramètres :

$$M = M_s + \sum_i \left(M_{g,i} + M_{b,i} \right) = \sum_i \left(M_{s,i} + M_{g,i} + M_{b,i} \right) .$$

Optimisation de l'approximation

Optimisation de l'approximation

- Minimiser $\|f - f_M\|^2$ pour un nombre fixé M de paramètres.

Optimisation de l'approximation

- Minimiser $\|f - f_M\|^2$ pour un nombre fixé M de paramètres.
- Approche lagrangienne : trouver le meilleur flot géométrique segmenté qui minimise

$$\|f - f_M\|^2 + T^2\, M \ .$$

Optimisation de l'approximation

- Minimiser $\|f - f_M\|^2$ pour un nombre fixé M de paramètres.
- Approche lagrangienne : trouver le meilleur flot géométrique segmenté qui minimise

$$\|f - f_M\|^2 + T^2\, M \ .$$

- Additivité du Lagrangien :

$$\|f - f_M\|^2 + T^2\, M = \sum_i \|f - f_M\|^2_{\Omega_i} + T^2\, M_i \quad .$$

Optimisation de l'approximation

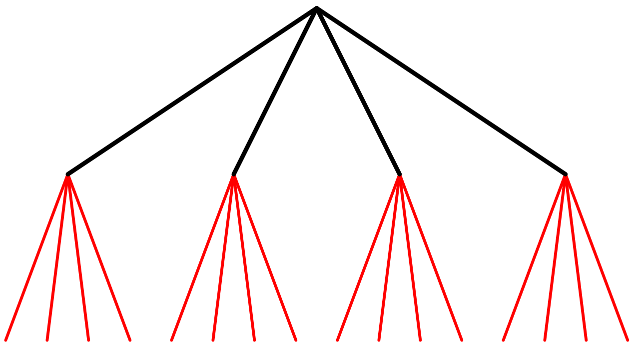
- Minimiser $\|f - f_M\|^2$ pour un nombre fixé M de paramètres.
- Approche lagrangienne : trouver le meilleur flot géométrique segmenté qui minimise

$$\|f - f_M\|^2 + T^2\, M\ .$$

- Additivité du Lagrangien :

$$\|f - f_M\|^2 + T^2\, M = \sum_i \|f - f_M\|^2_{\Omega_i} + T^2\, M_i \quad .$$

- Algorithme rapide (CART) : programmation dynamique de bas en haut sur la segmentation en arbre.



Optimisation de l'approximation

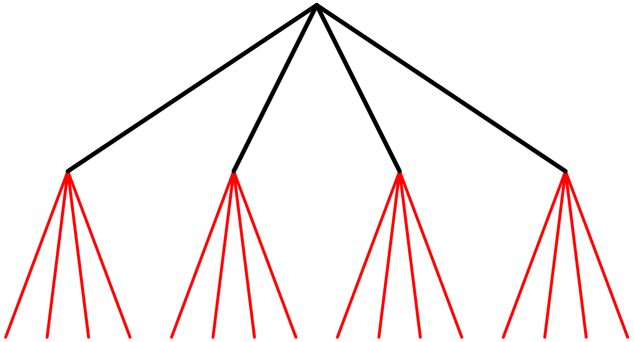
- Minimiser $\|f - f_M\|^2$ pour un nombre fixé M de paramètres.
- Approche lagrangienne : trouver le meilleur flot géométrique segmenté qui minimise

$$\|f - f_M\|^2 + T^2\, M \ .$$

- Additivité du Lagrangien :

$$\|f - f_M\|^2 + T^2\, M = \sum_i \|f - f_M\|^2_{\Omega_i} + T^2\, M_i \quad .$$

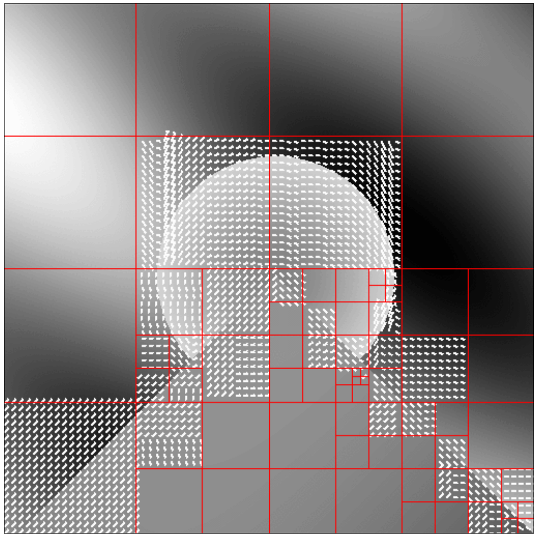
- Algorithme rapide (CART) : programmation dynamique de bas en haut sur la segmentation en arbre.
- Complexité polynomiale : nombre total de vecteurs de base et non pas le nombre total de bases



