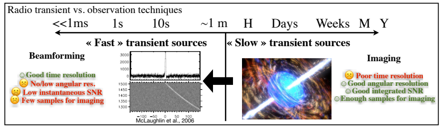
Sparse spatio-temporal imaging of radio transients

J. N. Girard, M. Jiang, J.-L. Starck and S. Corbel
AIM/IRFU/SAp/CEA-Saclay, Université Paris Diderot, Orme des Merisiers, 91191 Gif-Sur-Yvette, France
Email: julien.girard@cea.fr

Context

- **Huge continental radio facilities**: LOFAR, MeerKAT/SKA, ASKAP
- The need for **robust transient detection** tools at all temporal scales
- The recent framework of **Compressed Sensing / Sparse representation / Convex optimization**

sensitivity +++
angular resolution +++
data volume ++++++

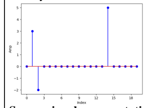


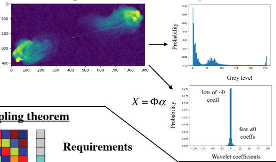
We are proposing a novel approach to address the deconvolution of interferometric images accounting for the spatial AND temporal dependency of the data.

Sparse representation ? Compressed Sensing?

Theoretical « sparse » signal very few non-zero entries

Natural signal ? A natural signal may not be sparse itself, but can be « sparsified » in a dictionary Φ



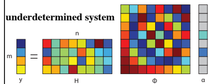


$X = \Phi\alpha$

Sparse signal representation

Compressed Sensing as a sampling theorem

underdetermined system



$y = \mathrm{H}x = \mathrm{H}\Phi\alpha = \mathrm{A}\alpha$

Requirements
- **Sparsity**
- **Incoherence** between the sensing basis H and the dictionary basis Φ

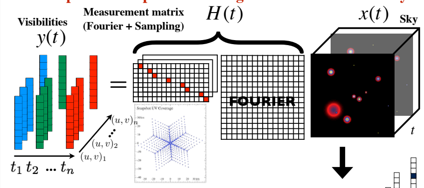
Sparse recovery

$$\min_{\alpha} \|\alpha\|_1 \quad \text{s.t.} \quad \|y - \mathrm{A}\alpha\| < \varepsilon$$

"Signals with exactly K components different from zero can be recovered perfectly from ~ K log N incoherent measurements"

(Candès et al. 2006, Donoho 2006) **Sampling with an interferometer ~ Compressed Sensing**

Simple 2D-1D sparse modeling for radio interferometry



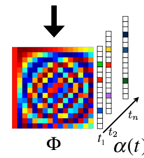
Both the sky and masking operator are time dependent
The time varying sky is supposed to be sparse in Φ

$$y = \mathbf{H}x + \mathbf{N} = \mathbf{MF}x + \mathbf{N}$$

ill-posed inverse problem

$x = \Phi\alpha$

e.g. Wavelet Tr. Sparse



Dictionaries for 2D-1D sparse signal representation

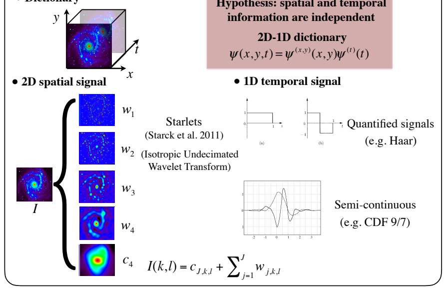
- **Dictionary**

Hypothesis: spatial and temporal information are independent

2D-1D dictionary

$$\psi(x, y, t) = \psi^{(x,y)}(x, y)\psi^{(t)}(t)$$

- **2D spatial signal**
- **1D temporal signal**



$$I(k, l) = c_{J,k,l} + \sum_{j=1}^{J} w_{j,k,l}$$

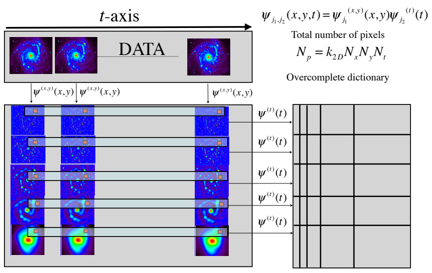
2D-1D wavelet decomposition

$$\psi_{j_1, j_2}(x, y, t) = \psi_{j_1}^{(x,y)}(x, y)\psi_{j_2}^{(t)}(t)$$

Total number of pixels

$$N_p = k_{2D} N_x N_y N_t$$

Overcomplete dictionary



Proximal algorithms to the rescue

If the problem has the form $\arg\min_{x} f(x) + g(x)$

Not solvable with gradient methods !

