The Lasso yields pairs of spikes on thin grids

Vincent Duval
INRIA Rocquencourt
MOKAPLAN

Gabriel Peyré
CNRS / Université Paris-Dauphine
MOKAPLAN/CEREMADE

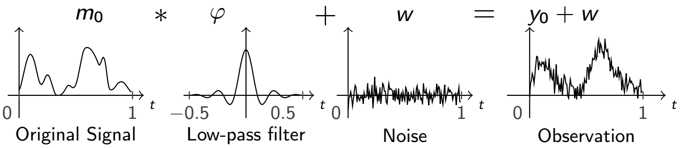
SPARS 2015
Cambridge

Observation model

- Consider a signal m_0 defined on $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ (i.e. $[0, 1)$ with periodic boundary condition).
- Observation :

$$\Phi m_0 + w = \int_{\mathbb{T}\times\mathbb{T}} \varphi(\cdot, y)\mathrm{d}m_0(y) + w \quad \text{where} \quad \varphi \text{ is smooth and } \textbf{known}.$$

- Example: Convolution

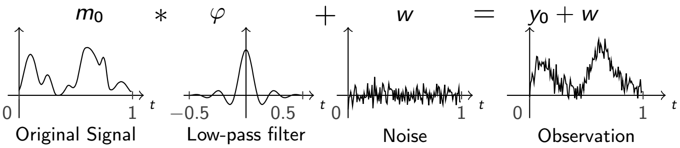


Observation model

- Consider a signal m_0 defined on $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ (i.e. $[0,1)$ with periodic boundary condition).
- Observation :

$$\Phi m_0 + w = \int_{\mathbb{T}\times\mathbb{T}} \varphi(\cdot, y)\mathrm{d}m_0(y) + w \quad \text{where} \quad \varphi \text{ is smooth and } \textbf{known}.$$

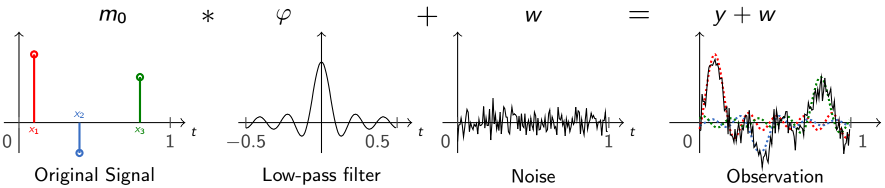
- Example: Convolution



- Goal: recover m_0 from the observation $y_0 + w = \Phi m_0 + w$ (or simply $y_0 = \Phi m_0$)
- Ill-posed problem:
  - the low pass filter might not be invertible ($\hat{\varphi}_n = 0$ for some frequency n)
  - even though, the problem is ill-conditioned ($|\hat{\varphi}_n| \ll |\hat{\varphi}_0|$ for high frequencies n)

Assumption

- **Assumption:** the signal m_0 is **sparse**



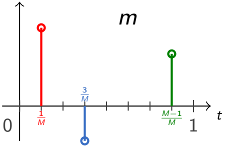
$$m_0 = \sum_{i=1}^{N} \alpha_i \delta_{x_i}, \quad \text{where} \begin{cases} \alpha_i \in \mathbb{R}, \\ x_i \in \mathbb{T}, \\ N \in \mathbb{N} \text{ is small.} \end{cases}$$

so that we observe $y + w = \sum_{i=1}^{N} \alpha_i \varphi(\cdot, x_i) + w$.

- **Idea:** Look for a **sparse** signal m such that $\Phi m \approx y_0 + w$ (or y_0).

Discretization

Define a finite grid $\mathcal{G} = \{\frac{i}{M};\ 0 \leqslant i \leqslant M-1\} \subset \mathbb{T}$, and consider signals of the form $m = \sum_{i=0}^{M-1} a_i \delta_{\frac{i}{M}}$.



Candidate Signal

- Write

$$\Phi m = \sum_{i=0}^{M-1} a_i \varphi\left(\cdot, \frac{i}{M}\right)$$

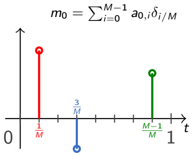
$$= \underbrace{\left(\varphi \;\middle|\; \varphi(\cdot, \tfrac{1}{M}) \;\middle|\; \dots \;\middle|\; \varphi(\cdot, \tfrac{M-1}{M}) \right)}_{\Phi_{\mathcal{G}}} \underbrace{\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{M-1} \end{pmatrix}}_{a}.$$

- Equivalent paradigm: Look for a **sparse** vector $a \in \mathbb{R}^M$ such that $\Phi_{\mathcal{G}} a \approx y_0$ (or $\Phi_{\mathcal{G}} a \approx y_0 + w$).

Discrete ℓ^1 regularization

Define

$$\|m\|_{\ell^1(\mathcal{G})} = \begin{cases} \sum_{i=0}^{M-1} |a_i| & \text{if } m = \sum_{i=0}^{M-1} a_i \delta_{i/M}, \\ +\infty & \text{otherwise.} \end{cases}$$



- **Basis Pursuit** [Chen & Donoho (94)]

$$\inf_{m \in \mathcal{M}(\mathbb{T})} \|m\|_{\ell^1(\mathcal{G})} \text{ such that } \Phi m = y_0 \qquad (\mathcal{P}_0^{\mathcal{G}}(y_0))$$

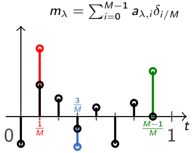
- **LASSO** [Tibshirani (96)] or **Basis Pursuit Denoising** [Chen et al. (99)]

$$\inf_{m \in \mathcal{M}(\mathbb{T})} \lambda \|m\|_{\ell^1(\mathcal{G})} + \frac{1}{2}\|\Phi m - (y_0 + w)\|_2^2 \qquad (\mathcal{P}_\lambda^{\mathcal{G}}(y_0 + w))$$

Discrete ℓ^1 regularization

Define

$$\|m\|_{\ell^1(\mathcal{G})} = \begin{cases} \sum_{i=0}^{M-1} |a_i| & \text{if } m = \sum_{i=0}^{M-1} a_i \delta_{i/M}, \\ +\infty & \text{otherwise.} \end{cases}$$



