Spike deconvolution with unknown noise level

Claire Boyer, Yohann de Castro, Joseph Salmon,

UPMC, LSTA, France

April 14, 2017

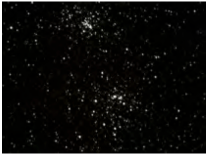
Motivations

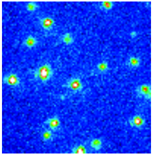
Point sources in applications

- Astronomy
- Microscopy (fluorescent molecules)
- Spectroscopy

Data

- Observed through a filter,
  in this talk, supposed to be a Fourier filter
- That would lead to a lower resolution version of the target signal





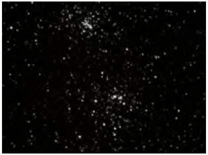
Motivations

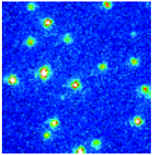
Point sources in applications

- Astronomy
- Microscopy (fluorescent molecules)
- Spectroscopy

Data

- Observed through a filter,
  in this talk, supposed to be a Fourier filter
- That would lead to a lower resolution version of the target signal





Underlying sparse objects?

- What about compressed sensing?

Setting for CS

- Ambient space: $\mathbb{R}^n$ of $\mathbb{C}^n$
- Data: $y \in \mathbb{C}^m$, m samples from spectrum
- Problem: reduce acquisition time by reducing the number of measurements while ensuring recovery
- Prior on the signal: sparsity

Setting for CS

- Ambient space: $\mathbb{R}^n$ of $\mathbb{C}^n$
- Data: $y \in \mathbb{C}^m$, m samples from spectrum
- Problem: reduce acquisition time by reducing the number of measurements while ensuring recovery
- Prior on the signal: sparsity

Minimizing the ℓ^1-norm?

- Convex relaxation of the ℓ^0-norm minimization
- Still enforce sparsity

CS problem

CS recovery: ℓ^1-minimization

To enforce sparsity of the solution

$$\min \|x\|_1 \text{ such that } \begin{cases} y = Ax & \text{(exact)} \\ \|y - Ax\|_2 \leq \eta & \text{(noisy)} \end{cases}$$

Lasso formulation

$$\min \frac{1}{2}\|y - Ax\|_2^2 + \lambda\|x\|_1$$

CS problem

CS recovery: ℓ^1-minimization

To enforce sparsity of the solution

$$\min \|x\|_1 \text{ such that } \begin{cases} y = Ax & \text{(exact)} \\ \|y - Ax\|_2 \leq \eta & \text{(noisy)} \end{cases}$$

Lasso formulation

$$\min \frac{1}{2}\|y - Ax\|_2^2 + \lambda\|x\|_1$$

- convex problem
- linear programming

CS problem

CS recovery: ℓ^1-minimization

To enforce sparsity of the solution

$$\min \|x\|_1 \text{ such that } \begin{cases} y = Ax & \text{(exact)} \\ \|y - Ax\|_2 \leq \eta & \text{(noisy)} \end{cases}$$

Lasso formulation

$$\min \frac{1}{2}\|y - Ax\|_2^2 + \lambda\|x\|_1$$

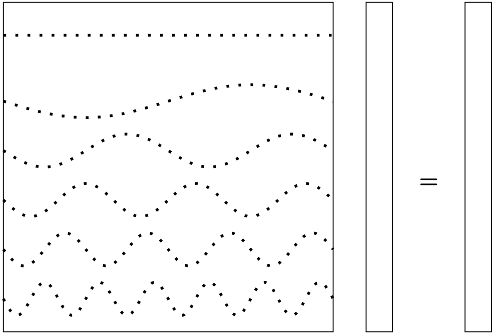
- convex problem
- linear programming

Using ℓ^1-minimization

- Can we recover the signal?
  - using ℓ^1-minimization
  - to enforce sparsity of the reconstruction
- How many measurements are required? Bound on m?

Partially measuring a sparse vector

1

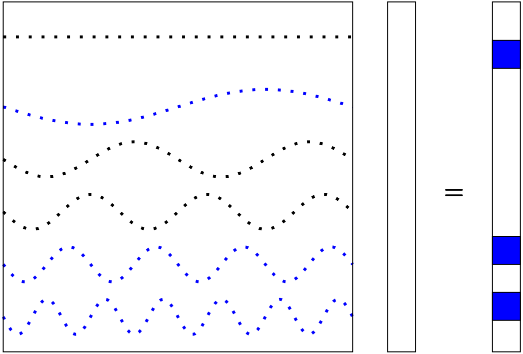


A full sensing matrix = Fourier matrix

[1] Courtesy of Carlos Fernandez-Granda

Partially measuring a sparse vector

1

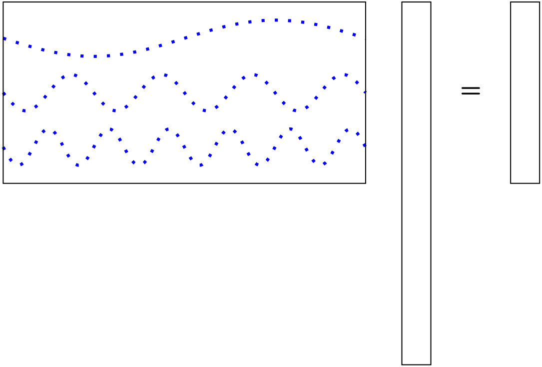


Limited number *m* of measurements

[1] Courtesy of Carlos Fernandez-Granda

Partially measuring a sparse vector

1

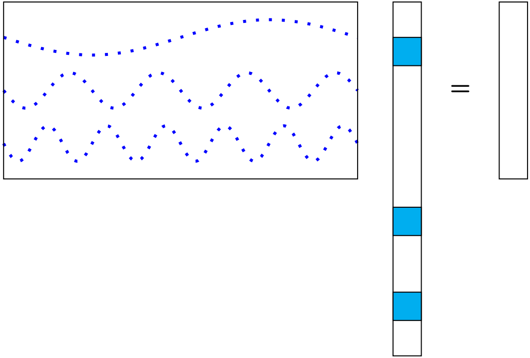


The sensing matrix (in $\mathbb{C}^{m \times n}$) is a submatrix of the Fourier one, with m rows

[1] Courtesy of Carlos Fernandez-Granda

Partially measuring a sparse vector

1

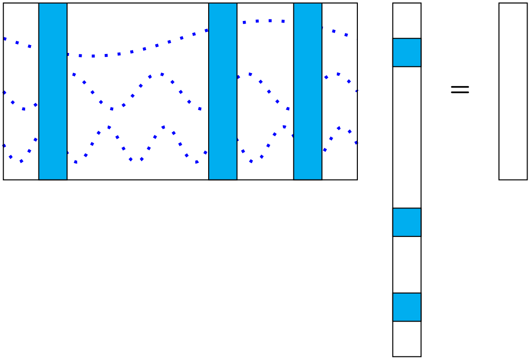


Is this sensing matrix A able to retrieve enough information on the target sparse vector x?

[1] Courtesy of Carlos Fernandez-Granda

Partially measuring a sparse vector

1

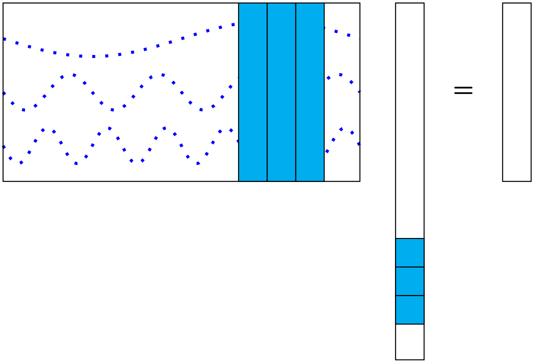


- If $S = \text{supp}(x)$, A_S is the extracted sensing matrix from A which columns are indexed by S.
- A_S should be able to keep the total energy of the input signal x.

