Orthogonal greedy algorithms for sparse reconstruction

IHES / Statistics & Machine Learning

Jan. 27, 2020

Charles SOUSSEN

CentraleSupélec, L2S / GPI

ANR Project BECOSE (2016-2020)

- Propose and analyze algorithms based on sparsity to solve ill-posed inverse problems

- Application fields :
  - Fluid mechanics
  - Nondestructive evaluation, optical spectroscopy, *etc.*

- GPI : Inverse Problems / sparsity, tomographic reconstruction

Main co-workers

- Jérôme Idier (LS2N, Nantes), El-Hadi Djermoune (CRAN, Nancy)
  Signal processing, inverse problems
- Cédric Herzet, Rémi Gribonval (INRIA Rennes)
  Sparsity, information theory
- Frédéric Champagnat (ONERA Palaiseau)
  Tomographic PIV

Summary

1. Inverse problems and sparse regularization

2. Greedy algorithms for ℓ_0 minimisation

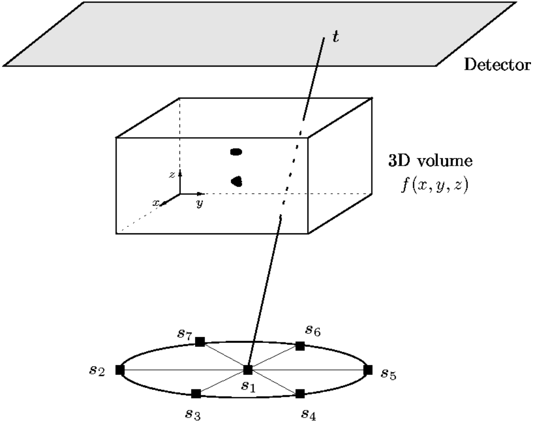
3. ℓ_0 continuation

4. Non-negative greedy algorithms

5. Exact recovery analyses

6. Conclusion

Example 1 : Computed Tomography

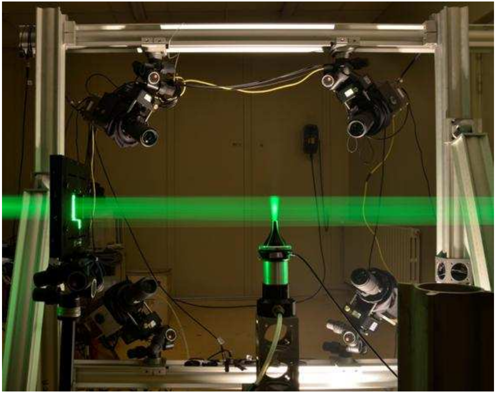


Tomographic reconstruction from limited projections

[S. & Idier, *J. Electr. Imag.*, 2008]

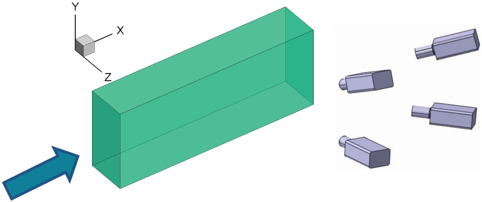
Tomographic PIV

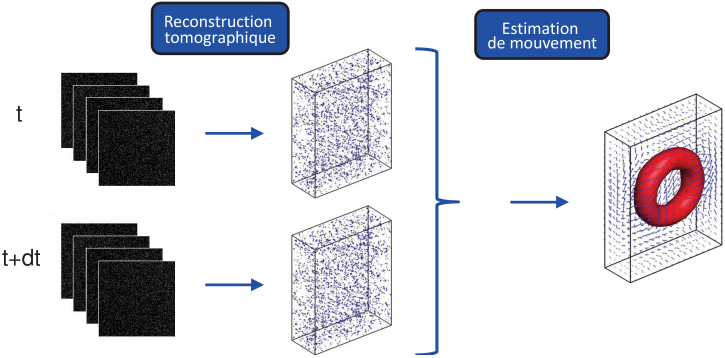
- Measure of moving fluids
- Multi-view stereo system



