Processing nonstationary data: representations, theory, algorithms and applications.

Barbara Pascal

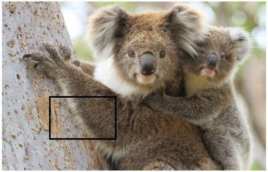
December 16$^{\text{th}}$ 2021

Laboratoire Grenoble Images Parole Signal Automatique (GIPSA-lab)

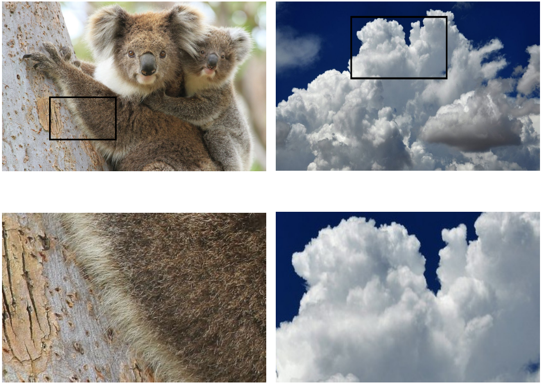
Pôle Géométries, Apprentissage, Information et Algorithmes (GAIA)

Part I: Texture segmentation based on fractal attributes

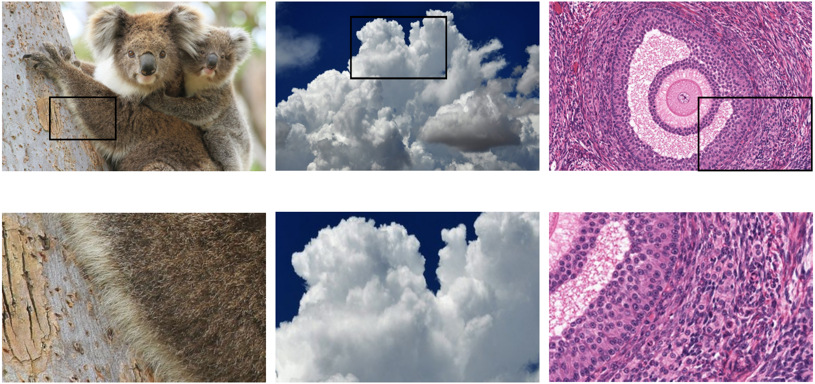
Textures



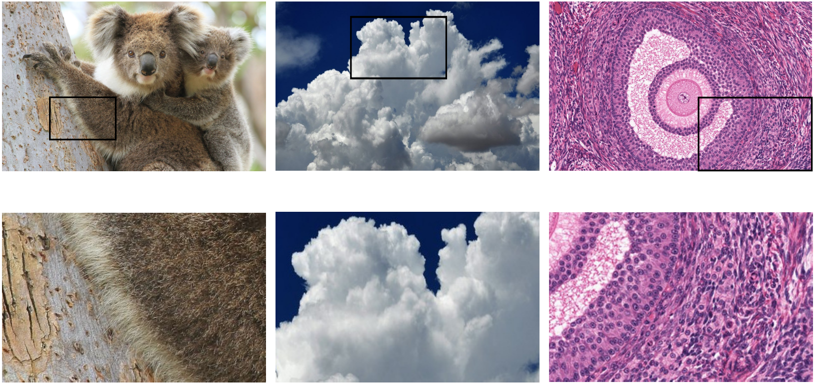
Textures



Textures

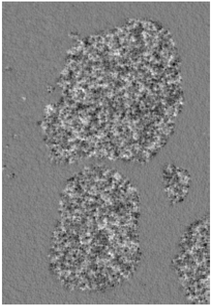


Textures

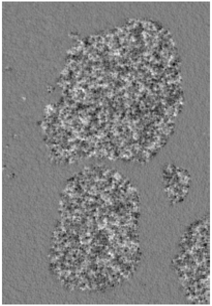


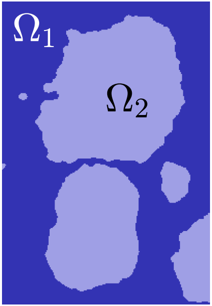
Crucial to describe real-world images

Textured image segmentation



Textured image segmentation

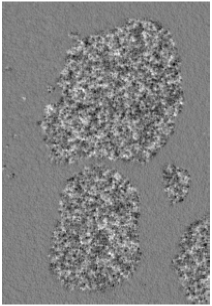


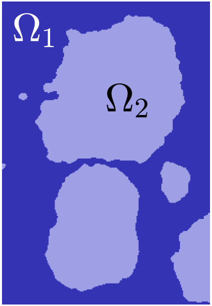


Goal: obtain a partition of the image into K homogeneous textures

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

Textured image segmentation

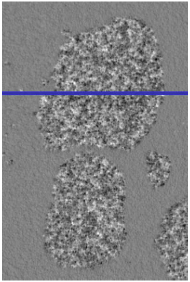


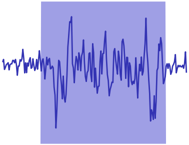


Goal: obtain a partition of the image into K **homogeneous textures**

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

Piecewise monofractal model

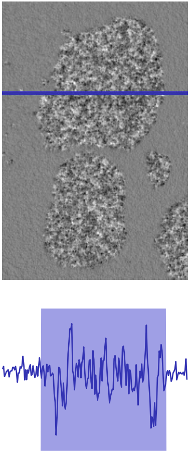




Piecewise monofractal model

Fractals attributes

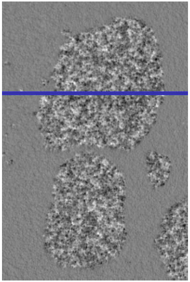
- variance σ^2 *amplitude of variations*

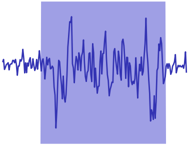


Piecewise monofractal model

Fractals attributes

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



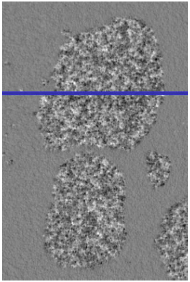


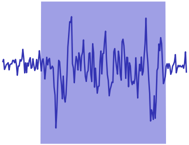
Piecewise monofractal model

Fractals attributes

- <u>variance σ^2</u> *amplitude of variations*
- <u>local regularity h</u> *scale invariance*

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



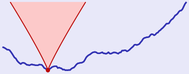


Piecewise monofractal model

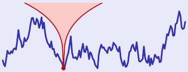
Fractals attributes

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

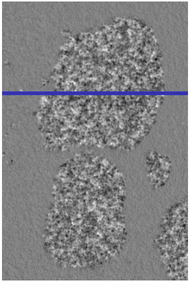
$$|f(x) - f(y)| \leq \sigma(x)|x - y|^{h(x)}$$

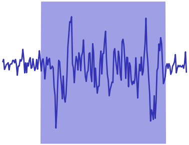


$h(x) \equiv h_1 = 0.9$



$h(x) \equiv h_2 = 0.3$



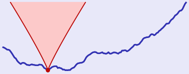


Piecewise monofractal model

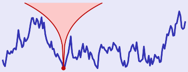
Fractals attributes

- <u>variance σ^2</u> *amplitude of variations*
- <u>local regularity h</u> *scale invariance*

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



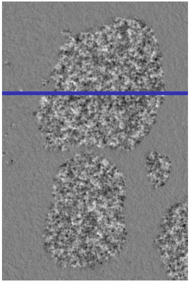
$h(x) \equiv h_1 = 0.9$

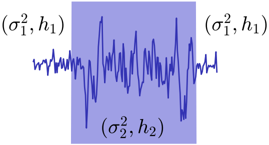


$h(x) \equiv h_2 = 0.3$

Segmentation

- σ^2 and h piecewise constant



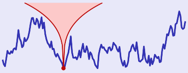


Piecewise monofractal model

Fractals attributes

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

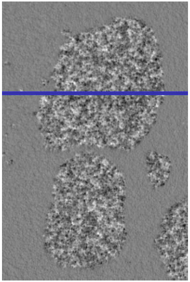
$$|f(x) - f(y)| \leq \sigma(x)|x - y|^{h(x)}$$

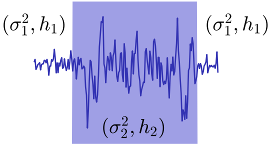


$h(x) \equiv h_1 = 0.9$ $\qquad$ $h(x) \equiv h_2 = 0.3$

Segmentation

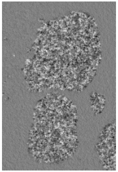
- σ^2 and h piecewise constant
- region Ω_k characterized by (σ_k^2, h_k)





