An introduction to compressive sensing & imaging applications

Invited talk
Observatoire Cote d Azur - Laboratoire Lagrange
Nice, April 28, 2015

Prof. Y. Wiaux
BASP - Biomedical and Astronomical Signal Processing Group
Heriot-Watt University, Edinburgh

PRESENTATION OVERVIEW

I. MOTIVATION

II. COMPRESSIVE SENSING THEORY

III. NEXT-GENERATION COMPRESSIVE IMAGING

Take-home message

I

MOTIVATION

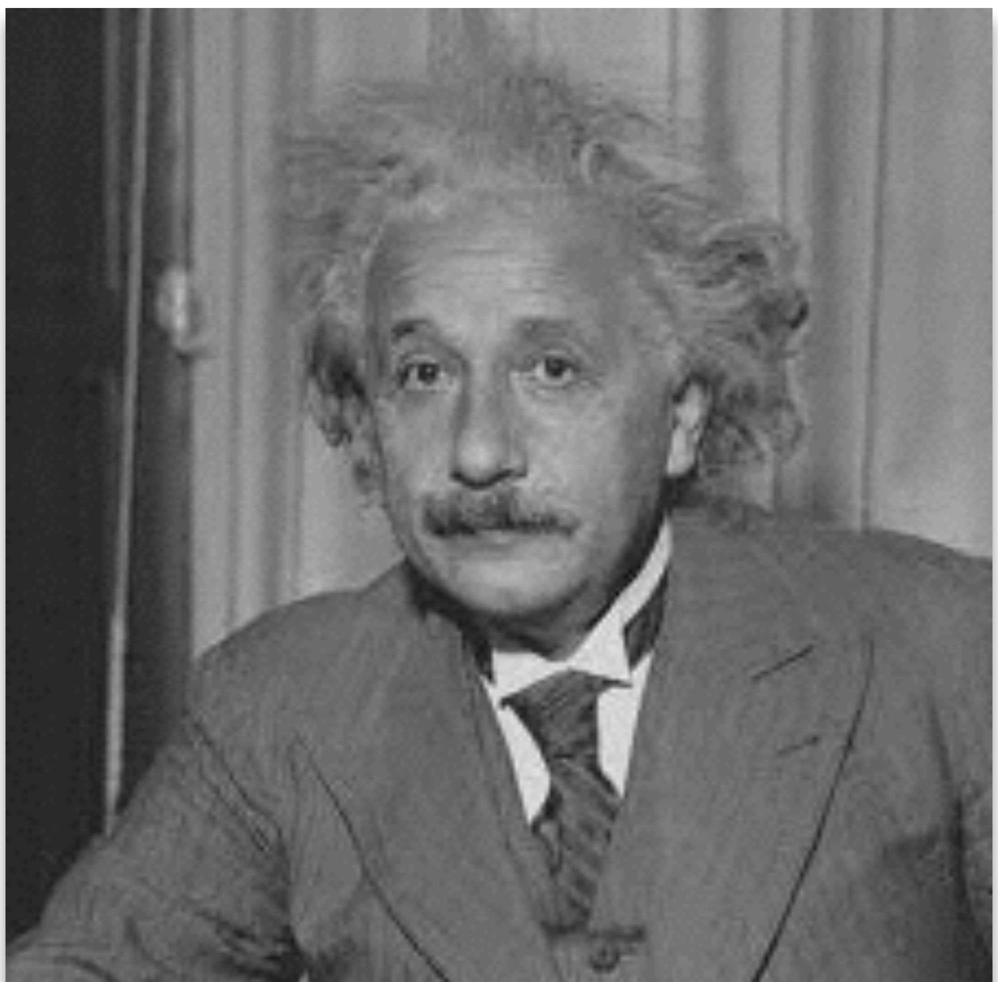
Enlightening demo...

AT THE BEGINNING WAS...

SPARSITY

Sparse signal model

* Wavelet or scale-space techniques provide a multi-scale representation of natural signals and images, generically characterized by a sparse structure!



Picture of Einstein

Through a wavelet transform, a signal is analysed by its **convolution with a wavelet filter at various scales**:

$$\langle \psi_{a,x_0} | f \rangle$$

Sparse signal model

* Wavelet or scale-space techniques provide a multi-scale representation of natural signals and images, generically characterized by a sparse structure!

Sparse multiscale representation

Signals are analysed by their **convolution with a wavelet filter at various scales**:

$$\langle \psi_{a,x_0} | f \rangle$$

This **multi-scale decomposition** provides a sparse representation, with numerous applications:
- denoising
- deconvolution
- compression
- etc.

Standard compression

* Signals are usually sampled at Nyquist rate to avoid aliasing and compressed **adaptively** after acquisition... typically leveraging sparsity!

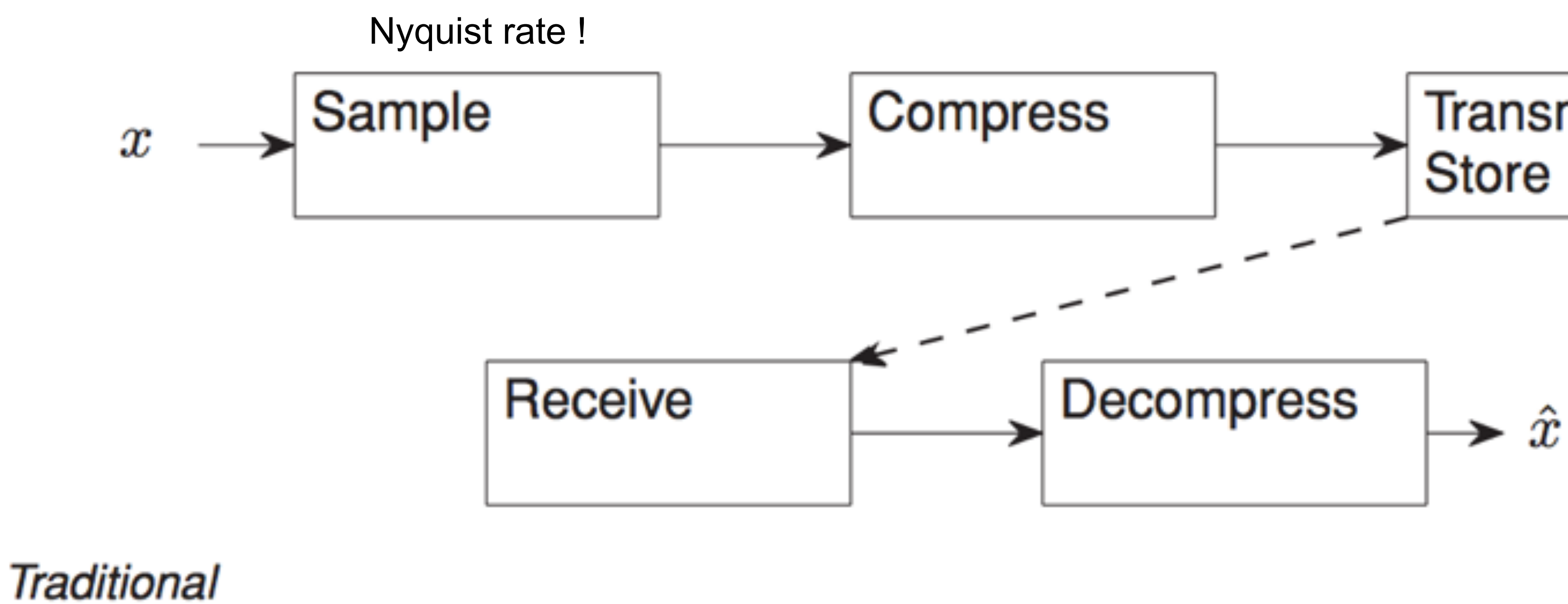
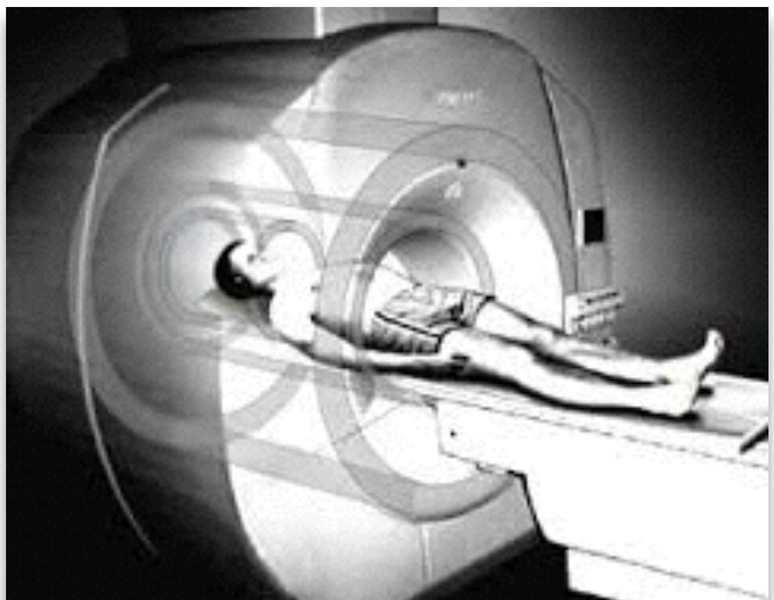


