



## Texture segmentation based on fractal attributes

Barbara Pascal  
([bpascal-fr.github.io](https://bpascal-fr.github.io))

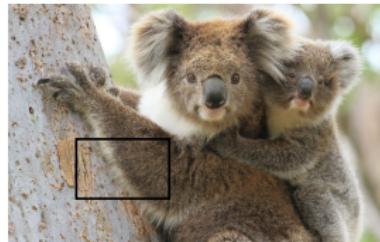
Joint work with Patrice Abry, Nelly Pustelnik, Valérie Vidal, and Samuel Vaïter

July 4<sup>th</sup> 2023

Harmonic and Multifractal Analyses: from Mathematics to Quantitative Neuroscience

Centre de Recherches Mathématiques (CRM), Montréal, Canada

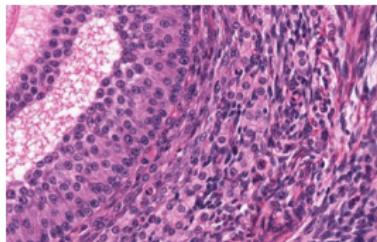
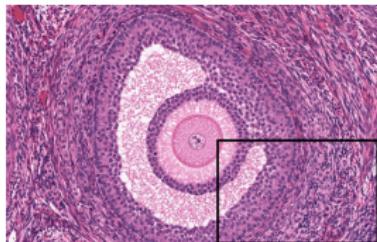
# Textures



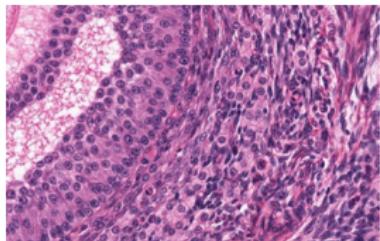
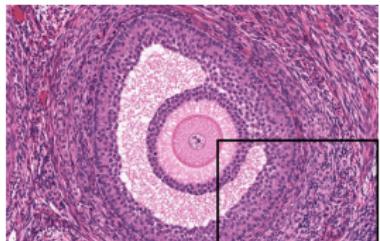
# Textures



# Textures

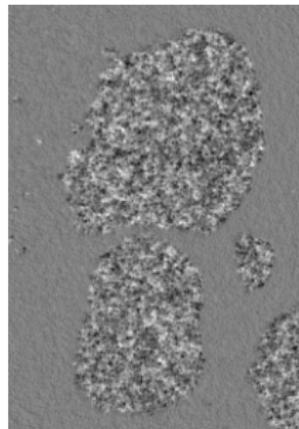


# Textures

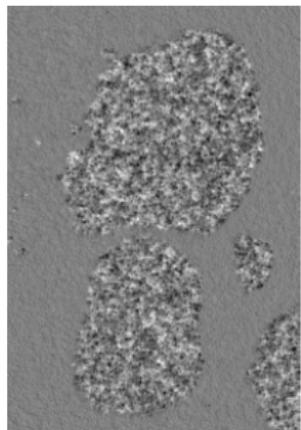


Crucial to describe real-world images

## Textured image segmentation



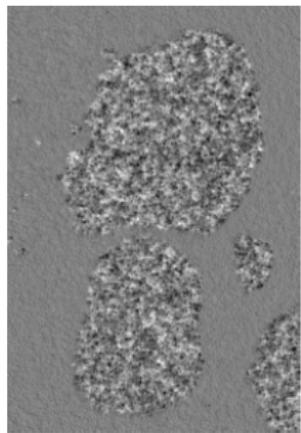
## Textured image segmentation



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

$$\Omega = \Omega_1 \sqcup \dots \sqcup \Omega_K$$

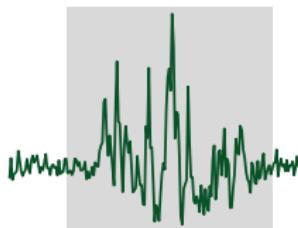
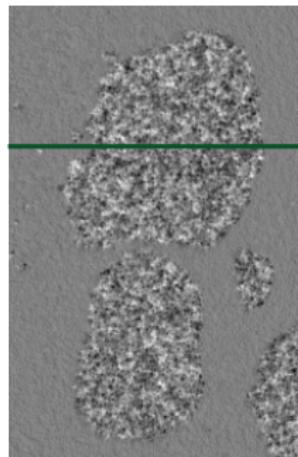
## Textured image segmentation



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

$$\Omega = \Omega_1 \sqcup \dots \sqcup \Omega_K$$

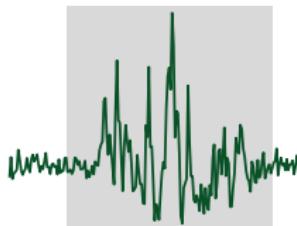
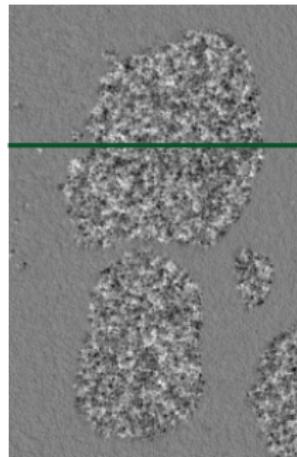
## Piecewise monofractal model



# Piecewise monofractal model

## Fractals attributes

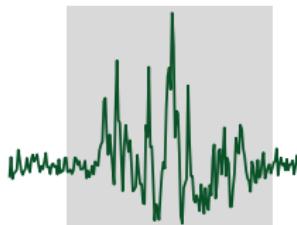
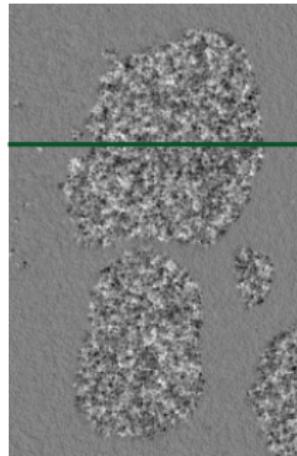
- variance  $\sigma^2$       *amplitude of variations*



# Piecewise monofractal model

## Fractals attributes

- variance  $\sigma^2$       *amplitude of variations*
- local regularity  $h$       *scale invariance*

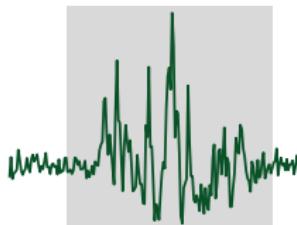
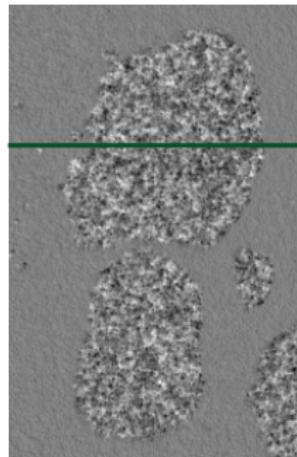


# Piecewise monofractal model

## Fractals attributes

- variance  $\sigma^2$       *amplitude of variations*
- local regularity  $h$       *scale invariance*

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



# Piecewise monofractal model

## Fractals attributes

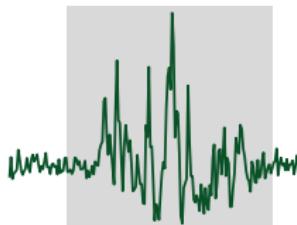
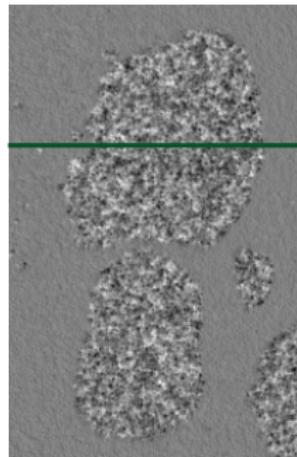
- variance  $\sigma^2$       *amplitude of variations*
- local regularity  $h$       *scale invariance*

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



$$h(x) \equiv h_1 = 0.9$$

$$h(x) \equiv h_2 = 0.3$$



# Piecewise monofractal model

## Fractals attributes

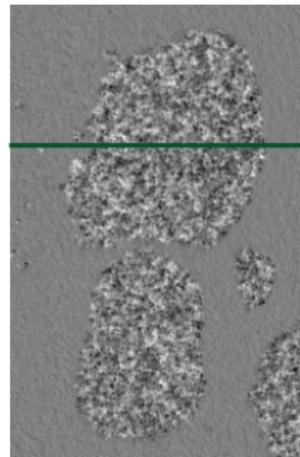
- variance  $\sigma^2$       *amplitude of variations*
- local regularity  $h$       *scale invariance*