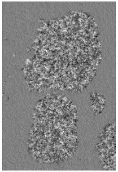
Multiscale analysis

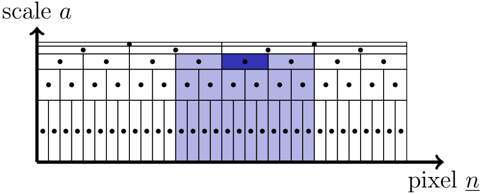
Textured image



Multiscale analysis

Textured image

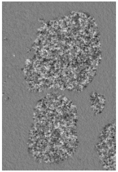




Local maximum of wavelet coefficients: $\mathcal{L}_{a,\cdot}$

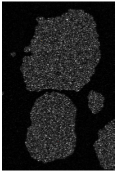
Multiscale analysis

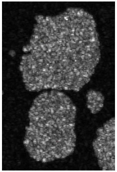
Textured image



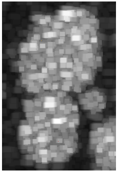
Local maximum of wavelet coefficients: $\mathcal{L}_{a,\cdot}$

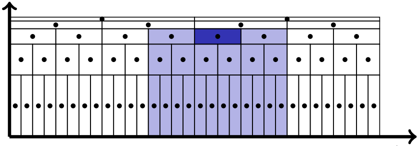
Scale $a = 2^1$ $a = 2^2$ $a = 2^5$





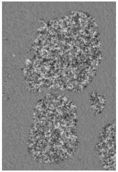
...





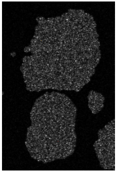
Multiscale analysis

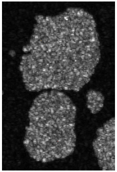
Textured image



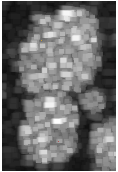
Local maximum of wavelet coefficients: $\mathcal{L}_{a,\cdot}$

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





$\cdots$

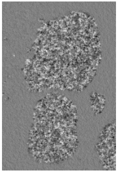


Proposition *(Jaffard, 2004), (Wendt, 2008)*

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

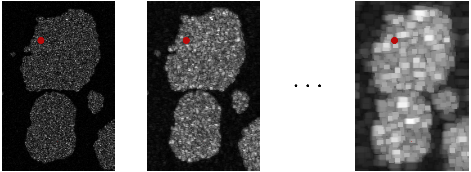
Multiscale analysis

Textured image



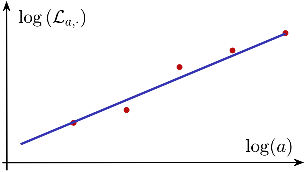
Local maximum of wavelet coefficients: $\mathcal{L}_{a,\cdot}$

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



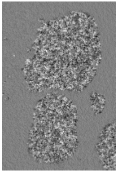
Proposition *(Jaffard, 2004), (Wendt, 2008)*

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



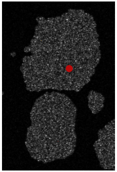
Multiscale analysis

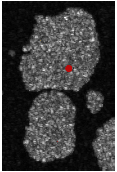
Textured image



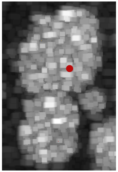
Local maximum of wavelet coefficients: $\mathcal{L}_{a,\cdot}$

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



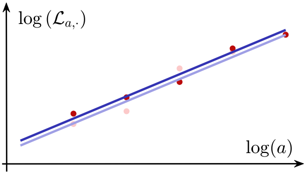


$\dots$



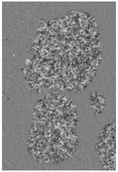
Proposition *(Jaffard, 2004), (Wendt, 2008)*

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



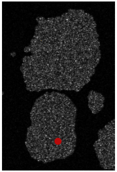
Multiscale analysis

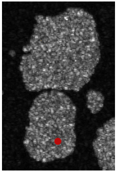
Textured image

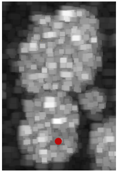


Local maximum of wavelet coefficients: $\mathcal{L}_{a,\cdot}$

Scale $a = 2^1$ $a = 2^2$ … $a = 2^5$

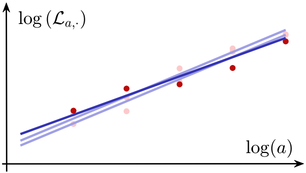






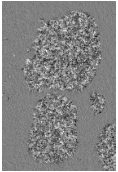
Proposition *(Jaffard, 2004), (Wendt, 2008)*

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



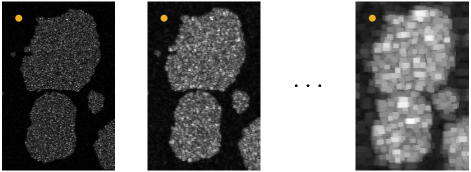
Multiscale analysis

Textured image



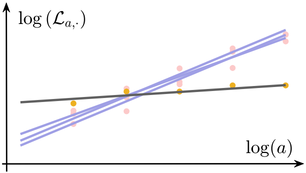
Local maximum of wavelet coefficients: $\mathcal{L}_{a,\cdot}$

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



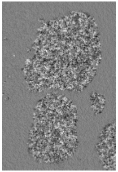
Proposition *(Jaffard, 2004), (Wendt, 2008)*

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



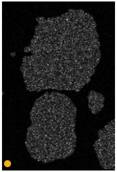
Multiscale analysis

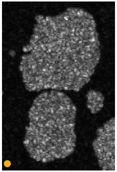
Textured image



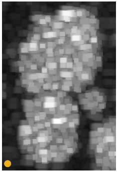
Local maximum of wavelet coefficients: $\mathcal{L}_{a,\cdot}$

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



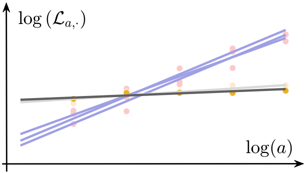


$\cdots$



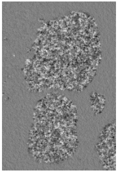
Proposition *(Jaffard, 2004), (Wendt, 2008)*

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



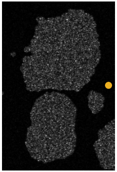
Multiscale analysis

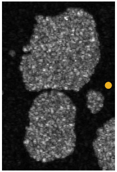
Textured image



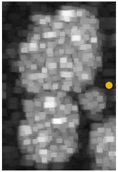
Local maximum of wavelet coefficients: $\mathcal{L}_{a,\cdot}$

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



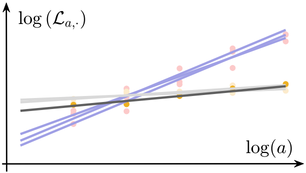


$\cdots$



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

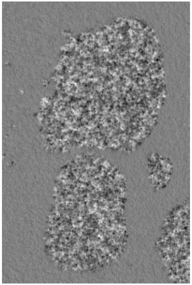
$$\log\left(\mathcal{L}_{a,\cdot}\right) \underset{a\to 0}{\simeq} \log(a) \underset{\text{regularity}}{\boldsymbol{h}} + \underset{\substack{\propto\log(\boldsymbol{\sigma}^2)\\ \text{(variance)}}}{\boldsymbol{v}}$$



Direct punctual 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

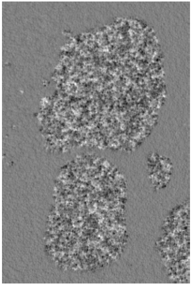


Direct punctual estimation

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

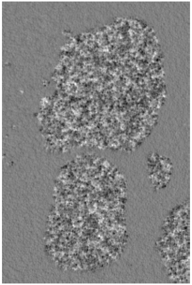


Direct punctual estimation

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

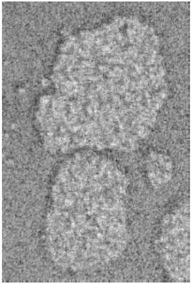
$$\left(\widehat{\boldsymbol{h}}^{\mathrm{LR}}, \widehat{\boldsymbol{v}}^{\mathrm{LR}}\right) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a=a_{\min}}^{a_{\max}} \left\| \log\left(\mathcal{L}_{a,\cdot}\right) - \log(a)\boldsymbol{h} - \boldsymbol{v} \right\|^2$$