Image Courtesy Han et al., 2013, Compressive Sensing for Wireless Networks

Acquisition, storage, transfer of full data sets can be costly or unaffordable!

MRI Fourier sampling

* MRI finds its superiority in that it is a non-invasive non-ionizing biomedical imaging modality offering multiple contrast mechanisms. It probes anatomy by Fourier sampling.

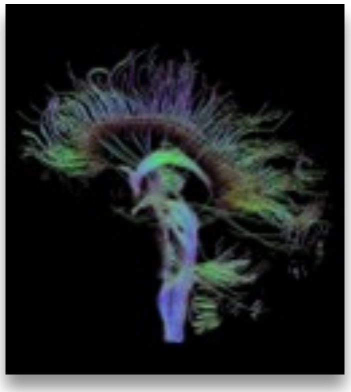


- Each measurement of an MRI sequence is a Fourier coefficient of the image of interest:

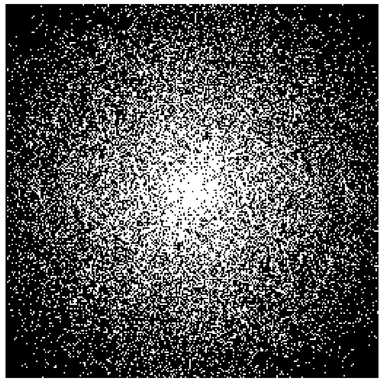
$$y\left(\boldsymbol{k}_i\right) \equiv \int_{\mathrm{D}_{\tau}} x\left(\boldsymbol{\tau}\right) \mathrm{e}^{-2\mathrm{i}\pi \boldsymbol{k}_i \cdot \boldsymbol{\tau}} \, \mathrm{d}\boldsymbol{\tau} \equiv \widehat{x}\left(\boldsymbol{k}_i\right).$$

- Full acquisition is prohibitively slow. Acceleration by incomplete sampling is a deep challenge in the field:

2D Structural MRI



6D Diffusion MRI



Fourier sampling

Interferometric Fourier sampling

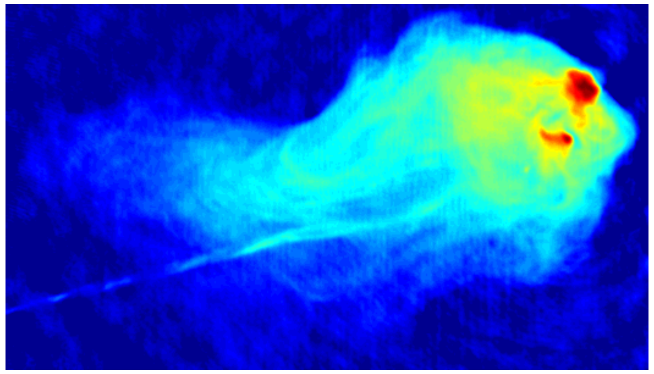
* Aperture synthesis in radio interferometry images the sky at extreme resolutions and sensitivities... through Fourier sampling.

Very Large Array, New Mexico

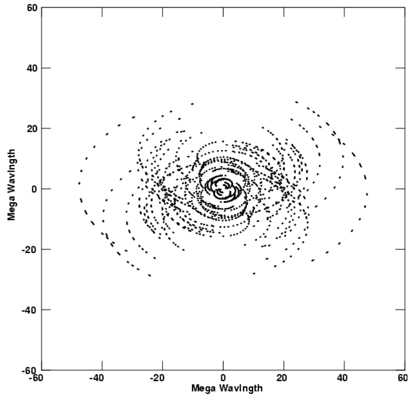
❖ Each Telescope pair probes the correlation of incoming electric fields from the source, called visibility:

$$y\left(\boldsymbol{b}_{\lambda}\right) \equiv \int_{\mathrm{S}^2} Ax\left(\boldsymbol{\tau}\right) \mathrm{e}^{-2\mathrm{i}\pi \boldsymbol{b}_{\lambda}\cdot\boldsymbol{\tau}}\, \mathrm{d}\Omega\left(\boldsymbol{\tau}\right),$$

... leading to incomplete Fourier sampling (monochromatic imaging, small FOV, incoherent source):



Cygnus A



Fourier sampling

Could we simply recover signals from sub-Nyquist sampling?

i.e. can we sample compressively?

?Compressive sensing?

* Compressive sensing would merge acquisition and compression... the compression being performed in sampling, it would have to be **non-adaptive**!

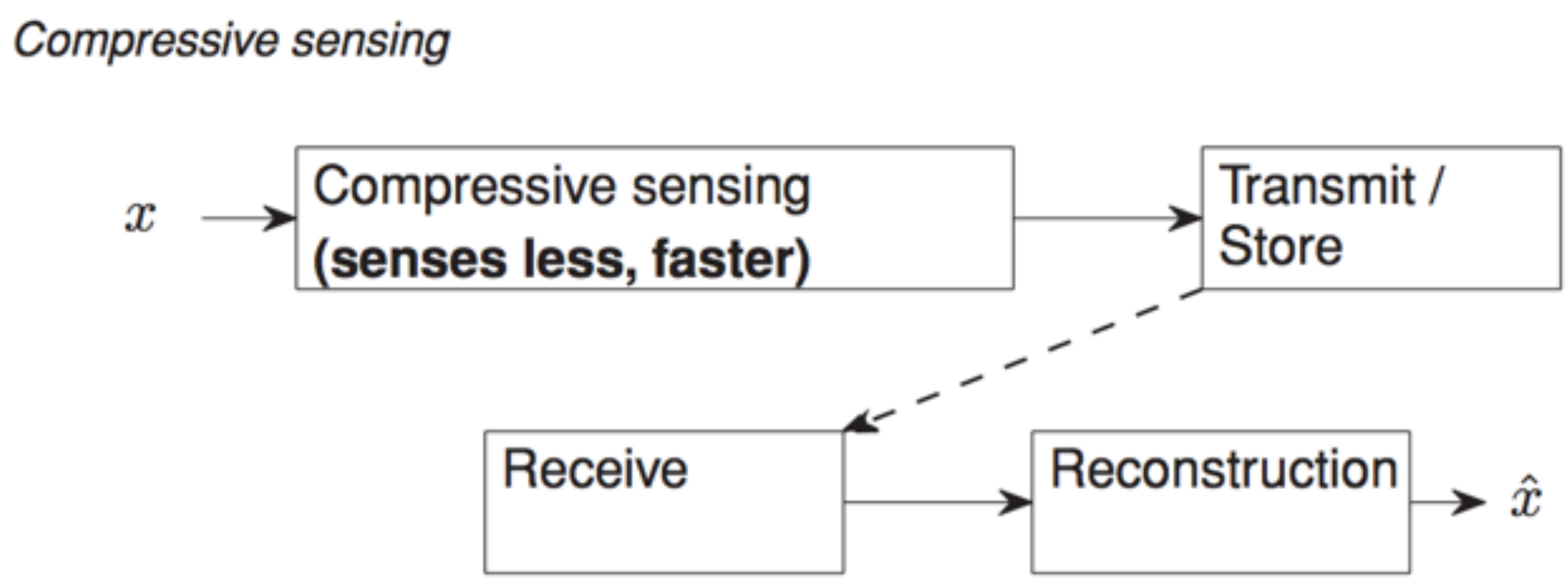


Image Courtesy Han et al., 2013, Compressive Sensing for Wireless Networks

II

COMPRESSIVE SENSING THEORY

E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," IEEE Trans. Inf. Theory, vol. 52, pp. 489-509, 2006.

