Texture segmentation based on fractal attributes

Convex functional minimization and generalized Stein formalism for automated regularization parameter selection

Barbara Pascal

April 7th 2023

Journées ANR Mistic

Joint work with **Patrice Abry**, **Nelly Pustelnik**, **Valérie Vidal** and **Samuel Vaiter**

Image segmentation

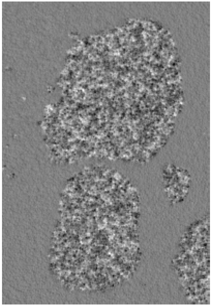
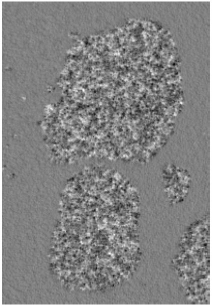
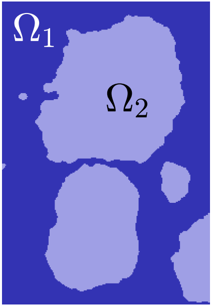


Image segmentation

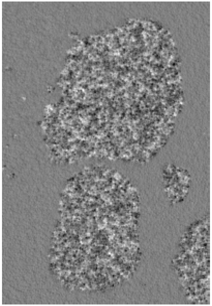


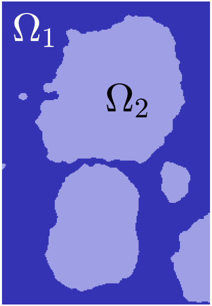


Goal : partitioning the image into K homogeneous regions

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

Image segmentation



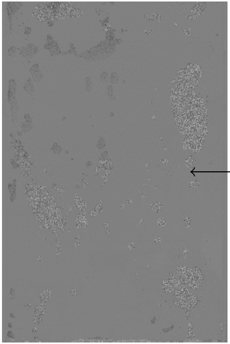


Goal : partitioning the image into K **homogeneous** regions

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

Multiphase flows through porous media

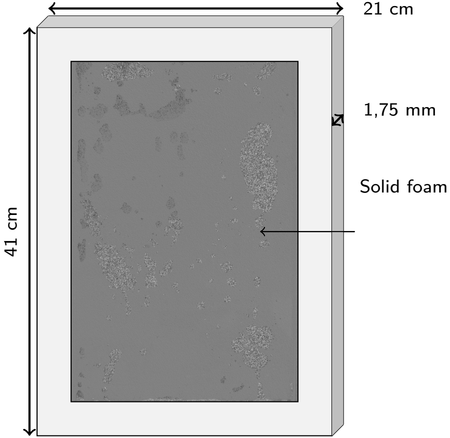
Laboratoire de Physique, ENS Lyon, V. Vidal, T. Busser, (M. Serres, IFPEN)



Solid foam

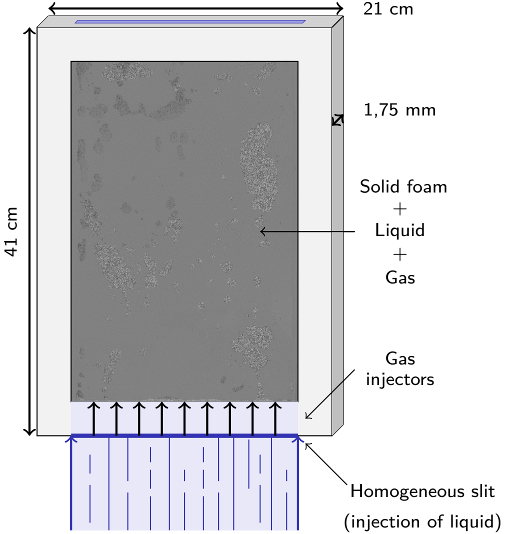
Multiphase flows through porous media

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



Multiphase flows through porous media

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



→ 1600 × 1100 pixels
→ video : ∼ 1000 images
→ phase space : ∼ 10 flow rates

Outline of the presentation

1. Texture characterization and synthesis

→ fractal attributes
- local variance σ^2
- local regularity h

2. Functional design for variational approaches

→ Penalized least squares
- free contours
- co-localized contours

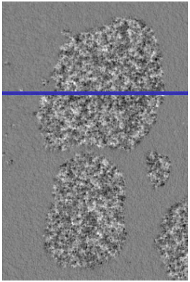
3. Accelerated minimization algorithms

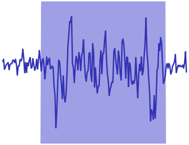
→ proximal splitting algorithms
- computation of proximity operators
- strong convexity acceleration

4. Hyperparameter tuning

→ SURE under correlated Gaussian noise
- projected estimation error

Piecewise monofractal model

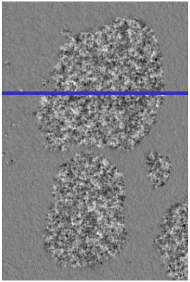


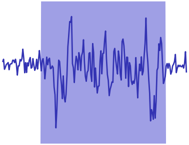


Piecewise monofractal model

Fractals attributes

- variance σ^2 — *amplitude of variations*

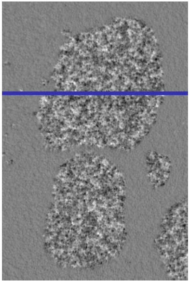


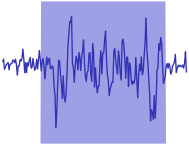


Piecewise monofractal model

Fractals attributes

- variance σ^2 — *amplitude of variations*
- local regularity h — *scale invariance*



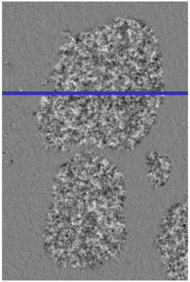


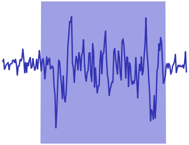
Piecewise monofractal model

Fractals attributes

- variance σ^2 *amplitude of variations*
- local regularity h *scale invariance*

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



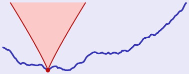


Piecewise monofractal model

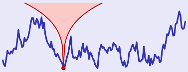
Fractals attributes

- variance σ^2 — *amplitude of variations*
- local regularity h — *scale invariance*

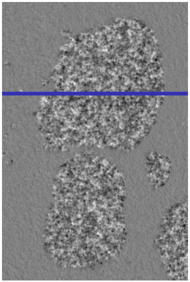
$$|f(x) - f(y)| \leq \sigma(x)|x-y|^{h(x)}$$

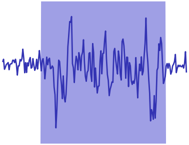


$h(x) \equiv h_1 = 0.9$



$h(x) \equiv h_2 = 0.3$



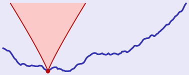


Piecewise monofractal model

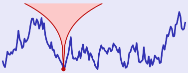
Fractals attributes

- variance σ^2 — *amplitude of variations*
- local regularity h — *scale invariance*

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



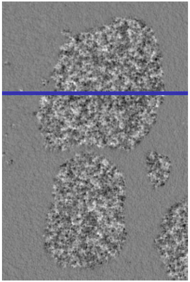
$h(x) \equiv h_1 = 0.9$

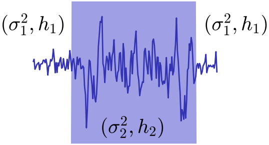


$h(x) \equiv h_2 = 0.3$

Segmentation

- h and σ^2 piecewise constant





Piecewise monofractal model

Fractals attributes

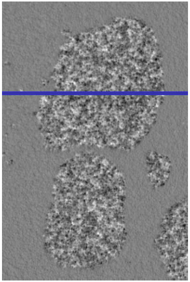
- variance σ^2 — *amplitude of variations*
- local regularity h — *scale invariance*

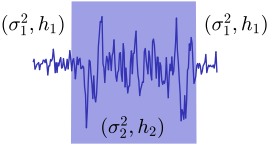
$$|f(x) - f(y)| \leq \sigma(x)|x - y|^{h(x)}$$

$h(x) \equiv h_1 = 0.9$ $\qquad$ $h(x) \equiv h_2 = 0.3$

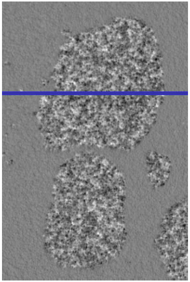
Segmentation

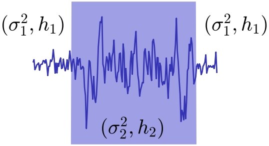
- h and σ^2 piecewise constant
- region Ω_k characterized by (h_k, σ_k^2)



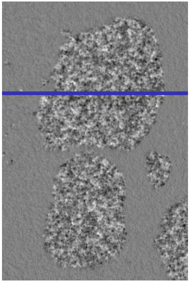


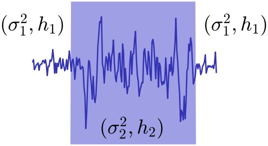
Synthesis of piecewise fractal textures





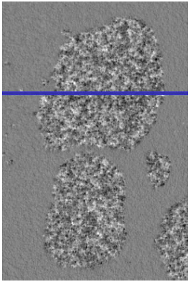
Synthesis of piecewise fractal textures

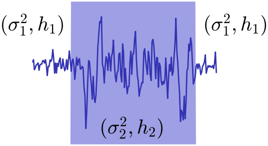




Which synthetic textures model

Synthesis of piecewise fractal textures

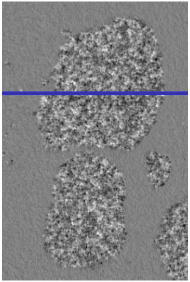


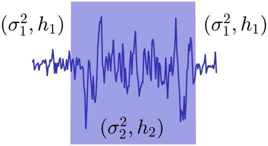


Which synthetic textures model

- resembling real-world textures,

Synthesis of piecewise fractal textures

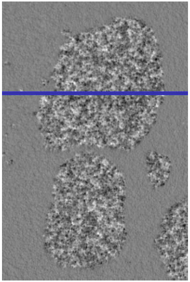


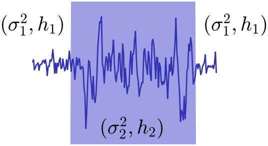


Which synthetic textures model

- resembling real-world textures,
- characterized by (h, σ^2),

Synthesis of piecewise fractal textures

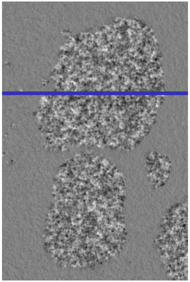


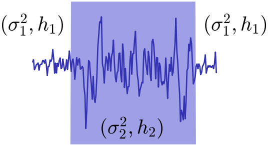


Which synthetic textures model

- resembling real-world textures,
- characterized by (h, σ^2),
- easy to "patch"?

Synthesis of piecewise fractal textures



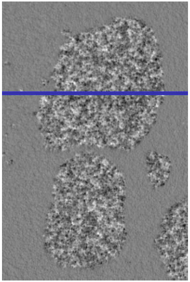


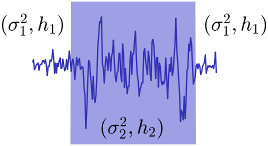
Which synthetic textures model

- resembling real-world textures,
- characterized by (h, σ^2),
- easy to "patch"?