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



$$h(x) \equiv h_1 = 0.9$$

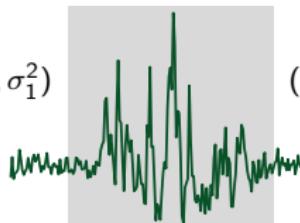
$$h(x) \equiv h_2 = 0.3$$



$$(h_2, \sigma_2^2)$$

$$(h_1, \sigma_1^2)$$

$$(h_1, \sigma_1^2)$$



## Segmentation

- $\sigma^2$  and  $h$  piecewise constant

# Piecewise monofractal model

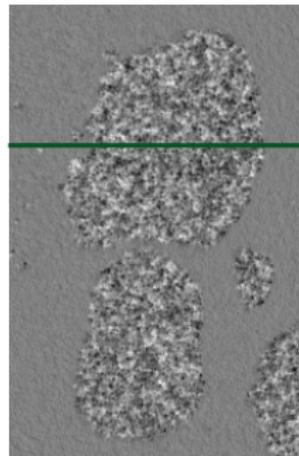
## Fractals attributes

- variance  $\sigma^2$       *amplitude of variations*
- local regularity  $h$       *scale invariance*

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

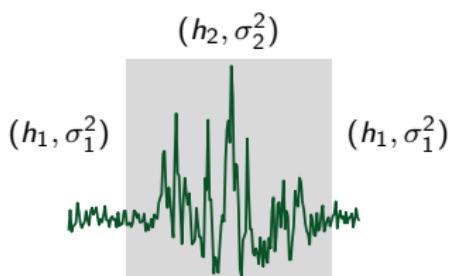


$$h(x) \equiv h_1 = 0.9 \quad h(x) \equiv h_2 = 0.3$$



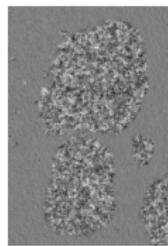
## Segmentation

- ▶  $\sigma^2$  and  $h$  piecewise constant
- ▶ region  $\Omega_k$  characterized by  $(\sigma_k^2, h_k)$



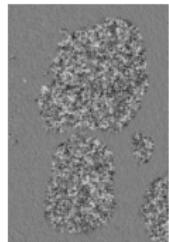
# Multiscale analysis

Textured image



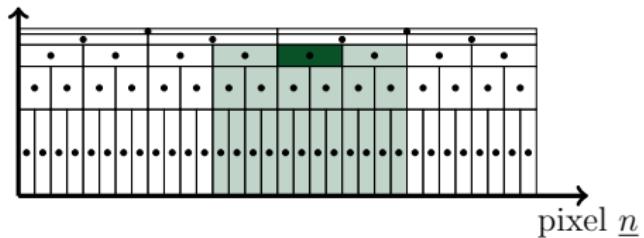
# Multiscale analysis

Textured image



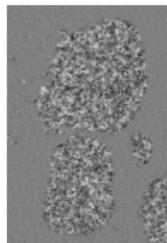
Local maximum of wavelet coefficients:  $\mathcal{L}_{a,\cdot}$

scale  $2^j$



# Multiscale analysis

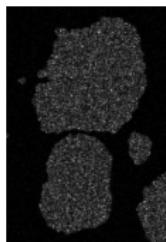
Textured image



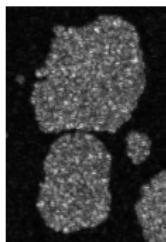
Local maximum of wavelet coefficients:  $\mathcal{L}_{a,\cdot}$

Scale

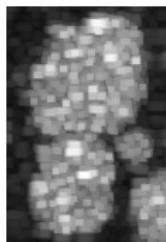
$a = 2^1$



$a = 2^2$

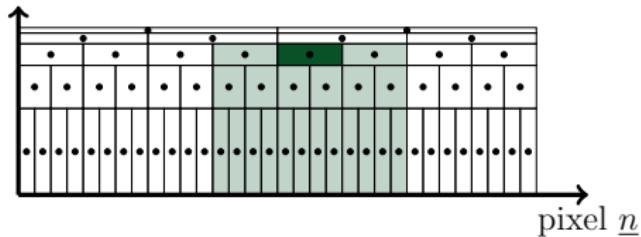


$a = 2^5$



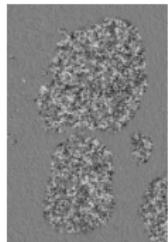
...

scale  $2^j$



# Multiscale analysis

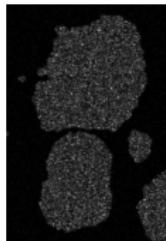
Textured image



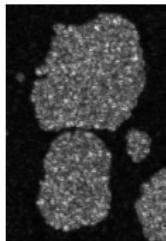
Local maximum of wavelet coefficients:  $\mathcal{L}_{a,\cdot}$

Scale

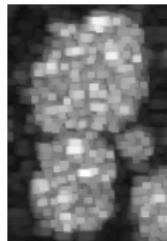
$a = 2^1$



$a = 2^2$



$a = 2^5$



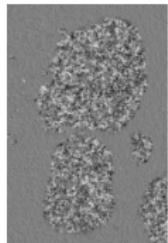
...

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

$$\log(\mathcal{L}_{a,\cdot}) \underset{a \rightarrow 0}{\simeq} \log(a) \underset{\text{regularity}}{h} + \underset{\propto \log(\sigma^2)}{\nu} \underset{\text{(variance)}}{}$$

# Multiscale analysis

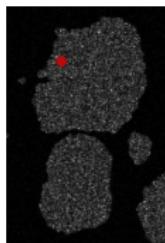
Textured image



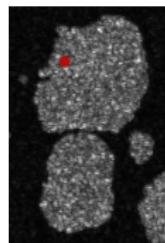
Local maximum of wavelet coefficients:  $\mathcal{L}_{a,\cdot}$

Scale

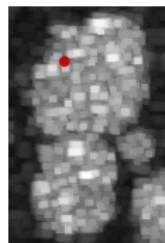
$a = 2^1$



$a = 2^2$



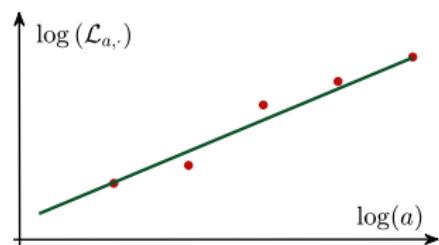
$a = 2^5$



...

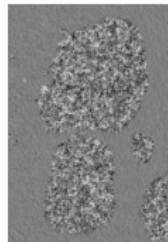
Proposition (Jaffard, 2004), (Wendt, 2008)

$$\log(\mathcal{L}_{a,\cdot}) \underset{a \rightarrow 0}{\simeq} \log(a) \begin{matrix} h \\ \text{regularity} \end{matrix} + \begin{matrix} v \\ \propto \log(\sigma^2) \\ (\text{variance}) \end{matrix}$$



# Multiscale analysis

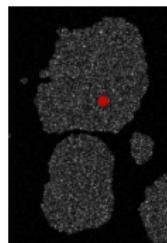
Textured image



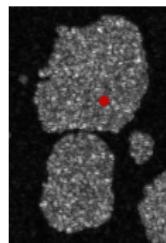
Local maximum of wavelet coefficients:  $\mathcal{L}_{a,\cdot}$

Scale

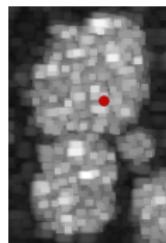
$a = 2^1$



$a = 2^2$



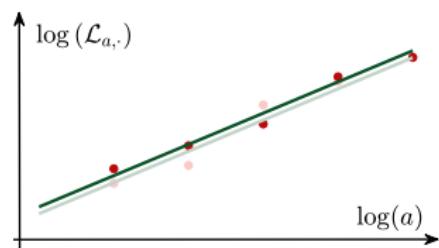
$a = 2^5$



...

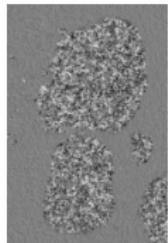
Proposition (Jaffard, 2004), (Wendt, 2008)

$$\log(\mathcal{L}_{a,\cdot}) \underset{a \rightarrow 0}{\simeq} \log(a) \begin{matrix} h \\ \text{regularity} \end{matrix} + \begin{matrix} v \\ \propto \log(\sigma^2) \\ (\text{variance}) \end{matrix}$$



# Multiscale analysis

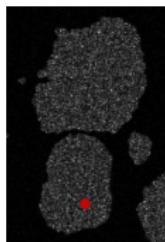
Textured image



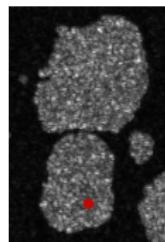
Local maximum of wavelet coefficients:  $\mathcal{L}_{a,\cdot}$

Scale

$a = 2^1$

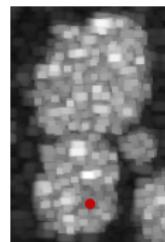


$a = 2^2$



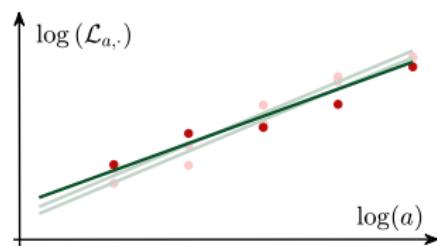
...