[1] Courtesy of Carlos Fernandez-Granda

Partially measuring a sparse vector

2



Actually, this kind of matrices are "good" sensing matrices for **any** sparse vector x!

[2] Courtesy of Carlos Fernandez-Granda

Restricted isometry property

Definition

A matrix A is said to satisfy the RIP property of order s if $\exists 0 < \delta_s < 1$ such that for any sparse vector x

$$(1-\delta_s)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1+\delta_s)\|x\|_2^2$$

Restricted isometry property

Definition

A matrix A is said to satisfy the RIP property of order s if $\exists 0 < \delta_s < 1$ such that for any sparse vector x

$$(1-\delta_s)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1+\delta_s)\|x\|_2^2$$

RIP for partial Fourier matrices

Random partial Fourier matrices are known to satisfy RIP with high probability for a number m of measurements of the order

$$m \geq Cs\log^4(n) \qquad (m \ll n)$$

Restricted isometry property

Definition

A matrix A is said to satisfy the RIP property of order s if $\exists 0 < \delta_s < 1$ such that for any sparse vector x

$$(1-\delta_s)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1+\delta_s)\|x\|_2^2$$

RIP for partial Fourier matrices

Random partial Fourier matrices are known to satisfy RIP with high probability for a number m of measurements of the order

$$m \geq Cs\log^4(n) \qquad (m \ll n)$$

- One can efficiently subsample with random Fourier measurements and ensure
  - exact recovery
  - robust recovery

Showing that ℓ^1-minimization work?

Let's focus on the exact recovery problem: given $y = Ax \in \mathbb{C}^m$, one wants to show that x is the unique solution of

$$\min_z \|z\|_1 \text{ such that } y = Az$$

First order optimality condition

x is solution if

$$0 \in \partial \|\cdot\|_1(x) + \mathrm{Im} A^*$$

Since

$$\partial\|\cdot\|_1(x) = \{g \in \mathbb{S}_\infty, \langle g, x\rangle = \|x\|_1\},$$

x is solution if $\exists v \in \mathrm{Im} A^*$, (i.e. $v = A^*h$) such that

$$\begin{cases} v_i & = \mathrm{sign}(x_i), \quad \forall i \in \mathrm{supp}(x), \\ \|v\|_\infty & \leq 1 \end{cases}$$

Showing that ℓ^1-minimization work?

Let's focus on the exact recovery problem: given $y = Ax \in \mathbb{C}^m$, one wants to show that x is the unique solution of

$$\min_z \|z\|_1 \text{ such that } y = Az$$

Uniqueness

x is solution if $\exists v \in \mathrm{Im} A^*$, (i.e. $v = A^* h$) such that

$$\begin{cases} v_i & = \mathrm{sign}(x_i), \quad \forall i \in T = \mathrm{supp}(x), \\ \|v_{T^c}\|_\infty & < 1 \end{cases}$$

Showing that ℓ^1-minimization work?

Let's focus on the exact recovery problem: given $y = Ax \in \mathbb{C}^m$, one wants to show that x is the unique solution of

$$\min_z \|z\|_1 \text{ such that } y = Az$$

Uniqueness

x is solution if $\exists v \in \mathrm{Im} A^*$, (i.e. $v = A^* h$) such that

$$\begin{cases} v_i & = \text{sign}(x_i), \quad \forall i \in T = \text{supp}(x), \\ \|v_{T^c}\|_\infty & < 1 \end{cases}$$

- v is called a dual certificate
- Proof based on the construction of such a v

UPMC SORBONNE UNIVERSITÉS

Showing that ℓ^1-minimization work?

Let's focus on the exact recovery problem: given $y = Ax \in \mathbb{C}^m$, one wants to show that x is the unique solution of

$$\min_z \|z\|_1 \text{ such that } y = Az$$

Uniqueness

x is solution if $\exists v \in \mathrm{Im} A^*$, (i.e. $v = A^* h$) such that

$$\begin{cases} v_i & = \mathrm{sign}(x_i), \quad \forall i \in T = \mathrm{supp}(x), \\ \|v_{T^c}\|_\infty & < 1 \end{cases}$$

- v is called a dual certificate
- Proof based on the construction of such a v

RIP ensures that a dual certificate exists and can be constructed

- with $m = O(s \log^4(n))$ measurements
- for any sparse vector and any sign pattern.

Conclusion on CS

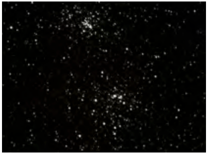
Some limitations

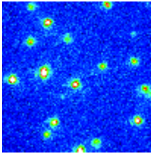
- Finite-dimensional setting: requires a grid
- One does not always have the choice of the sensing operator: it is often deterministic.

Conclusion on CS

Some limitations

- Finite-dimensional setting: requires a grid
- One does not always have the choice of the sensing operator: it is often deterministic.





Off-the-grid: motivations for super-resolution/spike deconvolution

- Point sources do not live on a cartesian grid.

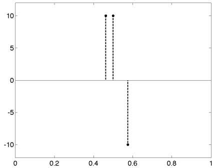
What is spike deconvolution?

Setting

- Aim: reconstruct a discrete measure

$$\mu^0 = \sum_{i=1}^{s} a_i^0 \delta_{t_i^0}$$

  - $\mathrm{supp}(\mu^0) = \{t_1^0, \ldots, t_s^0\} \subset \mathbb{T} \rightsquigarrow [0,1]$
  - $(a_i^0)_{1 \le i \le s} \in \mathbb{C}^s$



What is spike deconvolution?

Setting

- Aim: reconstruct a discrete measure

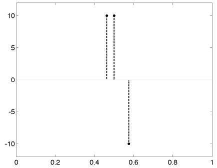
$$\mu^0 = \sum_{i=1}^{s} a_i^0 \delta_{t_i^0}$$

  - $\mathrm{supp}(\mu^0) = \{t_1^0, \dots, t_s^0\} \subset \mathbb{T} \rightsquigarrow [0,1]$
  - $(a_i^0)_{1 \leq i \leq s} \in \mathbb{C}^s$

- From linear measurements

$$y = \Phi\mu^0 + \varepsilon$$

  - Φ is a filter
  - ε is a complex Gaussian vector, $\varepsilon = \varepsilon^{(1)} + i\varepsilon^{(2)}$ with $\varepsilon^{(i)} \sim \mathcal{N}(0, \sigma_0^2 \mathrm{Id})$



What is spike deconvolution?

Setting

- Aim: reconstruct a discrete measure

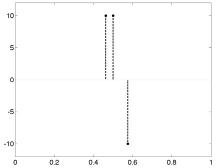
$$\mu^0 = \sum_{i=1}^{s} a_i^0 \delta_{t_i^0}$$

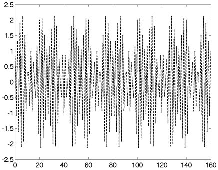
  - $\mathrm{supp}(\mu^0) = \{t_1^0, \ldots, t_s^0\} \subset \mathbb{T} \rightsquigarrow [0,1]$
  - $(a_i^0)_{1 \leq i \leq s} \in \mathbb{C}^s$

- From linear measurements

$$y = \Phi\mu^0 + \varepsilon$$

  - Φ is a filter
  - ε is a complex Gaussian vector, $\varepsilon = \varepsilon^{(1)} + i\varepsilon^{(2)}$ with $\varepsilon^{(i)} \sim \mathcal{N}(0, \sigma_0^2 \mathrm{Id})$

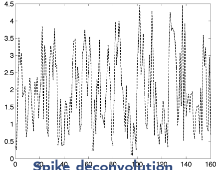




This talk

- Low-pass filter : $\Phi = \mathcal{F}_n$

$$\mathcal{F}_n(\mu) := (c_k(\mu))_{|k| \leq f_c}, \quad c_k(\mu) := \int_{\mathbb{T}} \exp(-2\pi \imath k t)\mu(\mathrm{d}t)$$



How to reconstruct μ^0?

- Is there a (continuous) analog of the ℓ^1-norm in the space of measures to ensure the measure to be discrete?
- That can be implemented efficiently?