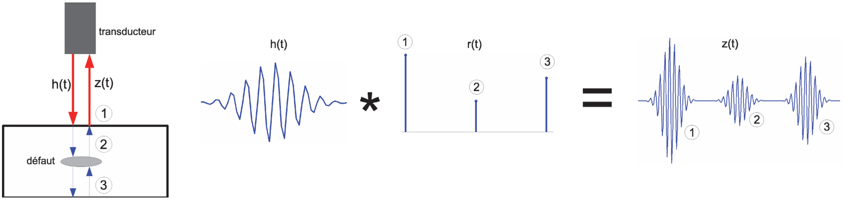
Credit : [ONERA DAAA]

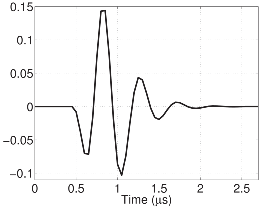
TomoPIV : principe global





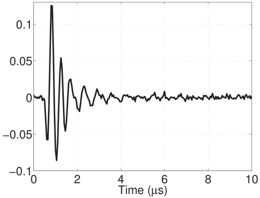
Example 2 : NDE using ultrasounds





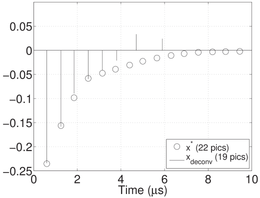
Acoustic wave

$\boldsymbol{h}$



Data

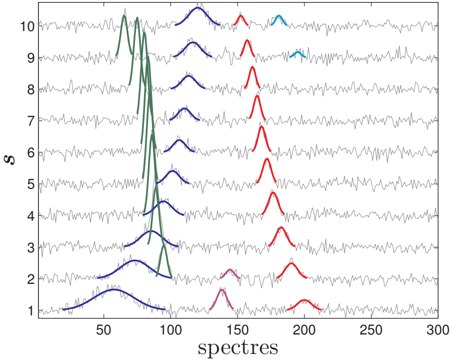
$\boldsymbol{z} \approx \boldsymbol{h} * \boldsymbol{x}^\star = \boldsymbol{A}\boldsymbol{x}^\star$



Detected pulses

$\boldsymbol{x}$

Example 3 : Signal decomposition

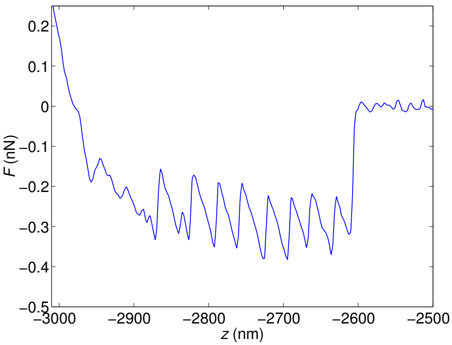


Optical spectroscopy

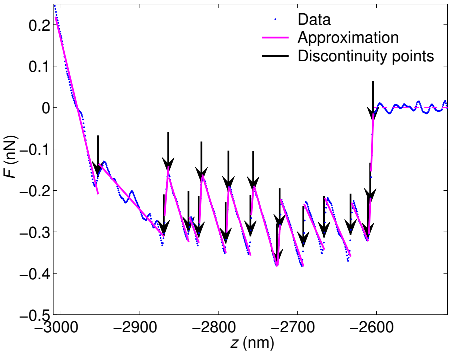
$\boldsymbol{y}_s = \boldsymbol{A}\boldsymbol{x}_s$ where $\boldsymbol{A}$: 2D dictionary (locations + wavelengths)

[Mordata *et al.*, *Sig. Process.*, 2019]

Example 4 : Signal segmentation



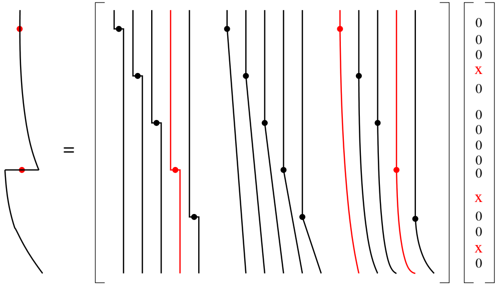
Raw data



Piecewise affine approximation

- Adaptive splines [Friedman, 1991]