Proposition of a Gaussian random field being

Synthesis of piecewise fractal textures





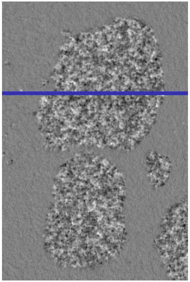
Which synthetic textures model

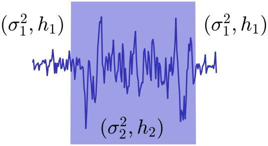
- resembling real-world textures,
- characterized by (h, σ^2),
- easy to "patch"?

Proposition of a Gaussian random field being

- isotropic,

Synthesis of piecewise fractal textures





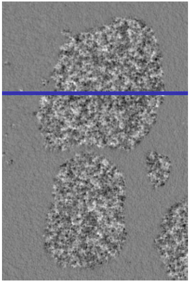
Which synthetic textures model

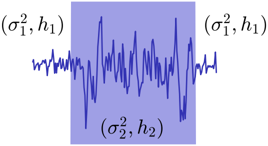
- resembling real-world textures,
- characterized by (h, σ^2),
- easy to "patch"?

Proposition of a Gaussian random field being

- isotropic,
- self-similar, with local regularity h,

Synthesis of piecewise fractal textures





Which synthetic textures model

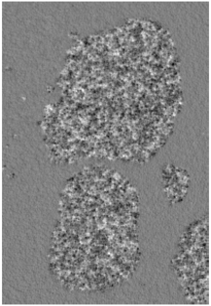
- resembling real-world textures,
- characterized by (h, σ^2),
- easy to "patch"?

Proposition of a Gaussian random field being

- isotropic,
- self-similar, with local regularity h,
- stationary, with variance σ^2.

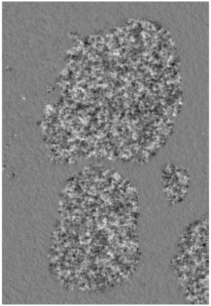
Synthesis of piecewise fractal textures

Real-world texture

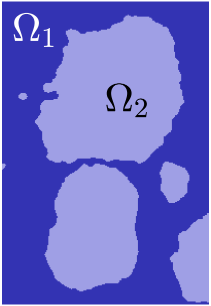


Synthesis of piecewise fractal textures

Real-world texture



Mask



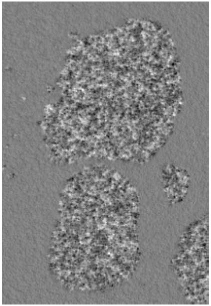
$\Omega_1 : (\sigma_1^2, h_1) \quad \Omega_2 : (\sigma_2^2, h_2)$

Synthesis of piecewise fractal textures

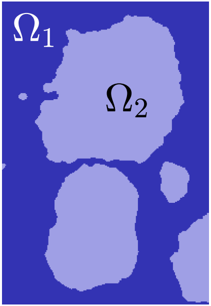
Fractional Brownian field

$$b_h(\underline{x}) = \sigma \int_{\mathbb{R}^2} \frac{\mathrm{e}^{-\mathrm{i}\langle \underline{x}, \underline{k} \rangle} - 1}{C_h^{1/2} \|\underline{k}\|^{H+1}} \, \widehat{w}(\mathrm{d}\underline{k})$$

Real-world texture



Mask



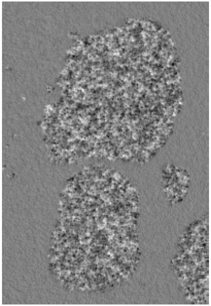
$\Omega_1 : (\sigma_1^2, h_1) \quad \Omega_2 : (\sigma_2^2, h_2)$

Synthesis of piecewise fractal textures

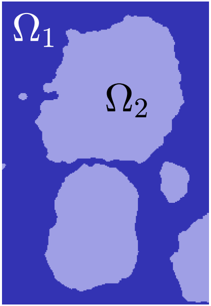
Fractional Brownian field

$$b_h(\underline{x}) = \sigma \int_{\mathbb{R}^2} \frac{\mathrm{e}^{-\mathrm{i}\langle \underline{x}, \underline{k} \rangle} - 1}{C_h^{1/2} \|\underline{k}\|^{H+1}} \, \widehat{w}(\mathrm{d}\underline{k})$$

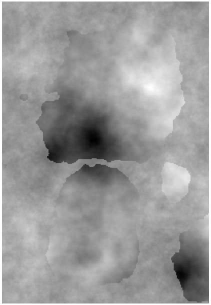
Real-world texture



Mask



Synthetic texture



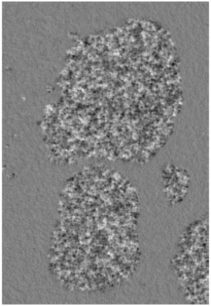
$\Omega_1 : (\sigma_1^2, h_1) \qquad \Omega_2 : (\sigma_2^2, h_2)$

Synthesis of piecewise fractal textures

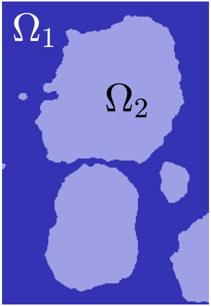
Fractional Gaussian field *stationary*

$$g_h(\underline{x}) = \frac{1}{2}\underbrace{\left(b_h(\underline{x}+\underline{e}_1) - b_h(\underline{x})\right)}_{\text{horizontal increment}} + \frac{1}{2}\underbrace{\left(b_h(\underline{x}+\underline{e}_2) - b_h(\underline{x})\right)}_{\text{vertical increment}}$$

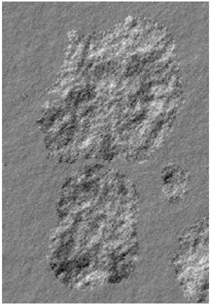
Real-world texture



Mask



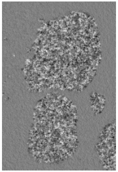
Synthetic texture



$\Omega_1 : (\sigma_1^2, h_1) \quad \Omega_2 : (\sigma_2^2, h_2)$

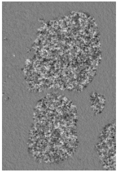
Multiscale analysis

Textured image

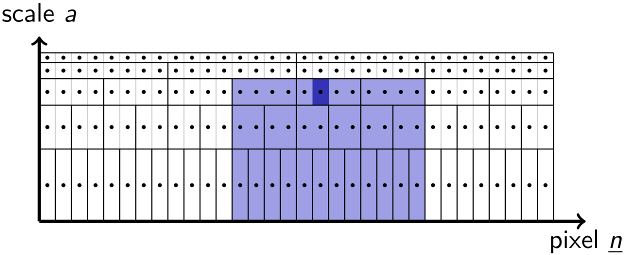


Multiscale analysis

Textured image

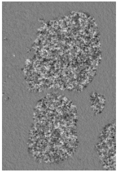


Wavelet coefficient local maximum: $\mathcal{L}_{a,\cdot}$



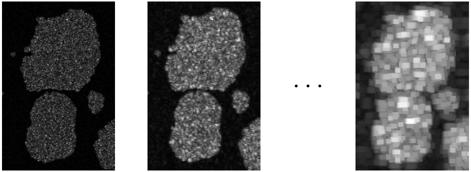
Multiscale analysis

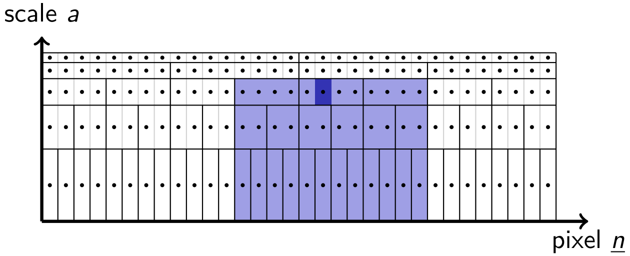
Textured image



Wavelet coefficient local maximum: $\mathcal{L}_{a,\cdot}$

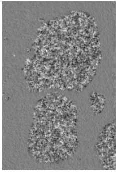
Scale $a = 2^1$ $a = 2^2$ … $a = 2^5$





Multiscale analysis

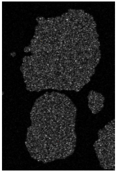
Textured image



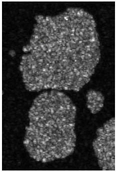
Wavelet coefficient local maximum: $\mathcal{L}_{a,\cdot}$

Scale

$a = 2^1$

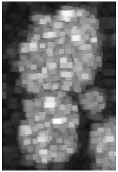


$a = 2^2$



$\ldots$

$a = 2^5$

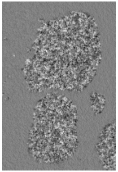


Proposition *(Jaffard, 2004), (Wendt, 2008)*

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

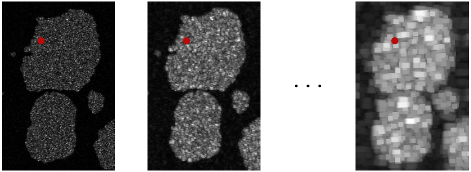
Multiscale analysis

Textured image



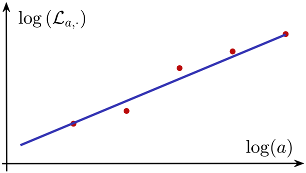
Wavelet coefficient local maximum: $\mathcal{L}_{a,\cdot}$

Scale $a = 2^1$ $a = 2^2$ $a = 2^5$



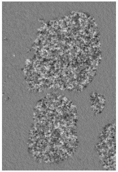
Proposition *(Jaffard, 2004), (Wendt, 2008)*

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



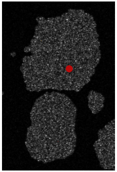
Multiscale analysis

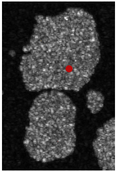
Textured image



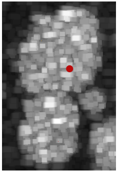
Wavelet coefficient local maximum: $\boldsymbol{\mathcal{L}}_{a,\cdot}$

Scale $a = 2^1$ $a = 2^2$ $a = 2^5$



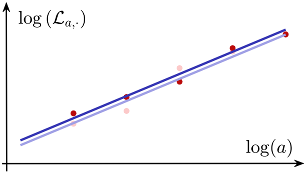


$\cdots$



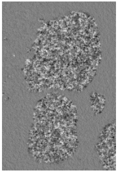
Proposition *(Jaffard, 2004), (Wendt, 2008)*

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



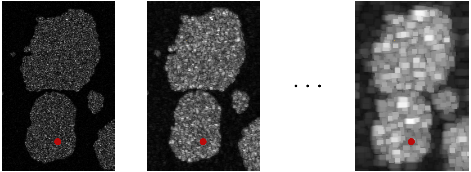
Multiscale analysis

Textured image



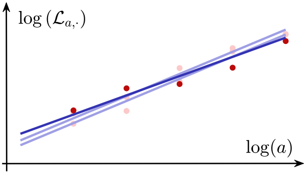
Wavelet coefficient local maximum: $\mathcal{L}_{a,\cdot}$