- **Basis Pursuit** [Chen & Donoho (94)]

$$\inf_{m \in \mathcal{M}(\mathbb{T})} \|m\|_{\ell^1(\mathcal{G})} \text{ such that } \Phi m = y_0 \qquad (\mathcal{P}_0^{\mathcal{G}}(y_0))$$

- **LASSO** [Tibshirani (96)] or **Basis Pursuit Denoising** [Chen et al. (99)]

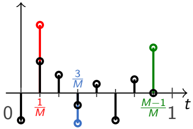
$$\inf_{m \in \mathcal{M}(\mathbb{T})} \lambda \|m\|_{\ell^1(\mathcal{G})} + \frac{1}{2}\|\Phi m - (y_0 + w)\|_2^2 \qquad (\mathcal{P}_\lambda^{\mathcal{G}}(y_0 + w))$$

ℓ^2-robustness (Grasmair et al. (2011))

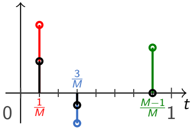
If $m_0 = \sum_i a_{0,i}\delta_{i/M}$ is the unique solution to $\mathcal{P}_0^{\mathcal{G}}(y_0)$, and $m_\lambda = \sum_i a_{\lambda,i}\delta_{i/M}$ is a solution to $\mathcal{P}_\lambda^{\mathcal{G}}(y_0 + w)$, then $\|a_\lambda - a_0\|_2 = \mathcal{O}(\|w\|_2)$ for $\lambda = C\|w\|_2$.

Robustness of the support (discrete problem)

$$m_0 = \sum_{k=0}^{M-1} a_{0,k}\delta_{k/M}$$



No support recovery

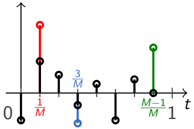


Support recovery

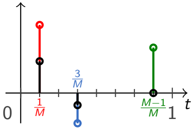
Can one guarantee that $\operatorname{Supp} m_\lambda = \operatorname{Supp} m_0$?

Robustness of the support (discrete problem)

$$m_0 = \sum_{k=0}^{M-1} a_{0,k}\delta_{k/M}$$



No support recovery



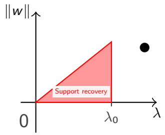
Support recovery

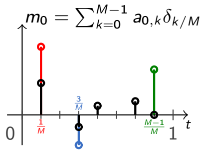
Can one guarantee that $\operatorname{Supp} m_\lambda = \operatorname{Supp} m_0$?

- Sufficient conditions [Tropp (06), Dossal & Mallat (05)],
- Almost necessary and sufficient [Fuchs (04)],
- Or look at the *minimal norm certificate*.

Fuchs theorem

For $m_0 = \sum_{i=1}^{M} a_{0,i}\delta_{x_{\mathbf{0},i}}$, assume that $\Phi_{x_0} \stackrel{\text{def.}}{=} (\varphi(\cdot, x_{0,1}), \ldots \varphi(\cdot, x_{0,N}))$ has full rank.





Theorem (Fuchs (04))

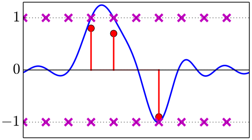
If $|\eta_F(\frac{k}{M})| < 1$ for all k such that $\frac{k}{M} \notin \{x_{0,1}, \ldots, x_{0,N}\}$, then m_0 is the unique solution to $\mathcal{P}_0^{\mathcal{G}}(y_0)$, and there exists $\gamma > 0$, $\lambda_0 > 0$ such that for $0 \leqslant \lambda \leqslant \lambda_0$ and $\|w\|_2 \leqslant \gamma\lambda$,

- *The solution m_λ to $\mathcal{P}_\lambda^{\mathcal{G}}(y_0 + w)$ is unique.*
- Supp m_λ = Supp m_0, *that is* $m_\lambda = \sum_{i=1}^{N} \alpha_{\lambda,i}\delta_{x_{\mathbf{0},i}}$, *and* $\operatorname{sign}(\alpha_{\lambda,i}) = \operatorname{sign}(\alpha_{0,i})$,
- $\alpha_\lambda = \alpha_0 + \Phi_{x_0}^{+} w - \lambda(\Phi_{x_0}^{*}\Phi_{x_0})^{-1}\operatorname{sign}(\alpha_0)$.

If $|\eta_F(\frac{k}{M})| > 1$ for some k, the support is not stable.

Fuchs theorem

For $m_0 = \sum_{i=1}^{M} a_{0,i}\delta_{x_{\mathbf{0},i}}$, assume that
$\Phi_{x_\mathbf{0}} \overset{\text{def.}}{=} (\varphi(\cdot, x_{0,1}), \ldots \varphi(\cdot, x_{0,N}))$
has full rank.



$$\eta_F \overset{\text{def.}}{=} \Phi^* p_F \quad \text{where} \quad p_F \overset{\text{def.}}{=} \operatorname{argmin}\{\|p\|_{L^2(\mathbb{T})}; (\Phi^* p)(x_{0,i}) = \operatorname{sign}(\alpha_{0,i})\}$$
$$= \Phi_{x_\mathbf{0}}^{+,*} s.$$

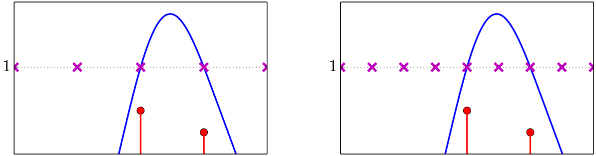
Theorem (Fuchs (04))

If $|\eta_F(\frac{k}{M})| < 1$ for all k such that $\frac{k}{M} \notin \{x_{0,1}, \ldots, x_{0,N}\}$, then m_0 is the unique solution to $\mathcal{P}_0^{\mathcal{G}}(y_0)$, and there exists $\gamma > 0$, $\lambda_0 > 0$ such that for $0 \leqslant \lambda \leqslant \lambda_0$ and $\|w\|_2 \leqslant \gamma\lambda$,

- *The solution m_λ to $\mathcal{P}_\lambda^{\mathcal{G}}(y_0 + w)$ is unique.*
- $\operatorname{Supp} m_\lambda = \operatorname{Supp} m_0$, *that is* $m_\lambda = \sum_{i=1}^{N} \alpha_{\lambda,i}\delta_{x_{\mathbf{0},i}}$, *and* $\operatorname{sign}(\alpha_{\lambda,i}) = \operatorname{sign}(\alpha_{0,i})$,
- $\alpha_\lambda = \alpha_0 + \Phi_{x_\mathbf{0}}^{+} w - \lambda(\Phi_{x_\mathbf{0}}^* \Phi_{x_\mathbf{0}})^{-1} \operatorname{sign}(\alpha_0)$.

