

Automated data-driven inverse problem resolution:
Applications in microfluidics and epidemiology

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

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

February 19, 2026

AI hackathon for women in the mathematical sciences

ICMS, Edinburgh, UK

Observation model

$$\mathbf{y} \sim \mathcal{B}(\Phi \bar{\mathbf{x}})$$

- $\mathbf{y} \in \mathbb{R}^P$: degraded observations;
- $\bar{\mathbf{x}} \in \mathbb{R}^N$: unknown quantity of interest;
- $\Phi : \mathbb{R}^N \rightarrow \mathbb{R}^P$: known deformation;
- \mathcal{B} : random measurement noise.

Observation model

$$\mathbf{y} \sim \mathcal{B}(\Phi \bar{\mathbf{x}})$$

- $\mathbf{y} \in \mathbb{R}^P$: degraded observations;
- $\bar{\mathbf{x}} \in \mathbb{R}^N$: unknown quantity of interest;
- $\Phi : \mathbb{R}^N \rightarrow \mathbb{R}^P$: known deformation;
- \mathcal{B} : random measurement noise.

Goal: Estimate $\bar{\mathbf{x}}$



Observation model

$$\mathbf{y} \sim \mathcal{B}(\Phi \bar{\mathbf{x}})$$

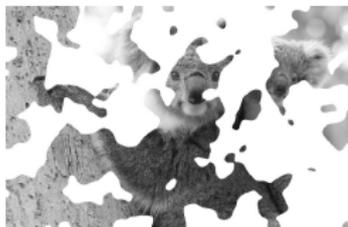
- $\mathbf{y} \in \mathbb{R}^P$: degraded observations;
- $\bar{\mathbf{x}} \in \mathbb{R}^N$: unknown quantity of interest;
- $\Phi : \mathbb{R}^N \rightarrow \mathbb{R}^P$: known deformation;
- \mathcal{B} : random measurement noise.

Goal: Estimate $\bar{\mathbf{x}}$



► ill-conditioned, rank deficient Φ

Inpainting



(Guillemot et al., 2013, *IEEE Sig. Process. Mag.*)

Super-resolution



(Marquina et al., 2008, *J. Sci. Comput.*)

Deblurring



(Pan, 2016, *IEEE Trans. Pattern Anal. Mach. Intell.*)

Observation model

$$\mathbf{y} \sim \mathcal{B}(\Phi \bar{\mathbf{x}})$$

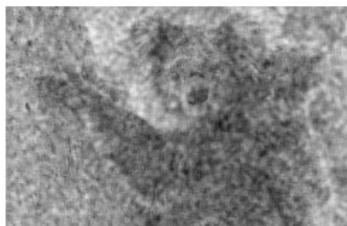
- $\mathbf{y} \in \mathbb{R}^P$: degraded observations;
- $\bar{\mathbf{x}} \in \mathbb{R}^N$: unknown quantity of interest;
- $\Phi : \mathbb{R}^N \rightarrow \mathbb{R}^P$: known deformation;
- \mathcal{B} : random measurement noise.

Goal: Estimate $\bar{\mathbf{x}}$



- ▶ ill-conditioned, rank deficient Φ
- ▶ correlated, data-dependent \mathcal{B}

Correlated



(Pascal et al., 2021, *J. Math. Imaging Vis.*)

Data-dependent



(Luisier et al., 2010, *IEEE Trans. Image Process.*)

Multiplicative



(Shama, 2016, *Appl. Math. Comput.*)

The variational framework: penalized log-likelihood

Variational estimator

$$\hat{\mathbf{x}}(\mathbf{y}; \lambda) \in \underset{\mathbf{x} \in \mathbb{R}^N}{\text{Argmin}} \mathcal{D}(\mathbf{y}, \Phi \mathbf{x})$$

- $\mathcal{D}(\mathbf{y}; \cdot) = -\log \mathbb{P}(\mathbf{y}|\cdot)$: negative log-likelihood

$$\text{Ex: } \mathcal{D}(\mathbf{y}; \mathbf{x}) = \|\mathbf{y} - \mathbf{x}\|_2^2$$

No regularization



$$\mathcal{R} = 0$$

The variational framework: penalized log-likelihood

Variational estimator

$$\hat{\mathbf{x}}(\mathbf{y}; \lambda) \in \underset{\mathbf{x} \in \mathbb{R}^N}{\text{Argmin}} \mathcal{D}(\mathbf{y}, \Phi \mathbf{x}) + \lambda \mathcal{R}(\mathbf{x})$$

- $\mathcal{D}(\mathbf{y}; \cdot) = -\log \mathbb{P}(\mathbf{y}|\cdot)$: negative log-likelihood
- \mathcal{R} : regularization term encoding a priori knowledge

$$\text{Ex: } \mathcal{D}(\mathbf{y}; \mathbf{x}) = \|\mathbf{y} - \mathbf{x}\|_2^2$$

$$\text{Ex: } \mathcal{R}(\mathbf{x}) = \|\mathbf{D}_1 \mathbf{x}\|_q^q$$

(Giovannelli & Idier, 2015, *Wiley*)

No regularization



$$\mathcal{R} = 0$$

Smooth



$$\mathcal{R}(\mathbf{x}) = \|\mathbf{D}_1 \mathbf{x}\|_2^2$$

(Tikhonov et al., 1977, *Wiley*)

Piecewise constant



$$\mathcal{R}(\mathbf{x}) = \|\mathbf{D}_1 \mathbf{x}\|_1$$

(Rudin et al., 1992, *Physica D*)

Fine-tuning of the regularization parameter

Example: $\hat{\mathbf{x}}(\mathbf{y}; \lambda) \in \underset{\mathbf{x} \in \mathbb{R}^N}{\text{Argmin}} \|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda \|\mathbf{D}_1 \mathbf{x}\|_2^2$ (Tikhonov)

Fine-tuning of the regularization parameter

Example: $\hat{\mathbf{x}}(\mathbf{y}; \lambda) \in \underset{\mathbf{x} \in \mathbb{R}^N}{\text{Argmin}} \|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda \|\mathbf{D}_1 \mathbf{x}\|_2^2$ (Tikhonov)

$\lambda = 1$



not enough

Fine-tuning of the regularization parameter

Example: $\hat{\mathbf{x}}(\mathbf{y}; \lambda) \in \underset{\mathbf{x} \in \mathbb{R}^N}{\text{Argmin}} \|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda \|\mathbf{D}_1 \mathbf{x}\|_2^2$ (Tikhonov)

$\lambda = 1$



not enough

$\lambda = 25$



too regularized

Fine-tuning of the regularization parameter

Example: $\hat{\mathbf{x}}(\mathbf{y}; \lambda) \in \underset{\mathbf{x} \in \mathbb{R}^N}{\text{Argmin}} \|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda \|\mathbf{D}_1 \mathbf{x}\|_2^2$ (Tikhonov)

$\lambda = 1$



not enough

$\lambda^\dagger = 5$



optimal

$\lambda = 25$



too regularized

Hyperparameter selection: bilevel optimization

Fine-tuning of the regularization parameter

Example: $\hat{\mathbf{x}}(\mathbf{y}; \lambda) \in \underset{\mathbf{x} \in \mathbb{R}^N}{\text{Argmin}} \|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda \|\mathbf{D}_1 \mathbf{x}\|_2^2$ (Tikhonov)

$\lambda = 1$



not enough

$\lambda^\dagger = 5$



optimal

$\lambda = 25$



too regularized

Oracle-based hyperparameter selection

$$\lambda^\dagger \in \underset{\lambda \in \Lambda}{\text{Argmin}} \mathcal{O}(\mathbf{y}; \lambda)$$

Hyperparameter selection: bilevel optimization

Fine-tuning of the regularization parameter

Example: $\hat{\mathbf{x}}(\mathbf{y}; \lambda) \in \underset{\mathbf{x} \in \mathbb{R}^N}{\text{Argmin}} \|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda \|\mathbf{D}_1 \mathbf{x}\|_2^2$ (Tikhonov)

$\lambda = 1$



not enough

$\lambda^\dagger = 5$



optimal

$\lambda = 25$



too regularized

Oracle-based hyperparameter selection

$$\lambda^\dagger \in \underset{\lambda \in \Lambda}{\text{Argmin}} \mathcal{O}(\mathbf{y}; \lambda)$$

- golden case: $\mathcal{O}(\mathbf{y}; \lambda) = \|\hat{\mathbf{x}}(\mathbf{y}; \lambda) - \bar{\mathbf{x}}\|^2 \implies$ efficient bi-level (\mathbf{x}, λ) minimization

Hyperparameter selection: bilevel optimization

Fine-tuning of the regularization parameter

Example: $\hat{\mathbf{x}}(\mathbf{y}; \lambda) \in \underset{\mathbf{x} \in \mathbb{R}^N}{\text{Argmin}} \|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda \|\mathbf{D}_1 \mathbf{x}\|_2^2$ (Tikhonov)

$\lambda = 1$



not enough

$\lambda^\dagger = 5$



optimal

$\lambda = 25$



too regularized

Oracle-based hyperparameter selection

$$\lambda^\dagger \in \underset{\lambda \in \Lambda}{\text{Argmin}} \mathcal{O}(\mathbf{y}; \lambda)$$

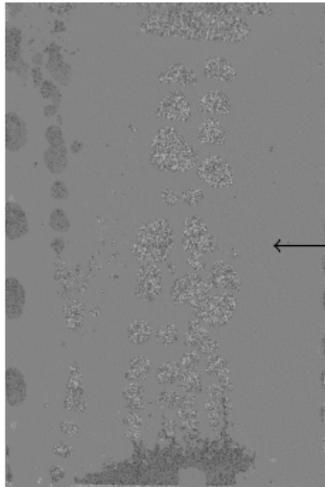
- golden case: $\mathcal{O}(\mathbf{y}; \lambda) = \|\hat{\mathbf{x}}(\mathbf{y}; \lambda) - \bar{\mathbf{x}}\|^2 \implies$ efficient bi-level (\mathbf{x}, λ) minimization
- practical case: ground truth $\bar{\mathbf{x}}$ **not available!** \implies data-driven $\mathcal{O}(\mathbf{y}; \lambda)$

Image processing:

Texture segmentation

Multiphase flow through porous media

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

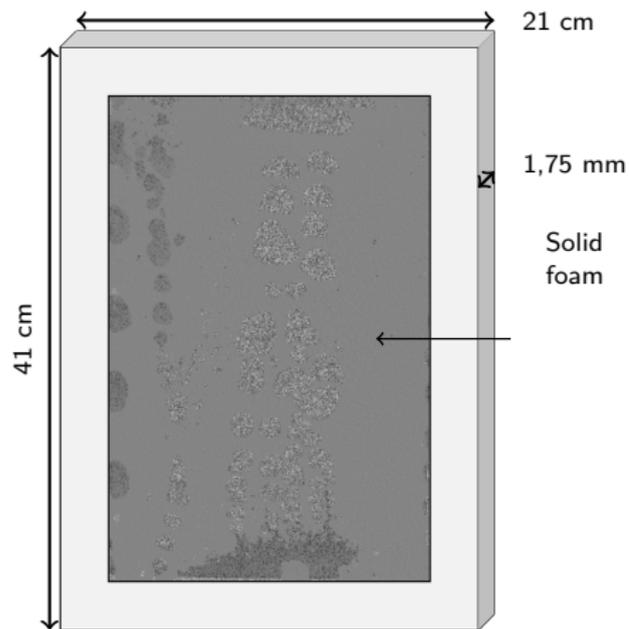


Solid
foam



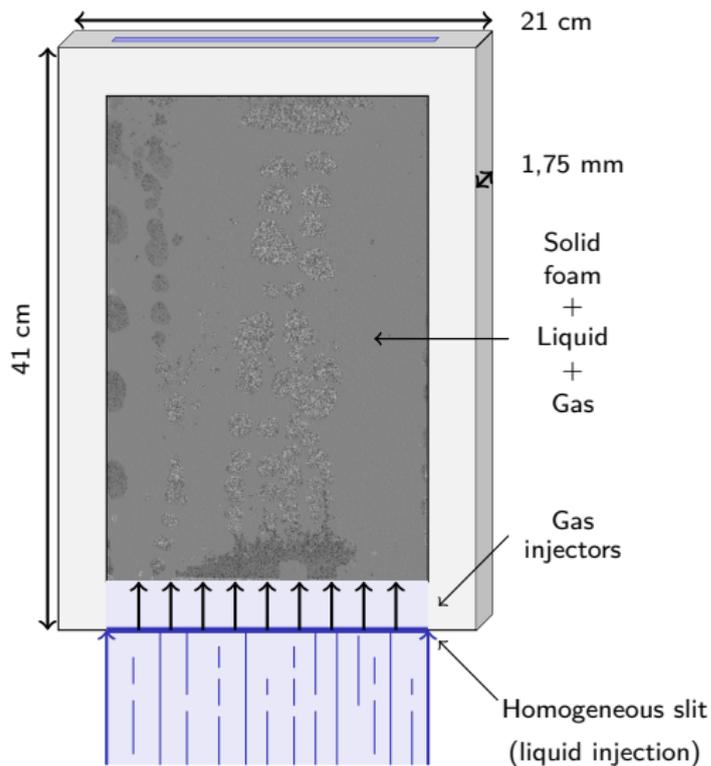
Multiphase flow through porous media

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



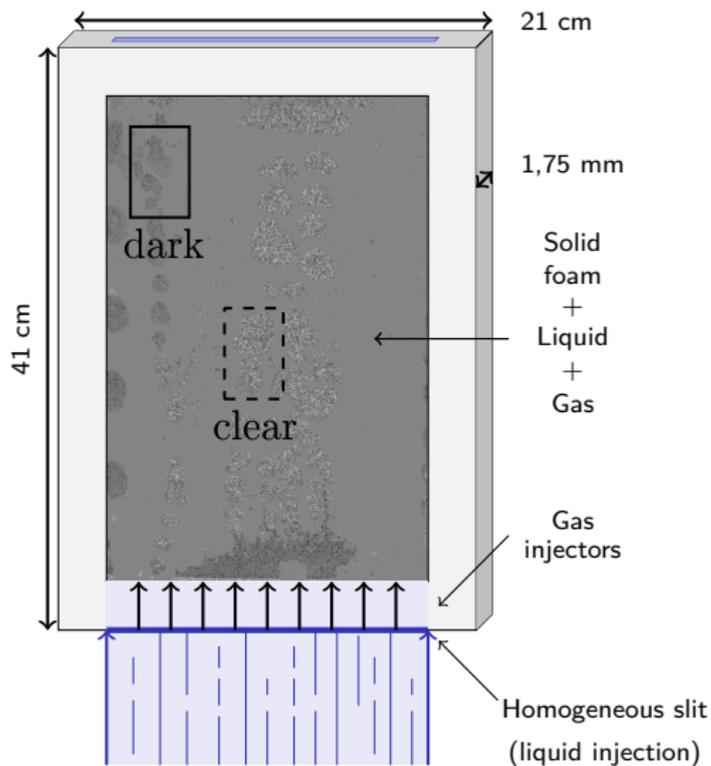
Multiphase flow through porous media

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



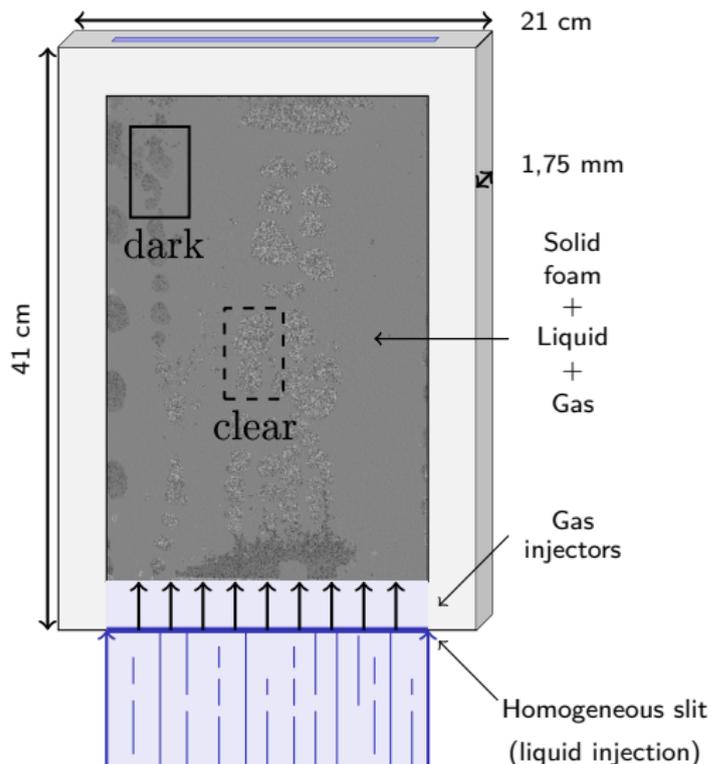
Multiphase flow through porous media

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

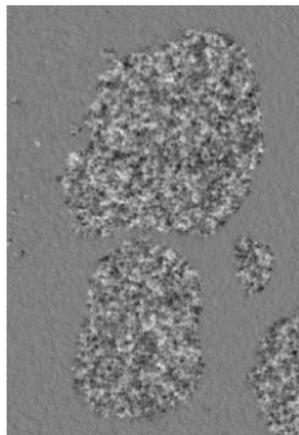


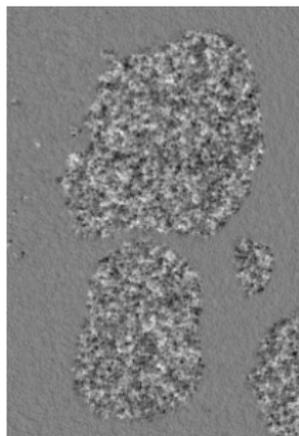
Multiphase flow through porous media

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



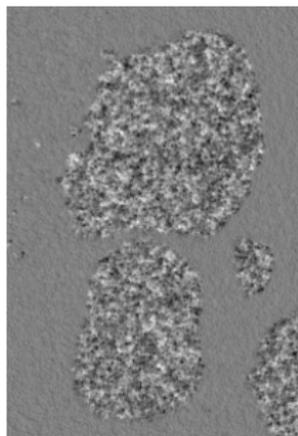
- 1600 × 1100 pixels
- video: ~ 1000 images
- phase diagram: ~ 10 flow rates





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

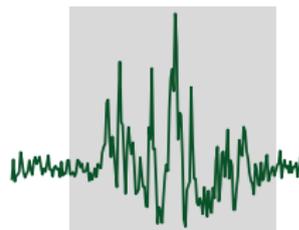
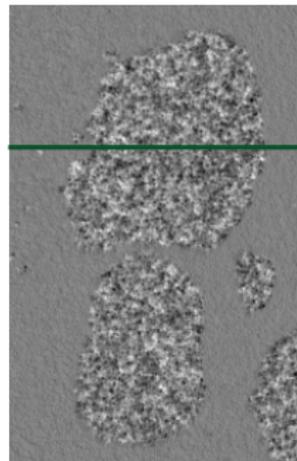
$$\Omega = \Omega_1 \sqcup \dots \sqcup \Omega_K$$



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

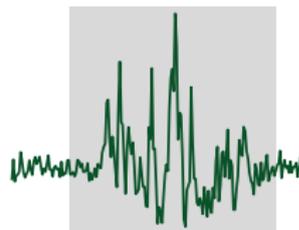
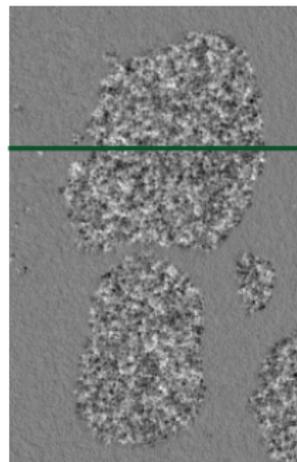
$$\Omega = \Omega_1 \sqcup \dots \sqcup \Omega_K$$

Piecewise monofractal model



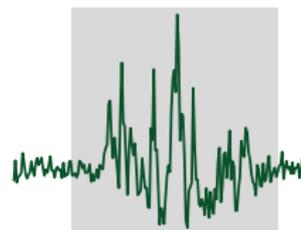
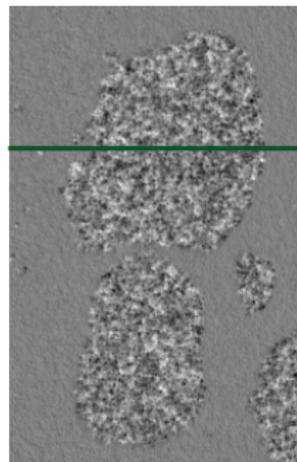
Fractal attributes

- variance σ^2 *amplitude of variations*



Fractal attributes

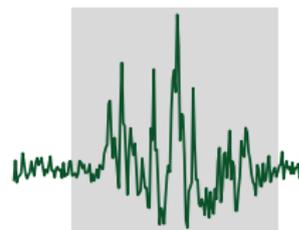
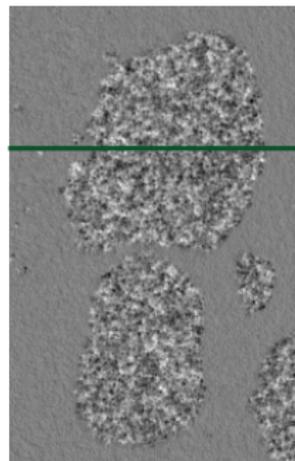
- variance σ^2 *amplitude of variations*
- local regularity h *scale invariance*



Fractal attributes

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

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



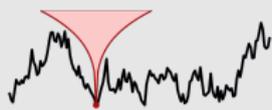
Fractal attributes

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

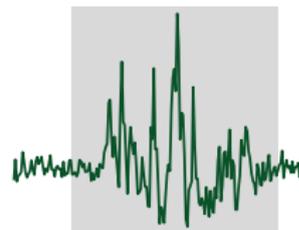
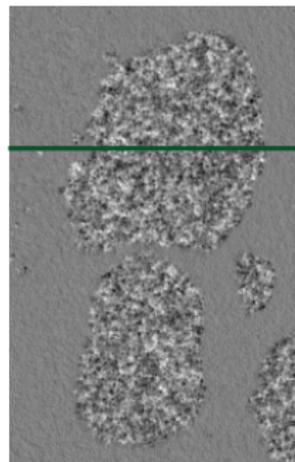
$$|f(x) - f(y)| \leq \sigma(x)|x - y|^{h(x)}$$



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



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



Piecewise monofractal model

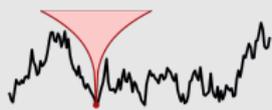
Fractal attributes

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

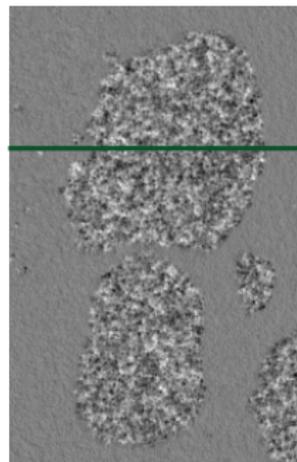
$$|f(x) - f(y)| \leq \sigma(x)|x - y|^{h(x)}$$



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

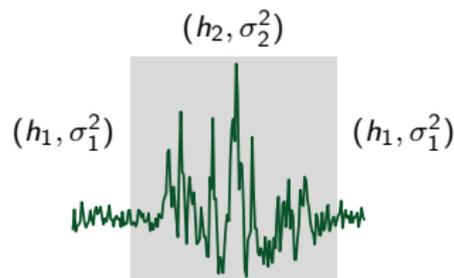


$$h(x) \equiv h_2 = 0.3$$



Segmentation

- ▶ σ^2 and h piecewise constant



Piecewise monofractal model

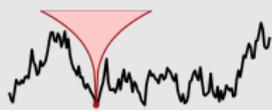
Fractal attributes

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

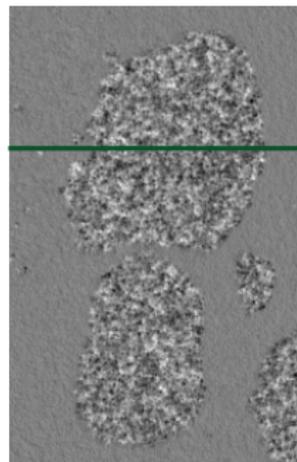
$$|f(x) - f(y)| \leq \sigma(x)|x - y|^{h(x)}$$



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

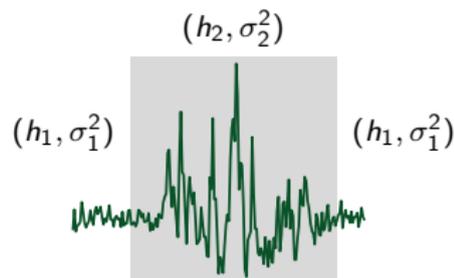


$$h(x) \equiv h_2 = 0.3$$

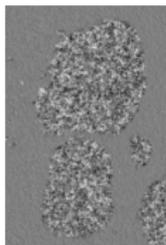


Segmentation

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



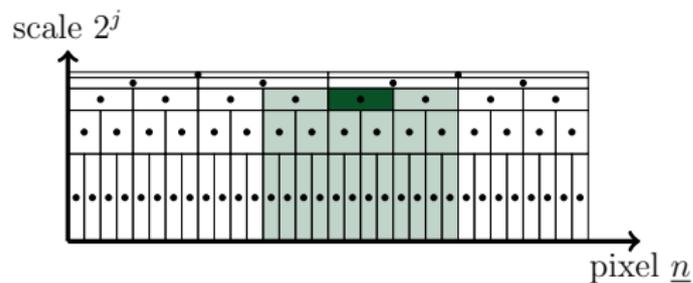
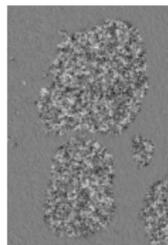
Textured image



Multiscale analysis

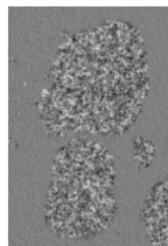
Textured image

Local maximum of wavelet coefficients: \mathcal{L}_a .



Multiscale analysis

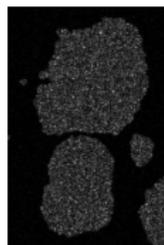
Textured image



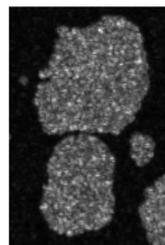
Local maximum of wavelet coefficients: \mathcal{L}_a .

Scale

$a = 2^1$

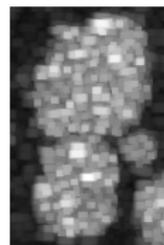


$a = 2^2$

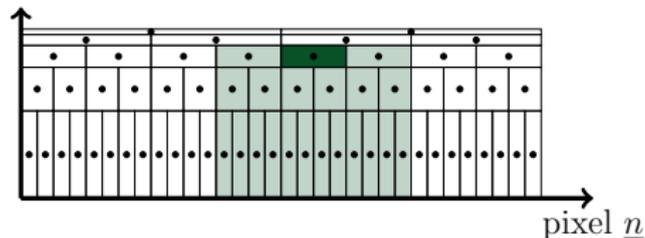


...