Piecewise polynomial approximation



$$\boldsymbol{y} \approx \boldsymbol{A} \boldsymbol{x} = \sum_{p=0}^{P} \boldsymbol{A}^p \boldsymbol{x}^p$$

- Overcomplete dictionary $\boldsymbol{A}$, deterministic, structured
- Correlated atoms

Summary

1. Inverse problems and sparse regularization

2. Greedy algorithms for ℓ_0 minimisation

3. ℓ_0 continuation

4. Non-negative greedy algorithms

5. Exact recovery analyses

6. Conclusion

ℓ_0 minimization

- Sparsity measure : $\|\boldsymbol{x}\|_0 = \text{Card}\{i : x_i \neq 0\}$
- NP-hard problems [Natarajan, 1995] :
  - $\min_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2$ s.t. $\|\boldsymbol{x}\|_0 \leq k$ $[\boldsymbol{C}]$
  - $\min_{\boldsymbol{x}} \left\{ \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \lambda \|\boldsymbol{x}\|_0 \right\}$ $[\boldsymbol{P}]$
- Combinatorial problems

$$[\boldsymbol{C}] \iff \min_{|S| \leq k} \|\boldsymbol{P}_S^{\perp} \boldsymbol{y}\|_2^2 \quad \text{with} \quad S = \text{supp}(\boldsymbol{x})$$

- Greedy algorithms : descent algorithms

ℓ_0 minimization

NP-hard problem : $\min_{\|\boldsymbol{x}\|_0 \leqslant k} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2$

Numerical approaches [Bruckstein *et al.*, 2009] :

1. Exact minimization
2. Convex (ℓ_1) relaxation
3. Non-convex continuous optimization
4. Heuristics for ℓ_0 minimization
   - Thresholding algorithms : CoSaMP, SP, *etc.*
   - Greedy algorithms : MP, OMP, OLS, *etc.*

Orthogonal greedy algorithms

Problem : $\min_{\|\boldsymbol{x}\|_0 \leq k} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2$

Descent algorithms : Orthogonal Matching Pursuit, Orthogonal Least Squares, *etc.*

$S \leftarrow \emptyset$

Repeat

- $S \leftarrow S \cup \{\ell\}$ [Selection]
- $\boldsymbol{x}_S \leftarrow \arg\min_{\boldsymbol{z}} \|\boldsymbol{y} - \boldsymbol{A}_S\boldsymbol{z}\|_2^2$ [Orthogonal projection]

OMP *vs* OLS

Repeat

- $S \leftarrow S \cup \{\ell\}$
- $\boldsymbol{x}_S \leftarrow \arg\min_{\boldsymbol{z}} \|\boldsymbol{y} - \boldsymbol{A}_S \boldsymbol{z}\|_2^2$

Selection rules

OMP : $\ell \in \arg\max_{i \notin S} |\langle \boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}, \boldsymbol{a}_i \rangle|$

OLS : $\ell \in \arg\min_{i \notin S} \left\{ \min_{\boldsymbol{z}} \|\boldsymbol{y} - \boldsymbol{A}_{S \cup \{i\}} \boldsymbol{z}\|_2^2 \right\}$

Bidirectional algorithms

- $\min_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 \quad \text{s.t.} \quad \|\boldsymbol{x}\|_0 \leq k$ $\qquad [\boldsymbol{C}]$

  Descent algorithms $\Longrightarrow$ forward schemes, $S \leftarrow S \cup \{\ell\}$

- $\min_{\boldsymbol{x}} \left\{ \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \lambda \|\boldsymbol{x}\|_0 \right\}$ $\qquad [\boldsymbol{P}]$

  Descent algorithms $\Longrightarrow$ bidirectional schemes, $S \leftarrow S \pm \{\ell\}$

[S. *et al.*, *IEEE TSP*, 2011]

Historic perspective

1) Restoration of Bernoulli-Gaussian signals

- $\boxed{x_i = q_i t_i}$
  - $\boldsymbol{q} \sim \mathcal{B}(\rho)$: support
  - $\boldsymbol{t} \,|\, \boldsymbol{q} \sim \mathcal{N}(\mathbf{0}; \sigma_x^2 \boldsymbol{I})$: amplitudes