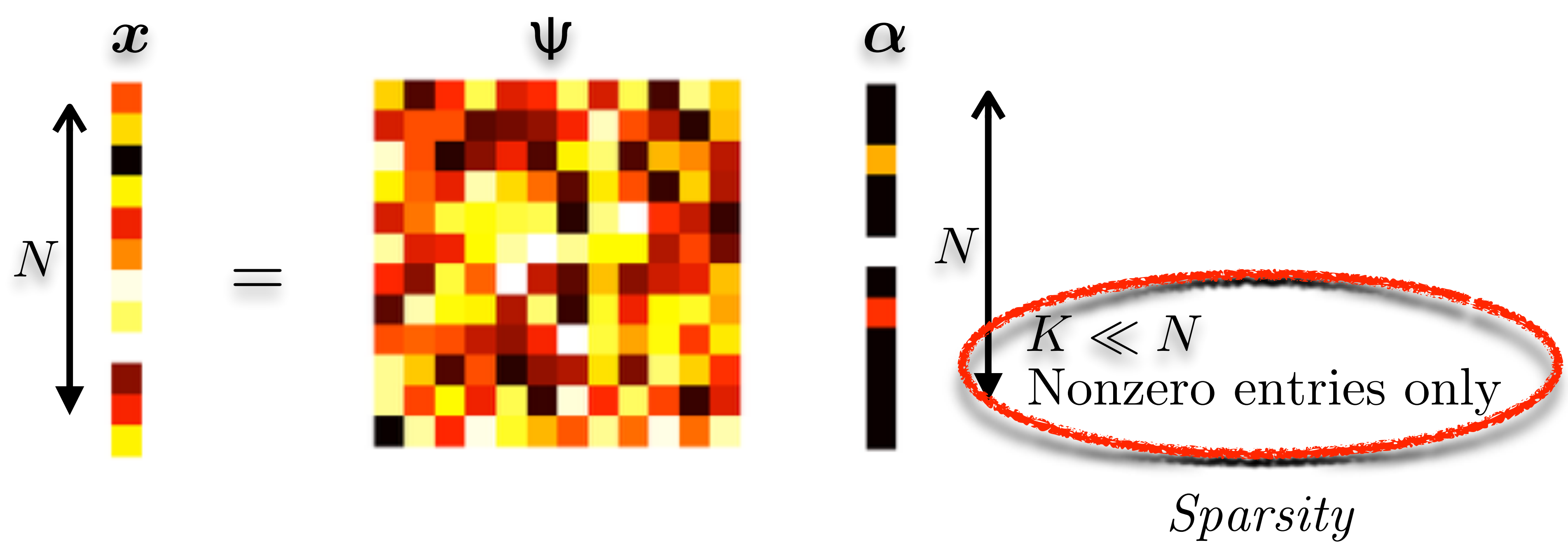
D. L. Donoho, "Compressed sensing," IEEE Trans. Inf. Theory, vol. 52, pp. 1289-1306, 2006.

E. J. Candès, Y. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," Applied Comput. Harm. Analysis vol. 31, pp. 59-73, 2010.

+ thousands of papers...

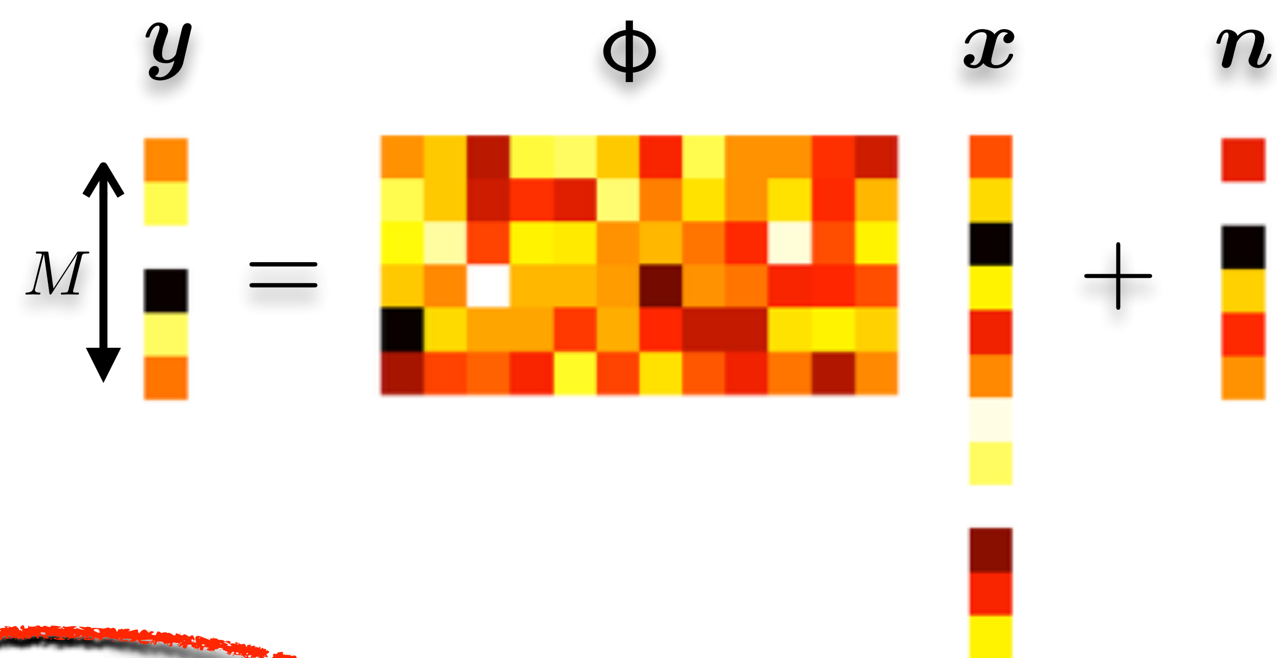
Sparse signal model !

$$\boldsymbol{x} = \Psi \boldsymbol{\alpha} \in \mathbb{R}^{\mathbb{N}}$$



Linear sensing model

$$\boldsymbol{y} = \Phi \boldsymbol{x} + \boldsymbol{n} \in \mathbb{R}^{\mathrm{M}}$$



$$M \ll N$$

Sub-Nyquist sampling

l1-norm minimisation problem

* Regularisation is needed to solve the ill-posed inverse problem. Leveraging sparsity…

❖ l0-norm minimisation promotes sparsity, but is non convex and combinatorial:

$$\min_{\boldsymbol{\alpha}} ||\boldsymbol{\alpha}||_0 \quad \text{subject to} \quad ||\boldsymbol{y} - \boldsymbol{\Phi\Psi\alpha}||_2 \leq \epsilon$$

Two approaches have been pursued: greedy approaches such as Matching Pursuit (MP), or convex relaxations…

❖ The theory of compressive sensing demonstrates that the convex l1-norm minimisation problem will recover sparse signals, from suitable (???) sub-Nyquist data:

$$\min_{\boldsymbol{\alpha}} ||\boldsymbol{\alpha}||_1 \quad \text{subject to} \quad ||\boldsymbol{y} - \boldsymbol{\Phi\Psi\alpha}||_2 \leq \epsilon$$