Textured image



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

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

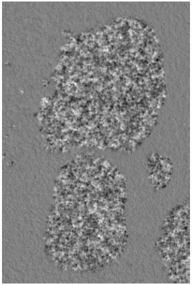


Direct punctual estimation

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

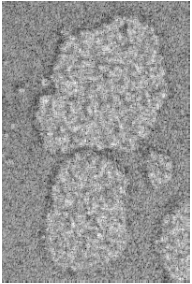
$$\left(\widehat{\boldsymbol{h}}^{\mathrm{LR}}, \widehat{\boldsymbol{v}}^{\mathrm{LR}}\right) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a=a_{\min}}^{a_{\max}} \left\|\log\left(\mathcal{L}_{a,\cdot}\right) - \log(a)\boldsymbol{h} - \boldsymbol{v}\right\|^2$$

Textured image



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

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



$\longrightarrow$ large estimation variance

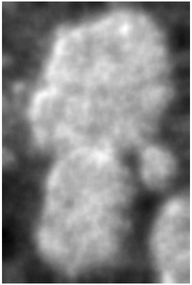
A posteriori regularization

Filter smoothing (linear)

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

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

Lissage



A posteriori regularization

Filter smoothing (linear)

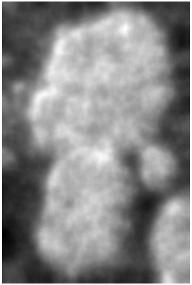
$$\left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \widehat{\boldsymbol{h}}^{\mathrm{LR}}$$

ROF denoising (nonlinear)

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

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

Lissage



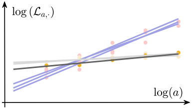
ROF

$\longrightarrow$ cumulative estimation variance and regularization bias

Functionals with either free or co-localized contours

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

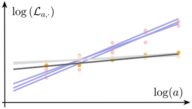
$\rightarrow$ fidelity to the log-linear model



Functionals with either free or co-localized contours

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

$\rightarrow$ fidelity to the log-linear model

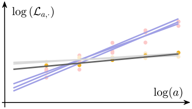


$\rightarrow$ favors piecewise constancy

Functionals with either free or co-localized contours

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

$\rightarrow$ fidelity to the log-linear model



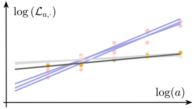
$\rightarrow$ favors piecewise constancy

Functionals with either free or co-localized contours

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

$\rightarrow$ fidelity to the log-linear model

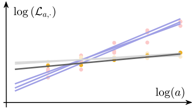
$\rightarrow$ favors piecewise constancy



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

Functionals with either free or co-localized contours

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \underbrace{\sum_a \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\substack{\textbf{Least-Squares} \\ \to \text{ fidelity to the log-linear model}}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\substack{\textbf{Total Variation} \\ \to \text{ favors piecewise constancy}}}$$



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

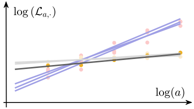
$$\mathcal{Q}_{\mathrm{F}}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha) = \alpha\|\mathbf{D}\boldsymbol{h}\|_{2,1} + \|\mathbf{D}\boldsymbol{v}\|_{2,1}$$

Functionals with either free or co-localized contours

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

→ fidelity to the log-linear model

→ 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

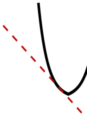
$$\mathcal{Q}_{\mathrm{F}}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha) = \alpha\|\mathbf{D}\boldsymbol{h}\|_{2,1} + \|\mathbf{D}\boldsymbol{v}\|_{2,1}$$

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

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

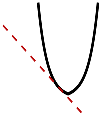
Functionals minimization

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



Functionals minimization

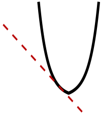
$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \ \sum_a \underbrace{\|\log \mathcal{L}_{a,.} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least-Squares}} \ + \ \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total Variation}}$$



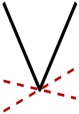
- gradient descent $\boldsymbol{x}^{n+1} = \boldsymbol{x}^n - \tau \nabla \varphi(\boldsymbol{x}^n)$

Functionals minimization

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



nonsmooth



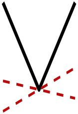
- gradient descent $\boldsymbol{x}^{n+1} = \boldsymbol{x}^n - \tau \nabla \varphi(\boldsymbol{x}^n)$
- implicit subgradient descent: proximal point algorithm

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

Functionals minimization

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

nonsmooth



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

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

- splitting proximal algorithm

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

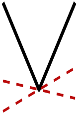
$$\boldsymbol{x}^{n+1} = \mathrm{prox}_{\tau\|\mathcal{L}-\Phi\cdot\|_2^2}\left(\boldsymbol{x}^n - \tau \mathbf{D}^\top \boldsymbol{y}^{n+1}\right), \qquad \Phi : (\boldsymbol{h}, \boldsymbol{v}) \mapsto \{\log(a)\boldsymbol{h} + \boldsymbol{v}\}_a$$

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

Functionals 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

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

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

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

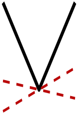
$$\boldsymbol{x}^{n+1} = \operatorname{prox}_{\tau\|\boldsymbol{\mathcal{L}} - \boldsymbol{\Phi}\cdot\|_2^2}\left(\boldsymbol{x}^n - \tau \mathbf{D}^\top \boldsymbol{y}^{n+1}\right), \qquad \boldsymbol{\Phi} : (\boldsymbol{h}, \boldsymbol{v}) \mapsto \{\log(a)\boldsymbol{h} + \boldsymbol{v}\}_a$$

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

Computation of proximal operators

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

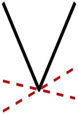
nonsmooth



Computation of proximal operators

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

nonsmooth

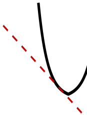


Ex. Mixed norm: for $\boldsymbol{z} = [\boldsymbol{z}_1; \ldots, ; \boldsymbol{z}_I]$

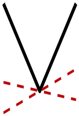
$$\mathcal{Q}(\boldsymbol{z}) = \|\boldsymbol{z}\|_{2,1} = \sum_{\underline{n}\in\Omega} \sqrt{\sum_{i=1}^{I} z_i^2(\underline{n})} = \sum_{\underline{n}\in\Omega} \|\boldsymbol{z}(\underline{n})\|_2$$

Computation of proximal operators

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



nonsmooth



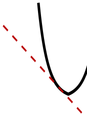
Ex. Mixed norm: for $\boldsymbol{z} = [\boldsymbol{z}_1; \ldots, ; \boldsymbol{z}_I]$

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

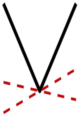
$$\boldsymbol{p} = \mathrm{prox}_{\lambda\|\cdot\|_{2,1}}(\boldsymbol{z}) \quad \Leftrightarrow \quad p_i(\underline{n}) = \max\left(0, 1 - \frac{\lambda}{\|\boldsymbol{z}(\underline{n})\|_2}\right) z_i(\underline{n})$$

Computation of proximal operators

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



nonsmooth

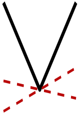


Least-Squares: $\|\log \boldsymbol{\mathcal{L}} - \Phi(\boldsymbol{h}, \boldsymbol{v})\|^2, \quad \Phi : (\boldsymbol{h}, \boldsymbol{v}) \mapsto \{\log(a)\boldsymbol{h} + \boldsymbol{v}\}_a$

Computation of proximal operators

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

nonsmooth



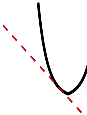
Least-Squares: $\|\log \boldsymbol{\mathcal{L}} - \boldsymbol{\Phi}(\boldsymbol{h}, \boldsymbol{v})\|^2, \quad \boldsymbol{\Phi} : (\boldsymbol{h}, \boldsymbol{v}) \mapsto \{\log(a)\boldsymbol{h} + \boldsymbol{v}\}_a$

Proposition *(Pascal, 2019)*

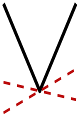
$$(\widetilde{\boldsymbol{h}}, \widetilde{\boldsymbol{v}}) = \mathrm{prox}_{\tau\|\boldsymbol{\mathcal{L}} - \boldsymbol{\Phi}\cdot\|^2}(\boldsymbol{h}, \boldsymbol{v}) \Longleftrightarrow (\widetilde{\boldsymbol{h}}, \widetilde{\boldsymbol{v}}) = \left(\mathbf{I} + \tau\boldsymbol{\Phi}^\top\boldsymbol{\Phi}\right)^{-1}\left((\boldsymbol{h},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