The total variation norm

- Formal definition:

$$\|\mu\|_{TV} = \sup \left| \sum_j \mu(B_j) \right|$$

over all finite partitions (B_j) of $[0, 1]$

- If $\mu = \sum_i a_i \delta_{t_i}$, then $\|\mu\|_{TV} = \sum_i |a_i|$
- $\neq$ the total variation in image processing

Beurling minimal extrapolation

Beurling minimal extrapolation (1938)

$$\mu^\star \in \underset{\mu}{\operatorname{argmin}} \|\mu\|_{TV} \quad \text{s.t.} \quad \int_{\mathbb{T}} \Phi d\mu = y$$

with $\Phi = (\varphi_1, \ldots, \varphi_n)$

- Infinite-dimensional variable μ
- Finitely many constraints

Beurling minimal extrapolation

Beurling minimal extrapolation (1938)

$$\mu^\star \in \underset{\mu}{\operatorname{argmin}} \|\mu\|_{TV} \quad \text{s.t.} \quad \int_{\mathbb{T}} \Phi d\mu = y$$

with $\Phi = (\varphi_1, \ldots, \varphi_n)$

- Infinite-dimensional variable μ
- Finitely many constraints

Fenchel dual program

$$\hat{u} \in \arg\max \langle y, u \rangle \quad \text{s.t.} \quad \left\| \sum_{k=0}^{n} u_k \varphi_k \right\|_\infty \leq 1$$

- Finite-dimensional variable
- Infinitely many constraints
- Strong duality

How to reconstruct μ^0?

ℓ^1-deconvolution: Beurling lasso (Blasso)

$$\min_{\mu} \frac{1}{2}\|y - \mathcal{F}_n(\mu)\|_2^2 + \lambda\|\mu\|_{TV}$$

- $\|\mu\|_{TV} = \sup |\sum_j \mu(B_j)|$ over all finite partitions (B_j) of $[0,1]$
- $\|\mu\|_{TV} = \sum_{i=1}^{s} |a_i|$ when the measure is discrete

Algorithms

- proximal-based [Bredies, Pikkarainen 2012]
- root-finding [Candés, Fernandez-Granda 2012]

How to reconstruct μ^0?

ℓ^1-deconvolution: Beurling lasso (Blasso)

$$\min_{\mu} \frac{1}{2}\|y - \mathcal{F}_n(\mu)\|_2^2 + \lambda\|\mu\|_{TV}$$

- $\|\mu\|_{TV} = \sup |\sum_j \mu(B_j)|$ over all finite partitions (B_j) of $[0,1]$
- $\|\mu\|_{TV} = \sum_{i=1}^{s} |a_i|$ when the measure is discrete

Algorithms

- proximal-based [Bredies, Pikkarainen 2012]
- root-finding [Candés, Fernandez-Granda 2012]

How to reconstruct μ^0?

ℓ^1-deconvolution: Beurling lasso (Blasso)

$$\min_{\mu} \frac{1}{2}\|y - \mathcal{F}_n(\mu)\|_2^2 + \lambda\|\mu\|_{TV}$$

- $\|\mu\|_{TV} = \sup |\sum_j \mu(B_j)|$ over all finite partitions (B_j) of $[0,1]$
- $\|\mu\|_{TV} = \sum_{i=1}^s |a_i|$ when the measure is discrete

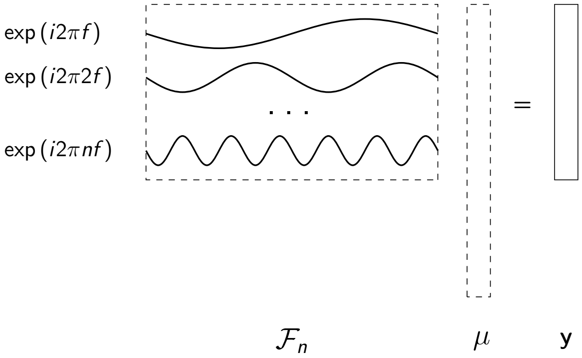
Algorithms

- proximal-based [Bredies, Pikkarainen 2012]
- root-finding [Candés, Fernandez-Granda 2012]

Others based on Prony's methods : MUSIC, ESPRIT, FRI (but lack of robustness...)

What kind of measures $\mathcal{F}_n$ is able to retrieve?

[3]

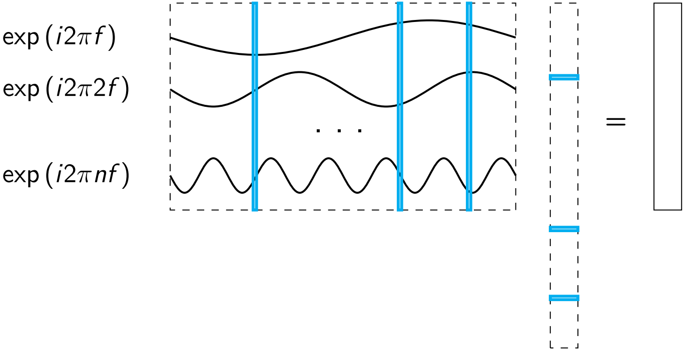


The sampling operator gives the first Fourier coefficients of the target measure

[3] Courtesy of Carlos Fernandez-Granda

What kind of measures $\mathcal{F}_n$ is able to retrieve?

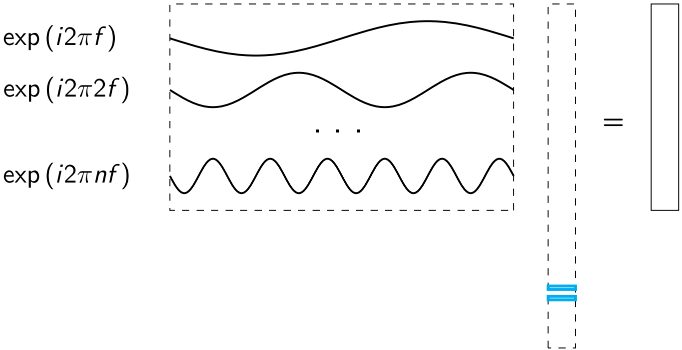
3



[3] Courtesy of Carlos Fernandez-Granda

What kind of measures $\mathcal{F}_n$ is able to retrieve?

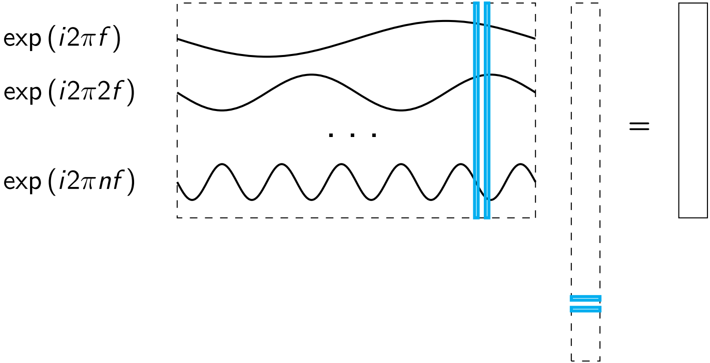
3



[3] Courtesy of Carlos Fernandez-Granda

What kind of measures $\mathcal{F}_n$ is able to retrieve?

3



For such a discrete measure the sampling operator will be ill-posed

[3]Courtesy of Carlos Fernandez-Granda

Key feature for super-resolution

- Deterministic sampling operator
- Should be well-posed for spread out supports

Definition: minimum separation

The minimum separation for a measure μ such that $\mathrm{supp}(\mu^0) = \{t_1^0, \dots, t_{s_0}^0\}$ is the following quantity

$$\Delta = \inf_{i \neq j} d(t_i^0, t_j^0).$$

Minimum separation condition

- If $\Delta < 2/(n-1)$, the problem is ill-posed
- If $\Delta > 2/(n-1)$, the problem is well-posed

Showing that Blasso works?