SMLR algorithm [Kormylo & Mendel, 1982] :

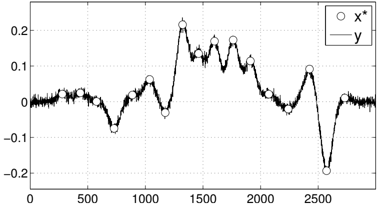
$$\min_{\boldsymbol{x}} \left\{ \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \lambda_{\rho,\sigma_b} \|\boldsymbol{x}\|_0 + \frac{\sigma_b^2}{\sigma_x^2} \|\boldsymbol{x}\|_2^2 \right\}$$

2) *Stepwise regression* [Efroymson, 1960], [Miller, 2002]

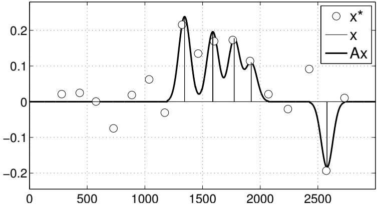
- Statistical inference, undercomplete dictionaries
- bidirectional extensions of forward selection (OLS)

Sparse signal restoration

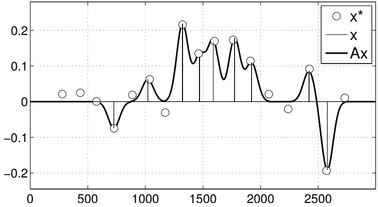
$$\boldsymbol{y} = \boldsymbol{h} * \boldsymbol{x}^\star + \boldsymbol{b}$$



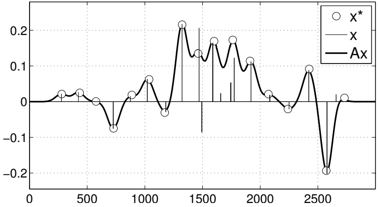
Data : $\|\boldsymbol{x}^\star\|_0 = 17$, RSB = 20 dB



$\lambda = 1$



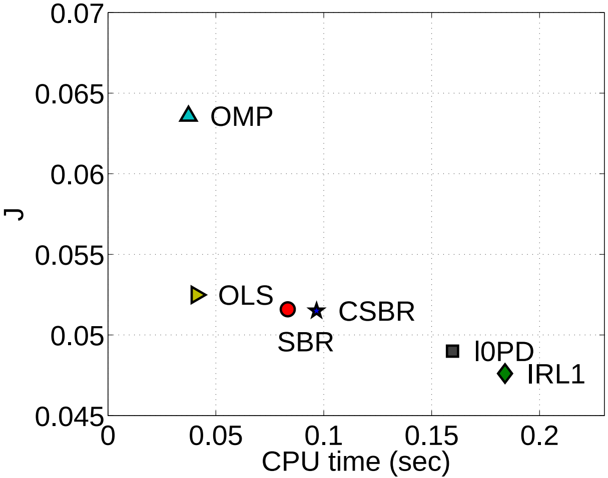
$\lambda = 0.15$



$\lambda = 0.01$

Tradeoff accuracy / CPU time

(noisy deconvolution, $SNR = 25, k = 10, m = 900, n = 756, \sigma = 24$)



Summary

1. Inverse problems and sparse regularization

2. Greedy algorithms for ℓ_0 minimisation

3. ℓ_0 continuation

4. Non-negative greedy algorithms

5. Exact recovery analyses

6. Conclusion

ℓ_0 continuation

- Regularisation path : $\lambda \mapsto \arg\min_{\boldsymbol{x}} \left\{ \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \lambda \|\boldsymbol{x}\|_0 \right\}$

- Reconstruct sparse solutions for a continuum of λ's

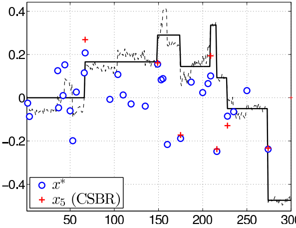
- Inspired from modified LARS for ℓ_1 problems :
  - $\forall \mu,\ \arg\min_{\boldsymbol{x}} \left\{ \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \mu \|\boldsymbol{x}\|_1 \right\}$
  - $\forall t,\ \arg\min_{\boldsymbol{x}} \{ \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 \}$ s.t. $\|\boldsymbol{x}\|_1 \leq t$