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

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 $\boldsymbol{\mathcal{G}} = \sum_a \log(a) \log \boldsymbol{\mathcal{L}}_a$, alors

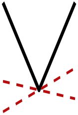
$$(\widetilde{\boldsymbol{h}}, \widetilde{\boldsymbol{v}}) = \operatorname{prox}_{\tau\|\boldsymbol{\mathcal{L}} - \boldsymbol{\Phi}\cdot\|^2}(\boldsymbol{h}, \boldsymbol{v}) \iff (\widetilde{\boldsymbol{h}}, \widetilde{\boldsymbol{v}}) = \left(\mathbf{I} + \tau\boldsymbol{\Phi}^\top\boldsymbol{\Phi}\right)^{-1}\left((\boldsymbol{h},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

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

Accelerated algorithm based on strong-convexity

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

nonsmooth



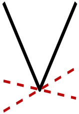
Primal-dual algorithm *(Chambolle, 2011)*

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

Accelerated algorithm based on strong-convexity

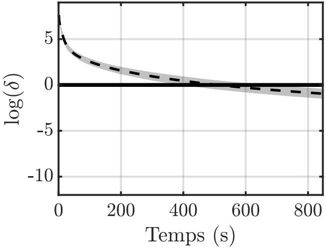
$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \sum_{a} \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,.} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least-Squares}} + \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total Variation}}$$

nonsmooth



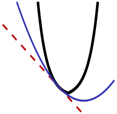
Primal-dual algorithm *(Chambolle, 2011)*

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



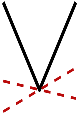
Convexity properties

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



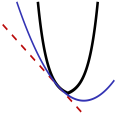
$\boldsymbol{\mu}$-strongly convex

nonsmooth



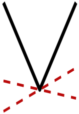
Convexity properties

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



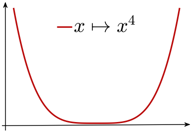
$\boldsymbol{\mu}$-strongly convex

nonsmooth

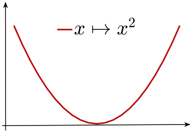


Strong-convexity

- φ $\boldsymbol{\mu}$-strongly convex iff $\varphi - \dfrac{\boldsymbol{\mu}}{2}\|\cdot\|^2$ convex



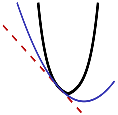
✓ strictly convex
✗ non strongly convex



✓ strictly convex
✓ 1-strongly convex

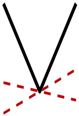
Convexity properties

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



$\boldsymbol{\mu}$-strongly convex

nonsmooth

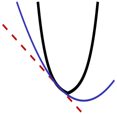


Strong-convexity

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

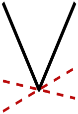
Convexity properties

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



μ-strongly convex

nonsmooth



Strong-convexity

- φ μ-strongly convex iff $\varphi - \dfrac{\mu}{2}\|\cdot\|^2$ convex
- φ $\mathcal{C}^2$ with Hessian matrix $\boldsymbol{H}\varphi \succeq 0 \implies \mu = \min \mathrm{Sp}(\boldsymbol{H}\varphi)$

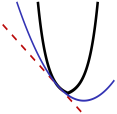
Proposition *(Pascal, 2019)*

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

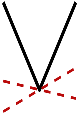
Accelerated algorithm based on strong-convexity

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



μ-strongly convex

nonsmooth



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

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

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

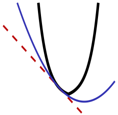
$$\boldsymbol{x}^{n+1} = \operatorname{prox}_{\tau_n\|\boldsymbol{\mathcal{L}} - \boldsymbol{\Phi}\cdot\|_2^2}\left(\boldsymbol{x}^n - \tau_n \mathbf{D}^\top \boldsymbol{y}^{n+1}\right)$$

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

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

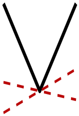
Algorithme accéléré par forte-convexité

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



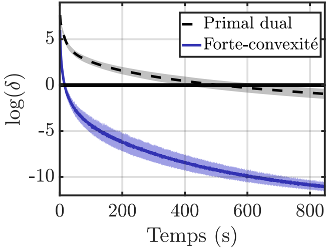
μ-strongly convex

nonsmooth



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

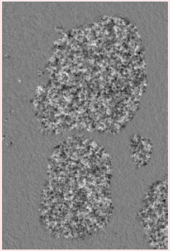
δ: duality gap, $\delta(\boldsymbol{x}^n, \boldsymbol{y}^n) \underset{n\to\infty}{\longrightarrow} 0$



Segmentation *via* iterated thresholding

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

Textured image

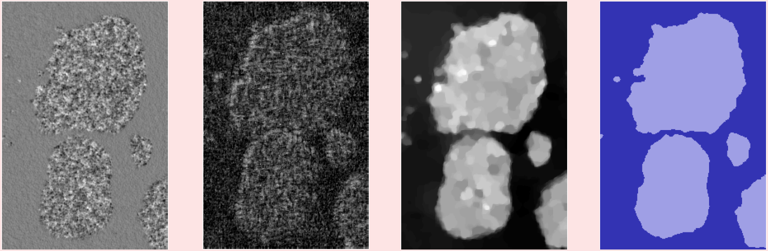


Lin. reg. $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$

Segmentation *via* iterated thresholding

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

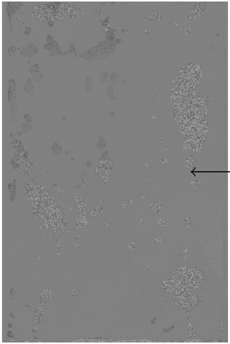
Textured image	Lin. reg. $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$	Co-localized contours $\widehat{\boldsymbol{h}}^{\mathrm{C}}$	Threshold estimate† $T\widehat{\boldsymbol{h}}^{\mathrm{C}}$



† *(Cai, 2013)*

Multiphase flow through porous media

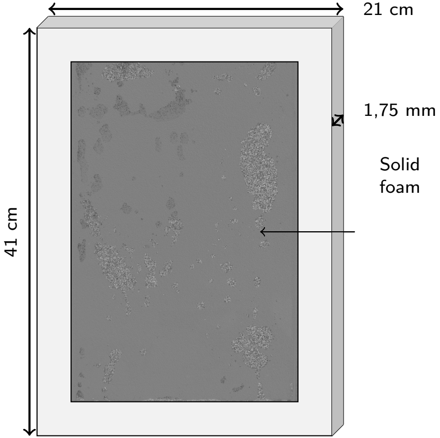
Laboratoire de Physique, ENS Lyon, V. Vidal, T. Busser, (M. Serres, IFPEN)