$a = 2^5$



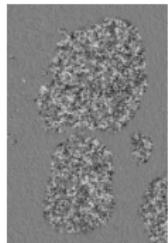
Proposition (Jaffard, 2004), (Wendt, 2008)

$$\log(\mathcal{L}_{a,\cdot}) \underset{a \rightarrow 0}{\simeq} \log(a) \begin{matrix} h \\ \text{regularity} \end{matrix} + \begin{matrix} v \\ \propto \log(\sigma^2) \\ (\text{variance}) \end{matrix}$$



# Multiscale analysis

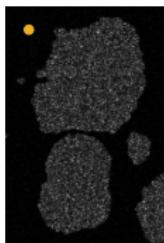
Textured image



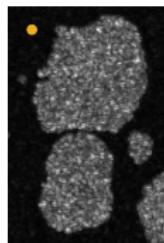
Local maximum of wavelet coefficients:  $\mathcal{L}_{a,\cdot}$

Scale

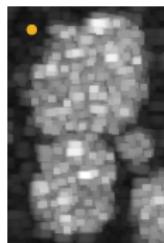
$a = 2^1$



$a = 2^2$



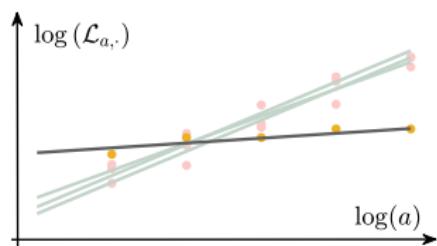
$a = 2^5$



...

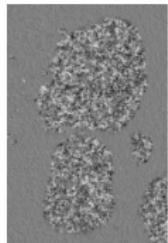
Proposition (Jaffard, 2004), (Wendt, 2008)

$$\log(\mathcal{L}_{a,\cdot}) \underset{a \rightarrow 0}{\simeq} \log(a) \underset{\text{regularity}}{h} + \underset{\propto \log(\sigma^2)}{\nu} \underset{\text{(variance)}}{}$$



# Multiscale analysis

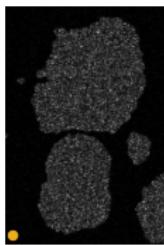
Textured image



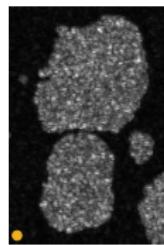
Local maximum of wavelet coefficients:  $\mathcal{L}_{a,\cdot}$

Scale

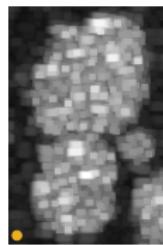
$a = 2^1$



$a = 2^2$



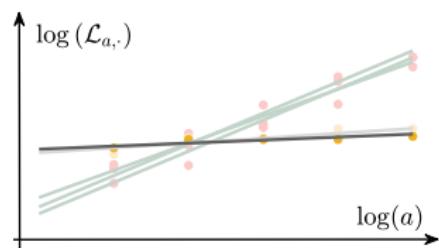
$a = 2^5$



...

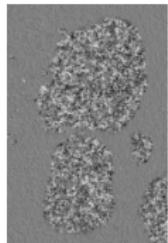
Proposition (Jaffard, 2004), (Wendt, 2008)

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



# Multiscale analysis

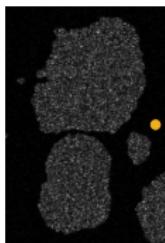
Textured image



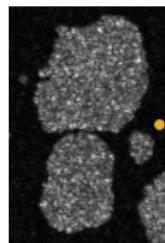
Local maximum of wavelet coefficients:  $\mathcal{L}_{a,\cdot}$

Scale

$a = 2^1$

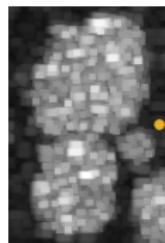


$a = 2^2$



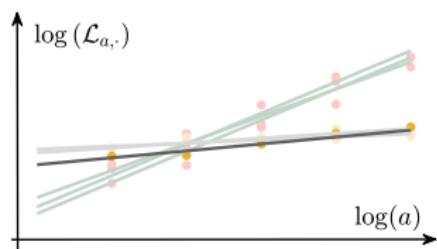
...

$a = 2^5$



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

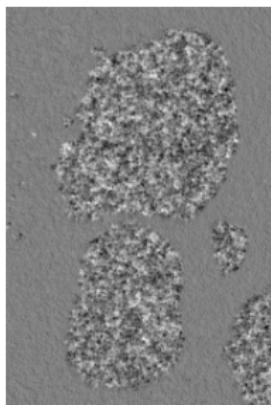
$$\log(\mathcal{L}_{a,\cdot}) \underset{a \rightarrow 0}{\simeq} \log(a) \underset{\text{regularity}}{h} + \underset{\propto \log(\sigma^2)}{\nu} \underset{\text{(variance)}}{}$$



## Direct punctual estimation

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

Textured image

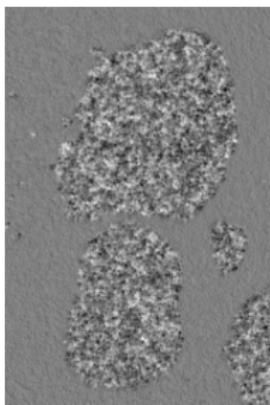


## Direct punctual estimation

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

$$(\hat{\textbf{h}}^{\text{LR}}, \hat{\textbf{v}}^{\text{LR}}) = \underset{\textbf{h}, \textbf{v}}{\text{argmin}} \sum_{a=a_{\min}}^{a_{\max}} \|\log(\mathcal{L}_{a,\cdot}) - \log(a)\textbf{h} - \textbf{v}\|^2$$

Textured image

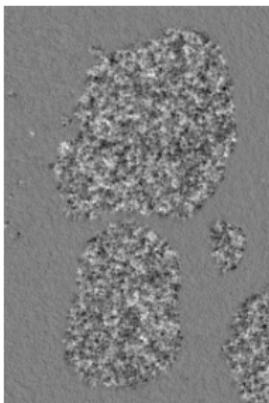


## Direct punctual estimation

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

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

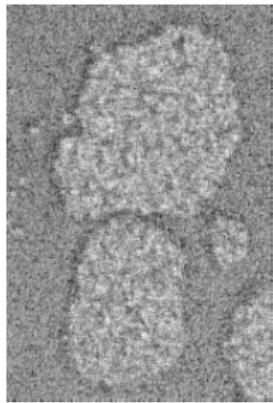
Textured image



Local regularity  $\hat{\boldsymbol{h}}^{\text{LR}}$



Local power  $\hat{\boldsymbol{v}}^{\text{LR}}$



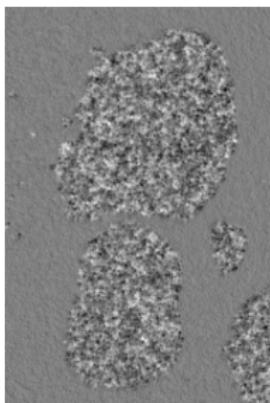
## Direct punctual estimation

### Linear regression

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

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

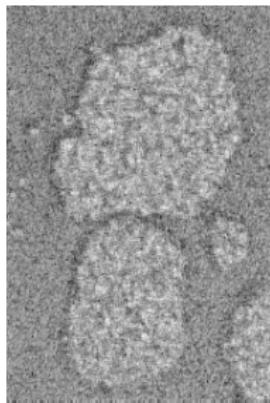
Textured image



Local regularity  $\hat{\boldsymbol{h}}^{\text{LR}}$



Local power  $\hat{\boldsymbol{v}}^{\text{LR}}$



→ large estimation variance

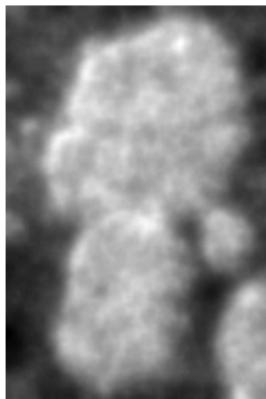
## Filter smoothing (linear)

$$\left( \mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D} \right)^{-1} \hat{\mathbf{h}}^{\text{LR}}$$

Linear regression  $\hat{\mathbf{h}}^{\text{LR}}$



Smoothing



**Filter smoothing** (linear)

$$\left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \hat{\mathbf{h}}^{\text{LR}}$$

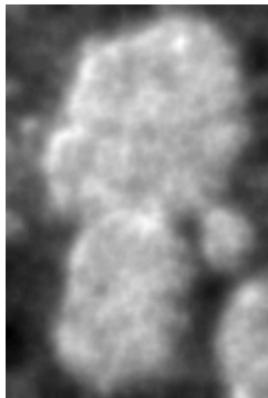
Linear regression  $\hat{\mathbf{h}}^{\text{LR}}$



**ROF denoising** (nonlinear)

$$\operatorname{argmin}_{\mathbf{h}} \|\mathbf{h} - \hat{\mathbf{h}}^{\text{LR}}\|^2 + \lambda \|\mathbf{D}\mathbf{h}\|_{2,1}$$

