Exact Support Recovery for Sparse Spikes Deconvolution

Vincent Duval   Gabriel Peyré

CEREMADE
CNRS - Université Paris-Dauphine

May 14, 2014

Deconvolution

Measuring devices have a non sharp impulse response: our observations are **blurred** of a ”true ideal scene”.

- Geophysics,
- Astronomy,
- Microscopy,
- Spectroscopy,
- . . .

Image courtesy of S. Ladjal

Goal: Obtain as much detail as we can from given measurements.

Outline

1. Mathematical model
2. Robustness for the continuous problem
3. Asymptotics for the discrete problems

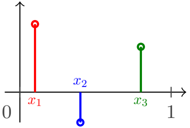
Outline

1. Mathematical model
2. Robustness for the continuous problem
3. Asymptotics for the discrete problems

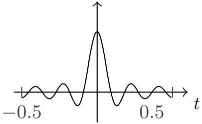
The Deconvolution Problem

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

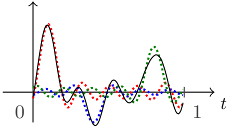
$m_0 \quad * \quad \varphi \quad = \quad y_0$



Original Signal



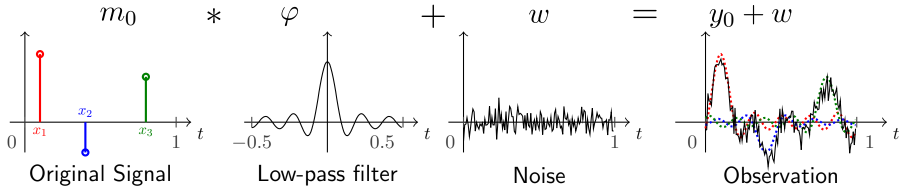
Low-pass filter



Observation

The Deconvolution Problem

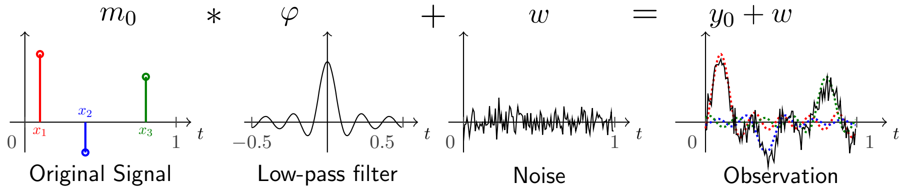
Consider a signal m_0 defined on $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ (i.e. $[0,1)$ with periodic boundary condition).



Recovery of m_0 = Ill-posed problem

The Deconvolution Problem

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



- Assumption on the signal: m_0 is a **sparse** Radon measure

$$m_0 = \sum_{i=1}^{N} a_{0,i}\delta_{x_{0,i}}, \quad \text{where} \begin{cases} a_{0,i} \in \mathbb{R}, \\ x_{0,i} \in \mathbb{T}, \\ N \in \mathbb{N} \text{ is small.} \end{cases}$$

so that we observe $y_0 + w = \sum_{i=1}^{N} a_{0,i}\varphi(\cdot - x_{0,i}) + w$.

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

Gridless Regularization

- Following (de Castro and Gamboa, 2012; Bredies and Pikkarainen, 2013; Candès and Fernandez-Granda, 2013b), we consider the continuous counterpart of the ℓ^1 norm, the **total variation of measures**:

$$\forall m \in \mathcal{M}(\mathbb{T}), \quad |m|(\mathbb{T}) = \sup \left\{ \int \psi \mathrm{d}m \; ; \; \psi \in C(\mathbb{T}), \|\psi\|_\infty \leqslant 1 \right\}.$$

  - if $dm = f(t)dt$, then $|m|(\mathbb{T}) = \int_{\mathbb{T}} |f(t)|dt$,
  - if $m = \sum_{i=1}^N a_i \delta_{x_i}$, then $|m|(\mathbb{T}) = \sum_{i=1}^N |a_i|$.