Solid
foam

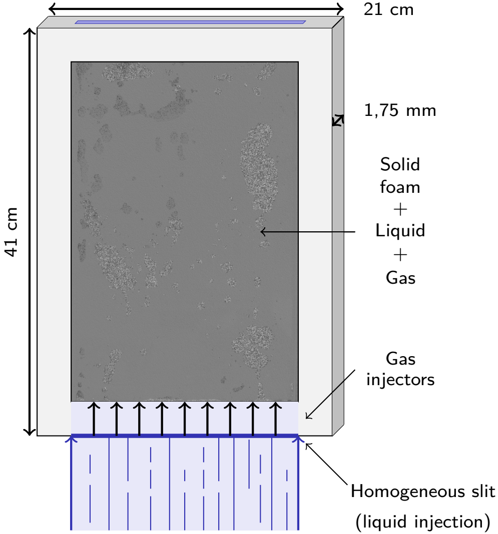
Multiphase flow through porous media

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



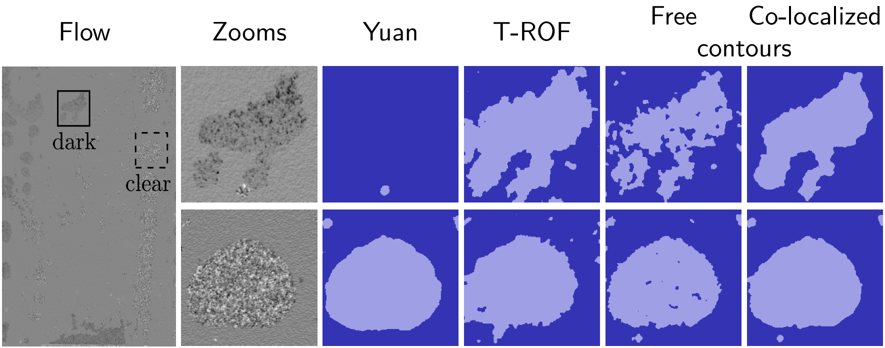
Multiphase flow through porous media

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

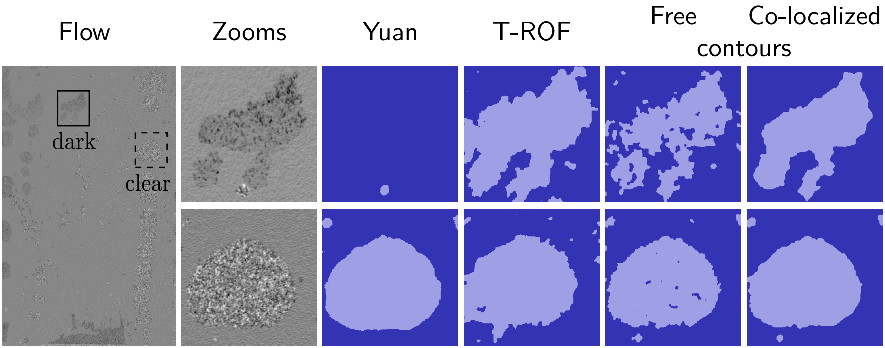


→ 1600×1100 pixels
→ video: $\sim$ 1000 images
→ phase diagram: $\sim$ 10 flow rates

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



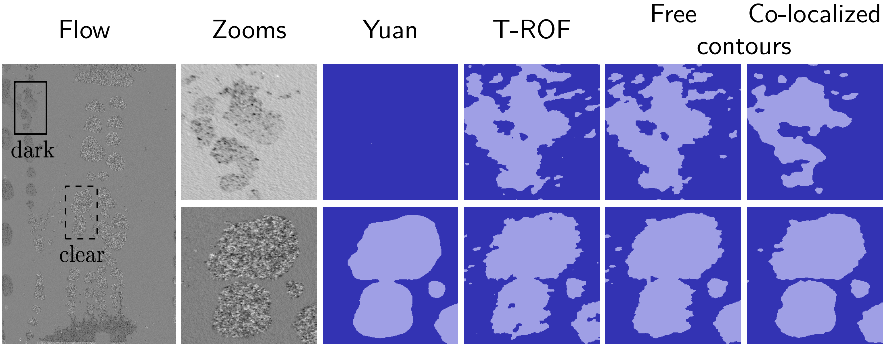
Low activity: $Q_{\mathrm{G}} = 300$mL/min - $Q_{\mathrm{L}} = 300$mL/min



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

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

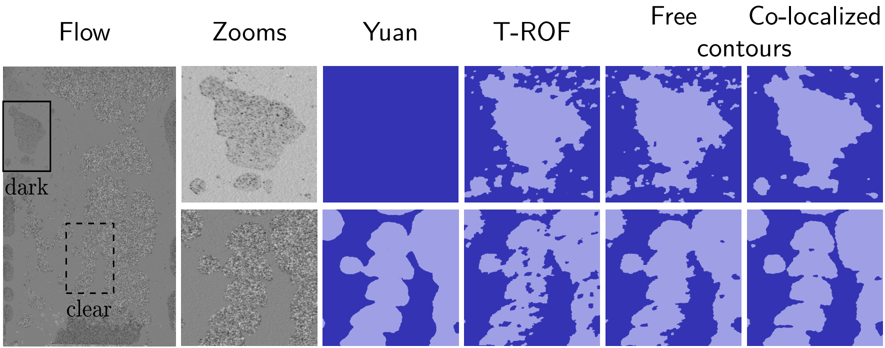
Transition: $Q_{\mathrm{G}} = 400\mathrm{mL}/\mathrm{min}$ - $Q_{\mathrm{L}} = 700\mathrm{mL}/\mathrm{min}$



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

Gas: $h_{\mathrm{G}} = 0.9 \quad \left| \begin{array}{ll} \sigma^2_{\mathrm{dark}} = 10^{-2} & \text{(dark bubbles)} \\ \sigma^2_{\mathrm{clear}} = 10^{-1} & \text{(clear bubbles).} \end{array} \right.$

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

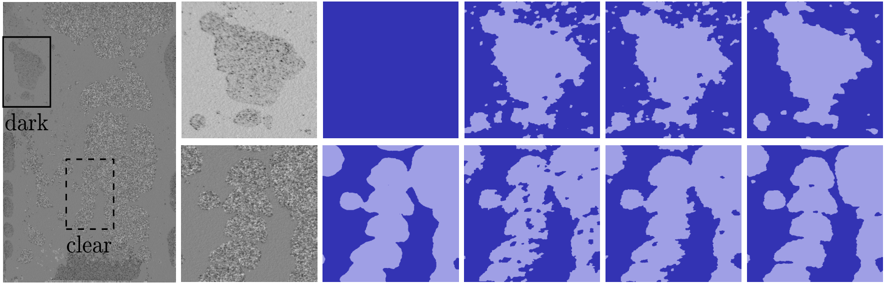


Liquid: $h_{\mathrm{L}} = 0.4 \qquad \sigma^2_{\mathrm{dark}} = 10^{-2}$

Gas: $h_{\mathrm{G}} = 0.9 \quad \left| \begin{array}{ll} \sigma^2_{\mathrm{dark}} = 10^{-2} & \text{(dark bubbles)} \\ \sigma^2_{\mathrm{clear}} = 10^{-1} & \text{(clear bubbles).} \end{array} \right.$

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

Flow	Zooms	Yuan	T-ROF	Free contours	Co-localized contours



Computational load	1s	12s	700s	**2100s**

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

Gas: $h_{\mathrm{G}} = 0.9 \quad \left| \begin{array}{ll} \sigma^2_{\mathrm{dark}} = 10^{-2} & \text{(dark bubbles)} \\ \sigma^2_{\mathrm{clear}} = 10^{-1} & \text{(clear bubbles).} \end{array} \right.$

Regularization parameters selection

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

Regularization parameters selection

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

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

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

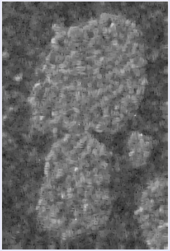
Regularization parameters selection

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

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

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

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



too small

Regularization parameters selection

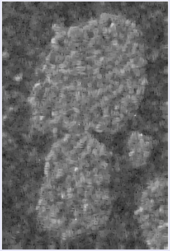
$$\left(\widehat{\boldsymbol{h}}, \widehat{\boldsymbol{v}}\right)(\mathcal{L}; \lambda, \alpha) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \mathcal{L}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v};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

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

Co-localized contours estimate $\widehat{\boldsymbol{h}}^{\mathrm{C}}$

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

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



too small

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

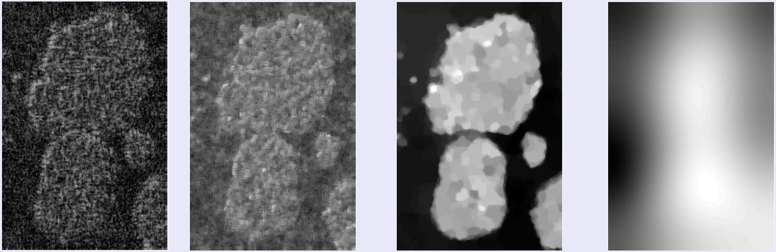
too large

Regularization parameters selection

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

Lin. reg. $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$ Co-localized contours 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 determine $\lambda^\dagger$ and $\alpha^\dagger$?

Parameter tuning (Grid search)

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

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

Parameter tuning (Grid 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

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

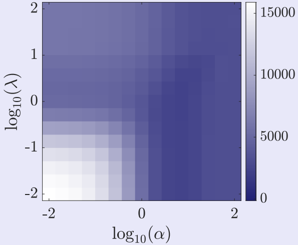
Parameter tuning (Grid search)

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

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

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

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



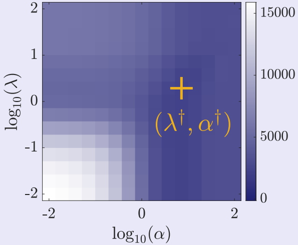
Parameter tuning (Grid search)

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

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

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

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



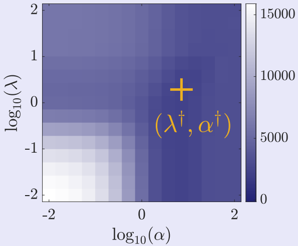
Parameter tuning (Grid search)

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

$\boldsymbol{h}$: *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!

