

Bandelettes : Maxiset et Réalité

E. Le Pennec / LPMA / Université Paris Diderot

S. Mallat, Ch. Dossal et G. Peyré,

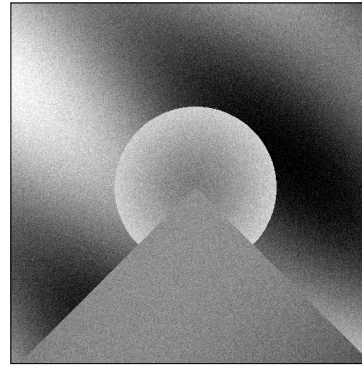
F. Autin, J.-M. Loubes et V. Rivoirard

5 juin 2007

Estimation, Bandelettes et Maxiset

Estimation, Bandelettes et Maxiset

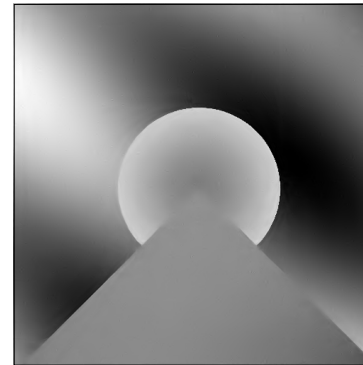
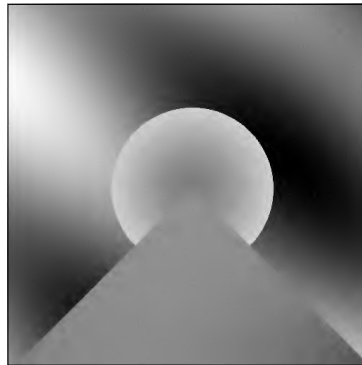
Bruitée



Estimation

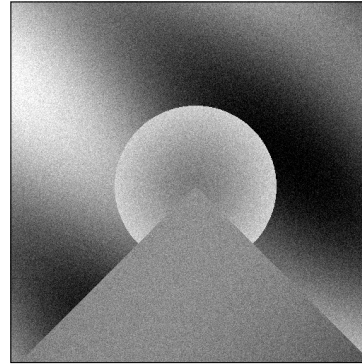
Sans géométrie

Avec géométrie



Estimation, Bandelettes et Maxiset

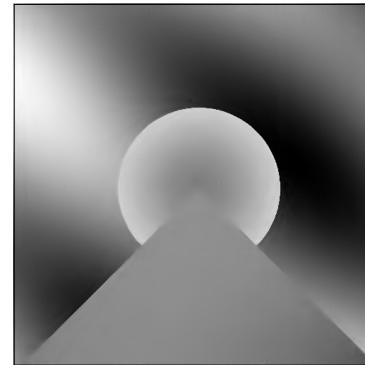
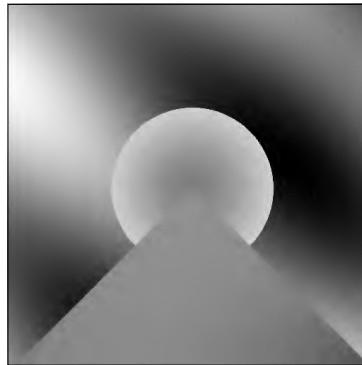
Bruitée



Estimation

Sans géométrie

Avec géométrie

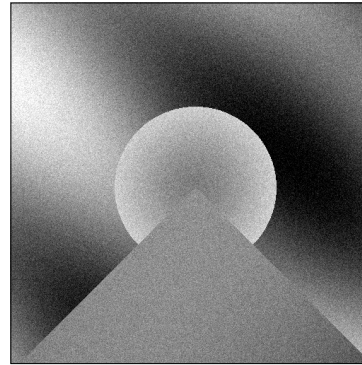


● Estimation dans un modèle de bruit blanc gaussien :

$$Y = f + \epsilon W = f + \frac{1}{\sqrt{N}} W \quad .$$

Estimation, Bandelettes et Maxiset

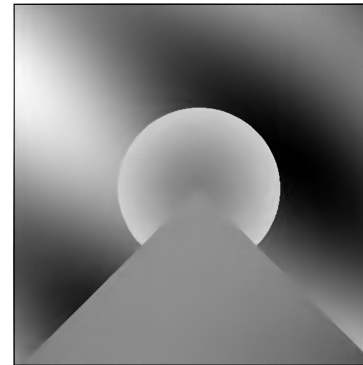
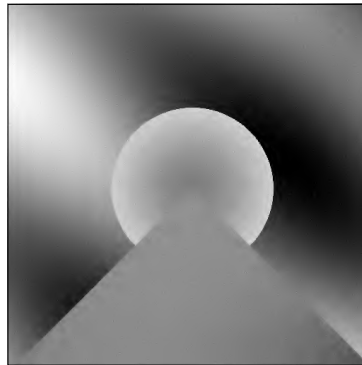
Bruitée



Estimation

Sans géométrie

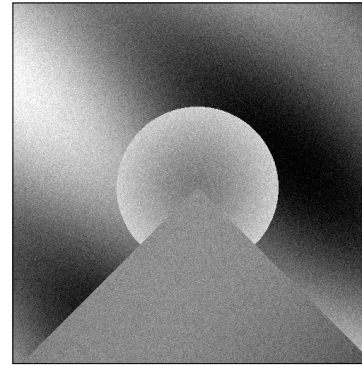
Avec géométrie



- Estimation dans un modèle de bruit blanc gaussien :
$$Y = f + \epsilon W = f + \frac{1}{\sqrt{N}} W \quad .$$
- Estimation de f dans des bases : seuillage et sélection de modèles.

Estimation, Bandelettes et Maxiset

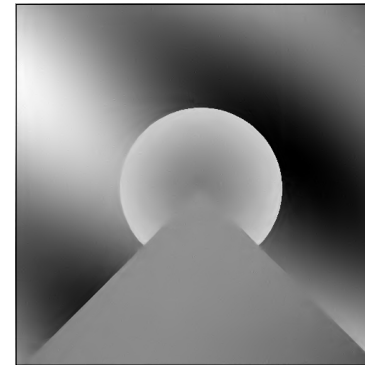
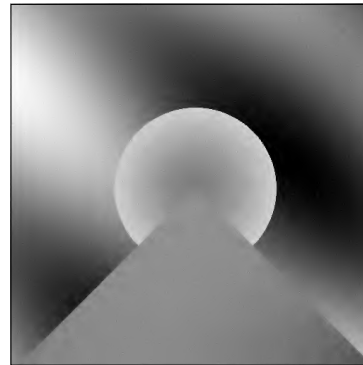
Bruitée



Estimation

Sans géométrie

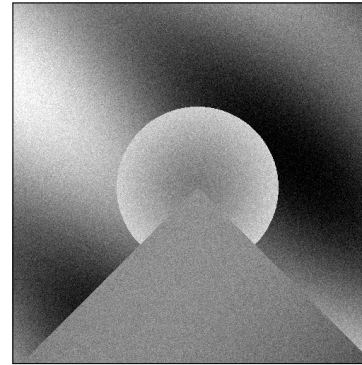
Avec géométrie



- Estimation dans un modèle de bruit blanc gaussien :
$$Y = f + \epsilon W = f + \frac{1}{\sqrt{N}} W \quad .$$
- Estimation de f dans des bases : seuillage et sélection de modèles.
- Maxiset : nécessité d'avoir une représentation creuse (approximation).

Estimation, Bandelettes et Maxiset

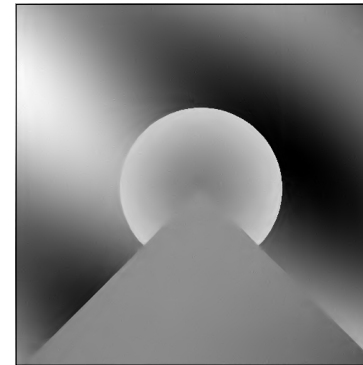
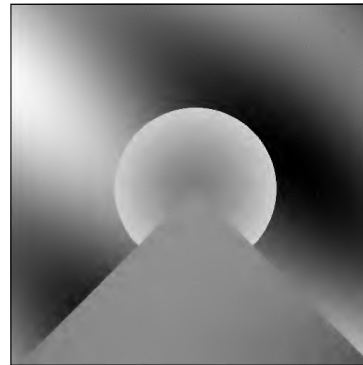
Bruitée



Estimation

Sans géométrie

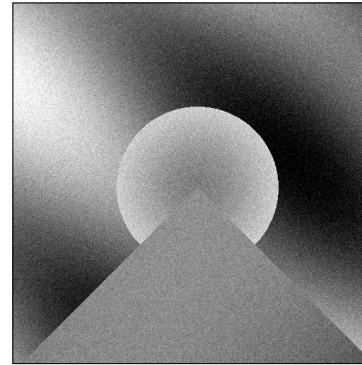
Avec géométrie



- Estimation dans un modèle de bruit blanc gaussien :
$$Y = f + \epsilon W = f + \frac{1}{\sqrt{N}} W \quad .$$
- Estimation de f dans des bases : seuillage et sélection de modèles.
- Maxiset : nécessité d'avoir une représentation creuse (approximation).
- Efficacité provient d'une forme de régularité (géométrie pour les images).

Estimation, Bandelettes et Maxiset

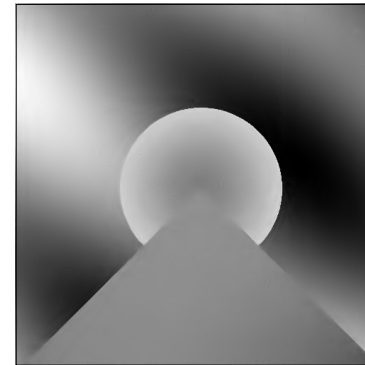
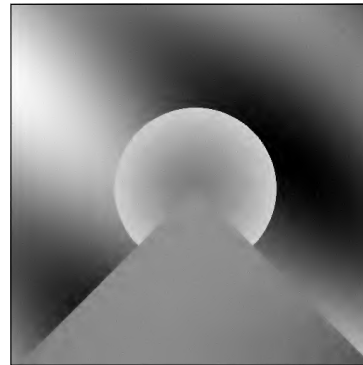
Bruitée



Estimation

Sans géométrie

Avec géométrie



- Estimation dans un modèle de bruit blanc gaussien :
$$Y = f + \epsilon W = f + \frac{1}{\sqrt{N}} W \quad .$$
- Estimation de f dans des bases : seuillage et sélection de modèles.
- Maxiset : nécessité d'avoir une représentation creuse (approximation).
- Efficacité provient d'une forme de régularité (géométrie pour les images).
- Bandelettes pour l'estimation géométrique des images.

Plan

Plan

- Estimation dans une base et approximation / Maxiset.

Plan

- Estimation dans une base et approximation / Maxiset.
- Images 2D et ondelettes.

Plan

- Estimation dans une base et approximation / Maxiset.
- Images 2D et ondelettes.
- Représentation des images géométriques.

Plan

- Estimation dans une base et approximation / Maxiset.
- Images 2D et ondelettes.
- Représentation des images géométriques.
- Bandelettes pour les images géométriques.

Plan

- Estimation dans une base et approximation / Maxiset.
- Images 2D et ondelettes.
- Représentation des images géométriques.
- Bandelettes pour les images géométriques.
- Sélection de modèles et optimalité.

Estimation oracle dans une base

Estimation oracle dans une base

- Décomposition de $Y = f + \frac{1}{\sqrt{N}}W$ dans une base orthonormée

$$Y = \sum_{b_n} \langle Y, b_n \rangle b_n = \sum_{b_n} \left(\langle f, b_n \rangle + \frac{1}{\sqrt{N}} \langle W, b_n \rangle \right) b_n \quad .$$

Estimation oracle dans une base

- Décomposition de $Y = f + \frac{1}{\sqrt{N}}W$ dans une base orthonormée

$$Y = \sum_{b_n} \langle Y, b_n \rangle b_n = \sum_{b_n} \left(\langle f, b_n \rangle + \frac{1}{\sqrt{N}} \langle W, b_n \rangle \right) b_n \quad .$$

- Estimateur F par sélection de coordonnées :

$$F = Y_\Gamma = \sum_{n \in \Gamma} \langle Y, b_n \rangle b_n \quad .$$

Estimation oracle dans une base

- Décomposition de $Y = f + \frac{1}{\sqrt{N}}W$ dans une base orthonormée

$$Y = \sum_{b_n} \langle Y, b_n \rangle b_n = \sum_{b_n} \left(\langle f, b_n \rangle + \frac{1}{\sqrt{N}} \langle W, b_n \rangle \right) b_n \quad .$$

- Estimateur F par sélection de coordonnées :

$$F = Y_\Gamma = \sum_{n \in \Gamma} \langle Y, b_n \rangle b_n \quad .$$

- Minimisation du risque quadratique :

$$E(\|f - F\|^2) = \sum_{n \notin \Gamma} |\langle f, b_n \rangle|^2 + \sum_{n \in \Gamma} \frac{1}{N} \quad .$$

Estimation oracle dans une base

- Décomposition de $Y = f + \frac{1}{\sqrt{N}}W$ dans une base orthonormée

$$Y = \sum_{b_n} \langle Y, b_n \rangle b_n = \sum_{b_n} \left(\langle f, b_n \rangle + \frac{1}{\sqrt{N}} \langle W, b_n \rangle \right) b_n \quad .$$

- Estimateur F par sélection de coordonnées :

$$F = Y_\Gamma = \sum_{n \in \Gamma} \langle Y, b_n \rangle b_n \quad .$$

- Minimisation du risque quadratique :

$$E(\|f - F\|^2) = \sum_{n \notin \Gamma} |\langle f, b_n \rangle|^2 + \sum_{n \in \Gamma} \frac{1}{N} \quad .$$

- Solution : $\Gamma_O = \{n, |\langle f, b_n \rangle| \geq \frac{1}{\sqrt{N}}\}$ et $F_O = Y_{\Gamma_O}$.

Estimation oracle dans une base

- Décomposition de $Y = f + \frac{1}{\sqrt{N}}W$ dans une base orthonormée

$$Y = \sum_{b_n} \langle Y, b_n \rangle b_n = \sum_{b_n} \left(\langle f, b_n \rangle + \frac{1}{\sqrt{N}} \langle W, b_n \rangle \right) b_n \quad .$$

- Estimateur F par sélection de coordonnées :

$$F = Y_\Gamma = \sum_{n \in \Gamma} \langle Y, b_n \rangle b_n \quad .$$

- Minimisation du risque quadratique :

$$E(\|f - F\|^2) = \sum_{n \notin \Gamma} |\langle f, b_n \rangle|^2 + \sum_{n \in \Gamma} \frac{1}{N} \quad .$$

- Solution : $\Gamma_O = \{n, |\langle f, b_n \rangle| \geq \frac{1}{\sqrt{N}}\}$ et $F_O = Y_{\Gamma_O}$.
- Problème : demande la connaissance de f ! (Oracle $\neq$ estimateur)

Oracle, risque et approximation

Oracle, risque et approximation

- Risque quadratique de l'estimateur oracle F_O :

$$E(\|f - F_O\|^2) = \sum_{n \notin \Gamma_O} |\langle f, b_n \rangle|^2 + \sum_{n \in \Gamma_O} \frac{1}{N}$$

$$E(\|f - F_O\|^2) = \|f - f_{\Gamma_O}\|^2 + \frac{1}{N} |\Gamma_O| \quad .$$

Oracle, risque et approximation

- Risque quadratique de l'estimateur oracle F_O :

$$E(\|f - F_O\|^2) = \sum_{n \notin \Gamma_O} |\langle f, b_n \rangle|^2 + \sum_{n \in \Gamma_O} \frac{1}{N}$$

$$E(\|f - F_O\|^2) = \|f - f_{\Gamma_O}\|^2 + \frac{1}{N} |\Gamma_O| \quad .$$

- Compromis entre erreur d'approximation et nombre de termes.

Oracle, risque et approximation

- Risque quadratique de l'estimateur oracle F_O :

$$E(\|f - F_O\|^2) = \sum_{n \notin \Gamma_O} |\langle f, b_n \rangle|^2 + \sum_{n \in \Gamma_O} \frac{1}{N}$$

$$E(\|f - F_O\|^2) = \|f - f_{\Gamma_O}\|^2 + \frac{1}{N} |\Gamma_O| \quad .$$

- Compromis entre erreur d'approximation et nombre de termes.
- Théorie de l'approximation :

$$E(\|f - F_O\|^2) = \|f - f_{\Gamma_O}\|^2 + \frac{1}{N} |\Gamma_O| \leq C \left(\frac{1}{N} \right)^{\frac{\beta}{\beta+1}}$$
$$\Leftrightarrow \|f - f_M\|^2 \leq CM^{-\beta} \Leftrightarrow f \in \mathcal{A}^{\beta} \quad .$$

Oracle, risque et approximation

- Risque quadratique de l'estimateur oracle F_O :

$$E(\|f - F_O\|^2) = \sum_{n \notin \Gamma_O} |\langle f, b_n \rangle|^2 + \sum_{n \in \Gamma_O} \frac{1}{N}$$
$$E(\|f - F_O\|^2) = \|f - f_{\Gamma_O}\|^2 + \frac{1}{N} |\Gamma_O| \quad .$$

- Compromis entre erreur d'approximation et nombre de termes.
- Théorie de l'approximation :

$$E(\|f - F_O\|^2) = \|f - f_{\Gamma_O}\|^2 + \frac{1}{N} |\Gamma_O| \leq C \left(\frac{1}{N} \right)^{\frac{\beta}{\beta+1}}$$
$$\Leftrightarrow \|f - f_M\|^2 \leq CM^{-\beta} \Leftrightarrow f \in \mathcal{A}^{\beta} \quad .$$

- Minimax : pour Θ , classe de fonctions, quelle base donne $\Theta \subset \mathcal{A}^{\beta}$ avec β optimal ? $((\frac{1}{N})^{\frac{\beta}{\beta+1}}$ vitesse minimax).