- Define $\Phi : \mathcal{M}(\mathbb{T}) \to \mathrm{L}^2(\mathbb{T})$ by

$$\Phi m : t \mapsto \int_{\mathbb{T}} \varphi(x - t) \mathrm{d}m(x).$$

Gridless Regularization

- Following (de Castro and Gamboa, 2012; Bredies and Pikkarainen, 2013; Candès and Fernandez-Granda, 2013b), we consider the continuous counterpart of the ℓ^1 norm, the **total variation of measures**:

$$\forall m \in \mathcal{M}(\mathbb{T}), \quad |m|(\mathbb{T}) = \sup \left\{ \int \psi \mathrm{d}m \; ; \; \psi \in C(\mathbb{T}), \|\psi\|_\infty \leqslant 1 \right\}.$$

Beurling minimal extrapolation (de Castro and Gamboa, 2012)

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

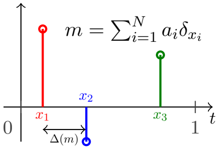
Beurling LASSO (Bredies and Pikkarainen, 2013; Azais et al., 2013)

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

Identifiability for discrete measures

Minimum separation distance of a measure m:

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



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 and Fernandez-Granda, 2013b))

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

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

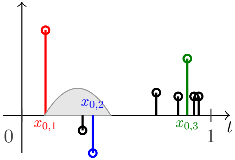
where $y = \Phi m_0$.

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

Outline

Robustness of the continuous problem?

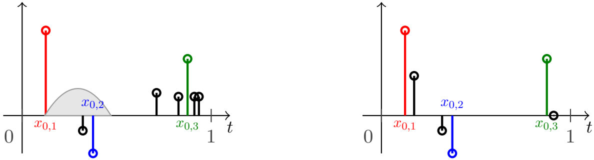
$m_0^* = \sum_{i=0}^N a_{0,i}^* \delta_{x_{0,i}^*}$



- (Bredies and Pikkarainen, 2013) the solutions to $\mathcal{P}_\lambda(y+w)$ converge to m_0 **for the weak-* topology** as $\lambda \to 0$, $\frac{\|w\|^2}{\lambda} \to 0$.
- (Azais et al., 2013): (ideal LPF) the spikes are in "boxes" around the original spikes

Robustness of the continuous problem?

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



No Support recovery

"Support recovery"

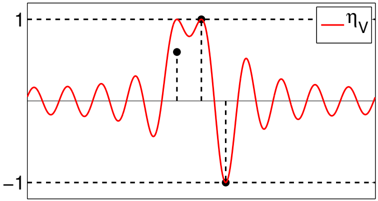
- (Bredies and Pikkarainen, 2013) the solutions to $\mathcal{P}_\lambda(y+w)$ converge to m_0 **for the weak-* topology** as $\lambda \to 0$, $\frac{\|w\|^2}{\lambda} \to 0$.
- (Azais et al., 2013): (ideal LPF) the spikes are in "boxes" around the original spikes

Subdifferential of the total variation

Subdifferential of the total variation

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

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



Subdifferential of the total variation

Subdifferential of the total variation

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

Optimality conditions

- For $\mathcal{P}_0(y_0)$:

$$\exists \eta \in \partial|m_0|(\mathbb{T}) \text{ such that } \eta \in \operatorname{Im} \Phi^*.$$

- For $\mathcal{P}_\lambda(y_0)$:

$$\exists \eta_\lambda \in \partial|m_\lambda|(\mathbb{T}) \text{ such that } \lambda\eta_\lambda + \Phi^*(\Phi m_\lambda - y_0) = 0.$$

η (resp. η_λ) is a **certificate** for m_0 (resp. m_λ), and

Find the support of m_0 (resp. m_λ) $\rightsquigarrow$ Find all t such that $\eta(t) = \pm 1$ (resp. $\eta_\lambda(t) = \pm 1$)