If $|\eta_F(\frac{k}{M})| > 1$ for some k, the support is not stable.

But. . .



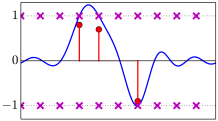
When the grid is too thin, the Fuchs criterion cannot hold
$\Rightarrow$ the support is *not stable*.

Question

What is the support at low noise when the Fuchs criterion does not hold?

Need to study the *minimal norm certificate*.

The minimal norm certificate



Assume that m_0 is a solution to $\mathcal{P}_0(y_0)$.

Define the *minimal norm certificate* on $\mathcal{G}$

$$\eta_0^{\mathcal{G}} \overset{\text{def.}}{=} \Phi^* p_0^{\mathcal{G}} \text{ where } p_0^{\mathcal{G}} \overset{\text{def.}}{=} \operatorname{argmin}\{\|p\|_{L^2(\mathbb{T})};\ (\Phi^* p)(x_{0,i}) = \operatorname{sign}(\alpha_{0,i}) \text{ for } 1 \leqslant i \leqslant N$$
$$\text{and } \left|(\Phi^* p)\left(\frac{k}{M}\right)\right| \leqslant 1 \text{ for } 0 \leqslant k \leqslant M-1\}.$$

General principle

- If $\left|\eta_0^{\mathcal{G}}(\frac{k}{M})\right| < 1$ for all k such that $k/M \notin \{x_{0,1}, \ldots, x_{0,N}\}$, there is a low noise regime with support recovery.
- If $\left|\eta_0^{\mathcal{G}}(\frac{k}{M})\right| = 1$ for some k such that $k/M \notin \{x_{0,1}, \ldots, x_{0,N}\}$, then for arbitrary small values of λ, $\|w\|_{L^2(\mathbb{T})}$, a spike may appear at k/M.

The set $\{k/M; \left|\eta_0^{\mathcal{G}}(\frac{k}{M})\right| = 1\}$ is called the **extended support** on $\mathcal{G}$ (see also [Dossal (07)]).

Working on thin grids

- Consider a sequence of refining grids with vanishing stepsize:

$$\mathcal{G}_n = \left\{ \frac{k}{M_n};\ 0 \leqslant k \leqslant M_n - 1 \right\} \subset \mathbb{T} \quad \text{with} \left\{ \begin{array}{l} \mathcal{G}_n \subset \mathcal{G}_{n+1} \text{ (e.g. } M_n = \frac{1}{2^n}), \\ \lim_{n\to+\infty} M_n = +\infty, \end{array} \right.$$

- Assume that $\operatorname{Supp} m_0 \subset \mathcal{G}_n$ for n large enough, *i.e.* $x_{0,i} \in \mathcal{G}_n$ for all $1 \leqslant i \leqslant N$.

Working on thin grids

- Consider a sequence of refining grids with vanishing stepsize:

$$\mathcal{G}_n = \left\{ \frac{k}{M_n};\ 0 \leqslant k \leqslant M_n - 1 \right\} \subset \mathbb{T} \quad \text{with} \left\{ \begin{array}{l} \mathcal{G}_n \subset \mathcal{G}_{n+1} \text{ (e.g. } M_n = \frac{1}{2^n}), \\ \lim_{n\to+\infty} M_n = +\infty, \end{array} \right.$$

- Assume that $\operatorname{Supp} m_0 \subset \mathcal{G}_n$ for n large enough, *i.e.* $x_{0,i} \in \mathcal{G}_n$ for all $1 \leqslant i \leqslant N$.

Proposition (Tang & Recht (13))

The solutions of $\mathcal{P}_0^{\mathcal{G}_n}(y_0)$ (resp. $\mathcal{P}_\lambda^{\mathcal{G}_n}(y_0+w)$) weakly converge (up to subsequences) towards the solutions of $\mathcal{P}_0(y_0)$ (resp. $\mathcal{P}_\lambda(y_0+w)$).*

- **Basis Pursuit for measures** [de Castro & Gamboa (12), Candes & Fernandez-Granda (13)],

$$\inf_{m\in\mathcal{M}(\mathbb{T})} |m|(\mathbb{T}) \text{ such that } \Phi m = y_0 \qquad (\mathcal{P}_0(y_0))$$

- **LASSO for measures** [Recht et al. (12), Bredies & Pikkarainen (13), Azais et al. (13)]

$$\inf_{m\in\mathcal{M}(\mathbb{T})} \lambda |m|(\mathbb{T}) + \frac{1}{2}\|\Phi m - (y_0+w)\|_2^2 \qquad (\mathcal{P}_\lambda(y_0+w))$$

Working on thin grids

- Consider a sequence of refining grids with vanishing stepsize:

$$\mathcal{G}_n = \left\{ \frac{k}{M_n};\ 0 \leqslant k \leqslant M_n - 1 \right\} \subset \mathbb{T} \quad \text{with} \left\{ \begin{array}{l} \mathcal{G}_n \subset \mathcal{G}_{n+1} \text{ (e.g. } M_n = \frac{1}{2^n}\text{)}, \\ \lim_{n \to +\infty} M_n = +\infty, \end{array} \right.$$

- Assume that $\operatorname{Supp} m_0 \subset \mathcal{G}_n$ for n large enough, *i.e.* $x_{0,i} \in \mathcal{G}_n$ for all $1 \leqslant i \leqslant N$.

Proposition (Duval & Peyré (13))

The minimal norm certificate $\eta_0^{\mathcal{G}_n}$ for $\mathcal{P}_0^{\mathcal{G}_n}(y_0)$ converges towards the minimal norm certificate of $\mathcal{P}_0(y_0)$.