Oracle, risque et approximation

- Risque quadratique de l'estimateur oracle F_O :

$$E(\|f - F_O\|^2) = \sum_{n \notin \Gamma_O} |\langle f, b_n \rangle|^2 + \sum_{n \in \Gamma_O} \frac{1}{N}$$
$$E(\|f - F_O\|^2) = \|f - f_{\Gamma_O}\|^2 + \frac{1}{N} |\Gamma_O| \quad .$$

- Compromis entre erreur d'approximation et nombre de termes.
- Théorie de l'approximation :

$$E(\|f - F_O\|^2) = \|f - f_{\Gamma_O}\|^2 + \frac{1}{N} |\Gamma_O| \leq C \left(\frac{1}{N} \right)^{\frac{\beta}{\beta+1}}$$
$$\Leftrightarrow \|f - f_M\|^2 \leq CM^{-\beta} \Leftrightarrow f \in \mathcal{A}^{\beta} \quad .$$

- Minimax : pour Θ , classe de fonctions, quelle base donne $\Theta \subset \mathcal{A}^{\beta}$ avec β optimal ? $\left(\left(\frac{1}{N} \right)^{\frac{\beta}{\beta+1}} \right)$ vitesse minimax).
- Maxiset : pour une base fixée, quelle est l'ensemble des fonctions estimées avec une vitesse $\left(\frac{1}{N} \right)^{\frac{\beta}{\beta+1}}$? Ici $\mathcal{A}^{\beta}$.

Estimateur par seuillage

Estimateur par seuillage

- Oracle : $\Gamma_O = \{n, |\langle f, b_n \rangle| \geq \frac{1}{\sqrt{N}}\}$ et $F_O = Y_{\Gamma_O}$.

Estimateur par seuillage

- Oracle : $\Gamma_O = \{n, |\langle f, b_n \rangle| \geq \frac{1}{\sqrt{N}}\}$ et $F_O = Y_{\Gamma_O}$.
- Stratégie : garder les grands coefficients.

Estimateur par seuillage

- Oracle : $\Gamma_O = \{n, |\langle f, b_n \rangle| \geq \frac{1}{\sqrt{N}}\}$ et $F_O = Y_{\Gamma_O}$.
- Stratégie : garder les grands coefficients.
- Seuillage : $\Gamma_S = \{n, |\langle Y, b_n \rangle| \geq T \left(\frac{1}{\sqrt{N}} \right)\}$ et $F_S = Y_{\Gamma_S}$.

Estimateur par seuillage

- Oracle : $\Gamma_O = \{n, |\langle f, b_n \rangle| \geq \frac{1}{\sqrt{N}}\}$ et $F_O = Y_{\Gamma_O}$.
- Stratégie : garder les grands coefficients.
- Seuillage : $\Gamma_S = \{n, |\langle Y, b_n \rangle| \geq T \left(\frac{1}{\sqrt{N}} \right)\}$ et $F_S = Y_{\Gamma_S}$.
- **Théorème (*Donoho, Johnstone*)** : Si $T \left(\frac{1}{\sqrt{N}} \right) = \lambda \sqrt{\frac{\log N}{N}}$, alors

$$E(\|f - F_S\|^2) \leq C(\log N)E(\|f - F_O\|^2)$$

$$E(\|f - F_S\|^2) \leq C \min_{\Gamma} \|f - f_{\Gamma}\|^2 + \lambda^2 \frac{\log N}{N} |\Gamma| \quad \text{plus fin.}$$

Estimateur par seuillage

- Oracle : $\Gamma_O = \{n, |\langle f, b_n \rangle| \geq \frac{1}{\sqrt{N}}\}$ et $F_O = Y_{\Gamma_O}$.
- Stratégie : garder les grands coefficients.
- Seuillage : $\Gamma_S = \{n, |\langle Y, b_n \rangle| \geq T \left(\frac{1}{\sqrt{N}} \right)\}$ et $F_S = Y_{\Gamma_S}$.
- **Théorème (Donoho, Johnstone)** : Si $T \left(\frac{1}{\sqrt{N}} \right) = \lambda \sqrt{\frac{\log N}{N}}$, alors

$$E(\|f - F_S\|^2) \leq C(\log N)E(\|f - F_O\|^2)$$

$$E(\|f - F_S\|^2) \leq C \min_{\Gamma} \|f - f_{\Gamma}\|^2 + \lambda^2 \frac{\log N}{N} |\Gamma| \quad \text{plus fin.}$$

- **Théorème (Maxiset) (Cohen, DeVore, Kerkycharian, Picard)** :

$$E(\|f - F_S\|^2) \leq C \left(\frac{\log N}{N} \right)^{\frac{\beta}{\beta+1}} \Leftrightarrow f \in V_{\frac{2\beta}{\beta+1}}^*$$

$$\Leftrightarrow \min_{\Gamma} \|f - f_{\Gamma}\|^2 + \lambda^2 T^2 |\Gamma| \leq C(T^2)^{\frac{\beta}{\beta+1}}$$

$$\Leftrightarrow \|f - f_M\|^2 \leq CM^{-\beta} \Leftrightarrow f \in \mathcal{A}^{\beta}.$$

Estimateur par seuillage

- Oracle : $\Gamma_O = \{n, |\langle f, b_n \rangle| \geq \frac{1}{\sqrt{N}}\}$ et $F_O = Y_{\Gamma_O}$.
- Stratégie : garder les grands coefficients.
- Seuillage : $\Gamma_S = \{n, |\langle Y, b_n \rangle| \geq T \left(\frac{1}{\sqrt{N}} \right)\}$ et $F_S = Y_{\Gamma_S}$.
- **Théorème (Donoho, Johnstone)** : Si $T \left(\frac{1}{\sqrt{N}} \right) = \lambda \sqrt{\frac{\log N}{N}}$, alors

$$E(\|f - F_S\|^2) \leq C(\log N)E(\|f - F_O\|^2)$$

$$E(\|f - F_S\|^2) \leq C \min_{\Gamma} \|f - f_{\Gamma}\|^2 + \lambda^2 \frac{\log N}{N} |\Gamma| \quad \text{plus fin.}$$

- **Théorème (Maxiset) (Cohen, DeVore, Kerkycharian, Picard)** :

$$\begin{aligned} E(\|f - F_S\|^2) &\leq C \left(\frac{\log N}{N} \right)^{\frac{\beta}{\beta+1}} \Leftrightarrow f \in V_{\frac{2\beta}{\beta+1}}^* \\ &\Leftrightarrow \min_{\Gamma} \|f - f_{\Gamma}\|^2 + \lambda^2 T^2 |\Gamma| \leq C(T^2)^{\frac{\beta}{\beta+1}} \\ &\Leftrightarrow \|f - f_M\|^2 \leq CM^{-\beta} \Leftrightarrow f \in \mathcal{A}^{\beta} . \end{aligned}$$

- Importance du choix de la base : en 1D, ondelettes mieux que Fourier.

Seuillage en ondelettes 2D

Seuillage en ondelettes 2D

● Maxiset associé à la vitesse $\left(\frac{\log N}{N}\right)^{\frac{\beta}{\beta+1}}$:

$$\mathcal{A}^{\beta} = \mathcal{W}B_{2/(2\beta+1), 2/(2\beta+1)}^{\beta} = \text{version faible de } B_{2/(2\beta+1), 2/(2\beta+1)}^{\beta}.$$

Seuillage en ondelettes 2D

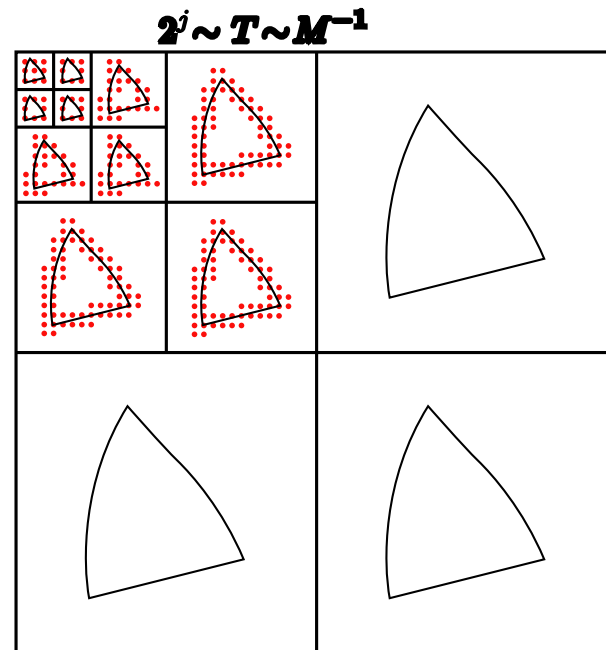
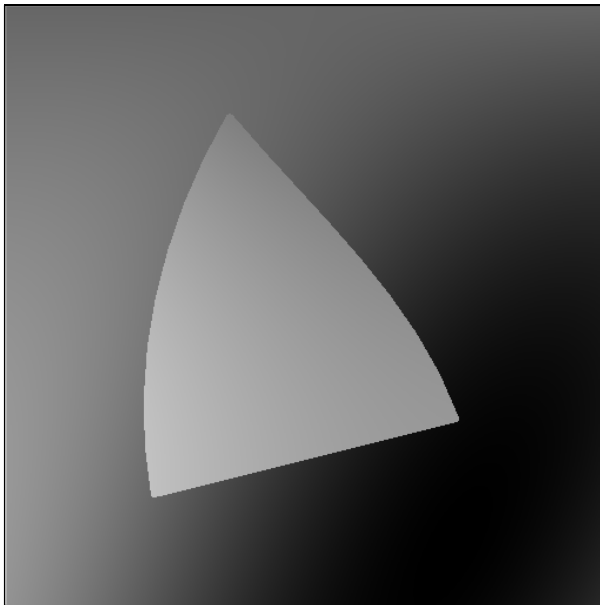
- Maxiset associé à la vitesse $\left(\frac{\log N}{N}\right)^{\frac{\beta}{\beta+1}}$:
- $\mathcal{A}^{\beta} = \mathcal{W}B_{2/(2\beta+1), 2/(2\beta+1)}^{\beta} = \text{version faible de } B_{2/(2\beta+1), 2/(2\beta+1)}^{\beta}$.
- Pour $f \in \mathbf{C}^{\alpha} - \mathbf{C}^{\alpha}$ ($\mathbf{C}^{\alpha}$ en dehors de contours $\mathbf{C}^{\alpha}$) (Korostelev, Tsybakov) : vitesse minimax $\left(\frac{1}{N}\right)^{\frac{\alpha}{\alpha+1}}$ ($\beta = \alpha$).

Seuillage en ondelettes 2D

- Maxiset associé à la vitesse $\left(\frac{\log N}{N}\right)^{\frac{\beta}{\beta+1}}$:
- $\mathcal{A}^{\beta} = \mathcal{W}B_{2/(2\beta+1), 2/(2\beta+1)}^{\beta}$ = version faible de $B_{2/(2\beta+1), 2/(2\beta+1)}^{\beta}$.
- Pour $f \in \mathbf{C}^{\alpha} - \mathbf{C}^{\alpha}$ ($\mathbf{C}^{\alpha}$ en dehors de contours $\mathbf{C}^{\alpha}$) (Korostelev, Tsybakov) : vitesse minimax $\left(\frac{1}{N}\right)^{\frac{\alpha}{\alpha+1}}$ ($\beta = \alpha$).
- Impossible pour les ondelettes : $\mathbf{C}^{\alpha} - \mathbf{C}^{\alpha} \not\subset \mathcal{W}B_{2/(2\alpha+1), 2/(2\alpha+1)}^{\alpha}$!

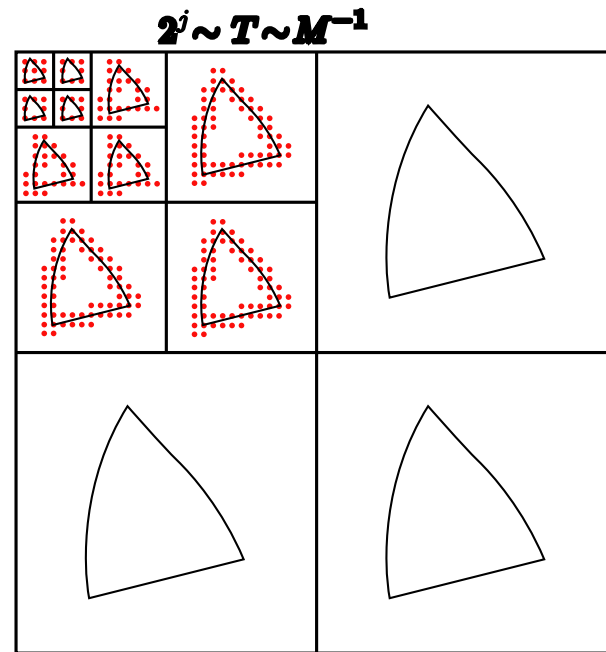
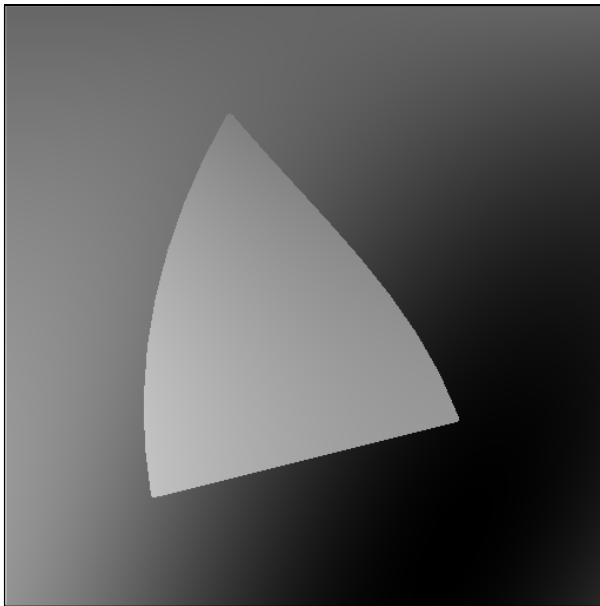
Seuillage en ondelettes 2D

- Maxiset associé à la vitesse $\left(\frac{\log N}{N}\right)^{\frac{\beta}{\beta+1}}$:
- $\mathcal{A}^\beta = \mathcal{W}B_{2/(2\beta+1), 2/(2\beta+1)}^\beta =$ version faible de $B_{2/(2\beta+1), 2/(2\beta+1)}^\beta$.
- Pour $f \in \mathbf{C}^\alpha - \mathbf{C}^\alpha$ ($\mathbf{C}^\alpha$ en dehors de contours $\mathbf{C}^\alpha$) (Korostelev, Tsybakov) : vitesse minimax $\left(\frac{1}{N}\right)^{\frac{\alpha}{\alpha+1}}$ ($\beta = \alpha$).
- Impossible pour les ondelettes : $\mathbf{C}^\alpha - \mathbf{C}^\alpha \not\subset \mathcal{W}B_{2/(2\alpha+1), 2/(2\alpha+1)}^\alpha$!