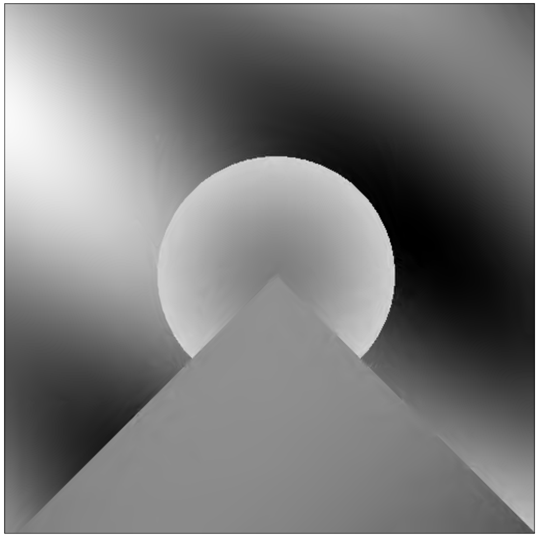
Fonction régulière par morceaux

Fonction régulière par morceaux

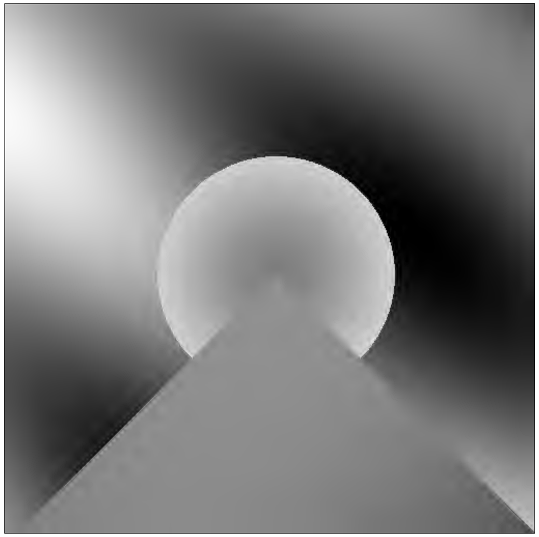
M=2650



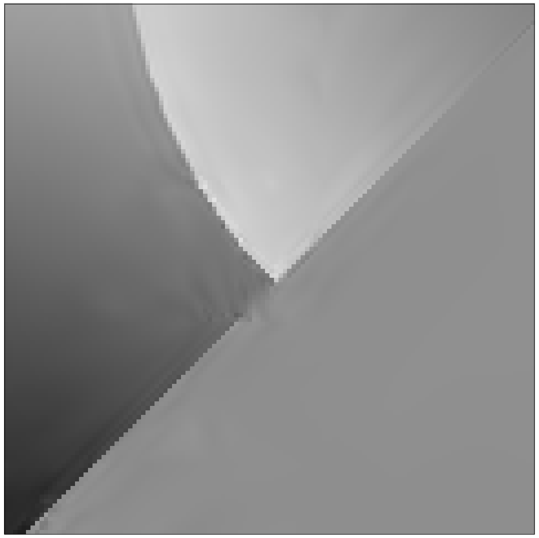
PSNR = $45{,}97$ dB



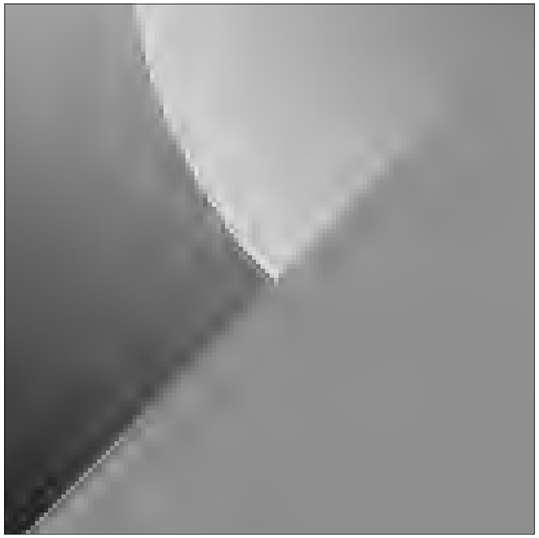
PSNR = $40{,}17$ dB



Bandelettes



Ondelettes

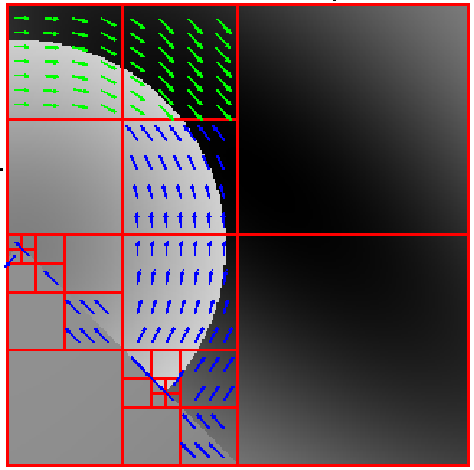


Théorème d'approximation

Théorème d'approximation

- **Théorème :** Si f est de régularité géométrique $\mathbf{C}^\alpha$ ($f = \tilde{f}$ ou $f = \tilde{f} \star h$ avec $\tilde{f}$ $\mathbf{C}^\alpha$ en dehors de courbes $\mathbf{C}^\alpha$ par morceaux avec des conditions de non tangence) alors

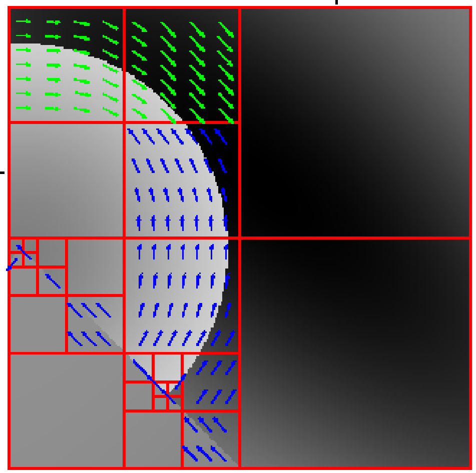
$$\|f - f_M\|^2 \leqslant C\,(\log M)^{\alpha+1} M^{-\alpha} \ .$$



Théorème d'approximation

- **Théorème :** Si f est de régularité géométrique $\mathbf{C}^\alpha$ ($f = \tilde{f}$ ou $f = \tilde{f} \star h$ avec $\tilde{f}$ $\mathbf{C}^\alpha$ en dehors de courbes $\mathbf{C}^\alpha$ par morceaux avec des conditions de non tangence) alors

$$\|f - f_M\|^2 \leqslant C\,(\log M)^{\alpha+1} M^{-\alpha}\ .$$

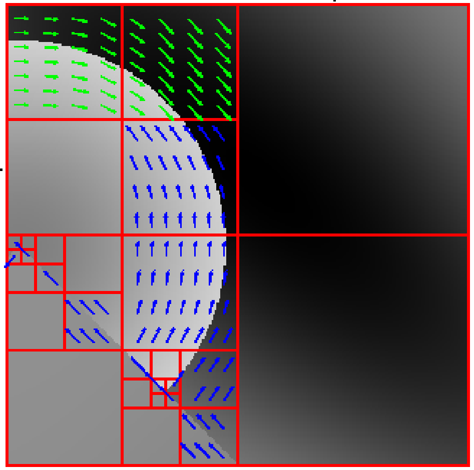


- Adaptivité : degré de régularité α inconnu.