$a = 2^5$

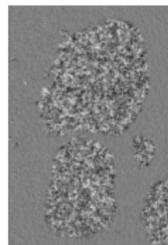


scale 2^j



Multiscale analysis

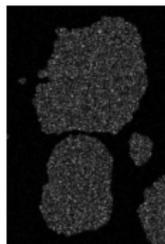
Textured image



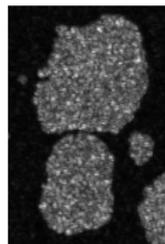
Scale

Local maximum of wavelet coefficients: \mathcal{L}_a .

$a = 2^1$

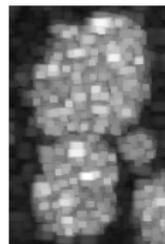


$a = 2^2$



...

$a = 2^5$

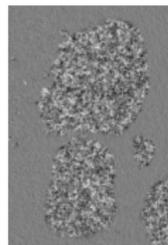


Proposition (Jaffard, 2004, *Proc. Symp. Pure Math.*; Wendt et al., 2009, *Signal Process.*)

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

Multiscale analysis

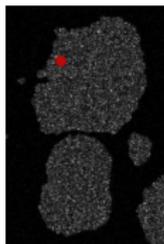
Textured image



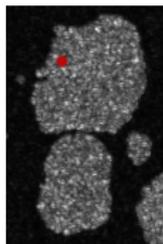
Scale

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

$a = 2^1$

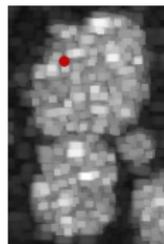


$a = 2^2$



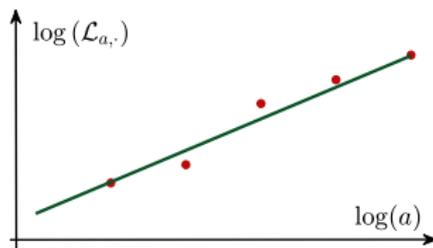
...

$a = 2^5$



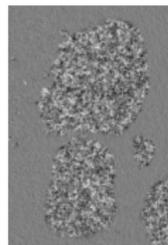
Proposition (Jaffard, 2004, *Proc. Symp. Pure Math.*; Wendt et al., 2009, *Signal Process.*)

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



Multiscale analysis

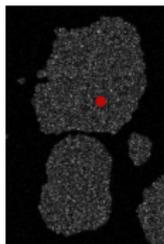
Textured image



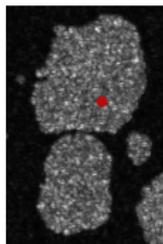
Scale

Local maximum of wavelet coefficients: \mathcal{L}_a .

$a = 2^1$

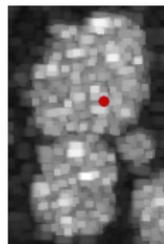


$a = 2^2$



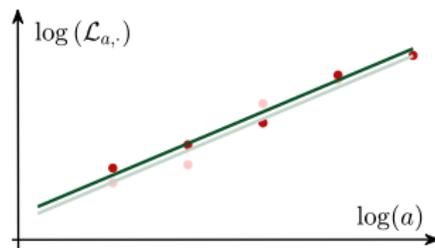
...

$a = 2^5$



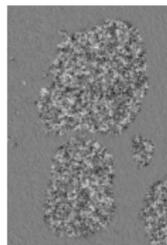
Proposition (Jaffard, 2004, *Proc. Symp. Pure Math.*; Wendt et al., 2009, *Signal Process.*)

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



Multiscale analysis

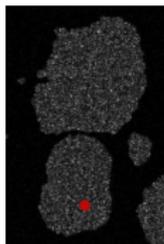
Textured image



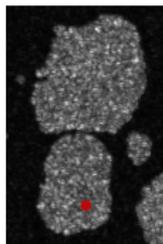
Scale

Local maximum of wavelet coefficients: \mathcal{L}_a .

$a = 2^1$

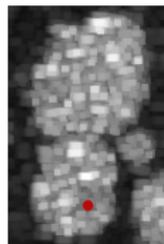


$a = 2^2$



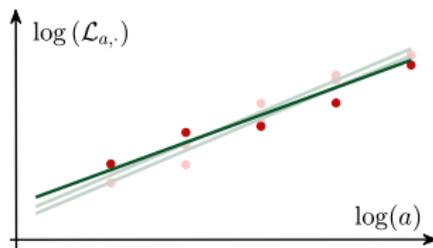
...

$a = 2^5$



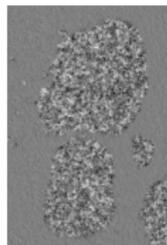
Proposition (Jaffard, 2004, *Proc. Symp. Pure Math.*; Wendt et al., 2009, *Signal Process.*)

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



Multiscale analysis

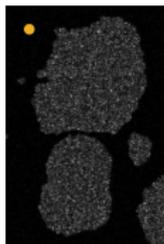
Textured image



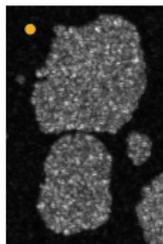
Scale

Local maximum of wavelet coefficients: \mathcal{L}_a .

$a = 2^1$

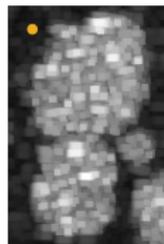


$a = 2^2$



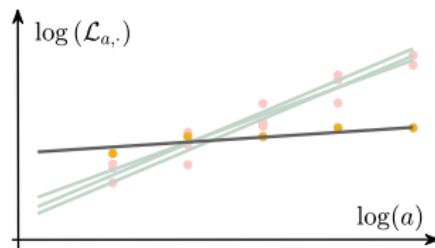
...

$a = 2^5$



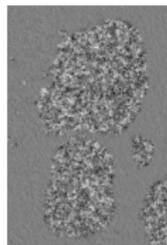
Proposition (Jaffard, 2004, *Proc. Symp. Pure Math.*; Wendt et al., 2009, *Signal Process.*)

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



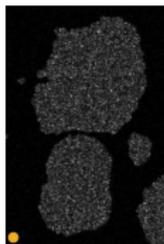
Multiscale analysis

Textured image

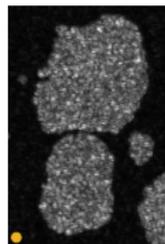


Scale

$a = 2^1$

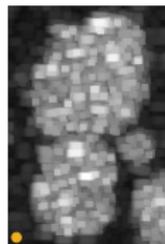


$a = 2^2$



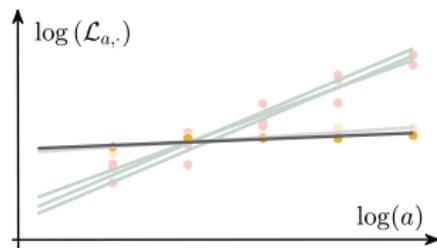
...

$a = 2^5$



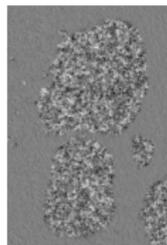
Proposition (Jaffard, 2004, *Proc. Symp. Pure Math.*; Wendt et al., 2009, *Signal Process.*)

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



Multiscale analysis

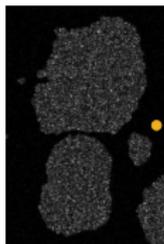
Textured image



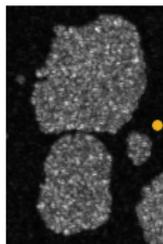
Scale

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

$a = 2^1$

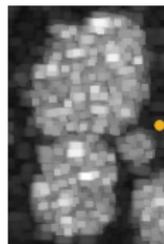


$a = 2^2$



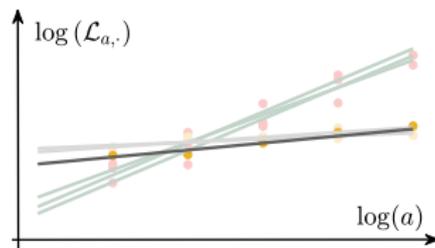
...

$a = 2^5$



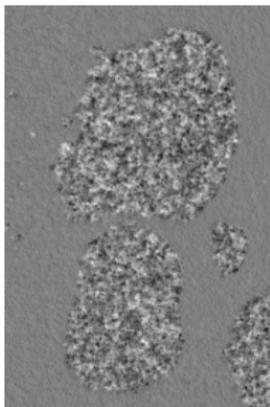
Proposition (Jaffard, 2004, *Proc. Symp. Pure Math.*; Wendt et al., 2009, *Signal Process.*)

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



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

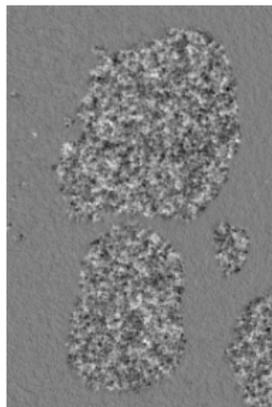
Textured image



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

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

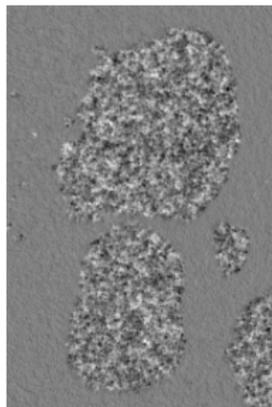
Textured image



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

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

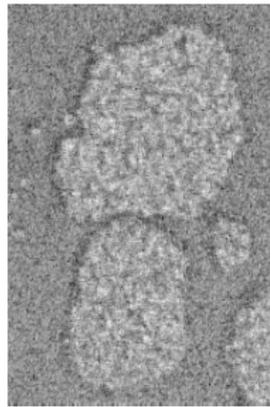
Textured image



Local regularity $\widehat{\mathbf{h}}^{\text{LR}}$



Local power $\widehat{\mathbf{v}}^{\text{LR}}$

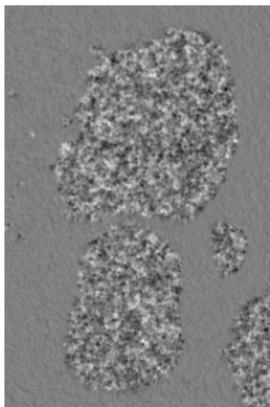


Direct punctual estimation

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

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

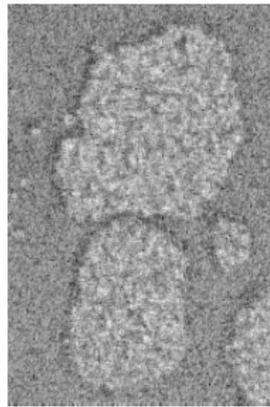
Textured image



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



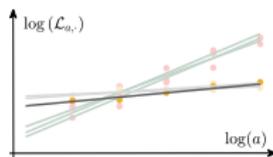
Local power $\hat{\mathbf{v}}^{\text{LR}}$



→ large estimation variance

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

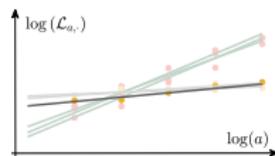
→ fidelity to the log-linear model



Functionals with either free or co-localized contours

$$\sum_a \frac{\|\log \mathcal{L}_{a,\cdot} - \log(a)\mathbf{h} - \mathbf{v}\|^2}{\text{Least-Squares}} + \lambda \frac{Q(\mathbf{D}_1\mathbf{h}, \mathbf{D}_1\mathbf{v}; \alpha)}{\text{Total Variation}}$$

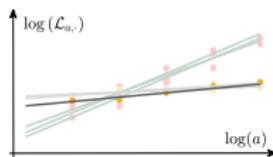
→ fidelity to the log-linear model → favors piecewise constancy



Functionals with either free or co-localized contours

$$\underset{h, v}{\text{minimize}} \quad \sum_a \frac{\|\log \mathcal{L}_{a,\cdot} - \log(a)h - v\|^2}{\text{Least-Squares}} + \lambda \frac{Q(\mathbf{D}_1 h, \mathbf{D}_1 v; \alpha)}{\text{Total Variation}}$$

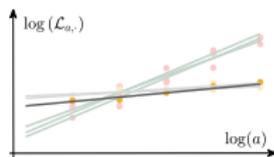
\rightarrow fidelity to the log-linear model \rightarrow favors piecewise constancy



Functionals with either free or co-localized contours

$$\underset{h, v}{\text{minimize}} \sum_a \frac{\|\log \mathcal{L}_{a, \cdot} - \log(a) \mathbf{h} - \mathbf{v}\|^2}{\text{Least-Squares}} + \lambda \frac{Q(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha)}{\text{Total Variation}}$$

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

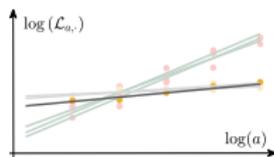


Finite differences $\mathbf{D}_1^{\rightarrow} \mathbf{x}$ (horizontal), $\mathbf{D}_1^{\uparrow} \mathbf{x}$ (vertical) at each pixel

Functionals with either free or co-localized contours

$$\underset{\mathbf{h}, \mathbf{v}}{\text{minimize}} \quad \sum_a \frac{\|\log \mathcal{L}_{a,\cdot} - \log(a)\mathbf{h} - \mathbf{v}\|^2}{\text{Least-Squares}} + \lambda \frac{Q(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha)}{\text{Total Variation}}$$

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



Finite differences $\mathbf{D}_1 \mathbf{x} = [\mathbf{D}_1^{\rightarrow} \mathbf{x}, \mathbf{D}_1^{\uparrow} \mathbf{x}]$

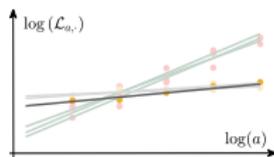
Free: \mathbf{h}, \mathbf{v} are **independently** piecewise constant

$$Q_F(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha) = \alpha \|\mathbf{D}_1 \mathbf{h}\|_{2,1} + \|\mathbf{D}_1 \mathbf{v}\|_{2,1}$$