- **Basis Pursuit for measures** [de Castro & Gamboa (12), Candes & Fernandez-Granda (13)],

$$\inf_{m \in \mathcal{M}(\mathbb{T})} |m|(\mathbb{T}) \text{ such that } \Phi m = y_0 \qquad (\mathcal{P}_0(y_0))$$

- **LASSO for measures** [Recht et al. (12), Bredies & Pikkarainen (13), Azais et al. (13)]

$$\inf_{m \in \mathcal{M}(\mathbb{T})} \lambda |m|(\mathbb{T}) + \frac{1}{2} \|\Phi m - (y_0 + w)\|_2^2 \qquad (\mathcal{P}_\lambda(y_0 + w))$$

Extended support on thin grids

If m_0 is "non-degenerate", for n large enough

$$\left\{ k/M ; \left| \eta_0^{\mathcal{G}} \left(\frac{k}{M} \right) \right| = 1 \right\} \subseteq \bigcup_{i=1}^{N} \left\{ x_{0,i}, x_{0,i} + \frac{\varepsilon_{n,i}}{M_n} \right\}.$$

where $\varepsilon_{n,i} \in \{-1, 1\}^N$.
Define the *natural shift* as

$$\rho \stackrel{\text{def.}}{=} (\Phi_{x_0}'^* \Pi \Phi_{x_0}')^{-1} \Phi_{x_0}'^* \Phi_{x_0}^{+,*} \operatorname{sign}(\alpha_0).$$

Theorem (D.-Peyré (15))

If $\rho_i \neq 0$ for all $1 \leqslant i \leqslant N$, then ε does not depend on n, and is given by

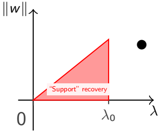
$$\varepsilon = (\operatorname{diag}(\operatorname{sign}(\alpha_0))) \operatorname{sign}(\rho).$$

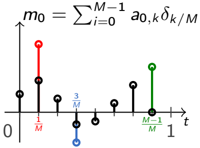
where Π is the orthogonal projector onto $(\operatorname{Im} \Phi_{x_0})^\perp$.
Moreover

$$\left\{ k/M ; \left| \eta_0^{\mathcal{G}} \left(\frac{k}{M} \right) \right| = 1 \right\} = \bigcup_{i=1}^{N} \left\{ x_{0,i}, x_{0,i} + \frac{\varepsilon_i}{M_n} \right\}.$$

Low noise "robustness" on thin grids

Under the same hypotheses:





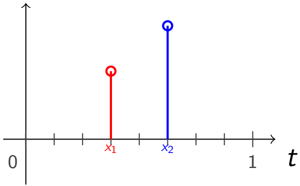
Theorem (D.-Peyré (15))

There exists $\gamma^{(n)} > 0$, $\lambda_0^{(n)} > 0$ such that for $0 \leqslant \lambda \leqslant \lambda_0^{(n)}$ and $\|w\|_2 \leqslant \gamma^{(n)}\lambda$,

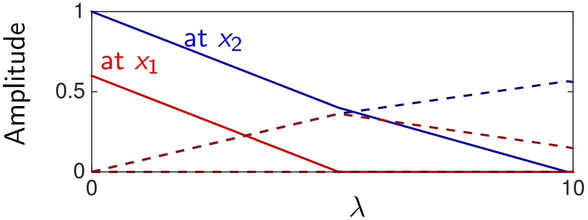
- *The solution $m_\lambda^{(n)}$ to $\mathcal{P}_\lambda^{\mathcal{G}_n}(y_0 + w)$ is unique.*
- $\operatorname{Supp} m_\lambda^{(n)} = \bigcup_{1\leqslant i\leqslant N} \left\{x_{0,i}, x_{0,i} + \frac{\varepsilon_i}{M_n}\right\}$, *that is*
  $m_\lambda^n = \sum_{i=1}^N \left(\alpha_{\lambda,i}^{(n)}\delta_{x_{0,i}} + \beta_{\lambda,i}^{(n)}\delta_{x_{0,i}+\frac{\varepsilon_i}{M_n}}\right)$, *and*
  $\operatorname{sign}(a_{\lambda,i}) = \operatorname{sign}(b_{\lambda,i}) = \operatorname{sign}(a_{0,i})$,
- $\begin{pmatrix}\alpha_\lambda^{(n)}\\ \beta_\lambda^{(n)}\end{pmatrix} = \begin{pmatrix}\alpha_0\\ 0\end{pmatrix} + \Phi^+_{x_0,x_0+\varepsilon} w - \lambda(\Phi^*_{x_0,x_0+\varepsilon}\Phi_{x_0,x_0+\varepsilon})^{-1}\begin{pmatrix}\operatorname{sign}(\alpha_{0,I})\\ \operatorname{sign}(\alpha_{0,I})\end{pmatrix}$.

In fact $\gamma^{(n)} = O(1)$ and $\lambda_0^{(n)} = O(\frac{1}{M_n})$.

Numerical example ($w = 0$)



Input Signal



Conclusion

- (Almost)-stability of the support on thin grids
- As the grid stepsize refines, stability decreases
- For a more stable "support recovery", use the continuous approach

Papers:

Exact Support Recovery for Sparse Spikes Deconvolution,
V. Duval & G. Peyré (JFoCM 2014)
Sparse Spikes Deconvolution on thin Grids
V. Duval & G. Peyré (ArXiv Preprint 2015)

Thank you for your attention!

