Spikes super-resolution with random Fourier sampling

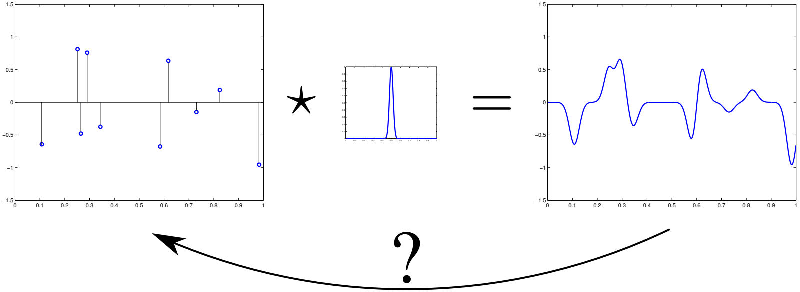
SPARS, 2017

Yann Traonmilin, **Nicolas Keriven**, Rémi Gribonval, Gilles Blanchard

INRIA Rennes

University of Potsdam

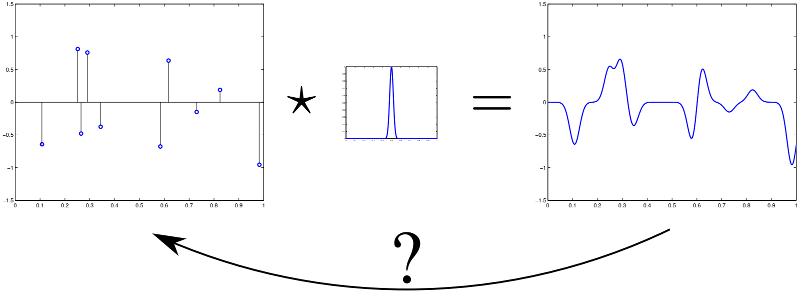
Spikes super-resolution (deconvolution)



State of the art : Guarantees for practical estimation methods (Shannon-type condition).

[Candès, De Castro, Duval, etc...]

Spikes super-resolution (deconvolution)

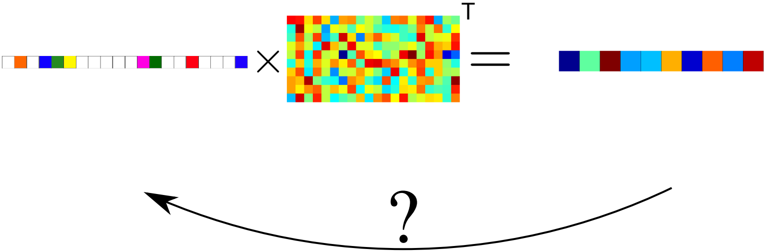


State of the art : Guarantees for practical estimation methods (Shannon-type condition).

[Candès, De Castro, Duval, etc...]

Equally spaced frequencies in Fourier domain.

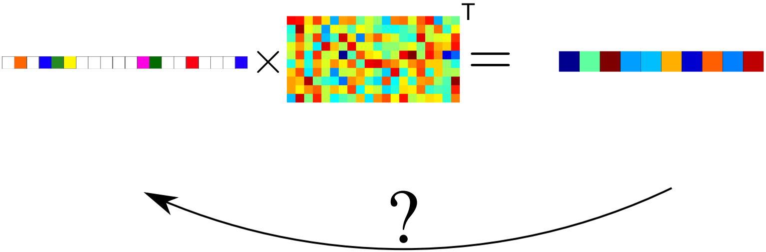
Compressive Sensing, sparsity



State of the art : Guarantees for practical dimensionality-reduction schemes and practical estimation methods (in Hilbert space)

[Candès, Donoho, Gribonval, Puy, Dirksen, Traonmilin, etc...]

Compressive Sensing, sparsity



State of the art : Guarantees for practical dimensionality-reduction schemes and practical estimation methods (in Hilbert space)

[Candès, Donoho, Gribonval, Puy, Dirksen, Traonmilin, etc...]

Random design of measurement matrix.