Scale $a = 2^1$ $a = 2^2$ $\cdots$ $a = 2^5$



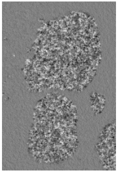
Proposition *(Jaffard, 2004), (Wendt, 2008)*

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



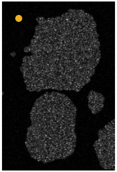
Multiscale analysis

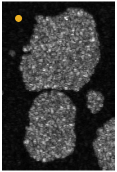
Textured image



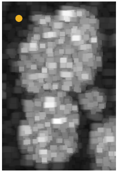
Wavelet coefficient local maximum: $\mathcal{L}_{a,\cdot}$

Scale $a = 2^1$ $a = 2^2$ $a = 2^5$



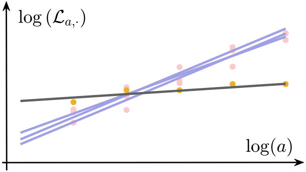


$\cdots$



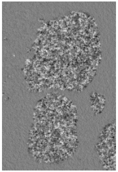
Proposition *(Jaffard, 2004), (Wendt, 2008)*

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



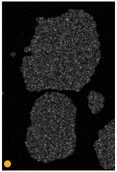
Multiscale analysis

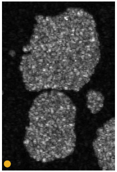
Textured image



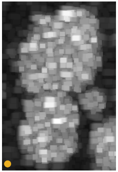
Wavelet coefficient local maximum: $\mathcal{L}_{a,\cdot}$

Scale $a = 2^1$ $a = 2^2$ $a = 2^5$



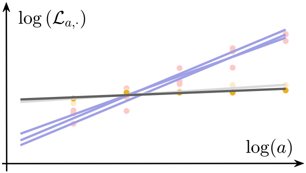


$\cdots$



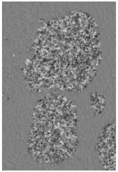
Proposition *(Jaffard, 2004), (Wendt, 2008)*

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



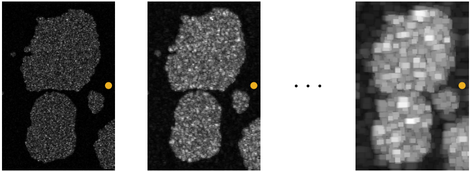
Multiscale analysis

Textured image



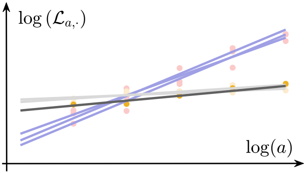
Wavelet coefficient local maximum: $\mathcal{L}_{a,\cdot}$

Scale $a = 2^1$ $a = 2^2$ … $a = 2^5$



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

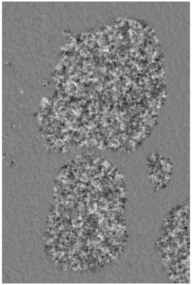
$$\log\left(\mathcal{L}_{a,\cdot}\right) \underset{a\to 0}{\simeq} \log(a) \underbrace{\boldsymbol{h}}_{\text{regularity}} + \underbrace{\boldsymbol{v}}_{\substack{\propto\log(\boldsymbol{\sigma}^2)\\ \text{(variance)}}}$$



Direct pixelwise estimation

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

Textured image

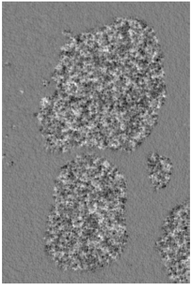


Direct pixelwise estimation

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

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

Textured image

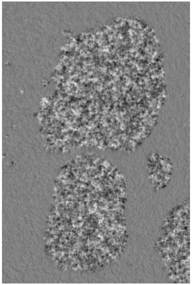


Direct pixelwise estimation

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

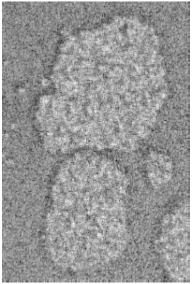
$$\left(\widehat{\boldsymbol{h}}^{\mathrm{LR}}, \widehat{\boldsymbol{v}}^{\mathrm{LR}}\right) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a=a_{\min}}^{a_{\max}} \|\log(\mathcal{L}_{a,\cdot}) - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2$$

Textured image



Local regularity locale $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$

Local power $\widehat{\boldsymbol{v}}^{\mathrm{LR}}$

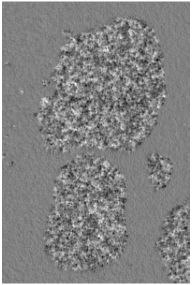


Direct pixelwise estimation

Linear regression $\underbrace{\mathbb{E}\log\left(\mathcal{L}_{a,\cdot}\right)}_{\text{expectation}} = \log(a)\underbrace{\bar{\boldsymbol{h}}}_{\text{regularity}} + \underbrace{\bar{\boldsymbol{v}}}_{\propto\log(\boldsymbol{\sigma}^2)}$

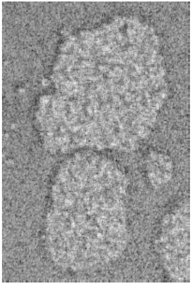
$$\left(\widehat{\boldsymbol{h}}^{\mathrm{LR}}, \widehat{\boldsymbol{v}}^{\mathrm{LR}}\right) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a=a_{\min}}^{a_{\max}} \left\|\log\left(\mathcal{L}_{a,\cdot}\right) - \log(a)\boldsymbol{h} - \boldsymbol{v}\right\|^2$$

Textured image



Local regularity locale $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$

Local power $\widehat{\boldsymbol{v}}^{\mathrm{LR}}$



$\longrightarrow$ high estimation variance

A posteriori regularization

Linear regression $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$

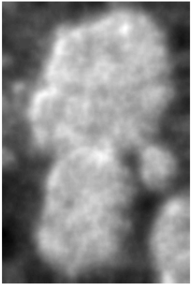
A posteriori regularization

Smoothing via filtering (linear)

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

Linear regression $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$

Smoothing



A posteriori regularization

Smoothing via filtering (linear)

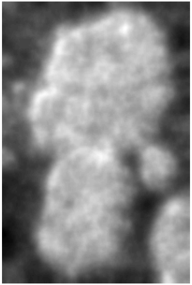
$$\left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \widehat{\boldsymbol{h}}^{\mathrm{LR}}$$

ROF denoising (nonlinear)

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

Linear regression $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$

Smoothing



ROF

A posteriori regularization

Smoothing via filtering (linear)

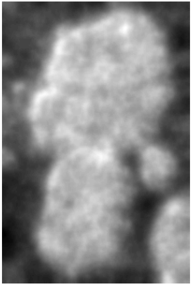
$$\left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \widehat{\boldsymbol{h}}^{\mathrm{LR}}$$

ROF denoising (nonlinear)

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

Linear regression $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$

Smoothing



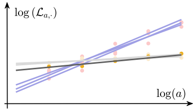
ROF

$\longrightarrow$ accumulation of estimation variance and regularization bias

Free contour and co-localized contour functionals

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

$\rightarrow$ fidelity to log-linear mdoel

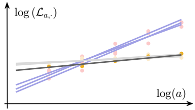


Free contour and co-localized contour functionals

$$\sum_a \underbrace{\|\log \mathcal{L}_{a,.} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} + \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$

$\rightarrow$ fidelity to log-linear mdoel

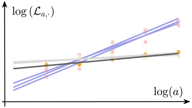
$\rightarrow$ favors piecewise constancy



Free contour and co-localized contour functionals

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \sum_{a} \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} \quad + \quad \lambda \, \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$

$\rightarrow$ fidelity to log-linear mdoel



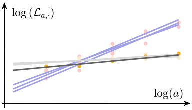
$\rightarrow$ favors piecewise constancy

Free contour and co-localized contour functionals

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \ \underbrace{\sum_a \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} + \underbrace{\lambda \ \mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$

→ fidelity to log-linear mdoel

→ favors piecewise constancy



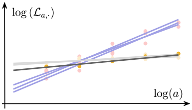
Finite differences $\mathbf{D}_1\boldsymbol{x}$ (horizontal), $\mathbf{D}_2\boldsymbol{x}$ (vertical) at each pixel

Free contour and co-localized contour functionals

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \sum_a \underbrace{\|\log \mathcal{L}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} + \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$

$\rightarrow$ fidelity to log-linear mdoel

$\rightarrow$ favors piecewise constancy



Finite differences $\mathbf{D}\boldsymbol{x} = [\mathbf{D}_1\boldsymbol{x}, \mathbf{D}_2\boldsymbol{x}]$

free: $\boldsymbol{h}$, $\boldsymbol{v}$ are **independently** piecewise constant

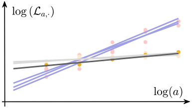
$$\mathcal{Q}_{\mathrm{L}}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha) = \alpha\|\mathbf{D}\boldsymbol{h}\|_{2,1} + \|\mathbf{D}\boldsymbol{v}\|_{2,1}$$

Free contour and co-localized contour functionals

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \; \underbrace{\sum_a \|\log \mathcal{L}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} + \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$

→ fidelity to log-linear mdoel

→ favors piecewise constancy



Finite differences $\mathbf{D}\boldsymbol{x} = [\mathbf{D}_1\boldsymbol{x}, \mathbf{D}_2\boldsymbol{x}]$

free: $\boldsymbol{h}$, $\boldsymbol{v}$ are **independently** piecewise constant

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

co-localized: $\boldsymbol{h}$, $\boldsymbol{v}$ are **concomitantly** piecewise constant

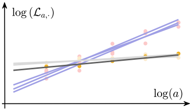
$$\mathcal{Q}_{\mathrm{C}}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha) = \|[\alpha\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}]\|_{2,1}$$

Free contour and co-localized contour functionals

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \underbrace{\sum_a \|\log \mathcal{L}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$

→ fidelity to log-linear model

→ favors piecewise constancy



Disjoint contours

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

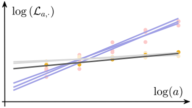
Common contours

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

Free contour and co-localized contour functionals

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \underbrace{\sum_a \|\log \mathcal{L}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} + \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$

Least squares
→ fidelity to log-linear model



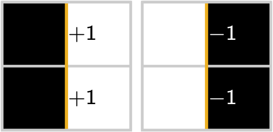
Total variation
→ favors piecewise constancy

Disjoint contours

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

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

Common contours



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

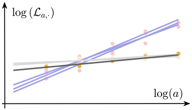
$\mathcal{Q}_L(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; 1) = 4$

Free contour and co-localized contour functionals

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \ \sum_a \underbrace{\|\log \mathcal{L}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} + \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$