Smoothing



ROF

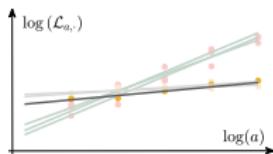


→ cumulative estimation variance and regularization bias

## Functionals with either free or co-localized contours

$$\sum_a \frac{\|\log \mathcal{L}_{a,.} - \log(a) \mathbf{h} - \mathbf{v}\|^2}{\text{Least-Squares}}$$

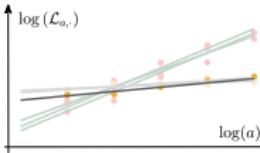
→ fidelity to the log-linear model



## Functionals with either free or co-localized contours

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

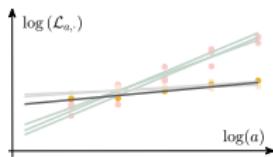
$\rightarrow$  fidelity to the log-linear model  
 $\rightarrow$  favors piecewise constancy



## Functionals with either free or co-localized contours

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

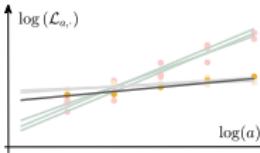
$\rightarrow$  fidelity to the log-linear model  
 $\rightarrow$  favors piecewise constancy



## Functionals with either free or co-localized contours

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

$\rightarrow$  fidelity to the log-linear model  
 $\rightarrow$  favors piecewise constancy

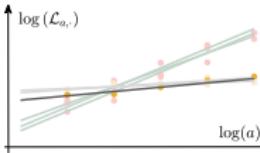


**Finite differences**  $\mathbf{D}_1 \mathbf{x}$  (horizontal),  $\mathbf{D}_2 \mathbf{x}$  (vertical) in each pixel

## Functionals with either free or co-localized contours

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

$\rightarrow$  fidelity to the log-linear model  
 $\rightarrow$  favors piecewise constancy



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

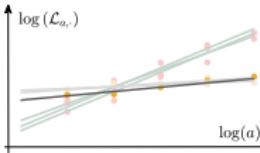
Free:  $\mathbf{h}$ ,  $\mathbf{v}$  are **independently** piecewise constant

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

## Functionals with either free or co-localized contours

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

$\rightarrow$  fidelity to the log-linear model  
 $\rightarrow$  favors piecewise constancy



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

Free:  $\mathbf{h}, \mathbf{v}$  are **independently** piecewise constant

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

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

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

## Functionals minimization

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



## Functionals minimization

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



- gradient descent  $\mathbf{x}^{n+1} = \mathbf{x}^n - \tau \nabla \varphi(\mathbf{x}^n)$   $\mathbf{x} = (\mathbf{h}, \mathbf{v})$

# Functionals minimization

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



nonsmooth



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

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

# Functionals minimization

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



nonsmooth



- ▶ gradient descent  $\mathbf{x}^{n+1} = \mathbf{x}^n - \tau \nabla \varphi(\mathbf{x}^n)$   $\mathbf{x} = (\mathbf{h}, \mathbf{v})$
- ▶ implicit subgradient descent: proximal point algorithm  
$$\mathbf{x}^{n+1} = \mathbf{x}^n - \tau \mathbf{u}^n, \quad \mathbf{u}^n \in \partial \varphi(\mathbf{x}^{n+1}) \Leftrightarrow \mathbf{x}^{n+1} = \text{prox}_{\tau \varphi}(\mathbf{x}^n)$$
- ▶ splitting proximal algorithm (Chambolle, 2011)

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

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

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

# Functionals minimization

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



nonsmooth



- ▶ gradient descent  $\mathbf{x}^{n+1} = \mathbf{x}^n - \tau \nabla \varphi(\mathbf{x}^n)$   $\mathbf{x} = (\mathbf{h}, \mathbf{v})$
- ▶ implicit subgradient descent: proximal point algorithm  
$$\mathbf{x}^{n+1} = \mathbf{x}^n - \tau \mathbf{u}^n, \quad \mathbf{u}^n \in \partial \varphi(\mathbf{x}^{n+1}) \Leftrightarrow \mathbf{x}^{n+1} = \text{prox}_{\tau \varphi}(\mathbf{x}^n)$$
- ▶ splitting proximal algorithm (Chambolle, 2011)

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

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

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

## Accelerated algorithm based on strong-convexity

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



nonsmooth



Primal-dual algorithm (Chambolle, 2011)

$$\delta: \text{duality gap}, \delta(\mathbf{x}^n, \mathbf{y}^n) \xrightarrow{n \rightarrow \infty} 0$$

# Accelerated algorithm based on strong-convexity

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

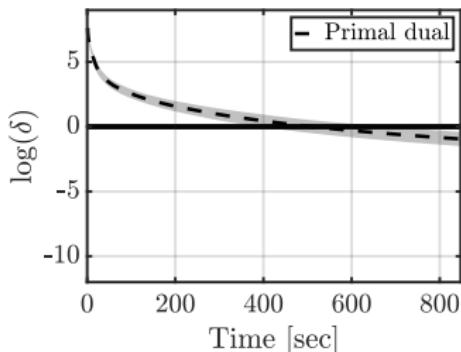


nonsmooth



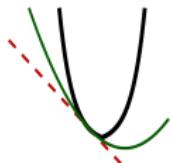
Primal-dual algorithm (Chambolle, 2011)

$\delta$ : duality gap,  $\delta(\mathbf{x}^n, \mathbf{y}^n) \rightarrow 0$



## Convexity properties

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



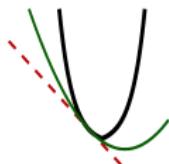
$\mu$ -strongly convex

nonsmooth



# Convexity properties

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



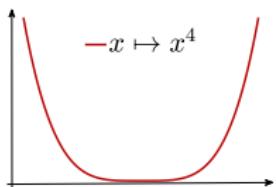
$\mu$ -strongly convex

nonsmooth

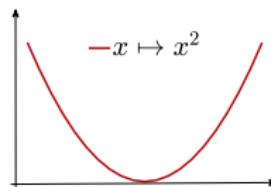


## Strong-convexity

- $\varphi$   $\mu$ -strongly convex iff  $\varphi - \frac{\mu}{2} \|\cdot\|^2$  convex



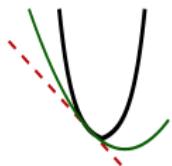
✓ strictly convex  
✗ non strongly convex



✓ strictly convex  
✓ 1-strongly convex

# Convexity properties

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



$\mu$ -strongly convex

nonsmooth

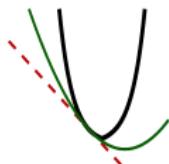


## Strong-convexity

- $\varphi$   $\mu$ -strongly convex iff  $\varphi - \frac{\mu}{2} \|\cdot\|^2$  convex
- $\varphi \in \mathcal{C}^2$  with Hessian matrix  $\mathbf{H}\varphi \succeq 0 \implies \mu = \min \text{Sp}(\mathbf{H}\varphi)$

# Convexity properties

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



$\mu$ -strongly convex

nonsmooth



## Strong-convexity

- $\varphi$   $\mu$ -strongly convex iff  $\varphi - \frac{\mu}{2} \|\cdot\|^2$  convex
- $\varphi \in \mathcal{C}^2$  with Hessian matrix  $\mathbf{H}\varphi \succeq 0 \implies \mu = \min \text{Sp}(\mathbf{H}\varphi)$

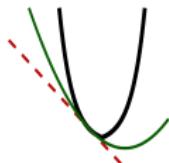
Proposition (Pascal, 2019)

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

$a_{\min} = 2^1, \quad a_{\max}$	$2^2$	$2^3$	$2^4$	$2^5$	$2^6$
$\mu = \min \text{Sp}(2\Phi^\top \Phi)$	0.29	<b>0.72</b>	1.20	1.69	2.20

# Accelerated algorithm based on strong-convexity

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



$\mu$ -strongly convex

nonsmooth



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

**for**  $n = 0, 1, \dots$   $\mathbf{x} = (\mathbf{h}, \mathbf{v})$

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

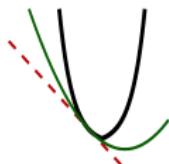
$$\mathbf{x}^{n+1} = \text{prox}_{\tau_n \|\mathcal{L} - \Phi\cdot\|_2^2} \left( \mathbf{x}^n - \tau_n \mathbf{D}^\top \mathbf{y}^{n+1} \right)$$

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

$$\bar{\mathbf{x}}^{n+1} = \mathbf{x}^{n+1} + \theta_n^{-1} (\mathbf{x}^{n+1} - \mathbf{x}^n)$$