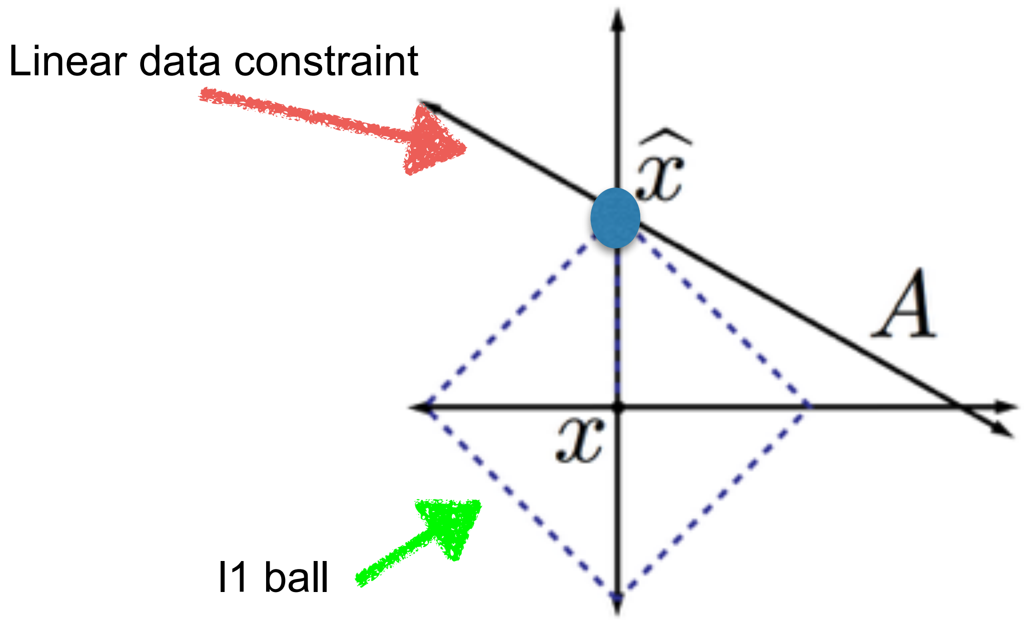
❖ Such formulation opens the door to the realm of convex optimisation, that offer an extremely versatile framework to solve many evolutions of this minimisation problem!

l1-norm minimisation problem

* Geometrical understanding of the l1-norm relaxation for the sparsity prior...

l1-norm

l2-norm



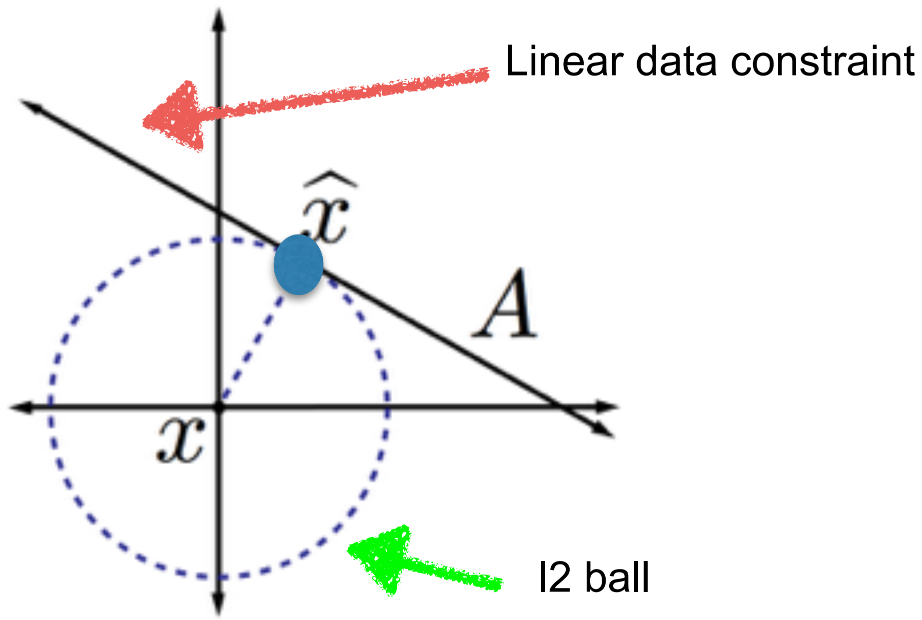


Image Courtesy Davenport et al. Introduction to Compressed Sensing

Restricted Isometry Property

* Theorem for l1-norm recovery. $\Theta = \Phi\Psi$ should be such that the distance between 2 sparse signals should be preserved in data space (Johnson-Lindenstrauss embedding)!!!

Definition:

$\Theta = \Phi\Psi$ satisfies the Restricted Isometry Property (RIP) of order K iff there exists some constant $\delta_K < 1$ such that $(1-\delta_K)\,||\boldsymbol{\alpha_K}||_2^2 \leq ||\Theta\boldsymbol{\alpha_K}||_2^2 \leq (1+\delta_K)\,||\boldsymbol{\alpha_K}||_2^2$ for all K-sparse vectors $\boldsymbol{\alpha_K}$.

Theorem:

If the RIP is satisfied with $\delta_{2K} < \dfrac{2}{3+\sqrt{7/4}} \simeq 0.4627$, then for all signals,

the BP minimization problem $\min_{\bar{\boldsymbol{\alpha}} \in \mathbb{C}^N} ||\bar{\boldsymbol{\alpha}}||_1$ subject to $||\boldsymbol{y} - \Theta\bar{\boldsymbol{\alpha}}||_2 \leq \epsilon$

provides accurate and stable reconstruction $\boldsymbol{\alpha}^*$: $||\boldsymbol{\alpha} - \boldsymbol{\alpha}^*||_2 \leq c\epsilon + \dfrac{d}{\sqrt{K}}||\boldsymbol{\alpha} - \boldsymbol{\alpha_K}||_1$,

for ϵ such that $||\boldsymbol{n}||_2 \leq \epsilon$, $\boldsymbol{\alpha_K}$ the best K-sparse approximation of the signal in the l1 sense, and c and d are constants depending only on δ_{2K} .

LATER ALSO CAME...

RANDOMNESS AND INCOHERENCE

Random sensing matrices

* Gaussian matrices satisfy the RIP for a number of measurements scaling with the signal sparsity rather than its dimension.

- A Random Gaussian Matrix $\Phi \in \mathbb{R}^{M \times N}$ can be constructed with entries sampled from i.i.d. Gaussian distributions $\mathcal{N}(0, M^{-1})$.

- **Theorem:**

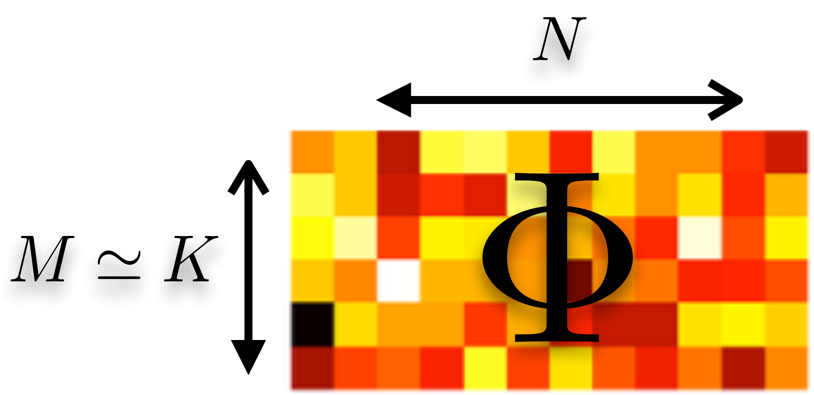
For a Gaussian measurement matrix, $\Theta = \Phi\Psi$ satisfies the RIP with overwhelming probability for a nearly **optimal** number of measurements

$$M > CK \ln(\frac{N}{K}),$$

for some constant C. The measurement scheme is **universal** in the sense that the result holds for all sparsity bases.

Random sensing matrices

* Gaussian matrices satisfy the RIP for a number of measurements scaling with the signal sparsity rather than its dimension.