Continuous framework for deconvolution

Using the **total variation of measures**:
$|m|(\mathbb{T}) = \sup \left\{ \int_{\mathbb{T}} \psi dm ; \psi \in C(\mathbb{T}), \|\psi\|_\infty \leqslant 1 \right\}$

- **Basis Pursuit for measures** [de Castro & Gamboa (12), Candes & Fernandez-Granda (13)],

$$\inf_{m \in \mathcal{M}(\mathbb{T})} |m|(\mathbb{T}) \text{ such that } \Phi m = y_0 \qquad (\mathcal{P}_0^\infty(y_0))$$

- **LASSO for measures** [Recht et al. (12), Bredies & Pikkarainen (13), Azais et al. (13)]

$$\inf_{m \in \mathcal{M}(\mathbb{T})} \lambda |m|(\mathbb{T}) + \frac{1}{2} \|\Phi m - (y_0 + w)\|_2^2 \qquad (\mathcal{P}_\lambda^\infty(y_0 + w))$$

$\exists$ numerical methods for solving $\mathcal{P}_0(y_0)$ and $\mathcal{P}_\lambda(y_0 + w)$, see [Bredies & Pikkarainen (13), Candes & Fernandez-Granda (13)]

Limit of the functionals

We say that $m_n \in \mathcal{M}(\mathbb{T})$ weakly * converges towards $m \in \mathcal{M}(\mathbb{T})$ if

$$\forall f \in C(\mathbb{T}), \ \lim_{n\to+\infty} \int_{\mathbb{T}} f \mathrm{d}m_n = \int_{\mathbb{T}} f \mathrm{d}m.$$

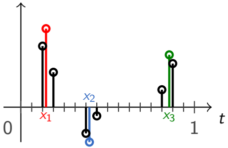
Consider a sequence $(m_n)_{n\in\mathbb{N}} \in \mathcal{M}(\mathbb{T})^{\mathbb{N}}$ such that each m_n is a minimizer of $\mathcal{P}_0^{M_n}(y_0)$ (resp. $\mathcal{P}_\lambda^{M_n}(y_0 + w)$).

Theorem ([Tang et al. 13])

*The sequence $(m_n)_{n\in\mathbb{N}}$ has convergent subsequences for the weak * convergence, and each limit point is a minimizer of $\mathcal{P}_0^{\infty}(y_0)$ (resp. $\mathcal{P}_\lambda^{\infty}(y_0 + w)$).*

Remark: In fact $\mathcal{P}_0^{M_n}(y_0)$ (resp. $\mathcal{P}_\lambda^{\infty}(y_0 + w)$ Γ-converges towards $\mathcal{P}_0^{\infty}(y_0)$ (resp. $\mathcal{P}_\lambda^{\infty}(y_0 + w)$).

Fine properties of the support



More precisely, if the solution $m^\infty = \sum_{i=1}^N \alpha_i \delta_{x_i}$ to $\mathcal{P}_\lambda^\infty(y_0 + w)$ (resp. $\mathcal{P}_0^\infty(y_0)$) is "non-degenerate",

- then the solution m_n to $\mathcal{P}_\lambda^\infty(y_0 + w)$ (resp. $\mathcal{P}_0(y_0)$) is made of pairs consecutive spikes:

$$m_n = \sum_{i=1}^N (a_i \delta_{k_i/M} + b_i \delta_{(k_i+\varepsilon_i)/M})$$

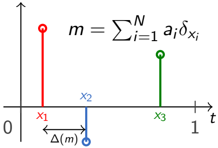
$$\text{with } \operatorname{sign}(a_i) = \operatorname{sign}(b_i) = \operatorname{sign}(\alpha_i), \varepsilon_i \in \{\pm 1\}$$

- At low noise, if the original measure is on the grid, pairs of consecutive spikes (including the original one) (see Section 1) .

Identifiability for discrete measures

Minimum separation distance of a measure m:

$$\Delta(m) = \min_{x,x' \in \mathrm{Supp}\, m, x \neq x'} |x - x'|$$



Ideal Low Pass filter: $\varphi(t) = \frac{\sin(2f_c+1)\pi t)}{\sin \pi t}$
i.e $\hat{\varphi}_n = 1$ for $|n| \leqslant f_c$, 0 otherwise.

Theorem (Candès & Fernandez-Granda (2013))

Let φ be the ideal low-pass filter. There exists a constant $C > 0$ such that, for any (discrete) measure m_0 with $\Delta(m_0) \geqslant \frac{C}{f_c}$, m_0 is the unique solution of

$$\inf_{m \in \mathcal{M}(\mathbb{T})} |m|(\mathbb{T}) \text{ such that } \Phi m = y_0 \qquad (\mathcal{P}_0(y_0))$$

where $y_0 = \Phi m_0$.

Remark: $1 \leqslant C \leqslant 1.87$.

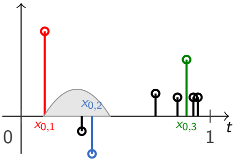
Robustness?

Weak-* robustness (Bredies & Pikkarainen (2013))

If $m_0 = \sum_i a_{0,i}\delta_{x_{0,i}}$ is the unique solution to $\mathcal{P}_0(y_0)$, m_λ is a solution to $\mathcal{P}_\lambda(y_0 + w)$, then $m_\lambda \overset{*}{\rightharpoonup} m_0$ as $\lambda \to 0^+$, $\|w\|_2^2/\lambda \to 0$.

(see also [Azais et al. (13), Fernandez-Granda (13)] for robustness of local averages in the case of the ideal LPF)

$$m_0 = \sum_{i=0}^{N} a_{0,i}\delta_{x_{0,i}}$$



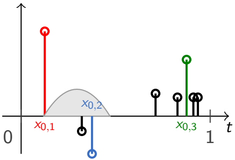
Robustness?

Weak-* robustness (Bredies & Pikkarainen (2013))

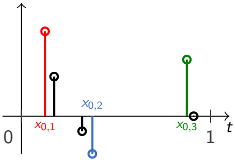
If $m_0 = \sum_i a_{0,i}\delta_{x_{0,i}}$ is the unique solution to $\mathcal{P}_0(y_0)$, m_λ is a solution to $\mathcal{P}_\lambda(y_0 + w)$, then $m_\lambda \overset{*}{\rightharpoonup} m_0$ as $\lambda \to 0^+$, $\|w\|_2^2/\lambda \to 0$.

