Off-the-grid dynamic super-resolution for fluorescence microscopy.

Bastien Laville

Team Morpheme: UCA, Inria SAM, I3S (CNRS, UCA)

6th December 2021

MORPHEME
MORPHOLOGIE ET IMAGES

Table of contents

Introduction

Biomedical context

Aim

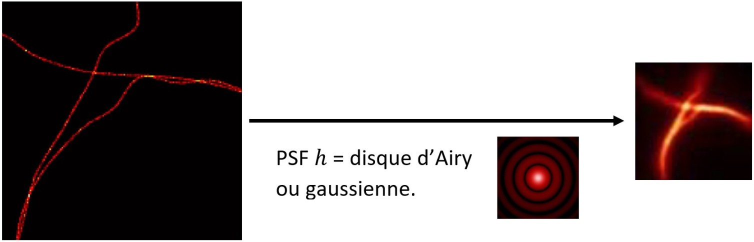
Image biological structures at small scales

Biomedical context

Aim

Image biological structures at small scales

Physical limitation due to diffraction for bodies < 200 nm: convolution by the microscope's *point spread function*.

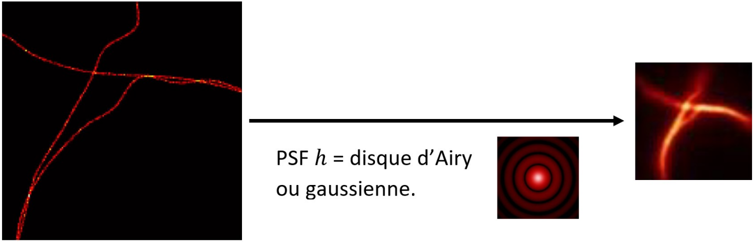


Biomedical context

Aim

Image biological structures at small scales

Physical limitation due to diffraction for bodies < 200 nm: convolution by the microscope's *point spread function*.



Reconstruction e.g. by fluorescence microscopy SMLM: acquisition stack with few lit fluorophores per image. Drawback: many images ($\approx 1 \times 10^4$, does not allow imaging of living cells).

An imagery solution: SOFI

SOFI imaging (Super-resolution optical fluctuation imaging).
Applications: imaging for localisation in fluorescence microscopy, etc. [Dertinger10].

An imagery solution: SOFI

SOFI imaging (Super-resolution optical fluctuation imaging). Applications: imaging for localisation in fluorescence microscopy, etc. [Dertinger10].

- many fluorophores lit at the same time;

An imagery solution: SOFI

SOFI imaging (Super-resolution optical fluctuation imaging).
Applications: imaging for localisation in fluorescence microscopy, etc. [Dertinger10].

- many fluorophores lit at the same time;
- temporal independence of the fluorophores' luminosity fluctuation;

An imagery solution: SOFI

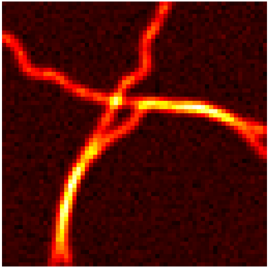
SOFI imaging (Super-resolution optical fluctuation imaging). Applications: imaging for localisation in fluorescence microscopy, etc. [Dertinger10].

- many fluorophores lit at the same time;
- temporal independence of the fluorophores' luminosity fluctuation;
- less harmful to the biological structures studied.

An imagery solution: SOFI

SOFI imaging (Super-resolution optical fluctuation imaging).
Applications: imaging for localisation in fluorescence microscopy, etc. [Dertinger10].

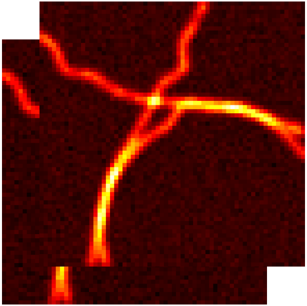
- many fluorophores lit at the same time;
- temporal independence of the fluorophores' luminosity fluctuation;
- less harmful to the biological structures studied.



An imagery solution: SOFI

SOFI imaging (Super-resolution optical fluctuation imaging).
Applications: imaging for localisation in fluorescence microscopy, etc. [Dertinger10].

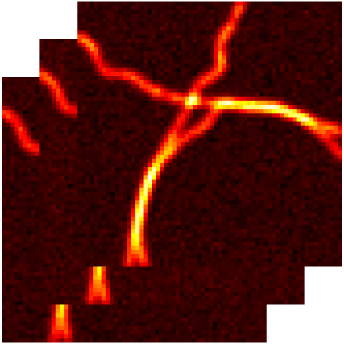
- many fluorophores lit at the same time;
- temporal independence of the fluorophores' luminosity fluctuation;
- less harmful to the biological structures studied.



An imagery solution: SOFI

SOFI imaging (Super-resolution optical fluctuation imaging).
Applications: imaging for localisation in fluorescence microscopy, etc. [Dertinger10].

- many fluorophores lit at the same time;
- temporal independence of the fluorophores' luminosity fluctuation;
- less harmful to the biological structures studied.

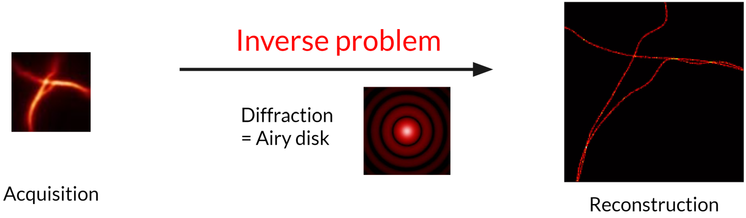


Inverse problem: from an acquisition, we reconstruct spikes' positions and amplitudes. Super-resolution grid problem:

Acquisition

Reconstruction

Inverse problem: from an acquisition, we reconstruct spikes' positions and amplitudes. Super-resolution grid problem:



- off-the-grid deconvolution can be understood as the 'limit' of an increasingly fine grid;
- not limited by the fine grid but still limited by the noise.