[Efron *et al.*, 2004 ; Donoho & Tsaig, 2008]

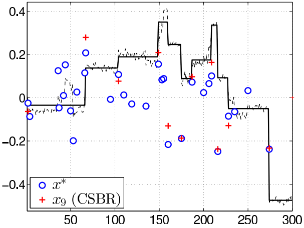
- Model order selection using *e.g., Minimum Description Length*

[S. *et al.*, *IEEE TSP*, 2015]

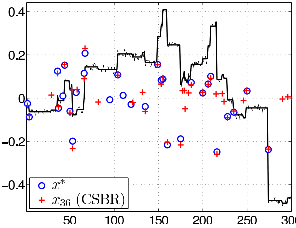
Example : CSBR, problem $k = 30$, SNR = 25 dB



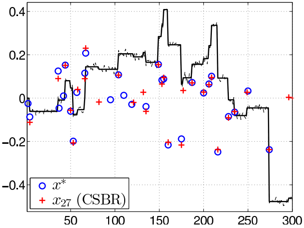
$|S_5| = 7$, $\lambda_5 = 7.8e^{-2}$



$|S_9| = 11$, $\lambda_9 = 4.1e^{-2}$



$|S_{36}| = 41$, $\lambda_{36} = 3.4e^{-4}$



MDLc solution : $|S_{27}| = 29$

Summary

1. Inverse problems and sparse regularization

2. Greedy algorithms for ℓ_0 minimisation

3. ℓ_0 continuation

4. Non-negative greedy algorithms

5. Exact recovery analyses

6. Conclusion

Reconstruction of non-negative sparse signals

- *Inverse problems* $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{n}$

  $\boldsymbol{x}$: sparse and non-negative

- *Goal* : reconstruct x with an accurate and fast algorithm

- *Applications* : Optical spectroscopy, hyperspectral unmixing, tomographic PIV

Orthogonal greedy algorithms

Problem $\min_{\|\boldsymbol{x}\|_0 \leq k} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2$

Algorithm

$S \leftarrow \emptyset$.

Repeat

- $S \leftarrow S \cup \{\ell\}$ [Selection]
- $\hat{\boldsymbol{x}} \leftarrow \arg\min_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 \text{ s.t. } \operatorname{supp}(\boldsymbol{x}) \subset S$ [Unconstrained LS]

OMP : $\ell \in \arg\max_{i \notin S} |\langle \boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{x}}, \boldsymbol{a}_i \rangle|$

Greedy algorithms under NN constraints

Repeat

- $S \leftarrow S \cup \{\ell\}$ [Selection]
- $\hat{\boldsymbol{x}} \leftarrow \arg\min_{\boldsymbol{x} \geq \boldsymbol{0}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 \text{ s.t. } \mathrm{supp}(\boldsymbol{x}) \subset S$ [Projection : NNLS]

Greedy algorithms under NN constraints

Repeat :

- $S \leftarrow S \cup \{\ell\}$ [Selection]
- $\hat{\boldsymbol{x}} \leftarrow \arg\min_{\boldsymbol{x} \geq \mathbf{0}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 \text{ s.t. } \operatorname{supp}(\boldsymbol{x}) \subset S$ [NNLS]
- $S \leftarrow \operatorname{supp}(\hat{\boldsymbol{x}})$ [Compression]

NNOMP : $\ell \in \arg\max_{i \notin S} \langle \boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{x}}, \boldsymbol{a}_i \rangle$

Structural properties induced by support compression

- $S = \text{supp}\,(\hat{\boldsymbol{x}})$
- The current iterate is an ULS solution :

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 \;\text{ s.t. }\; \text{supp}(\boldsymbol{x}) \subset S$$

$$\Longrightarrow \quad \forall i \in S,\ (\boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{x}}) \perp \boldsymbol{a}_i$$