Previous work on learning

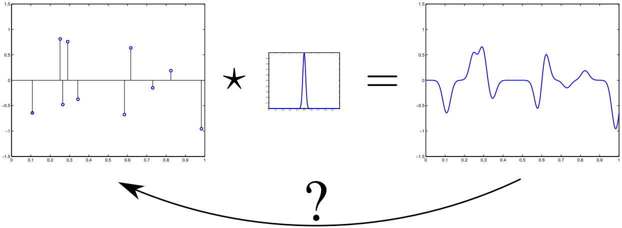
Compressive k-means [Keriven et al., ICASSP 2017]

Random Moments for... [Keriven et al., SPARS 2017] :

- We can do k-means from the sketch of a database ...
- ... by **recovering linear combination of Diracs from random linear measurements**
- The Compressive Learning-OMP (CL-OMP) heuristic (Keriven et al., SPARS 2015, ICASSP 2016) performs well in practice
  - OMP + non-convex updates

Goal

Consequence for super-resolution?



Layout

Introduction

Sparse recovery in infinite-dimensional spaces

Experiments

Conclusion

Inverse problem

- Measurements

$$y = Ax_0 + e$$

- Finite dimension: classical Signal Processing.
  $A =$ convolution, sub-sampling, etc....
- Infinite dimension (Hilbert) : "generalized" sampling (Adcock and Hansen, Traonmilin and Gribonval)
- **Infinite dimension (Banach) :** spikes super-resolution, $A =$ "low-pass" filter

Dimension reduction and low-complexity

- A is dimension reducing : regularity comes from "low-complexity" models Σ

Dimension reduction and low-complexity

- A is dimension reducing : regularity comes from "low-complexity" models Σ
- Sparsity : $\Sigma = \Sigma_k$ = set of k-sparse vectors

Dimension reduction and low-complexity

- A is dimension reducing : regularity comes from "low-complexity" models Σ
- Sparsity : $\Sigma = \Sigma_k =$ set of k-sparse vectors
- Super-resolution: $\Sigma = \Sigma_{k,\epsilon} =$ set of sums of k Diracs with supports separated by ϵ (in a bounded domain)

$$\Sigma_{k,\epsilon} = \left\{ \sum_{i=1,k} a_i \delta_{t_i} : \forall r \neq l, \|t_r - t_l\|_2 \geq \epsilon, \|t_l\|_\infty \leq 1, \mathbf{a} \in \mathcal{C}_a \right\}$$

Measurement methods

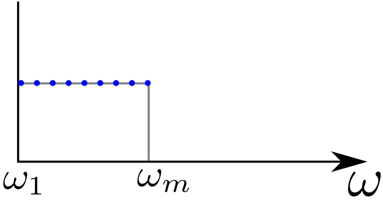
$$Ax_0 = \left(\int_t x_0(t) f_i(t) dt \right)_{i=1,m}$$

où $f_i(t) = e^{j\langle \omega_i, t \rangle}$, $(\omega_i)_{i=1,m} \subset \mathbb{R}^d$.

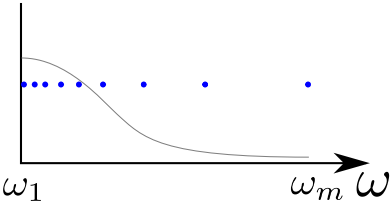
Measurement methods

$$Ax_0 = \left(\int_t x_0(t) f_i(t) dt \right)_{i=1,m}$$

où $f_i(t) = e^{j\langle \omega_i, t\rangle}$, $(\omega_i)_{i=1,m} \subset \mathbb{R}^d$.



Uniform (A_U)



Random (A_R)

Measurement methods (1)

$$A x_0 = \left(\int_t x_0(t) f_i(t) dt \right)_{i=1,m}$$

où $f_i(t) = e^{j\langle \omega_i, t \rangle}$, $(\omega_i)_{i=1,m} \subset \mathbb{R}^d$.

- A_U**: Uniform Fourier sampling (low pass filter)**: frequencies $(\omega_i)_{i=1,m}$ taken uniformly in $[-\frac{\pi q}{2}, \frac{\pi q}{2}]^d$ where q is an integer and $m = (2q+1)^d$.
- Estimation of x_0 possible if $m \geq \frac{2}{\epsilon}$ (Work of Candès, De Castro, Duval ... !!! Results are usually given on the torus !!!)

Measurement methods (2)

$$Ax_0 = \left(\int_t x_0(t) f_i(t) dt \right)_{i=1,m}$$

où $f_i(t) = e^{j\langle \omega_i, t\rangle} / c_{\omega_i}$, $(\omega_i)_{i=1,m} \subset \mathbb{R}^d$.