Théorème d'approximation

- **Théorème :** Si f est de régularité géométrique $\mathbf{C}^\alpha$ ($f = \tilde{f}$ ou $f = \tilde{f} \star h$ avec $\tilde{f}$ $\mathbf{C}^\alpha$ en dehors de courbes $\mathbf{C}^\alpha$ par morceaux avec des conditions de non tangence) alors

$$\|f - f_M\|^2 \leqslant C\,(\log M)^{\alpha+1} M^{-\alpha} \ .$$

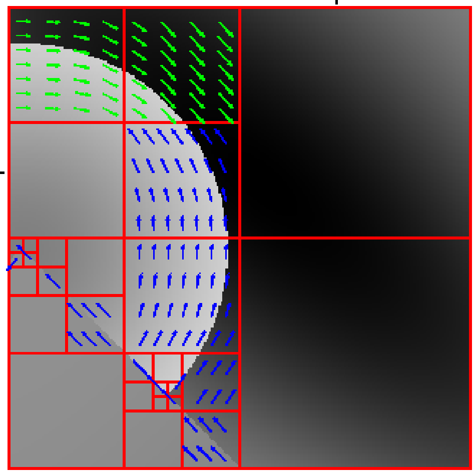


- Adaptivité : degré de régularité α inconnu.
- Optimalité : exposant de décroissance α.

Théorème d'approximation

- **Théorème :** Si f est de régularité géométrique $\mathbf{C}^\alpha$ ($f = \tilde{f}$ ou $f = \tilde{f} \star h$ avec $\tilde{f}$ $\mathbf{C}^\alpha$ en dehors de courbes $\mathbf{C}^\alpha$ par morceaux avec des conditions de non tangence) alors

$$\|f - f_M\|^2 \leqslant C\,(\log M)^{\alpha+1} M^{-\alpha} \ .$$



- Adaptivité : degré de régularité α inconnu.
- Optimalité : exposant de décroissance α.
- Comparaison :
  - Ondelettes isotropes : $\|f - f_M\|^2 \leqslant C\,M^{-1}$.
  - Curvelets : $\|f - f_M\|^2 \leqslant C\,(\log M)^3\,M^{-2}$.

Sélection de modèle

Sélection de modèle

- Approximation non linéaire : minimisation de $\|f - f_M\|^2 + T^2 M$.

Sélection de modèle

- Approximation non linéaire : minimisation de $\|f - f_M\|^2 + T^2 M$.
- Pour une base, équivalence avec du seuillage.

Sélection de modèle

- Approximation non linéaire : minimisation de $\|f - f_M\|^2 + T^2 M$.
- Pour une base, équivalence avec du seuillage.
- Algorithme de pénalisation pour la sélection de modèle :

$$F = \underset{\tilde{f} = P_{\mathcal{M}} Y}{\operatorname{argmin}} \|Y - \tilde{f}\|^2 + T^2 \operatorname{dim}(\mathcal{M})$$

où $\mathcal{M}$ parcourt les espace engendrés par des sous familles de bases orthonormées d'un sous espace de dimension N.

Sélection de modèle

- Approximation non linéaire : minimisation de $\|f - f_M\|^2 + T^2 M$.
- Pour une base, équivalence avec du seuillage.
- Algorithme de pénalisation pour la sélection de modèle :

$$F = \operatorname*{argmin}_{\tilde{f} = P_{\mathcal{M}} Y} \|Y - \tilde{f}\|^2 + T^2 \,\mathrm{dim}(\mathcal{M})$$

  où $\mathcal{M}$ parcourt les espace engendrés par des sous familles de bases orthonormées d'un sous espace de dimension N.
- **Théorème** (Donoho, Birge, Massart, Nowak,. . .) :
  Si le nombre total de vecteurs dans les différentes bases est au plus polynomial en N alors, pour λ suffisamment grand et $T = \lambda\sqrt{\log N}\epsilon$,

$$E(\|F - f\|^2) \leqslant C \log N (E(\|F_O - f\|^2) + \epsilon/N)$$

Sélection de modèle

- Approximation non linéaire : minimisation de $\|f - f_M\|^2 + T^2 M$.
- Pour une base, équivalence avec du seuillage.
- Algorithme de pénalisation pour la sélection de modèle :

$$F = \underset{\tilde{f} = P_{\mathcal{M}} Y}{\text{argmin}} \|Y - \tilde{f}\|^2 + T^2 \,\text{dim}(\mathcal{M})$$

  où $\mathcal{M}$ parcourt les espace engendrés par des sous familles de bases orthonormées d'un sous espace de dimension N.
- **Théorème** (Donoho, Birge, Massart, Nowak,...) :
  Si le nombre total de vecteurs dans les différentes bases est au plus polynomial en N alors, pour λ suffisamment grand et $T = \lambda \sqrt{\log N} \epsilon$,

$$E(\|F - f\|^2) \leqslant C \log N (E(\|F_O - f\|^2) + \epsilon/N)$$

- Quasi optimalité de la méthode.

Sélection de modèle

- Approximation non linéaire : minimisation de $\|f - f_M\|^2 + T^2 M$.
- Pour une base, équivalence avec du seuillage.
- Algorithme de pénalisation pour la sélection de modèle :

$$F = \underset{\tilde{f}=P_{\mathcal{M}}Y}{\text{argmin}} \|Y - \tilde{f}\|^2 + T^2 \,\text{dim}(\mathcal{M})$$

  où $\mathcal{M}$ parcourt les espace engendrés par des sous familles de bases orthonormées d'un sous espace de dimension N.
- **Théorème** (Donoho, Birge, Massart, Nowak,. . .) :
  Si le nombre total de vecteurs dans les différentes bases est au plus polynomial en N alors, pour λ suffisamment grand et $T = \lambda\sqrt{\log N}\epsilon$,

$$E(\|F - f\|^2) \leqslant C \log N (E(\|F_O - f\|^2) + \epsilon/N)$$

- Quasi optimalité de la méthode.
- Possibilité pratique de faire la minimisation ?

Esquisse de démonstration

Esquisse de démonstration

- Preuve d'une inégalité similaire en Probabilité.

Esquisse de démonstration

- Preuve d'une inégalité similaire en Probabilité.
- Inégalité de concentration pour les processus gaussiens (Borell, Cirel'son, Ibragimov, Sudakov).