# Accelerated algorithm based on strong-convexity

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



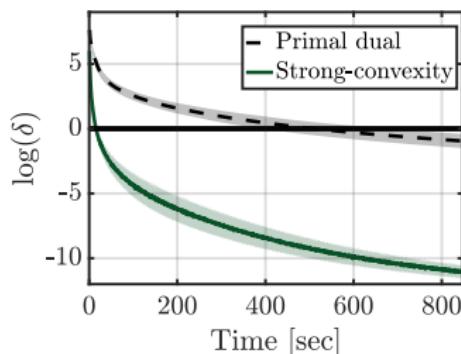
$\mu$ -strongly convex

nonsmooth



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

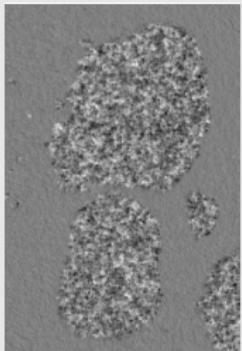
$\delta$ : duality gap,  $\delta(\mathbf{x}^n, \mathbf{y}^n) \rightarrow 0$



## Segmentation via iterated thresholding

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

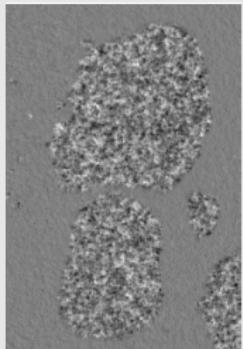
Textured image    Lin. reg.  $\hat{\mathbf{h}}^{\text{LR}}$



## Segmentation via iterated thresholding

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

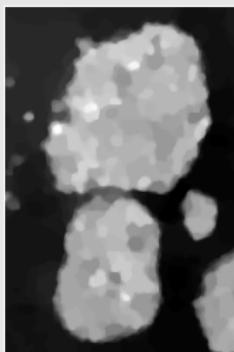
Textured image



Lin. reg.  $\hat{\mathbf{h}}^{\text{LR}}$



Co-localized  
contours  $\hat{\mathbf{h}}^{\text{C}}$



Threshold  
estimate<sup>†</sup>  $T\hat{\mathbf{h}}^{\text{C}}$



<sup>†</sup>(Cai, 2013)

## Threshold-ROF on $\hat{h}^{\text{LR}}$

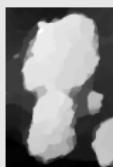
(Naftornita, 2014), (Pustelnik, 2016)

$$\operatorname{argmin}_{\mathbf{h}} \|\mathbf{h} - \hat{\mathbf{h}}^{\text{LR}}\|^2 + \lambda \|\mathbf{D}\mathbf{h}\|_{2,1}$$

Lin. reg.



ROF



Threshold



Only based on regularity  $\mathbf{h}$ .

## Threshold-ROF on $\hat{h}^{\text{LR}}$

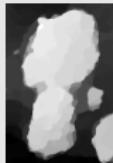
(Naftornita, 2014), (Pustelnik, 2016)

$$\operatorname{argmin}_{\mathbf{h}} \|\mathbf{h} - \hat{\mathbf{h}}^{\text{LR}}\|^2 + \lambda \|\mathbf{D}\mathbf{h}\|_{2,1}$$

Lin. reg.



ROF



Threshold

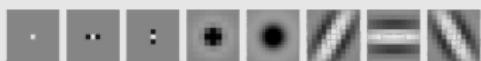


Only based on regularity  $\mathbf{h}$ .

## Factorization based segmentation<sup>†</sup>

(Yuan, 2015)

(i) local histograms



(ii) matrix factorization

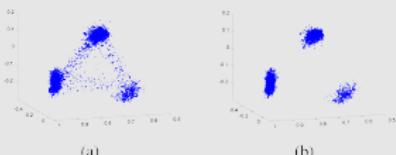
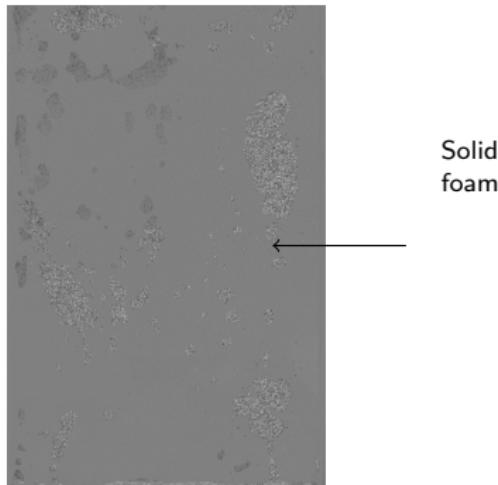


Fig. 2. Scatterplot of features in subspace. (a) Scatterplot of features projected onto the 3-d subspace. (b) Scatterplot after removing features with high edgeness.