Focus on

$$\min \|\mu\|_{TV} \text{ such that } \mathcal{F}_n(\mu) = y$$

$\mu = \sum_i a_i \delta_{t_i}$ with support T is the unique solution if there exists a trigonometric polynomial p,

$$p = \mathcal{F}_n^* c = \sum_\ell c_\ell e^{-2i\pi\ell\cdot}, \text{ for some } c \in \mathbb{C}^n$$

that satisfies

$$(*) \begin{cases} p(t_i) = \frac{a_i}{|a_i|} & \text{if } t_i \in T \\ |p(t)| < 1 & \text{if } t_i \notin T \end{cases}$$

$$(*) \begin{cases} \Re\left(\int_{\mathbb{T}} \overline{p}(t)\mu(\mathrm{d}t)\right) & = \|\mu\|_{\mathrm{TV}} \quad (p \text{ subgradient of } \|\cdot\|_{TV} \text{ at } \mu) \\ \|p\|_\infty & \leq 1 \end{cases}$$

- Q is called a dual polynomial
- Proof based on constructing such a polynomial

Dual polynomial and root-finding

Focus on

$$\min \|\mu\|_{TV} \text{ such that } \mathcal{F}_n(\mu) = y$$

$\mu = \sum_i a_i \delta_{t_i}$ with support T is the unique solution if there exists a trigonometric polynomial p,

$$p = \mathcal{F}_n^* c = \sum_\ell c_\ell e^{-2i\pi\ell\cdot}, \text{ for some } c \in \mathbb{C}^n$$

that satisfies

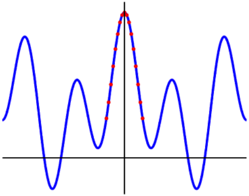
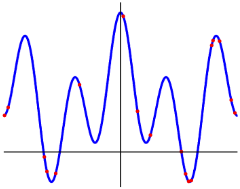
$$(*) \begin{cases} p(t_i) = \frac{a_i}{|a_i|} & \text{if } t_i \in T \\ |p(t)| < 1 & \text{if } t_i \notin T \end{cases} \quad \rightsquigarrow p'(t_i) = 0$$

Root-finding

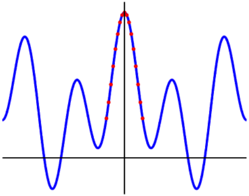
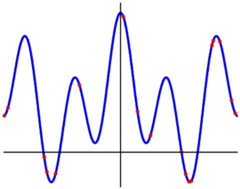
- Once the dual polynomial constructed
- The support of μ is included in the set of the roots of the polynomial derivative!

	Spike deconvolution	Compressed sensing
Setting	∞-dim	finite-dim (or easy ∞-dim)
Object of interest	a discrete measure $\mu = \sum_{i=1}^{s} a_i \delta_{t_i}$	a sparse signal (mainly) $x \in \mathbb{C}^n$

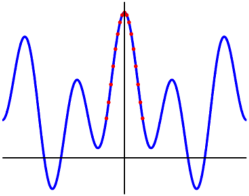
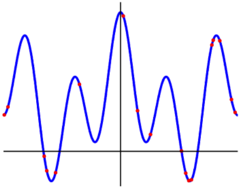
[4] illustrations from Carlos Fernandez-Granda

	Spike deconvolution	Compressed sensing
Setting	∞-dim	finite-dim (or easy ∞-dim)
Object of interest	a discrete measure $\mu = \sum_{i=1}^{s} a_i \delta_{t_i}$	a sparse signal (mainly) $x \in \mathbb{C}^n$
Measurements	low-pass filter  [4] spectrum extrapolation	random  spectrum interpolation

[4] illustrations from Carlos Fernandez-Granda

	Spike deconvolution	Compressed sensing
Setting	∞-dim	finite-dim (or easy ∞-dim)
Object of interest	a discrete measure $\mu = \sum_{i=1}^{s} a_i \delta_{t_i}$	a sparse signal (mainly) $x \in \mathbb{C}^n$
Measurements	low-pass filter  [4] spectrum extrapolation	random  spectrum interpolation
Resolution	$\min_{\mu} \frac{1}{2}\\|y - \mathcal{F}_n(\mu)\\|_2^2 + \lambda \\|\mu\\|_{TV}$ Beurling Lasso estimator	$\min_{x} \frac{1}{2}\\|y - \mathcal{F}_{\mathrm{rd}}(x)\\|_2^2 + \lambda \\|x\\|_1$ Lasso
Key feature	Minimum separation	Measurement incoherence + degree of sparsity

[4] illustrations from Carlos Fernandez-Granda

	Spike deconvolution	Compressed sensing
Setting	∞-dim	finite-dim (or easy ∞-dim)
Object of interest	a discrete measure $\mu = \sum_{i=1}^{s} a_i \delta_{t_i}$	a sparse signal (mainly) $x \in \mathbb{C}^n$
Measurements	low-pass filter [4] spectrum extrapolation	random  spectrum interpolation
Resolution	$\min_\mu \frac{1}{2}\|y - \mathcal{F}_n(\mu)\|_2^2 + \lambda \|\mu\|_{TV}$ Beurling Lasso estimator	$\min_x \frac{1}{2}\|y - \mathcal{F}_{\mathrm{rd}}(x)\|_2^2 + \lambda \|x\|_1$ Lasso
Key feature	Minimum separation	Measurement incoherence + degree of sparsity
Algorithm	SDP Root-finding (...)	LP

[4] illustrations from Carlos Fernandez-Granda

Contributions

- Handling unknown noise level

Contributions

- Handling unknown noise level
- Assessing the noise level using the Rice method for a non-Gaussian process

Contributions

- Handling unknown noise level
- Assessing the noise level using the Rice method for a non-Gaussian process
- Prediction & strong localization accuracy

Contributions

- Handling unknown noise level
- Assessing the noise level using the Rice method for a non-Gaussian process
- Prediction & strong localization accuracy
- Numerics : the root-finding can always be done (open question of [1, Eq. (4.5)])

E. J. Candès and C. Fernandez-Granda.
Towards a mathematical theory of super-resolution.
Communications on Pure and Applied Mathematics, 67(6):906–956, 2014.

Contributions

- Handling unknown noise level
- Assessing the noise level using the Rice method for a non-Gaussian process
- Prediction & strong localization accuracy
- Numerics : the root-finding can always be done (open question of [1, Eq. (4.5)])

Claire Boyer, Yohann de Castro, and Joseph Salmon.
Adapting to unknown noise level in sparse deconvolution.
Information and Inference, abs/1606.04760, 2017.

E. J. Candès and C. Fernandez-Granda.
Towards a mathematical theory of super-resolution.
Communications on Pure and Applied Mathematics, 67(6):906–956, 2014.

The proposed method

Data

$$y = \mathcal{F}_n \mu^0 + \varepsilon$$

- $y \in \mathbb{C}^n$ with $n = 2f_c + 1$
- ε is a complex Gaussian vector, $\varepsilon = \varepsilon^{(1)} + i\varepsilon^{(2)}$ with $\varepsilon^{(j)} \sim \mathcal{N}(0, \sigma_0^2 \mathrm{Id})$

The proposed method

Data

$$y = \mathcal{F}_n \mu^0 + \varepsilon$$

- $y \in \mathbb{C}^n$ with $n = 2f_c + 1$
- ε is a complex Gaussian vector, $\varepsilon = \varepsilon^{(1)} + i\varepsilon^{(2)}$ with $\varepsilon^{(j)} \sim \mathcal{N}(0, \sigma_0^2 \mathrm{Id})$

For handling the unknown noise level, we propose the CBLasso for Concomitant Beurling Lasso.

The proposed method

Data

$$y = \mathcal{F}_n \mu^0 + \varepsilon$$

- $y \in \mathbb{C}^n$ with $n = 2f_c + 1$
- ε is a complex Gaussian vector, $\varepsilon = \varepsilon^{(1)} + i\varepsilon^{(2)}$ with $\varepsilon^{(j)} \sim \mathcal{N}(0, \sigma_0^2 \mathrm{Id})$

For handling the unknown noise level, we propose the CBLasso for Concomitant Beurling Lasso.