Esquisse de démonstration

- Preuve d'une inégalité similaire en Probabilité.
- Inégalité de concentration pour les processus gaussiens (Borell, Cirel'son, Ibragimov, Sudakov).
- Tableau. . .

Quasi optimalité des Bandelettes

Quasi optimalité des Bandelettes

- Contrôle polynomial sur le nombre d'éléments.

Quasi optimalité des Bandelettes

- Contrôle polynomial sur le nombre d'éléments.
- Possibilité d'effectuer la minimisation par l'algorithme de programmation dynamique.

Quasi optimalité des Bandelettes

- Contrôle polynomial sur le nombre d'éléments.
- Possibilité d'effectuer la minimisation par l'algorithme de programmation dynamique.
- Quasi optimalité de la représentation pour l'approximation non linéaire.

Quasi optimalité des Bandelettes

- Contrôle polynomial sur le nombre d'éléments.
- Possibilité d'effectuer la minimisation par l'algorithme de programmation dynamique.
- Quasi optimalité de la représentation pour l'approximation non linéaire.
- **Théorème :** Pour $N \simeq \epsilon^{-1}$ et $T = \lambda\sqrt{\log N}\epsilon$,

$$E(\|F - f\|^2) \leqslant C|\log \epsilon|^{\alpha+2}\,\epsilon^{2\alpha/(\alpha+1)}$$

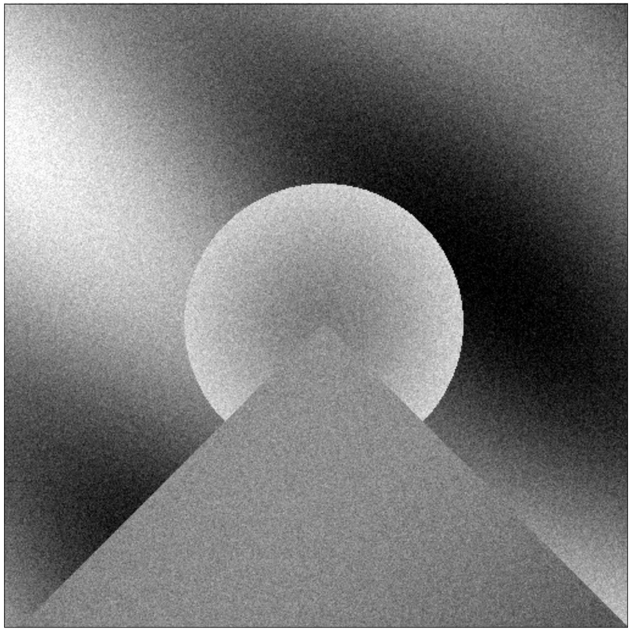
Quasi optimalité des Bandelettes

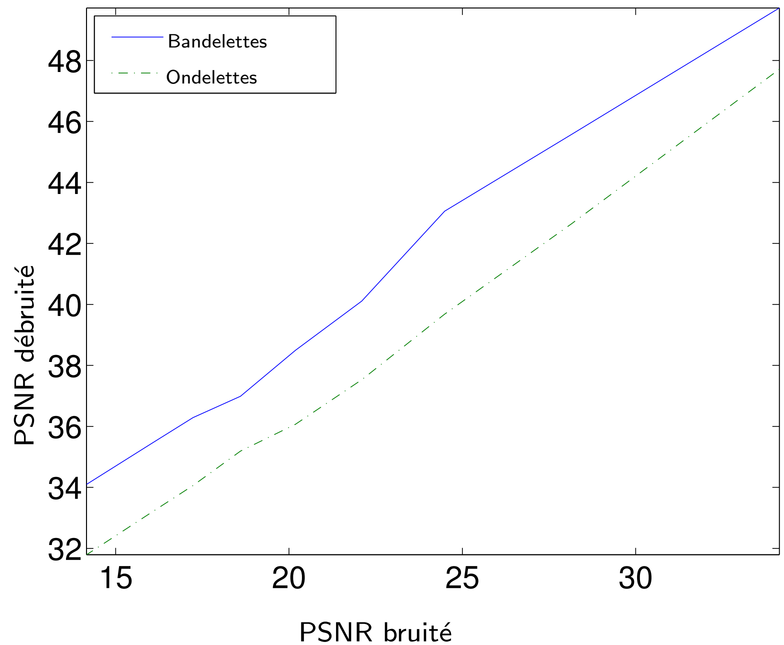
- Contrôle polynomial sur le nombre d'éléments.
- Possibilité d'effectuer la minimisation par l'algorithme de programmation dynamique.
- Quasi optimalité de la représentation pour l'approximation non linéaire.
- **Théorème :** Pour $N \simeq \epsilon^{-1}$ et $T = \lambda\sqrt{\log N}\epsilon$,

$$E(\|F - f\|^2) \leqslant C|\log \epsilon|^{\alpha+2}\, \epsilon^{2\alpha/(\alpha+1)}$$

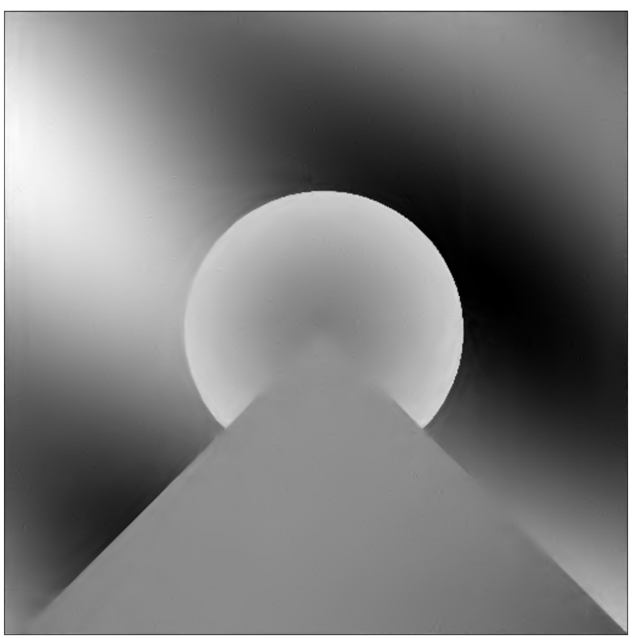
- Adaptivité automatique.