Seuillage en ondelettes 2D

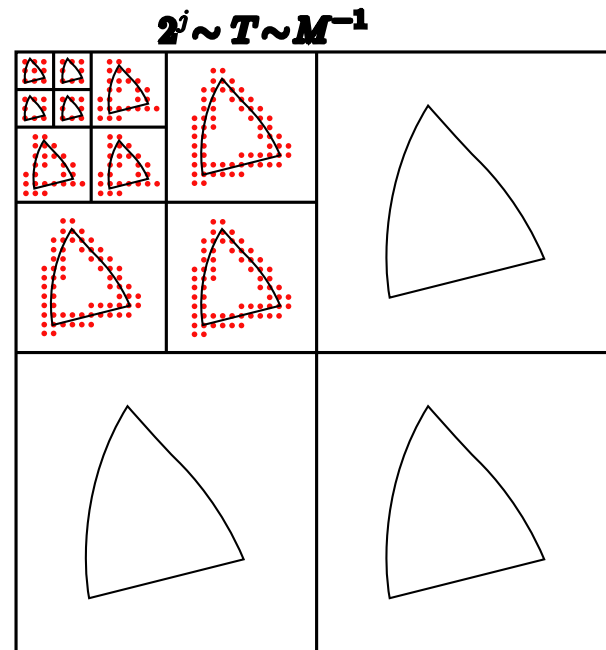
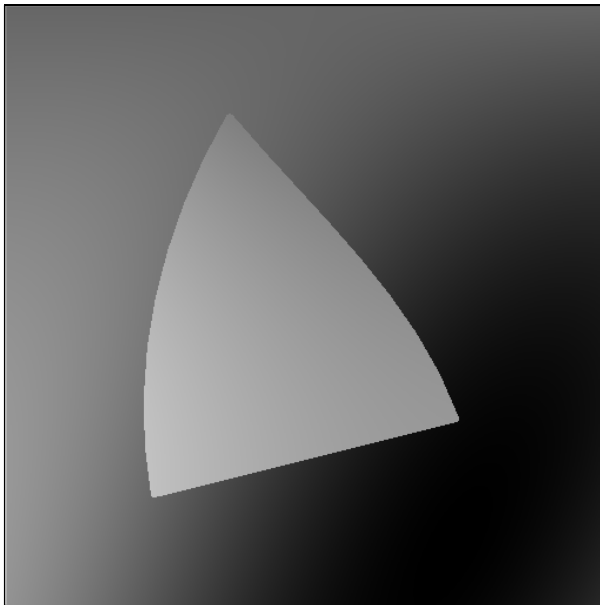
- Maxiset associé à la vitesse $\left(\frac{\log N}{N}\right)^{\frac{\beta}{\beta+1}}$:
- $\mathcal{A}^\beta = \mathcal{W}B_{2/(2\beta+1), 2/(2\beta+1)}^\beta =$ version faible de $B_{2/(2\beta+1), 2/(2\beta+1)}^\beta$.
- Pour $f \in \mathbf{C}^\alpha - \mathbf{C}^\alpha$ ($\mathbf{C}^\alpha$ en dehors de contours $\mathbf{C}^\alpha$) (Korostelev, Tsybakov) : vitesse minimax $\left(\frac{1}{N}\right)^{\frac{\alpha}{\alpha+1}}$ ($\beta = \alpha$).
- Impossible pour les ondelettes : $\mathbf{C}^\alpha - \mathbf{C}^\alpha \not\subset \mathcal{W}B_{2/(2\alpha+1), 2/(2\alpha+1)}^\alpha$!



- Pb d'approximation connu : avec M ondelettes $\|f - f_M\|^2 \leq C M^{-1}$.

Seuillage en ondelettes 2D

- Maxiset associé à la vitesse $\left(\frac{\log N}{N}\right)^{\frac{\beta}{\beta+1}}$:
- $\mathcal{A}^\beta = \mathcal{W}B_{2/(2\beta+1), 2/(2\beta+1)}^\beta = \text{version faible de } B_{2/(2\beta+1), 2/(2\beta+1)}^\beta$.
- Pour $f \in \mathbf{C}^\alpha - \mathbf{C}^\alpha$ ($\mathbf{C}^\alpha$ en dehors de contours $\mathbf{C}^\alpha$) (Korostelev, Tsybakov) : vitesse minimax $\left(\frac{1}{N}\right)^{\frac{\alpha}{\alpha+1}}$ ($\beta = \alpha$).
- Impossible pour les ondelettes : $\mathbf{C}^\alpha - \mathbf{C}^\alpha \not\subset \mathcal{W}B_{2/(2\alpha+1), 2/(2\alpha+1)}^\alpha$!

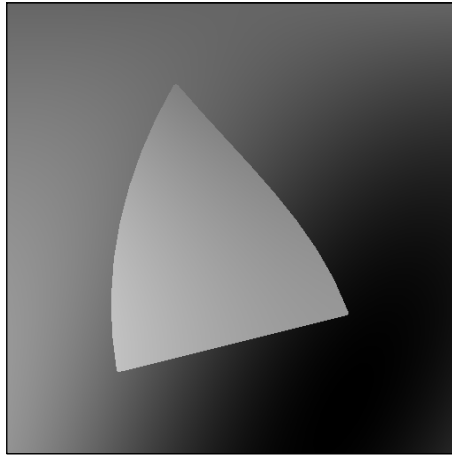


- Pb d'approximation connu : avec M ondelettes $\|f - f_M\|^2 \leq C M^{-1}$.
- Besoin de $\|f - f_M\|^2 \leq C M^{-\alpha}$ pour le risque minimax.

Éléments géométriques pour les contours

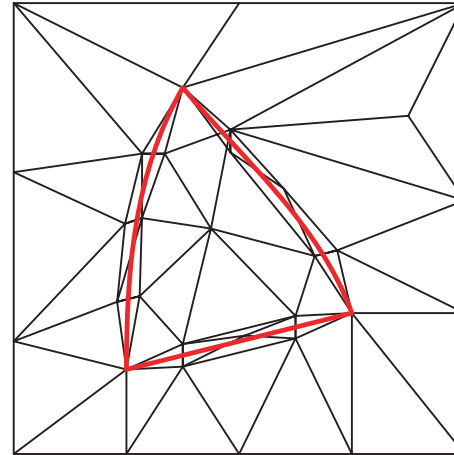
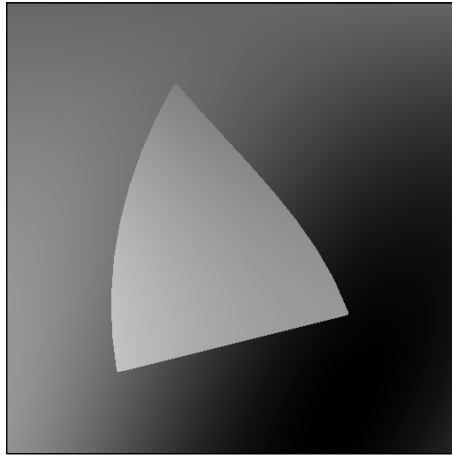
Éléments géométriques pour les contours

- Approximation de f qui est C^α en dehors de contours C^α :

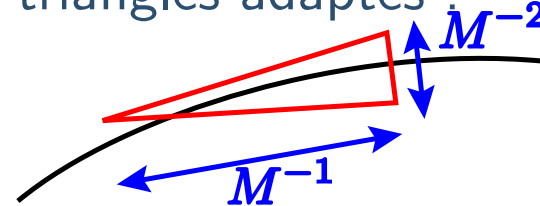


Éléments géométriques pour les contours

- Approximation de f qui est C^α en dehors de contours C^α :

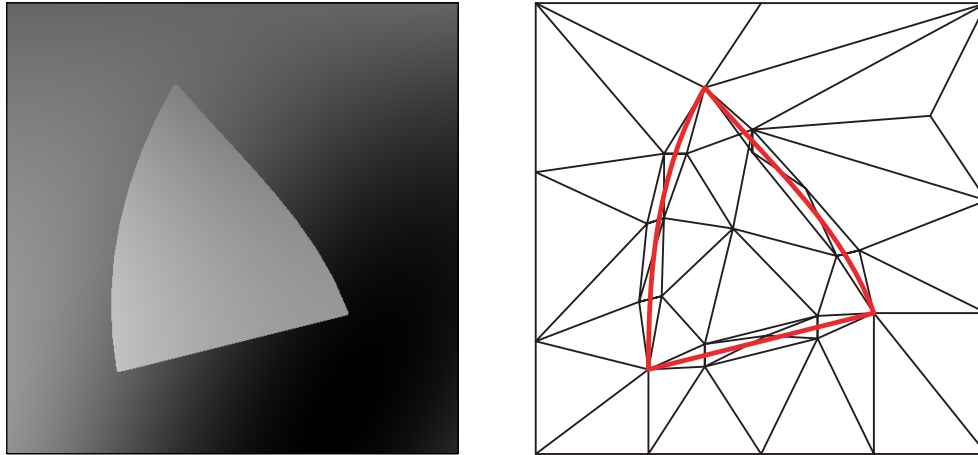


- Approximation linéaire par morceau sur M triangles adaptés :
si $\alpha \geq 2$ alors $\|f - f_M\|^2 \leq C M^{-2}$.

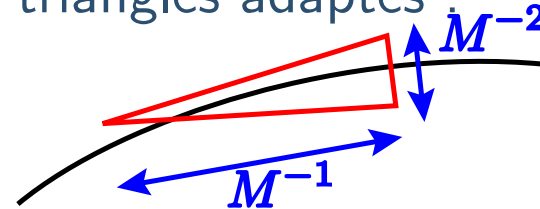


Éléments géométriques pour les contours

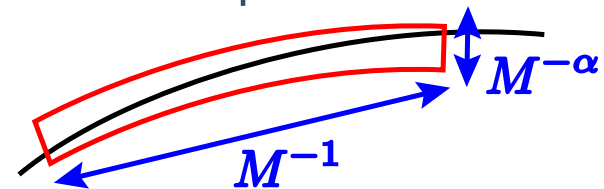
- Approximation de f qui est C^α en dehors de contours C^α :



- Approximation linéaire par morceau sur M triangles adaptés :
si $\alpha \geq 2$ alors $\|f - f_M\|^2 \leq C M^{-2}$.

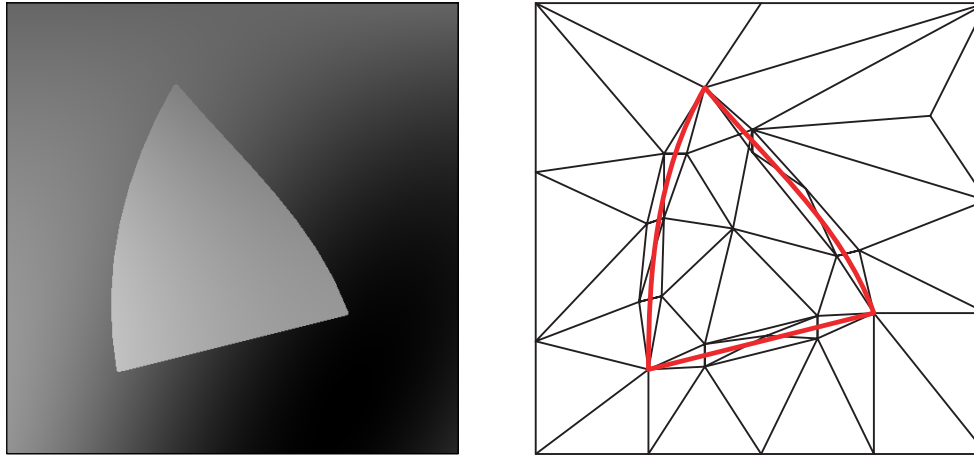


- Approximation d'ordre élevé avec M “éléments” adaptés :
 $\|f - f_M\|^2 \leq C M^{-\alpha}$.

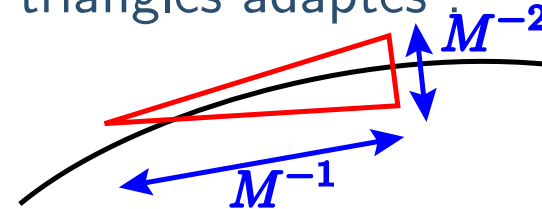


Éléments géométriques pour les contours

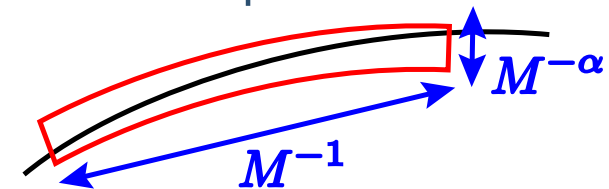
- Approximation de f qui est C^α en dehors de contours C^α :



- Approximation linéaire par morceau sur M triangles adaptés :
si $\alpha \geq 2$ alors $\|f - f_M\|^2 \leq C M^{-2}$.



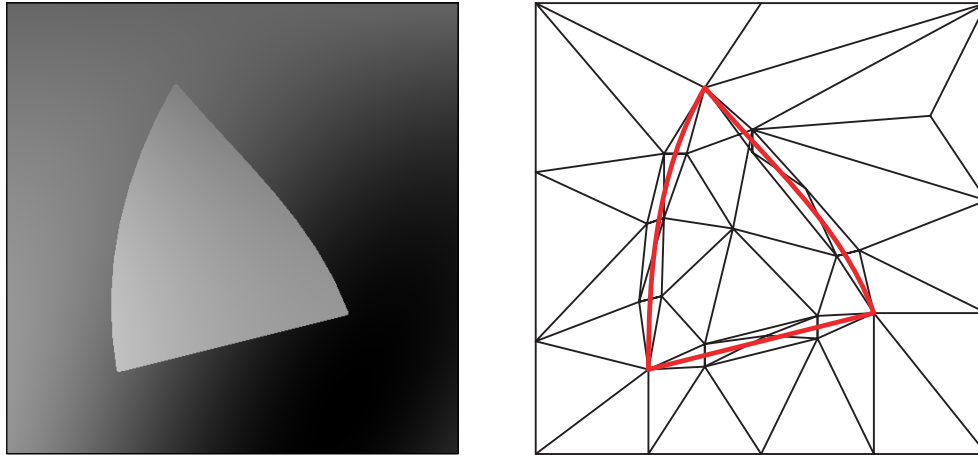
- Approximation d'ordre élevé avec M "éléments" adaptés :
 $\|f - f_M\|^2 \leq C M^{-\alpha}$.



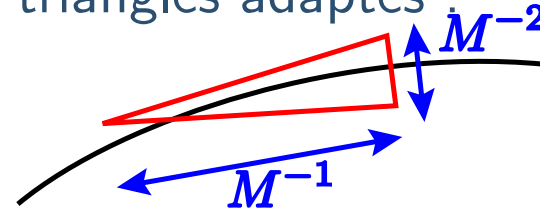
- Pas de bases et optimisation difficile.

Éléments géométriques pour les contours

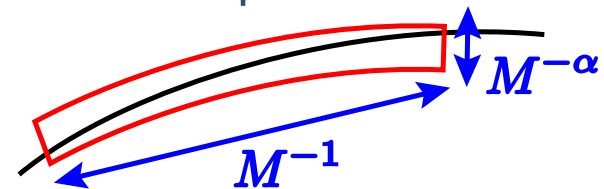
- Approximation de f qui est C^α en dehors de contours C^α :



- Approximation linéaire par morceau sur M triangles adaptés :
si $\alpha \geq 2$ alors $\|f - f_M\|^2 \leq C M^{-2}$.



- Approximation d'ordre élevé avec M “éléments” adaptés :
 $\|f - f_M\|^2 \leq C M^{-\alpha}$.



- Pas de bases et optimisation difficile.
- Cas $\alpha = 2$: tight frame de curvelets parfaitement adapté.

Bandelettes

Bandelettes

Bandelettes

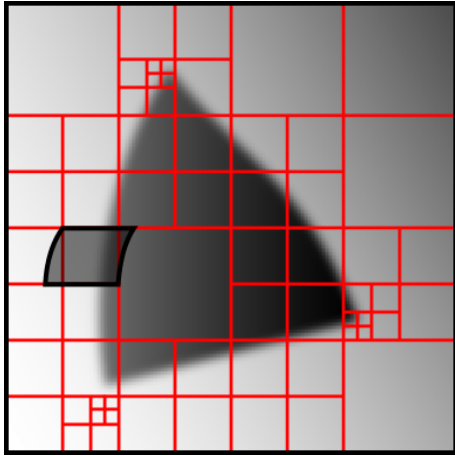
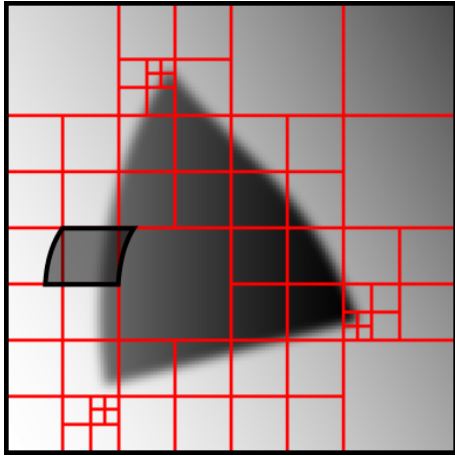


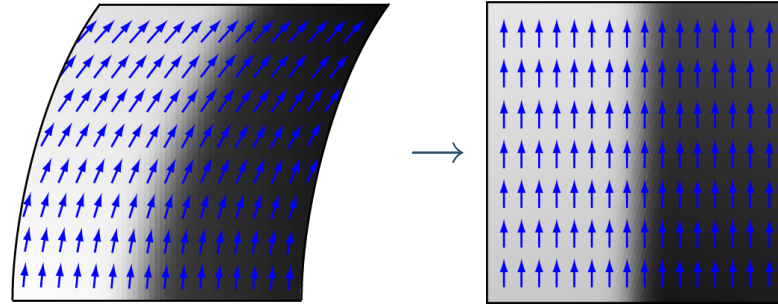
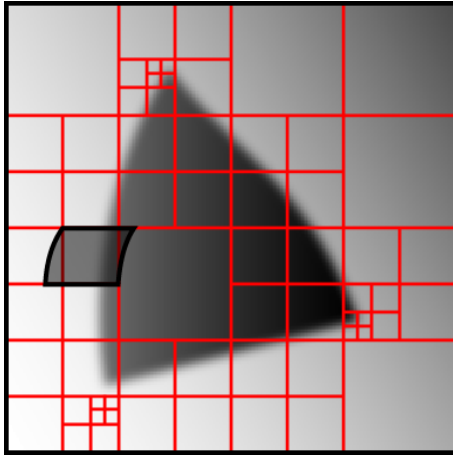
Image C^α – C^α simple

Bandelettes



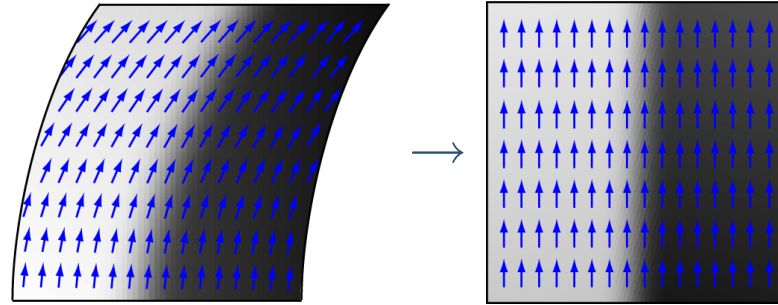
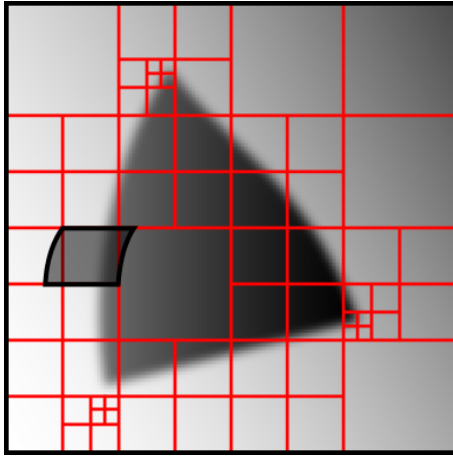
● Image $C^\alpha - C^\alpha$ simple par morceaux.