Minimal norm certificate

Define the **minimal norm certificate** for m_0 as

$$\eta_0 = \Phi^* p_0, \quad \text{where} \quad p_0 = \operatorname{argmin}\{\|p\|_{L^2(\mathbb{T})}; \Phi^* p \in \partial|m_0|(\mathbb{T})\}.$$

Proposition (D.P.'2013)

Let η_λ be the certificate for $\mathcal{P}_\lambda(y_0)$. Then η_λ converges to the minimal norm certificate η_0 as $\lambda \to 0^+$,

$$\lim_{\lambda \to 0^+} \eta_\lambda^{(k)} = \eta_0^{(k)}.$$

Minimal norm certificate

Define the **minimal norm certificate** for m_0 as

$$\eta_0 = \Phi^* p_0, \quad \text{where} \quad p_0 = \operatorname{argmin}\{\|p\|_{L^2(\mathbb{T})}; \Phi^* p \in \partial|m_0|(\mathbb{T})\}.$$

Proposition (D.P.'2013)

Let η_λ be the certificate for $\mathcal{P}_\lambda(y_0)$. Then η_λ converges to the minimal norm certificate η_0 as $\lambda \to 0^+$,

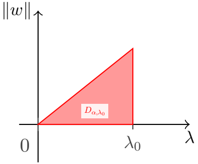
$$\lim_{\lambda \to 0^+} \eta_\lambda^{(k)} = \eta_0^{(k)}.$$

Definition (Non Degenerate Source Condition)

We say that $m_0 = \sum_{i=1}^N a_{0,i} x_{0,i}$ satisfies the **Non Degenerate Source Condition** (NDSC) if η_0 exists (= *source condition*) and

- for all $t \in \mathbb{T} \setminus \{x_{0,1}, \ldots x_{0,N}\}$, $\eta_0(t) < 1$,
- for all $i \in \{1, \ldots N\}$, $\eta_0''(x_{0,i}) \neq 0$.

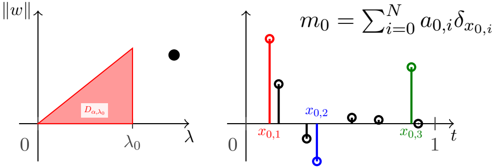
Robustness of the support



Define, for $m_0 = \sum_{i=1}^N a_{0,i}\delta_{x_{0,i}}$,

- the "matrix"
  $\Gamma_{x_0} = (\varphi(\cdot - x_{0,1}), \dots \varphi(\cdot - x_{0,N}), \varphi'(\cdot - x_{0,1}), \dots \varphi'(\cdot - x_{0,N}))$,
- the set $D_{\alpha,\lambda_0} = \{(\lambda, w) \in \mathbb{R} \times L^2(\mathbb{T}), 0 \leqslant \lambda \leqslant \lambda_0 \text{ and } \|w\| \leqslant \alpha\lambda\}$.

Robustness of the support



Theorem ((D.-P. 2013))

Let $m_0 = \sum_{i=1}^N a_i \delta_{x_{0,i}}$ *such that* m_0 *satisfies the* **Non Degenerate Source Condition** *and that* Γ_{x_0} *has full rank. Then there exists,* $\alpha > 0$, $\lambda_0 > 0$ *such that on* D_{α,λ_0}.

- *the solution* $\tilde{m}_\lambda$ *to* $\mathcal{P}_\lambda(y+w)$ *is unique and has exactly* N *spikes,* $\tilde{m}_\lambda = \sum_{i=1}^N \tilde{a}_{\lambda,i}\delta_{\tilde{x}_{\lambda,i}}$,
- *the mapping* $(\lambda, w) \mapsto (\tilde{a}_\lambda, \tilde{x}_\lambda)$ *is* C^1.

In particular, for $\lambda = \frac{1}{\alpha}\|w\|$, $|x_{0,i} - \tilde{x}_{\lambda,i}| = \mathcal{O}(\|w\|)$ and $|a_{0,i} - \tilde{a}_{\lambda,i}| = \mathcal{O}(\|w\|)$ for $1 \leqslant i \leqslant N$.