→ fidelity to log-linear model

→ favors piecewise constancy



Disjoint contours

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

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

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

Common contours

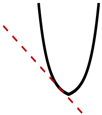
$\boldsymbol{h} \in \mathbb{R}^{2\times 2}$ $\quad$ $\boldsymbol{v} \in \mathbb{R}^{2\times 2}$

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

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

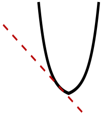
Functional minimization

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \;\; \sum_{a} \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,.} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} \;\; + \;\; \lambda \, \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$



Functional minimization

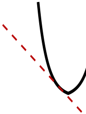
$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \ \sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,.} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} + \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$



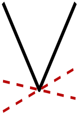
- gradient descent $\boldsymbol{x}^{n+1} = \boldsymbol{x}^n - \tau \nabla \varphi(\boldsymbol{x}^n)$

Functional minimization

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \ \sum_a \underbrace{\|\log \mathcal{L}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} \ + \ \lambda \ \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$



nonsmooth

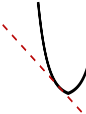


- gradient descent $\boldsymbol{x}^{n+1} = \boldsymbol{x}^n - \tau \nabla \varphi(\boldsymbol{x}^n)$
- implicit subgradient descent: proximal point algorithm

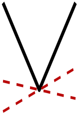
$$\boldsymbol{x}^{n+1} = \boldsymbol{x}^n - \boldsymbol{u}^n, \ \boldsymbol{u}^n \in \partial\varphi(\boldsymbol{x}^{n+1}) \ \Leftrightarrow \ \boldsymbol{x}^{n+1} = \mathrm{prox}_{\tau\varphi}(\boldsymbol{x}^n)$$

Functional minimization

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \; \sum_a \underbrace{\|\log \mathcal{L}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} \; + \; \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$



nonsmooth



- gradient descent $\boldsymbol{x}^{n+1} = \boldsymbol{x}^n - \tau \nabla \varphi(\boldsymbol{x}^n)$
- implicit subgradient descent: proximal point algorithm

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

- proximal splitting algorithm

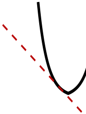
$$\boldsymbol{y}^{n+1} = \operatorname{prox}_{\sigma(\lambda\mathcal{Q})^*}\left(\boldsymbol{y}^n + \sigma \mathbf{D}\bar{\boldsymbol{x}}^n\right)$$

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

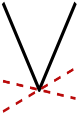
$$\bar{\boldsymbol{x}}^{n+1} = 2\boldsymbol{x}^{n+1} - \boldsymbol{x}^n$$

Functional minimization

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} \quad + \quad \lambda \; \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$



nonsmooth



- gradient descent $\boldsymbol{x}^{n+1} = \boldsymbol{x}^n - \tau \nabla \varphi(\boldsymbol{x}^n)$
- implicit subgradient descent: proximal point algorithm

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

- proximal splitting algorithm $\qquad \mathrm{prox}_{\tau\varphi}(\boldsymbol{x}) = \underset{\boldsymbol{u}}{\mathrm{argmin}} \, \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{u}\|^2 + \tau\varphi(\boldsymbol{u})$

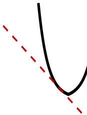
$$\boldsymbol{y}^{n+1} = \mathrm{prox}_{\sigma(\lambda\mathcal{Q})^*}\left(\boldsymbol{y}^n + \sigma \mathbf{D}\bar{\boldsymbol{x}}^n\right)$$

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

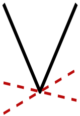
$$\bar{\boldsymbol{x}}^{n+1} = 2\boldsymbol{x}^{n+1} - \boldsymbol{x}^n$$

Computation of proximity operators

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \ \sum_{a} \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,.} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} \ + \ \lambda \ \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$



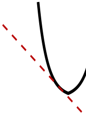
nonsmooth



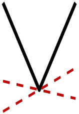
Least squares : $\|\log \boldsymbol{\mathcal{L}} - \Phi(\boldsymbol{h}, \boldsymbol{v})\|^2, \quad \Phi : (\boldsymbol{h}, \boldsymbol{v}) \mapsto \{\log(a)\boldsymbol{h} + \boldsymbol{v}\}_a$

Computation of proximity operators

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} + \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$



nonsmooth



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

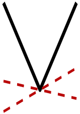
Proposition *(Pascal, 2019)*

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

Computation of proximity operators

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \; \sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} \; + \; \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$

nonsmooth



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

Proposition *(Pascal, 2019)*

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

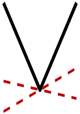
$$(\widetilde{\boldsymbol{h}}, \widetilde{\boldsymbol{v}}) = \mathrm{prox}_{\tau\|\boldsymbol{\mathcal{L}} - \boldsymbol{\Phi}\cdot\|^2}(\boldsymbol{h}, \boldsymbol{v}) \iff (\widetilde{\boldsymbol{h}}, \widetilde{\boldsymbol{v}}) = \left(\mathbf{I} + \tau \boldsymbol{\Phi}^\top \boldsymbol{\Phi}\right)^{-1}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

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

Strong convexity accelerated algorithm

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{minimize}} \;\; \sum_a \underset{\textbf{Least squares}}{\underline{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}} \;+\; \lambda\, \underset{\textbf{Total variation}}{\underline{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}}$$

nonsmooth

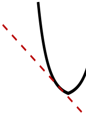


primal-dual algorithm *(Chambolle, 2011)*

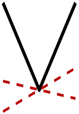
δ: duality gap, $\delta(\boldsymbol{x}^n, \boldsymbol{y}^n) \underset{n\to+\infty}{\longrightarrow} 0$

Strong convexity accelerated algorithm

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \ \sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} + \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$

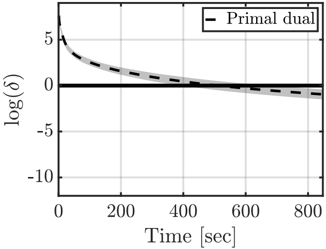


nonsmooth



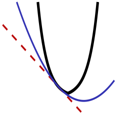
primal-dual algorithm *(Chambolle, 2011)*

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



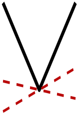
Convexity properties

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \; \sum_a \underset{\textbf{Least squares}}{\underline{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}} \; + \; \lambda \; \underset{\textbf{Total variation}}{\underline{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}}$$



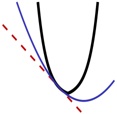
$\boldsymbol{\mu}$-strongly convex

nonsmooth



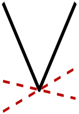
Convexity properties

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \; \sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} \; + \; \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$



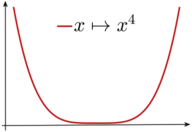
$\boldsymbol{\mu}$-strongly convex

nonsmooth

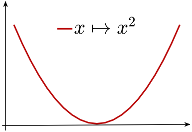


Forte convexité

- φ $\boldsymbol{\mu}$-strongly convex ssi $\varphi - \frac{\boldsymbol{\mu}}{2}\|\cdot\|^2$ convex



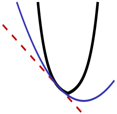
✓ strictly convex
✗ not strongly convex



✓ strictly convex
✓ 1-strongly convex

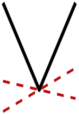
Convexity properties

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \; \underbrace{\sum_a \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} \; + \; \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$



$\boldsymbol{\mu}$-strongly convex

nonsmooth

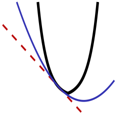


Forte convexité

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

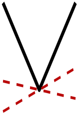
Convexity properties

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \; \sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} \; + \; \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$



μ-strongly convex

nonsmooth



Forte convexité

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

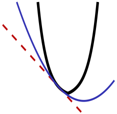
Proposition *(Pascal, 2019)*

$\sum_a \|\log \boldsymbol{\mathcal{L}} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2$ is μ-strongly convex.

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

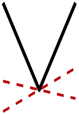
Strong convexity accelerated algorithm

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$



μ-strongly convex

nonsmooth



Accelerated primal-dual algorithm *(Chambolle, 2011)*

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

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

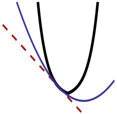
$$\boldsymbol{x}^{n+1} = \operatorname{prox}_{\tau_n\|\boldsymbol{\mathcal{L}} - \boldsymbol{\Phi}\cdot\|_2^2}\left(\boldsymbol{x}^n - \tau_n \mathbf{D}^\top \boldsymbol{y}^{n+1}\right)$$

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

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

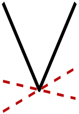
Strong convexity accelerated algorithm

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \; \sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} \; + \; \lambda \, \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$



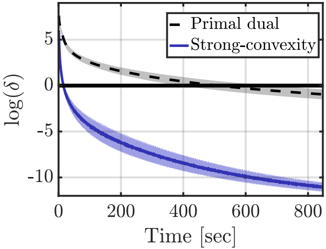
μ-strongly convex

nonsmooth



Accelerated primal-dual algorithm *(Chambolle, 2011)*

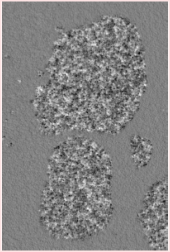
δ: duality gap, $\delta(\boldsymbol{x}^n, \boldsymbol{y}^n) \underset{n\to+\infty}{\longrightarrow} 0$



Segmentation by iterative thresholding

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \underbrace{\sum_a \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$

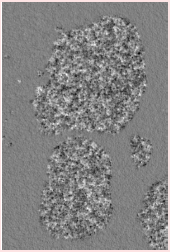
Textured image



Segmentation by iterative thresholding

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \; \sum_a \underbrace{\|\log \mathcal{L}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} \; + \; \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$

Textured image

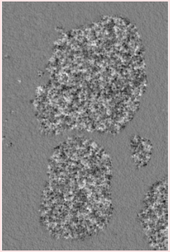


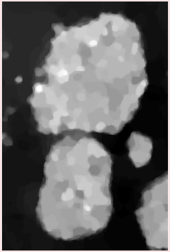
Lin. reg. $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$

Segmentation by iterative thresholding

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \; \sum_a \underbrace{\|\log \mathcal{L}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} \; + \; \lambda \, \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$

Textured image | Lin. reg. $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$ | Co-localized contours $\widehat{\boldsymbol{h}}^{\mathrm{C}}$

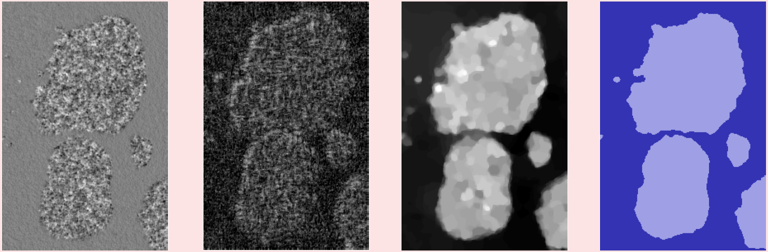




Segmentation by iterative thresholding

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \; \underbrace{\sum_a \|\log \mathcal{L}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} \; + \; \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$

Textured image | Lin. reg. $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$ | Co-localized contours $\widehat{\boldsymbol{h}}^{\mathrm{C}}$ | Thresholded estimate† $S\widehat{\boldsymbol{h}}^{\mathrm{C}}$



† *(Cai, 2013)*

State-of-the-art methods for texture segmentation

Threshold–ROF on $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$

(Nafornita, 2014), (Pustelnik, 2016)

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

Based only on local regularity $\boldsymbol{h}$.