Bandelettes



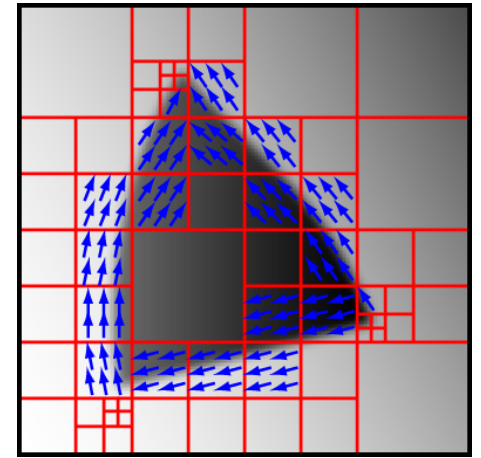
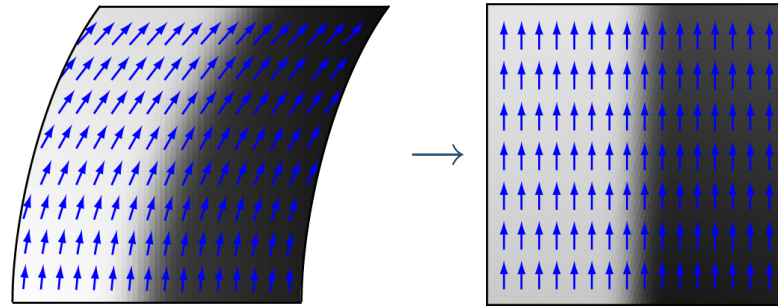
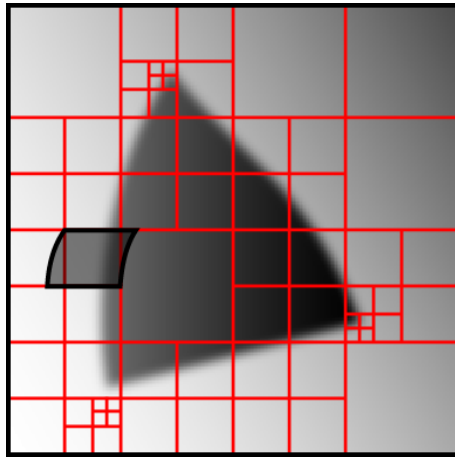
- Image $C^\alpha - C^\alpha$ simple par morceaux.
- Déformation locale $\implies$ singularité verticale/horizontale .

Bandelettes



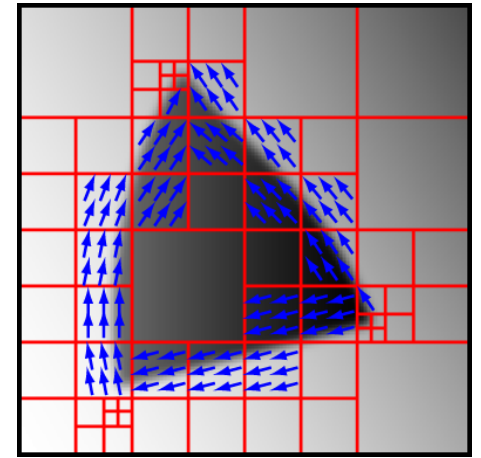
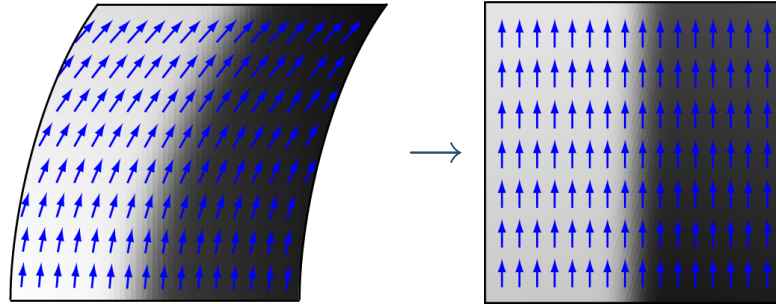
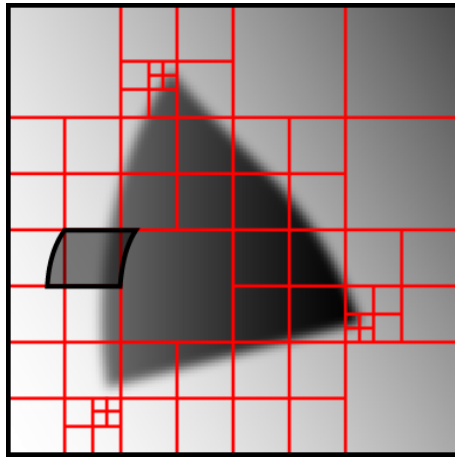
- Image $C^\alpha - C^\alpha$ simple par morceaux.
- Déformation locale $\implies$ singularité verticale/horizontale .
- Bandelettes locales : préimage d'une base adaptée.

Bandelettes



- Image $C^\alpha - C^\alpha$ simple par morceaux.
- Déformation locale $\implies$ singularité verticale/horizontale .
- Bandelettes locales : préimage d'une base adaptée.
- Base de bandelettes : segmentation dyadique + une géométrie par carré.

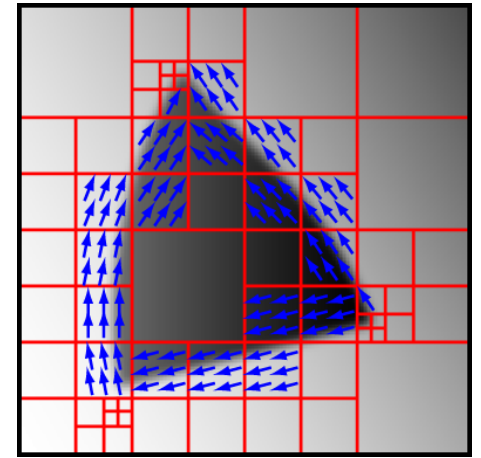
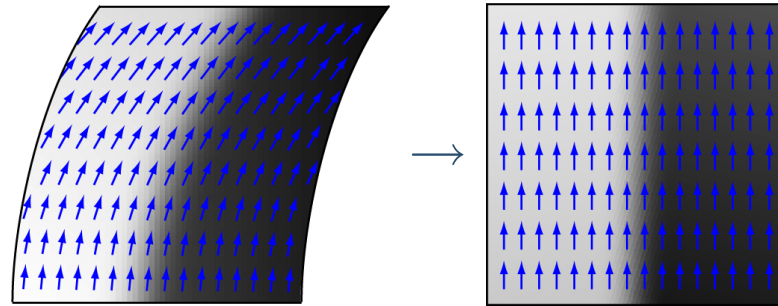
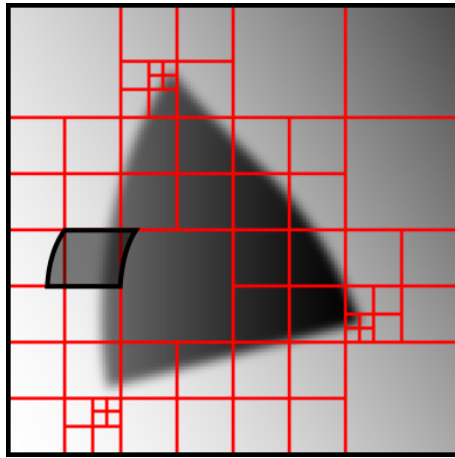
Bandelettes



- Image $C^\alpha - C^\alpha$ simple par morceaux.
- Déformation locale $\implies$ singularité verticale/horizontale .
- Bandelettes locales : préimage d'une base adaptée.
- Base de bandelettes : segmentation dyadique + une géométrie par carré.
- **Théorème** : Si f est $C^\alpha - C^\alpha$, alors, dans la meilleure base,

$$\|f - f_M\|^2 \leq C(\log M)M^{-\alpha} .$$

Bandelettes

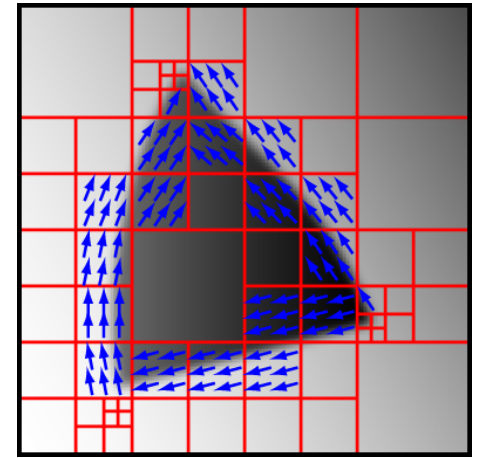
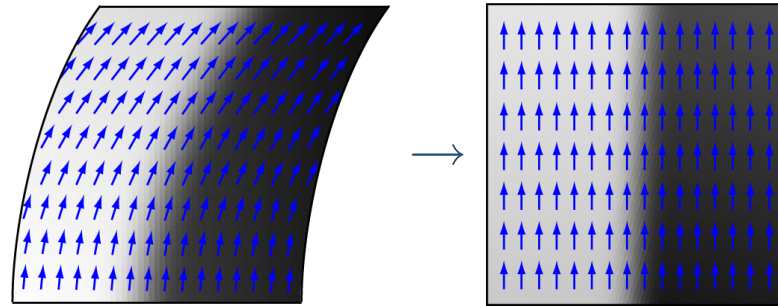
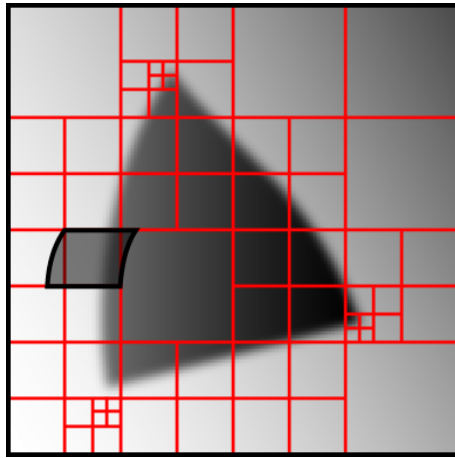


- Image $C^\alpha - C^\alpha$ simple par morceaux.
- Déformation locale $\implies$ singularité verticale/horizontale .
- Bandelettes locales : préimage d'une base adaptée.
- Base de bandelettes : segmentation dyadique + une géométrie par carré.
- **Théorème** : Si f est $C^\alpha - C^\alpha$, alors, dans la meilleure base,

$$\|f - f_M\|^2 \leq C(\log M)M^{-\alpha} .$$

- Minimisation du Lagrangien $\|f - f_M\|^2 + T^2 M$: seuillage à base fixée (facile) et recherche de meilleure base (difficile).

Bandelettes

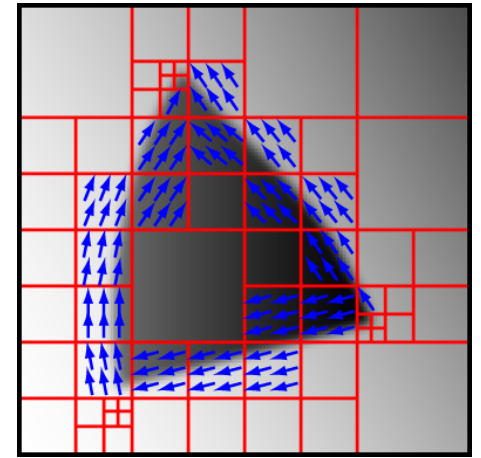
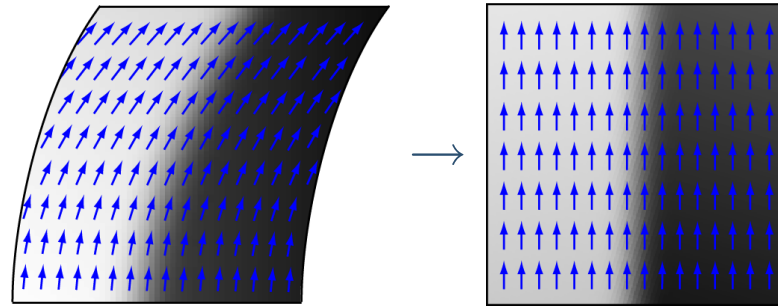
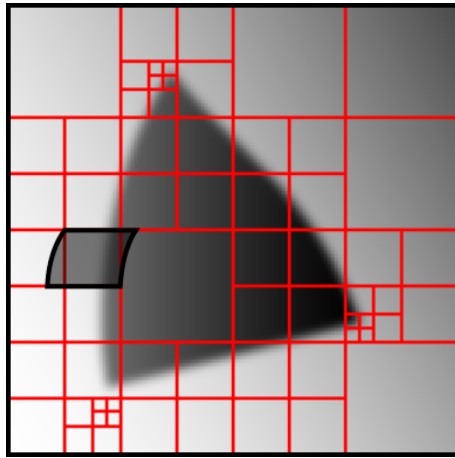


- Image $C^\alpha - C^\alpha$ simple par morceaux.
- Déformation locale $\implies$ singularité verticale/horizontale .
- Bandelettes locales : préimage d'une base adaptée.
- Base de bandelettes : segmentation dyadique + une géométrie par carré.
- **Théorème** : Si f est $C^\alpha - C^\alpha$, alors, dans la meilleure base,

$$\|f - f_M\|^2 \leq C(\log M)M^{-\alpha} .$$

- Minimisation du Lagrangien $\|f - f_M\|^2 + T^2 M$: seuillage à base fixée (facile) et recherche de meilleure base (difficile).
- Structure hiérarchique de la partition et additivité de la fonction à minimiser : algorithme de meilleure base de Wickerhauser (CART).