Functionals with either free or co-localized contours

$$\underset{\mathbf{h}, \mathbf{v}}{\text{minimize}} \sum_a \frac{\|\log \mathcal{L}_{a,\cdot} - \log(a)\mathbf{h} - \mathbf{v}\|^2}{\text{Least-Squares}} + \lambda \frac{\mathcal{Q}(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha)}{\text{Total Variation}}$$

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



Finite differences $\mathbf{D}_1 \mathbf{x} = [\mathbf{D}_1^{\rightarrow} \mathbf{x}, \mathbf{D}_1^{\uparrow} \mathbf{x}]$

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

$$\mathcal{Q}_F(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha) = \alpha \|\mathbf{D}_1 \mathbf{h}\|_{2,1} + \|\mathbf{D}_1 \mathbf{v}\|_{2,1}$$

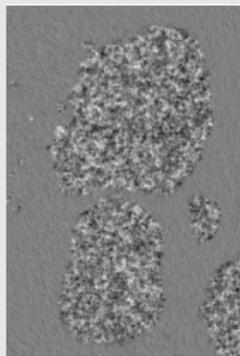
Co-localized: \mathbf{h}, \mathbf{v} are **concomitantly** piecewise constant

$$\mathcal{Q}_C(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha) = \|[\alpha \mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}]\|_{2,1}$$

Segmentation *via* iterated thresholding

$$\underset{\mathbf{h}, \mathbf{v}}{\text{minimize}} \quad \sum_a \underbrace{\|\log \mathcal{L}_{a,\cdot} - \log(a)\mathbf{h} - \mathbf{v}\|^2}_{\text{Least-Squares}} + \lambda \underbrace{Q(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha)}_{\text{Total Variation}}$$

Textured image



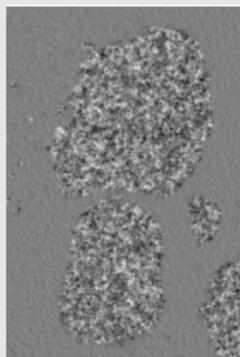
Lin. reg. $\hat{\mathbf{h}}^{\text{LR}}$



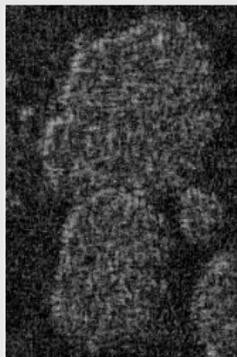
Segmentation *via* iterated thresholding

$$\underset{\mathbf{h}, \mathbf{v}}{\text{minimize}} \quad \sum_a \frac{\|\log \mathcal{L}_{a,\cdot} - \log(a)\mathbf{h} - \mathbf{v}\|^2}{\text{Least-Squares}} + \lambda \frac{Q(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha)}{\text{Total Variation}}$$

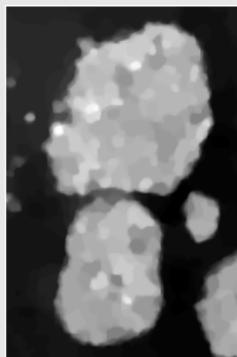
Textured image



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



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



Threshold
estimate[†] $T\hat{\mathbf{h}}^{\text{C}}$



[†](Cai et al., 2013, *J. Sci. Comput.*)

$$\left(\hat{\mathbf{h}}, \hat{\mathbf{v}}\right) (\mathcal{L}; \lambda, \alpha) = \operatorname{argmin}_{\mathbf{h}, \mathbf{v}} \sum_a \|\log \mathcal{L}_{a,\cdot} - \log(a)\mathbf{h} - \mathbf{v}\|^2 + \lambda Q(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha)$$

$$(\hat{\mathbf{h}}, \hat{\mathbf{v}})(\mathcal{L}; \lambda, \alpha) = \operatorname{argmin}_{\mathbf{h}, \mathbf{v}} \sum_a \|\log \mathcal{L}_{a,\cdot} - \log(a)\mathbf{h} - \mathbf{v}\|^2 + \lambda Q(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha)$$

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

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



$$(\hat{\mathbf{h}}, \hat{\mathbf{v}})(\mathcal{L}; \lambda, \alpha) = \underset{\mathbf{h}, \mathbf{v}}{\operatorname{argmin}} \sum_a \|\log \mathcal{L}_{a,\cdot} - \log(a)\mathbf{h} - \mathbf{v}\|^2 + \lambda Q(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha)$$

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

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



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

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



too small

Regularization parameters selection

$$(\hat{\mathbf{h}}, \hat{\mathbf{v}})(\mathcal{L}; \lambda, \alpha) = \operatorname{argmin}_{\mathbf{h}, \mathbf{v}} \sum_a \|\log \mathcal{L}_{a,\cdot} - \log(a)\mathbf{h} - \mathbf{v}\|^2 + \lambda Q(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha)$$

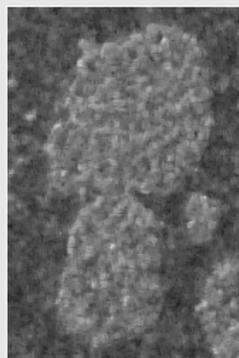
Lin. reg. $\hat{\mathbf{h}}^{\text{LR}}$

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



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

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



too small

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



too large

Regularization parameters selection

$$(\hat{\mathbf{h}}, \hat{\mathbf{v}})(\mathcal{L}; \lambda, \alpha) = \operatorname{argmin}_{\mathbf{h}, \mathbf{v}} \sum_a \|\log \mathcal{L}_{a,\cdot} - \log(a)\mathbf{h} - \mathbf{v}\|^2 + \lambda Q(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha)$$

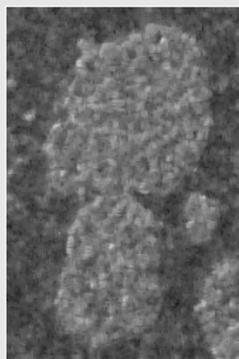
Lin. reg. $\hat{\mathbf{h}}^{\text{LR}}$

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



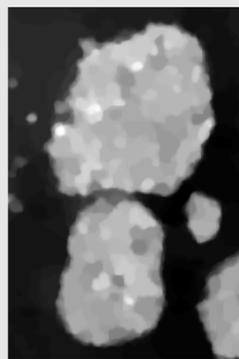
Co-localized contours estimate $\hat{\mathbf{h}}^{\text{C}}$

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



too small

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



optimal

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



too large

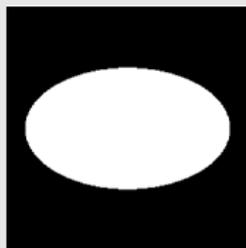
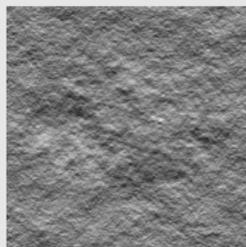
What *optimal* means? How to determine λ^\dagger and α^\dagger ?

$$\left(\hat{\mathbf{h}}, \hat{\mathbf{v}}\right) (\mathcal{L}; \lambda, \alpha) = \operatorname{argmin}_{\mathbf{h}, \mathbf{v}} \sum_a \|\log \mathcal{L}_{a,\cdot} - \log(a) \mathbf{h} - \mathbf{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha)$$

Parameter tuning (Grid search)

$$\left(\hat{\mathbf{h}}, \hat{\mathbf{v}}\right) (\mathcal{L}; \lambda, \alpha) = \underset{\mathbf{h}, \mathbf{v}}{\operatorname{argmin}} \sum_a \|\log \mathcal{L}_{a, \cdot} - \log(a) \mathbf{h} - \mathbf{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha)$$

Example



Parameter tuning (Grid search)

$$\left(\hat{\mathbf{h}}, \hat{\mathbf{v}}\right)(\mathcal{L}; \lambda, \alpha) = \underset{\mathbf{h}, \mathbf{v}}{\operatorname{argmin}} \sum_a \left\| \log \mathcal{L}_{a..} - \log(a) \mathbf{h} - \mathbf{v} \right\|^2 + \lambda \mathcal{Q}(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha)$$

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

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

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

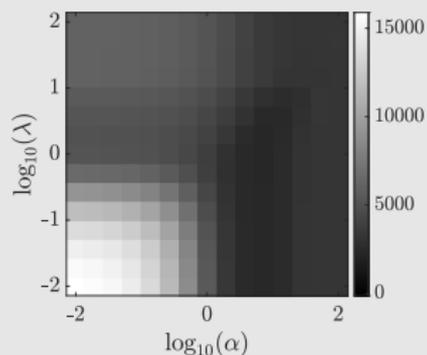
Parameter tuning (Grid search)

$$\left(\hat{\mathbf{h}}, \hat{\mathbf{v}}\right)(\mathcal{L}; \lambda, \alpha) = \underset{\mathbf{h}, \mathbf{v}}{\operatorname{argmin}} \sum_a \|\log \mathcal{L}_{a..} - \log(a)\mathbf{h} - \mathbf{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha)$$

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

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

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



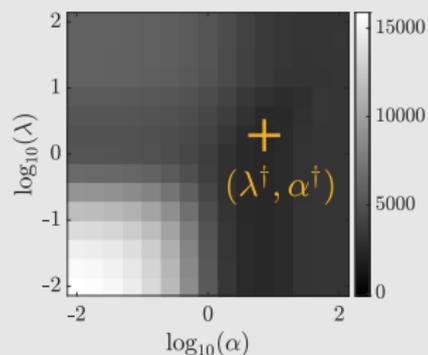
Parameter tuning (Grid search)

$$\left(\hat{\mathbf{h}}, \hat{\mathbf{v}}\right)(\mathcal{L}; \lambda, \alpha) = \underset{\mathbf{h}, \mathbf{v}}{\operatorname{argmin}} \sum_a \left\| \log \mathcal{L}_{a..} - \log(a) \mathbf{h} - \mathbf{v} \right\|^2 + \lambda Q(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha)$$

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

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

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



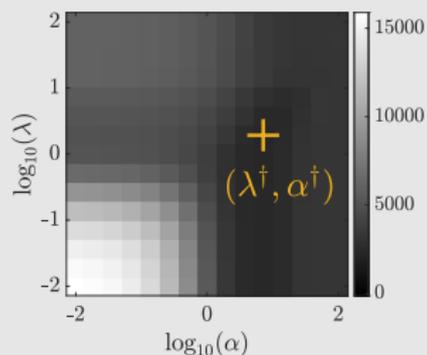
Parameter tuning (Grid search)

$$\left(\hat{\mathbf{h}}, \hat{\mathbf{v}}\right)(\mathcal{L}; \lambda, \alpha) = \underset{\mathbf{h}, \mathbf{v}}{\operatorname{argmin}} \sum_a \left\| \log \mathcal{L}_{a,\cdot} - \log(a) \mathbf{h} - \mathbf{v} \right\|^2 + \lambda Q(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha)$$

\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!

?

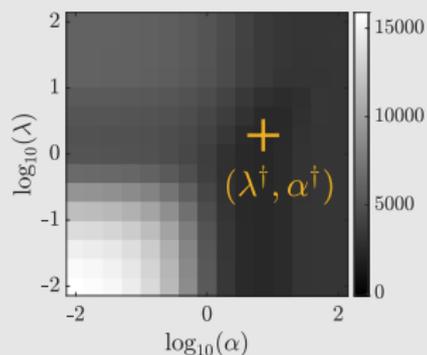
Parameter tuning (Grid search)

$$\left(\hat{\mathbf{h}}, \hat{\mathbf{v}}\right)(\mathcal{L}; \lambda, \alpha) = \underset{\mathbf{h}, \mathbf{v}}{\operatorname{argmin}} \sum_a \left\| \log \mathcal{L}_{a,\cdot} - \log(a) \mathbf{h} - \mathbf{v} \right\|^2 + \lambda Q(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha)$$

\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 (Principe)

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

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

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

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

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

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

$\bar{\mathbf{x}}$ unknown

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

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

Theorem (Stein, 1981, *Ann. Stat.*)

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. $\mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$.

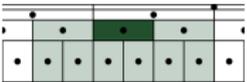
$$\begin{aligned} \hat{R}(\mathbf{y}; \lambda) &\triangleq \|\hat{\mathbf{x}}(\mathbf{y}; \lambda) - \mathbf{y}\|^2 + 2\rho^2 \operatorname{tr}(\partial_{\mathbf{y}} \hat{\mathbf{x}}(\mathbf{y}; \lambda)) - \rho^2 P \\ &\implies R(\lambda) = \mathbb{E}_{\boldsymbol{\zeta}} [\hat{R}(\mathbf{y}; \lambda)].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

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

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

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

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

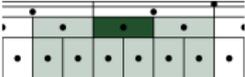
Projected estimation error $R_{\Pi}(\Lambda) \triangleq \mathbb{E}_{\zeta} \|\Pi \hat{\mathbf{x}}(\mathbf{y}; \Lambda) - \Pi \bar{\mathbf{x}}\|^2$

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

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

$$\Phi : (\mathbf{h}, \mathbf{v}) \mapsto \{\log(a)\mathbf{h} + \mathbf{v}\}_a \quad \Pi : (\mathbf{h}, \mathbf{v}) \mapsto (\mathbf{h}, \mathbf{0})$$


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

Theorem (Pascal et al., 2021, *J. Math. Imaging Vis.*)

Let $(\mathbf{y}; \Lambda) \mapsto \hat{\mathbf{x}}(\mathbf{y}; \Lambda)$ be 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. $\mathcal{N}(\mathbf{0}, \mathcal{S})$.