CBLasso

$$(\hat{\mu}, \hat{\sigma}) \in \underset{(\mu,\sigma) \in E^* \times \mathbb{R}_{++}}{\operatorname{argmin}} \frac{1}{2n\sigma} \|y - \mathcal{F}_n(\mu)\|_2^2 + \frac{\sigma}{2} + \lambda \|\mu\|_{\mathrm{TV}} ,$$

- convex
- Owen 07, Antoniadis 10, Belloni Chernozhukov Wang 11, Sun Zhang 12, Chretien Darses 14, van de Geer 15
  $\rightsquigarrow$ square-root lasso and scaled-lasso

Where is the square-root?

Why square-root?

The CBLasso reads ([2] $\sim$ scaled lasso)

$$(\hat{\mu},\hat{\sigma}) \in \operatorname*{argmin}_{(\mu,\sigma)\in E^*\times\mathbb{R}_{++}} \frac{1}{2n\sigma}\|y-\mathcal{F}_n(\mu)\|_2^2+\frac{\sigma}{2}+\lambda\|\mu\|_{\mathrm{TV}} \ .$$

When the solution is reached for $\hat{\sigma}>0$, ([1] $\sim$ square-root lasso)

$$\hat{\sigma}=\|y-\mathcal{F}_n(\hat{\mu})\|_2/\sqrt{n}$$

$$\hat{\mu}\in\operatorname*{argmin}_{\mu\in E^*}\|y-\mathcal{F}_n(\mu)\|_2/\sqrt{n}+\lambda\|\mu\|_{\mathrm{TV}}$$

A. Belloni, V. Chernozhukov, and L. Wang.
Square-root Lasso: Pivotal recovery of sparse signals via conic programming.
Biometrika, 98(4):791–806, 2011.

T. Sun and C.-H. Zhang.
Scaled sparse linear regression.
Biometrika, 99(4):879–898, 2012.

Compatibility limits

Sufficient conditions for oracle inequalities

- RIP $\Longrightarrow$ REC $\Longrightarrow$ Compatibility
- Compatibility condition :

$$\mathcal{C}(L,S) > 0$$

$$\mathcal{C}(L,S) := \inf\left\{|S|\|\mathcal{F}_n(\nu)\|_2^2/n \quad \text{s.t.} \quad \text{supp}(\nu) = S,\ \|\nu_S\|_{\mathrm{TV}} = 1,\ \|\nu_{S^c}\|_{\mathrm{TV}} \leq L\right\} \quad .$$

- Consider ν as a difference of two measures: $\hat{\mu}$ and μ^0

Compatibility limits

Sufficient conditions for oracle inequalities

- RIP $\Longrightarrow$ REC $\Longrightarrow$ Compatibility
- Compatibility condition :

$$\mathcal{C}(L,S) > 0$$

$$\mathcal{C}(L,S) := \inf\left\{|S|\|\mathcal{F}_n(\nu)\|_2^2/n \quad \text{s.t.} \quad \text{supp}(\nu) = S,\ \|\nu_S\|_{\text{TV}} = 1,\ \|\nu_{S^c}\|_{\text{TV}} \leq L\right\} \quad .$$

- Consider ν as a difference of two measures: $\hat{\mu}$ and μ^0

Compatibility does not hold

- $\nu_\varepsilon = \delta_{-\varepsilon} + \delta_\varepsilon$
- ~~Compatibility~~ $\Longrightarrow$ ~~REC~~ $\Longrightarrow$ ~~RIP~~
- highly coherent designs: close Dirac masses share almost the same Fourier coefficients

Non-uniform approach

- Measure-dependent reconstruction

Standard assumptions

Assumption (Sampling rate condition).

$$\lambda \mathrm{SNR} \leq \frac{\sqrt{17}-4}{2} \simeq 0.0616$$

with $\mathrm{SNR} := \dfrac{\|\mu^0\|_{\mathrm{TV}}}{\sqrt{\mathbb{E}[\|\varepsilon\|_2^2]/n}} = \dfrac{\|\mu^0\|_{\mathrm{TV}}}{\sqrt{2}\sigma_0}$.

- $\lambda \geq 2\sqrt{2\log n/n} \quad \Longrightarrow \quad n/\log n \geq C\,\mathrm{SNR}^2$

Standard assumptions

Assumption (Sampling rate condition).

$$\lambda \mathrm{SNR} \leq \frac{\sqrt{17}-4}{2} \simeq 0.0616$$

with $\mathrm{SNR} := \frac{\|\mu^0\|_{\mathrm{TV}}}{\sqrt{\mathbb{E}[\|\varepsilon\|_2^2]/n}} = \frac{\|\mu^0\|_{\mathrm{TV}}}{\sqrt{2}\sigma_0}$.

- $\lambda \geq 2\sqrt{2\log n/n} \quad \Longrightarrow \quad n/\log n \geq C\,\mathrm{SNR}^2$

Assumption (No-overfitting).

$$\hat{\sigma} > 0$$

Standard assumptions

Assumption (Sampling rate condition).

$$\lambda \mathrm{SNR} \leq \frac{\sqrt{17}-4}{2} \simeq 0.0616$$

with $\mathrm{SNR} := \frac{\|\mu^0\|_{\mathrm{TV}}}{\sqrt{\mathbb{E}[\|\varepsilon\|_2^2]/n}} = \frac{\|\mu^0\|_{\mathrm{TV}}}{\sqrt{2}\sigma_0}.$

- $\lambda \geq 2\sqrt{2\log n/n} \quad \Longrightarrow \quad n/\log n \geq C\,\mathrm{SNR}^2$

Assumption (No-overfitting).

$$\hat{\sigma} > 0$$

Assumption (Separation condition).

For $\mathrm{supp}(\mu^0) = \{t_1^0, \ldots, t_{s_0}^0\}$,

$$\forall i \neq j, \qquad d(t_i^0, t_j^0) \geq \frac{1.26}{f_c}$$

Prediction result

Theorem (B., de Castro, Salmon, 2017).

Let $C > 2\sqrt{2}$. Set $C' > 0$, that may depend on C. Assume

- *the sampling rate condition,*
- *the separation condition.*

The estimator $\hat{\mu}$ solution to CBLasso with a choice $\lambda \geq C\sqrt{\log n/n}$ satisfies

$$\frac{1}{n}\|\mathcal{F}_n(\hat{\mu} - \mu^0)\|_2^2 \leq C' s_0 \lambda^2 \sigma_0,$$

with high probability.

Prediction result

Theorem (B., de Castro, Salmon, 2017).

Let $C > 2\sqrt{2}$. Set $C' > 0$, that may depend on C. Assume

- *the sampling rate condition,*
- *the separation condition.*

The estimator $\hat{\mu}$ solution to CBLasso with a choice $\lambda \geq C\sqrt{\log n/n}$ satisfies

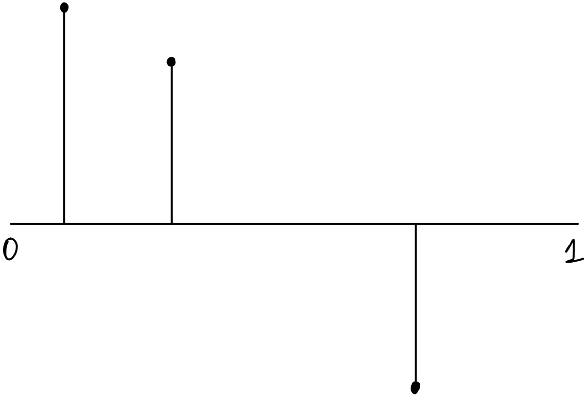
$$\frac{1}{n}\|\mathcal{F}_n(\hat{\mu} - \mu^0)\|_2^2 \leq C' s_0 \lambda^2 \sigma_0^2 = O\left(\frac{s_0 \sigma_0^2 \log n}{n}\right),$$

with high probability.