Bandelettes



- Image $C^\alpha - C^\alpha$ simple par morceaux.
- Déformation locale $\implies$ singularité verticale/horizontale .
- Bandelettes locales : préimage d'une base adaptée.
- Base de bandelettes : segmentation dyadique + une géométrie par carré.
- **Théorème** : Si f est $C^\alpha - C^\alpha$, alors, dans la meilleure base,

$$\|f - f_M\|^2 \leq C(\log M)M^{-\alpha} .$$

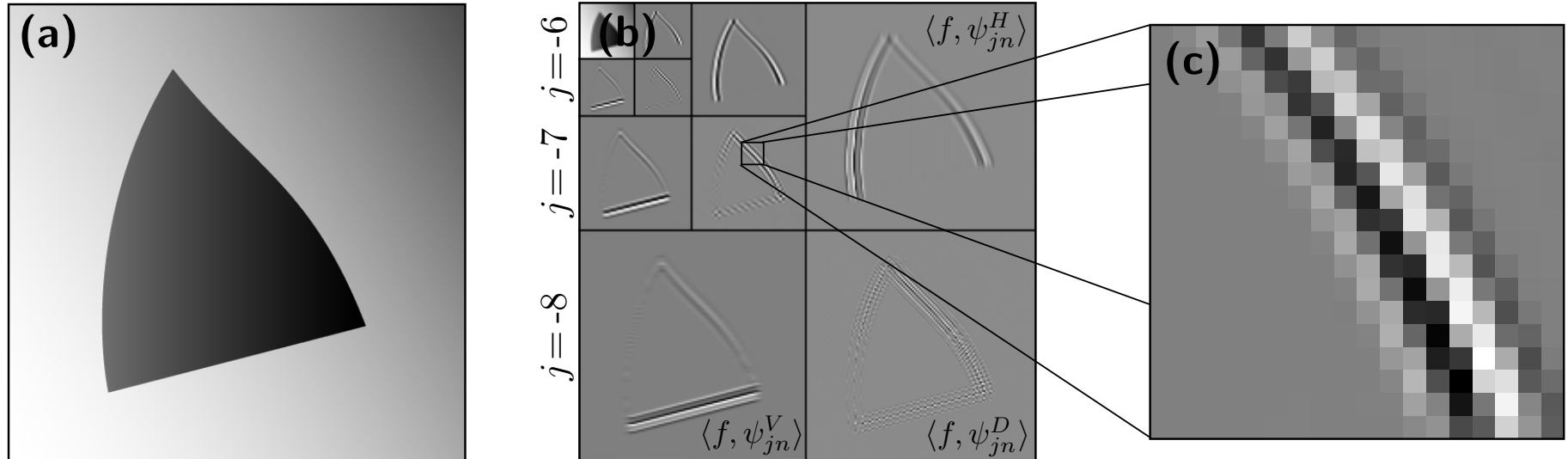
- Minimisation du Lagrangien $\|f - f_M\|^2 + T^2 M$: seuillage à base fixée (facile) et recherche de meilleure base (difficile).
- Structure hiérarchique de la partition et additivité de la fonction à minimiser : algorithme de meilleure base de Wickerhauser (CART).
- Exploration exhaustive des géométries dans chaque carré.

Bandelettes et géométrie multiéchelle

Bandelettes et géométrie multiéchelle

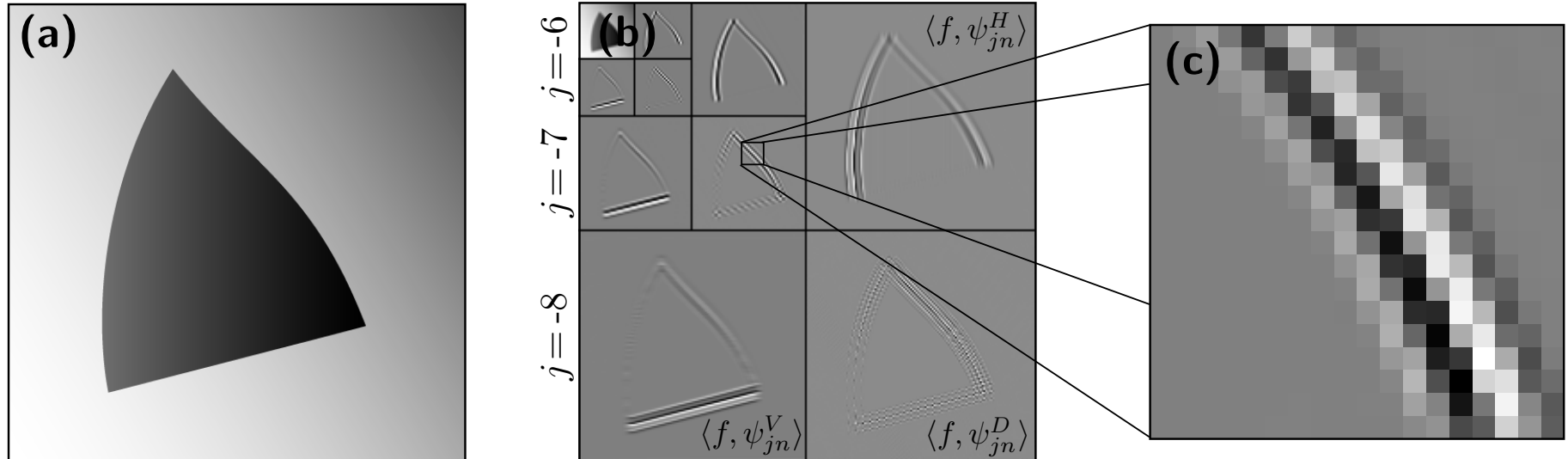
- Bandelettes : représentation à géométrie monoéchelle.

Bandelettes et géométrie multiéchelle



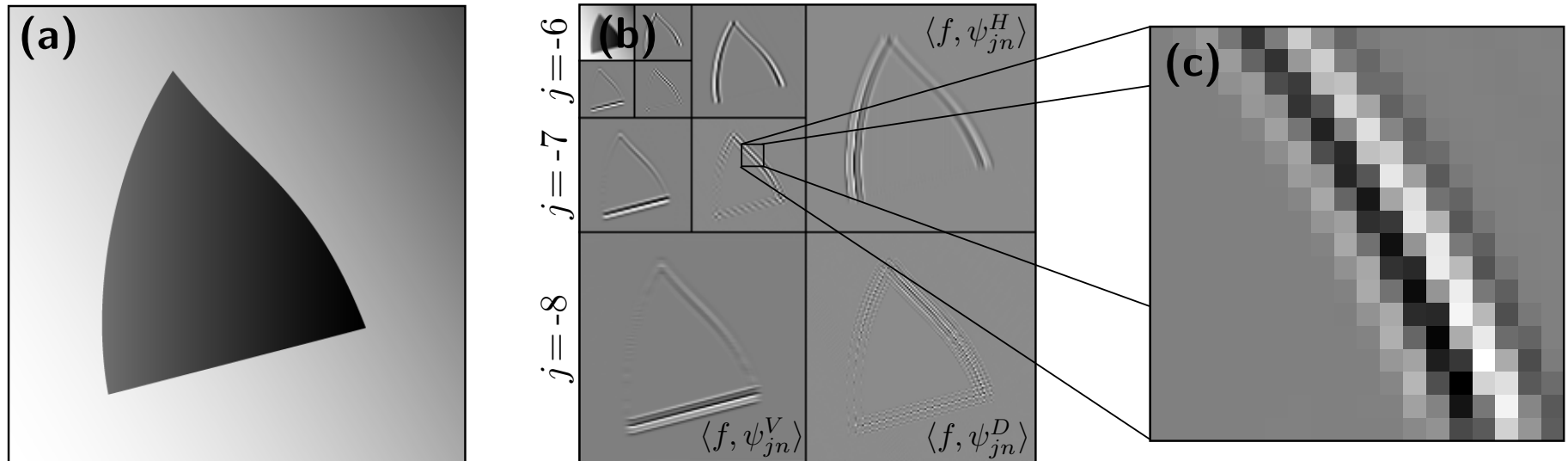
- Bandelettes : représentation à géométrie monoéchelle.
- Ondelettes : bonne représentation multiéchelles.

Bandelettes et géométrie multiéchelle



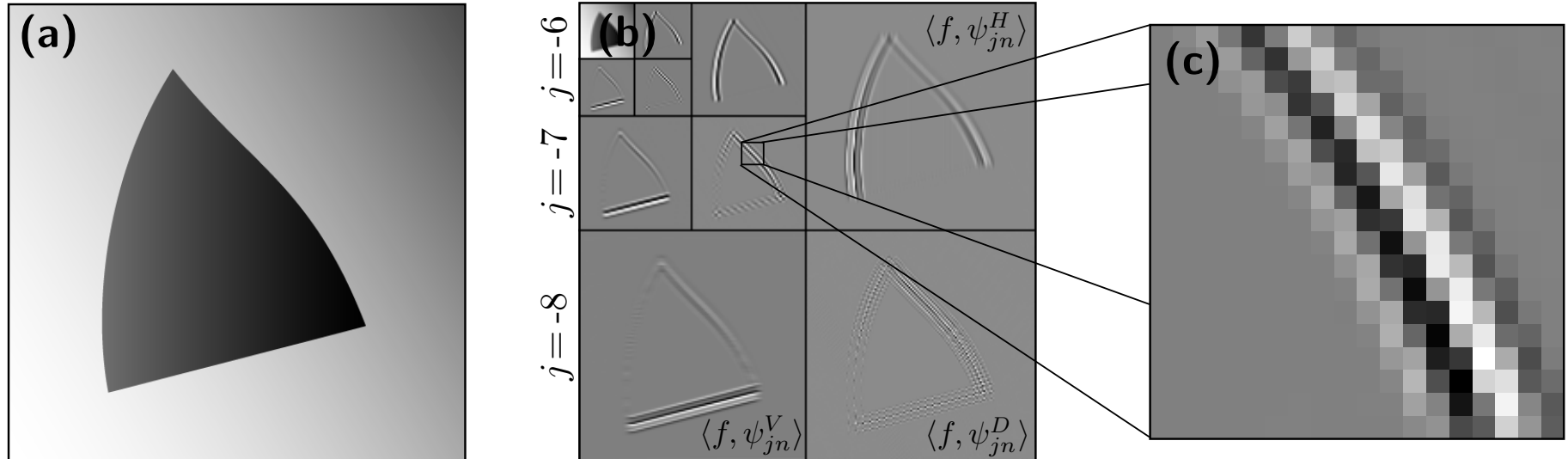
- Bandelettes : représentation à géométrie monoéchelle.
- Ondelettes : bonne représentation multiéchelles.
- Comment combiner les deux ?

Bandelettes et géométrie multiéchelle



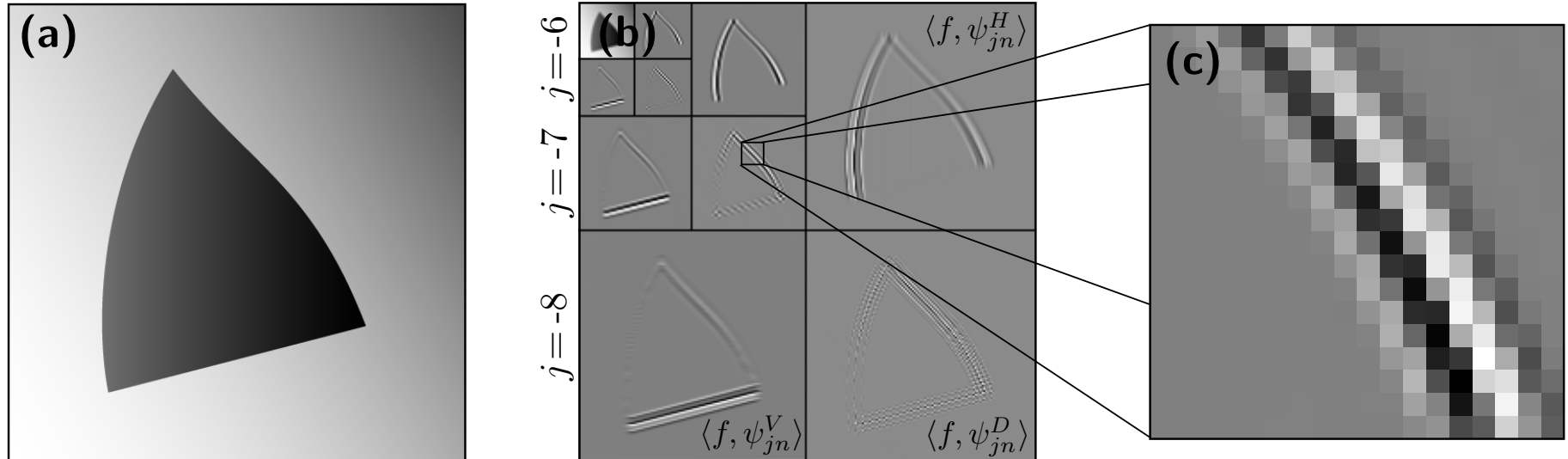
- Bandelettes : représentation à géométrie monoéchelle.
- Ondelettes : bonne représentation multiéchelles.
- Comment combiner les deux ?
- Analogie avec le système visuel.

Bandelettes et géométrie multiéchelle



- Bandelettes : représentation à géométrie monoéchelle.
- Ondelettes : bonne représentation multiéchelles.
- Comment combiner les deux ?
- Analogie avec le système visuel.
- Bandelettes sur les coefficients !

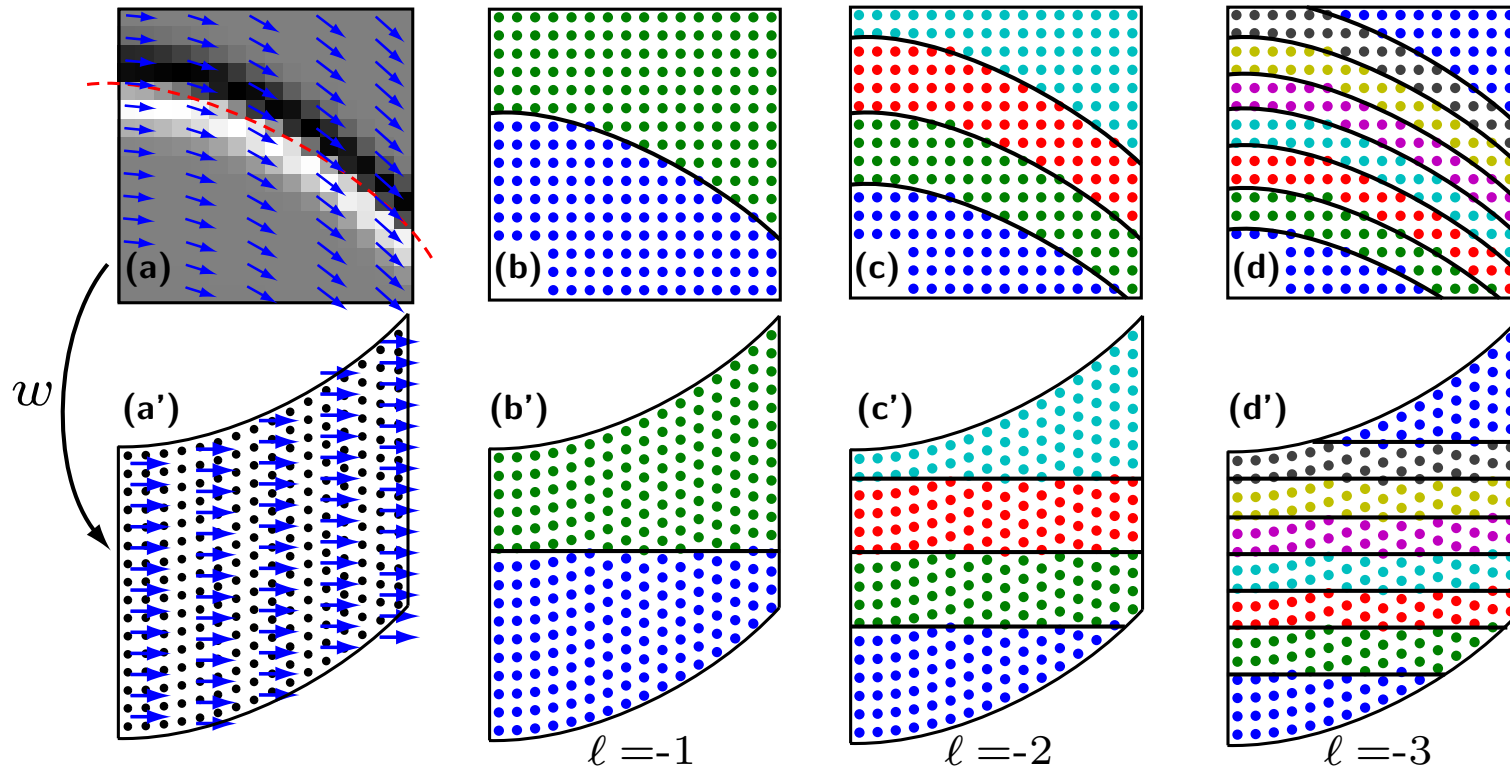
Bandelettes et géométrie multiéchelle



- Bandelettes : représentation à géométrie monoéchelle.
- Ondelettes : bonne représentation multiéchelles.
- Comment combiner les deux ?
- Analogie avec le système visuel.
- Bandelettes sur les coefficients !
- Similarités avec JPEG2000, Edgeprint,...