State-of-the-art methods for texture segmentation

Threshold–ROF on $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$

(Nafornita, 2014), (Pustelnik, 2016)

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

Lin. reg.

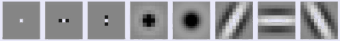
ROF

Threshold

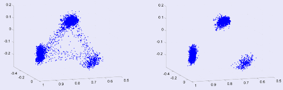
Based only on local regularity $\boldsymbol{h}$.

Matrix factorization based segmentation† *(Yuan, 2015)*

(i) local histograms



(ii) matrix factorization



(a) (b)

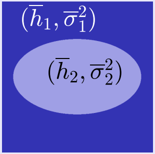
Fig. 2. Scatterplot of features in subspace. (a) Scatterplot of features projected onto the 3-d subspace. (b) Scatterplot after removing features with high edgeness.

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

Compared performances on synthetic textures

Piecewise monofractal texture synthesis *(Pascal, 2019)*

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

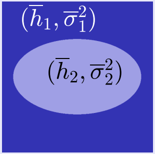


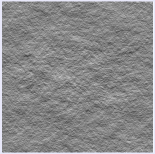
Compared performances on synthetic textures

Piecewise monofractal texture synthesis *(Pascal, 2019)*

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

E.g., $\bar{h}_1 = 0.5$, $\bar{\sigma}_1^2 = 0.6$
$\bar{h}_2 = 0.6$, $\bar{\sigma}_2^2 = 0.7$



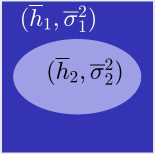


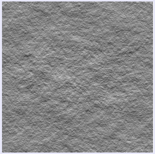
Compared performances on synthetic textures

Piecewise monofractal texture synthesis *(Pascal, 2019)*

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

E.g., $\overline{h}_1 = 0.5,\ \overline{\sigma}^2_1 = 0.6$
$\overline{h}_2 = 0.6,\ \overline{\sigma}^2_2 = 0.7$



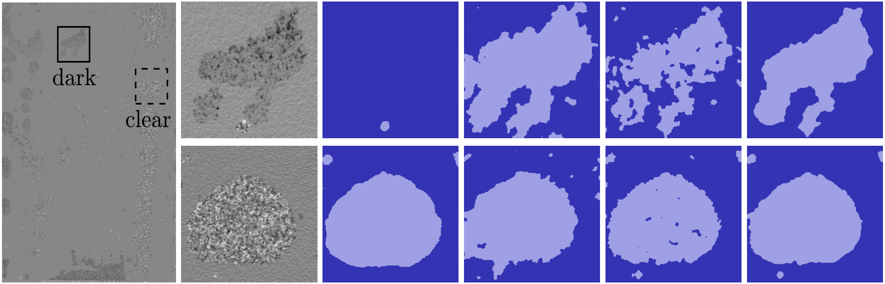


Averaged segmentation performance over 5 realizations

Yuan	T-ROF	Free	Co-localized
		contours	
			
$71.1 \pm 1.3\%$	$78.5 \pm 1.1\%$	$90.2 \pm 1.9\%$	$91.1 \pm 1.5\%$

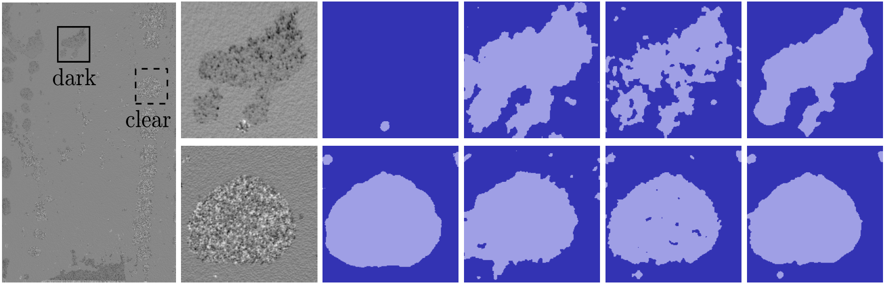
Low activity : $Q_\mathrm{G} = 300\mathrm{mL}/\mathrm{min}$ - $Q_\mathrm{L} = 300\mathrm{mL}/\mathrm{min}$

Flow | Zooms | Yuan | T-ROF | free contours | co-localized contours



Low activity : $Q_{\mathrm{G}} = 300\mathrm{mL}/\mathrm{min}$ - $Q_{\mathrm{L}} = 300\mathrm{mL}/\mathrm{min}$

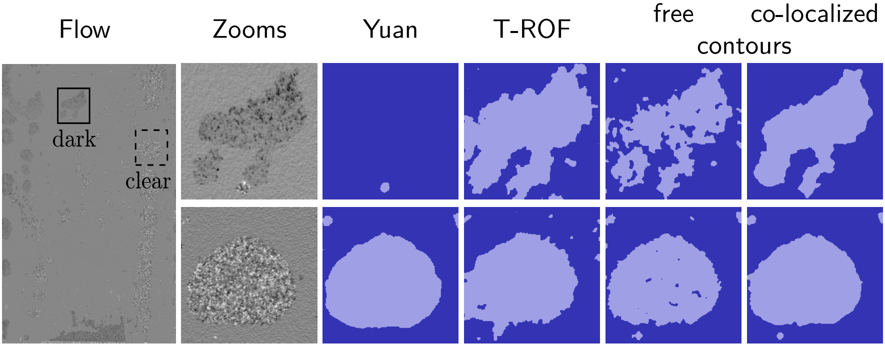
Flow Zooms Yuan T-ROF free contours co-localized



Liquid: $h_{\mathrm{L}} = 0.4$

Gas: $h_{\mathrm{G}} = 0.9$

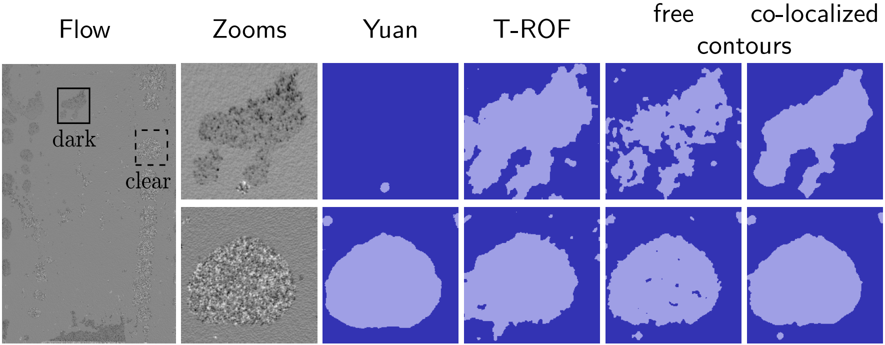
Low activity : $Q_{\mathrm{G}} = 300\text{mL}/\text{min}$ - $Q_{\mathrm{L}} = 300\text{mL}/\text{min}$



Liquid: $h_{\mathrm{L}} = 0.4$ $\sigma^2_{\mathrm{dark}} = 10^{-2}$

Gas: $h_{\mathrm{G}} = 0.9$

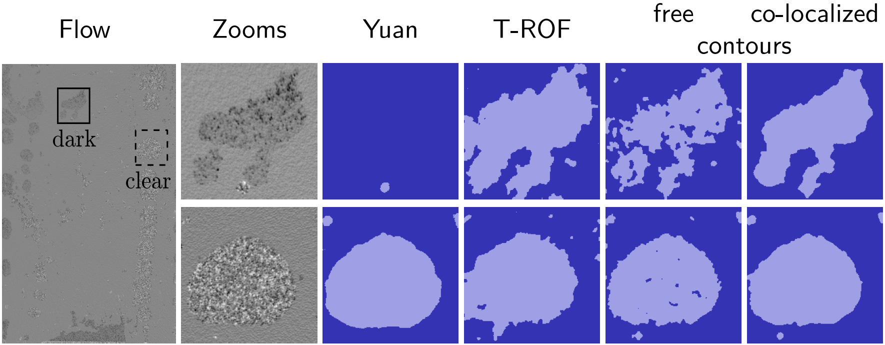
Low activity : $Q_\mathrm{G} = 300\mathrm{mL}/\mathrm{min}$ - $Q_\mathrm{L} = 300\mathrm{mL}/\mathrm{min}$



Liquid: $h_\mathrm{L} = 0.4$ $\sigma^2_\mathrm{dark} = 10^{-2}$

Gas: $h_\mathrm{G} = 0.9$

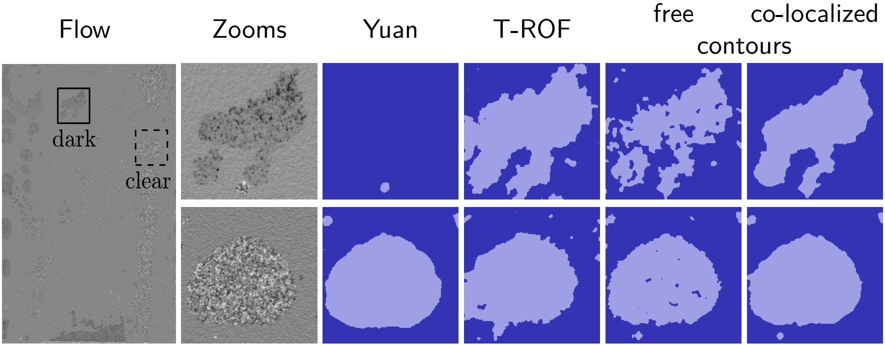
Low activity : $Q_{\mathrm{G}} = 300\text{mL}/\text{min}$ - $Q_{\mathrm{L}} = 300\text{mL}/\text{min}$



Liquid: $h_{\mathrm{L}} = 0.4$ $\sigma^2_{\text{dark}} = 10^{-2}$

Gas: $h_{\mathrm{G}} = 0.9$ | $\sigma^2_{\text{dark}} = 10^{-2}$ (dark bubbles)

Low activity : $Q_{\mathrm{G}} = 300\mathrm{mL}/\mathrm{min}$ - $Q_{\mathrm{L}} = 300\mathrm{mL}/\mathrm{min}$

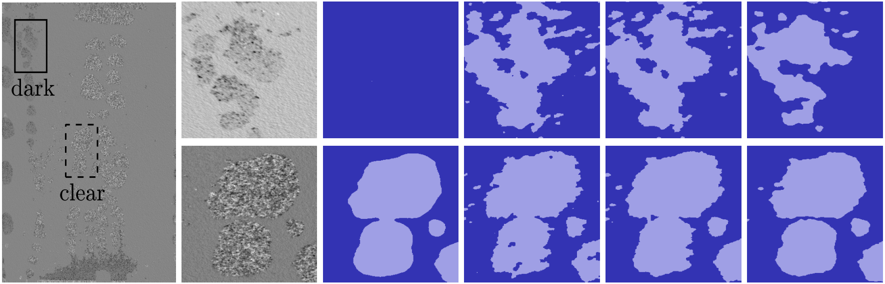


Liquid: $h_{\mathrm{L}} = 0.4$ $\quad \sigma^2_{\mathrm{dark}} = 10^{-2}$

Gas: $h_{\mathrm{G}} = 0.9$ $\quad \sigma^2_{\mathrm{dark}} = 10^{-2}$ (dark bubbles)
$\quad \sigma^2_{\mathrm{clear}} = 10^{-1}$ (clear bubbles)