<sup>†</sup><https://sites.google.com/site/factorizationsegmentation/>

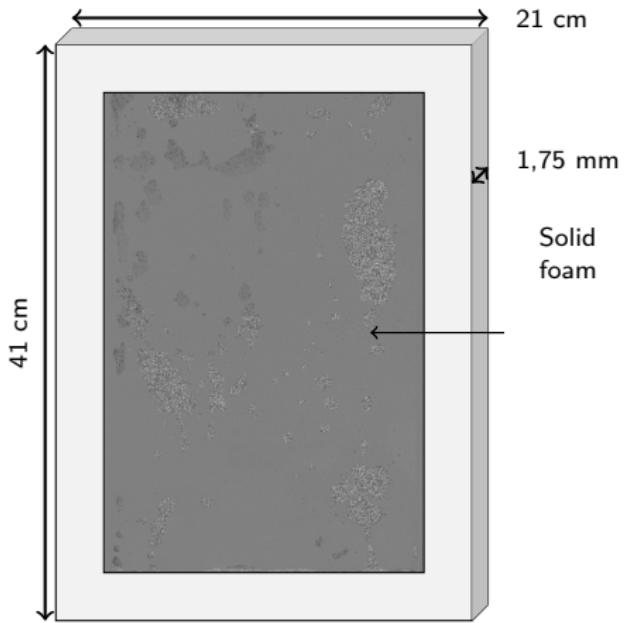
# Multiphase flow through porous media

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



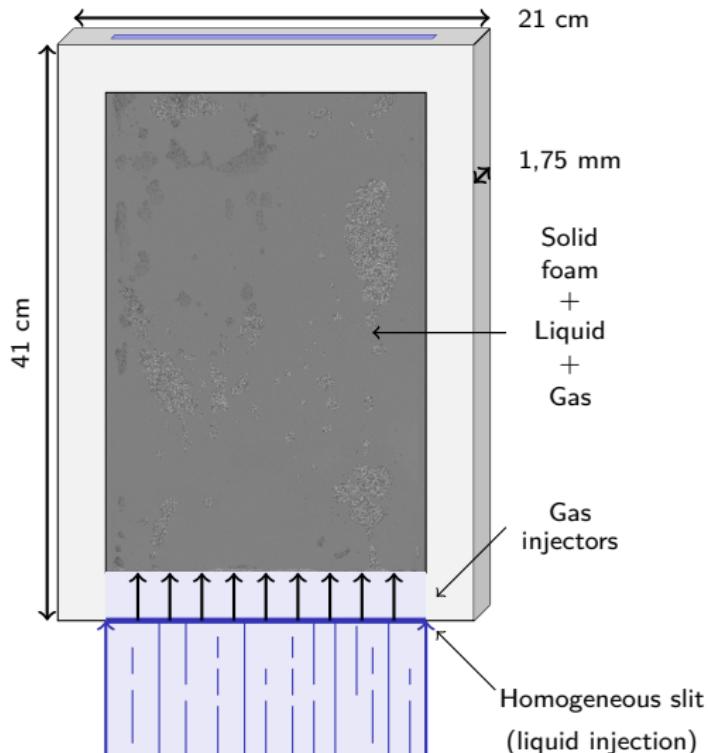
# Multiphase flow through porous media

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



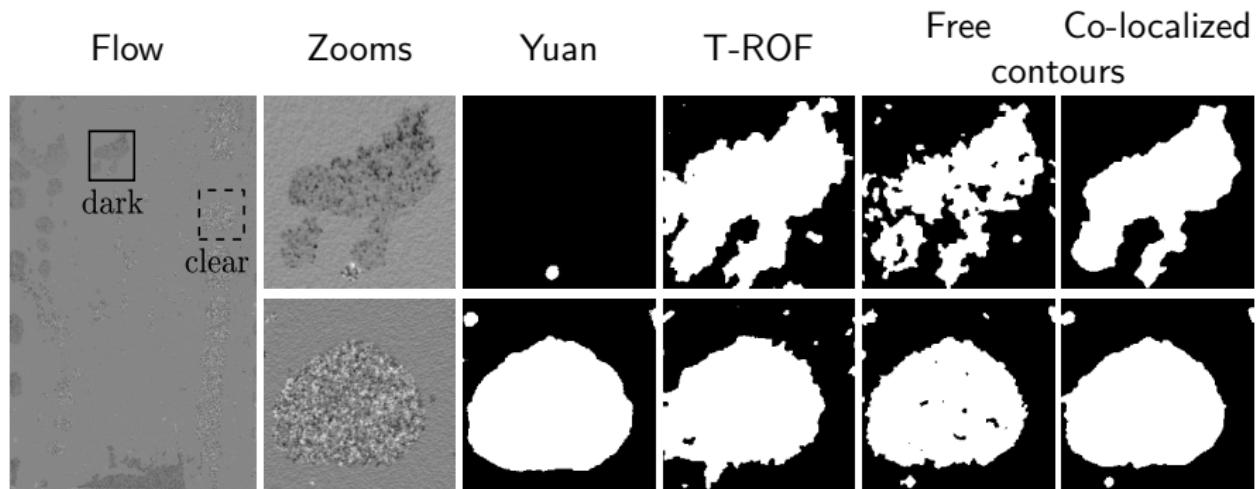
# Multiphase flow through porous media

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

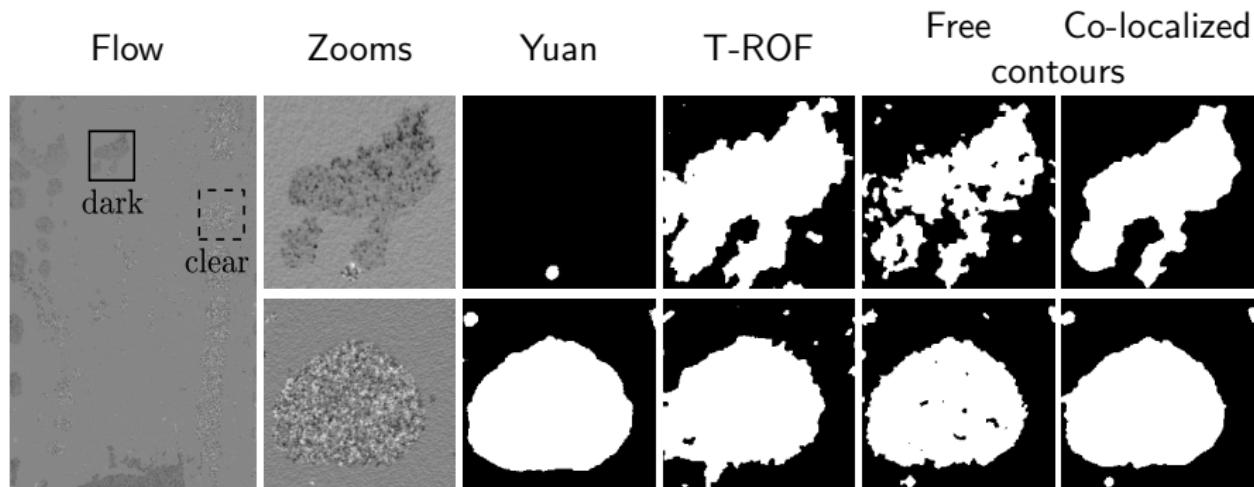


- 1600 × 1100 pixels
- video: ~ 1000 images
- phase diagram: ~ 10 flow rates

Low activity:  $Q_G = 300\text{mL/min}$  -  $Q_L = 300\text{mL/min}$



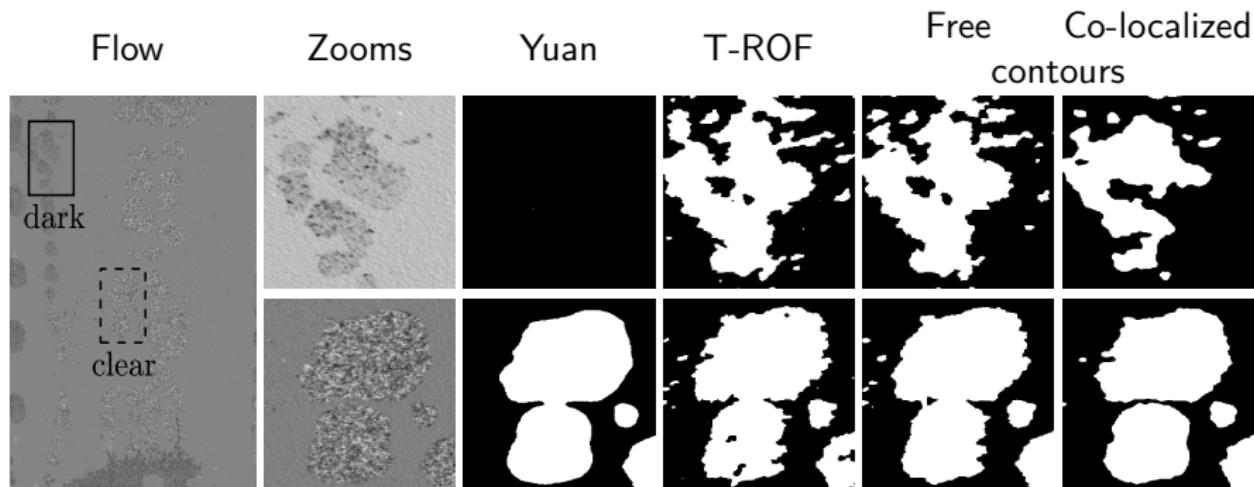
Low activity:  $Q_G = 300\text{mL/min}$  -  $Q_L = 300\text{mL/min}$



Liquid:  $h_L = 0.4$        $\sigma_{\text{dark}}^2 = 10^{-2}$

Gas:       $h_G = 0.9$        $\left| \begin{array}{l} \sigma_{\text{dark}}^2 = 10^{-2} \quad (\text{dark bubbles}) \\ \sigma_{\text{clear}}^2 = 10^{-1} \quad (\text{clear bubbles}) \end{array} \right.$

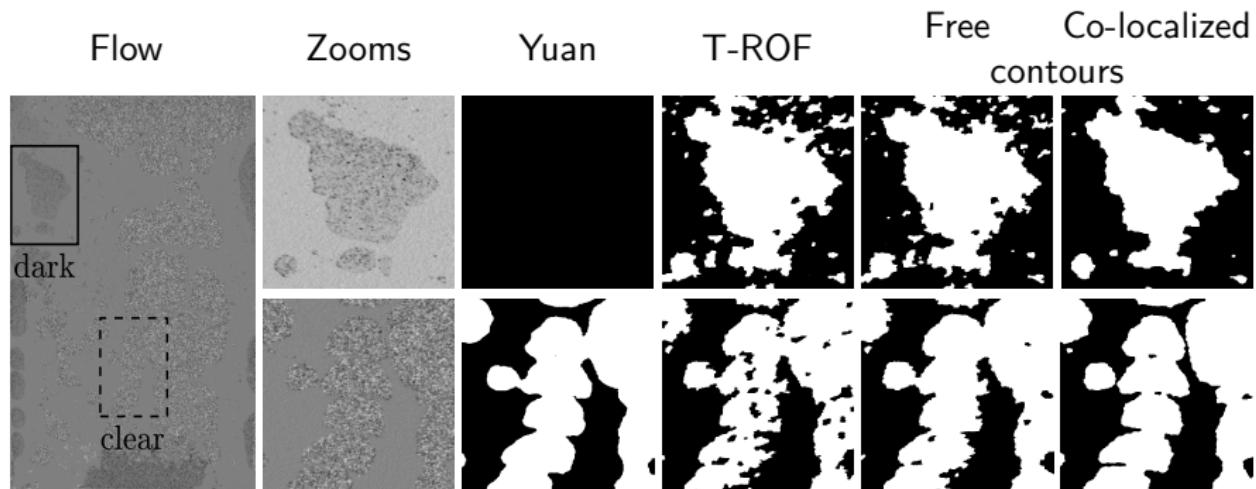
Transition:  $Q_G = 400\text{mL/min}$  -  $Q_L = 700\text{mL/min}$



Liquid:  $h_L = 0.4$        $\sigma_{\text{dark}}^2 = 10^{-2}$

Gas:       $h_G = 0.9$        $\left| \begin{array}{l} \sigma_{\text{dark}}^2 = 10^{-2} \quad (\text{dark bubbles}) \\ \sigma_{\text{clear}}^2 = 10^{-1} \quad (\text{clear bubbles}). \end{array} \right.$