Link with grid problem

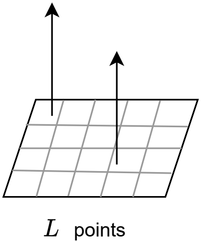
$$\min_{a\in\mathbb{R}^L} \|y - \Phi_L a\|_{\mathcal{H}}^2 + \lambda \|a\|_1 \quad \text{(LASSO)}$$

$$\min_{m\in\mathcal{M}(\mathcal{X})} \|y - \Phi m\|_{\mathcal{H}}^2 + \lambda \|m\|_{\mathrm{TV}} \quad \text{(BLASSO)}$$

Link with grid problem

$$\min_{a \in \mathbb{R}^L} \|y - \Phi_L a\|_{\mathcal{H}}^2 + \lambda \|a\|_1 \qquad \text{(LASSO)}$$

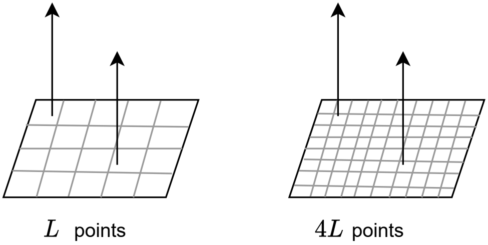
$$\min_{m \in \mathcal{M}(\mathcal{X})} \|y - \Phi m\|_{\mathcal{H}}^2 + \lambda \|m\|_{\mathrm{TV}} \qquad \text{(BLASSO)}$$



Link with grid problem

$$\min_{a \in \mathbb{R}^L} \|y - \Phi_L a\|_{\mathcal{H}}^2 + \lambda \|a\|_1 \qquad \text{(LASSO)}$$

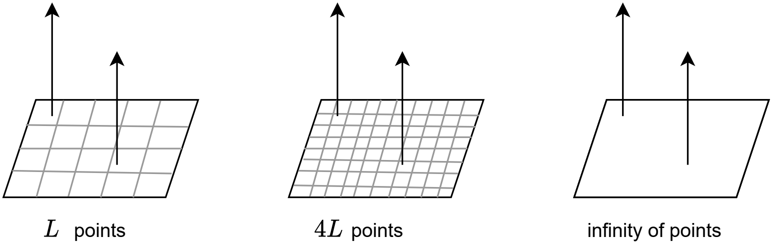
$$\min_{m \in \mathcal{M}(\mathcal{X})} \|y - \Phi m\|_{\mathcal{H}}^2 + \lambda \|m\|_{\mathrm{TV}} \qquad \text{(BLASSO)}$$



Link with grid problem

$$\min_{a\in\mathbb{R}^L} \|y - \Phi_L a\|_{\mathcal{H}}^2 + \lambda\|a\|_1 \quad \text{(LASSO)}$$

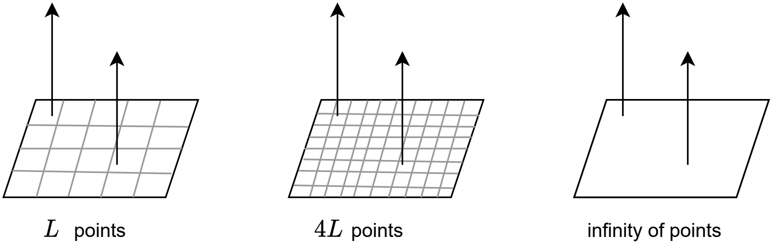
$$\min_{m\in\mathcal{M}(\mathcal{X})} \|y - \Phi m\|_{\mathcal{H}}^2 + \lambda\|m\|_{\mathrm{TV}} \quad \text{(BLASSO)}$$



Link with grid problem

$$\min_{a \in \mathbb{R}^L} \|y - \Phi_L a\|_{\mathcal{H}}^2 + \lambda \|a\|_1 \qquad \text{(LASSO)}$$

$$\min_{m \in \mathcal{M}(\mathcal{X})} \|y - \Phi m\|_{\mathcal{H}}^2 + \lambda \|m\|_{\mathrm{TV}} \qquad \text{(BLASSO)}$$



BLASSO is the functiunal limit of the LASSO problem for $L \to +\infty$.

Discrete case

- the reconstructed peaks are necessarily on the fine grid;
- (Non-)convex combinatorial optimisation;
- fast numerical computation;
- large literature.

Discrete case

- the reconstructed peaks are necessarily on the fine grid;
- (Non-)convex combinatorial optimisation;
- fast numerical computation;
- large literature.

Off-the-grid case

- not limited by the grid;
- convexity of the functional on an infinite dimensional space;
- existence and uniqueness guarantees;
- recent field of research.

Off-the-grid digest

Elementary bricks :

- $\mathcal{X}$ is a compact of $\mathbb{R}^d$;

Elementary bricks :

- $\mathcal{X}$ is a compact of $\mathbb{R}^d$;
- how to model spikes? Dirac measures δ_x, elements of $\mathcal{M}(\mathcal{X})$ the *space of signed Radon measures*;

Elementary bricks :

- $\mathcal{X}$ is a compact of $\mathbb{R}^d$;
- how to model spikes? Dirac measures δ_x, elements of $\mathcal{M}(\mathcal{X})$ the *space of signed Radon measures*;
- topological dual of $\mathscr{C}_0(\mathcal{X})$ ($= \mathscr{C}(\mathcal{X})^*$ here) for $\langle f, m \rangle = \int_{\mathcal{X}} f \, \mathrm{d}m$. Generalisation of $\mathrm{L}^1(\mathcal{X})$ since $\mathrm{L}^1(\mathcal{X}) \hookrightarrow \mathcal{M}(\mathcal{X})$;

Elementary bricks :

- $\mathcal{X}$ is a compact of $\mathbb{R}^d$;
- how to model spikes? Dirac measures δ_x, elements of $\mathcal{M}(\mathcal{X})$ the *space of signed Radon measures*;
- topological dual of $\mathscr{C}_0(\mathcal{X})$ ($= \mathscr{C}(\mathcal{X})^*$ here) for $\langle f, m \rangle = \int_{\mathcal{X}} f \, \mathrm{d}m$. Generalisation of $\mathrm{L}^1(\mathcal{X})$ since $\mathrm{L}^1(\mathcal{X}) \hookrightarrow \mathcal{M}(\mathcal{X})$;
- Banach for the TV-norm: $m \in \mathcal{M}(\mathcal{X})$,

$$|m|(\mathcal{X}) \stackrel{\text{def.}}{=} \sup \left(\int_{\mathcal{X}} f \, \mathrm{d}m \,\middle|\, f \in \mathscr{C}(\mathcal{X}), \|f\|_{\infty,\mathcal{X}} \leq 1 \right).$$

If $m = \sum_{i=1}^N a_i \delta_{x_i}$ is a discrete measure then $|m|(\mathcal{X}) = \sum_{i=1}^N |a_i|$.