How to check the Non Degenerate Source Condition?

Let $m_0 = \sum_{i=1}^{N} a_{i,0} x_{i,0}$, and define the vanishing derivatives precertificate:

$$\eta_V = \Phi^* \Gamma_x^{+,*} \begin{pmatrix} s \\ 0_N \end{pmatrix},$$

where $\quad s = (\mathrm{sign}(a_{0,i}))_{1 \leqslant i \leqslant N} \quad$ and $\quad 0_N = (0, \ldots, 0)^T \in \mathbb{R}^N.$

Proposition (Non Degenerate Source Condition)

m_0 *satisfies the* **Non Degenerate Source Condition** *(NDSC) if and only if*

- *for all* $t \in \mathbb{T} \setminus \{x_{0,1}, \ldots x_{0,N}\}$, $\eta_V(t) < 1$,
- *for all* $i \in \{1, \ldots N\}$, $\eta_V''(x_{0,i}) \neq 0$.

In that case $\eta_0 = \eta_V$.

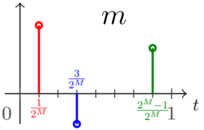
Outline

1. Mathematical model
2. Robustness for the continuous problem
3. Asymptotics for the discrete problems

Discretization

Define a finite grid $\mathcal{G}_M = \{\frac{k}{2^M};\ 0 \leqslant k \leqslant 2^M - 1\} \subset \mathbb{T}$, and impose that the support of the solution is on the grid

$$m = \sum_{k=0}^{2^M-1} b_k \delta_{\frac{k}{2^M}}.$$



Candidate Signal

- **Basis Pursuit** (Chen and Donoho, 1994)

$$\inf_{m \in \mathcal{M}(\mathbb{T})} |m|(\mathbb{T}) \text{ such that } \Phi m = y_0 \quad \text{and} \quad \operatorname{Supp} m \subset \mathcal{G}_M \qquad (\mathcal{P}_0^M(y_0))$$

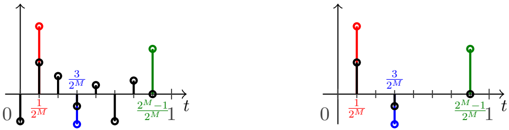
- **LASSO** (Tibshirani, 1996) or **Basis Pursuit Denoising** (Chen et al., 1999)

$$\inf_{m \in \mathcal{M}(\mathbb{T})} \lambda |m|(\mathbb{T}) + \frac{1}{2}\|\Phi m - (y_0 + w)\|_2^2 \quad \text{and} \quad \operatorname{Supp} m \subset \mathcal{G}_M \qquad (\mathcal{P}_\lambda^M(y_0 + w))$$

Robustness

- ℓ^2-robustness (Grasmair et al. (2011))
- Support recovery
  - **Exact Recovery Principle (ERC)** (Tropp, 2006)
  - **Weak Exact Recovery Principle (W-ERC)** (Dossal and Mallat, 2005)
  - **Fuchs criterion** (Fuchs, 2004)

$m_0^* = \sum_{i=0}^{2^M-1} a_{0,i}^* \delta_{i/2^M}$



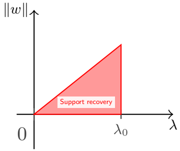
No support recovery | Support recovery

Fuchs theorem

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

$$\eta_F = \Phi^* p_F, \text{ where}$$

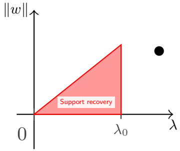
$$p_F = \operatorname{argmin}\{\|p\|_{L^2(\mathbb{T})}; (\Phi^* p)(x_{0,i}) = \operatorname{sign}(a_{i,0})\}$$
$$= \Phi^{+,*} s.$$

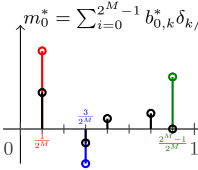


Fuchs theorem

For $m_0 = \sum_{i=1}^{N} a_{i0}\delta_{x_{i,0}}$, define

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





Theorem (Fuchs (2004))