?

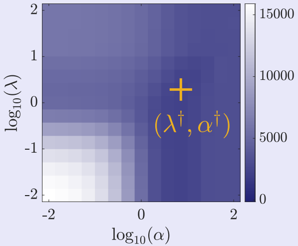
Parameter tuning (Grid search)

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

$\boldsymbol{h}$: *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!

?

Stein Unbiased Risk Estimate
(SURE)

Stein Unbiased Risk Estimate (Principe)

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

Stein Unbiased Risk Estimate (Principe)

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

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

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

Stein Unbiased Risk Estimate (Principe)

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

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

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

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

Stein Unbiased Risk Estimate (Principe)

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

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

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

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

Theorem *(Stein, 1981)*

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

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

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

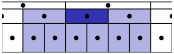
Generalized *Stein Unbiased Risk Estimate*

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. the estimators $\widehat{h}(\mathcal{L}; \lambda, \alpha)$ with free or co-localized contours

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

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



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

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

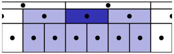
Generalized *Stein Unbiased Risk Estimate*

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. the estimators $\widehat{\boldsymbol{h}}(\mathcal{L}; \lambda, \alpha)$ with free or co-localized contours

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

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



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

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

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

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

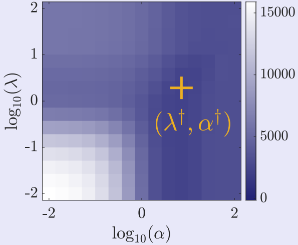
$$\Longrightarrow R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) = \mathbb{E}_\zeta[\widehat{R}(\boldsymbol{\Lambda})].$$

Parameter tuning (Grid 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!

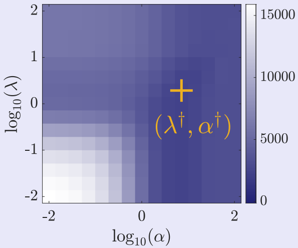
$$\widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha | \boldsymbol{\mathcal{S}})$$

Parameter tuning (Grid search)

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

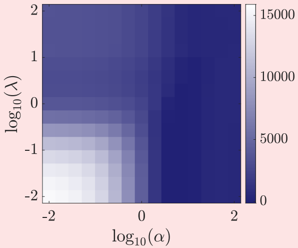
$\bar{\boldsymbol{h}}$: *true* regularity

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



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

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

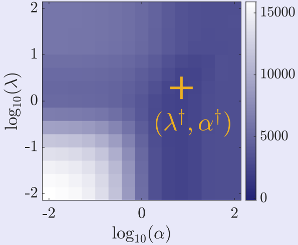


Parameter tuning (Grid search)

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

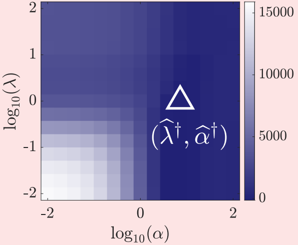
$\bar{\boldsymbol{h}}$: *true* regularity

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



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

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

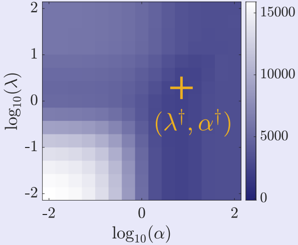


Parameter tuning (Grid search)

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

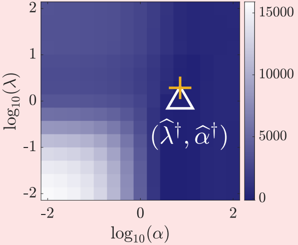
$\bar{\boldsymbol{h}}$: *true* regularity

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



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

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

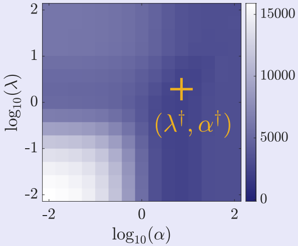


Parameter tuning (Automatic selection)

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

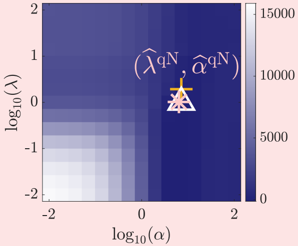
$\bar{\boldsymbol{h}}$: *true* regularity

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



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

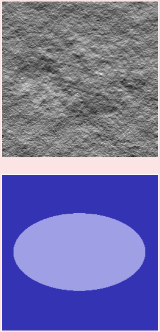
$$\widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{\mathcal{L}};\lambda,\alpha|\,\boldsymbol{\mathcal{S}})$$



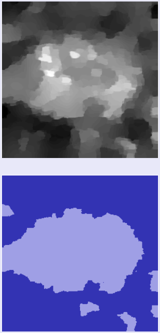
Automated selection of regularization parameters

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

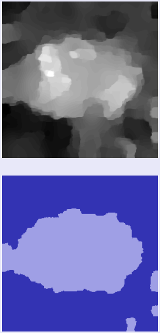
Example



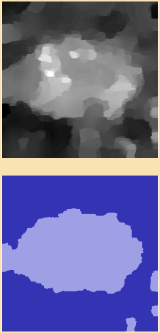
$\widehat{\boldsymbol{h}}^{\mathsf{F}}(\boldsymbol{\mathcal{L}}; \lambda^{\dagger}, \alpha^{\dagger})$
(grid)



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



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



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

Part I: Fractal texture segmentation

Take home messages

- Fractal texture model based on local *regularity* and *variance*
  - ∗ appropriate for real-world texture characterization
  - ∗ complementary attributes able to finely discriminate

Part I: Fractal texture segmentation

Take home messages

- Fractal texture model based on local *regularity* and *variance*
  - ∗ appropriate for real-world texture characterization
  - ∗ complementary attributes able to finely discriminate
- Simultaneous estimation and regularization
  - ∗ significant decrease of the estimation error
  - ∗ accurate and regular contours thanks to co-localized penalization

Part I: Fractal texture segmentation

Take home messages

- Fractal texture model based on local *regularity* and *variance*
  - ∗ appropriate for real-world texture characterization
  - ∗ complementary attributes able to finely discriminate

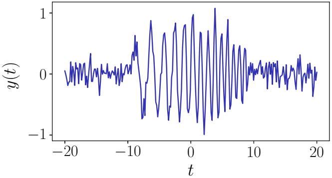
- Simultaneous estimation and regularization
  - ∗ significant decrease of the estimation error
  - ∗ accurate and regular contours thanks to co-localized penalization

- Fast algorithms for automated tuning of hyperparameters
  - ∗ possibility to manage huge amount of data
  - ∗ amenable to process data corrupted by *correlated* Gaussian noise
  - ∗ ensured objectivity and reproducibility

Part II: Point processes in time-frequency analysis

Harmonic analysis of temporal signals

Standard modeling of a "signal": $y : \mathbb{R} \to \mathbb{C}$ function of time t.



- electrical cardiac activity,
- audio recording,
- seismic activity,
- light intensity on a photosensor
- ...

Information of interest:

- time events, e.g., an earthquake and its replica
- frequency content, e.g., monitoring of the heart beating rate

time	frequency
ever-changing world	waves, oscillations, rhythms
marker of events and evolutions	intrinsic mechanisms

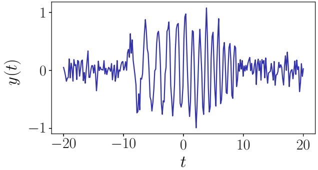
Harmonic analysis of temporal signals

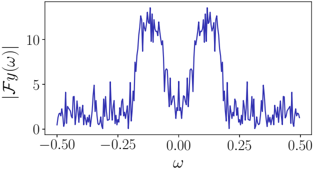
Noisy chirp: transient waveform modulated in amplitude and frequency

$$y(t) = e_\nu(t) \sin\left(2\pi\left(f_1 + (f_2 - f_1)\frac{t+\nu}{2\nu}\right)t\right) + \sigma n(t)$$

Time or frequency

Fourier transform: $\mathcal{F}y(\omega) \triangleq \int_{\mathbb{R}} \overline{y(t)} \exp(-\mathrm{i}\omega t)\,\mathrm{d}t$



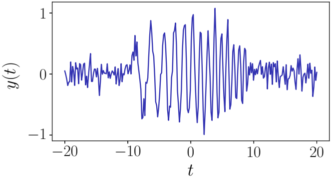


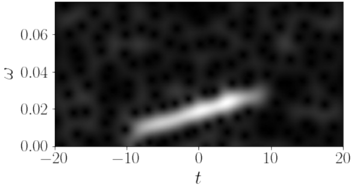
In the Fourier representation, the temporal information is **lost**.