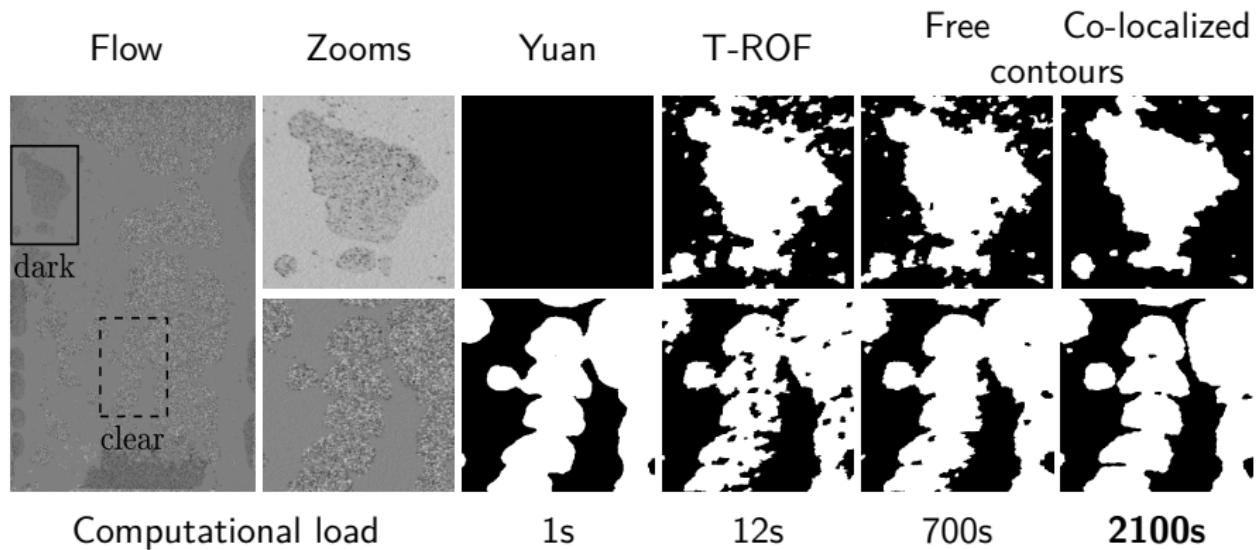
High activity:  $Q_G = 1200\text{mL/min}$  -  $Q_L = 300\text{mL/min}$



Liquid:  $h_L = 0.4$        $\sigma_{\text{dark}}^2 = 10^{-2}$

Gas:       $h_G = 0.9$        $\left| \begin{array}{l} \sigma_{\text{dark}}^2 = 10^{-2} \quad (\text{dark bubbles}) \\ \sigma_{\text{clear}}^2 = 10^{-1} \quad (\text{clear bubbles}). \end{array} \right.$

High activity:  $Q_G = 1200\text{mL/min}$  -  $Q_L = 300\text{mL/min}$



Liquid:  $h_L = 0.4$        $\sigma_{\text{dark}}^2 = 10^{-2}$

Gas:       $h_G = 0.9$        $\left| \begin{array}{l} \sigma_{\text{dark}}^2 = 10^{-2} \quad (\text{dark bubbles}) \\ \sigma_{\text{clear}}^2 = 10^{-1} \quad (\text{clear bubbles}). \end{array} \right.$

## Regularization parameters selection

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

## Regularization parameters selection

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

Lin. reg.  $\hat{\mathbf{h}}^{\text{LR}}$

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



## Regularization parameters selection

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

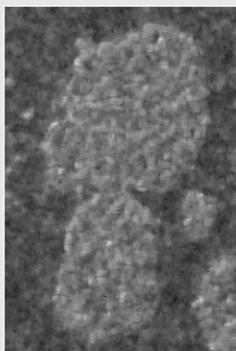
Lin. reg.  $\hat{\mathbf{h}}^{\text{LR}}$

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



Co-localized contours estimate  $\hat{\mathbf{h}}^C$

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



too small

## Regularization parameters selection

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

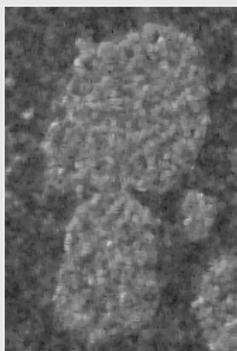
Lin. reg.  $\hat{\mathbf{h}}^{\text{LR}}$

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



Co-localized contours estimate  $\hat{\mathbf{h}}^C$

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



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



too small

too large

## Regularization parameters selection

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

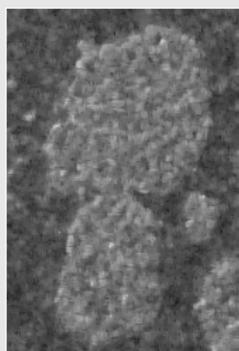
Lin. reg.  $\hat{\mathbf{h}}^{\text{LR}}$

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

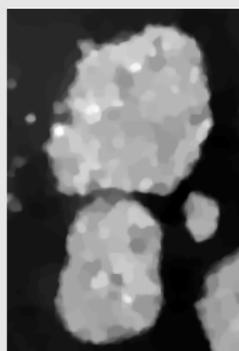


Co-localized contours estimate  $\hat{\mathbf{h}}^C$

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



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



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



too small

optimal

too large

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

## Parameter tuning (Grid search)

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

$\mathbf{h}$ : discriminant,  $\mathbf{v}$ : auxiliary

## Parameter tuning (Grid search)

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

$\mathbf{h}$ : discriminant,  $\mathbf{v}$ : auxiliary

$\bar{\mathbf{h}}$ : true regularity

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

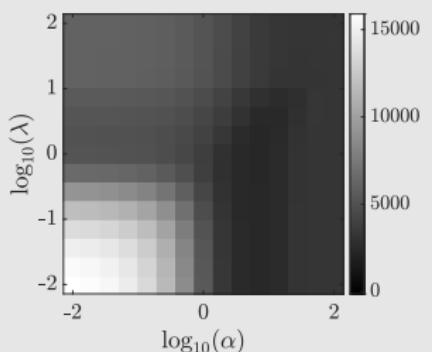
## Parameter tuning (Grid search)

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

$\mathbf{h}$ : discriminant,  $\mathbf{v}$ : auxiliary

$\bar{\mathbf{h}}$ : true regularity

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



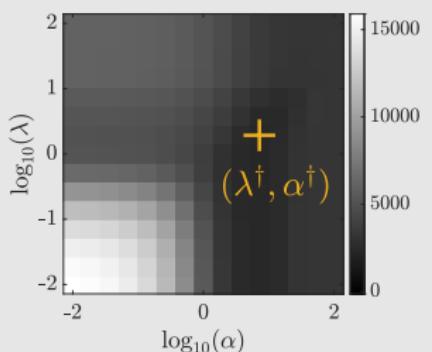
## Parameter tuning (Grid search)

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

$\mathbf{h}$ : discriminant,  $\mathbf{v}$ : auxiliary

$\bar{\mathbf{h}}$ : true regularity

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



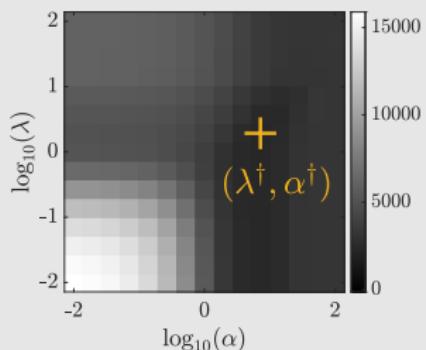
## Parameter tuning (Grid search)

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

$\mathbf{h}$ : discriminant,  $\mathbf{v}$ : auxiliary

$\bar{\mathbf{h}}$ : true regularity

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



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

?

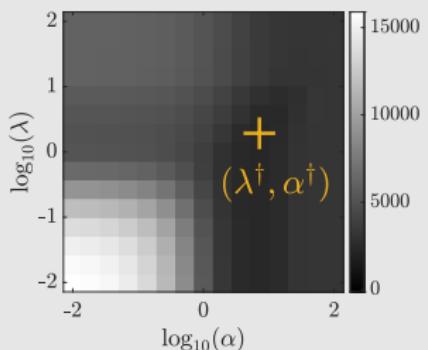
## Parameter tuning (Grid search)

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

$\mathbf{h}$ : discriminant,  $\mathbf{v}$ : auxiliary

$\bar{\mathbf{h}}$ : true regularity

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



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

?

Stein Unbiased Risk Estimate  
(SURE)

## *Stein Unbiased Risk Estimate (Principle)*

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

## Stein Unbiased Risk Estimate (Principle)

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

**Parametric estimator**  $(\mathbf{y}; \lambda) \mapsto \hat{\mathbf{x}}(\mathbf{y}; \lambda)$

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

## Stein Unbiased Risk Estimate (Principe)

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

**Parametric estimator**  $(\mathbf{y}; \lambda) \mapsto \hat{\mathbf{x}}(\mathbf{y}; \lambda)$

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

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

# Stein Unbiased Risk Estimate (Principe)

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

**Parametric estimator**  $(\mathbf{y}; \lambda) \mapsto \hat{\mathbf{x}}(\mathbf{y}; \lambda)$

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

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

**Theorem** (Stein, 1981)

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

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

$$\begin{aligned} \widehat{R}(\mathbf{y}; \lambda) &\triangleq \|\hat{\mathbf{x}}(\mathbf{y}; \lambda) - \mathbf{y}\|^2 + 2\rho^2 \operatorname{tr}(\partial_{\mathbf{y}} \hat{\mathbf{x}}(\mathbf{y}; \lambda)) - \rho^2 P \\ &\implies R(\lambda) = \mathbb{E}_{\boldsymbol{\zeta}} [\widehat{R}(\mathbf{y}; \lambda)]. \end{aligned}$$