BLASSO

Let $m_{a_0,x_0} \overset{\text{def.}}{=} \sum_{i=1}^{N} a_i \delta_{x_i}$ be a discrete measure,

BLASSO

Let $m_{a_0,x_0} \overset{\text{def.}}{=} \sum_{i=1}^N a_i \delta_{x_i}$ be a discrete measure, $\Phi : \mathcal{M}(\mathcal{X}) \to \mathrm{L}^2(\mathcal{X})$ forward operator (e.g. $\Phi m_{a_0,x_0} \overset{\text{def.}}{=} \sum_{i=1}^N a_i h(x - x_i)$ the Gaussian kernel)

BLASSO

Let $m_{a_0,x_0} \stackrel{\text{def.}}{=} \sum_{i=1}^N a_i \delta_{x_i}$ be a discrete measure, $\Phi : \mathcal{M}(\mathcal{X}) \to \mathrm{L}^2(\mathcal{X})$ forward operator (e.g. $\Phi m_{a_0,x_0} \stackrel{\text{def.}}{=} \sum_{i=1}^N a_i h(x - x_i)$ the Gaussian kernel) and $w \in \mathrm{L}^2(\mathcal{X})$ noise:

$$y \stackrel{\text{def.}}{=} \Phi m_{a_0,x_0} + w.$$

BLASSO

Let $m_{a_0,x_0} \stackrel{\text{def.}}{=} \sum_{i=1}^N a_i \delta_{x_i}$ be a discrete measure, $\Phi : \mathcal{M}(\mathcal{X}) \to \mathrm{L}^2(\mathcal{X})$ forward operator (e.g. $\Phi m_{a_0,x_0} \stackrel{\text{def.}}{=} \sum_{i=1}^N a_i h(x - x_i)$ the Gaussian kernel) and $w \in \mathrm{L}^2(\mathcal{X})$ noise:

$$y \stackrel{\text{def.}}{=} \Phi m_{a_0,x_0} + w.$$

We call **BLASSO** the optimisation problem [Castro12, Duval15] for $\lambda > 0$:

$$\operatorname*{argmin}_{m \in \mathcal{M}(\mathcal{X})} \frac{1}{2} \|y - \Phi m\|^2_{\mathrm{L}^2(\mathcal{X})} + \lambda |m|(\mathcal{X}) \qquad (\mathcal{P}_\lambda(y))$$

BLASSO

Let $m_{a_0,x_0} \stackrel{\text{def.}}{=} \sum_{i=1}^N a_i \delta_{x_i}$ be a discrete measure, $\Phi : \mathcal{M}(\mathcal{X}) \to \mathrm{L}^2(\mathcal{X})$ forward operator (e.g. $\Phi m_{a_0,x_0} \stackrel{\text{def.}}{=} \sum_{i=1}^N a_i h(x - x_i)$ the Gaussian kernel) and $w \in \mathrm{L}^2(\mathcal{X})$ noise:

$$y \stackrel{\text{def.}}{=} \Phi m_{a_0,x_0} + w.$$

We call **BLASSO** the optimisation problem [Castro12, Duval15] for $\lambda > 0$:

$$\underset{m \in \mathcal{M}(\mathcal{X})}{\operatorname{argmin}} \frac{1}{2} \|y - \Phi m\|_{\mathrm{L}^2(\mathcal{X})}^2 + \lambda |m|(\mathcal{X}) \qquad (\mathcal{P}_\lambda(y))$$

The optimisation space $\mathcal{M}(\mathcal{X})$ is an infinite dimensional space, reflexive only for weak-* topology: a difficult problem.

Dual certificate

To solve this problem, we have the strong (Fenchel) duality:

$$\eta_\lambda \in \partial|m|(\mathcal{X}) \cap \Phi^* y$$

defines a simpler dual problem to study.

[1] *Non-Degenerate Source Condition e.g.*

Dual certificate

To solve this problem, we have the strong (Fenchel) duality:

$$\eta_\lambda \in \partial|m|(\mathcal{X}) \cap \Phi^* y$$

defines a simpler dual problem to study.
Let m be a minimum of $(\mathcal{P}_\lambda(y))$. The optimality of the measure is characterised by this dual certificate:

$$\eta_\lambda = \frac{1}{\lambda}\Phi^*(\Phi m - y), \quad \text{such that } \|\eta\|_{\infty,\mathcal{X}} \leq 1.$$

[1] *Non-Degenerate Source Condition e.g.*

Dual certificate

To solve this problem, we have the strong (Fenchel) duality:

$$\eta_\lambda \in \partial|m|(\mathcal{X}) \cap \Phi^* y$$

defines a simpler dual problem to study.
Let m be a minimum of $(\mathcal{P}_\lambda(y))$. The optimality of the measure is characterised by this dual certificate:

$$\eta_\lambda = \frac{1}{\lambda}\Phi^*(\Phi m - y), \quad \text{such that } \|\eta\|_{\infty,\mathcal{X}} \leq 1.$$

Conditions[1] on the dual certificate + on the operator = reconstruction guarantees.

[1] *Non-Degenerate Source Condition e.g.*

Numerical computation

Numerical computation

- $\mathcal{M}(\mathcal{X})$ is 'only' a Banach space: it is then hard to use a proximal algorithm such as [Beck09];

Numerical computation

- $\mathcal{M}(\mathcal{X})$ is 'only' a Banach space: it is then hard to use a proximal algorithm such as [Beck09];
- greedy algorithm such as the conditional gradient algorithm does not require a Hilbertian structure;

Numerical computation

- $\mathcal{M}(\mathcal{X})$ is 'only' a Banach space: it is then hard to use a proximal algorithm such as [Beck09];
- greedy algorithm such as the conditional gradient algorithm does not require a Hilbertian structure;
- we use its enhanced version *Sliding Frank-Wolfe* [Denoyelle19] but differs in the final non-convex step.