Bruité (20,19 dB)

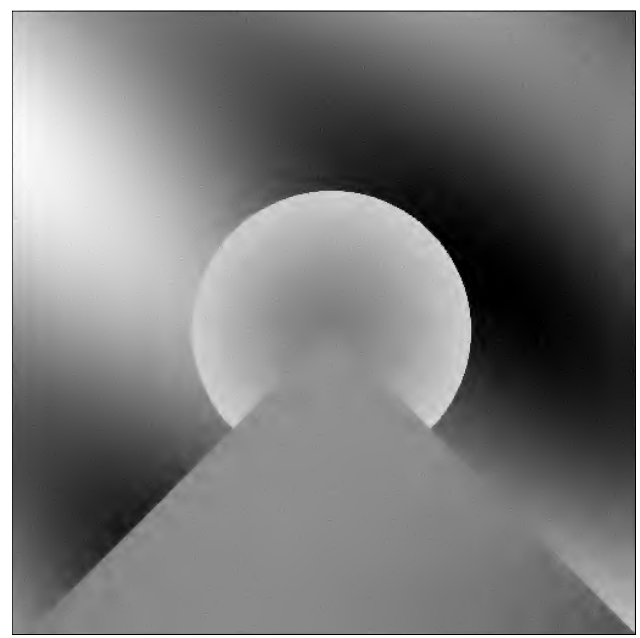




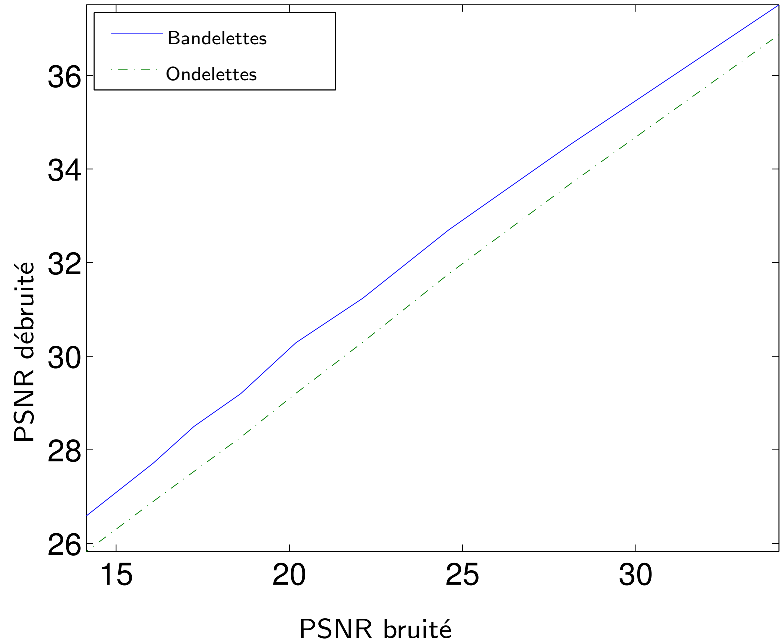
Bandelettes (30,29 dB)



Ondelettes (28,21 dB)