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

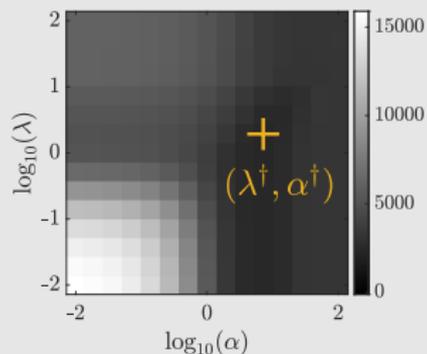
$$\implies R_{\Pi}(\Lambda) = \mathbb{E}_{\zeta} [\hat{R}(\Lambda)].$$

Parameter tuning (Grid search)

$$\left(\hat{\mathbf{h}}, \hat{\mathbf{v}}\right)(\mathcal{L}; \lambda, \alpha) = \operatorname{argmin}_{\mathbf{h}, \mathbf{v}} \sum_a \left\| \log \mathcal{L}_{a..} - \log(a) \mathbf{h} - \mathbf{v} \right\|^2 + \lambda Q(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha)$$

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

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



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

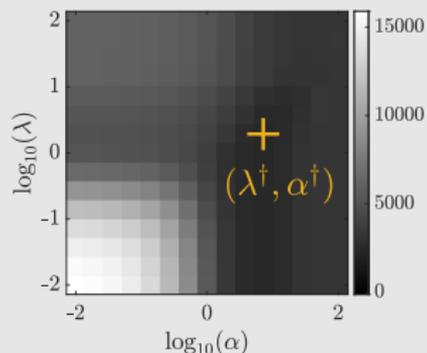
$$\hat{R}_{\nu, \epsilon}(\mathcal{L}; \lambda, \alpha | \mathcal{S})$$

Parameter tuning (Grid search)

$$\left(\hat{\mathbf{h}}, \hat{\mathbf{v}}\right)(\mathcal{L}; \lambda, \alpha) = \operatorname{argmin}_{\mathbf{h}, \mathbf{v}} \sum_a \left\| \log \mathcal{L}_{a,\cdot} - \log(a) \mathbf{h} - \mathbf{v} \right\|^2 + \lambda Q(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha)$$

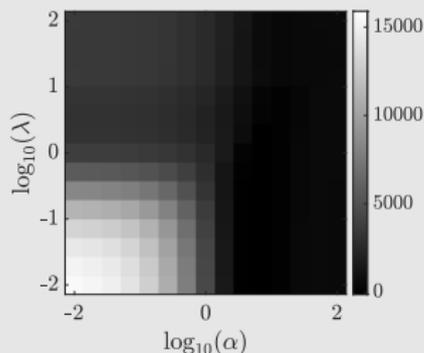
$\bar{\mathbf{h}}$: true regularity

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



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

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

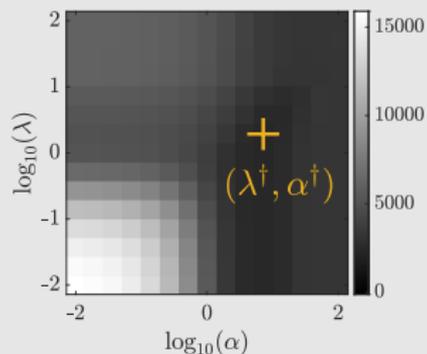


Parameter tuning (Grid search)

$$\left(\hat{\mathbf{h}}, \hat{\mathbf{v}}\right) (\mathcal{L}; \lambda, \alpha) = \operatorname{argmin}_{\mathbf{h}, \mathbf{v}} \sum_a \left\| \log \mathcal{L}_{a,\cdot} - \log(a) \mathbf{h} - \mathbf{v} \right\|^2 + \lambda Q(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha)$$

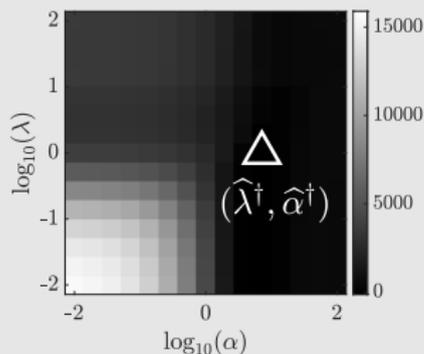
$\bar{\mathbf{h}}$: true regularity

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



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

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

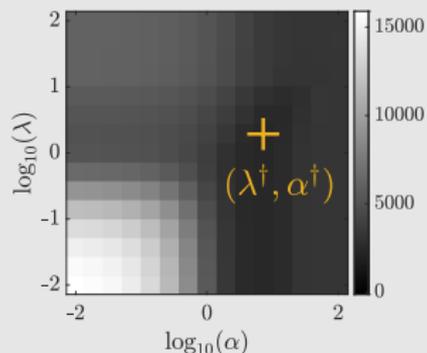


Parameter tuning (Grid search)

$$\left(\hat{\mathbf{h}}, \hat{\mathbf{v}}\right)(\mathcal{L}; \lambda, \alpha) = \operatorname{argmin}_{\mathbf{h}, \mathbf{v}} \sum_a \left\| \log \mathcal{L}_{a,\cdot} - \log(a) \mathbf{h} - \mathbf{v} \right\|^2 + \lambda Q(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha)$$

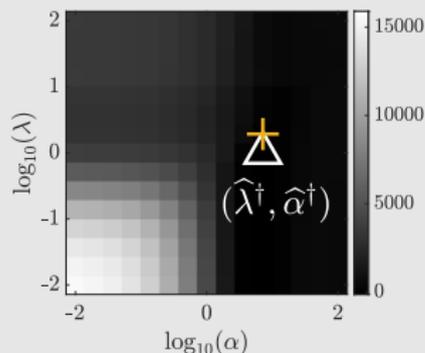
$\bar{\mathbf{h}}$: true regularity

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



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

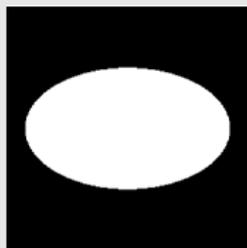
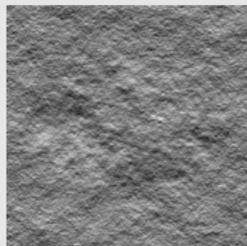
$$\hat{R}_{\nu, \epsilon}(\mathcal{L}; \lambda, \alpha | \mathcal{S})$$



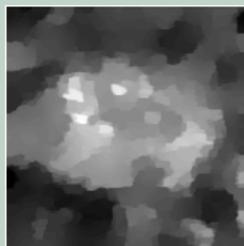
Parameter tuning (Grid search)

$$\left(\hat{\mathbf{h}}, \hat{\mathbf{v}}\right) (\mathcal{L}; \lambda, \alpha) = \underset{\mathbf{h}, \mathbf{v}}{\operatorname{argmin}} \sum_a \|\log \mathcal{L}_{a,\cdot} - \log(a) \mathbf{h} - \mathbf{v}\|^2 + \lambda Q(\mathbf{D}_1 \mathbf{h}, \mathbf{D}_1 \mathbf{v}; \alpha)$$

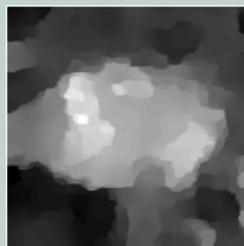
Example



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

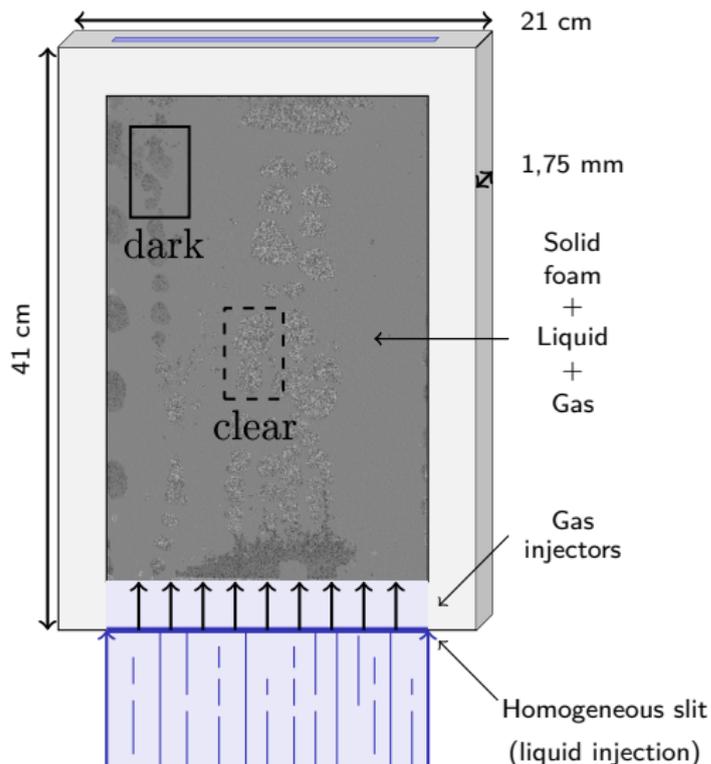


$\hat{\mathbf{h}}^F(\mathcal{L}; \hat{\lambda}^\dagger, \hat{\alpha}^\dagger)$
(grid)



Multiphase flow through porous media

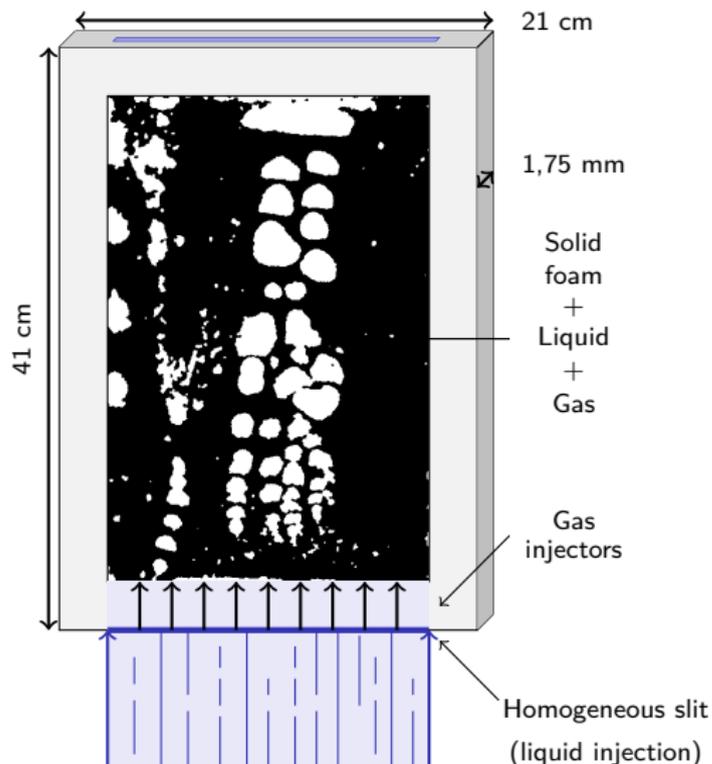
Laboratoire de Physique, ENS Lyon, V. Vidal, T. Busser, (M. Serres, IFPEN)



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

Multiphase flow through porous media

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

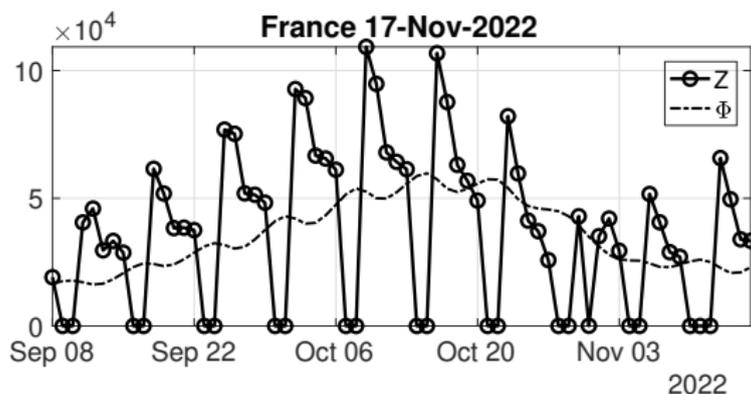


- 1600×1100 pixels
- video: ~ 1000 images
- phase diagram: ~ 10 flow rates

Time series analysis:

Epidemiological indicator estimation

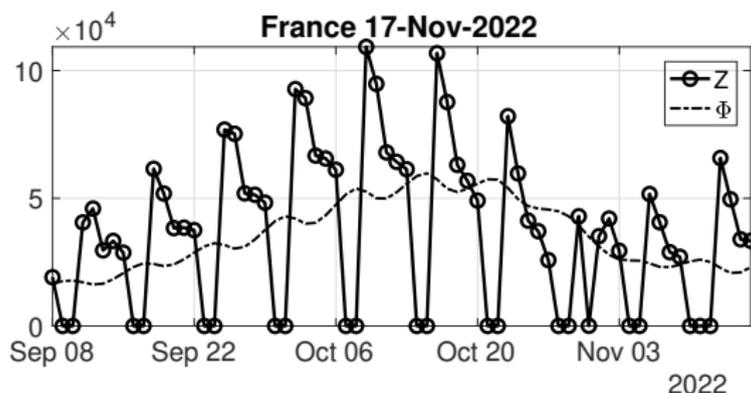
Counts of daily new infections



data from National Health Agencies collected by Johns Hopkins University

\implies number of cases not informative enough: need to capture the **dynamics**

Counts of daily new infections



data from National Health Agencies collected by Johns Hopkins University

⇒ number of cases not informative enough: need to capture the **dynamics**

Design adapted counter measures and evaluate their effectiveness

→ efficient monitoring tools

→ robust to low quality of the data

*epidemiological model,
managing erroneous counts.*

Reproduction number in Cori model

“averaged number of secondary cases generated by a typical infectious individual”

(Cori et al., 2013, *Am. Journal of Epidemiology*; Liu et al., 2018, *PNAS*)

Reproduction number in Cori model

“averaged number of secondary cases generated by a typical infectious individual”

(Cori et al., 2013, *Am. Journal of Epidemiology*; Liu et al., 2018, *PNAS*)

Interpretation: at day t

$R_t > 1$ the virus propagates at exponential speed,

$R_t < 1$ the epidemic shrinks with an exponential decay,

$R_t = 1$ the epidemic is stable.