- A_R**: Random (weighted) Fourier sampling**: ω_i drawn at random from $\Lambda \propto c_\omega^2 e^{-\sigma^2 \|\omega\|_2^2/2}$ (with scale parameter σ).
  - use of "smoothing" weights c_ω
- CL-OMP heuristic for estimating x_0 (Keriven et al. 2016,2017)

Ideal decoder

- With A_R, the "ideal" decoder is :

$$x^* \in \arg\min_{x \in \Sigma} \|Ax - y\|_2$$

- Information preservation guarantees?

$$\|x^* - x_0\| \lesssim \|e\|_2 + d(x_0, \Sigma)$$

Information Preservation Guarantee

Theorem (Blanchard, Gribonval, Keriven, Traonmilin) : Assume

$$m \geq \; O(k^2 d^2 (\text{polylog}(k, d) + \log(1/\epsilon))).$$

Then with high probability on A_R, for all x_0 and $y = A_R x_0 + e$, we have

$$\|x^* - x_0\|_h \lesssim \|e\| + d_h(x_0, \Sigma)$$

where $d_h(x_0, \Sigma_{k,\epsilon}) = \inf_{x \in \Sigma_{k,\epsilon}} \|x_0 - x\|_h$ is the modelisation error ($= 0$ if x_0 is exactly a sum of Diracs).

Restricted Isometry Property

For $x \in \Sigma - \Sigma$:

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

- Sufficient condition on A to guarantee success of the ideal decoder (and convex relaxation in classical compressive sensing)
- Sub-gaussian matrices have this for many Σ (Puy et al. 2015).
- **RIP in super-resolution framework?**

Kernel, Hilbert space

- In the Banach space of finite-signed measures, the low-pass filter (A_U) does not satisfies the RIP for the natural metric $\|\cdot\| = \|\cdot\|_{TV}$ (total variation of measures)

Kernel, Hilbert space

- In the Banach space of finite-signed measures, the low-pass filter (A_U) does not satisfies the RIP for the natural metric $\|\cdot\| = \|\cdot\|_{TV}$ (total variation of measures)

- One can build kernel norm to get a Hilbert structure. In our case it is actually linked to the chosen resolution :

$$\|\cdot\| := \|\cdot\|_h = \|h \star \cdot\|_2 \tag{1}$$

  where $h(t) = e^{-\frac{\|t\|_2^2}{2\sigma^2}}$ (gaussian kernel, σ scale parameter used for defining A_R).

- This metric can be seen as a **distance at some resolution** in the space of finite signed measures.

Does A_R satisfy the RIP on $\Sigma_{k,\epsilon}$?

Classical **two-steps** proof of the RIP

- Pointwise concentration : for $x_1, x_2 \in \Sigma_{k,\epsilon}$,

$$\|A(x_1 - x_2)\| \approx \|x_1 - x_2\|_h \qquad (2)$$

(Bernstein concentration inequality)

Does A_R satisfy the RIP on $\Sigma_{k,\epsilon}$?

Classical **two-steps** proof of the RIP

- Pointwise concentration : for $x_1, x_2 \in \Sigma_{k,\epsilon}$,

$$\|A(x_1 - x_2)\| \approx \|x_1 - x_2\|_h \tag{2}$$

  (Bernstein concentration inequality)

- The normalized secant set $\mathcal{S} = \{\frac{u}{\|u\|_h} : u \in \Sigma_{k,\epsilon} - \Sigma_{k,\epsilon}\}$ has finite covering numbers (finite "upper box counting" dimension):

$$N(\mathcal{S}, \alpha) \leq \left(\frac{C}{\alpha}\right)^{-dim(\mathcal{S})} \tag{3}$$

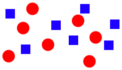
Key principle

The result comes from the ϵ-separation condition and the definition of the kernel.

Let $u \in \Sigma - \Sigma$.

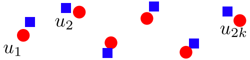
Without separation

$$u = x_1 - x_2$$



With separation

$$u = x_1 - x_2 = u_1 + u_2 + \ldots + u_{2k}$$



Pythagore-like bound :

$$1 - \beta \leq \frac{\left\| \sum_{l=1}^{2k} u_l \right\|_h^2}{\sum_{l=1}^{\ell} \left\| u_l \right\|_h^2} \leq 1 + \beta.$$

Discussion

Measurement scheme	Uniform frequencies	Random frequencies
Number of meas. m	$O(1/\epsilon)$	$O(k^2 d^2 \text{polylog}(k,d) \log(1/\epsilon))$

- Dependency in ϵ improved
- Close to case *with grid*
  - grid size $\mathcal{O}(1/\epsilon^d)$, sparse recovery: sparsity times log of grid size $\mathcal{O}(kd \log(1/\epsilon))$
- !!! Technically speaking, Gaussian random frequencies are not bounded !!!