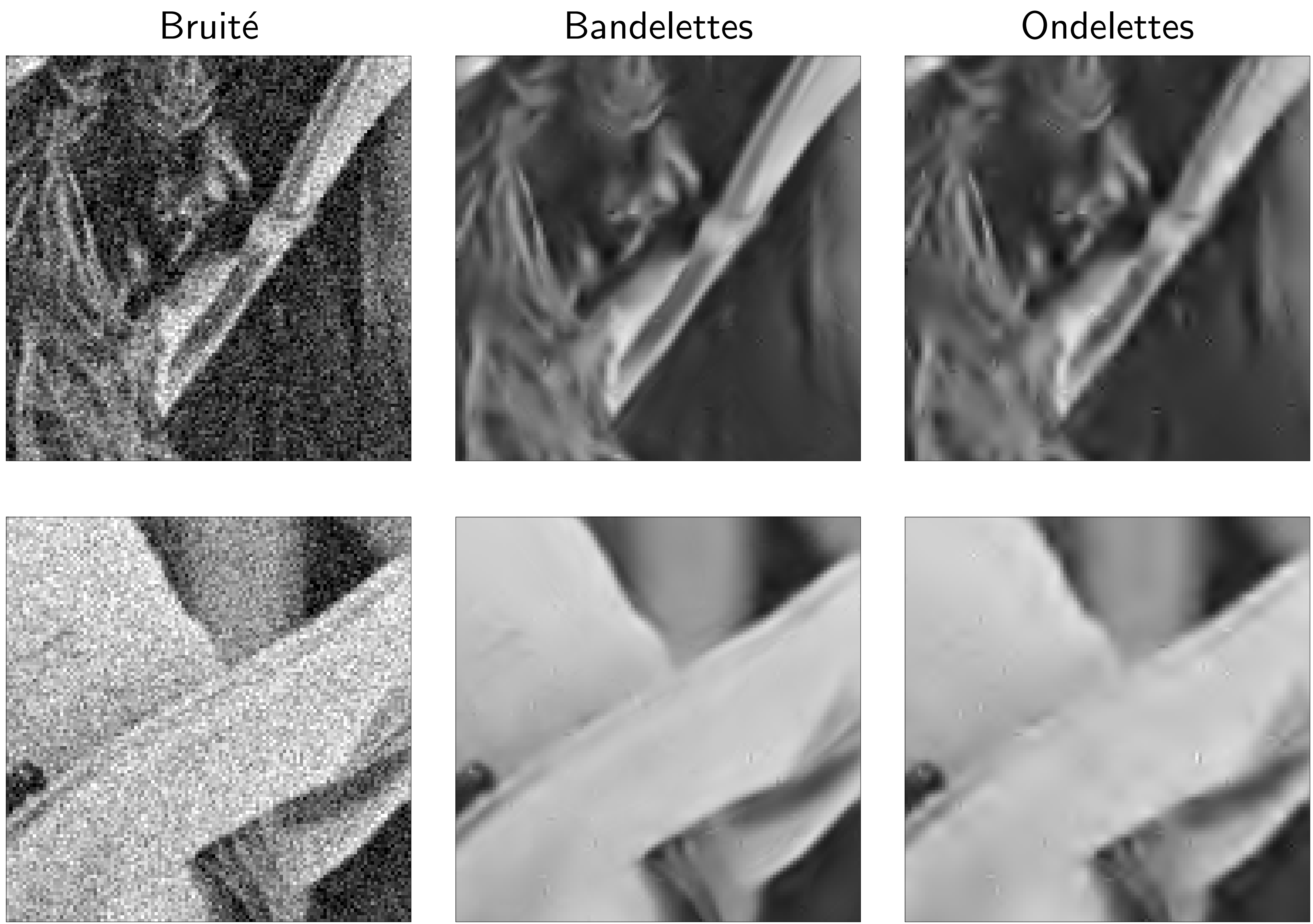


Bruité $(20{,}19\,\mathrm{dB})$

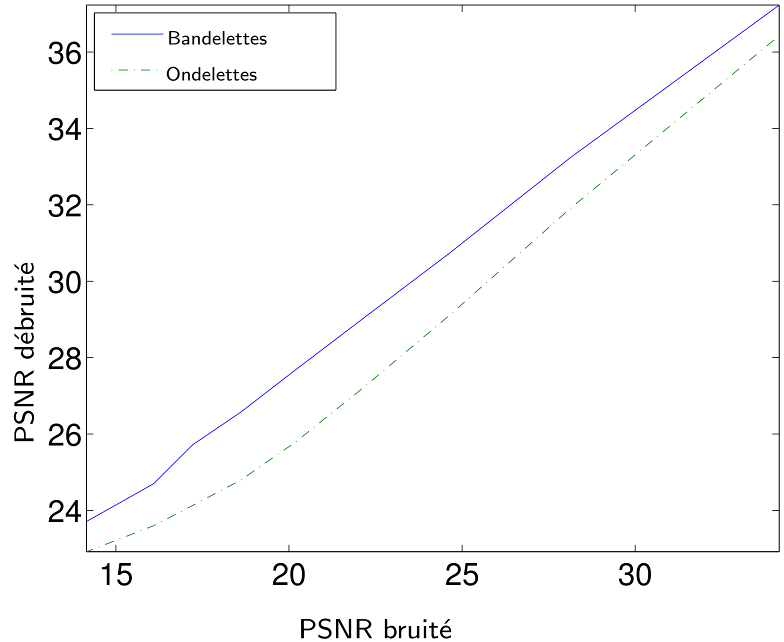


Bandelettes $(30{,}29\,\mathrm{dB})$

Ondelettes $(28{,}21\,\mathrm{dB})$

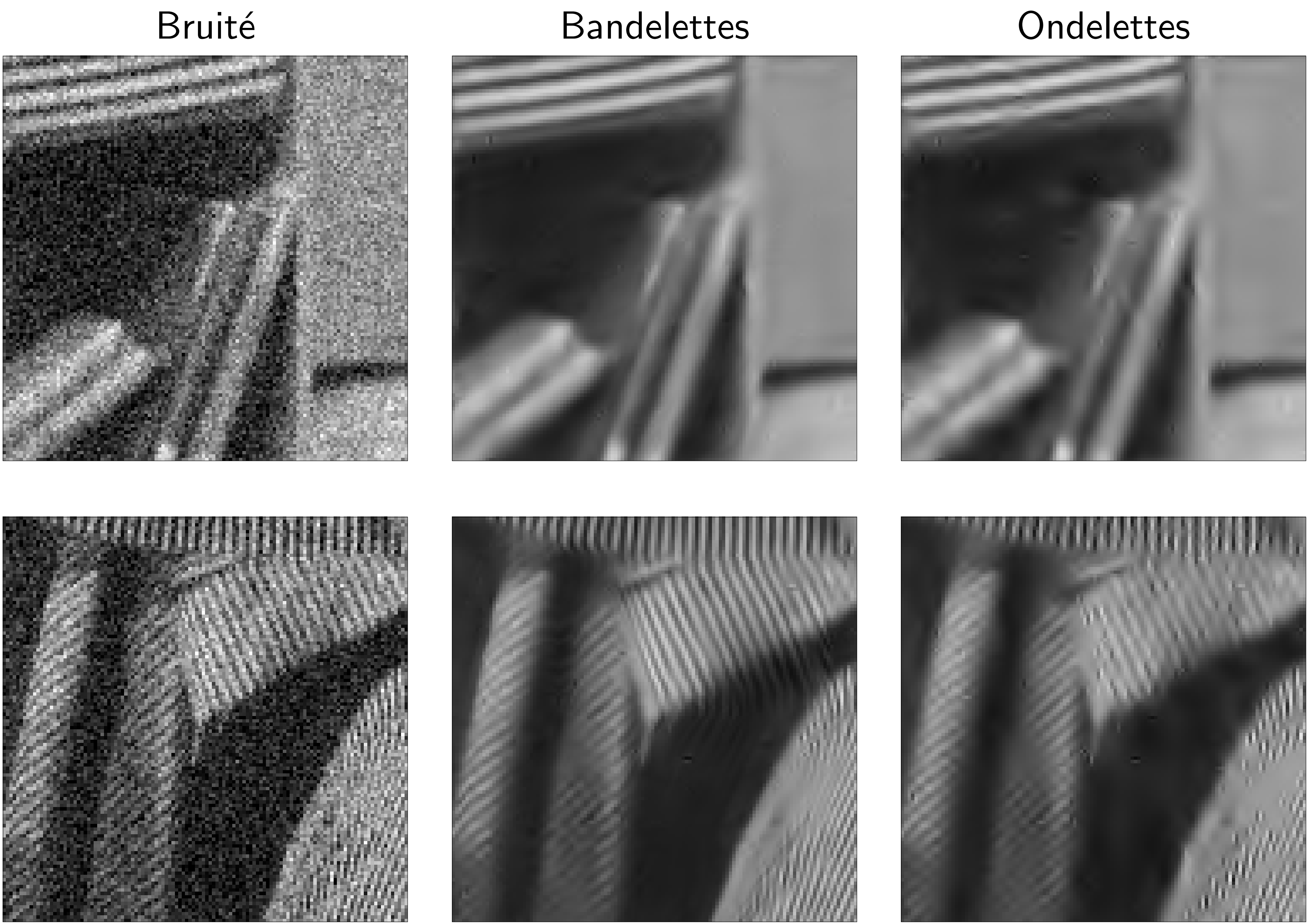

Bruité
Bandelettes
Ondelettes

Bruité $(20{,}19\,\mathrm{dB})$



Bandelettes $(27{,}68\,\mathrm{dB})$

Ondelettes $(25{,}79\,\mathrm{dB})$


Bruité
Bandelettes
Ondelettes

Conclusion

Conclusion

- Estimation à l'aide de la théorie de la sélection de modèle.