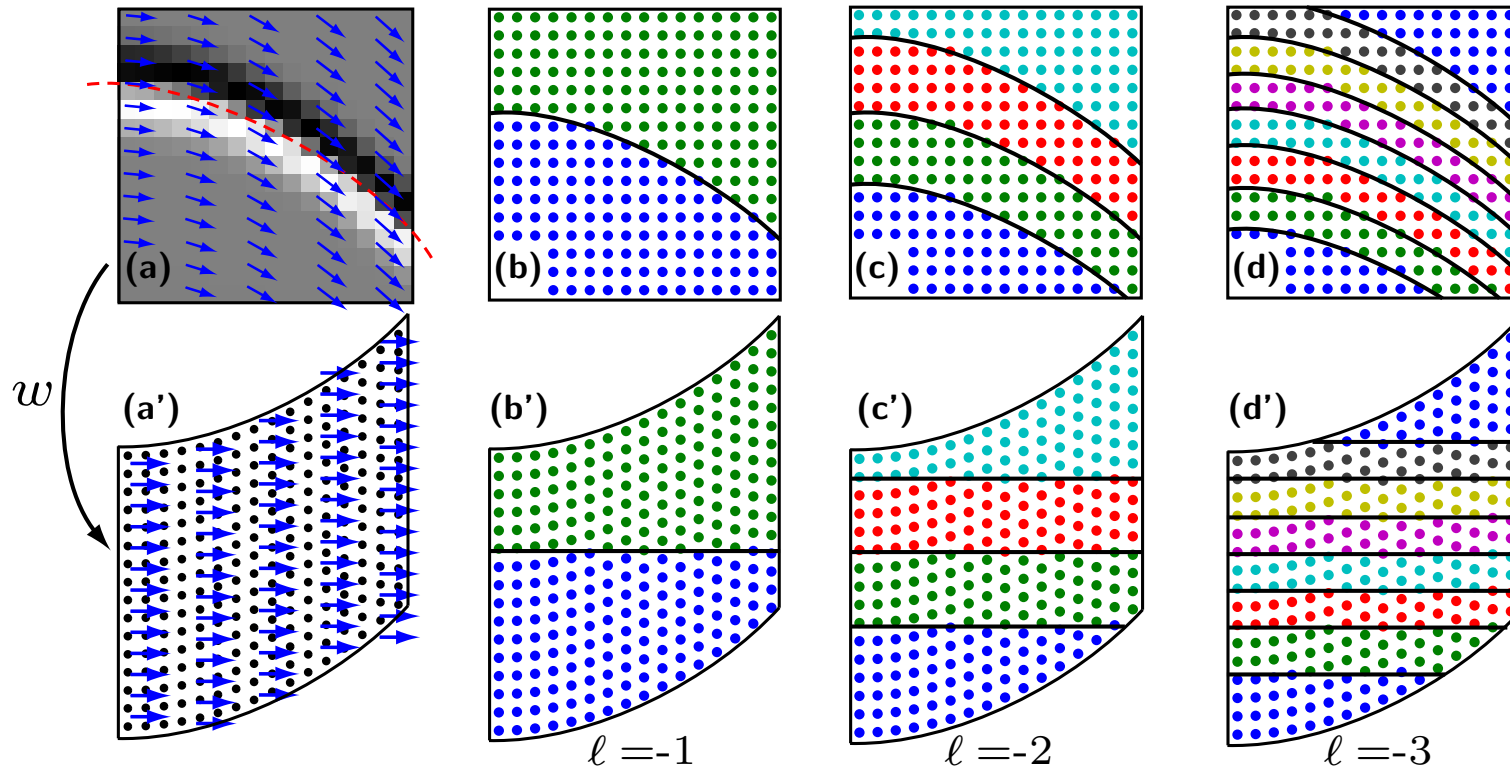
Base locale de bandelettes

Base locale de bandelettes



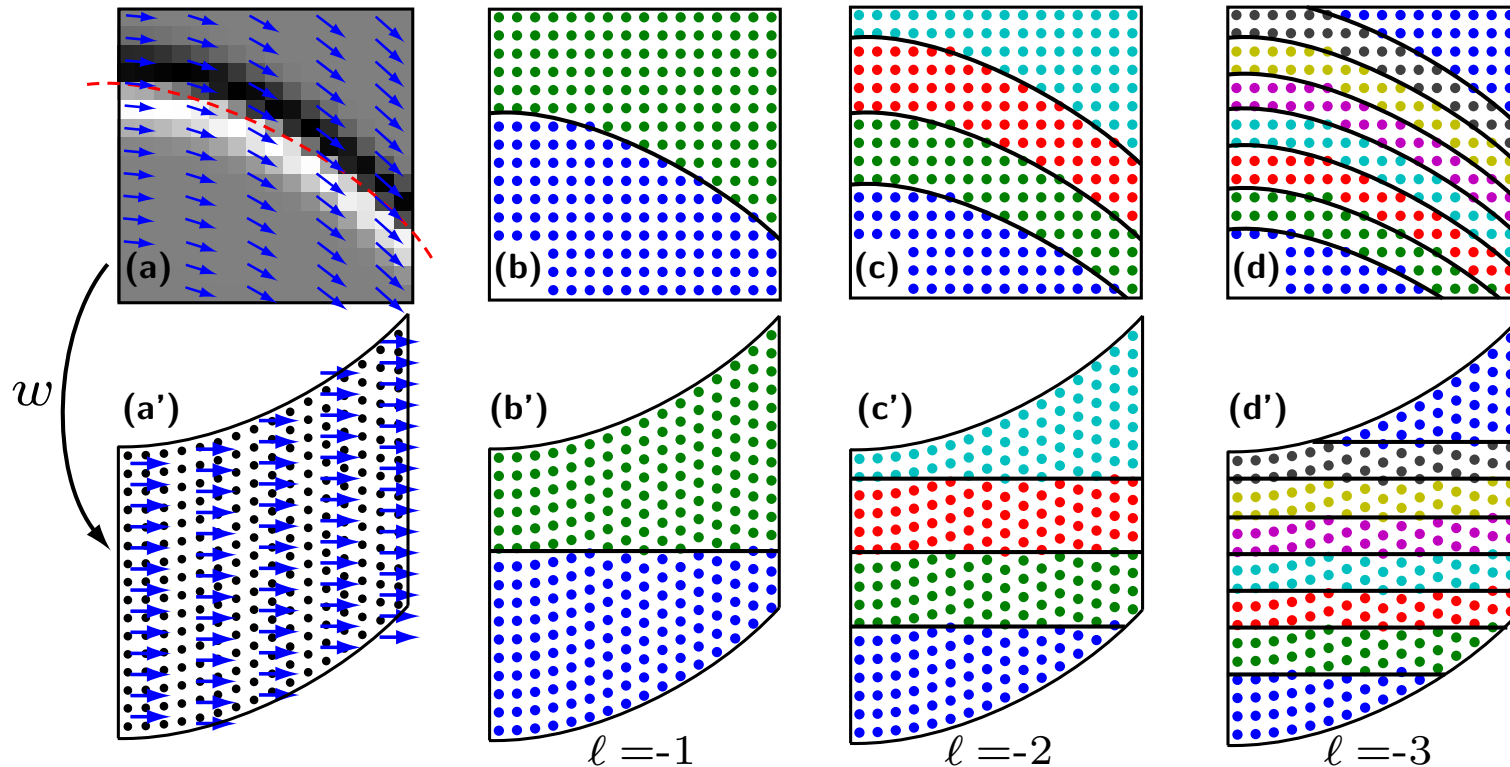
Bandelettes 2G (*Peyré*) : Changement de base orthogonale adapté à la géométrie sur les coefficients d'ondelette.

Base locale de bandelettes



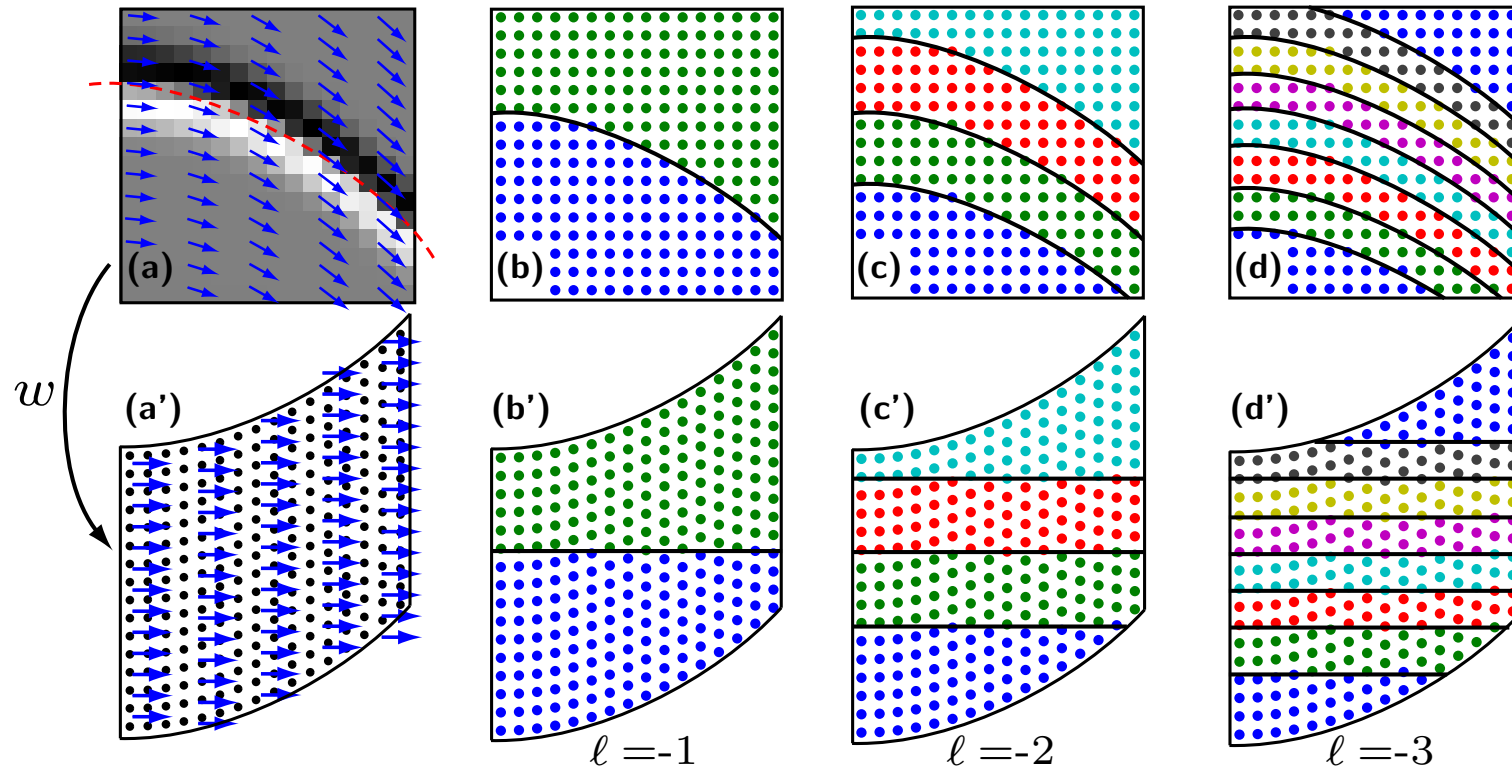
- Bandelettes 2G (*Peyré*) : Changement de base orthogonale adapté à la géométrie sur les coefficients d'ondelette.
- Multirésolution d'espaces d'approximations polynomiales par morceaux.

Base locale de bandelettes



- Bandelettes 2G (*Peyré*) : Changement de base orthogonale adapté à la géométrie sur les coefficients d'ondelette.
- Multirésolution d'espaces d'approximations polynomiales par morceaux.
- Base des compléments orthogonaux de ces espaces (*Alpert*).

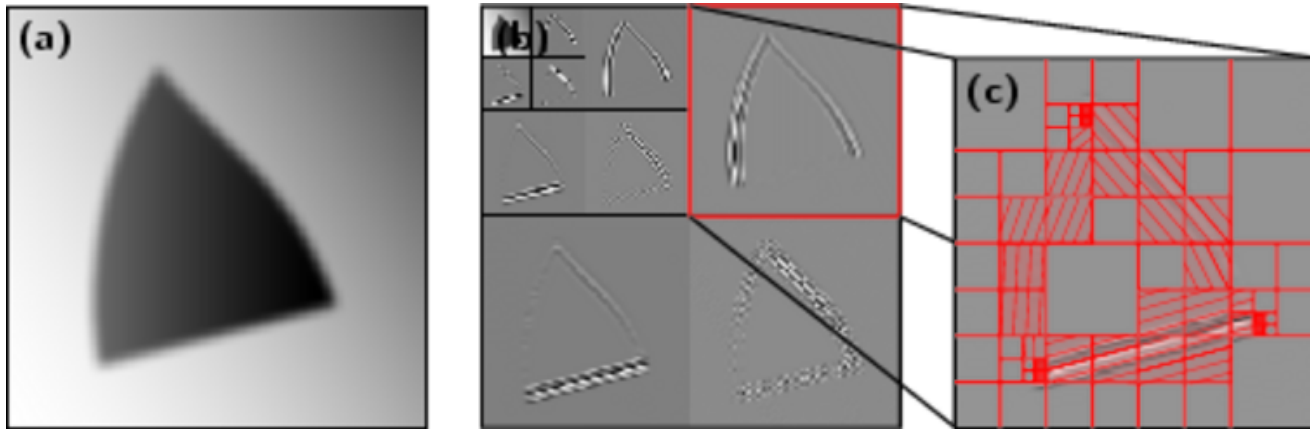
Base locale de bandelettes



- Bandelettes 2G (*Peyré*) : Changement de base orthogonale adapté à la géométrie sur les coefficients d'ondelette.
- Multirésolution d'espaces d'approximations polynomiales par morceaux.
- Base des compléments orthogonaux de ces espaces (*Alpert*).
- Image des ondelettes par ce changement de bases : bandelettes 2G.

Base de bandelettes 2G

Base de bandelettes 2G



Base de bandelettes :

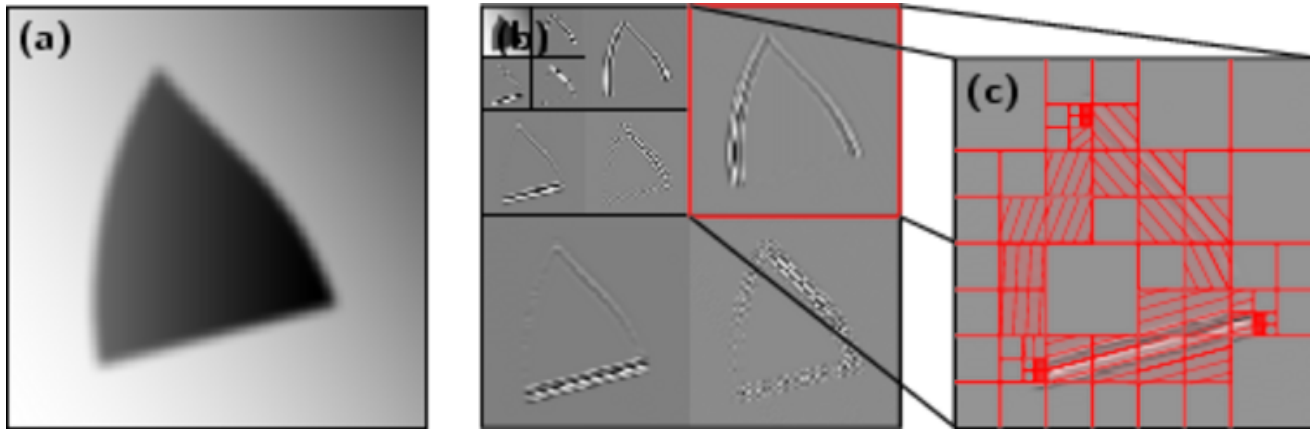


segmentation dyadique des sous-bandes,



géométrie dans chaque carrés.