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

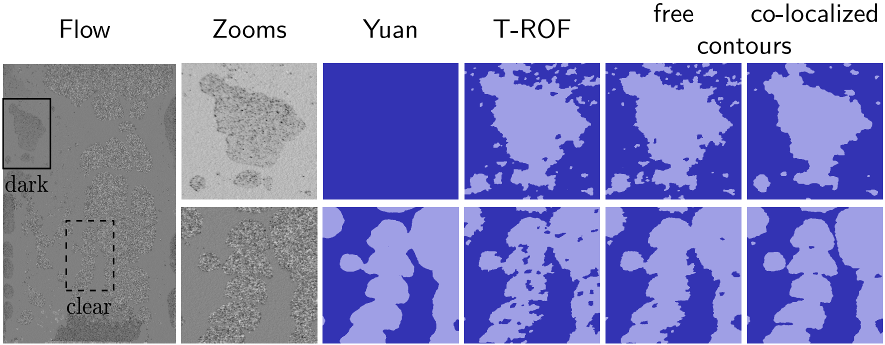
Flow | Zooms | Yuan | T-ROF | free contours | co-localized contours



Liquid: $h_{\mathrm{L}} = 0.4$ $\sigma^2_{\mathrm{dark}} = 10^{-2}$

Gas: $h_{\mathrm{G}} = 0.9$ $\sigma^2_{\mathrm{dark}} = 10^{-2}$ (dark bubbles)
$\sigma^2_{\mathrm{clear}} = 10^{-1}$ (clear bubbles).

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

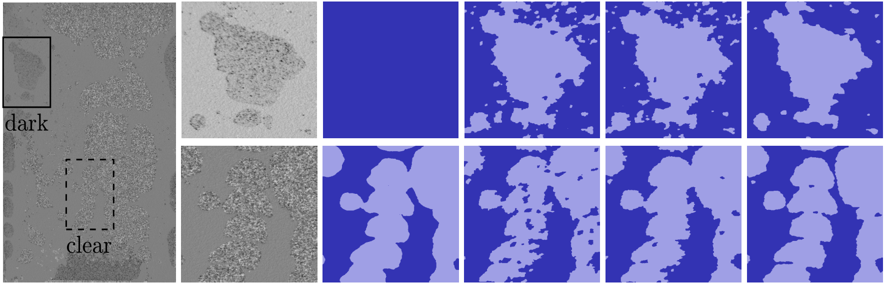


Liquid: $h_{\mathrm{L}} = 0.4$ $\quad \sigma^2_{\mathrm{dark}} = 10^{-2}$

Gas: $h_{\mathrm{G}} = 0.9$ $\quad \Big| \quad \sigma^2_{\mathrm{dark}} = 10^{-2}$ (dark bubbles)
$\sigma^2_{\mathrm{clear}} = 10^{-1}$ (clear bubbles).

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

Flow	Zooms	Yuan	T-ROF	free contours	co-localized contours



Computational time		1s	12s	700s	**2100s**

Liquid: $h_{\mathrm{L}} = 0.4$ $\sigma^2_{\mathrm{dark}} = 10^{-2}$

Gas: $h_{\mathrm{G}} = 0.9$ | $\sigma^2_{\mathrm{dark}} = 10^{-2}$ (dark bubbles)
$\sigma^2_{\mathrm{clear}} = 10^{-1}$ (clear bubbles).

Selection of regularization parameters

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

Selection of regularization parameters

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

Lin. reg. $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$

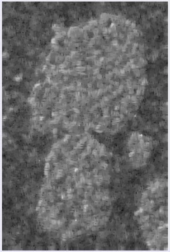
$(\lambda; \alpha) = (0; 0)$

Selection of regularization parameters

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

Lin. reg. $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$ Co-localized contour estimate $\widehat{\boldsymbol{h}}^{\mathrm{C}}$

$(\lambda; \alpha) = (0; 0)$ $(\lambda, \alpha) = (0.5; 0.5)$



too small

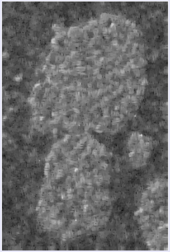
Selection of regularization parameters

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

Lin. reg. $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$ | Co-localized contour estimate $\widehat{\boldsymbol{h}}^{\mathrm{C}}$

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

$(\lambda, \alpha) = (0.5; 0.5)$



too small

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

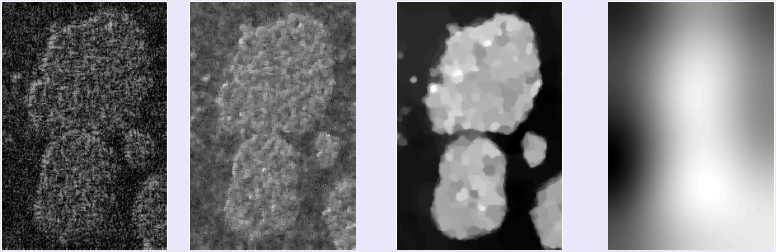
too large

Selection of regularization parameters

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

Lin. reg. $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$ | Co-localized contour estimate $\widehat{\boldsymbol{h}}^{\mathrm{C}}$

$(\lambda; \alpha) = (0; 0)$ | $(\lambda, \alpha) = (0.5; 0.5)$ | $(\lambda^{\dagger}, \alpha^{\dagger}) = (11.5; 0.8)$ | $(\lambda; \alpha) = (500; 500)$



too small | optimal | too large

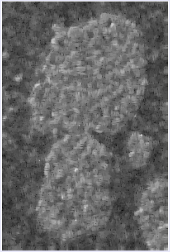
Selection of regularization parameters

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

Lin. reg. $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$ Co-localized contour estimate $\widehat{\boldsymbol{h}}^{\mathrm{C}}$

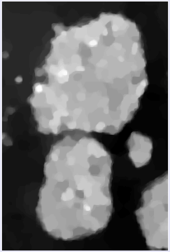
$(\lambda; \alpha) = (0; 0)$

$(\lambda, \alpha) = (0.5; 0.5)$



too small

$(\lambda^\dagger, \alpha^\dagger) = (11.5; 0.8)$



optimal

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

too large

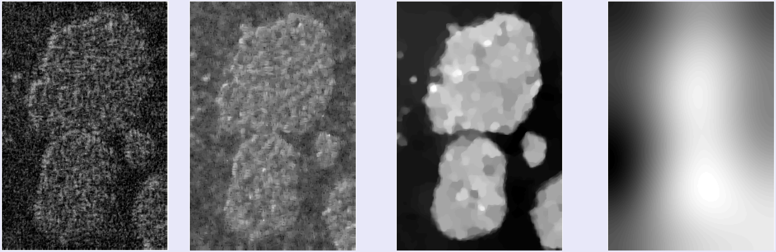
What *optimal* means?

Selection of regularization parameters

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

Lin. reg. $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$ Co-localized contour estimate $\widehat{\boldsymbol{h}}^{\mathrm{C}}$

$(\lambda;\alpha) = (0;0)$ $(\lambda,\alpha) = (0.5;0.5)$ $(\lambda^\dagger,\alpha^\dagger) = (11.5;0.8)$ $(\lambda;\alpha) = (500;500)$



too small optimal too large

What *optimal* means? How to find $\lambda^\dagger$ and $\alpha^\dagger$?

Parameter tuning (systematic search)

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

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

Parameter tuning (systematic search)

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

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

$\bar{\boldsymbol{h}}$: *true* regularity

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

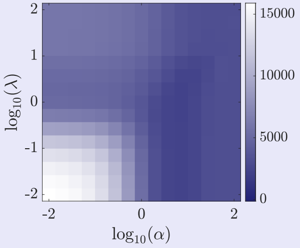
Parameter tuning (systematic search)

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

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

$\bar{\boldsymbol{h}}$: *true* regularity

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



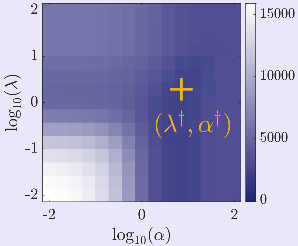
Parameter tuning (systematic search)

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

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

$\bar{\boldsymbol{h}}$: *true* regularity

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



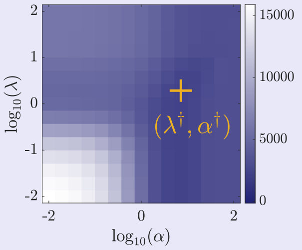
Parameter tuning (systematic search)

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

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

$\bar{\boldsymbol{h}}$: *true* regularity

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



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

?

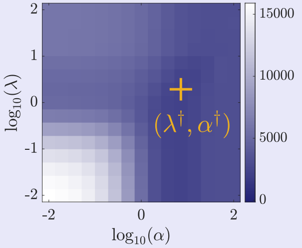
Parameter tuning (systematic search)

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

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

$\bar{\boldsymbol{h}}$: *true* regularity

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



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

?

Stein Unbiased Risk Estimate
(SURE)

Stein Unbiased Risk Estimate (Principle)

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

Stein Unbiased Risk Estimate (Principle)

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

Parametric estimator $(\boldsymbol{y}; \lambda) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)$

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

Stein Unbiased Risk Estimate (Principle)

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

Parametric estimator $(\boldsymbol{y}; \lambda) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)$

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

Quadratic error $R(\lambda) \triangleq \mathbb{E}_{\boldsymbol{\zeta}} \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \bar{\boldsymbol{x}}\|^2 \stackrel{?}{=} \mathbb{E}_{\boldsymbol{\zeta}} \widehat{R}(\boldsymbol{y}; \lambda)$ $\bar{\boldsymbol{x}}$ unknown

Stein Unbiased Risk Estimate (Principle)

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

Parametric estimator $(\boldsymbol{y}; \lambda) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)$

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

Quadratic error $R(\lambda) \triangleq \mathbb{E}_{\boldsymbol{\zeta}} \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \bar{\boldsymbol{x}}\|^2 \stackrel{?}{=} \mathbb{E}_{\boldsymbol{\zeta}} \widehat{R}(\boldsymbol{y}; \lambda)$ $\bar{\boldsymbol{x}}$ unknown

Theorem *(Stein, 1981)*

Let $(\boldsymbol{y}; \lambda) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)$ be an estimator of $\bar{\boldsymbol{x}}$

- weakly differentiable with respect to $\boldsymbol{y}$,
- such that $\boldsymbol{\zeta} \mapsto \langle \widehat{\boldsymbol{x}}(\bar{\boldsymbol{x}} + \boldsymbol{\zeta}; \lambda), \boldsymbol{\zeta} \rangle$ is integrable w. r. t. $\mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$.