Conclusion

- Estimation à l'aide de la théorie de la sélection de modèle.
- Construction d'une représentation adaptée à la nature géométrique des images (bandelettes).

Conclusion

- Estimation à l'aide de la théorie de la sélection de modèle.
- Construction d'une représentation adaptée à la nature géométrique des images (bandelettes).
- Bonnes classes de fonctions ?

Conclusion

- Estimation à l'aide de la théorie de la sélection de modèle.
- Construction d'une représentation adaptée à la nature géométrique des images (bandelettes).
- Bonnes classes de fonctions ?
- Exemple de traitement d'image.

Conclusion

- Estimation à l'aide de la théorie de la sélection de modèle.
- Construction d'une représentation adaptée à la nature géométrique des images (bandelettes).
- Bonnes classes de fonctions ?
- Exemple de traitement d'image.
- Théorie de l'approximation et le principe de concision ont de nombreuses autres applications : compression, apprentissage,. . .

Preuve – 1

Preuve – 1

- $F = \text{argmin}\, \|Y - \tilde{f}\|^2 + \lambda^2 \log \nu \sigma^2 M$ avec M_F coefficients.

Preuve – 1

- $F = \text{argmin}\,\|Y - \tilde{f}\|^2 + \lambda^2 \log \nu \sigma^2 M$ avec M_F coefficients.
- $f_\lambda = \text{argmin}\,\|f - \tilde{f}\|^2 + \lambda^2 \log \nu \sigma^2 M$ avec M_λ coefficients.

Preuve – 1

- $F =$ argmin $\|Y - \tilde{f}\|^2 + \lambda^2 \log \nu\sigma^2 M$ avec M_F coefficients.
- $f_\lambda =$ argmin $\|f - \tilde{f}\|^2 + \lambda^2 \log \nu\sigma^2 M$ avec M_λ coefficients.
- $\|Y - g\|^2 = \|Y - f\|^2 + 2\langle Y - f, f - g\rangle + \|f - g\|^2$.

Preuve – 1

- $F = \text{argmin}\, \|Y - \tilde{f}\|^2 + \lambda^2 \log \nu \sigma^2 M$ avec M_F coefficients.
- $f_\lambda = \text{argmin}\, \|f - \tilde{f}\|^2 + \lambda^2 \log \nu \sigma^2 M$ avec M_λ coefficients.
- $\|Y - g\|^2 = \|Y - f\|^2 + 2\langle Y - f, f - g\rangle + \|f - g\|^2$.
- En insérant ceci dans

$$\|Y - F\|^2 + \lambda^2 \log \nu \sigma^2 M_F \leqslant \|Y - f_\lambda\|^2 + \lambda^2 \log \nu \sigma^2 M_\lambda$$

on obtient

$$\begin{aligned} \|f - F\|^2 + \lambda^2 \log \nu \sigma^2 M_F \leqslant\ & \|f - f_\lambda\|^2 + \lambda^2 \log \nu \sigma^2 M_\lambda \\ & + 2\langle Y - f, f_\lambda - F\rangle \quad . \end{aligned}$$

Preuve – 1

- $F = \text{argmin}\, \|Y - \tilde{f}\|^2 + \lambda^2 \log \nu\sigma^2 M$ avec M_F coefficients.
- $f_\lambda = \text{argmin}\, \|f - \tilde{f}\|^2 + \lambda^2 \log \nu\sigma^2 M$ avec M_λ coefficients.
- $\|Y - g\|^2 = \|Y - f\|^2 + 2\langle Y - f, f - g\rangle + \|f - g\|^2$.
- En insérant ceci dans

$$\|Y - F\|^2 + \lambda^2 \log \nu\sigma^2 M_F \leqslant \|Y - f_\lambda\|^2 + \lambda^2 \log \nu\sigma^2 M_\lambda$$

  on obtient

$$\begin{aligned}\|f - F\|^2 + \lambda^2 \log \nu\sigma^2 M_F \leqslant{} & \|f - f_\lambda\|^2 + \lambda^2 \log \nu\sigma^2 M_\lambda \\ & + 2\langle Y - f, f_\lambda - F\rangle \quad .\end{aligned}$$

- On conclut en prouvant qu'avec grande probabilité,

$$|\langle Y - f, f_\lambda - F\rangle| \leqslant (C/\lambda)(\|f - F\|^2 + \lambda^2 \log \nu\sigma^2 M_F) \quad .$$

  qui implique $C/\lambda < 1$.

Preuve – 2

Preuve – 2

- $|\langle Y - f, f_\lambda - F\rangle| \leqslant \|P_{\mathcal{M}_\lambda \cup \mathcal{M}_F} W\| \, \|f_\lambda - F\| \quad .$

Preuve – 2

- $|\langle Y - f, f_\lambda - F\rangle| \leqslant \|P_{\mathcal{M}_\lambda \cup \mathcal{M}_F} W\| \, \|f_\lambda - F\| \quad .$
- $\|f_\lambda - F\| \leqslant \|f_\lambda - f\| + \|f - F\|$

  $\|f_\lambda - F\| \leqslant 2(\|f - F\|^2 + \lambda^2 \log \nu \sigma^2 M_F)^{1/2} \quad .$

Preuve – 2

- $|\langle Y - f, f_\lambda - F\rangle| \leqslant \|P_{\mathcal{M}_\lambda \cup \mathcal{M}_F} W\| \, \|f_\lambda - F\| \quad .$
- $\|f_\lambda - F\| \leqslant \|f_\lambda - f\| + \|f - F\|$
  $\|f_\lambda - F\| \leqslant 2(\|f - F\|^2 + \lambda^2 \log \nu \sigma^2 M_F)^{1/2} \quad .$
- Inégalité de concentration :

$$P\left(\forall \mathcal{M}, \quad \|P_{\mathcal{M}} W\| \leqslant \sqrt{12 \log \nu \sigma^2 \mathsf{dim}(\mathcal{M})}\right) \geqslant 1 - e/\nu \quad .$$