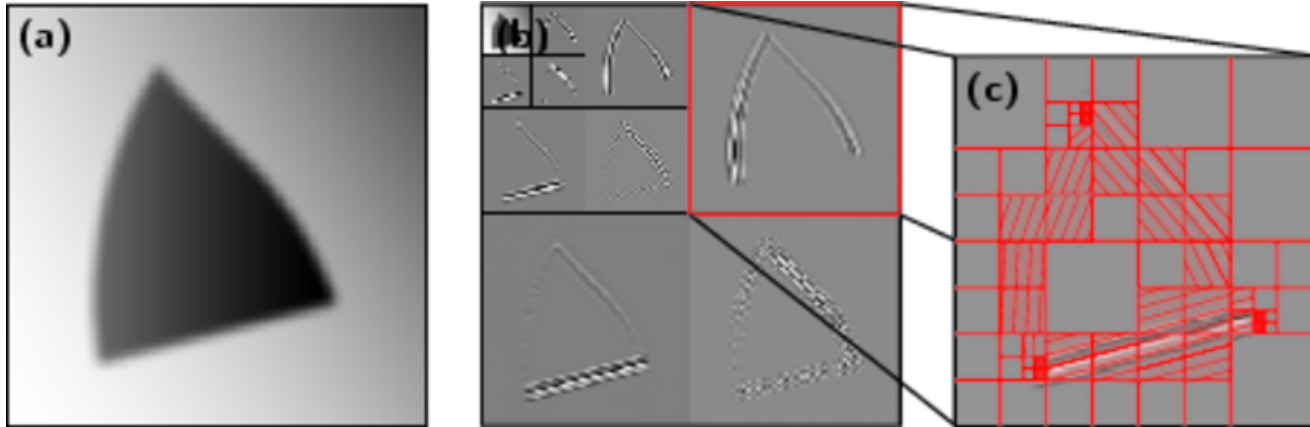
Base de bandelettes 2G



- Base de bandelettes :
 - segmentation dyadique des sous-bandes,
 - géométrie dans chaque carré.
- **Théorème** : Si f est $C^\alpha - C^\alpha$, alors, dans la meilleure base,

$$\|f - f_M\|^2 \leq CM^{-\alpha}$$

Base de bandelettes 2G



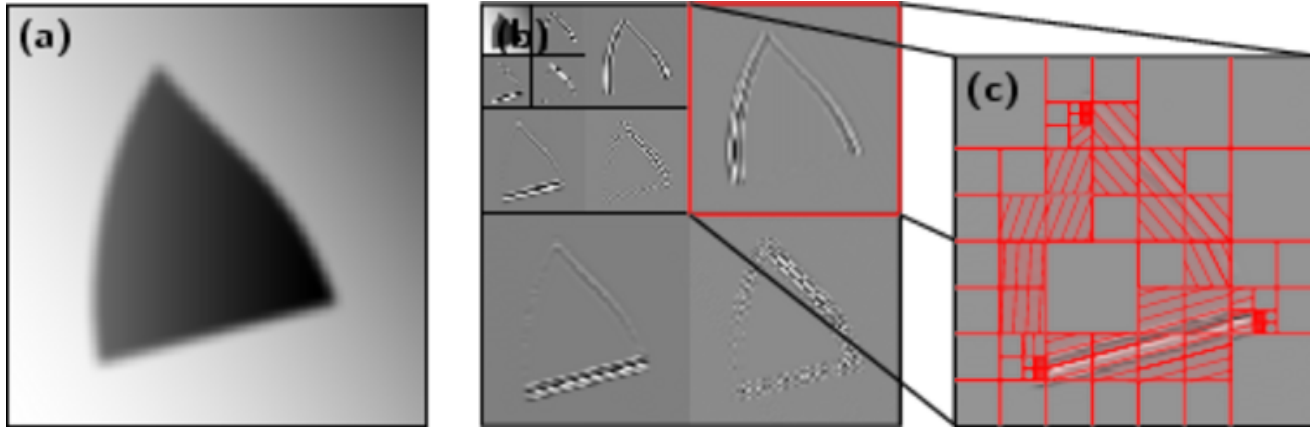
- Base de bandelettes :
 - segmentation dyadique des sous-bandes,
 - géométrie dans chaque carré.
- **Théorème** : Si f est $C^\alpha - C^\alpha$, alors, dans la meilleure base,

$$\|f - f_M\|^2 \leq CM^{-\alpha}$$

- Bases structurées et nombre polynomiale de bandelettes : algorithme d'optimisation par programmation dynamique (CART) de

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

Base de bandelettes 2G



- Base de bandelettes :
 - segmentation dyadique des sous-bandes,
 - géométrie dans chaque carré.
- **Théorème** : Si f est $C^\alpha - C^\alpha$, alors, dans la meilleure base,

$$\|f - f_M\|^2 \leq CM^{-\alpha}$$

- Bases structurées et nombre polynomiale de bandelettes : algorithme d'optimisation par programmation dynamique (CART) de

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

- Comment utiliser ces bases pour de l'estimation ?

Seuillage et sélection de modèles

Seuillage et sélection de modèles

- Retour sur l'estimateur oracle.

Seuillage et sélection de modèles

- Retour sur l'estimateur oracle.
- Risque oracle : $\Gamma_0 = \operatorname{argmin}_{\Gamma} \|f - f_{\Gamma}\|^2 + \frac{1}{N}|\Gamma|$.

Seuillage et sélection de modèles

- Retour sur l'estimateur oracle.
- Risque oracle : $\Gamma_0 = \operatorname{argmin}_{\Gamma} \|f - f_{\Gamma}\|^2 + \frac{1}{N}|\Gamma|$.
- Analogue empirique : $\Gamma_S = \operatorname{argmin}_{\Gamma} \|Y - Y_{\Gamma}\|^2 + \frac{\lambda_N}{N}|\Gamma|$.

Seuillage et sélection de modèles

- Retour sur l'estimateur oracle.
- Risque oracle : $\Gamma_0 = \operatorname{argmin}_{\Gamma} \|f - f_{\Gamma}\|^2 + \frac{1}{N}|\Gamma|$.
- Analogue empirique : $\Gamma_S = \operatorname{argmin}_{\Gamma} \|Y - Y_{\Gamma}\|^2 + \frac{\lambda_N}{N}|\Gamma|$.
- Minimisation : $\Gamma_S = \{n, |\langle Y, b_n \rangle| \geq \sqrt{\frac{\lambda_N}{N}}\}$ (seuillage) et $F_S = Y_{\Gamma_S}$.

Seuillage et sélection de modèles

- Retour sur l'estimateur oracle.
- Risque oracle : $\Gamma_0 = \operatorname{argmin}_{\Gamma} \|f - f_{\Gamma}\|^2 + \frac{1}{N}|\Gamma|$.
- Analogue empirique : $\Gamma_S = \operatorname{argmin}_{\Gamma} \|Y - Y_{\Gamma}\|^2 + \frac{\lambda_N}{N}|\Gamma|$.
- Minimisation : $\Gamma_S = \{n, |\langle Y, b_n \rangle| \geq \sqrt{\frac{\lambda_N}{N}}\}$ (seuillage) et $F_S = Y_{\Gamma_S}$.
- Sélection de modèles :

$$F_S = \operatorname{argmin}_{P_m Y, m \in \mathcal{M}_N} \|Y - P_m Y\|^2 + \operatorname{pen}(m) = \operatorname{argmin}_{P_m Y, m \in \mathcal{M}_N} \|Y - P_m Y\|^2 + \frac{\lambda_N}{N} .$$

avec $\mathcal{M}_N$, collection des modèles m = ensemble des sous-espaces engendrés par les N vecteurs de base.

Seuillage et sélection de modèles

- Retour sur l'estimateur oracle.
- Risque oracle : $\Gamma_0 = \operatorname{argmin}_{\Gamma} \|f - f_{\Gamma}\|^2 + \frac{1}{N}|\Gamma|$.
- Analogue empirique : $\Gamma_S = \operatorname{argmin}_{\Gamma} \|Y - Y_{\Gamma}\|^2 + \frac{\lambda_N}{N}|\Gamma|$.
- Minimisation : $\Gamma_S = \{n, |\langle Y, b_n \rangle| \geq \sqrt{\frac{\lambda_N}{N}}\}$ (seuillage) et $F_S = Y_{\Gamma_S}$.
- Sélection de modèles :

$$F_S = \operatorname{argmin}_{P_m Y, m \in \mathcal{M}_N} \|Y - P_m Y\|^2 + \operatorname{pen}(m) = \operatorname{argmin}_{P_m Y, m \in \mathcal{M}_N} \|Y - P_m Y\|^2 + \frac{\lambda_N}{N} .$$

avec $\mathcal{M}_N$, collection des modèles m = ensemble des sous-espaces engendrés par les N vecteurs de base.

- **Théorème (*Barron, Birgé, Massart*) $\simeq$ (*Donoho-Johnstone*)** : Pour $\lambda_N = \lambda \log N$ avec λ assez grand,

$$E(\|f - F_S\|^2) \leq C \min_{m \in \mathcal{M}_N} \|f - P_m f\|^2 + \lambda \frac{\log N}{N} \dim(m) .$$

Seuillage et sélection de modèles

- Retour sur l'estimateur oracle.
- Risque oracle : $\Gamma_0 = \operatorname{argmin}_{\Gamma} \|f - f_{\Gamma}\|^2 + \frac{1}{N}|\Gamma|$.
- Analogue empirique : $\Gamma_S = \operatorname{argmin}_{\Gamma} \|Y - Y_{\Gamma}\|^2 + \frac{\lambda_N}{N}|\Gamma|$.
- Minimisation : $\Gamma_S = \{n, |\langle Y, b_n \rangle| \geq \sqrt{\frac{\lambda_N}{N}}\}$ (seuillage) et $F_S = Y_{\Gamma_S}$.
- Sélection de modèles :

$$F_S = \operatorname{argmin}_{P_m Y, m \in \mathcal{M}_N} \|Y - P_m Y\|^2 + \operatorname{pen}(m) = \operatorname{argmin}_{P_m Y, m \in \mathcal{M}_N} \|Y - P_m Y\|^2 + \frac{\lambda_N}{N} .$$

avec $\mathcal{M}_N$, collection des modèles m = ensemble des sous-espaces engendrés par les N vecteurs de base.

- **Théorème (*Barron, Birgé, Massart*) $\simeq$ (*Donoho-Johnstone*)** : Pour $\lambda_N = \lambda \log N$ avec λ assez grand,

$$E(\|f - F_S\|^2) \leq C \min_{m \in \mathcal{M}_N} \|f - P_m f\|^2 + \lambda \frac{\log N}{N} \dim(m) .$$

- Cadre permettant de travailler dans plusieurs bases à la fois...

Sélection de modèles

Sélection de modèles

- Cadre simplifié : $\mathcal{M}_N$ collection de modèles m engendrés par des vecteurs choisis parmi κ vecteurs différents (a priori $\kappa \neq N$).

Sélection de modèles

- Cadre simplifié : $\mathcal{M}_N$ collection de modèles m engendrés par des vecteurs choisis parmi κ vecteurs différents (a priori $\kappa \neq N$).
- **Théorème (*Barron, Birgé, Massart*) :**

$$F_S = \operatorname{argmin}_{P_m Y, m \in \mathcal{M}_N} \|Y - P_m Y\|^2 + \lambda \frac{\log \kappa}{N} \dim(m)$$

satisfait

$$E(\|f - F_S\|^2) \leq C \min_{m \in \mathcal{M}_N} \|f - P_m f\|^2 + \lambda \frac{\log \kappa}{N} \dim(m) \quad .$$

Sélection de modèles

- Cadre simplifié : $\mathcal{M}_N$ collection de modèles m engendrés par des vecteurs choisis parmi κ vecteurs différents (a priori $\kappa \neq N$).
- **Théorème (*Barron, Birgé, Massart*) :**

$$F_S = \operatorname{argmin}_{P_m Y, m \in \mathcal{M}_N} \|Y - P_m Y\|^2 + \lambda \frac{\log \kappa}{N} \dim(m)$$

satisfait

$$E(\|f - F_S\|^2) \leq C \min_{m \in \mathcal{M}_N} \|f - P_m f\|^2 + \lambda \frac{\log \kappa}{N} \dim(m) \quad .$$

- **Théorème (Maxiset) :**

$$E(\|f - F_S\|^2) \leq C \left(\frac{\log \kappa}{N} \right)^{\frac{\beta}{\beta+1}}$$

$$\Leftrightarrow \min_{m \in \mathcal{M}_N} \|f - P_m f\|^2 + \lambda \frac{\log \kappa}{N} \dim(m) \leq C \left(\frac{\log \kappa}{N} \right)^{\frac{\beta}{\beta+1}}$$

$$\Leftrightarrow \min_{m \in \mathcal{M}_N} \|f - P_m f\|^2 + T^2 \dim(m) \leq C (T^2)^{\frac{\beta}{\beta+1}}$$

$$\Leftrightarrow \|f - f_M\|^2 \leq C M^{-\beta} \Leftrightarrow f \in \mathcal{A}^{\beta} \quad .$$

Estimation géométrique par bandelettes

Estimation géométrique par bandelettes

- Contrôle polynomial sur le nombre total κ de bandelettes :
 $\log \kappa = \gamma \log N$.

Estimation géométrique par bandelettes

- Contrôle polynomial sur le nombre total κ de bandelettes :
 $\log \kappa = \gamma \log N$.
- Sélection de bandelettes :

$$F_S = \operatorname{argmin}_{m \in \mathcal{M}} \|Y - P_m Y\|^2 + \lambda \frac{\log N}{N} \dim(m)$$

avec $\mathcal{M}$ qui contient les sous-espaces des bases de bandelettes.

Estimation géométrique par bandelettes

- Contrôle polynomial sur le nombre total κ de bandelettes :
 $\log \kappa = \gamma \log N$.

- Sélection de bandelettes :

$$F_S = \operatorname{argmin}_{m \in \mathcal{M}} \|Y - P_m Y\|^2 + \lambda \frac{\log N}{N} \dim(m)$$

avec $\mathcal{M}$ qui contient les sous-espaces des bases de bandelettes.

- Même algorithme de minimisation que pour l'approximation.

Estimation géométrique par bandelettes

- Contrôle polynomial sur le nombre total κ de bandelettes :
 $\log \kappa = \gamma \log N$.

- Sélection de bandelettes :

$$F_S = \operatorname{argmin}_{m \in \mathcal{M}} \|Y - P_m Y\|^2 + \lambda \frac{\log N}{N} \dim(m)$$

avec $\mathcal{M}$ qui contient les sous-espaces des bases de bandelettes.

- Même algorithme de minimisation que pour l'approximation.
- **Théorème (Quasi optimalité minimax)** : si $f \in \mathbf{C}^\alpha - \mathbf{C}^\alpha$ alors

$$E(\|f - F_S\|^2) \leq C \left(\frac{\log N}{N} \right)^{\frac{\alpha}{\alpha+1}} .$$

Estimation géométrique par bandelettes

- Contrôle polynomial sur le nombre total κ de bandelettes :
 $\log \kappa = \gamma \log N$.

- Sélection de bandelettes :

$$F_S = \operatorname{argmin}_{m \in \mathcal{M}} \|Y - P_m Y\|^2 + \lambda \frac{\log N}{N} \dim(m)$$

avec $\mathcal{M}$ qui contient les sous-espaces des bases de bandelettes.

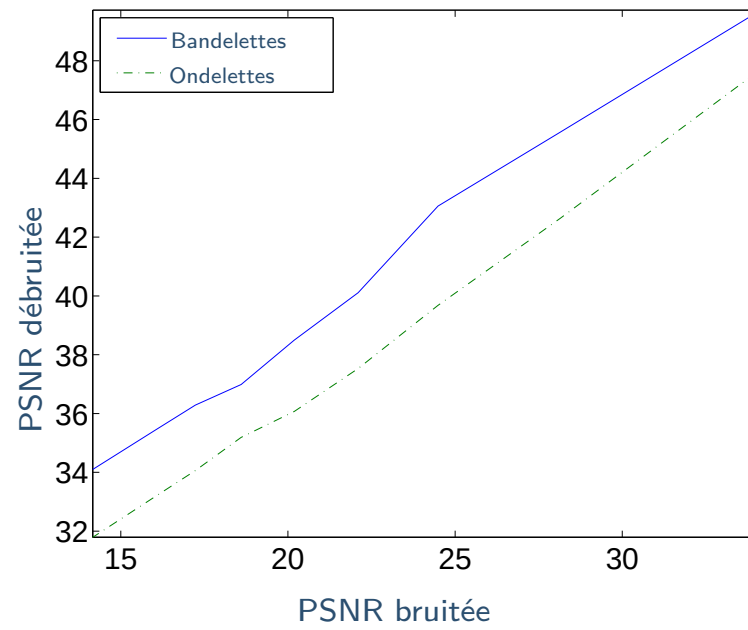
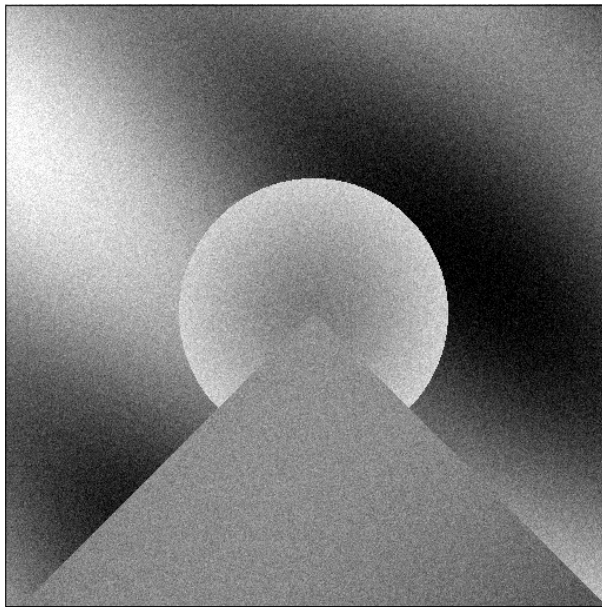
- Même algorithme de minimisation que pour l'approximation.
- **Théorème (Quasi optimalité minimax)** : si $f \in \mathbf{C}^\alpha - \mathbf{C}^\alpha$ alors

$$E(\|f - F_S\|^2) \leq C \left(\frac{\log N}{N} \right)^{\frac{\alpha}{\alpha+1}} .$$