Numerical computation

- $\mathcal{M}(\mathcal{X})$ is 'only' a Banach space: it is then hard to use a proximal algorithm such as [Beck09];
- greedy algorithm such as the conditional gradient algorithm does not require a Hilbertian structure;
- we use its enhanced version *Sliding Frank-Wolfe* [Denoyelle19] but differs in the final non-convex step.

Numerous other algorithms:

	pros	cons
moments method [Lasserre10]		difficult nD case
conditional gradient [Bredies13]	convergence guarantees	difficult iteration
particle gradient flow [Chizat20]	quick to compute	not robust wrt noise

Illustration of Conic Particle Gradient Flow reconstruction [Chizat20]:

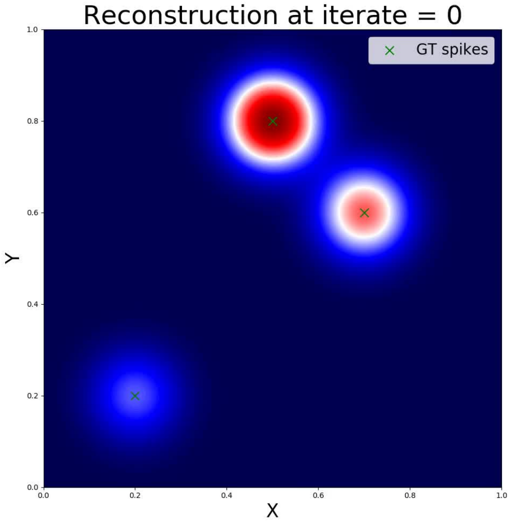
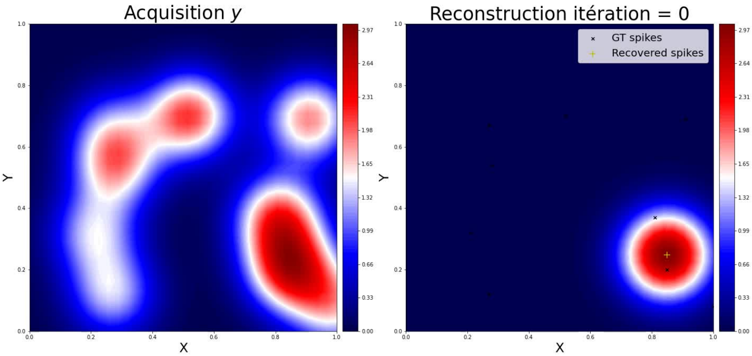
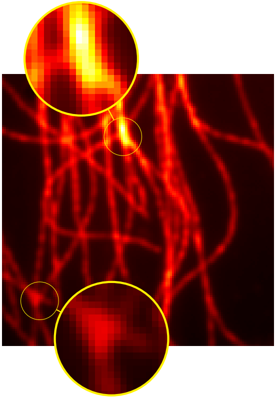


Illustration of Sliding Frank-Wolfe [Denoyelle19] iterative reconstruction:

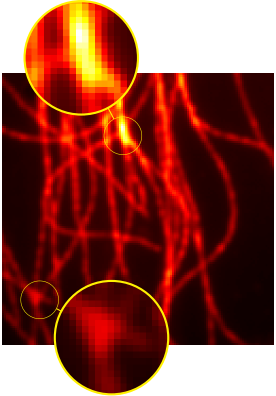


On an EPFL SMLM Challenge stack (10000 images, high density):

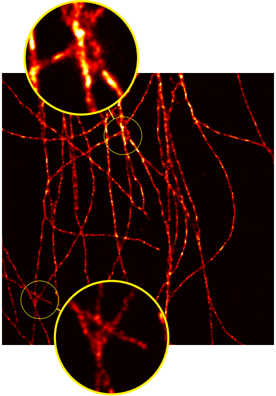


Mean of stack

On an EPFL SMLM Challenge stack (10000 images, high density):

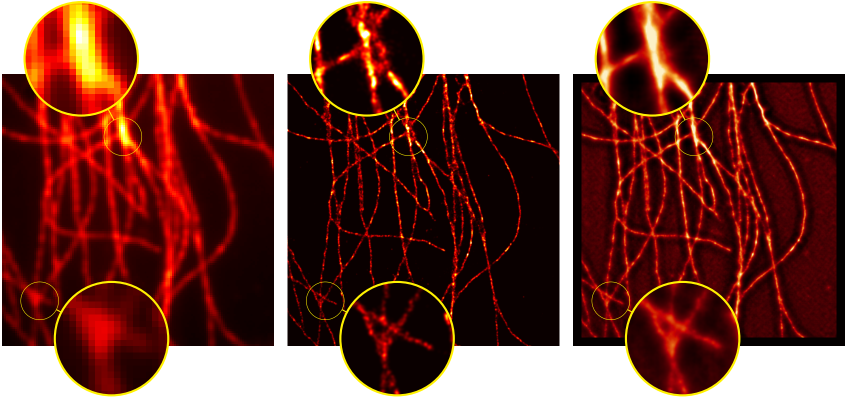


Mean of stack



Off-the-grid SFW

On an EPFL SMLM Challenge stack (10000 images, high density):



Mean of stack — Off-the-grid SFW — SRRF [Culley18]

Dynamic off-the-grid

Quantities at stake

- acquisition stack (images in $\mathrm{L}^2(\mathcal{X})$) during $[0, T]$;

Quantities at stake

- acquisition stack (images in $\mathrm{L}^2(\mathcal{X})$) during $[0, T]$;
- we define $y : [0, T] \to \mathrm{L}^2(\mathcal{X})$ the SOFI acquisition stack ;

Quantities at stake

- acquisition stack (images in $\mathrm{L}^2(\mathcal{X})$) during $[0, T]$;
- we define $y : [0, T] \to \mathrm{L}^2(\mathcal{X})$ the SOFI acquisition stack ;
- we aim to reconstruct the *dynamic* measure:

$$t \mapsto \mu(t) \overset{\text{def.}}{=} \sum_{i=1}^{N} a_i(t)\delta_{x_i} \in \mathrm{L}^2(0, T; \mathcal{M}(\mathcal{X}))$$

generating a.e. $t \in [0, T] : y(t) = \Phi\mu(t)$.

