Off-the-grid covariance-based super-resolution.

Bastien Laville, Laure Blanc-Féraud, Gilles Aubert

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

ICASSP 2022
Table of contents

1. Introduction
2. Off-the-grid digest
3. Dynamic off-the-grid
4. Conclusion
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*.

\[ \text{PSF } h = \text{disque d’Airy ou gaussienne.} \]
Biomedical context

Aim

Image biological structures at small scales

Physical limitation due to diffraction for bodies $< 200 \text{ nm}$: 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).
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
On an EPFL SMLM Challenge stack (10000 images, high density):

Mean of stack  
Off-the-grid  
Deep-STORM
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.
Off-the-grid digest
Inverse problem: from an acquisition, we reconstruct spike positions and amplitudes. Super-resolution grid problem:
Inverse problem: from an acquisition, we reconstruct spike positions and amplitudes. Super-resolution grid problem:

- off-the-grid super-resolution can be understood as the 'limit' of an increasingly fine grid;
- not limited by a (fine) grid, but still limited by the noise.
<table>
<thead>
<tr>
<th>Introduction</th>
<th>Off-the-grid digest</th>
<th>Dynamic off-the-grid</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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.
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.
Introduction

Off-the-grid digest

Dynamic off-the-grid

Conclusion

Elementary bricks:

\(X\) is a compact of \(R\);

how to model spikes? Dirac measures \(\delta_x\), elements of \(M(X)\), the space of signed Radon measures; topological dual of \(C(X)\) for \(\langle f, m \rangle = \int_X f \, dm\).

Generalisation of \(L_1(X)\) since \(L_1(X) \hookrightarrow M(X)\); Banach for the TV-norm:

\[|m|_{\text{TV}} = \sup_{f \in C(X), \|f\|_{\infty}, X \leq 1} \left| \int_X f \, dm \right|\]

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

\[|m|_{\text{TV}} = \sum_{i=1}^{N} |a_i|\].
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 $\delta_x$, 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 $\delta_x$, elements of $\mathcal{M}(\mathcal{X})$ the space of signed Radon measures;
- topological dual of $C(\mathcal{X})$ for $\langle f, m \rangle = \int_{\mathcal{X}} f \, dm$. Generalisation of $L^1(\mathcal{X})$ since $L^1(\mathcal{X}) \hookrightarrow \mathcal{M}(\mathcal{X})$. 
Elementary bricks:

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

$$|m|(\mathcal{X}) \overset{\text{def.}}{=} \sup \left( \int_{\mathcal{X}} f \, dm \mid f \in \mathcal{C}(\mathcal{X}), \|f\|_\infty, x \leq 1 \right).$$

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,
Let $m_{a_0,x_0} \overset{\text{def.}}{=} \sum_{i=1}^{N} a_i \delta_{x_i}$ be a discrete measure, $\Phi : \mathcal{M}(\mathcal{X}) \to L^2(\mathcal{X})$ 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)
Let $m_{a_0,x_0} \overset{\text{def.}}{=} \sum_{i=1}^{N} a_i \delta_{x_i}$ be a discrete measure, $\Phi : \mathcal{M}(\mathcal{X}) \to L^2(\mathcal{X})$ 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) and $w \in L^2(\mathcal{X})$ noise:

$$y \overset{\text{def.}}{=} \Phi m_{a_0,x_0} + w.$$
Let $m_{a_0,x_0} \overset{\text{def.}}{=} \sum_{i=1}^{N} a_i \delta_{x_i}$ be a discrete measure, $\Phi : \mathcal{M}(\mathcal{X}) \rightarrow L^2(\mathcal{X})$ 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) and $w \in L^2(\mathcal{X})$ noise:

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

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

$$\arg\min_{m \in \mathcal{M}(\mathcal{X})} \frac{1}{2} \|y - \Phi m\|_{L^2(\mathcal{X})}^2 + \lambda |m|(\mathcal{X}) \quad (\mathcal{P}_\lambda(y))$$
Let $m_{a_0, x_0} \overset{\text{def.}}{=} \sum_{i=1}^{N} a_i \delta_{x_i}$ be a discrete measure, $\Phi : \mathcal{M}(\mathcal{X}) \to L^2(\mathcal{X})$ 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) and $w \in L^2(\mathcal{X})$ noise:

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

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

$$\arg\min_{m \in \mathcal{M}(\mathcal{X})} \frac{1}{2} \|y - \Phi m\|_{L^2(\mathcal{X})}^2 + \lambda |m|(\mathcal{X}) \quad (P_\lambda(y))$$

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

- acquisition stack (images in $L^2(X)$) during $[0, T]$;
Quantities at stake

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

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

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

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

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

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

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

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

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

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

Cumulants are a tool to reconstruct the positions $x_i$. Example: temporal mean $\bar{y} \overset{\text{def.}}{=} \frac{1}{T} \int_0^T y(\cdot, t) \, dt$. One have $\Phi m_{a,x} = \bar{y}$ where $m_{a,x} \overset{\text{def.}}{=} \sum_{i=1}^{N} \bar{a}_i \delta_{x_i}$ 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:
Let $R_y$ be the temporal covariance, $\forall u, v \in \mathcal{X}$ we get:

$$R_y(u, v) \overset{\text{def.}}{=} \frac{1}{T} \int_0^T (y(u, t) - \bar{y}(u)) (y(v, t) - \bar{y}(v)) \, dt$$