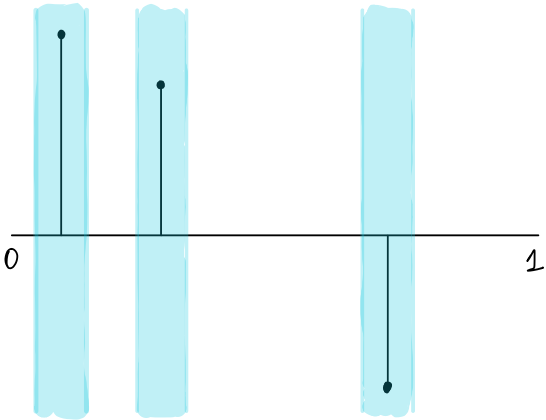
- "fast rate of convergence" of [1]

G. Tang, B. N. Bhaskar, and B. Recht.
Near minimax line spectral estimation.
Information Theory, IEEE Transactions on, 61(1):499–512, 2015.

Localization results



Localization results



- Near regions: $N = \left\{ t : \exists t_j^0, d(t, t_j^0) \leq c_1 / f_c \right\}$
- Far regions: $\mathbb{T} \setminus N$

Localization results

Near regions

$$\forall j \in [s_0], \quad N_j := \left\{ t \in [0,1];\ d(t, t_j^0) \leq \frac{c_1}{f_c} \right\},$$

with $0 < c_1 < 1.26/2$.

Far region

$$F := [0,1] \setminus \bigcup_{j \in [s_0]} N_j$$

Localization results

Near regions

$$\forall j \in [s_0], \quad N_j := \left\{ t \in [0,1];\ d(t, t_j^0) \leq \frac{c_1}{f_c} \right\},$$

with $0 < c_1 < 1.26/2$.

Far region

$$F := [0,1] \setminus \bigcup_{j \in [s_0]} N_j$$

Theorem (B., de Castro, Salmon, 2017).

Let $C > 2\sqrt{2}$. Let $C' > 0$ may depend on C. The estimator $\hat{\mu}$, solution to CBLasso with a choice $\lambda \geq C\sqrt{\log n / n}$, satisfies

1. $\forall j \in [s_0], \quad \left| a_j^0 - \sum_{\{k:\ \hat{t}_k \in N_j\}} \hat{a}_k \right| \leq C' \sigma_0 s_0 \sqrt{\log n / n}$,

Localization results

Near regions

$$\forall j \in [s_0], \quad N_j := \left\{ t \in [0,1];\ d(t, t_j^0) \le \frac{c_1}{f_c} \right\} ,$$

with $0 < c_1 < 1.26/2$.

Far region

$$F := [0,1] \setminus \bigcup_{j \in [s_0]} N_j$$

Theorem (B., de Castro, Salmon, 2017).

Let $C > 2\sqrt{2}$. Let $C' > 0$ may depend on C. The estimator $\hat{\mu}$, solution to CBLasso with a choice $\lambda \ge C\sqrt{\log n/n}$, satisfies

1. $\forall j \in [s_0], \quad \left| a_j^0 - \sum_{\{k:\ \hat{t}_k \in N_j\}} \hat{a}_k \right| \le C' \sigma_0 s_0 \sqrt{\log n/n}$,
2. $\forall j \in [s_0], \quad \sum_{\{k:\ \hat{t}_k \in N_j\}} |\hat{a}_k| \mathrm{d}^2(t_j^0, \hat{t}_k) \le C' \sigma_0 s_0 \sqrt{\log n/n}/n^2$,

Localization results

Near regions

$$\forall j \in [s_0], \quad N_j := \left\{ t \in [0,1];\ d(t, t_j^0) \leq \frac{c_1}{f_c} \right\},$$

with $0 < c_1 < 1.26/2$.

Far region

$$F := [0,1] \setminus \bigcup_{j \in [s_0]} N_j$$

Theorem (B., de Castro, Salmon, 2017).

Let $C > 2\sqrt{2}$. Let $C' > 0$ may depend on C. The estimator $\hat{\mu}$, solution to CBLasso with a choice $\lambda \geq C\sqrt{\log n / n}$, satisfies

1. $\forall j \in [s_0], \quad \left| a_j^0 - \sum_{\{k:\ \hat{t}_k \in N_j\}} \hat{a}_k \right| \leq C' \sigma_0 s_0 \sqrt{\log n / n}$,
2. $\forall j \in [s_0], \quad \sum_{\{k:\ \hat{t}_k \in N_j\}} |\hat{a}_k| \mathrm{d}^2(t_j^0, \hat{t}_k) \leq C' \sigma_0 s_0 \sqrt{\log n / n} / n^2$,
3. $\sum_{\{k:\ \hat{t}_k \in F\}} |\hat{a}_k| \leq C' \sigma_0 s_0 \sqrt{\log n / n}$,

with high probability.

Localization results

Corollary (B., de Castro, Salmon, 2017).

For any t_j^0 *in the support of* μ^0 *such that* $a_j^0 > C'\sigma_0 s_0 \lambda$*, there exists an element* $\hat{t}_k$ *in the support of* $\hat{\mu}$ *such that*

$$\mathrm{d}(t_j^0, \hat{t}_k) \leq \sqrt{\frac{C'\sigma_0 s_0 \lambda}{|a_j^0| - C'\sigma_0 s_0 \lambda}\,\frac{1}{n}}\ ,$$

with high probability.

- independent of the other spikes magnitude

Noise level estimation

Proposition (B., de Castro, Salmon, 2017).

Under the sampling rate assumption,

$$\left| \frac{\sqrt{n}\hat{\sigma}}{\|\varepsilon\|_2} - 1 \right| \leq \frac{1}{2} \ ,$$

with probability larger than $1 - \exp(-n/9)\left(\frac{2\sqrt{2}}{n} + \frac{2\sqrt{3}+3}{3}\right)$.

Some words about the proof

KKT condition

$\hat{\mu} = \sum_i \hat{a}_i \delta_{t_i}$ with support $\hat{T}$ is solution of CBLasso
if there exists a dual polynomial $\hat{p}$

$$\frac{1}{n}\mathcal{F}_n^*(y - \mathcal{F}_n(\hat{\mu})) = \hat{\sigma}\lambda\hat{p} \qquad \begin{cases} \hat{p}(t_i) = \frac{\hat{a}_i}{\hat{a}_i|} & \text{if } t_i \in \hat{T} \\ |\hat{p}(t)| < 1 & \text{if } t_i \notin \hat{T} \end{cases}$$

Some words about the proof

KKT condition

$\hat{\mu} = \sum_i \hat{a}_i \delta_{t_i}$ with support $\hat{T}$ is solution of CBLasso
if there exists a dual polynomial $\hat{p}$

$$\frac{1}{n}\mathcal{F}_n^*(y - \mathcal{F}_n(\hat{\mu})) = \hat{\sigma}\lambda\hat{p} \qquad \begin{cases} \hat{p}(t_i) = \frac{\hat{a}_i}{\hat{a}_i|} & \text{if } t_i \in \hat{T} \\ |\hat{p}(t)| < 1 & \text{if } t_i \notin \hat{T} \end{cases}$$

Some words about the proof

KKT condition

$\hat{\mu} = \sum_i \hat{a}_i \delta_{t_i}$ with support $\hat{T}$ is solution of CBLasso
if there exists a dual polynomial $\hat{p}$

$$\frac{1}{n}\mathcal{F}_n^*(y - \mathcal{F}_n(\hat{\mu})) = \hat{\sigma}\lambda\hat{p} \qquad \begin{cases} \hat{p}(t_i) = \frac{\hat{a}_i}{\hat{a}_i|} & \text{if } t_i \in \hat{T} \\ |\hat{p}(t)| < 1 & \text{if } t_i \notin \hat{T} \end{cases}$$

- Proofs are based on generalizations of the dual certificate for the noiseless problem

Sketch of proof

Spike localization

- Proofs are based on the work of [1, 2, 3]
- amended by Rice method

Prediction

- adapt the proof of [3] to our setting
- Rice method for a non-Gaussian process

J.-M. Azaïs, Y. De Castro, and F. Gamboa.
Spike detection from inaccurate samplings.
Applied and Computational Harmonic Analysis, 38(2):177–195, 2015.