⇒ one single indicator accounting for the overall pandemic mechanism

Pandemic study: modeling at the service of monitoring

Reproduction number in Cori model

“averaged number of secondary cases generated by a typical infectious individual”

(Cori et al., 2013, *Am. Journal of Epidemiology*; Liu et al., 2018, *PNAS*)

Interpretation: at day t

$R_t > 1$ the virus propagates at exponential speed,

$R_t < 1$ the epidemic shrinks with an exponential decay,

$R_t = 1$ the epidemic is stable.

⇒ one single indicator accounting for the overall pandemic mechanism

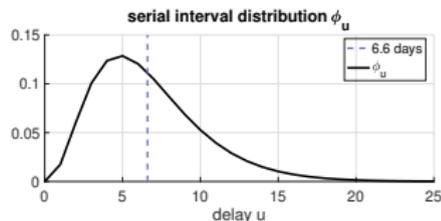
Principle: Z_t new infections at day t

$$\mathbb{E}[Z_t] = R_t \Phi_t, \quad \Phi_t = \sum_{u=1}^{\tau_\Phi} \phi_u Z_{t-u}$$

with Φ_t global “infectiousness” in the population

$\{\phi_u\}_{u=1}^{\tau_\Phi}$ distribution of delay between onset of symptoms in primary and secondary cases

Gamma distribution truncated at 25 days, of mean 6.6 days and standard deviation 3.5 days



Pandemic study: modeling at the service of monitoring

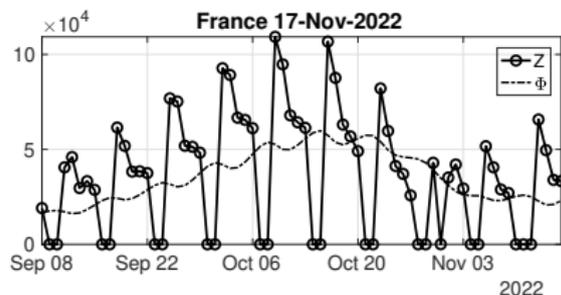
Data: daily counts $\mathbf{Z} = (Z_1, \dots, Z_T)$

Model: **scaled** Poisson distribution

$$\frac{Z_t}{\gamma} \mid Z_{t-\tau_\phi}, \dots, Z_{t-1}; R_t \sim \mathcal{P} \left(\frac{R_t \phi_t}{\gamma} \right)$$

$\gamma > 0$ scaling parameter: controls **variance**

(Pascal & Vaiter, 2025, *Signal Process.*)



Pandemic study: modeling at the service of monitoring

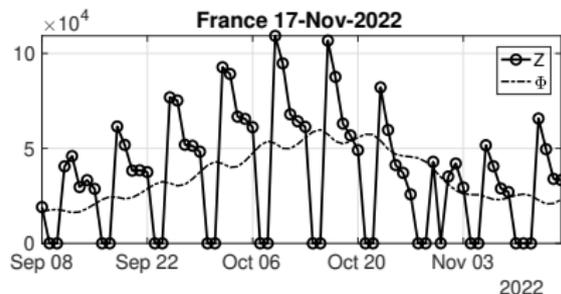
Data: daily counts $\mathbf{Z} = (Z_1, \dots, Z_T)$

Model: **scaled** Poisson distribution

$$\frac{Z_t}{\gamma} \mid Z_{t-\tau_\phi}, \dots, Z_{t-1}; R_t \sim \mathcal{P} \left(\frac{R_t \Phi_t}{\gamma} \right)$$

$\gamma > 0$ scaling parameter: controls **variance**

(Pascal & Vaiter, 2025, *Signal Process.*)



Inverse problem formalism:

$$\frac{\mathbf{Z}}{\gamma} \sim \mathcal{P} \left(\frac{\Phi \mathbf{R}}{\gamma} \right)$$

- $\mathbf{Z} \in \mathbb{N}^T$: reported infection counts,
- $\mathbf{R} = (R_1, \dots, R_T) \in \mathbb{R}_+^T$: daily unknown reproduction number,
- $\Phi = \text{diag}(\Phi_1, \dots, \Phi_T)$: linear operator,
- \mathcal{P} : data-dependent Poisson noise, with scale parameter γ

$$\implies \mathcal{D}(\mathbf{Z}, \Phi \mathbf{R}) = -\log \mathbb{P}(\mathbf{Z} | \mathbf{R})$$

Pandemic study: modeling at the service of monitoring

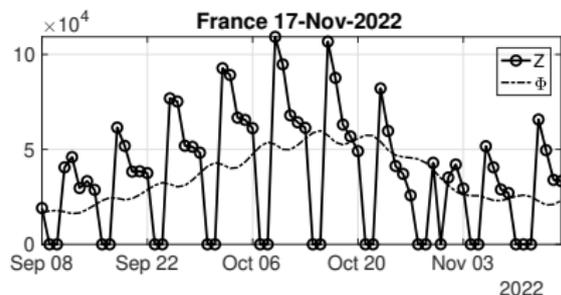
Data: daily counts $\mathbf{Z} = (Z_1, \dots, Z_T)$

Model: **scaled** Poisson distribution

$$\frac{Z_t}{\gamma} \mid \mathbf{Z}_{t-\tau_\phi, \dots, Z_{t-1}}; R_t \sim \mathcal{P} \left(\frac{R_t \phi_t}{\gamma} \right)$$

$\gamma > 0$ scaling parameter: controls **variance**

(Pascal & Vaiter, 2025, *Signal Process.*)



Maximum Likelihood Estimate (MLE)

$$\ln (\mathbb{P}(Z_t \mid \mathbf{Z}_{t-\tau_\phi:t-1}, R_t))$$

$$\stackrel{\text{(def.)}}{=} -\frac{1}{\gamma} d_{\text{KL}}(Z_t \mid R_t \phi_t) \quad (\text{Kullback-Leibler})$$

Pandemic study: modeling at the service of monitoring

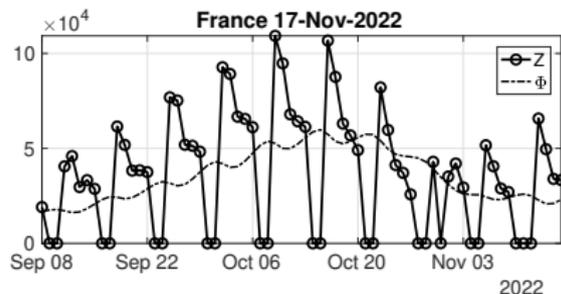
Data: daily counts $\mathbf{Z} = (Z_1, \dots, Z_T)$

Model: **scaled** Poisson distribution

$$\frac{Z_t}{\gamma} \mid Z_{t-\tau_\phi}, \dots, Z_{t-1}; R_t \sim \mathcal{P} \left(\frac{R_t \phi_t}{\gamma} \right)$$

$\gamma > 0$ scaling parameter: controls **variance**

(Pascal & Vaiter, 2025, *Signal Process.*)



Maximum Likelihood Estimate (MLE)

$$\ln (\mathbb{P}(Z_t \mid \mathbf{Z}_{t-\tau_\phi:t-1}, R_t))$$

$$\stackrel{\text{(def.)}}{=} -\frac{1}{\gamma} d_{\text{KL}}(Z_t \mid R_t \phi_t) \quad (\text{Kullback-Leibler})$$

$$\Rightarrow \hat{R}_t^{\text{MLE}} = Z_t / \phi_t = Z_t / \sum_{u=1}^{\tau_\phi} \phi_u Z_{t-u}$$

ratio of moving averages

Pandemic study: modeling at the service of monitoring

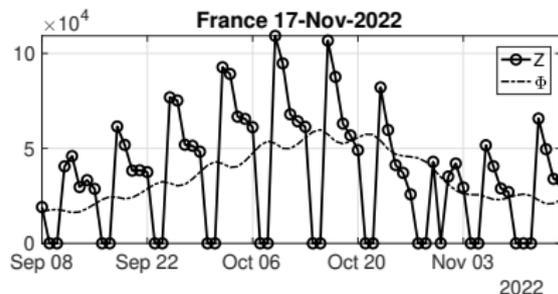
Data: daily counts $\mathbf{Z} = (Z_1, \dots, Z_T)$

Model: **scaled** Poisson distribution

$$\frac{Z_t}{\gamma} \mid Z_{t-\tau_\phi}, \dots, Z_{t-1}; R_t \sim \mathcal{P} \left(\frac{R_t \phi_t}{\gamma} \right)$$

$\gamma > 0$ scaling parameter: controls **variance**

(Pascal & Vaiter, 2025, *Signal Process.*)



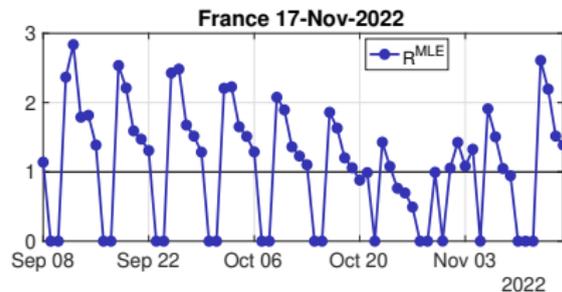
Maximum Likelihood Estimate (MLE)

$$\ln (\mathbb{P}(Z_t \mid \mathbf{Z}_{t-\tau_\phi:t-1}, R_t))$$

$$\stackrel{\text{(def.)}}{=} -\frac{1}{\gamma} d_{\text{KL}}(Z_t \mid R_t \phi_t) \quad (\text{Kullback-Leibler})$$

$$\Rightarrow \hat{R}_t^{\text{MLE}} = Z_t / \phi_t = Z_t / \sum_{u=1}^{\tau_\phi} \phi_u Z_{t-u}$$

ratio of moving averages



- huge variability along time / no local trend
- not robust to pseudo-periodicity / misreported counts

Penalized likelihood: regularization through nonlinear filtering

$$\hat{\mathbf{R}}^{\text{PKL}} = \underset{\mathbf{R} \in \mathbb{R}_+^T}{\operatorname{argmin}} \sum_{t=1}^T \frac{1}{\gamma} d_{\text{KL}}(\mathbf{Z}_t | \mathbf{R}_t \Phi_t) + \lambda \mathcal{R}(\mathbf{R}) \quad (\text{penalized Kullback-Leibler})$$

with $\mathcal{R}(\mathbf{R})$ favoring some temporal regularity

(Abry et al., 2020, *PLoSOne*)

Penalized likelihood: regularization through nonlinear filtering

$$\hat{\mathbf{R}}^{\text{PKL}} = \underset{\mathbf{R} \in \mathbb{R}_+^T}{\operatorname{argmin}} \sum_{t=1}^T \frac{1}{\gamma} d_{\text{KL}}(Z_t | R_t \Phi_t) + \lambda \mathcal{R}(\mathbf{R}) \quad (\text{penalized Kullback-Leibler})$$

with $\mathcal{R}(\mathbf{R})$ favoring some temporal regularity

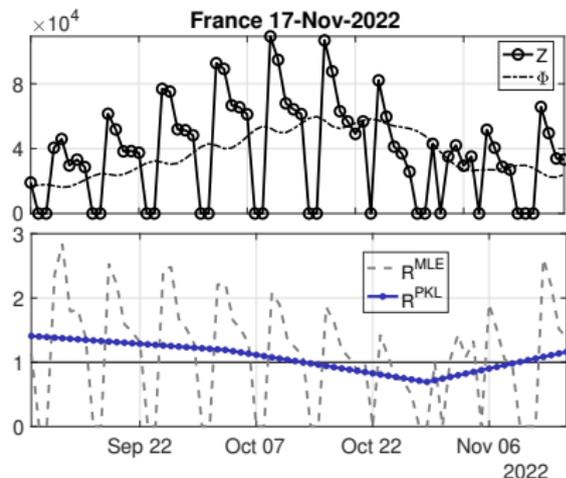
(Abry et al., 2020, *PLOSOne*)

$$\mathcal{R}(\mathbf{R}) = \|\mathbf{D}_2 \mathbf{R}\|_1$$

$$(\mathbf{D}_2 \mathbf{R})_t = R_{t+1} - 2R_t + R_{t-1}$$

2nd order derivative & ℓ_1 -norm

\Rightarrow piecewise linearity



captures global **trend**, **smooth** temporal behavior, **no pseudo-oscillations**

Penalized Kullback-Leibler estimator:

$$\hat{\mathbf{R}}(\mathbf{Z}; \lambda) = \operatorname{argmin}_{\mathbf{R} \in \mathbb{R}_+^T} \sum_{t=1}^T \frac{1}{\gamma} d_{\text{KL}}(Z_t | R_t \Phi_t) + \lambda \|\mathbf{D}_2 \mathbf{R}\|_1$$

with $\mathcal{R}(\mathbf{R})$ favoring some temporal regularity

(Abry et al., 2020, *PLOSOne*)