Quantities at stake

- acquisition stack (images in $\mathrm{L}^2(\mathcal{X})$) during $[0,T]$;
- we define $y : [0,T] \to \mathrm{L}^2(\mathcal{X})$ the SOFI acquisition stack ;
- we aim to reconstruct the *dynamic* measure:

$$t \mapsto \mu(t) \stackrel{\text{def.}}{=} \sum_{i=1}^{N} a_i(t)\delta_{x_i} \in \mathrm{L}^2(0,T;\mathcal{M}(\mathcal{X}))$$

generating a.e. $t \in [0,T]$: $y(t) = \Phi\mu(t)$. In the convolution case for PSF h, $\Phi\mu(t) = \sum_{i=1}^{N} a_i(t) \int_{\mathcal{X}} h(x - x_i)\,\mathrm{d}x$.

Quantities at stake

- acquisition stack (images in $\mathrm{L}^2(\mathcal{X})$) during $[0, T]$;
- we define $y : [0, T] \to \mathrm{L}^2(\mathcal{X})$ the SOFI acquisition stack ;
- we aim to reconstruct the *dynamic* measure:

$$t \mapsto \mu(t) \stackrel{\text{def.}}{=} \sum_{i=1}^{N} a_i(t)\delta_{x_i} \in \mathrm{L}^2(0, T; \mathcal{M}(\mathcal{X}))$$

generating a.e. $t \in [0, T] : y(t) = \Phi\mu(t)$. In the convolution case for PSF h, $\Phi\mu(t) = \sum_{i=1}^{N} a_i(t) \int_{\mathcal{X}} h(x - x_i)\, \mathrm{d}x$.

Cumulants are a tool to reconstruct the positions x_i. Example : temporal mean $\bar{y} \stackrel{\text{def.}}{=} \frac{1}{T} \int_0^T y(\cdot, t)\, \mathrm{d}t$. One have $\Phi m_{a,x} = \bar{y}$ where $m_{a,x} \stackrel{\text{def.}}{=} \sum_{i=1}^{N} \bar{a}_i \delta_{x_i}$ and $\bar{a}_i$ is the mean of $a_i(\cdot)$.

Build the variational problem

Let R_y be the temporal covariance, $\forall u, v \in \mathcal{X}$ we get:

Build the variational problem

Let R_y be the temporal covariance, $\forall u, v \in \mathcal{X}$ we get:

$$\begin{aligned}
R_y(u,v) &\overset{\text{def.}}{=} \frac{1}{T}\int_0^T \left(y(u,t)-\bar{y}(u)\right)\left(y(v,t)-\bar{y}(v)\right)\,\mathrm{d}t \\
&= \dots \quad \text{(independence of fluctuations [Dertinger10])} \\
&= \sum_{i=1}^{N} \underbrace{M_i}_{a_i \text{ variance}} \; h(u-x_i)h(v-x_i) \\
&= \int_{\mathcal{X}} h(u-x)h(v-x)\,\mathrm{d}m_{M,x}(x) \\
&= \Lambda m_{M,x}(u,v).
\end{aligned}$$

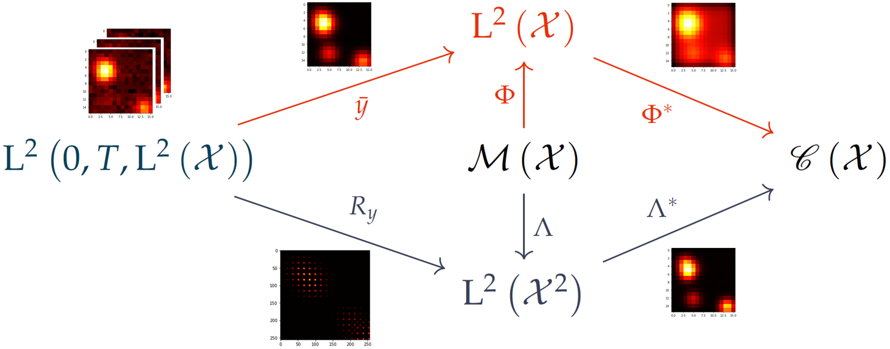
Build the variational problem

Let R_y be the temporal covariance, $\forall u, v \in \mathcal{X}$ we get:

$$
\begin{aligned}
R_y(u,v) &\overset{\text{def.}}{=} \frac{1}{T}\int_0^T \left(y(u,t) - \bar{y}(u)\right)\left(y(v,t) - \bar{y}(v)\right) \, \mathrm{d}t \\
&= \dots \quad \text{(independence of fluctuations [Dertinger10])} \\
&= \sum_{i=1}^{N} \underbrace{M_i}_{a_i \text{ variance}} \; h(u - x_i)h(v - x_i) \\
&= \int_{\mathcal{X}} h(u-x)h(v-x) \, \mathrm{d}m_{M,x}(x) \\
&= \Lambda m_{M,x}(u,v).
\end{aligned}
$$

$m_{M,x} \overset{\text{def.}}{=} \sum_{i=1}^{N} M_i \delta_{x_i}$ shares the same positions as $\mu = \sum_{i=1}^{N} a_i(t)\delta_{x_i}$, we call $\Lambda : \mathcal{M}(\mathcal{X}) \to \mathrm{L}^2(\mathcal{X}^2)$ this « covariance operator ».

Quantities digest



Legend: dynamic part, temporal mean part $\bar{y}$ and covariance R_y.