Fast and exact implementation

1. Active-set algorithm (NNLS) [Lawson & Hanson, 74]

   $\Rightarrow$ 100 % greedy implementation

   $\Rightarrow$ Recursive scheme

2. Active-set algorithm with *warm start*

3. Unified fast implementation of several NN extensions :
   NNOMP, NNOLS, SNNOLS

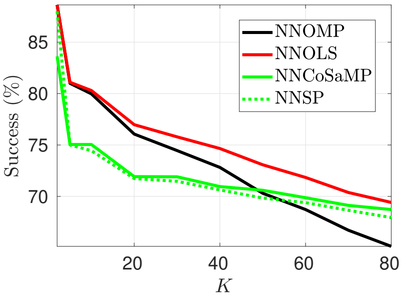
[Nguyen *et al.*, *IEEE TSP*, 2019]

Computational complexity

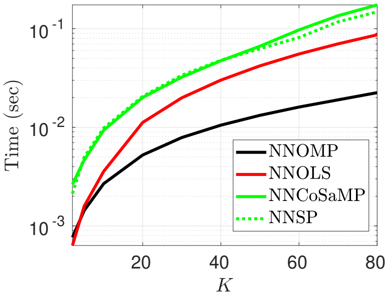
K	Overcost	
	NNOMP / OMP	NNOLS / OLS
60	1.1	2.1
80	1.1	4.1

(dictionary 1200×1140)

Numercial comparisons



True positives



CPU time

Summary

1. Inverse problems and sparse regularization

2. Greedy algorithms for ℓ_0 minimisation

3. ℓ_0 continuation

4. Non-negative greedy algorithms

5. Exact recovery analyses

6. Conclusion

Exact recovery analysis

- *Direct model* : $y = Ax$

- *Exact identification of the support of* x

  - Worst case analyses

  - Probabilistic analyses [Foucart & Rauhut, 2013] :
    - for well-conditioned random dictionnaries
    - for convex minimisation

  - Continuous dictionaries (convex formulation) [Duval & Peyré, 2015]

k-step analysis of greedy algorithms

- $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ where $\boldsymbol{x}$ is k-sparse
- Worst-case ($\forall \boldsymbol{A}, \boldsymbol{x}$) necessary and sufficient condition of reconstruction of $S = \mathrm{supp}(\boldsymbol{x})$:

$$\mu(\boldsymbol{A}) := \max_{i \neq j} |\langle \boldsymbol{a}_i, \boldsymbol{a}_j \rangle| \quad < \quad \frac{1}{2k-1}$$

[Tropp, 2004]

- Guaranteed reconstruction with OMP [Tropp, 2004] and OLS [S. *et al*, 2013]

k-step analysis of greedy algorithms for a given $\boldsymbol{A}$

- Worst-case ($\forall \boldsymbol{x}$) exact recovery of $S = \mathrm{supp}\,(\boldsymbol{x})$ from $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$:

$$\max_{j \notin S} \|\boldsymbol{A}_S^{\dagger} \boldsymbol{a}_j\|_1 < 1 \qquad \text{(ERC) [Tropp, 2004]}$$

- Analysis of OMP [Tropp, 2004] and OLS [S. *et al*, IEEE-IT, 2013]

Analysis of non-negative greedy algorithms

- Sign preservation by OMP/OLS if $\mu < \frac{1}{2k-1}$

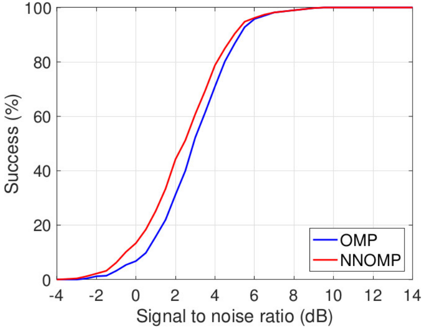
  For $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$, at each iteration, $\mathrm{sgn}(\hat{\boldsymbol{x}}) = \mathrm{sgn}(\boldsymbol{x})$

- *Case* $\boldsymbol{x} \geq \mathbf{0}$

  $\mu < \frac{1}{2k-1} \implies$ *k*-step exact recovery using $\begin{cases} \textit{NNOMP} \\ \textit{NNOLS} \end{cases}$