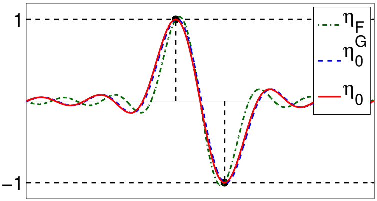
Assume that $\{\varphi(\cdot - x_{0,1}), \dots \varphi(\cdot - x_{0,N})\}$ *has full rank.*
If $|\eta_F(\frac{k}{2^M})| < 1$ *for all* k *such that* $\frac{k}{2^M} \notin \{x_{0,1}, \dots, x_{0,N}\}$, *then* m_0 *is the unique solution to* $\mathcal{P}_0^M(y)$, *and there exists* $\alpha > 0$, $\lambda_0 > 0$ *such that for* $0 \leqslant \lambda \leqslant \lambda_0$ *and* $\|w\|_2 \leqslant \alpha\lambda$,

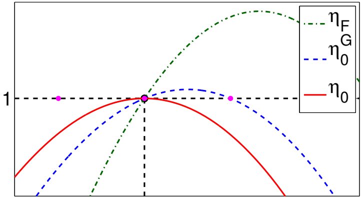
- *The solution* m_λ^M *to* $\mathcal{P}_\lambda^M(y + w)$ *is unique.*
- $\operatorname{Supp} m^M = \operatorname{Supp} m_0$, *that is* $m_\lambda^M = \sum_{i=1}^{N} a_{\lambda,i}^M \delta_{x_{0,i}}$, *and* $(\operatorname{sign} a_{\lambda,i}^*) = \operatorname{sign} a_{0,i})$,
- $a_{\lambda,I}^M = a_{0,I} + \Phi_{x_I}^+ w - \lambda(\Phi_{x_I}^* \Phi_{x_I})^{-1} \operatorname{sign}(a_{0,I})$.

If $|\eta_F(\frac{k}{2^M})| > 1$ *for some* k, *the support is not stable.*

Fuchs criterion is not valid for thin grids

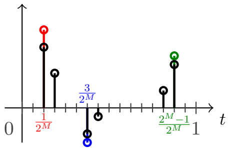
- η_F interpolates the sign
- Fuchs criterion: $|\eta_F(\frac{k}{2^M})| < 1$ for all k such that $\frac{k}{2^M} \notin \{x_{0,1}, \ldots, x_{0,N}\}$





In general, the support **cannot be stable** for thin grids, but...

Almost-robustness of the support



Theorem ((D.-P. 2013))

Let $m_0 = \sum_{0=1}^N a_{0,i}\delta_{x_{0,i}}$ *be a discrete measure, such that* $x_{0,i} \in \{\frac{k}{2^{M_0}}, 0 \leqslant k \leqslant 2^{M_0}-1\}$. *Assume that* m_0 *satisfies the* **Non Degenerate Source Condition**. *Then, for* M *large enough, there exists,* $\alpha^M > 0$, $\lambda_0^M > 0$ *such that for* $(\lambda, w) \in D_{\alpha^M,\lambda_0^M}$, *for all* m_λ^N *solution of* $\mathcal{P}_\lambda^N(y_0+w)$,

$$\operatorname{Supp} m_\lambda^N \subset \bigcup_{i=1}^N \{x_{0,i}, x_{0,i} + \frac{\varepsilon}{2^M}\} \text{ where } \varepsilon_i \in \{-1, 0, +1\}.$$

The support for small noise is made at most of **pairs of consecutive spikes**: the original one and one of its immediate neighbours.

Conclusion

- Support recovery for the continuous problem.
- Asymptotics of the discrete problems: "almost robustness" of the support
- Experiment yourself the Sparse Spikes Deconvolution on Numerical tours!

www.numerical-tours.com

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. (2013a). Super-resolution from noisy data. *To appear in Journal of Fourier Analysis and Applications*.

Candès, E. J. and Fernandez-Granda, C. (2013b). 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.