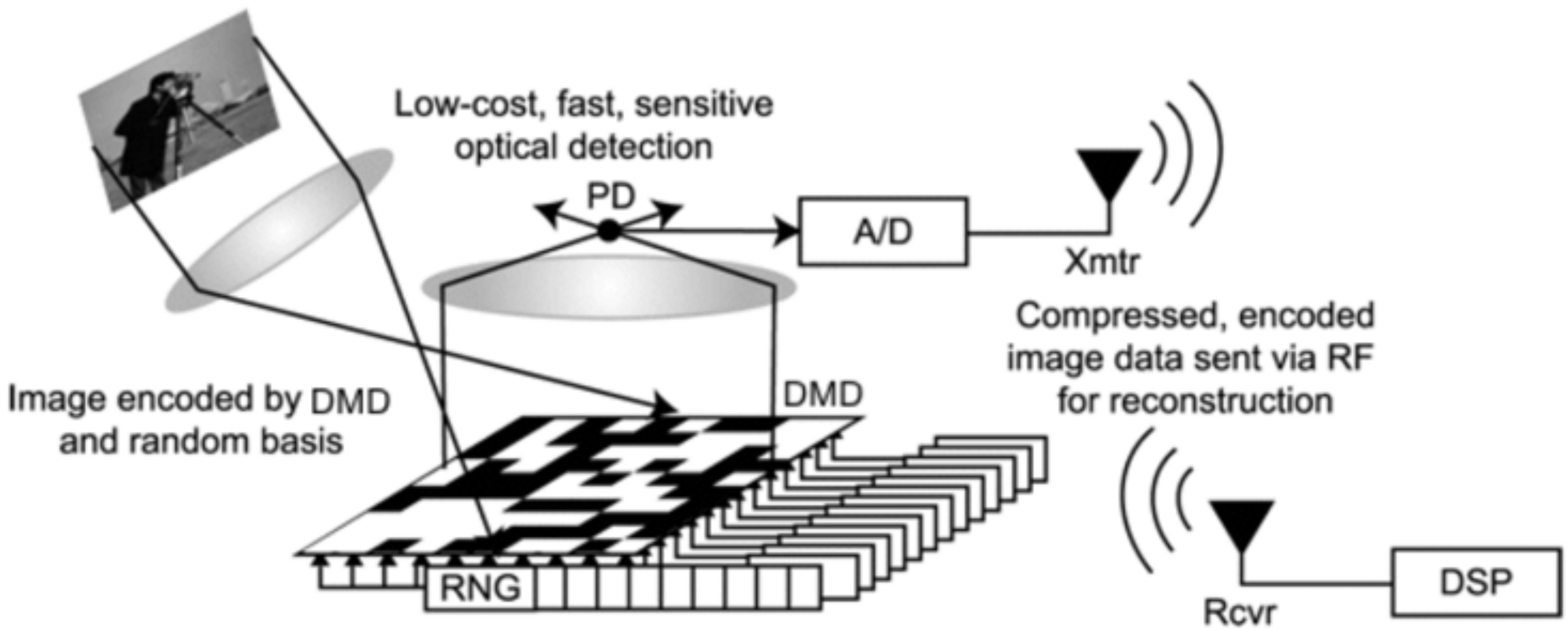
$$M \simeq K \ll N$$

❖However, purely random matrices are often **unrealistic** due to constraints of the physics of acquisition. It is also **inefficient** algorithmically due to the absence fast matrix multiplication (also leading to storage issues).

Random sensing matrices

* Identical results hold for Bernoulli matrices... solving the inefficiency problem due to the sparse binary nature of the measurement matrix.

❖ The Rice single pixel camera (image from http://dsp.rice.edu/cscamera) is based on such a sensing matrix. The typical advantage over CCD etc. is that you can image at any wavelength...



Digital Micromirror Device (DMD)

Structured random matrices

* Structured random matrices typically built by random sampling in an o.n. basis, incoherent with the sparsity basis, satisfy the RIP for low number of measurements.

- A structured random matrix can be built as $\Phi \equiv \mathsf{M}\Omega^{\dagger} \in \mathbb{C}^{M\times N}$, where $\Omega \in \mathbb{C}^{N\times N}$ is an o.n. basis and $\mathsf{M} \in \mathbb{C}^{M\times N}$ identifies a binary mask operating the random selection of M vectors in Ω . (Identical results also hold for general bounded o.n. systems).

- **Theorem:**

For such a measurement scheme, $\Theta = \Phi\Psi$ satisfies the RIP with overwhelming probability for a measurement number

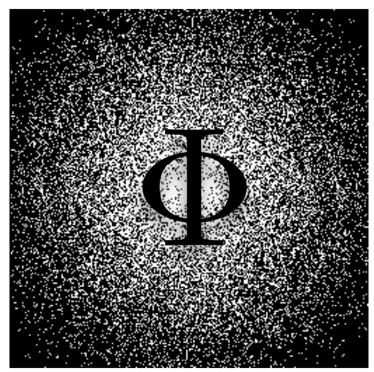
$$M > C\mu^2 K \ln^4 N,$$

for some constant C, and where μ stands for the coherence between the measurement and sparsity bases:

$$\mu = \sqrt{N} \max_{i,j} |\langle \Omega_{\cdot i} | \Psi_{\cdot j} \rangle|.$$

Structured random matrices

* Structured random matrices typically built by random sampling in an o.n. basis, incoherent with the sparsity basis, satisfy the RIP for low number of measurements.



Incoherence between sparsity and sensing

$$M \simeq \mu^2 K$$

❖ The measurement scheme is **realistic** … see all Fourier imaging applications including radio-interferometric imaging, and magnetic resonance imaging!

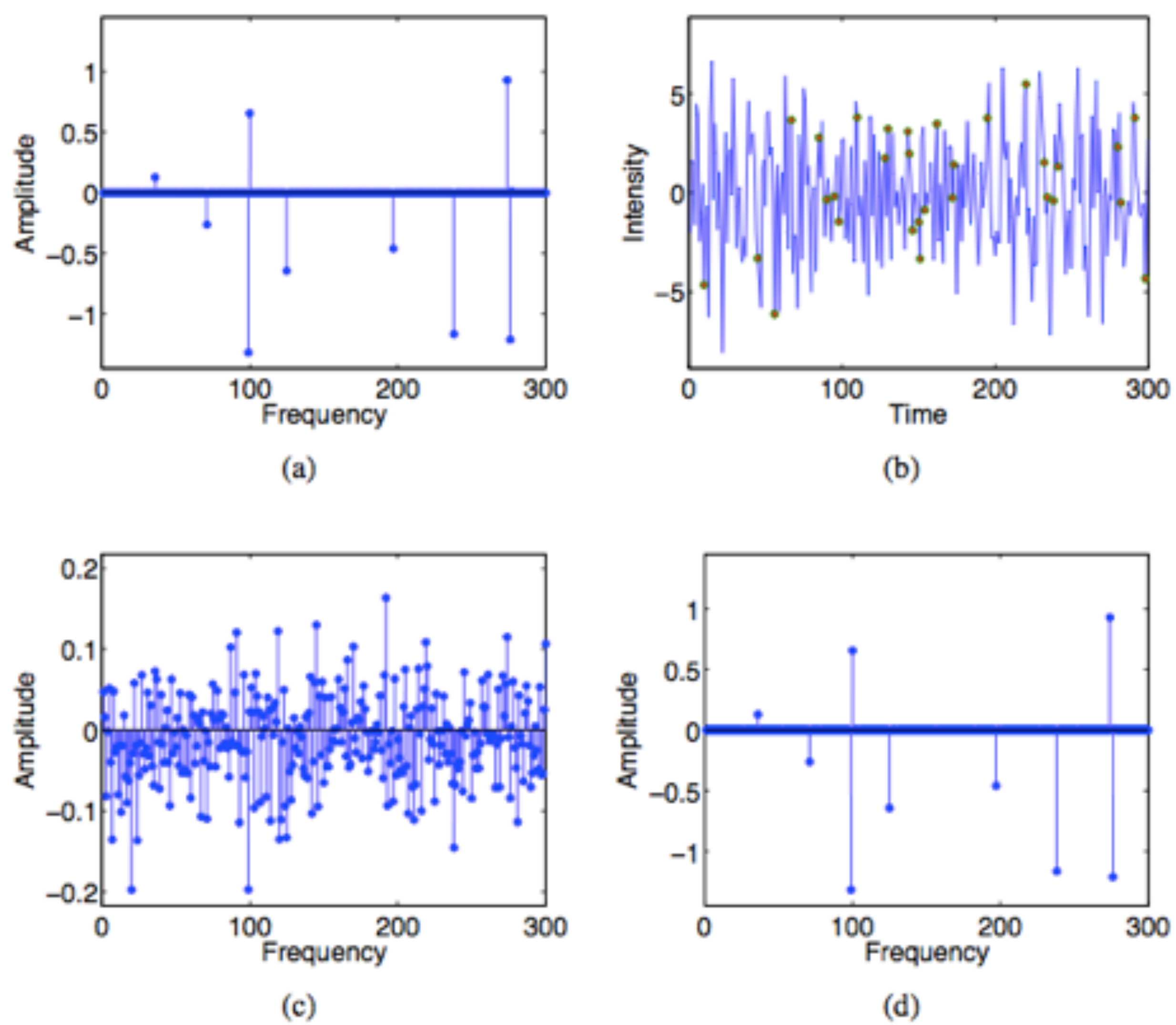
❖ This scheme is also **efficient** as fast matrix multiplications are often available. It is however **not universal**. **Optimality** is reached for the Fourier-Dirac case due to optimal incoherence: $\mu = 1$.