BLASSO on cumulants

Let $\lambda > 0$,
covariance problem writes down:

$$\underset{m \in \mathcal{M}(\mathcal{X})}{\operatorname{argmin}} T_\lambda(m) \stackrel{\text{def.}}{=} \frac{1}{2} \left\| R_y - \Lambda(m) \right\|^2_{\mathrm{L}^2(\mathcal{X}^2)} + \lambda |m|(\mathcal{X}) \qquad (\mathcal{Q}_\lambda(y))$$

BLASSO on cumulants

Let $\lambda > 0$,
covariance problem writes down:

$$\operatorname*{argmin}_{m \in \mathcal{M}(\mathcal{X})} T_\lambda(m) \stackrel{\text{def.}}{=} \frac{1}{2} \| R_y - \Lambda(m) \|^2_{\mathrm{L}^2(\mathcal{X}^2)} + \lambda |m|(\mathcal{X}) \qquad (\mathcal{Q}_\lambda(y))$$

while mean reconstruction is:

$$\operatorname*{argmin}_{m \in \mathcal{M}(\mathcal{X})} S_\lambda(m) \stackrel{\text{def.}}{=} \frac{1}{2} \| \bar{y} - \Phi(m) \|^2_{\mathrm{L}^2(\mathcal{X})} + \lambda |m|(\mathcal{X}) \qquad (\mathcal{P}_\lambda(\bar{y}))$$

BLASSO on cumulants

Let $\lambda > 0$,
covariance problem writes down:

$$\underset{m \in \mathcal{M}(\mathcal{X})}{\operatorname{argmin}} T_\lambda(m) \overset{\text{def.}}{=} \frac{1}{2} \|R_y - \Lambda(m)\|^2_{\mathrm{L}^2(\mathcal{X}^2)} + \lambda |m|(\mathcal{X}) \qquad (\mathcal{Q}_\lambda(y))$$

while mean reconstruction is:

$$\underset{m \in \mathcal{M}(\mathcal{X})}{\operatorname{argmin}} S_\lambda(m) \overset{\text{def.}}{=} \frac{1}{2} \|\bar{y} - \Phi(m)\|^2_{\mathrm{L}^2(\mathcal{X})} + \lambda |m|(\mathcal{X}) \qquad (\mathcal{P}_\lambda(\bar{y}))$$

Let $\Delta \overset{\text{def.}}{=} \min_{i \neq j} \left| x_i - x_j \right|$ be the *minimum* separation distance

Proposition

Support of a ***real*** *Radon measure in noiseless setting is reconstructed:*

- *for* $(\mathcal{P}_\lambda(\bar{y}))$ *if* $\Delta \gtrsim 1{,}1\sigma$ *[Bendory16] ;*
- *for* $(\mathcal{Q}_\lambda(y))$ *if* $\Delta \gtrsim 1{,}1\sigma/\sqrt{2}$*:* ***better!*** *.*

Numerical results 1D

Implementation in an OOP module in `python`:

- to use Radon measures, certificates, optimisation algorithm, ($\mathcal{Q}_\lambda(y)$) et ($\mathcal{P}_\lambda(\bar{y})$);
- written in `PyTorch` + CUDA (GPU);
- question of quality metrics. L^2 distance is not suitable, we prefer the *flat metric* (or Kantorovitch-Rubinstein metric).

Numerical results 1D

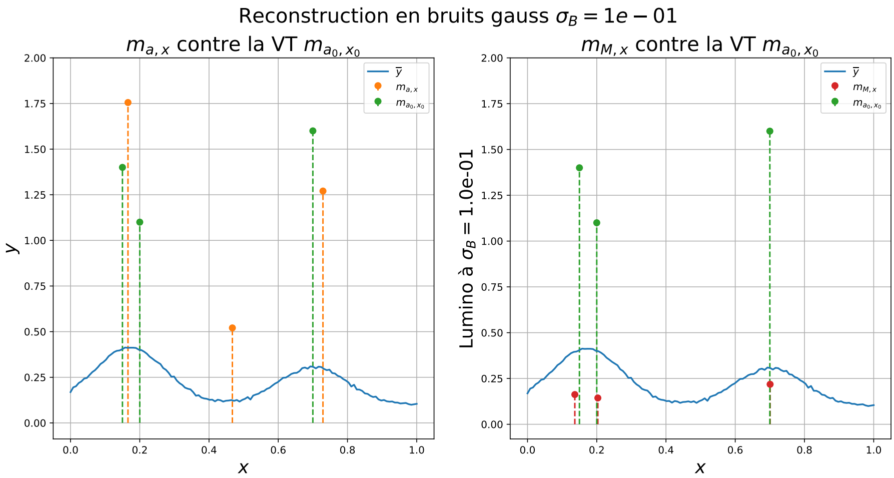
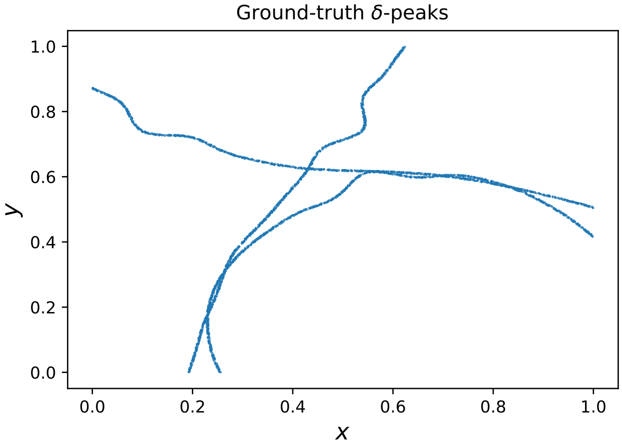


Figure 1: $\mathcal{W}_1(m_{a_0,x_0}, m_{a,x}) \approx 1 \times 10^{-1}$ et $\mathcal{W}_1(m_{a_0,x_0}, m_{M,x}) \approx 5 \times 10^{-3}$.

Numerical 2D results SOFItool

Test on 2D tubulins from ISBI challenge 2016:

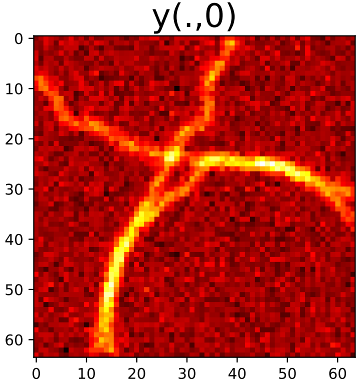
- stack of 1000 acquisitions 64×64 simulated by SOFItool;
- 8700 emitters scattered along the tubulins; **high** background noise + Poisson noise at 4 + Gaussian noise at 1×10^{-2}. SNR $\approx$ 10 db.