Time-frequency analysis

Time *and* frequency Short-Time Fourier Transform with window h:

$$V_h y(t,\omega) \triangleq \int_{-\infty}^{\infty} \overline{y(u)} h(u-t) \exp(-\mathrm{i}\omega u) \, \mathrm{d}u$$





Energy density interpretation $S_h y(t,\omega) = |V_h y(t,\omega)|^2$ the *spectrogram*

$$\int \int_{-\infty}^{+\infty} S_h y(t,\omega) \, \mathrm{d}t \frac{\mathrm{d}\omega}{2\pi} = \int_{-\infty}^{+\infty} |x(t)|^2 \, \mathrm{d}t \quad \text{if} \quad \|h\|_2^2 = 1$$

Signal, i.e., information of interest: regions of maximal energy.

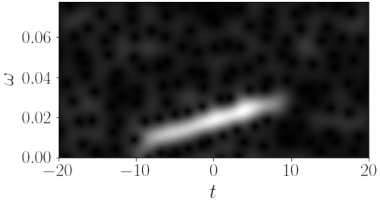
Standard: denoising based on the spectrogram maxima

<u>Inversion formula</u> $$y(t) = \int\int_{-\infty}^{+\infty} \overline{V_h y(u,\omega)} h(t-u) \exp(\mathrm{i}\omega u)\, \mathrm{d}u \frac{\mathrm{d}\omega}{2\pi}$$

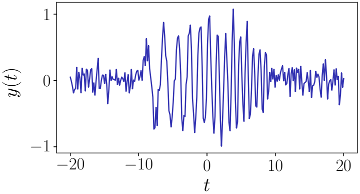
Standard: denoising based on the spectrogram maxima

Inversion formula
$$y(t) = \int \int_{-\infty}^{+\infty} \overline{V_h y(u,\omega)} h(t-u) \exp(\mathrm{i}\omega u) \, \mathrm{d}u \frac{\mathrm{d}\omega}{2\pi}$$

spectrogram



noisy observation

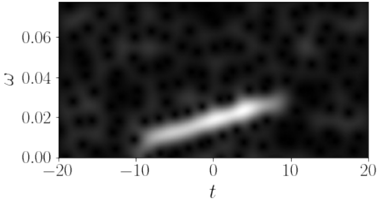


Standard: denoising based on the spectrogram maxima

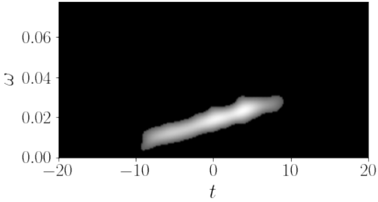
<u>Inversion formula</u>

$$y(t) = \int \int_{-\infty}^{+\infty} \overline{V_h y(u,\omega)} h(t-u) \exp(\mathrm{i}\omega u) \, \mathrm{d}u \frac{\mathrm{d}\omega}{2\pi}$$

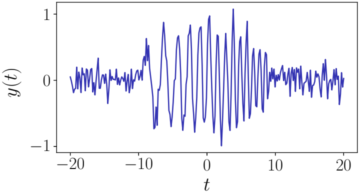
spectrogram



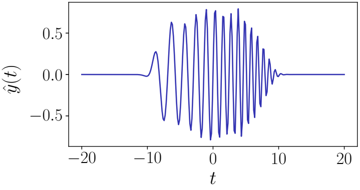
thresholded



noisy observation



denoised signal

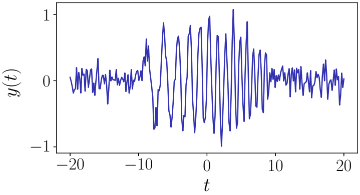


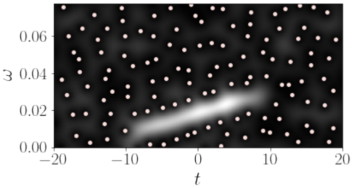
Unorthodox: focus on the spectrogram zeros

Restriction to the *circular Gaussian window*: $g(t) = \pi^{-1/4} e^{-t^2/2}$

Look for the (t_i, ω_i) such that $S_g(t_i, \omega_i) = 0$.

[Flandrin, 2015]



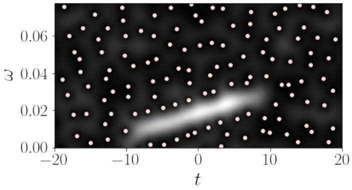


<u>Observations:</u>

- zeros are "repelled" by the signal,
- in the "noise" region, zeros are evenly spread,
- short-range repulsion between zeros.

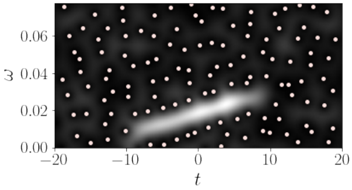
Unorthodox: theoretical study of the spectrogram zeros

Idea assimilate the time-frequency plane with $\mathbb{C}$ through $z = \omega + \mathrm{i}t$



Unorthodox: theoretical study of the spectrogram zeros

Idea assimilate the time-frequency plane with $\mathbb{C}$ through $z = \omega + \mathrm{i}t$



Bargmann factorization

$$V_g y(t,\omega) = \mathrm{e}^{-|z|^2/4}\mathrm{e}^{-\mathrm{i}\omega t/2}\mathcal{B}y(z/\sqrt{2})$$

where the Bargmann transform of the signal y, defined as

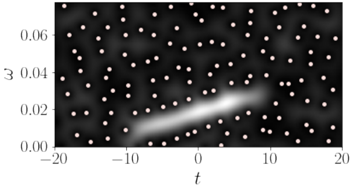
$$\mathcal{B}y(z) \triangleq \pi^{-1/4}\mathrm{e}^{-z^2/2}\int_{\mathbb{R}} \overline{y(u)}\exp\left(\sqrt{2}uz - u^2/2\right)\,\mathrm{d}u,$$

is an **entire** function, almost characterized by its infinitely many zeros:

$$\mathcal{B}y(z) = z^m \mathrm{e}^{C_0 + C_1 z + C_2 z^2}\prod_{n\in\mathbb{N}}\left(1 - \frac{z}{z_n}\right)\exp\left(\frac{z}{z_n} + \frac{1}{2}\left(\frac{z}{z_n}\right)^2\right).$$

Unorthodox: theoretical study of the spectrogram zeros

Idea assimilate the time-frequency plane with $\mathbb{C}$ through $z = \omega + \mathrm{i}t$



Bargmann factorization

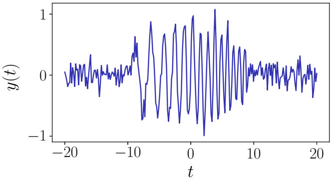
$$V_g y(t,\omega) = \mathrm{e}^{-|z|^2/4}\mathrm{e}^{-\mathrm{i}\omega t/2}\mathcal{B}y(z/\sqrt{2})$$

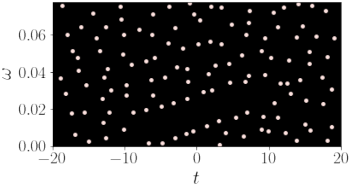
Theorem The zeros of the Gaussian spectrogram $V_g y(t,\omega)$

- coincide with the zeros of $\mathcal{B}y(\cdot/\sqrt{2})$, which is an **entire** function
- hence are **isolated** and constitute a **random point process**,
- which almost completely **characterizes** the spectrogram.

[Flandrin, 2015]

Unorthodox: the point pattern of the spectrogram zeros

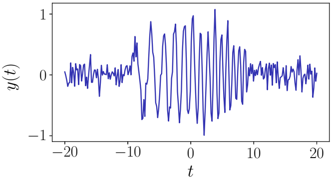


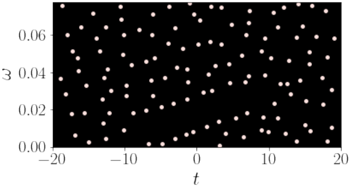


Advantages of working with the zeros

- easy to find compared to *relative* maxima,
- require little memory space for storage,
- use of the tools of **stochastic geometry**.