(see also [Azais et al. (13), Fernandez-Granda (13)] for robustness of local averages in the case of the ideal LPF)

$$m_0 = \sum_{i=0}^{N} a_{0,i}\delta_{x_{0,i}}$$



No Support recovery

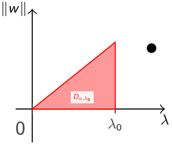


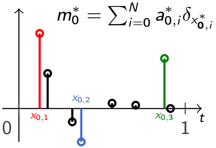
"Support recovery"

Robustness of the support (continuous problem)

For $m_0 = \sum_{i=1}^N a_{i_0} \delta_{x_{0,i}}$, define

$$\Gamma_{x_0} = \left(\varphi(\cdot - x_{0,1}), \dots \varphi(\cdot - x_{0,N}), \varphi'(\cdot - x_{0,1}), \dots \varphi'(\cdot - x_{0,N})\right)$$





Theorem (D.-Peyré 2013)

Assume that Γ_{x_0} has full rank, and that m_0 is non-degenerate.
Then there exists, $\alpha > 0$, $\lambda_0 > 0$ such that for $0 \leqslant \lambda \leqslant \lambda_0$ and $\|w\|_2 \leqslant \alpha\lambda$,

- *the solution m_λ to $\mathcal{P}_\lambda(y + w)$ is unique and has exactly N spikes, $m_\lambda = \sum_{i=1}^N a_{\lambda,i} \delta_{x_{\lambda,i}}$,*
- *the mapping $(\lambda, w) \mapsto (a_\lambda, x_\lambda)$ is C^1.*
- *the solution has the Taylor expansion*

$$\begin{pmatrix} a_\lambda \\ x_\lambda \end{pmatrix} = \begin{pmatrix} a_0 \\ x_0 \end{pmatrix} + \begin{pmatrix} I & 0 \\ 0 & \operatorname{diag} a_0^{-1} \end{pmatrix} (\Gamma_{x_0}^* \Gamma_{x_0})^{-1} \left[\begin{pmatrix} \operatorname{sign}(a_0) \\ 0 \end{pmatrix} \lambda - \Gamma_{x_0}^* w \right] + o\begin{pmatrix} \lambda \\ w \end{pmatrix}$$

Duality

$$\inf_{m\in\mathcal{M}(\mathbb{T})} |m|(\mathbb{T}) \text{ s.t. } \Phi m = y \qquad (\mathcal{P}_0(y))$$

$$(\mathcal{D}_0(y))$$

$$\inf_{m\in\mathcal{M}(\mathbb{T})} \lambda|m|(\mathbb{T}) + \frac{1}{2}\|\Phi m - y\|_2 \qquad (\mathcal{P}_\lambda(y))$$

$$(\mathcal{D}_\lambda(y))$$

Duality

$$\inf_{m \in \mathcal{M}(\mathbb{T})} |m|(\mathbb{T}) \text{ s.t. } \Phi m = y \qquad (\mathcal{P}_0(y))$$

$$\sup_{\|\Phi^* p\|_\infty \leqslant 1} \langle y,\ p \rangle \qquad (\mathcal{D}_0(y))$$

$$\inf_{m \in \mathcal{M}(\mathbb{T})} \lambda |m|(\mathbb{T}) + \frac{1}{2}\|\Phi m - y\|_2 \qquad (\mathcal{P}_\lambda(y))$$

$$\sup_{\|\Phi^* p\|_\infty \leqslant 1} \langle y,\ p \rangle - \frac{\lambda}{2}\|p\|_2^2 \qquad (\mathcal{D}_\lambda(y))$$

Duality

$$\inf_{m \in \mathcal{M}(\mathbb{T})} |m|(\mathbb{T}) \text{ s.t. } \Phi m = y \qquad (\mathcal{P}_0(y))$$

$$\sup_{\|\Phi^* p\|_\infty \leqslant 1} \langle y,\, p \rangle \qquad (\mathcal{D}_0(y))$$

$$\inf_{m \in \mathcal{M}(\mathbb{T})} \lambda |m|(\mathbb{T}) + \frac{1}{2}\|\Phi m - y\|_2 \qquad (\mathcal{P}_\lambda(y))$$

$$\sup_{\|\Phi^* p\|_\infty \leqslant 1} \langle y,\, p \rangle - \frac{\lambda}{2}\|p\|_2^2 \qquad (\mathcal{D}_\lambda(y))$$

Extremality relation:

$$\Phi^* p \in \partial |m|(\mathbb{T})$$

$$\begin{cases} \Phi^* p_\lambda \in \partial |m_\lambda|(\mathbb{T}), \\ -p_\lambda = \frac{1}{\lambda}(\Phi m_\lambda - y). \end{cases}$$

$\eta := \Phi^* p$ is a **certificate** for m

Duality

$$\inf_{m\in\mathcal{M}(\mathbb{T})} |m|(\mathbb{T}) \text{ s.t. } \Phi m = y \qquad (\mathcal{P}_0(y))$$

$$\sup_{\|\Phi^* p\|_\infty \leqslant 1} \langle y,\, p\rangle \qquad (\mathcal{D}_0(y))$$

$$\inf_{m\in\mathcal{M}(\mathbb{T})} \lambda|m|(\mathbb{T}) + \frac{1}{2}\|\Phi m - y\|_2 \qquad (\mathcal{P}_\lambda(y))$$

$$\sup_{\|\Phi^* p\|_\infty \leqslant 1} \langle y,\, p\rangle - \frac{\lambda}{2}\|p\|_2^2 \qquad (\mathcal{D}_\lambda(y))$$

Extremality relation:

$$\Phi^* p \in \partial|m|(\mathbb{T})$$

$$\begin{cases} \Phi^* p_\lambda \in \partial|m_\lambda|(\mathbb{T}), \\ -p_\lambda = \frac{1}{\lambda}(\Phi m_\lambda - y). \end{cases}$$

$\eta := \Phi^* p$ is a **certificate** for m

Subdifferential of the total variation

$$\begin{aligned}\partial|m|(\mathbb{T}) &= \{\eta \in C(\mathbb{T});\quad \forall m' \in \mathcal{M}(\mathbb{T}),\ |m'|(\mathbb{T}) \geqslant |m|(\mathbb{T}) + \langle \eta, m' - m\rangle\} \\ &= \{\eta \in C(\mathbb{T});\ \|\eta\|_\infty \leqslant 1, \forall t \in \operatorname{Supp} m_+\ \eta(t) = 1, \\ &\qquad \text{and } \forall t \in \operatorname{Supp} m_-\ \eta(t) = -1\}\end{aligned}$$