Structured random matrices

* Illustration: exact BP recovery of a signal of length 300, 10-sparse in Fourier, from 30 time samples (no noise).

original

sampling



12-minimization

11-minimization

Images: courtesy H. Rauhut

III

NEXT-GENERATION COMPRESSIVE IMAGING

G. Puy, P. Vandergheynst, R. Gribonval, and Y. Wiaux, "Universal and efficient compressed sensing by spread spectrum and application to realistic Fourier imaging techniques," EURASIP J. Adv. Signal Process. 6, 2012.

R. E. Carrillo, J. D. McEwen, D. Van De Ville, J.Ph. Thiran, and Y. Wiaux, "Sparsity averaging for compressive imaging," IEEE Signal Process. Lett., vol. 20, pp. 591-594, 2013, preprint arXiv:1208.2330 [cs.IT].

Y. Wiaux, L. Jacques, G. Puy, A. M. M. Scaife, and P. Vandergheynst, "Compressed sensing imaging techniques for radio interferometry," Mon. Not. R. Astron. Soc., vol. 395, pp. 1733-1742, 2009, preprint arXiv:0812.4933v2 [astro-ph].

Y. Wiaux, G. Puy, Y. Boursier, and P. Vandergheynst, "Spread spectrum for imaging techniques in radio interferometry," Mon. Not. R. Astron. Soc., vol. 400, pp. 1029-1038, 2009, preprint arXiv:0907.0944v2 [astro-ph.IM].

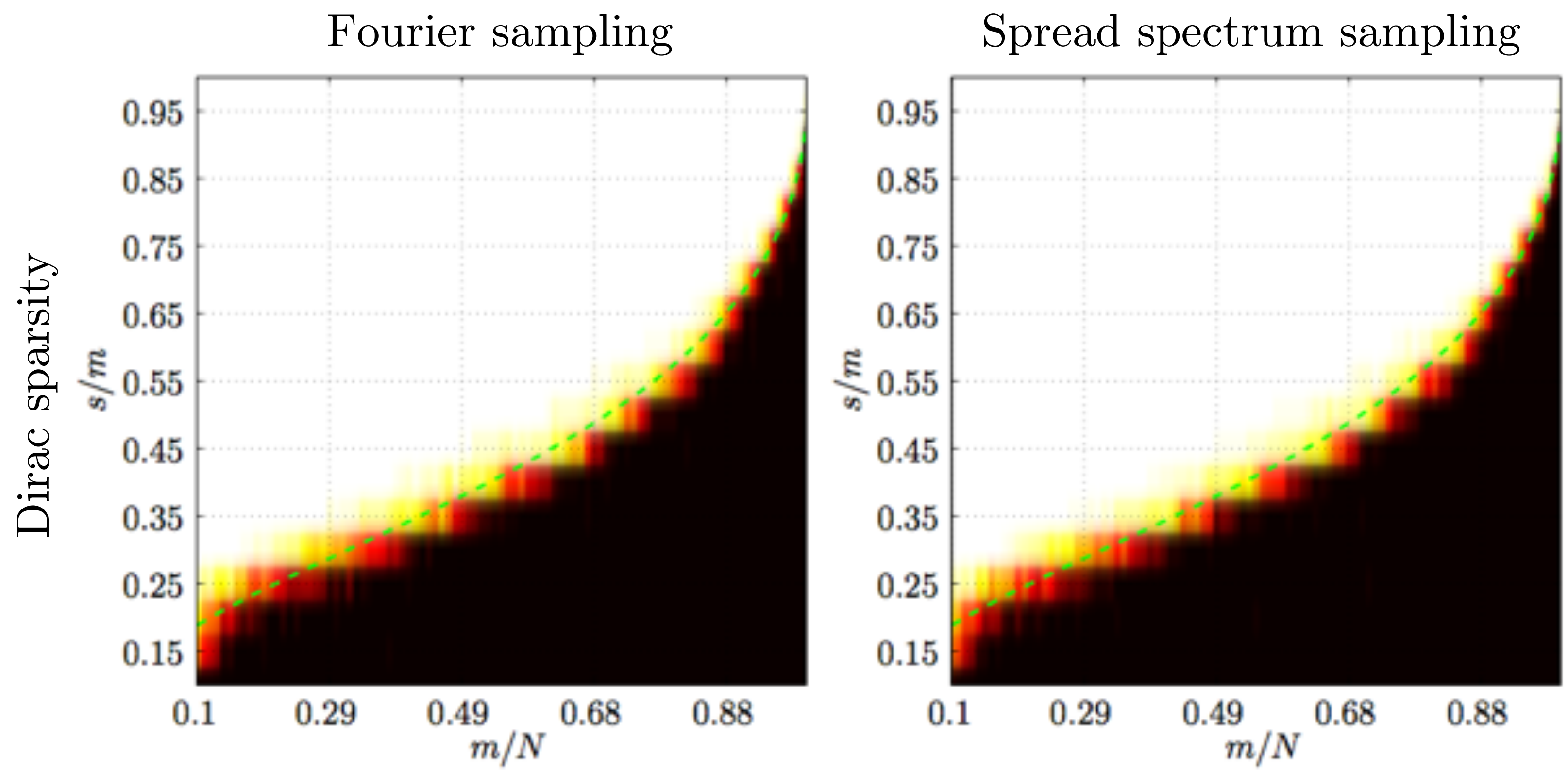
R. E. Carrillo, J. D. McEwen, and Y. Wiaux, "Sparsity averaging Reweighted Analysis (SARA): a novel algorithm for radio-interferometric imaging," Mon. Not. R. Astron. Soc., vol. 426, pp. 1223-1234, 2012, preprint arXiv:1205.3123 [astro-ph.IM].

Spread spectrum

* This **compressive sampling technique** relies on a random pre-modulation prior to Fourier-like random under-sampling... thus spreading the signal spectrum.

- **Efficient** (fast Fourier transforms) and **realistic.**
- Proved **optimal** and **universal** for o.n. sparsity bases, **and** redundant dictionaries.

Illustration for random s-sparse signals of size N=1024:



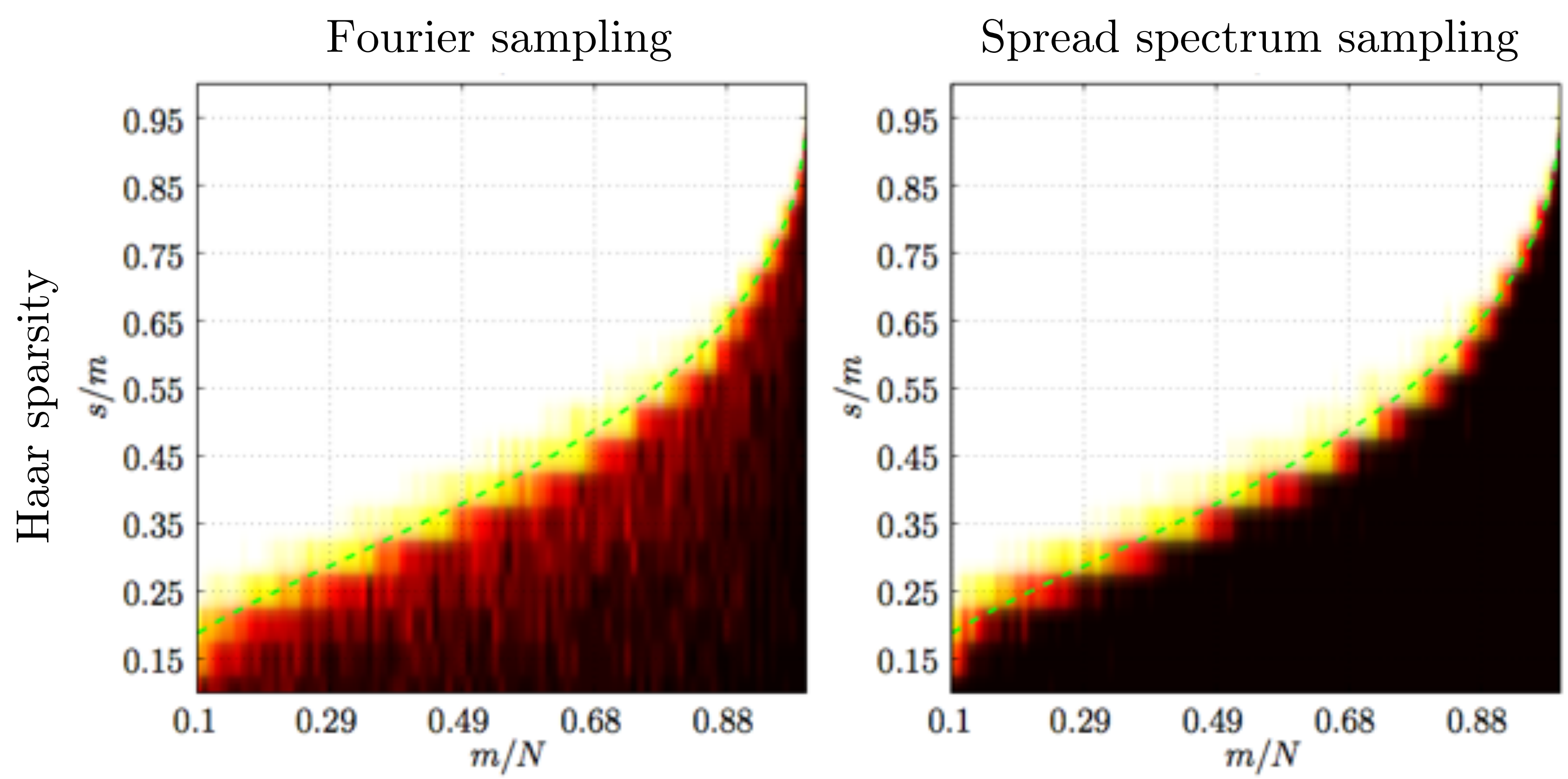
l1 phase transition diagrams

Spread spectrum

* This **compressive sampling technique** relies on a random pre-modulation prior to Fourier-like random under-sampling... thus spreading the signal spectrum.

- **Efficient** (fast Fourier transforms) and **realistic.**
- Proved **optimal** and **universal** for o.n. sparsity bases, **and** redundant dictionaries.

Illustration for random s-sparse signals of size N=1024:



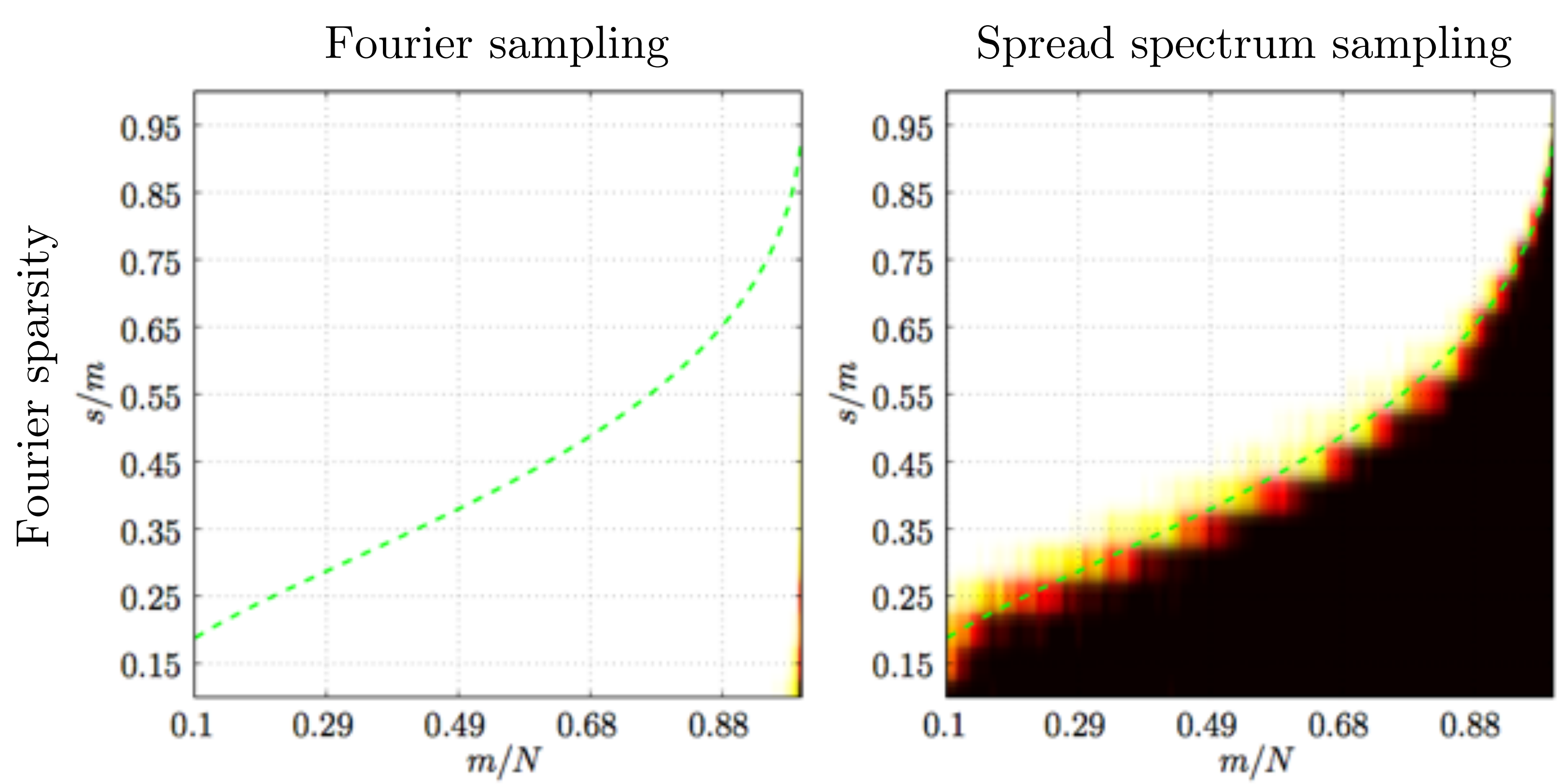
l1 phase transition diagrams

Spread spectrum

* This **compressive sampling technique** relies on a random pre-modulation prior to Fourier-like random under-sampling... thus spreading the signal spectrum.

- **Efficient** (fast Fourier transforms) and **realistic.**
- Proved **optimal** and **universal** for o.n. sparsity bases, **and** redundant dictionaries.

Illustration for random s-sparse signals of size N=1024:



l1 phase transition diagrams

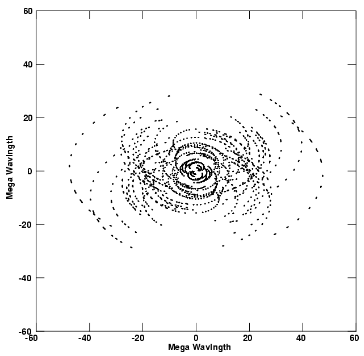
Impact in Radio interferometry...

* A spread spectrum phenomenon naturally appears in the radio-interferometric measurement equation through to the ‘w-component’ of the baselines on large FOV, and more generally due to the DDEs, due to their convolutional nature.

Fornax A radio emission
Image courtesy NRAO & Uson

Impact in Radio interferometry...