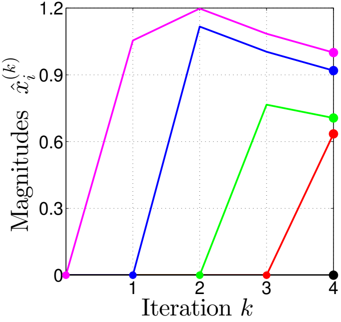
[Nguyen *et al.*, 2019]

Numerical evaluations



Exact reconstruction rate

for $\mu(\boldsymbol{A}) = \dfrac{7}{k-1}$



Non-monotony

of coefficients (OMP, OLS)

Summary

1. Inverse problems and sparse regularization

2. Greedy algorithms for ℓ_0 minimisation

3. ℓ_0 continuation

4. Non-negative greedy algorithms

5. Exact recovery analyses

6. Conclusion

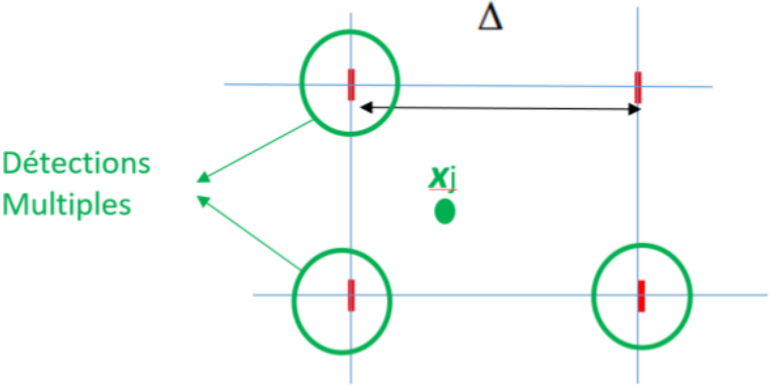
Conclusions and perspectives

- Unified vision of greedy algorithms
- Fast implementations, Matlab codes :

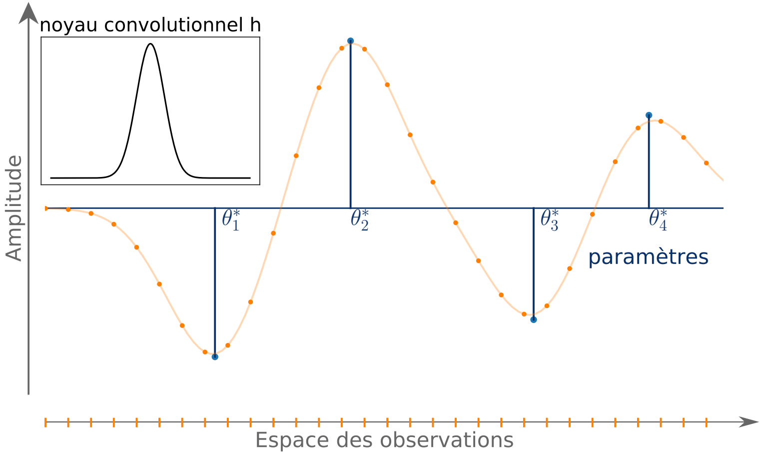
  `http://webpages.lss.supelec.fr/charles.soussen`

- Exact recovery analyses
- High-resolution : fine grids *vs* off-the-grid approaches
  - Convex relaxation (TV-norm minimization) and support identification [de Castro & Gamboa, 2012]
  - Frank-Wolfe based algorithm [Denoyelles *et al.*, 2019]
  - Contious form of OMP [Elvira *et al.*, 2019]

Discrete *vs* continuous dictionaries



Off-the-grid sparse deconvolution



« Dictionary » composed of a continuum of atoms $\mathcal{A} = \{h(t - \theta),\ \theta \in \mathbb{R}\}$

Ongoing work

Analysis of a continuous version of OMP

$$\mathcal{A} = \{\boldsymbol{a}(\theta),\, \theta \in \mathbb{R}\} \quad \text{tel que} \quad \langle \boldsymbol{a}(\theta_1), \boldsymbol{a}(\theta_2) \rangle = K(\|\theta_1 - \theta_2\|)$$

Mutual incoherence based analysis

- Inspired by [Tropp 2004]
- Redefine the concept of exact recovery

[Elvira *et al.*, 2019]

Thank you
for
your attention !