Duality

$$\inf_{m\in\mathcal{M}(\mathbb{T})} |m|(\mathbb{T}) \text{ s.t. } \Phi m = y \qquad (\mathcal{P}_0(y))$$

$$\sup_{\|\Phi^* p\|_\infty \leqslant 1} \langle y,\, p\rangle \qquad (\mathcal{D}_0(y))$$

$$\inf_{m\in\mathcal{M}(\mathbb{T})} \lambda |m|(\mathbb{T}) + \frac{1}{2}\|\Phi m - y\|_2 \qquad (\mathcal{P}_\lambda(y))$$

$$\sup_{\|\Phi^* p\|_\infty \leqslant 1} \langle y,\, p\rangle - \frac{\lambda}{2}\|p\|_2^2 \qquad (\mathcal{D}_\lambda(y))$$

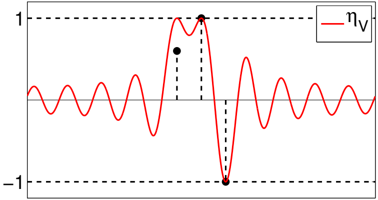
Extremality relation:

$$\Phi^* p \in \partial |m|(\mathbb{T})$$

$$\begin{cases} \Phi^* p_\lambda \in \partial |m_\lambda|(\mathbb{T}), \\ -p_\lambda = \frac{1}{\lambda}(\Phi m_\lambda - y). \end{cases}$$

Find the support of m $\leftrightsquigarrow$ Find all t such that $\eta(t) := (\Phi^* p)(t) = \pm 1$

For $m = \sum_{i=1}^N a_i \delta_{x_i}$,
$\partial |m|(\mathbb{T}) = \{\eta \in C(\mathbb{T});\, \|\eta\|_\infty \leqslant 1 \text{ and } \eta(x_i) = \operatorname{sign}(a_i) \text{ for } 1 \leqslant i \leqslant N\}$.



How to solve $\mathcal{P}_0(y)$ in the case of the ideal LPF? $\longrightarrow$ use the **Fourier coefficients**.

- Solve

$$\sup_{p \in L^2(\mathbb{T})} \langle y, p \rangle \quad \text{s.t.} \quad \sup_{t \in \mathbb{T}} |(\Phi^* p)(t)| \leqslant 1.$$

Algorithm (Candès & Fernandez-Granda)

How to solve $\mathcal{P}_0(y)$ in the case of the ideal LPF? $\longrightarrow$ use the **Fourier coefficients**.

- Solve

$$\sup_{c \in \mathbb{R}^{2f_c+1}} \Re\langle \hat{y}, c\rangle \quad \text{s.t.} \quad \sup_{t\in\mathbb{T}} \left| \sum_{n=-f_c}^{f_c} c_n e^{2i\pi nt} \right| \leqslant 1.$$

Lemma (Dumitrescu)

A causal trigonometric polynomial $\sum_{n=0}^{M-1} c_n e^{2i\pi nt}$ *is bounded by one in magnitude if and only if there exists a Hermitian matrix* $Q \in \mathbb{C}^{M\times M}$ *such that*

$$\begin{bmatrix} Q & c \\ c^* & 1 \end{bmatrix} \succeq 0 \text{ and } \sum_{i=1}^{M-j} Q_{i,i+j} = \begin{cases} 1, j = 0 \\ 0, j = 1, 2 \dots M-1 \end{cases}$$

Algorithm (Candès & Fernandez-Granda)

How to solve $\mathcal{P}_0(y)$ in the case of the ideal LPF? $\longrightarrow$ use the **Fourier coefficients**.

- Solve

$$\sup_{c \in \mathbb{R}^{2f_c+1}, Q \in \mathcal{H}_{2f_c+1}} \Re\langle \hat{y}, c\rangle \quad \text{s.t.} \begin{bmatrix} Q & c \\ c^* & 1 \end{bmatrix} \succeq 0 \text{ and } \ldots$$

Algorithm (Candès & Fernandez-Granda)

How to solve $\mathcal{P}_0(y)$ in the case of the ideal LPF? $\longrightarrow$ use the **Fourier coefficients**.

- Solve

$$\sup_{c\in\mathbb{R}^{2f_c+1}, Q\in\mathcal{H}_{2f_c+1}} \Re\langle \hat{y}, c\rangle \quad \text{s.t. } \begin{bmatrix} Q & c \\ c^* & 1 \end{bmatrix} \succeq 0 \text{ and } \ldots$$

- Find the roots of $\left|\sum_{n=-f_c}^{f_c} c_n X^{f_c+n}\right|^2 - 1$ **on the unit circle**: $e^{2i\pi x_1}, \ldots, e^{2i\pi x_N}$.

Algorithm (Candès & Fernandez-Granda)

How to solve $\mathcal{P}_0(y)$ in the case of the ideal LPF? $\longrightarrow$ use the **Fourier coefficients**.

- Solve

$$\sup_{c \in \mathbb{R}^{2f_c+1}, Q \in \mathcal{H}_{2f_c+1}} \Re\langle \hat{y}, c\rangle \quad \text{s.t.} \begin{bmatrix} Q & c \\ c^* & 1 \end{bmatrix} \succeq 0 \text{ and } \ldots$$

- Find the roots of $\left|\sum_{n=-f_c}^{f_c} c_n X^{f_c+n}\right|^2 - 1$ **on the unit circle**: $e^{2i\pi x_1}, \ldots, e^{2i\pi x_N}$.
- Solve the system $\sum_{n=1}^{N} a_n e^{2i\pi k x_n} = \hat{y}_k$ for $-f_c \leqslant k \leqslant f_c$

Algorithm (Candès & Fernandez-Granda)

How to solve $\mathcal{P}_0(y)$ in the case of the ideal LPF? $\longrightarrow$ use the **Fourier coefficients**.