Preuve – 2

- $|\langle Y - f, f_\lambda - F\rangle| \leqslant \|P_{\mathcal{M}_\lambda \cup \mathcal{M}_F} W\| \, \|f_\lambda - F\| \quad .$
- $\|f_\lambda - F\| \leqslant \|f_\lambda - f\| + \|f - F\|$
  $\|f_\lambda - F\| \leqslant 2(\|f - F\|^2 + \lambda^2 \log \nu\sigma^2 M_F)^{1/2} \quad .$
- Inégalité de concentration :

$$P\left(\forall \mathcal{M}, \quad \|P_{\mathcal{M}} W\| \leqslant \sqrt{12 \log \nu\sigma^2 \mathrm{dim}(\mathcal{M})}\right) \geqslant 1 - e/\nu \quad .$$

- Avec $P \geqslant 1 - e/\nu$,

$$\|P_{\mathcal{M}_\lambda \cup \mathcal{M}_F} W\| \leqslant \sqrt{12 \log \nu\sigma^2 (M_\lambda + M_F)}$$

$$\|P_{\mathcal{M}_\lambda \cup \mathcal{M}_F} W\| \leqslant \sqrt{12/\lambda^2}(\|f - F\|^2 + \lambda^2 \log \nu\sigma^2 M_F)^{1/2} \quad .$$

Preuve – 2

- $|\langle Y-f, f_\lambda - F\rangle| \leqslant \|P_{\mathcal{M}_\lambda \cup \mathcal{M}_F} W\| \, \|f_\lambda - F\| \quad .$
- $\|f_\lambda - F\| \leqslant \|f_\lambda - f\| + \|f - F\|$
  $\|f_\lambda - F\| \leqslant 2(\|f-F\|^2 + \lambda^2 \log \nu \sigma^2 M_F)^{1/2} \quad .$
- Inégalité de concentration :

$$P\left(\forall \mathcal{M}, \quad \|P_{\mathcal{M}} W\| \leqslant \sqrt{12 \log \nu \sigma^2 \mathrm{dim}(\mathcal{M})}\right) \geqslant 1 - e/\nu \quad .$$

- Avec $P \geqslant 1 - e/\nu$,

$$\|P_{\mathcal{M}_\lambda \cup \mathcal{M}_F} W\| \leqslant \sqrt{12 \log \nu \sigma^2 (M_\lambda + M_F)}$$

$$\|P_{\mathcal{M}_\lambda \cup \mathcal{M}_F} W\| \leqslant \sqrt{12/\lambda^2} (\|f-F\|^2 + \lambda^2 \log \nu \sigma^2 M_F)^{1/2} \quad .$$

- $|\langle Y-f, f_\lambda - Fp\rangle| \leqslant \sqrt{48/\lambda^2} (\|f-F\|^2 + \lambda^2 \log \nu \sigma^2 M_F) \quad .$

Preuve – 3

Preuve – 3

- Pour chaque sous-espace $\mathcal{M} = \mathrm{vect}\{b_{\gamma_n}\}$,

$$P\left(\|P_{\mathcal{M}}W\| \geqslant E\|P_{\mathcal{M}}W\| + t\right) \leqslant e^{-t^2/(2\sigma^2)}$$

Preuve – 3

- Pour chaque sous-espace $\mathcal{M} = \text{vect}\{b_{\gamma_n}\}$,

$$P\left(\|P_{\mathcal{M}}W\| \geqslant E\|P_{\mathcal{M}}W\| + t\right) \leqslant e^{-t^2/(2\sigma^2)}$$

- $E\|P_{\mathcal{M}}W\| \leqslant (E(\|P_{\mathcal{M}}W\|^2))^{1/2} \leqslant \sqrt{M\sigma^2}$ avec $M = \dim(\mathcal{M})$.

Preuve – 3

- Pour chaque sous-espace $\mathcal{M} = \text{vect}\{b_{\gamma_n}\}$,

$$P\left(\|P_{\mathcal{M}}W\| \geqslant E\|P_{\mathcal{M}}W\| + t\right) \leqslant e^{-t^2/(2\sigma^2)}$$

- $E\|P_{\mathcal{M}}W\| \leqslant (E(\|P_{\mathcal{M}}W\|^2))^{1/2} \leqslant \sqrt{M\sigma^2}$ avec $M = \dim(\mathcal{M})$.
- $t = \sqrt{6 \log \nu\sigma^2 M}, \quad P\left(\|P_{\mathcal{M}}W\| \geqslant (1+\sqrt{6})\sqrt{\log \nu\sigma^2 M}\right) \leqslant \nu^{-3M}$

Preuve – 3

- Pour chaque sous-espace $\mathcal{M} = \text{vect}\{b_{\gamma_n}\}$,

$$P\left(\|P_{\mathcal{M}}W\| \geqslant E\|P_{\mathcal{M}}W\| + t\right) \leqslant e^{-t^2/(2\sigma^2)}$$

- $E\|P_{\mathcal{M}}W\| \leqslant (E(\|P_{\mathcal{M}}W\|^2))^{1/2} \leqslant \sqrt{M\sigma^2}$ avec $M = \dim(\mathcal{M})$.
- $t = \sqrt{6\log \nu\sigma^2 M}, \quad P\left(\|P_{\mathcal{M}}W\| \geqslant (1+\sqrt{6})\sqrt{\log \nu\sigma^2 M}\right) \leqslant \nu^{-3M}$
- Contrôle sur le nombre de sous-espaces possibles :

$$\begin{aligned} P\Big(\forall \mathcal{M}, \quad & \|P_{\mathcal{M}}W\| \geqslant \sqrt{6\log \nu\sigma^2 M}\Big) \\ & \leqslant \sum_M \sum_{\mathcal{M}, \dim \mathcal{M} = M} P\left(\|P_{\mathcal{M}}W\| \geqslant \sqrt{12\log \nu\sigma^2 M}\right) \\ & \leqslant \sum_M \binom{\nu}{M} \nu^{-3M} \leqslant \sum_M 1/(M!)\nu^{-2M} \leqslant e\nu^{-1} \end{aligned}$$