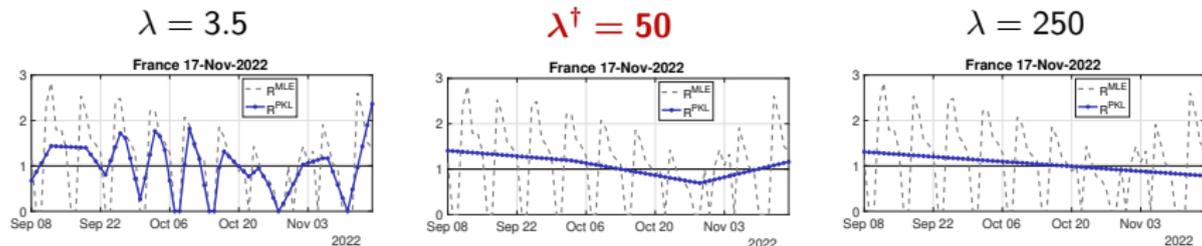
Penalized Kullback-Leibler estimator:

$$\widehat{\mathbf{R}}(\mathbf{Z}; \lambda) = \operatorname{argmin}_{\mathbf{R} \in \mathbb{R}_+^T} \sum_{t=1}^T \frac{1}{\gamma} d_{\text{KL}}(Z_t | R_t \Phi_t) + \lambda \|\mathbf{D}_2 \mathbf{R}\|_1$$

with $\mathcal{R}(\mathbf{R})$ favoring some temporal regularity

(Abry et al., 2020, *PLOSOne*)

Fine tuning of the regularization parameter:



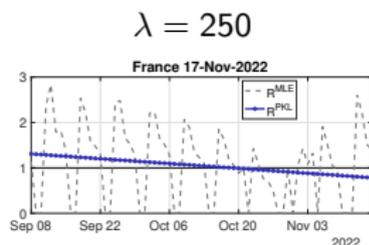
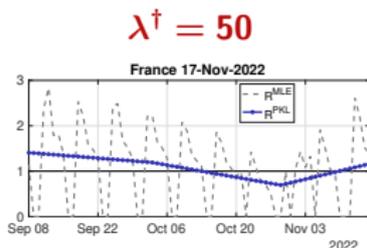
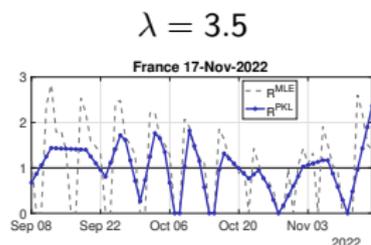
Penalized Kullback-Leibler estimator:

$$\widehat{\mathbf{R}}(\mathbf{Z}; \lambda) = \underset{\mathbf{R} \in \mathbb{R}_+^T}{\operatorname{argmin}} \sum_{t=1}^T \frac{1}{\gamma} d_{\text{KL}}(Z_t | R_t \Phi_t) + \lambda \|\mathbf{D}_2 \mathbf{R}\|_1$$

with $\mathcal{R}(\mathbf{R})$ favoring some temporal regularity

(Abry et al., 2020, *PLOSOne*)

Fine tuning of the regularization parameter:



Data-driven oracle minimization

$$\lambda^\dagger \in \underset{\lambda \in \Lambda}{\operatorname{Argmin}} \mathcal{O}(\mathbf{Z}; \lambda)$$

\Rightarrow **Goal:** \mathcal{O} data-driven proxy for $\|\widehat{\mathbf{R}}(\mathbf{Z}; \lambda) - \overline{\mathbf{R}}\|_2^2$

Goal: \mathcal{O} data-driven proxy for $\|\widehat{\mathbf{R}}(\mathbf{Z}; \lambda) - \overline{\mathbf{R}}\|_2^2$

Strategy: Unbiased Risk Estimate $\mathbb{E}_{\mathbf{Z}} [\mathcal{O}(\mathbf{Z}; \lambda)] = \mathbb{E}_{\mathbf{Z}} \left[\|\widehat{\mathbf{R}}(\mathbf{Z}; \lambda) - \overline{\mathbf{R}}\|_2^2 \right]$

Goal: \mathcal{O} data-driven proxy for $\|\widehat{\mathbf{R}}(\mathbf{Z}; \lambda) - \bar{\mathbf{R}}\|_2^2$

Strategy: Unbiased Risk Estimate $\mathbb{E}_{\mathbf{Z}} [\mathcal{O}(\mathbf{Z}; \lambda)] = \mathbb{E}_{\mathbf{Z}} \left[\|\widehat{\mathbf{R}}(\mathbf{Z}; \lambda) - \bar{\mathbf{R}}\|_2^2 \right]$

Reminder:

$$\frac{\mathbf{Z}}{\gamma} \sim \mathcal{P} \left(\frac{\Phi \mathbf{R}}{\gamma} \right)$$

- $\mathbf{Z} \in \mathbb{N}^T$: reported infection counts,
- $\mathbf{R} = (R_1, \dots, R_T) \in \mathbb{R}_+^T$: daily unknown reproduction number,
- $\Phi = \text{diag}(\Phi_1, \dots, \Phi_T)$: linear operator,
- \mathcal{P} : data-dependent Poisson noise, with scale parameter γ .

Goal: \mathcal{O} data-driven proxy for $\|\widehat{\mathbf{R}}(\mathbf{Z}; \lambda) - \overline{\mathbf{R}}\|_2^2$

Strategy: Unbiased Risk Estimate $\mathbb{E}_{\mathbf{Z}} [\mathcal{O}(\mathbf{Z}; \lambda)] = \mathbb{E}_{\mathbf{Z}} \left[\|\widehat{\mathbf{R}}(\mathbf{Z}; \lambda) - \overline{\mathbf{R}}\|_2^2 \right]$

Reminder:

$$\frac{\mathbf{Z}}{\gamma} \sim \mathcal{P} \left(\frac{\Phi \mathbf{R}}{\gamma} \right)$$

- $\mathbf{Z} \in \mathbb{N}^T$: reported infection counts,
- $\mathbf{R} = (R_1, \dots, R_T) \in \mathbb{R}_+^T$: daily unknown reproduction number,
- $\Phi = \text{diag}(\Phi_1, \dots, \Phi_T)$: linear operator,
- \mathcal{P} : data-dependent Poisson noise, with scale parameter γ .

Challenges:

- ▶ **Poisson model: Stein lemma does not apply** (Eldar, 2008, *IEEE Trans. Signal Process.*; Luisier et al., 2010, *IEEE Trans. Image Process.*; Li et al., 2017, *IEEE Trans. Image Process.*)

Goal: \mathcal{O} data-driven proxy for $\|\widehat{\mathbf{R}}(\mathbf{Z}; \lambda) - \overline{\mathbf{R}}\|_2^2$

Strategy: Unbiased Risk Estimate $\mathbb{E}_{\mathbf{Z}} [\mathcal{O}(\mathbf{Z}; \lambda)] = \mathbb{E}_{\mathbf{Z}} \left[\|\widehat{\mathbf{R}}(\mathbf{Z}; \lambda) - \overline{\mathbf{R}}\|_2^2 \right]$

Reminder:

$$\frac{\mathbf{Z}}{\gamma} \sim \mathcal{P} \left(\frac{\Phi(\mathbf{Z})\mathbf{R}}{\gamma} \right)$$

- $\mathbf{Z} \in \mathbb{N}^T$: reported infection counts,
- $\mathbf{R} = (R_1, \dots, R_T) \in \mathbb{R}_+^T$: daily unknown reproduction number,
- $\Phi = \text{diag}(\Phi_1, \dots, \Phi_T)$: linear operator,
- \mathcal{P} : data-dependent Poisson noise, with scale parameter γ .

Challenges:

- ▶ Poisson model: Stein lemma does not apply (Eldar, 2008, *IEEE Trans. Signal Process.*; Luisier et al., 2010, *IEEE Trans. Image Process.*; Li et al., 2017, *IEEE Trans. Image Process.*)
- ▶ Nonstationary driven autoregressive model: $(\Phi\mathbf{R})_t = R_t \sum_{s=1}^{T\Phi} \phi_s \mathbf{Z}_{t-s}$
 \implies Novel counterpart of Stein lemma for driven autoregressive Poisson model

Goal: \mathcal{O} data-driven proxy for $\|\widehat{\mathbf{R}}(\mathbf{Z}; \lambda) - \overline{\mathbf{R}}\|_2^2$

Strategy: Unbiased Risk Estimate $\mathbb{E}_{\mathbf{Z}} [\mathcal{O}(\mathbf{Z}; \lambda)] = \mathbb{E}_{\mathbf{Z}} \left[\|\widehat{\mathbf{R}}(\mathbf{Z}; \lambda) - \overline{\mathbf{R}}\|_2^2 \right]$

Reminder:

$$\frac{\mathbf{Z}}{\gamma} \sim \mathcal{P} \left(\frac{\Phi(\mathbf{Z})\mathbf{R}}{\gamma} \right)$$

- $\mathbf{Z} \in \mathbb{N}^T$: reported infection counts,
- $\mathbf{R} = (R_1, \dots, R_T) \in \mathbb{R}_+^T$: daily unknown reproduction number,
- $\Phi = \text{diag}(\Phi_1, \dots, \Phi_T)$: linear operator,
- \mathcal{P} : data-dependent Poisson noise, with scale parameter γ .

Challenges:

- ▶ Poisson model: Stein lemma does not apply (Eldar, 2008, *IEEE Trans. Signal Process.*; Luisier et al., 2010, *IEEE Trans. Image Process.*; Li et al., 2017, *IEEE Trans. Image Process.*)
- ▶ Nonstationary driven autoregressive model: $(\Phi\mathbf{R})_t = R_t \sum_{s=1}^{T_\Phi} \phi_s \mathbf{Z}_{t-s}$
 \implies Novel counterpart of Stein lemma for driven autoregressive Poisson model

Autoregressive Poisson Unbiased Risk Estimate (APURE)

Model

$$\frac{Z_t}{\gamma} \sim \mathcal{P} \left(\frac{\bar{R}_t \Phi_t(Z_1, \dots, Z_{t-1})}{\gamma} \right)$$

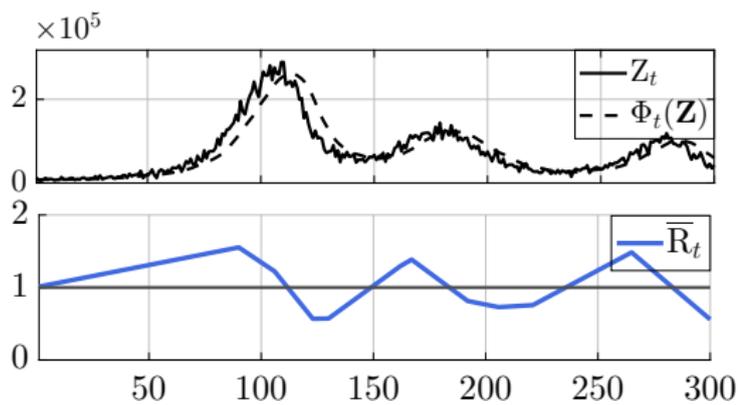
with scale parameter $\log_{10} \gamma \equiv 3$

Data-driven hyperparameter selection on synthetic Poisson data

Model

$$\frac{Z_t}{\gamma} \sim \mathcal{P} \left(\frac{\bar{R}_t \Phi_t(Z_1, \dots, Z_{t-1})}{\gamma} \right)$$

with scale parameter $\log_{10} \gamma \equiv 3$



Data-driven hyperparameter selection on synthetic Poisson data

Model

$$\frac{Z_t}{\gamma} \sim \mathcal{P} \left(\frac{\bar{R}_t \Phi_t(Z_1, \dots, Z_{t-1})}{\gamma} \right)$$

with scale parameter $\log_{10} \gamma \equiv 3$

True prediction \mathcal{P} & estimation \mathcal{E} errors

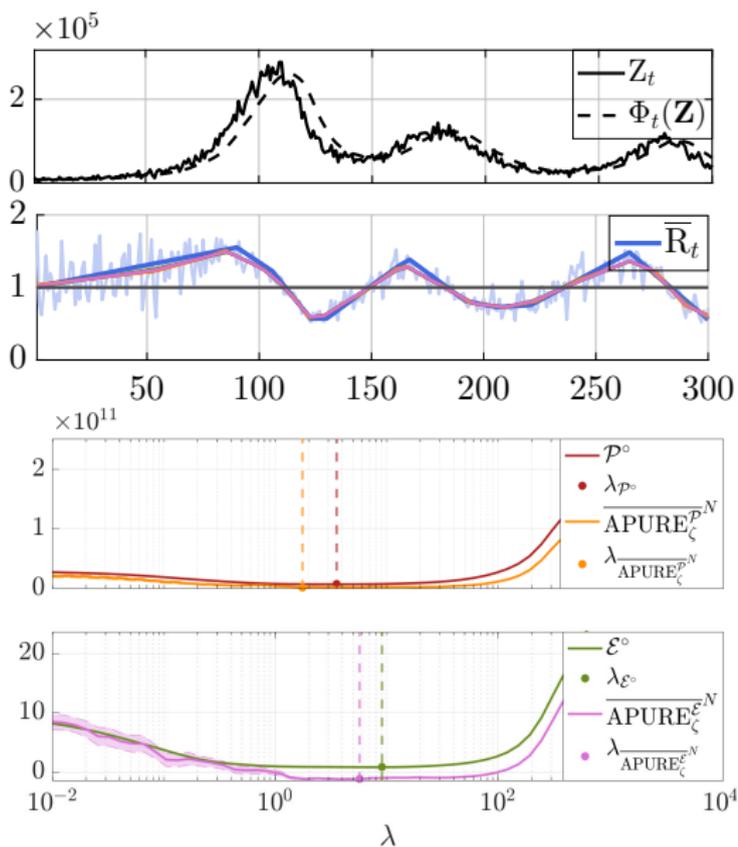
$$\mathcal{P}^\circ(\mathbf{Z}; \theta) = \|\hat{\mathbf{R}}(\mathbf{Z}; \theta) \odot \Phi(\mathbf{Z}) - \bar{\mathbf{R}} \odot \Phi(\mathbf{Z})\|_2^2$$

$$\mathcal{E}^\circ(\mathbf{Z}; \theta) = \|\hat{\mathbf{R}}(\mathbf{Z}; \theta) - \bar{\mathbf{R}}\|_2^2$$