## Generalized Stein Unbiased Risk Estimate

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

**E.g. the estimators  $\hat{\mathbf{h}}(\mathcal{L}; \lambda, \alpha)$  with free or co-localized contours**

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

$$\Phi : (\mathbf{h}, \mathbf{v}) \mapsto \{\log(a)\mathbf{h} + \mathbf{v}\}_a \quad \begin{array}{|c|c|c|c|c|c|c|c|} \hline & \cdot & \cdot & \cdot & \textcolor{darkgreen}{\cdot} & \cdot & \cdot & \cdot \\ \hline \cdot & \cdot \\ \hline \end{array} \quad \Pi : (\mathbf{h}, \mathbf{v}) \mapsto (\mathbf{h}, \mathbf{0})$$

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

# Generalized Stein Unbiased Risk Estimate

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

**E.g. the estimators  $\hat{\mathbf{h}}(\mathcal{L}; \lambda, \alpha)$  with free or co-localized contours**

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

$$\Phi : (\mathbf{h}, \mathbf{v}) \mapsto \{\log(a)\mathbf{h} + \mathbf{v}\}_a \quad \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|} \hline & \cdot \\ \hline \cdot & \cdot \\ \hline \end{array} \quad \Pi : (\mathbf{h}, \mathbf{v}) \mapsto (\mathbf{h}, \mathbf{0})$$

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

**Theorem** (Pascal, 2020)

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

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

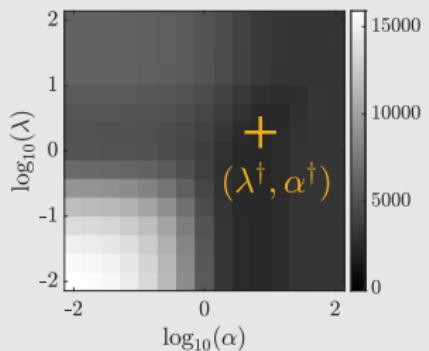
$$\begin{aligned} \hat{R}(\Lambda) &\triangleq \|\mathbf{A}(\Phi \hat{\mathbf{x}}(\mathbf{y}; \Lambda) - \mathbf{y})\|^2 + 2\text{tr} \left( \mathcal{S} \mathbf{A}^\top \Pi \partial_{\mathbf{y}} \hat{\mathbf{x}}(\mathbf{y}; \Lambda) \right) - \text{tr} \left( \mathbf{A} \mathcal{S} \mathbf{A}^\top \right) \\ &\implies R_{\Pi}(\Lambda) = \mathbb{E}_{\zeta} [\hat{R}(\Lambda)]. \end{aligned}$$

## Parameter tuning (Grid search)

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

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

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



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

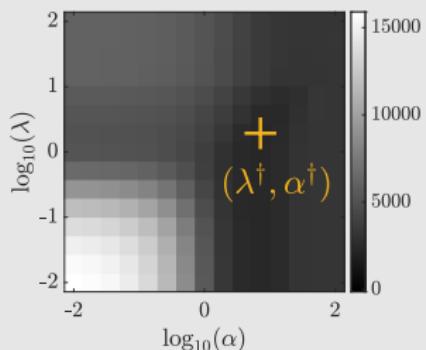
$$\widehat{R}_{\nu, \epsilon}(\mathcal{L}; \lambda, \alpha | \mathcal{S})$$

## Parameter tuning (Grid search)

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

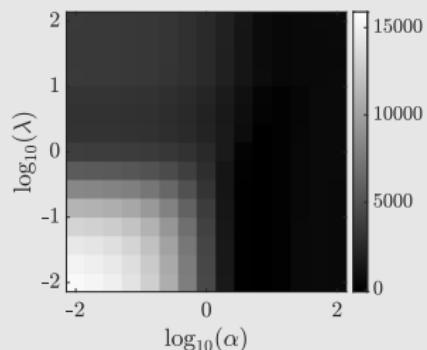
$\bar{\mathbf{h}}$ : true regularity

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



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

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

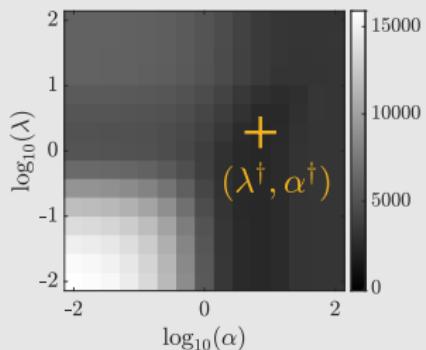


## Parameter tuning (Grid search)

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

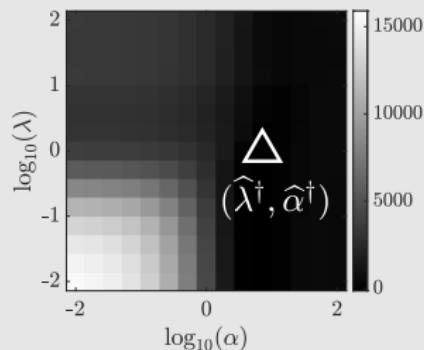
$\bar{\mathbf{h}}$ : true regularity

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



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

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

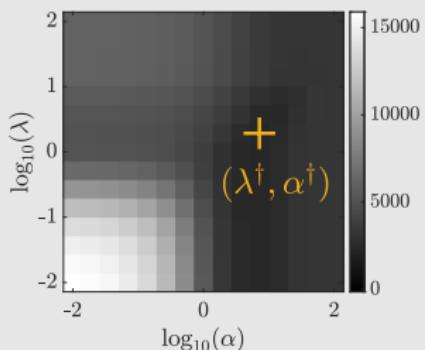


## Parameter tuning (Grid search)

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

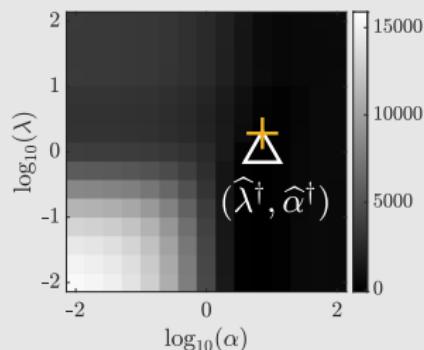
$\bar{\mathbf{h}}$ : true regularity

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



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

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

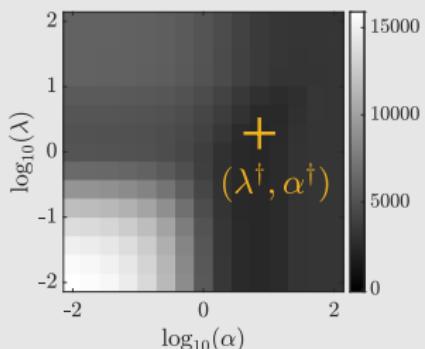


## Parameter tuning (Automatic selection)

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

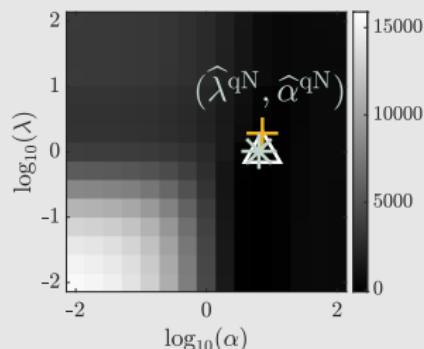
$\bar{\mathbf{h}}$ : true regularity

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



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

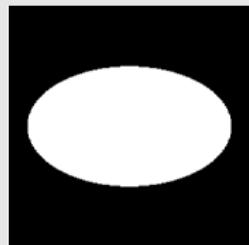
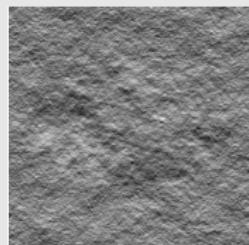
$$\widehat{R}_{\nu, \varepsilon}(\mathcal{L}; \lambda, \alpha | \mathcal{S})$$



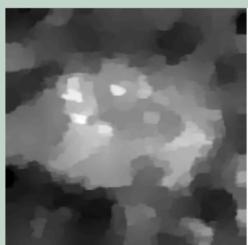
# Automated selection of regularization parameters

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

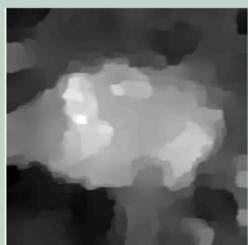
Example



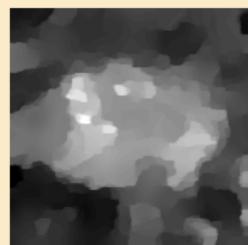
$\hat{\mathbf{h}}^F(\mathcal{L}; \lambda^\dagger, \alpha^\dagger)$   
(grid)



$\hat{\mathbf{h}}^F(\mathcal{L}; \hat{\lambda}^\dagger, \hat{\alpha}^\dagger)$   
(grid)



$\hat{\mathbf{h}}^F(\mathcal{L}; \hat{\lambda}^{qN}, \hat{\alpha}^{qN})$   
(quasi-Newton)



225 calls of the estimator over the grid v.s. 40 for quasi-Newton

## Take home messages

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

## 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

- ▶ 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

- ▶ 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

MATLAB toolbox available: [github.com/bpascal-fr/gsugar](https://github.com/bpascal-fr/gsugar)