$f(x)$ convex differentiable function

e.g Data fidelity $\frac{1}{2}\| y - x \|_2^2$

$g(x)$ convex but not differentiable function

e.g Sparsity constraint e.g $\|\bullet\|_1$

Hopefully, we now have tools to address this problem:

- **Proximal calculus**

At a point x, the proximity operator is defined by

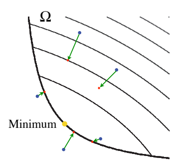
$$\mathrm{prox}_{\lambda g}(x) = \arg\min_{y \in \mathcal{H}} g(x) + \frac{1}{2\lambda}\|y - x\|_2^2$$

The prox operator is a « generalized » projection operator

when $g(x)$ is differentiable

$$\mathrm{prox}_{\lambda g}(x) = x - \lambda \nabla g(x)$$

a gradient step of the function g



x is in domain Ω update step = gradient step towards the minimum

x out of domain Ω update step = orthogonal projection on the boundary

- **Forward-Backward** (Combettes and Wajs 2005)

Loop i

$x^{(i+1)} = x^{(i)} - \mu \nabla f(x^{(i)})$ **Forward**

$x^{(i+1)} = \mathrm{prox}_{\mu g}(x^{(i+1)})$ **Backward**

Problems with this method

1) Not explicit proximity operator in analysis Framework $g(\bullet) = \|\Phi^t \bullet\|_1$, $\mathrm{prox}_g = ?$

2) How to include more constraints? (e.g. positivity)

- **Condat-Vu primal-dual algorithm** (Condat 2013; Vu 2013)

- Deal with two constraints (sparsity, positivity)
- Avoid the implicit evaluation of analysis sparsity operator

Application $\min_{x} \frac{1}{2}\| y - \mathbf{H}x \|_2^2 + \| \Lambda \odot \Phi^t x \|_1 + i_{\mathbb{R}_+}(x)$ $\quad \mathbf{H} = \mathbf{MF}$

Initialize $x^{(0)}, u^{(0)}$

Evaluate x: $x^{(i+1)} = \mathrm{Proj}_{\mathbb{R}_+}(x^{(i)} - \tau(\Phi u^{(i)} - \tau \mathbf{H}^t(y - \mathbf{H}x^{(i)}))$

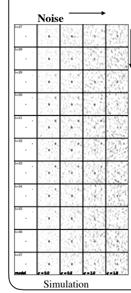
Evaluate u: $u^{(i+1)} = (\mathrm{Id} - \mathrm{ST}_\lambda)(u^{(i)} + \eta \Phi^T (2x^{(i+1)} - x^{(i)}))$

Λ **automatically set for each scale, derived from a noise-driven strategy from the residuals.**

M.Jiang, J.Girard, J-L.Starck, S.Corbel and C.Tasse, "Compressed Sensing and Radio Interferometry", EUSIPCO 2015, Nice

Numerical experiments

Data set simulation



Detection robustness towards
- **noise level** σ — *detection problem*
- **num of frames** N_t — *dilution problem*