C. Fernandez-Granda.
Support detection in super-resolution.
arXiv preprint arXiv:1302.3921, 2013.

G. Tang, B. N. Bhaskar, and B. Recht.
Near minimax line spectral estimation.
Information Theory, IEEE Transactions on, 61(1):499–512, 2015.

Primal-dual

Proposition.

Denoting $\mathcal{D}_n = \left\{c \in \mathbb{C}^n : \|\mathcal{F}_n^*(c)\|_\infty \leq 1, n\lambda^2\|c\|^2 \leq 1\right\}$*, the dual formulation of the CBLasso reads*

$$\hat{c} \in \arg\max_{c \in \mathcal{D}_n} \lambda \langle y, c \rangle \quad . \tag{1}$$

Then, we have the link-equation between primal and dual solutions

$$y = n\lambda\hat{\sigma}\hat{c} + \mathcal{F}_n(\hat{\mu}) \quad . \tag{2}$$

as well as a link between the coefficient and the polynomial

$$\mathcal{F}_n^*(\hat{c}) = \hat{p} \quad . \tag{3}$$

Primal-dual

Proposition.

Denoting $\mathcal{D}_n = \left\{c \in \mathbb{C}^n : \|\mathcal{F}_n^*(c)\|_\infty \leq 1, n\lambda^2\|c\|^2 \leq 1\right\}$, *the dual formulation of the CBLasso reads*

$$\hat{c} \in \arg\max_{c \in \mathcal{D}_n} \lambda \langle y, c\rangle \quad . \tag{1}$$

Then, we have the link-equation between primal and dual solutions

$$y = n\lambda\hat{\sigma}\hat{c} + \mathcal{F}_n(\hat{\mu}) \quad . \tag{2}$$

as well as a link between the coefficient and the polynomial

$$\mathcal{F}_n^*(\hat{c}) = \hat{p} \quad . \tag{3}$$

If the dual polynomial $\hat{p}$ associated is not constant

- the support of $\hat{\mu}$ is finite,
- the support of $\hat{\mu}$ is included in the set of its derivative roots, i.e. where the polynomial saturates at 1.
- the proof based on a dual certificate construction is possible.

Primal-dual : comparison with Blasso

Dual of the Blasso

$$\hat{c} \in \underset{\|\mathcal{F}_n^*(c)\|_\infty \leq 1}{\arg\max} \langle y, c \rangle \quad .$$

Dual of the CBLasso

$$\hat{c} \in \underset{\substack{\|\mathcal{F}_n^*(c)\|_\infty \leq 1 \\ n\lambda^2 \|c\|^2 \leq 1}}{\arg\max} \lambda \langle y, c \rangle \quad .$$

Primal-dual : comparison with Blasso

Dual of the Blasso

$$\hat{c} \in \operatorname*{arg\,max}_{\|\mathcal{F}_n^*(c)\|_\infty \leq 1} \langle y, c \rangle \quad .$$

Dual of the CBLasso

$$\hat{c} \in \operatorname*{arg\,max}_{\substack{\|\mathcal{F}_n^*(c)\|_\infty \leq 1 \\ n\lambda^2 \|c\|^2 \leq 1}} \lambda \langle y, c \rangle \quad .$$

Open question

When the dual polynomial is non-constant ?

For the CBLasso

We answer this.

Showing that the dual polynomial is non-constant

$|\hat{p}|^2$ is of constant modulus 1
$\Rightarrow \hat{p} = v\varphi_k$ with $v \in \mathbb{C}$ and $\varphi_k(\cdot) = \exp(2\pi\imath k\cdot)$ for some $-f_c \leq k \leq f_c$.

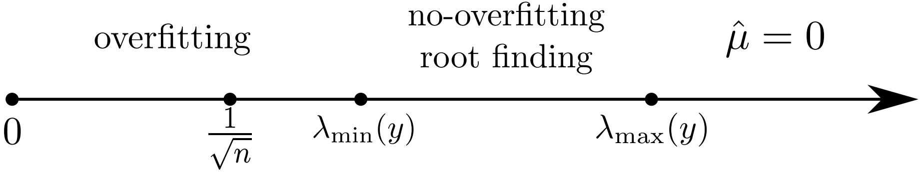
- if $|v| < 1$, using Holder's inequality on

$$\Re\left(\int_{\mathbb{T}} \overline{\hat{p}}(t)\hat{\mu}(\mathrm{d}t)\right) = \|\hat{\mu}\|_{\mathrm{TV}}$$

  leads to $\hat{\mu} = 0$.
- if $|v| = 1$, we also have $\hat{c} \in \mathcal{D}_n$, so $\|\hat{c}\|_2 \leq 1/(\sqrt{n}\lambda)$, leading to $|v| \leq 1/(\sqrt{n}\lambda)$. Since $\lambda > 2\sqrt{logn}/\sqrt{n} \Rightarrow |v| < 1$, which contradicts $|v| = 1$.

- One can then conclude that a dual polynomial of constant modulus never occurs in the CBLasso setup

Discussion on the value of λ



SDP formulation of the CBLasso

One can represent the dual feasible set $\mathcal{D}_n$ as an SDP condition.
The dual problem can be cast as follows

$$\max_{c \in \mathbb{C}^n} \lambda \langle y, c \rangle \quad \text{such that}$$

$$\exists Q \in \mathbb{C}^{n \times n} \begin{pmatrix} Q & c \\ c^* & 1 \end{pmatrix} \succcurlyeq 0 \qquad \sum_{i=1}^{n-j} Q_{i,i+j} = \left\{ \begin{array}{ll} 1 & \text{if } j = 0 \\ 0 & j = 1, \ldots, n-1. \end{array} \right.$$

SDP formulation of the CBLasso

One can represent the dual feasible set $\mathcal{D}_n$ as an SDP condition.
The dual problem can be cast as follows

$$\max_{c \in \mathbb{C}^n} \lambda \langle y, c \rangle \quad \text{such that}$$

$$\exists Q \in \mathbb{C}^{n \times n} \begin{pmatrix} Q & c \\ c^* & 1 \end{pmatrix} \succcurlyeq 0 \qquad \sum_{i=1}^{n-j} Q_{i,i+j} = \begin{cases} 1 & \text{if } j = 0 \\ 0 & j = 1, \dots, n-1. \end{cases}$$

$$\begin{pmatrix} I_n & \lambda\sqrt{n} c \\ \lambda\sqrt{n} c^* & 1 \end{pmatrix} \succcurlyeq 0$$

- The dual problem is a tractable SDP program

Proposed algorithm

Algorithm

Given the data $y \in \mathbb{C}^n$

1. solve the dual problem to find the coefficients $\hat{c}$ of the dual polynomial $\hat{p}$ (cvx Matlab toolbox);

Proposed algorithm

Algorithm

Given the data $y \in \mathbb{C}^n$

1. solve the dual problem to find the coefficients $\hat{c}$ of the dual polynomial $\hat{p}$ (cvx Matlab toolbox);
2. identify $\mathrm{supp}(\hat{\mu}) = \{\hat{t}_1, \ldots, \hat{t}_{\hat{s}}\}$ using the roots of $1 - |\hat{p}|^2$;

Proposed algorithm

Algorithm

Given the data $y \in \mathbb{C}^n$

1. solve the dual problem to find the coefficients $\hat{c}$ of the dual polynomial $\hat{p}$ (cvx Matlab toolbox);
2. identify $\mathrm{supp}(\hat{\mu}) = \{\hat{t}_1, \ldots, \hat{t}_{\hat{s}}\}$ using the roots of $1 - |\hat{p}|^2$;
3. fix the design matrix $X \in \mathbb{C}^{n \times \hat{s}}$, defined by $X_{k,j} = \overline{\varphi_k}(\hat{t}_j)$;