- Solve

$$\sup_{c\in\mathbb{R}^{2f_c+1}, Q\in\mathcal{H}_{2f_c+1}} \Re\langle \hat{y}, c\rangle \quad \text{s.t.} \begin{bmatrix} Q & c \\ c^* & 1 \end{bmatrix} \succeq 0 \text{ and } \dots$$

- Find the roots of $\left|\sum_{n=-f_c}^{f_c} c_n X^{f_c+n}\right|^2 - 1$ **on the unit circle**: $e^{2i\pi x_1}, \dots, e^{2i\pi x_N}$.
- Solve the system $\sum_{n=1}^{N} a_n e^{2i\pi k x_n} = \hat{y}_k$ for $-f_c \leqslant k \leqslant f_c$

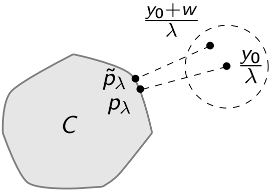
There is a variant for $\mathcal{P}_\lambda(y)$ (Azais et al., 2013)

Example

EXAMPLE

The dual as a projection

The set $C = \{p \in L^2(\mathbb{T}),\ \|\Phi^* p\|_\infty \leqslant 1\}$ is non-empty closed and convex.



Rewrite $$\sup_{\|\Phi^* p\|_\infty \leqslant 1} \langle y,\ p\rangle - \frac{\lambda}{2}\|p\|_2^2 \qquad (\mathcal{D}_\lambda(y))$$

as $$\inf_{p \in C} \left\| p - \frac{y}{\lambda} \right\|_2^2 \qquad (\mathcal{D}'_\lambda(y))$$

$\rightarrow$ projection of $\frac{y}{\lambda}$ onto C!

$$C = \bigcap_{t \in \mathbb{T}, \varepsilon = \pm 1} \left\{ p \in L^2(\mathbb{T});\ \langle \varphi(\cdot - t), p\rangle \leqslant 1 \right\}$$

Consequence

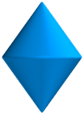
- The mapping $\frac{y}{\lambda} \mapsto p_\lambda$ is non-expansive, and $\|\tilde{p}_\lambda - p_\lambda\|_2 \leqslant \frac{\|w\|_2}{\lambda}$.
- Each "face" of C corresponds to a set of active constraints $(\Phi^* p)(t) = \pm 1$ and hence to a (signed) **support** for the solution $\tilde{m}_\lambda$.

Projection onto convex sets

C_3

C_4

$\dots$



$C = \bigcap_{n \in \mathbb{N}} C_n$

Discrete case

$C_M = \{p \in L^2(\mathbb{T});\ \left|(\Phi^* p)(\tfrac{i}{M})\right| \leqslant 1 \text{ for } 0 \leqslant i \leqslant M-1\}$ is a convex polytope.
The support is **locally constant**.

Continuous case

$C = \left\{p \in L^2(\mathbb{T}); |(\Phi^* p)(t)| \leqslant 1 \text{ for } t \in \mathbb{T}\right\}$ is convex, piecewise smooth.
The support **varies smoothly**.

Limit for the dual problem

Asumption : there is a solution to $\mathcal{D}_0(y)$ (OK if $\dim(\operatorname{Im}\Phi) < +\infty$) .

Lemma ((D.-Peyré. 2013))

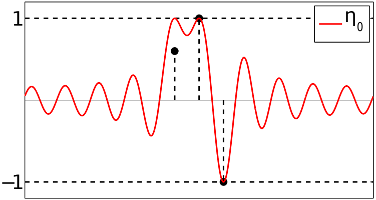
Let p_λ the unique solution of $\mathcal{D}_\lambda(y)$, and p_0 be the solution of $\mathcal{D}_0(y)$ ***with minimal norm****. Then*

$$\lim_{n\to+\infty} p_\lambda = p_0 \text{ in } L^2 \text{(strongly)}.$$

Moreover, the dual certificate $\eta_\lambda = \Phi^* p_\lambda$ and its derivatives $\eta_\lambda^{(k)}$ ($0 \leqslant k \leqslant 2$) satify:

$$\lim_{\lambda\to 0} \eta_\lambda^{(k)} = \eta_0^{(k)} \text{ in the sense of the uniform convergence.}$$

The Non Degenerate Source Condition



Definition

A measure $m_0 = \sum_{i=1}^{N} a_{0,i}\delta_{x_{\mathbf{0},i}}$ satisfies the **Non Degenerate Source Condition** if

- There exists $\eta \in \operatorname{Im}\Phi^*$ such that $\eta \in \partial|m_0|(\mathbb{T})$, or equivalently:
  - there exists a solution p to $\mathcal{D}_0(y)$,
  - m_0 is a solution to $\mathcal{P}_0(y)$
- The minimal norm certificate $\eta_0 = \Phi^* p_0$ satisfies
  - For all $s \in \mathbb{T} \setminus \{x_{0,1}, \dots x_{0,N}\}$, $|\eta_0(s)| < 1$,
  - For all $i \in \{1, \dots N\}$, $\eta_0''(x_{0,i}) \neq 0$.

Conclusion

- (Almost)-stability of the support for the deconvolution problem
- As the grid stepsize refines, stability decreases
- Try the grid free approaches the Sparse Spikes Deconvolution on Numerical tours!

www.numerical-tours.com

Papers:

Exact Support Recovery for Sparse Spikes Deconvolution,
V. Duval & G. Peyré (JFoCM 2014)
Sparse Spikes Deconvolution on thin Grids
V. Duval & G. Peyré (ArXiv Preprint 2015)

Azais, J.-M., De Castro, Y., and Gamboa, F. (2013). Spike detection from inaccurate samplings. Technical report.

Bredies, K. and Pikkarainen, H. (2013). Inverse problems in spaces of measures. *ESAIM: Control, Optimisation and Calculus of Variations*, 19:190–218.

Candès, E. J. and Fernandez-Granda, C. (2013). Towards a mathematical theory of super-resolution. *Communications on Pure and Applied Mathematics. To appear.*

Chen, S. and Donoho, D. (1994). Basis pursuit. Technical report, Stanford University.

Chen, S., Donoho, D., and Saunders, M. (1999). Atomic decomposition by basis pursuit. *SIAM journal on scientific computing*, 20(1):33–61.

de Castro, Y. and Gamboa, F. (2012). Exact reconstruction using beurling minimal extrapolation. *Journal of Mathematical Analysis and Applications*, 395(1):336–354.

Dossal, C. and Mallat, S. (2005). Sparse spike deconvolution with minimum scale. In *Proceedings of SPARS*, pages 123–126.