Unorthodox: the point pattern of the spectrogram zeros





Advantages of working with the zeros

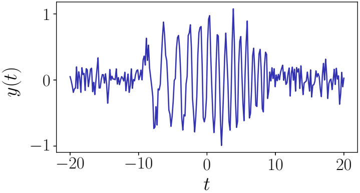
- easy to find compared to *relative* maxima,
- require little memory space for storage,
- use of the tools of **stochastic geometry**.

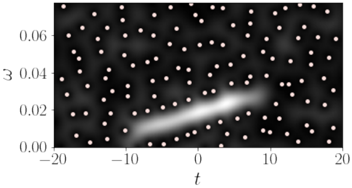
Application: hypothesis testing for signal detection

- H_0 white noisy only, i.e., $y(t) = n(t)$
- H_1 presence of a signal i.e., $y(t) = x(t) + \sigma n(t)$

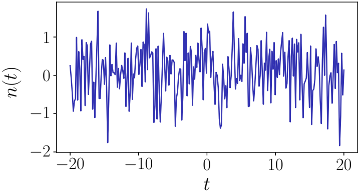
Unorthodox path: signal detection from the zeros

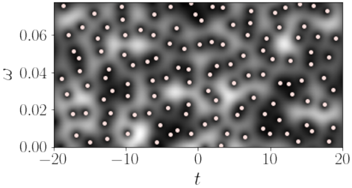
Noisy chirp H_1



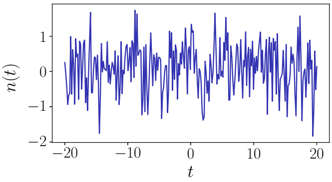


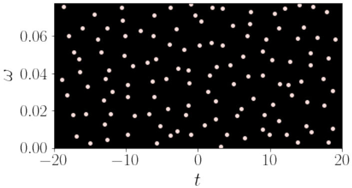
White noise only H_0





Unorthodox path: zeros of the spectrogram of white noise





Complex white noise $\quad \xi(t) = \sum_{k=0}^{\infty} \xi_k h_k(t), \ \xi_k \sim \mathcal{N}_{\mathbb{C}}(0,1)$

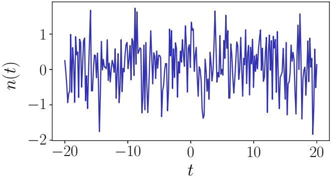
$\{h_k, k = 0, 1, \ldots\}$ the Hermite functions, Hilbertian basis of $L^2(\mathbb{R})$

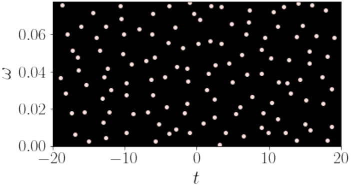
Theorem

$$V_g\xi(t,\omega) = \mathrm{e}^{-|z|^2/4}\mathrm{e}^{-\mathrm{i}\omega t/2} \sum_{k=0}^{\infty} \xi_k \frac{1}{\sqrt{k!}} \left(\frac{z}{\sqrt{2}}\right)^k$$

[Bardenet & Hardy, 2021]

Unorthodox path: zeros of Gaussian Analytic Functions





$$V_g\xi(t,\omega) = \mathrm{e}^{-|z|^2/4}\mathrm{e}^{-\mathrm{i}\omega t/2}\mathrm{GAF}_{\mathbb{C}}(z/\sqrt{2})$$

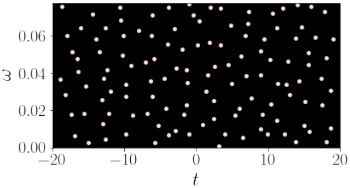
where $\mathrm{GAF}_{\mathbb{C}}(z) = \sum_{k=0}^{\infty} \xi_k \frac{z^k}{\sqrt{k!}}, \quad \xi_k \sim \mathcal{N}_{\mathbb{C}}(0,1)$

<u>Zeros of the *Planar* Gaussian Analytic Function (GAF)</u>

$$\mathcal{Z}(\mathrm{GAF}_{\mathbb{C}}) \overset{\text{(def.)}}{=} \{z_i, \text{ s.t. } \mathrm{GAF}_{\mathbb{C}}(z_i) = 0\}$$

Spatial statistics of the point process $\mathcal{Z}(\mathrm{GAF}_{\mathbb{C}})$ known explicitly.

Unorthodox path: zeros of Gaussian Analytic Functions



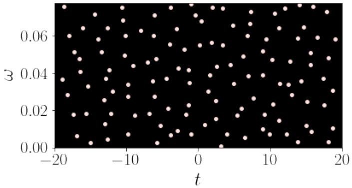
$$V_g\xi(t,\omega) \propto \mathrm{GAF}_{\mathbb{C}}(z/\sqrt{2})$$
$$z = \omega + \mathrm{i}t$$

Properties of the point process $\mathcal{Z}(\mathrm{GAF}_{\mathbb{C}})$:

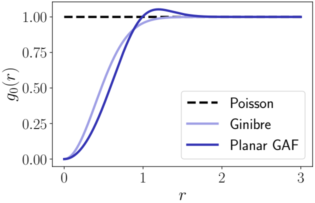
- invariant under the isometries of $\mathbb{C}$, i.e., **stationary**,
- has a uniform density $\rho^{(1)}(z) = \rho^{(1)} = 1/\pi$,
- explicit pair correlation function $\rho^{(2)}(z, z') = g_0(|z - z'|)$,
- scaling of the *hole probability*: $r^{-4} \log p_r \to -3\mathrm{e}^2/4$, as $r \to \infty$

$$p_r = \mathbb{P}\,(\text{no point in the disk of center 0 and radius } r)$$

Unorthodox path: zeros of Gaussian Analytic Functions



$$V_g\xi(t,\omega) \propto \mathrm{GAF}_{\mathbb{C}}(z/\sqrt{2})$$
$$z = \omega + \mathrm{i}t$$

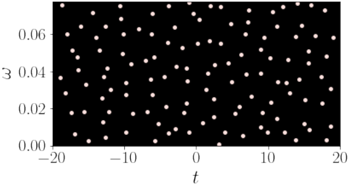


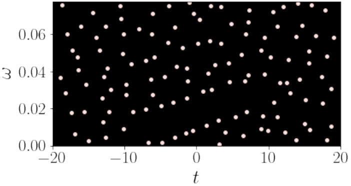
<u>Pair correlation</u> $\rho^{(2)}(z,z')\mathrm{d}z\mathrm{d}z' =$

$\mathbb{P}\,($1 point in $B(z,\mathrm{d}z)$ and 1 in $B(z',\mathrm{d}z'))$

The point process of the zeros of the spectrogram is not **determinantal**.

Unorthodox path: zeros of Gaussian Analytic Functions





Signal detection based on spatial statistics:

- the K-function

$$K(r) = 2\pi \int_0^r s g_0(s)\,\mathrm{d}s : \ \textbf{\# pairs at distance less than } r$$

- the F-function

$$F(r) = \mathbb{P}\left(\inf_{z_i \in \mathcal{Z}} \mathrm{d}(z_0, z_i) < r \right) : \ \textbf{empty space function}$$

Unorthodox: other GAF, other transforms

Spherical Gaussian Analytic Function

$$\mathrm{GAF}_{\mathbb{S}}(z) = \sum_{k=0}^{N} \xi_k \sqrt{\binom{N}{k}} z^k, \quad \xi_k \sim \mathcal{N}_{\mathbb{C}}(0,1)$$

" **Kravchuk transform**" of a **discrete** signal $y = \{y_k,\ k = 0, 1, \ldots, N\}$

$$K_y(\vartheta, \varphi) = \sum_{n=0}^{N} Ty_n \sqrt{\binom{N}{n}} \left(\cos\frac{\vartheta}{2}\right)^n \left(\sin\frac{\vartheta}{2}\right)^{N-n} \mathrm{e}^{\mathrm{i}\varphi n},\ z = \cot\vartheta/2\mathrm{e}^{\mathrm{i}\varphi}$$

with $Ty_n = \langle y, k_n \rangle$, $\{k_n, n = 0, 1, ..., N\}$ the *Kravchuk functions*.