Proposed algorithm

Algorithm

Given the data $y \in \mathbb{C}^n$

1. solve the dual problem to find the coefficients $\hat{c}$ of the dual polynomial $\hat{p}$ (cvx Matlab toolbox);
2. identify $\mathrm{supp}(\hat{\mu}) = \{\hat{t}_1, \ldots, \hat{t}_{\hat{s}}\}$ using the roots of $1 - |\hat{p}|^2$;
3. fix the design matrix $X \in \mathbb{C}^{n \times \hat{s}}$, defined by $X_{k,j} = \overline{\varphi_k}(\hat{t}_j)$;
4. solve now the following finite-dimensional problem

$$(\hat{a}, \hat{\sigma}) \in \operatorname*{argmin}_{(a,\sigma) \in \mathbb{C}^{\hat{s}} \times \mathbb{R}_{++}} \frac{1}{2n\sigma} \|y - Xa\|_2^2 + \frac{\sigma}{2} + \lambda \|a\|_1 \ ,$$

Proposed algorithm

Algorithm

Given the data $y \in \mathbb{C}^n$

1. solve the dual problem to find the coefficients $\hat{c}$ of the dual polynomial $\hat{p}$ (cvx Matlab toolbox);
2. identify $\text{supp}(\hat{\mu}) = \{\hat{t}_1, \ldots, \hat{t}_{\hat{s}}\}$ using the roots of $1 - |\hat{p}|^2$;
3. fix the design matrix $X \in \mathbb{C}^{n \times \hat{s}}$, defined by $X_{k,j} = \overline{\varphi_k}(\hat{t}_j)$;
4. solve now the following finite-dimensional problem

$$(\hat{a}, \hat{\sigma}) \in \underset{(a,\sigma) \in \mathbb{C}^{\hat{s}} \times \mathbb{R}_{++}}{\text{argmin}} \frac{1}{2n\sigma} \|y - Xa\|_2^2 + \frac{\sigma}{2} + \lambda \|a\|_1 \ ,$$

as follows: for an initial value of $\hat{\sigma}$, until some stopping criterion, alternate the following steps

- solve the previous problem for a fixed $\hat{\sigma}$ to compute $\hat{a}$,

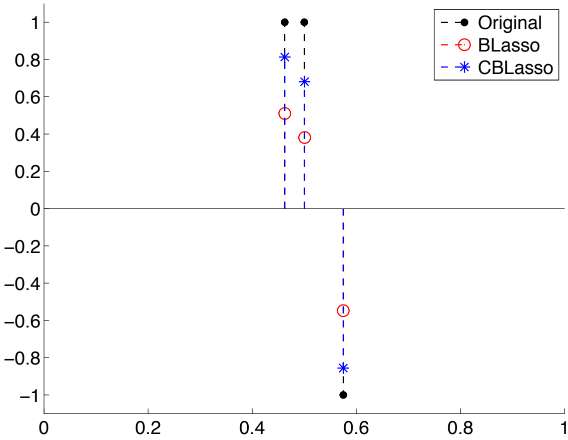
$$\hat{a} = X^+ y - \lambda \hat{\sigma} (X^* X)^{-1} \, \text{sign}(X^* \hat{c})$$

- update $\hat{\sigma} = \|y - X\hat{a}\|_2 / \sqrt{n}$ using the new value of $\hat{a}$,

Code available at

https://github.com/claireBoyer/CBLasso
http://www.lsta.upmc.fr/boyer/

Measure reconstruction



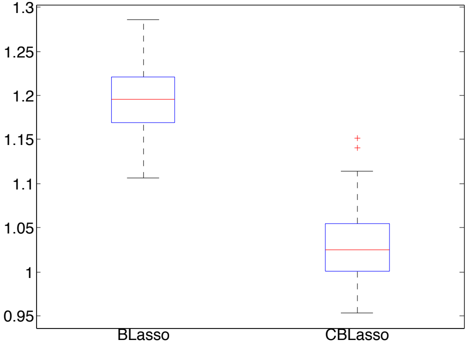
Reconstruction of a discrete measure. The original measure μ^0 is composed of 3 spikes (in black). The reconstructed measure $\hat{\mu}$ using our proposed CBLasso (in blue). In comparison, we plot the reconstructed measure using the BLasso, (in red).

Noise level estimation

$\varepsilon = \varepsilon^{(1)} + i\varepsilon^{(2)}$ with $\varepsilon^{(j)} \sim \mathcal{N}(0, \sigma_0 \mathrm{Id})$ with $\sigma_0 = 1/\sqrt{2}$.

Noise level estimation

$\varepsilon = \varepsilon^{(1)} + i\varepsilon^{(2)}$ with $\varepsilon^{(j)} \sim \mathcal{N}(0, \sigma_0 \mathrm{Id})$ with $\sigma_0 = 1/\sqrt{2}$.



Boxplot on $\hat{\sigma}$ for 100 CBLasso consistent estimations of $\sqrt{2}\sigma_0 = 1$. We compare our method to

$$\hat{\sigma}^{\mathrm{BLasso}} = \frac{\|y - \mathcal{F}_n(\hat{\mu}^{\mathrm{BLasso}})\|_2}{\sqrt{n - \hat{s}^{\mathrm{BLasso}}}}$$

proposed in [1] where $\hat{\mu}^{\mathrm{BLasso}}$ is the reconstructed measure supported on $\hat{s}^{\mathrm{BLasso}}$ spikes via BLasso.

S. Reid, R. Tibshirani, and J. Friedman.
A study of error variance estimation in lasso regression.
arXiv preprint arXiv:1311.5274, 2014.

Bias

Noise level estimation

$$\left|\frac{\sqrt{n}\hat{\sigma}}{\|\varepsilon\|_2} - 1\right| \leq \frac{1}{2} \ ,$$

with high probability.

Bias

$$\hat{\sigma} \simeq \sqrt{2}\sigma_0 \times \frac{\mathbb{E}\|g\|_2}{\sqrt{2n}} = \sqrt{2}\sigma_0 \times \frac{\Gamma(n+1/2)}{\sqrt{n}\Gamma(n)} \to \sqrt{2}\sigma_0 \,,$$

showing that $\hat{\sigma}/\sqrt{2}$ is a consistent estimator of σ_0.

Conclusion

The CBLasso

- new approach to handle unknown noise level in spike detection
- theoretical contributions :
  - prediction for CBLasso
  - localization for CBLasso
  - closing the question of constant polynomial in this setting
- numerical method is also proposed to tackle the CBLasso

Conclusion

The CBLasso

- new approach to handle unknown noise level in spike detection
- theoretical contributions :
  - prediction for CBLasso
  - localization for CBLasso
  - closing the question of constant polynomial in this setting
- numerical method is also proposed to tackle the CBLasso

Perspectives

- Choice of λ (impossible cross-validation)
- Towards nD results
- Other filters?

Conclusion

The CBLasso

- new approach to handle unknown noise level in spike detection
- theoretical contributions :
  - prediction for CBLasso
  - localization for CBLasso
  - closing the question of constant polynomial in this setting
- numerical method is also proposed to tackle the CBLasso

Perspectives

- Choice of λ (impossible cross-validation)
- Towards nD results
- Other filters?

Claire Boyer, Yohann de Castro, and Joseph Salmon.
Adapting to unknown noise level in sparse deconvolution.
Information and Inference, abs/1606.04760, 2017.

Conclusion

The CBLasso

- new approach to handle unknown noise level in spike detection
- theoretical contributions :
  - prediction for CBLasso
  - localization for CBLasso
  - closing the question of constant polynomial in this setting
- numerical method is also proposed to tackle the CBLasso

Perspectives

- Choice of λ (impossible cross-validation)
- Towards nD results
- Other filters?

Claire Boyer, Yohann de Castro, and Joseph Salmon.
Adapting to unknown noise level in sparse deconvolution.
Information and Inference, abs/1606.04760, 2017.

For the frustrated optimizers...

A subgradient descent