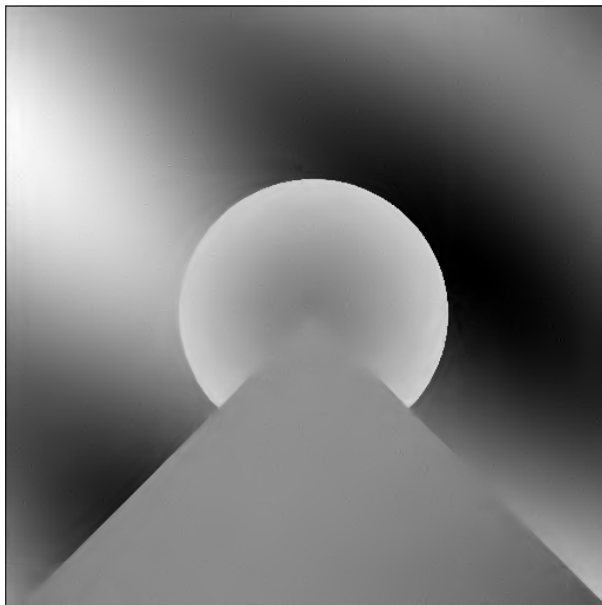
- **Théorème (Maxiset)** :

$$E(\|f - F_S\|^2) \leq C \left(\frac{\log N}{N} \right)^{\frac{\alpha}{\alpha+1}} \Leftrightarrow f \in \mathcal{A}^\alpha \Leftrightarrow \forall M, \exists \mathcal{B}, \|f - f_M\|^2 \leq C M^{-\alpha}$$

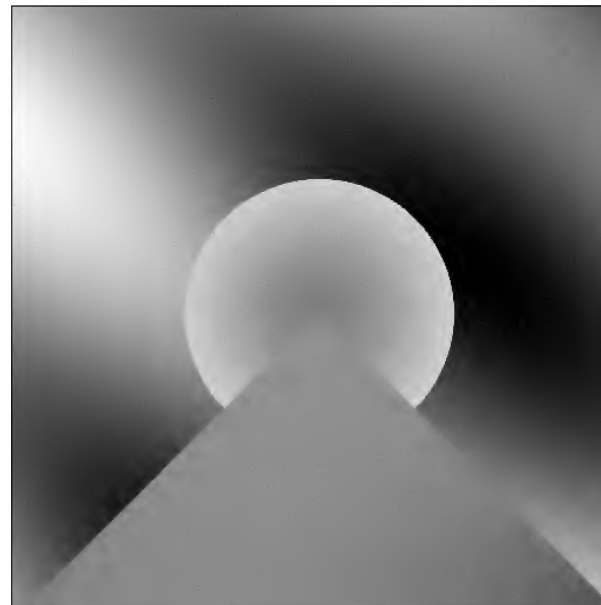
Bruitée (20,19 dB)



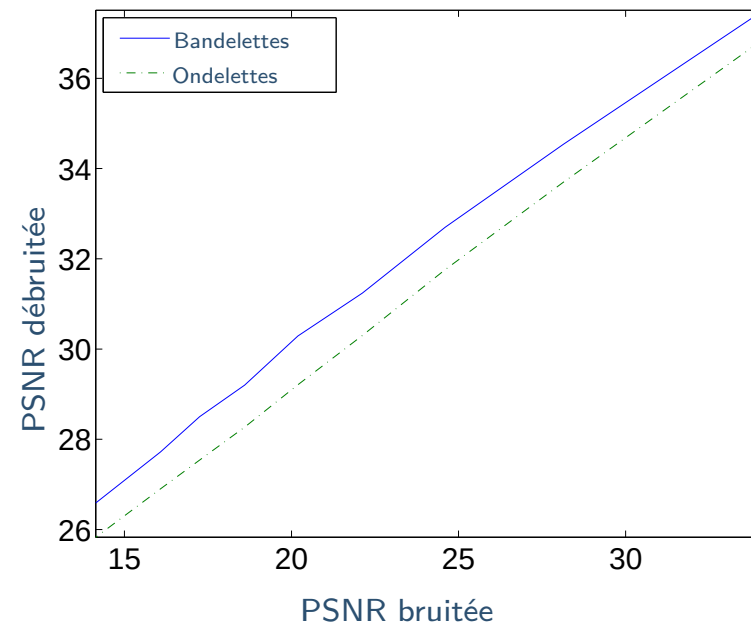
Bandelettes (30,29 dB)



Ondelettes (28,21 dB)



Bruitée (20,19 dB)



Bandelettes (30,29 dB)



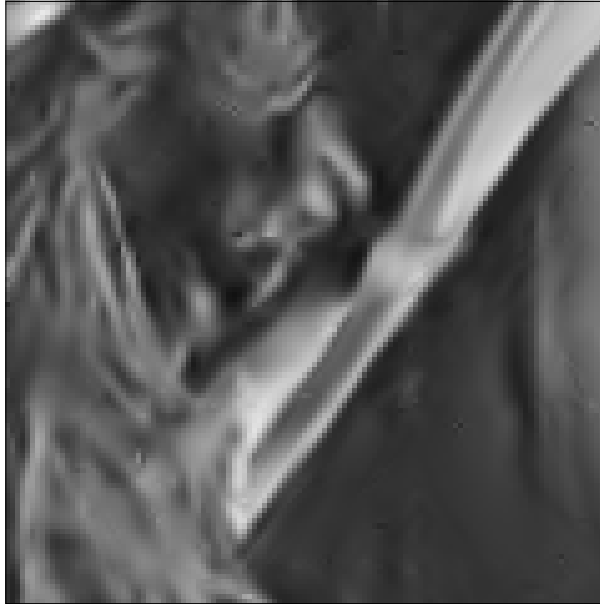
Ondelettes (28,21 dB)



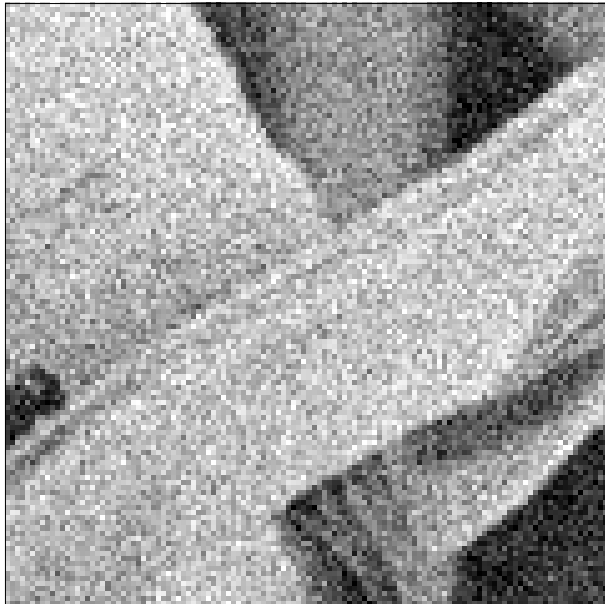
Bruitée



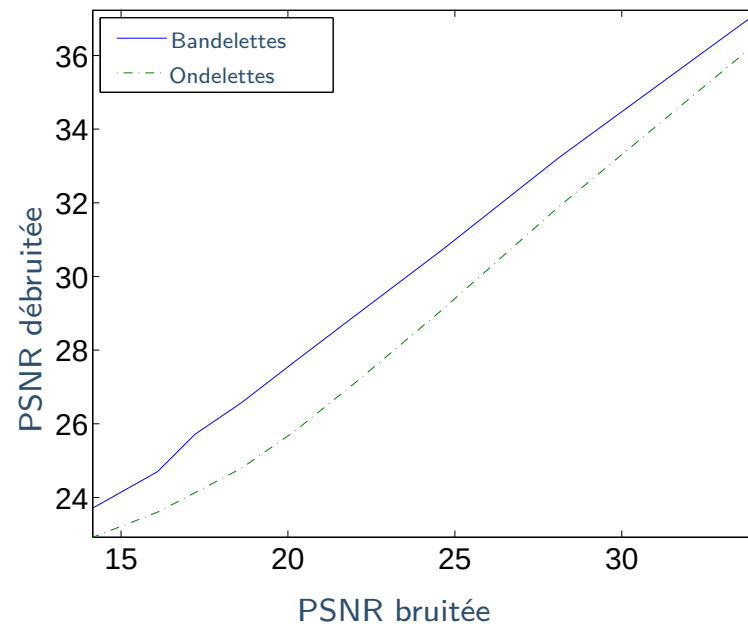
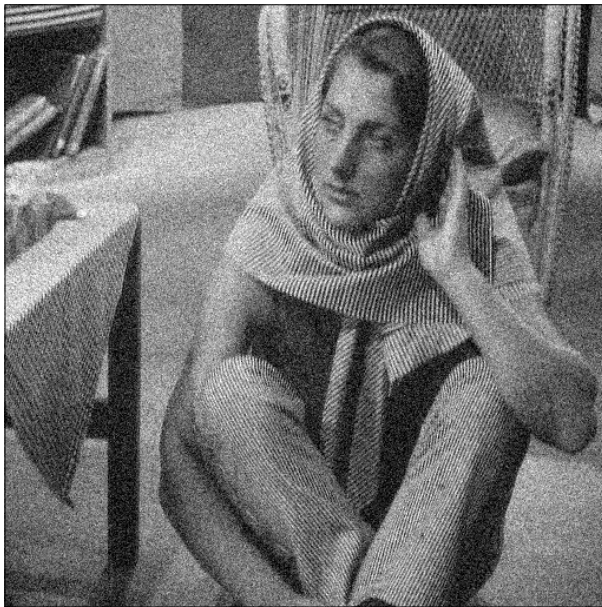
Bandelettes



Ondelettes



Bruitée (20,19 dB)



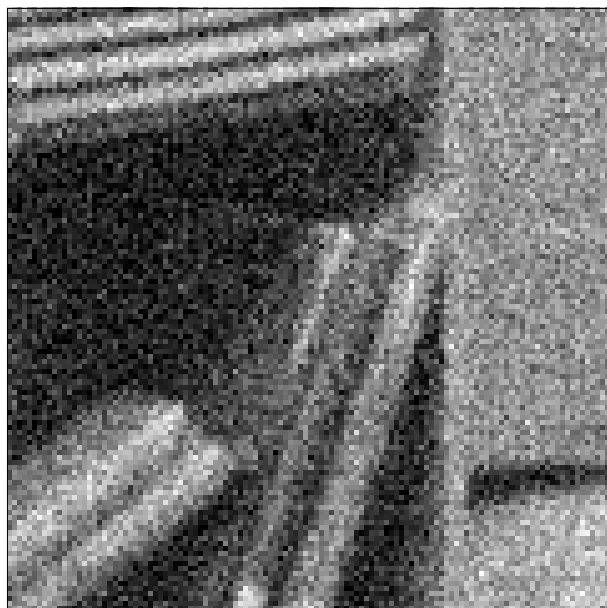
Bandelettes (27,68 dB)



Ondelettes (25,79 dB)



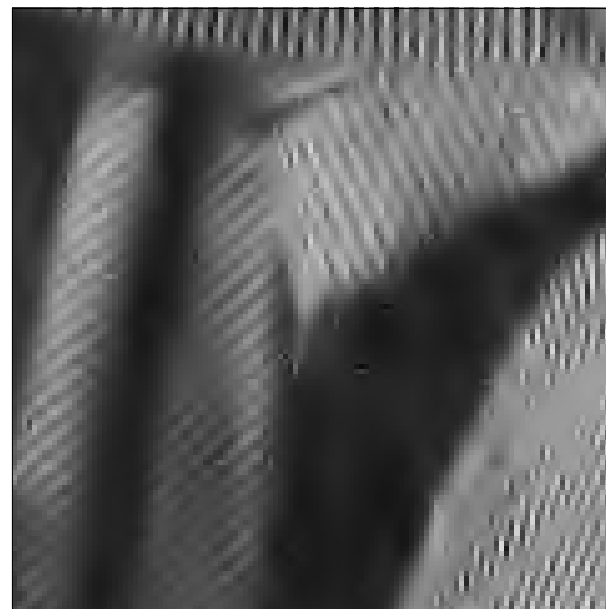
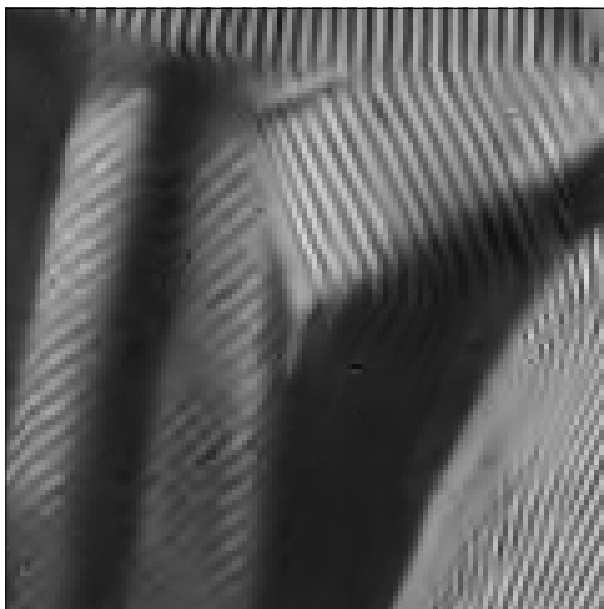
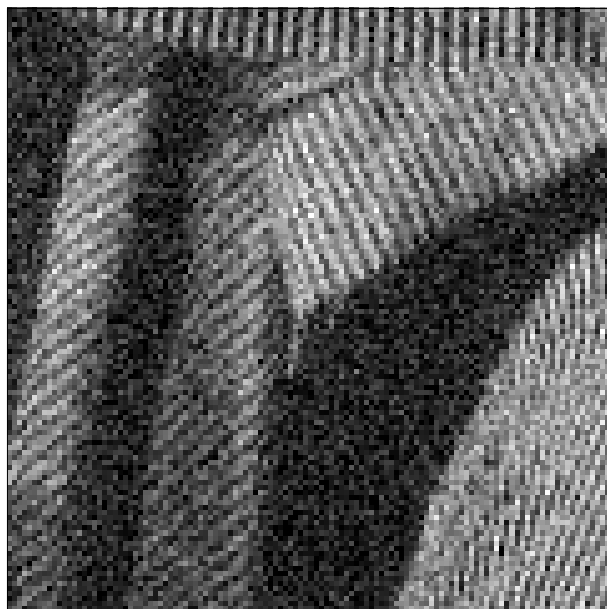
Bruitée



Bandelettes



Ondelettes



Conclusion

Conclusion

- Rôle crucial de l'approximation en traitement du signal utilisant des bases (estimation par seuillage/sélection de modèles, compression, ...).

Conclusion

- Rôle crucial de l'approximation en traitement du signal utilisant des bases (estimation par seuillage/sélection de modèles, compression, ...).
- Bandelettes : une représentation adaptée à la géométrie des images.

Conclusion

- Rôle crucial de l'approximation en traitement du signal utilisant des bases (estimation par seuillage/sélection de modèles, compression, ...).
- Bandelettes : une représentation adaptée à la géométrie des images.
- Qu'estiment bien les bandelettes ?

Conclusion

- Rôle crucial de l'approximation en traitement du signal utilisant des bases (estimation par seuillage/sélection de modèles, compression, ...).
- Bandelettes : une représentation adaptée à la géométrie des images.
- Qu'estiment bien les bandelettes ?
 - Maxiset : les images qui s'approchent bien en bandelettes !

Conclusion

- Rôle crucial de l'approximation en traitement du signal utilisant des bases (estimation par seuillage/sélection de modèles, compression, ...).
- Bandelettes : une représentation adaptée à la géométrie des images.
- Qu'estiment bien les bandelettes ?
 - Maxiset : les images qui s'approchent bien en bandelettes !
 - Réalité : images naturelles ?

Conclusion

- Rôle crucial de l'approximation en traitement du signal utilisant des bases (estimation par seuillage/sélection de modèles, compression, ...).
- Bandelettes : une représentation adaptée à la géométrie des images.
- Qu'estiment bien les bandelettes ?
 - Maxiset : les images qui s'approchent bien en bandelettes !
 - Réalité : images naturelles ?
- Problèmes ouverts :

Conclusion

- Rôle crucial de l'approximation en traitement du signal utilisant des bases (estimation par seuillage/sélection de modèles, compression, ...).
- Bandelettes : une représentation adaptée à la géométrie des images.
- Qu'estiment bien les bandelettes ?
 - Maxiset : les images qui s'approchent bien en bandelettes !
 - Réalité : images naturelles ?
- Problèmes ouverts :
 - Caractérisation fonctionnelle des espaces d'approximation en bandelettes,

Conclusion

- Rôle crucial de l'approximation en traitement du signal utilisant des bases (estimation par seuillage/sélection de modèles, compression, ...).
- Bandelettes : une représentation adaptée à la géométrie des images.
- Qu'estiment bien les bandelettes ?
 - Maxiset : les images qui s'approchent bien en bandelettes !
 - Réalité : images naturelles ?
- Problèmes ouverts :
 - Caractérisation fonctionnelle des espaces d'approximation en bandelettes,
 - Autres pénalisations ($l_1, \dots$),

Conclusion

- Rôle crucial de l'approximation en traitement du signal utilisant des bases (estimation par seuillage/sélection de modèles, compression, ...).
- Bandelettes : une représentation adaptée à la géométrie des images.
- Qu'estiment bien les bandelettes ?
 - Maxiset : les images qui s'approchent bien en bandelettes !
 - Réalité : images naturelles ?
- Problèmes ouverts :
 - Caractérisation fonctionnelle des espaces d'approximation en bandelettes,
 - Autres pénalisations ($l_1, \dots$),
 - Combinaisons/agrégations d'estimateurs...