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

Stein Unbiased Risk Estimate generalized

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

Stein Unbiased Risk Estimate generalized

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

E.g., free and co-localized contour estimators $\widehat{\boldsymbol{h}}(\mathcal{L}; \lambda, \alpha)$

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

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

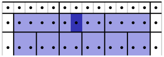
Stein Unbiased Risk Estimate generalized

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

E.g., free and co-localized contour estimators $\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha)$

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

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



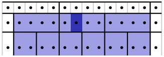
Stein Unbiased Risk Estimate generalized

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

E.g., free and co-localized contour estimators $\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha)$

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

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



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

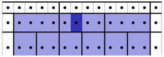
Stein Unbiased Risk Estimate generalized

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

E.g., free and co-localized contour estimators $\widehat{\boldsymbol{h}}(\mathcal{L}; \lambda, \alpha)$

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

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



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

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

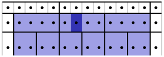
Stein Unbiased Risk Estimate generalized

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

E.g., free and co-localized contour estimators $\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha)$

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

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



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

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

Theorem *(Pascal, 2020)*

Let $\mathbf{A} \triangleq \boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top$ and $(\boldsymbol{y}; \boldsymbol{\Lambda}) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda})$ an estimator of $\bar{\boldsymbol{x}}$

Stein Unbiased Risk Estimate generalized

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

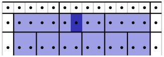
E.g., free and co-localized contour estimators $\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha)$

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

$\boldsymbol{\zeta} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\mathcal{S}})$

$\mathcal{R} = \|\widehat{\boldsymbol{h}} - \bar{\boldsymbol{h}}\|^2$

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



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

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

Theorem *(Pascal, 2020)*

Let $\mathbf{A} \triangleq \boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top$ and $(\boldsymbol{y}; \boldsymbol{\Lambda}) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda})$ an estimator of $\bar{\boldsymbol{x}}$

- weakly differentiable with respect to $\boldsymbol{y}$,
- such that $\boldsymbol{\zeta} \mapsto \langle \boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\bar{\boldsymbol{x}} + \boldsymbol{\zeta}; \lambda), \mathbf{A}\boldsymbol{\zeta}\rangle$ is integrable w.r.t. $\mathcal{N}(\mathbf{0}, \boldsymbol{\mathcal{S}})$.

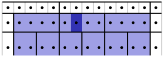
Stein Unbiased Risk Estimate generalized

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

E.g., free and co-localized contour estimators $\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha)$

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

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



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

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

Theorem *(Pascal, 2020)*

Let $\mathbf{A} \triangleq \boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top$ and $(\boldsymbol{y}; \boldsymbol{\Lambda}) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda})$ an estimator of $\bar{\boldsymbol{x}}$

- weakly differentiable with respect to $\boldsymbol{y}$,
- such that $\boldsymbol{\zeta} \mapsto \langle \boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\bar{\boldsymbol{x}} + \boldsymbol{\zeta}; \lambda), \mathbf{A}\boldsymbol{\zeta}\rangle$ is integrable w.r.t. $\mathcal{N}(\mathbf{0}, \boldsymbol{\mathcal{S}})$.

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

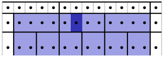
Stein Unbiased Risk Estimate generalized

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

E.g., free and co-localized contour estimators $\widehat{\boldsymbol{h}}(\mathcal{L}; \lambda, \alpha)$

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

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



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

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

Theorem *(Pascal, 2020)*

Let $\mathbf{A} \triangleq \boldsymbol{\Pi}(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top$ and $(\boldsymbol{y}; \boldsymbol{\Lambda}) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda})$ an estimator of $\bar{\boldsymbol{x}}$

- weakly differentiable with respect to $\boldsymbol{y}$,
- such that $\zeta \mapsto \langle \boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\bar{\boldsymbol{x}} + \zeta; \lambda), \mathbf{A}\zeta \rangle$ is integrable w.r.t. $\mathcal{N}(\mathbf{0}, \mathcal{S})$.

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

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

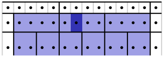
Stein Unbiased Risk Estimate generalized

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

E.g., free and co-localized contour estimators $\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha)$

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

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



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

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

Theorem *(Pascal, 2020)*

Let $\mathbf{A} \triangleq \boldsymbol{\Pi}(\boldsymbol{\Phi}^\top \boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\top$ and $(\boldsymbol{y}; \boldsymbol{\Lambda}) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda})$ an estimator of $\bar{\boldsymbol{x}}$

- weakly differentiable with respect to $\boldsymbol{y}$,
- such that $\boldsymbol{\zeta} \mapsto \langle \boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\bar{\boldsymbol{x}} + \boldsymbol{\zeta}; \lambda), \mathbf{A}\boldsymbol{\zeta} \rangle$ is integrable w.r.t. $\mathcal{N}(\boldsymbol{0}, \boldsymbol{\mathcal{S}})$.

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

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

Stein Unbiased Risk Estimate (Calcul)

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

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

Generalized Finite Differences Monte Carlo SURE

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

Stein Unbiased Risk Estimate (Calcul)

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

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

Generalized Finite Differences Monte Carlo SURE

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

Theorem *(Pascal, 2020)*

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

- uniformly Lipschitz w.r.t. $\boldsymbol{y}$,
- such that $\forall \boldsymbol{\Lambda} \in \mathbb{R}^L$, $\widehat{\boldsymbol{x}}(\mathbf{0}_P;\boldsymbol{\Lambda}) = \mathbf{0}_N$. Then

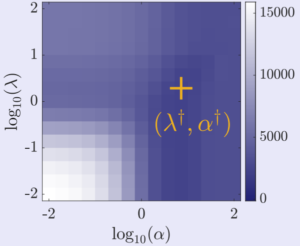
$$R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) = \lim_{\nu\to 0} \mathbb{E}_{\boldsymbol{\zeta},\boldsymbol{\varepsilon}}\left[\widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{y};\boldsymbol{\Lambda}\,|\,\boldsymbol{\mathcal{S}})\right]$$

Parameter tuning (Systematic search)

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

$\bar{\boldsymbol{h}}$: *true* regularity

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



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

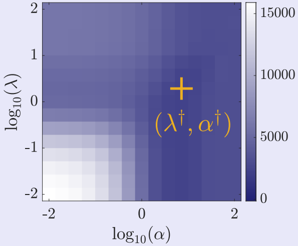
$$\widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha | \boldsymbol{\mathcal{S}})$$

Parameter tuning (Systematic search)

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

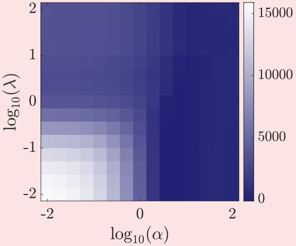
$\bar{\boldsymbol{h}}$: *true* regularity

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



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

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

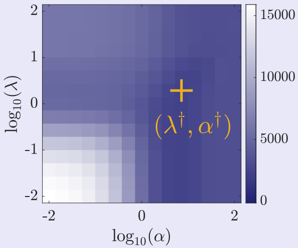


Parameter tuning (Systematic search)

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

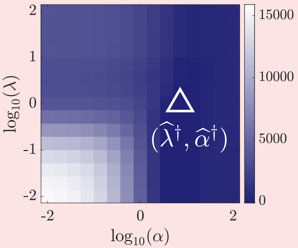
$\bar{\boldsymbol{h}}$: *true* regularity

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



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

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

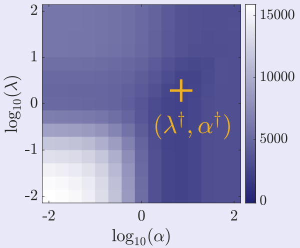


Parameter tuning (Systematic search)

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

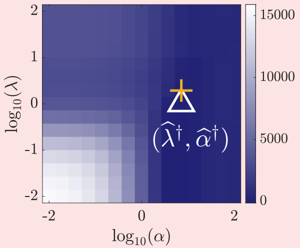
$\bar{\boldsymbol{h}}$: *true* regularity

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



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

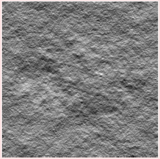
$$\widehat{R}_{\nu,\varepsilon}(\mathcal{L}; \lambda, \alpha | \mathcal{S})$$

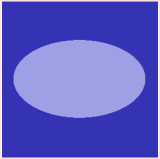


Systematic search of regularization parameters

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

Example

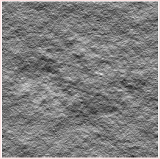


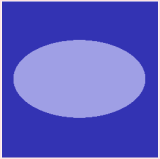


Systematic search of regularization parameters

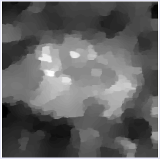
$$\left(\widehat{\boldsymbol{h}}^{\mathsf{L}}, \widehat{\boldsymbol{v}}^{\mathsf{L}}\right)(\boldsymbol{\mathcal{L}}; \lambda, \alpha) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_a \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda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

Example





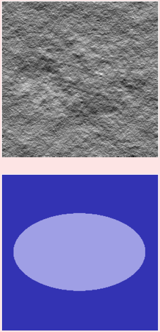
$\widehat{\boldsymbol{h}}^{\mathsf{L}}(\boldsymbol{\mathcal{L}}; \lambda^\dagger, \alpha^\dagger)$
(grid)



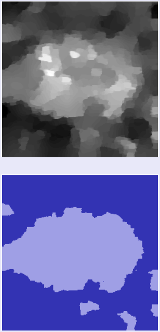
Systematic search of regularization parameters

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

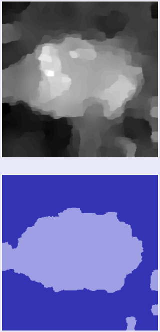
Example



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

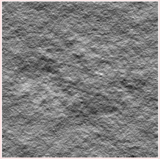
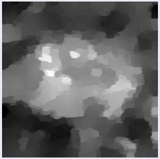
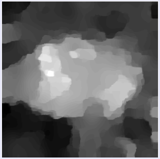
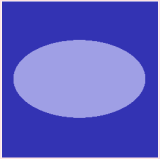


$\widehat{\boldsymbol{h}}^{\mathsf{L}}(\boldsymbol{\mathcal{L}}; \widehat{\lambda}^\dagger, \widehat{\alpha}^\dagger)$ (grid)



Systematic search of regularization parameters

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

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

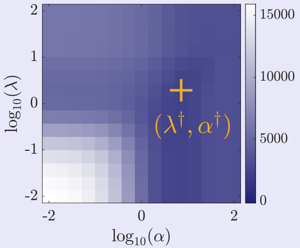
$15 \times 15 = \mathbf{225}$ parameters $\Rightarrow$ grid search is very costly!