Numerical 2D results SOFItool

Test on 2D tubulins from ISBI challenge 2016:

- stack of 1000 acquisitions 64×64 simulated by SOFItool;
- 8700 emitters scattered along the tubulins; **high** background noise + Poisson noise at 4 + Gaussian noise at 1×10^{-2}. SNR $\approx$ 10 db.



Numerical 2D results SOFItool

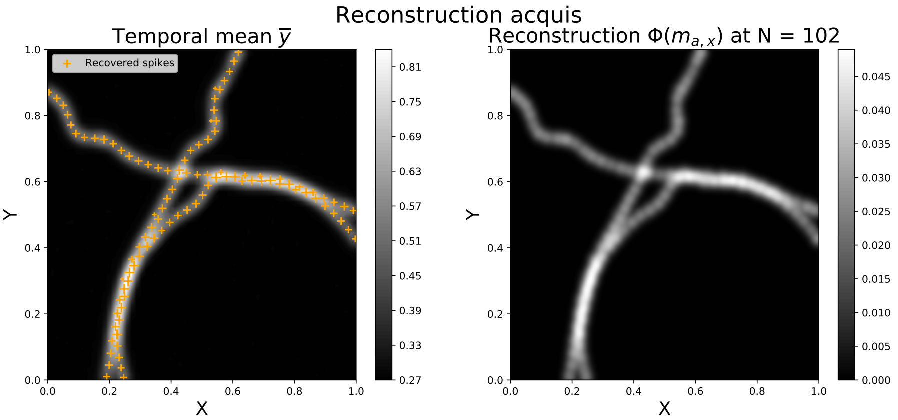


Figure 2: Reconstruction by $(\mathcal{P}_\lambda(\bar{y}))$.

Numerical 2D results SOFItool

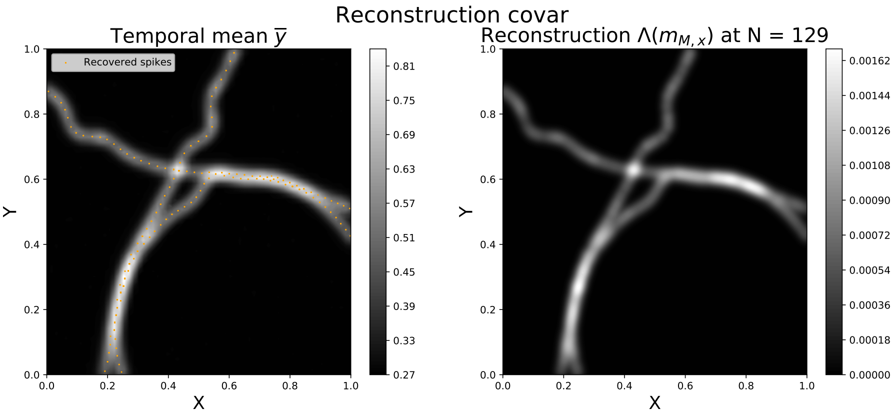


Figure 2: Reconstruction by $(\mathcal{Q}_\lambda(y))$.

Numerical 2D results SOFItool

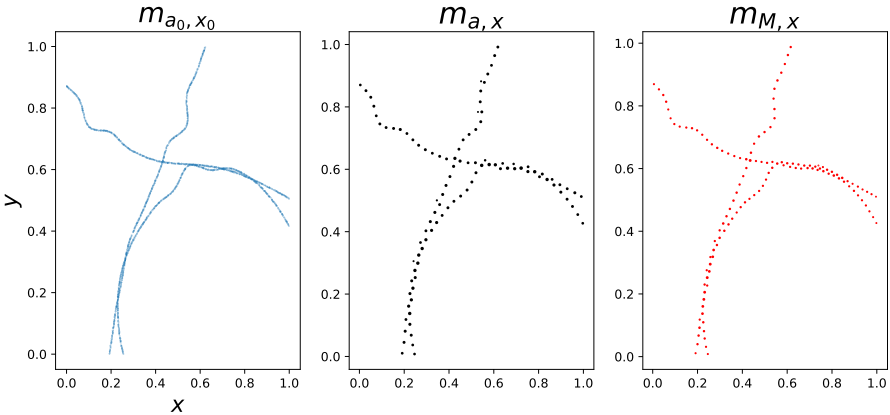


Figure 2: Comparison between ground-truth and solutions of both $(\mathcal{P}_\lambda(\bar{y}))$ and $(\mathcal{Q}_\lambda(y))$.

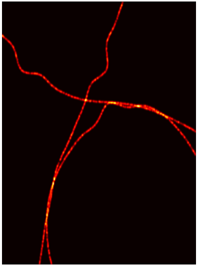


Figure 3: Ground-truth

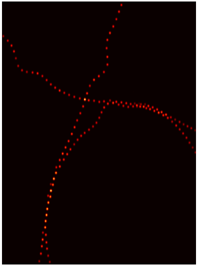


Figure 4: $(\mathcal{Q}_\lambda(y))$

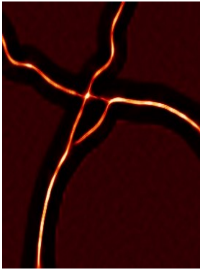


Figure 5: Grid: SRRF [Culley18]

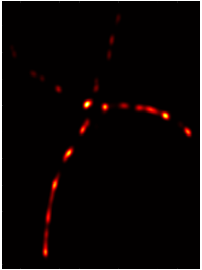


Figure 6: Grid: SPARCOM [Solomon18]

Conclusion

Take home statements

- off-the-grid methods squeeze all the 'information' out the acquisition y: no discretisation drawback;
- strong results for existence and uniqueness of BLASSO solution;
- only one efficient numerical algorithm: Sliding Frank-Wolfe;

Take home statements

- off-the-grid methods squeeze all the 'information' out the acquisition y: no discretisation drawback;
- strong results for existence and uniqueness of BLASSO solution;
- only one efficient numerical algorithm: Sliding Frank-Wolfe;

Outlook

- results only for sparse spike problem with known forward operator at the moment: extendable to other imagery problems? (Obviously yes, gridless compressed sensing, etc.)
- theory only suited for spikes: what about other source structures?
- quite costly numerical algorithms: learning approaches?