Unbiased risk estimators

$$\overline{\text{APURE}}_{\zeta}^{\mathcal{P}} = \frac{1}{N} \sum_{n=1}^N \text{APURE}_{\zeta^{(n)}}^{\mathcal{P}}$$

$$\overline{\text{APURE}}_{\zeta}^{\mathcal{E}} = \frac{1}{N} \sum_{n=1}^N \text{APURE}_{\zeta^{(n)}}^{\mathcal{E}}$$



Data-driven hyperparameter selection on synthetic Poisson data

Model

$$\frac{Z_t}{\gamma} \sim \mathcal{P} \left(\frac{\bar{R}_t \Phi_t(Z_1, \dots, Z_{t-1})}{\gamma} \right)$$

with scale parameter $\log_{10} \gamma \equiv 3.5$

True prediction \mathcal{P} & estimation \mathcal{E} errors

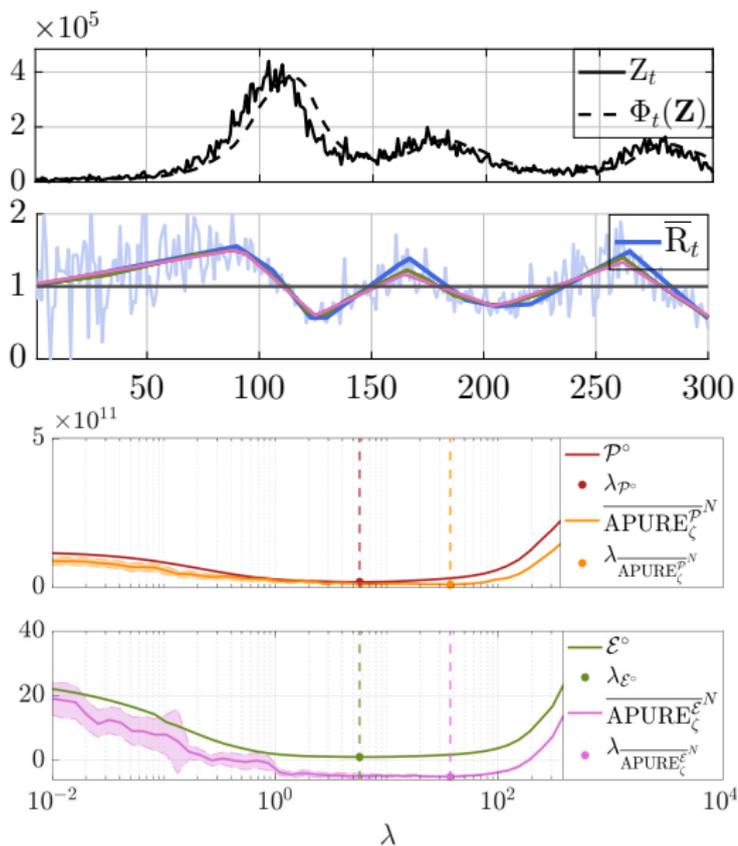
$$\mathcal{P}^\circ(\mathbf{Z}; \theta) = \|\hat{\mathbf{R}}(\mathbf{Z}; \theta) \odot \Phi(\mathbf{Z}) - \bar{\mathbf{R}} \odot \Phi(\mathbf{Z})\|_2^2$$

$$\mathcal{E}^\circ(\mathbf{Z}; \theta) = \|\hat{\mathbf{R}}(\mathbf{Z}; \theta) - \bar{\mathbf{R}}\|_2^2$$

Unbiased risk estimators

$$\overline{\text{APURE}}_{\zeta}^{\mathcal{P}^{\circ}} = \frac{1}{N} \sum_{n=1}^N \text{APURE}_{\zeta^{(n)}}^{\mathcal{P}^{\circ}}$$

$$\overline{\text{APURE}}_{\zeta}^{\mathcal{E}^{\circ}} = \frac{1}{N} \sum_{n=1}^N \text{APURE}_{\zeta^{(n)}}^{\mathcal{E}^{\circ}}$$



Data-driven hyperparameter selection on synthetic Poisson data

Model

$$\frac{Z_t}{\gamma} \sim \mathcal{P} \left(\frac{\bar{R}_t \Phi_t(Z_1, \dots, Z_{t-1})}{\gamma} \right)$$

with scale parameter $\log_{10} \gamma \equiv 4$

True prediction \mathcal{P} & estimation \mathcal{E} errors

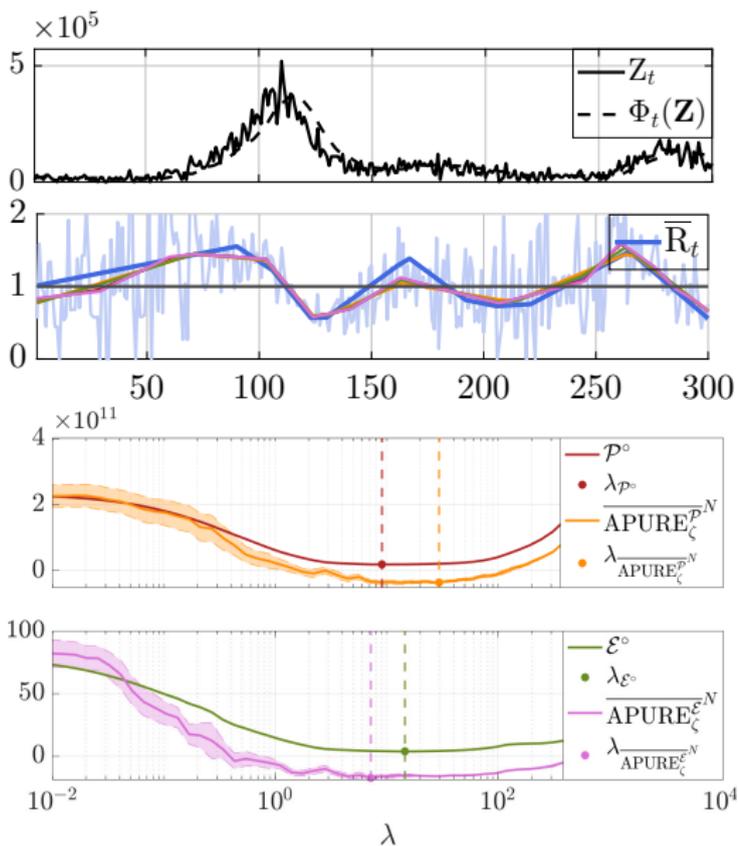
$$\mathcal{P}^\circ(\mathbf{Z}; \theta) = \|\hat{\mathbf{R}}(\mathbf{Z}; \theta) \odot \Phi(\mathbf{Z}) - \bar{\mathbf{R}} \odot \Phi(\mathbf{Z})\|_2^2$$

$$\mathcal{E}^\circ(\mathbf{Z}; \theta) = \|\hat{\mathbf{R}}(\mathbf{Z}; \theta) - \bar{\mathbf{R}}\|_2^2$$

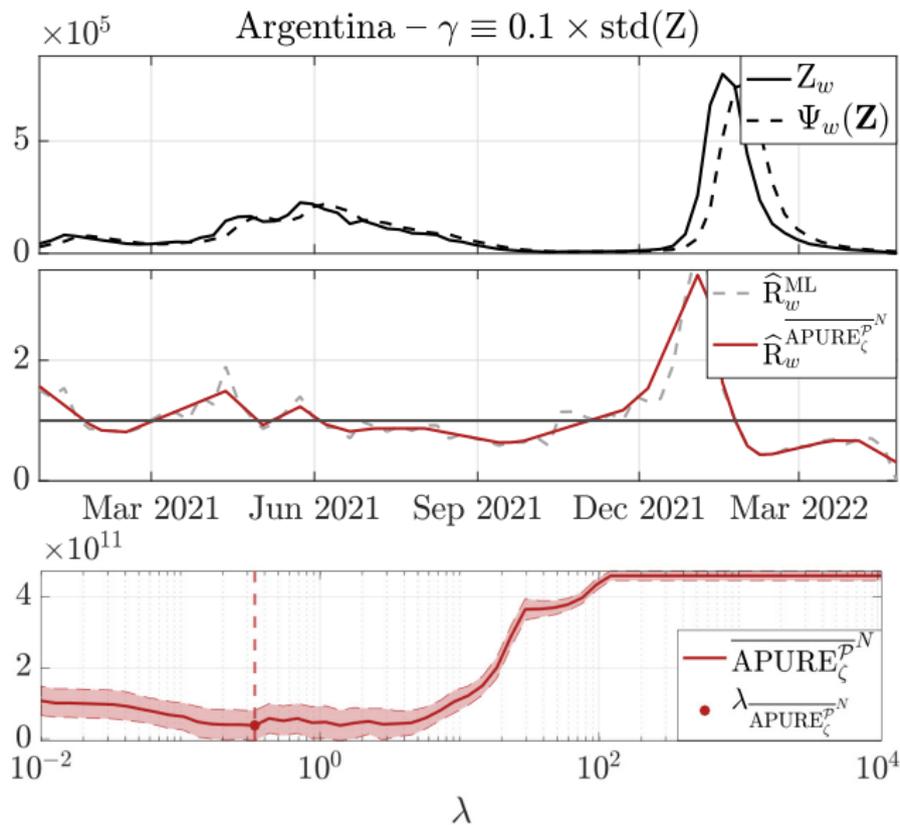
Unbiased risk estimators

$$\overline{\text{APURE}}_{\zeta}^{\mathcal{P}^{\circ}} = \frac{1}{N} \sum_{n=1}^N \text{APURE}_{\zeta^{(n)}}^{\mathcal{P}^{\circ}}$$

$$\overline{\text{APURE}}_{\zeta}^{\mathcal{E}^{\circ}} = \frac{1}{N} \sum_{n=1}^N \text{APURE}_{\zeta^{(n)}}^{\mathcal{E}^{\circ}}$$



Data-driven hyperparameter selection on weekly COVID-19 counts



Pascal & Vaïter, 2024, *Preprint arXiv:2409.14937*

Codes: github.com/bpascal-fr/APURE-Estim-Epi

Inverse problem

$$\mathbf{y} \sim \mathcal{B}(\Phi \bar{\mathbf{x}})$$

$$\lambda^\dagger \in \underset{\lambda \in \Lambda}{\text{Argmin}} \mathcal{O}(\mathbf{y}; \lambda), \quad \text{for} \quad \hat{\mathbf{x}}(\mathbf{y}; \lambda) \in \underset{\mathbf{x} \in \mathbb{R}^N}{\text{Argmin}} \mathcal{D}(\mathbf{y}, \Phi \mathbf{x}) + \lambda \mathcal{R}(\mathbf{x})$$

Inverse problem

$$\mathbf{y} \sim \mathcal{B}(\Phi \bar{\mathbf{x}})$$

$$\lambda^\dagger \in \underset{\lambda \in \Lambda}{\text{Argmin}} \mathcal{O}(\mathbf{y}; \lambda), \quad \text{for} \quad \hat{\mathbf{x}}(\mathbf{y}; \lambda) \in \underset{\mathbf{x} \in \mathbb{R}^N}{\text{Argmin}} \mathcal{D}(\mathbf{y}, \Phi \mathbf{x}) + \lambda \mathcal{R}(\mathbf{x})$$

Data-driven parameter selection

$\implies \mathcal{O}$: Unbiased Risk Estimate (Stein, 1981, *Ann. Stat.*; Eldar, 2008, *IEEE Trans. Signal Process.*; Luisier et al., 2010, *IEEE Trans. Image Process.*; Deledalle et al., 2014, *SIAM J. Imaging Sci.*; Pascal et al., 2021, *J. Math. Imaging Vis.*; Lucas et al., 2023, *Signal, Image Video Process.*)

- ▶ Texture segmentation: additive correlated Gaussian noise;
- ▶ Epidemic monitoring: driven autoregressive data-dependent Poisson noise.

Inverse problem

$$\mathbf{y} \sim \mathcal{B}(\Phi \bar{\mathbf{x}})$$

$$\lambda^\dagger \in \underset{\lambda \in \Lambda}{\text{Argmin}} \mathcal{O}(\mathbf{y}; \lambda), \quad \text{for} \quad \hat{\mathbf{x}}(\mathbf{y}; \lambda) \in \underset{\mathbf{x} \in \mathbb{R}^N}{\text{Argmin}} \mathcal{D}(\mathbf{y}, \Phi \mathbf{x}) + \lambda \mathcal{R}(\mathbf{x})$$

Data-driven parameter selection

$\implies \mathcal{O}$: Unbiased Risk Estimate (Stein, 1981, *Ann. Stat.*; Eldar, 2008, *IEEE Trans. Signal Process.*; Luisier et al., 2010, *IEEE Trans. Image Process.*; Deledalle et al., 2014, *SIAM J. Imaging Sci.*; Pascal et al., 2021, *J. Math. Imaging Vis.*; Lucas et al., 2023, *Signal, Image Video Process.*)

- ▶ Texture segmentation: additive correlated Gaussian noise;
- ▶ Epidemic monitoring: driven autoregressive data-dependent Poisson noise.

Extensions and perspectives

- ▶ Efficient and robust scheme for nonconvex $\mathcal{R}(\mathbf{x})$;
- ▶ Generalization to other noise models: speckle noise in medical imaging;
- ▶ Unsupervised learning for $\hat{\mathbf{x}}(\mathbf{y}; \lambda) = \mathbf{NN}_\theta(\mathbf{y})$ with loss $\mathcal{O}(\mathbf{y}; \theta)$.