$$= \ldots \quad \text{(independence of fluctuations [Dertinger10])}$$

$$= \sum_{i=1}^N M_i \quad a_i \quad \text{variance}$$

$$= \int_\mathcal{X} h(u - x) h(v - x) \, dm_{M,x}(x)$$

$$= \Lambda m_{M,x}(u, v).$$
Build the variational problem

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

$$R_y(u, v) \stackrel{\text{def.}}{=} \frac{1}{T} \int_0^T (y(u, t) - \bar{y}(u)) (y(v, t) - \bar{y}(v)) \, dt$$

$$= \ldots \quad \text{(independence of fluctuations [Dertinger10])}$$

$$= \sum_{i=1}^N M_i \quad \text{variance}$$

$$= \int_{\mathcal{X}} h(u - x) h(v - x) \, dm_{M,x}(x)$$

$$= \Lambda m_{M,x}(u, v).$$

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

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

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

$$
\arg\min_{m \in \mathcal{M}(\mathcal{X})} T_\lambda(m) \overset{\text{def.}}{=} \frac{1}{2} \| R_y - \Lambda(m) \|_{L^2(\mathcal{X}^2)}^2 + \lambda |m|(\mathcal{X}) \quad (Q_\lambda(y))
$$
BLASSO on cumulants

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

$$\arg\min_{m \in M(\mathcal{X})} T_{\lambda}(m) \overset{\text{def.}}{=} \frac{1}{2} \| R_y - \Lambda(m) \|_{L^2(\mathcal{X}^2)}^2 + \lambda |m|(\mathcal{X})$$  \hfill (Q_{\lambda}(y))

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

**Proposition**

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

- if $\Delta \gtrsim 1, 1\sigma$ [Bendory16] for mean-based reconstruction;
- if $\Delta \gtrsim 1, 1\sigma / \sqrt{2}$ for covariance-based reconstruction $(Q_{\lambda}(y))$: **better!**
Numerical 2D results SOFItool

Test on 2D tubulins from ISBI challenge 2016:
- stack of 1000 acquisitions $64 \times 64$ simulated by SOFItool;
- 8700 emitters scattered along the tubulins; **high** background noise + Poisson noise at $4 + \text{Gaussian noise at } 1 \times 10^{-2}$. SNR $\approx 10$ db.
Numerical 2D results SOFItool

Test on 2D tubulins from ISBI challenge 2016:

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

Off-the-grid digest

Dynamic off-the-grid

Conclusion

Figure 1: Ground-truth

Figure 2: \((Q_\lambda(y))\)

Figure 3: SRRF [Culley18]
Conclusion
Take home statements:

- off-the-grid methods squeeze all the 'information' out of the acquisition $y$: no discretisation drawback;
- allow performing structural assumption on the minimiser;
- strong results for existence and uniqueness of BLASSO solution;
- one efficient numerical algorithm: Sliding Frank-Wolfe.
Take home statements:

- off-the-grid methods squeeze all the ’information’ out of the acquisition $y$: no discretisation drawback;
- allow performing structural assumption on the minimiser;
- strong results for existence and uniqueness of BLASSO solution;
- one efficient numerical algorithm: Sliding Frank-Wolfe.

Outlook:

- spikes moving in position w.r.t time: *tracking* problem;
- theory only suited for spikes and sets: what about other source structures (e.g. curves)?


Lenaic Chizat. Sparse Optimization on Measures with Over-parameterized Gradient Descent. 2020


Bibliography IV

Algorithm 1: Sliding Frank-Wolfe.

Entrées: Acquisition $y \in H$, nombre d’itérations $K$, $\lambda > 0$

1 Initialisation : $m[0] = 0 \ N[k] = 0$

2 for Récurrence pour l’étape $k$, $0 \leq k \leq K$ do
3 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 :
4 $$x_∗[k] \in \arg\max_{x \in \mathcal{X}} \eta[k](x) \quad \text{où} \quad \eta[k](x) \overset{\text{def.}}{=} \frac{1}{\lambda} \Phi^∗(\Phi m[k] - y),$$
5 if $|\eta[k](x_∗[k])| < 1$ then
6 $m[k]$ est la solution du BLASSO. Stop.
7 else
8 Calculer $m[k+1/2] = \sum_{i=1}^{N[k]} a_i[k+1/2] \delta_{x_i[k+1/2]} + a_∗[k+1/2] \delta_{x_∗[k+1/2]}$ telle que :
9 $$a_i[k+1/2] \in \arg\min_{a \in \mathbb{R}^{N[k]+1}} \frac{1}{2} \|y - \Phi x_i[k+1/2](a)\|_H^2 + \lambda \|a\|_1$$
10 pour $x_i[k+1/2] \overset{\text{def.}}{=} (x_1[k], \ldots, x_N[k], x_∗[k])$.
11 Calculer $m[k+1] = \sum_{i=1}^{N[k+1]} a_i[k+1] \delta_{x_i[k+1]}$ telle que :
12 $$(a_i[k+1], x_i[k+1]) \in \arg\max_{(a,x) \in \mathbb{R}} \frac{1}{2} \|y - \Phi x_i[k+1/2](a)\|_H^2 + \lambda \|a\|_1$$
13 end
9 end

Sortie: Mesure discrète $m[k]$ pour $k$ l’itération d’arrêt.