Parameter tuning (Automated search)

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

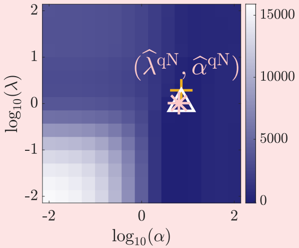
$\bar{\boldsymbol{h}}$: *true* regularity

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



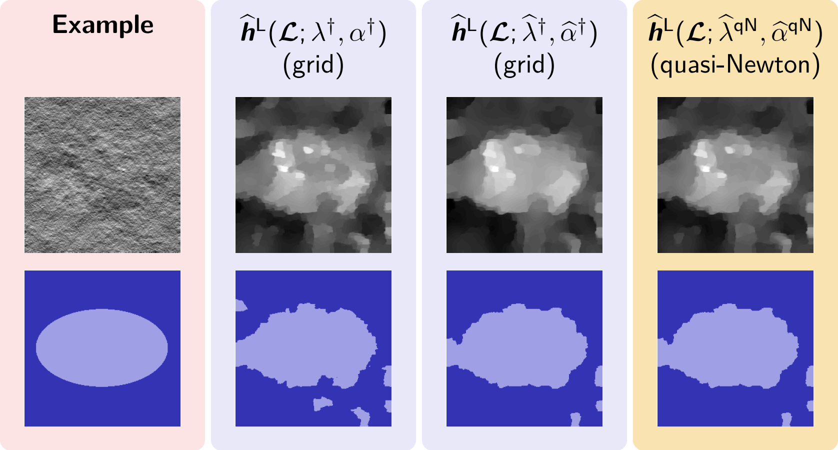
$\bar{\boldsymbol{h}}$: unknown!

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



Automated selection of regularization parameters

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



40 calls of the estimator v.s. **225** over a grid

Conclusion

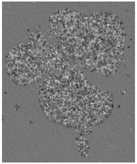
- Local regularity and local variance
  - able to characterize real-world textures
  - complementary attributes $\rightarrow$ ability to discriminate finely

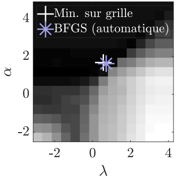
Conclusion

- Local regularity and local variance
  - able to characterize real-world textures
  - complementary attributes $\rightarrow$ ability to discriminate finely
- Simultaneous estimation and regularization
  - significant reduction of the estimation error
  - accurate and regular contours thanks to the co-localized penalization

Conclusion

- Local regularity and local variance
  - able to characterize real-world textures
  - complementary attributes → ability to discriminate finely
- Simultaneous estimation and regularization
  - significant reduction of the estimation error
  - accurate and regular contours thanks to the co-localized penalization
- Fast and automated algorithms
  - capacity to handle large amount of data
  - objectivity and reproducibility



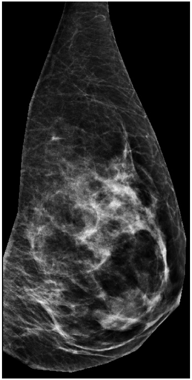


Perspectives: application to anticipate breast cancer

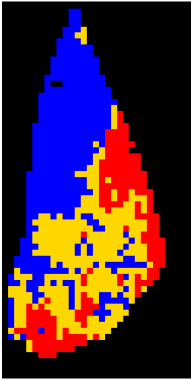
Observation: local regularity reflects health of breast tissues

(Marin et al., *Medical Physics*, 2017)

Mammogram



Local regularity (WTMM)



Microenvironment analysis via density fluctuations

Healthy $\Leftrightarrow$ structured

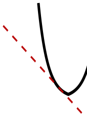
- $h < 0.5$: negative correlations
- $h > 0.5$: positive correlations

Prone to cancer $\Leftrightarrow$ unstructured

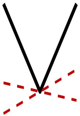
- $h = 0.5$: absence of correlations

Computation of proximity operators

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \sum_{a} \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,.} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} \quad + \quad \lambda \, \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$

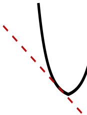


nonsmooth

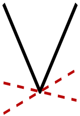


Computation of proximity operators

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \; \sum_{a} \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} \; + \; \lambda \, \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$



nonsmooth

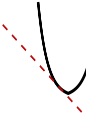


E.g., Mixed norm: for $\boldsymbol{z} = [\boldsymbol{z}_1; \ldots, ; \boldsymbol{z}_I]$

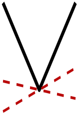
$$\mathcal{Q}(\boldsymbol{z}) = \|\boldsymbol{z}\|_{2,1} = \sum_{\underline{n}\in\Omega} \sqrt{\sum_{i=1}^{I} z_i^2(\underline{n})} = \sum_{\underline{n}\in\Omega} \|\boldsymbol{z}(\underline{n})\|_2$$

Computation of proximity operators

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \ \sum_a \underbrace{\|\log \mathcal{L}_{a,.} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least squares}} \ + \ \lambda \ \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total variation}}$$



nonsmooth



E.g., **Mixed norm:** for $\boldsymbol{z} = [\boldsymbol{z}_1; \ldots; \boldsymbol{z}_I]$

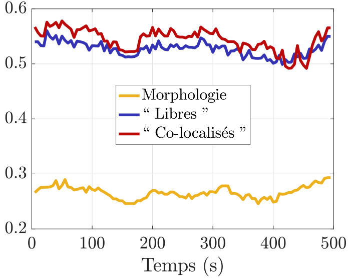
$$\mathcal{Q}(\boldsymbol{z}) = \|\boldsymbol{z}\|_{2,1} = \sum_{\underline{n}\in\Omega} \sqrt{\sum_{i=1}^{I} z_i^2(\underline{n})} = \sum_{\underline{n}\in\Omega} \|\boldsymbol{z}(\underline{n})\|_2$$

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

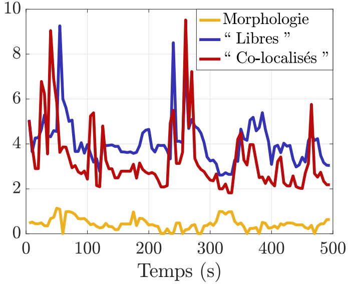
Écoulement multiphasiques en milieu poreux

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

Fraction de gaz dans la cellule



Périmètre d'interface



Computation of degree of freedom

Degrees of freedom $\qquad \mathrm{dof} \triangleq \mathrm{tr}\left(\boldsymbol{\mathcal{S}}\mathbf{A}^\top\boldsymbol{\Pi}\partial_{\boldsymbol{y}}\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda})\right)$

Computation of degree of freedom

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

- **Monte Carlo strategy** (MC) **large size $\mathbf{M} \in \mathbb{R}^{P\times P}$**

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

Computation of degree of freedom

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

- **Monte Carlo strategy** (MC) **large size M** $\in \mathbb{R}^{P\times P}$

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

- **Finite Differences** (FD) **inaccessible Jacobian**

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

Computation of degree of freedom

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

- **Monte Carlo strategy** (MC) **large size $\mathbf{M} \in \mathbb{R}^{P\times P}$**

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

- **Finite Differences** (FD) **inaccessible Jacobian**

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

Proposition *(Pascal, 2020)*

Let $(\boldsymbol{y}; \boldsymbol{\Lambda}) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda})$ an estimator of $\bar{\boldsymbol{x}}$

- uniformly Lipschitz with respect to $\boldsymbol{y}$,
- such that $\forall \boldsymbol{\Lambda} \in \mathbb{R}^L$, $\widehat{\boldsymbol{x}}(\mathbf{0}_P; \boldsymbol{\Lambda}) = \mathbf{0}_N$. Then

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

Automated minimization of SURE

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

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

Automated minimization of SURE

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

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

Goal : $\underset{\boldsymbol{\Lambda}}{\text{minimize}}\ \widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{y}; \boldsymbol{\Lambda} \mid \boldsymbol{\mathcal{S}})$ for given $\boldsymbol{y}$, $\boldsymbol{\mathcal{S}}$

Automated minimization of SURE

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

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

Goal : $\operatorname*{minimize}_{\boldsymbol{\Lambda}} \ \widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{y}; \boldsymbol{\Lambda} \,|\, \boldsymbol{\mathcal{S}}) \equiv \widehat{R}(\boldsymbol{\Lambda})$ for given $\boldsymbol{y}$, $\boldsymbol{\mathcal{S}}$

Automated minimization of SURE

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

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

Goal : $\underset{\boldsymbol{\Lambda}}{\mathrm{minimize}}\ \widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{y}; \boldsymbol{\Lambda} \,|\, \boldsymbol{\mathcal{S}}) \equiv \widehat{R}(\boldsymbol{\Lambda})$ for given $\boldsymbol{y}$, $\boldsymbol{\mathcal{S}}$

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

for $t = 0, 1, \dots$

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

$\alpha^{[t]} \in \underset{\alpha\in\mathbb{R}}{\mathrm{Argmin}}\ \widehat{R}(\boldsymbol{\Lambda}^{[t]} + \alpha \boldsymbol{d}^{[t]})$ *line search*

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

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

$\boldsymbol{H}^{[t+1]} = \mathrm{BFGS}(\boldsymbol{H}^{[t]}, \boldsymbol{d}^{[t]}, \boldsymbol{u}^{[t]})$ *inverse Hessian update*

Automated minimization of SURE

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

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

Goal : $\underset{\boldsymbol{\Lambda}}{\mathrm{minimize}}\ \widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{y}; \boldsymbol{\Lambda} \,|\, \boldsymbol{\mathcal{S}}) \equiv \widehat{R}(\boldsymbol{\Lambda})$ for given $\boldsymbol{y}$, $\boldsymbol{\mathcal{S}}$

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

for $t = 0, 1, \ldots$

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

$\alpha^{[t]} \in \underset{\alpha \in \mathbb{R}}{\mathrm{Argmin}}\ \widehat{R}(\boldsymbol{\Lambda}^{[t]} + \alpha \boldsymbol{d}^{[t]})$ *line search*

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

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

$\boldsymbol{H}^{[t+1]} = \mathrm{BFGS}(\boldsymbol{H}^{[t]}, \boldsymbol{d}^{[t]}, \boldsymbol{u}^{[t]})$ *inverse Hessian update*

Stein Unbiased GrAdient Risk estimate

Generalized FDMC SURE

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

Stein Unbiased GrAdient Risk estimate

Generalized FDMC SURE

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

Generalized Finite Difference Monte Carlo SUGAR

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

Stein Unbiased GrAdient Risk estimate

Generalized FDMC SURE

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

Generalized Finite Difference Monte Carlo SUGAR

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

Theorem *(Pascal, 2020)*

Let $(\boldsymbol{y};\boldsymbol{\Lambda}) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\Lambda})$ be an estimator of $\bar{\boldsymbol{x}}$

- uniformly Lipschitz with respect to $\boldsymbol{y}$
- such that $\forall \boldsymbol{\Lambda} \in \mathbb{R}^L$, $\widehat{\boldsymbol{x}}(\mathbf{0}_P;\boldsymbol{\Lambda}) = \mathbf{0}_N$,
- uniformly L-Lipschitz with respect to $\boldsymbol{\Lambda}$, L indep. of $\boldsymbol{y}$. Then

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