Bibliographie de l'exposé I

- Yohann de Castro and Fabrice Gamboa. *Exact reconstruction using Beurling minimal extrapolation.* Journal of Mathematical Analysis and Applications, Elsevier BV, 2012, 395, 336-354

- Quentin Denoyelle, Vincent Duval, Gabriel Peyré, Emmanuel Soubies. *The Sliding Frank-Wolfe Algorithm and its Application to Super-Resolution Microscopy*. Inverse Problems, IOP Publishing, In press.

- Lenaic Chizat, Francis Bach. *On the Global Convergence of Gradient Descent for Over-parameterized Models using Optimal Transport*. Advances in Neural Information Processing Systems (NIPS), Dec 2018, Montréal, Canada.

- Tamir Bendory, Shai Dekel, Arie Feuer, Robust recovery of stream of pulses using convex optimization, Journal of Mathematical Analysis and Applications, Volume 442, Issue 2, 2016.

Bibliographie de l'exposé II

- A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems, Amir Beck and Marc Teboulle, SIAM J. IMAGING SCIENCES, 2009.
- Marguerite Frank et Philip Wolfe, « An algorithm for quadratic programming », Naval Research Logistics Quarterly, vol. 3, 1956.
-] J. B. Lasserre, « Moments, positive polynomials and their applications », Imperial College Press Optimization Series, vol. 1, pp. xxii+361, 2010
- K. Bredies and H. K. Pikkarainen, « Inverse problems in spaces of measures », ESAIM Control Optim. Calc. Var., vol. 19, no. 1, pp. 190-218, 2013.
- Lenaic Chizat, Francis Bach. On the Global Convergence of Gradient Descent for Over-parameterized Models using Optimal Transport. Advances in Neural Information Processing Systems (NIPS), Dec 2018, Montréal, Canada.

Bibliographie de l'exposé III

- Lenaic Chizat. Sparse Optimization on Measures with Over-parameterized Gradient Descent. 2020
- Candès, Emmanuel & Fernandez-Granda, Carlos. (2014). Towards a Mathematical Theory of Super-Resolution. Communications on Pure and Applied Mathematics. 67. 10.1002/cpa.21455.
- Vincent Duval, Gabriel Peyré. Exact Support Recovery for Sparse Spikes Deconvolution. Foundations of Computational Mathematics, Springer Verlag, 2015, 15 (5), pp.1315-1355.
- Thomas Dertinger, Ryan Colyer, Robert Vogel, Jörg Enderlein, and Shimon Weiss, "Achieving increased resolution and more pixels with Superresolution Optical Fluctuation Imaging (SOFI)," Opt. Express 18, 18875-18885 (2010)
- Culley S, Tosheva KL, Matos Pereira P, Henriques R. SRRF: Universal live-cell super-resolution microscopy. Int J Biochem Cell Biol. 2018;101:74-79. doi:10.1016/j.biocel.2018.05.014

Bibliographie de l'exposé IV

Oren Solomon, Maor Mutzafi, Mordechai Segev, and Yonina C. Eldar, "Sparsity-based super-resolution microscopy from correlation information," Opt. Express 26, 18238-18269 (2018)

Sliding Frank-Wolfe

Algorithm 1: *Sliding Frank-Wolfe.*

Entrées: Acquisition $y \in \mathcal{H}$, nombre d'itérations K, $\lambda > 0$

Initialisation : $m^{[0]} = 0$ $N^{[k]} = 0$

for *Récurrence pour l'étape k,* $0 \leq k \leq K$ **do**

Pour $m^{[k]} = \sum_{i=1}^{N^{[k]}} a_i^{[k]} \delta_{x_i^{[k]}}$ telle que $a_i^{[k]} \in \mathbb{R}$, $x_i^{[k]} \in \mathcal{X}$, trouver $x_*^{[k]} \in \mathcal{X}$ tel que :

$$x_*^{[k]} \in \underset{x \in \mathcal{X}}{\operatorname{argmax}} \left| \eta^{[k]}(x) \right| \qquad \text{où} \quad \eta^{[k]}(x) \stackrel{\text{def.}}{=} \frac{1}{\lambda} \Phi^*(\Phi m^{[k]} - y),$$

if $\left| \eta^{[k]}(x_*^{[k]}) \right| < 1$ **then**

$m^{[k]}$ est la solution du BLASSO. Stop.

else

Calculer $m^{[k+1/2]} = \sum_{i=1}^{N^{[k]}} a_i^{[k+1/2]} \delta_{x_i^{[k+1/2]}} + a_{N^{[k]}+1}^{[k+1/2]} \delta_*^{[k+1/2]}$ telle que :

$$a_i^{[k+1/2]} \in \underset{a \in \mathbb{R}^{N^{[k]}+1}}{\operatorname{argmin}} \frac{1}{2} \|y - \Phi_{x^{[k+1/2]}}(a)\|_{\mathcal{H}}^2 + \lambda \|a\|_1$$

$$\text{pour} \quad x^{[k+1/2]} \stackrel{\text{def.}}{=} \left(x_1^{[k]}, \ldots, x_{N^{[k]}}^{[k]}, x_*^{[k]} \right).$$

Calculer $m^{[k+1]} = \sum_{i=1}^{N^{[k+1]}} a_i^{[k+1]} \delta_{x_i^{[k+1]}}$ telle que :

$$(a_i^{[k+1]}, x_i^{[k+1]}) \in \underset{(a,x) \in R}{\operatorname{argmax}} \frac{1}{2} \|y - \Phi_{x^{[k+1/2]}}(a)\|_{\mathcal{H}}^2 + \lambda \|a\|_1$$

end

end

Sortie: Mesure discrète $m^{[k]}$ pour k l'itération d'arrêt.

Choice of λ

λ is the only tuning parameter in BLASSO: it drives the number N of reconstructed spikes.

How do we choose it?

- Cross-validation.
- Homotopy algorithm.
- experimental choice.