Layout

Introduction

Sparse recovery in infinite-dimensional spaces

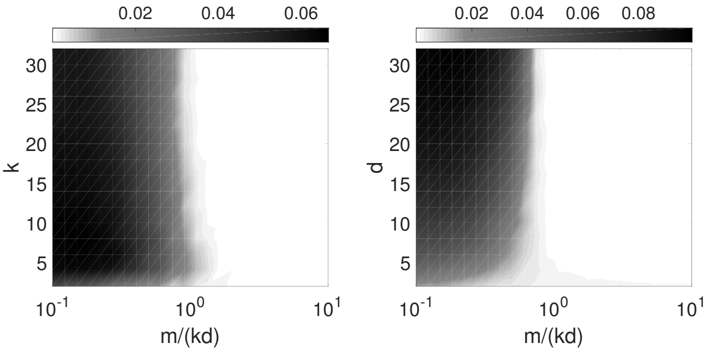
Experiments

Conclusion

In practice

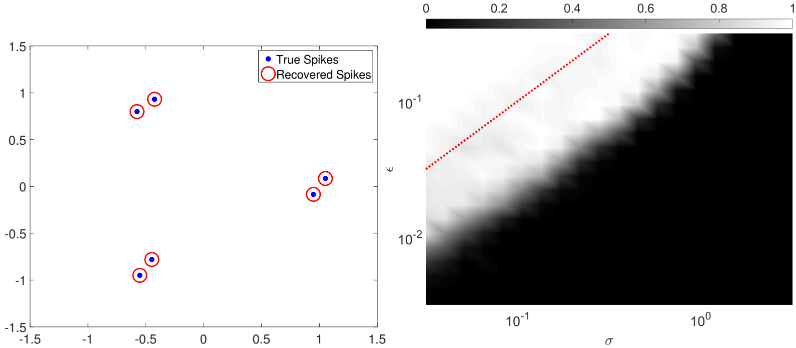
- Ideal decoder non-convex.
  - convex relaxation sometimes possible with TV norm [Candès, De Castro, Duval...], difficult in high dimension
- Heuristic: Compressive Learning-OMP (CL-OMP)
  - Greedy approach + non-convex gradient descent updates
  - `sketchml.gforge.inria.fr`

Number of measurements



Phase transition: $m \approx \mathcal{O}(kd)$ *seems sufficient (left* $d = 10$*, right* $k = 10$*).*

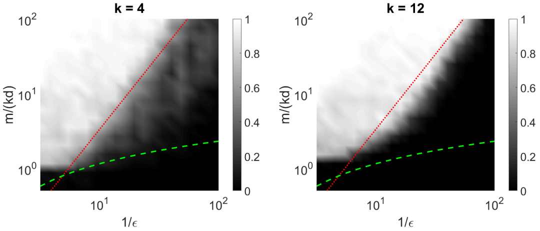
Choice of A_R (frequency distribution)



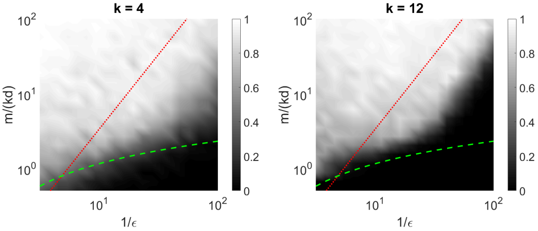
ϵ separation w.r.t. scale parameter σ

Toward compressive super-resolution?

Uniform Fourier



Random Fourier



Layout

Introduction

Sparse recovery in infinite-dimensional spaces

Experiments

Conclusion

Conclusion

- What we have done:
    - RIP in the space of finite signed measures
    - Information preservation guarantees
    - Encouraging practical results
- Outlooks
    - Practical random acquisition?
    - Extend comparison with existing results (what about kernel norms?)
    - Algorithms with guarantees : convex relaxation in any dimension? basin of attraction with the RIP?

Thank you !

yanntraonmilin.wordpress.com

people.irisa.fr/Nicolas.Keriven
sketchml.gforge.inria.fr

!!!Preprint online very soon!!!

Questions?

Introduction

Sparse recovery in infinite-dimensional spaces

Experiments

Conclusion