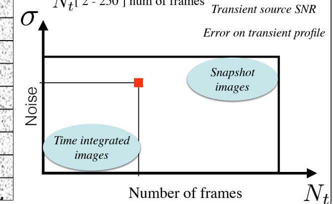
σ [0. - 2.0] arb. units

N_t [2 - 250] num of frames

Metrics

Transient source SNR

Error on transient profile



Test #1: SNR

Dirty cubes

Marginal improvement of SNR with time

SNR turn-over below N_t=40

→ temporal dilution

CLEANed cubes

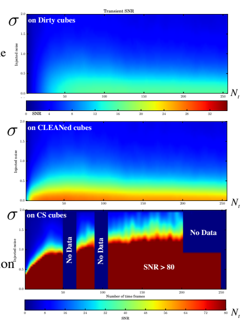
Higher SNR at low noise

Temporal dilution effect reduced

2D-1D Sparse cube

Higher SNR at low & high noise

Slow decrease of SNR due to dilution



One order of magnitude improvement in SNR

Test #2: Temporal profiles

Dirty & CLEAN

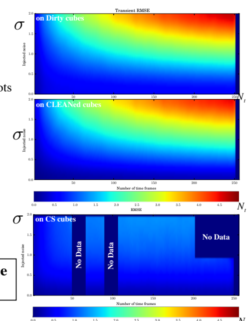
Higher RMSE in high noise snapshots

Similar error

2D-1D Sparse cube

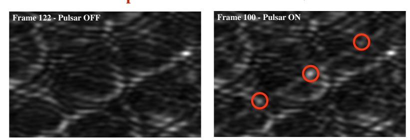
Higher RMSE in high noise snapshots

Better overall profile reconstruction



Factor of ~3 reduction of the RMSE of transient profile

Test #3: Real data: pulsar B0355+55 with the VLA

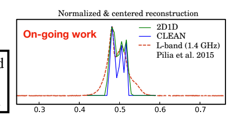


data $\Delta t = 1.03s$

res. $\delta t = 5ms$

Period $P_{\mathrm{B0355+55}} = 156ms$

- **Transient** takes the form of the PSF
- **PSF sidelobes** = false positive !



Peaks are successfully detected
CS 2D-1D less biased than
frame-by-frame reconstruction

Conclusions/Perspectives

- **Novel spatio-temporal sparse method**
  - Based on Condat-Vu primal-dual method
  - 2D-1D dictionary for sparse signal representation
  - Automatic setting of relaxation parameters
- **Radio transients detection**
  - Sensitivity: one order of magnitude improvement in SNR
  - Temporal profile: factor of 3 improvement in the profile reconstruction error
  - Preliminary validation on real data
- **Improvements/Applications**
  - Validation on spatially resolved transient sources (e.g. VLBI radio cores)
  - Code acceleration: Embarrassingly parallelisable => HPC
  - Integrate into radio transients detection pipeline

References

Candes, E., Romberg, J., & Tao, T. 2006, Information Theory, IEEE, 52, 489
Candes, E. J., Wakin, M. B., & Boyd, S. P. 2007, ArXiv:0711.1612
Condat, L. 2013, Journal of Optimization Theory and Applications, 158, 460
Dewdney, et al., 2009, Proceedings of the IEEE, 97, 1482
Donoho, D. 2006, IEEE Transactions on Information Theory, 52, 1289
Fender, R. P. & Bell, M. E. 2011, Bulletin of the ASI, 39, 315
Garsden, H., Girard, J. N., Starck, J. L., et al. 2015, A&A, 575, A90
van Haarlem, M. P., Wise, M. W., Gunst, A. W., & et al. 2013, A&A, 556, A2
Hogbom, J. A. 1974, A&AS, 15, 417
Jiang, M., Girard, J. N., Starck, J. L., Corbel, S. and Tasse, C., 2015, SF2A
M.Jiang, J.Girard, J-L.Starck, S.Corbel and C.Tasse, "Compressed Sensing and Radio Interferometry", EUSIPCO 2015, Nice, IEEE, 1646 - 1650
McLaughlin et al., 2006, Nature 439, 817
Hogbom, J. A. 1974, A&AS, 15, 417
Lorimer, et al..2007, Science, 318, 777
Lorimer, et al. 2013, MNRAS, 436, L5
Petrov, E., Bailes, M., Barr, E. D., et al. 2015, MNRAS, 447, 246
Spitler, L. et al. 2014, ApJ, 790, 101
Starck, J.-L. et al., 2009, A&A, 504, 641
Starck, J.-L. et al., 2011, Astronomical Data Processing (Springer), 1489-1531
Swinbank, J. D., Staley, T. D., Molenaar, G. J., et al. 2015, A&C, 11, 25
Thornton, D., Stappers, B., Bailes, M., et al. 2013, Science, 341, 53
Vu, B. C. 2013, Advances in Computational Mathematics, 38, 667