* Imaging quality at 10% sampling for random Fourier acquisition (30 dB input noise):

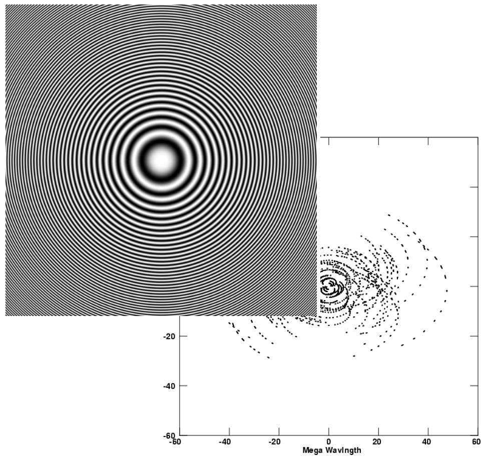


Fourier sampling

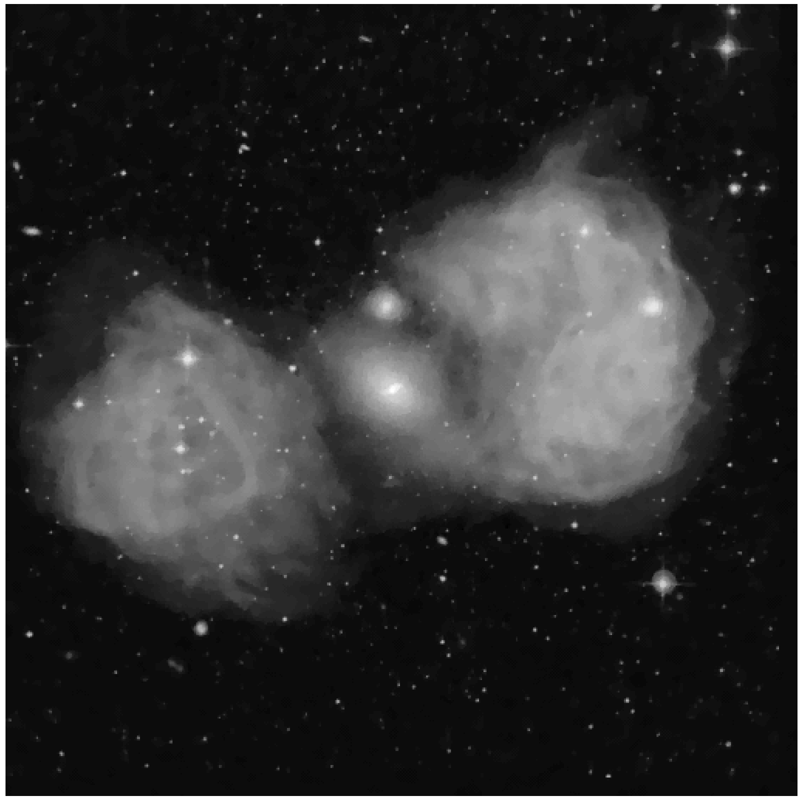
TV reconstruction (0.278 dB)

Impact in Radio interferometry. . .

* Quantum jump in imaging quality at 10% sampling for spread spectrum acquisition (30 dB input noise):



Spread spectrum sampling



TV reconstruction (21.8 dB)

SARA

* An **enhanced signal model** can be designed by acknowledging that natural signals often exhibit small average sparsity over multiple coherent frames...

❖ One is often left with a difficult choice to find the best sparsity basis as multiple bases offer 'some' sparsity. We can **enhance the signal model** and promote **average sparsity** over a large number of coherent frames:

$$\Psi \equiv \frac{1}{\sqrt{q}}[\Psi_1, \Psi_2, \ldots \Psi_q].$$

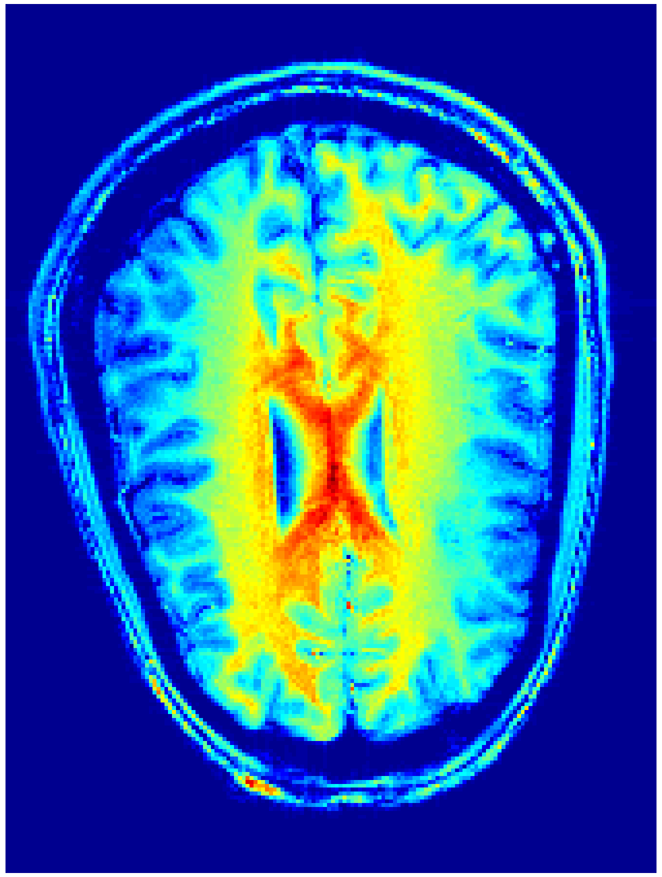
❖ The analysis prior consisting in minimising the l0 norm of the signal projection in Ψ simply corresponds to minimising the average (or total) sparsity over all frames:

$$\|\Psi^\dagger \bar{\boldsymbol{x}}\|_0 \sim \frac{1}{q}\sum_{b=1}^{q} \|\Psi_b^\dagger \bar{\boldsymbol{x}}\|_0,$$

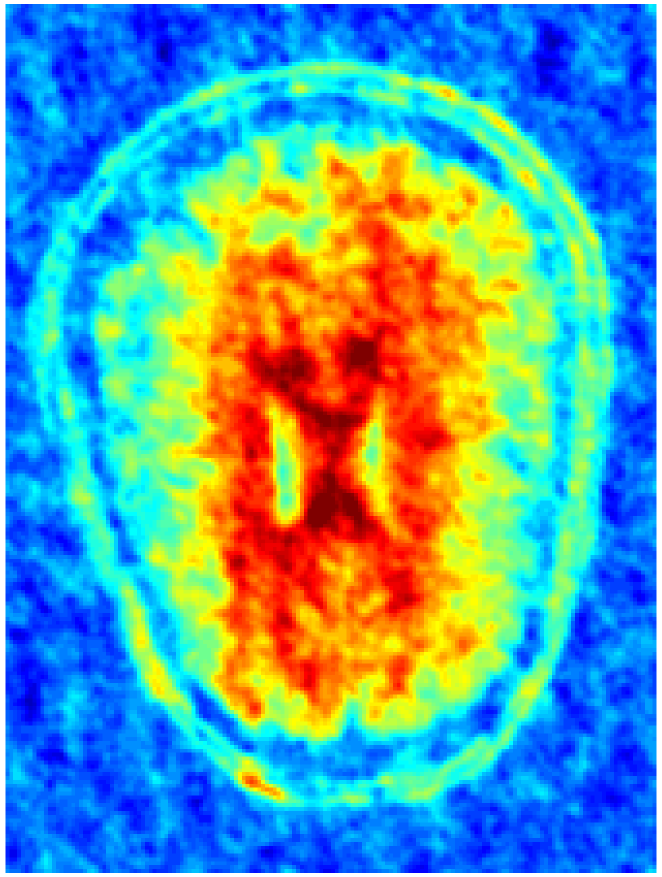
with
$$\Psi^\dagger \bar{\boldsymbol{x}} \equiv \bar{\boldsymbol{\alpha}} \equiv [\bar{\boldsymbol{\alpha}}_1^\dagger, \cdots, \bar{\boldsymbol{\alpha}}_q^\dagger]^\dagger.$$

Impact for MR imaging

* Superiority of SARA for Fourier acquisition on a brain image, for 5% under-sampling, i.e. 20-fold acceleration, with 30 dB input noise: visual assessment.



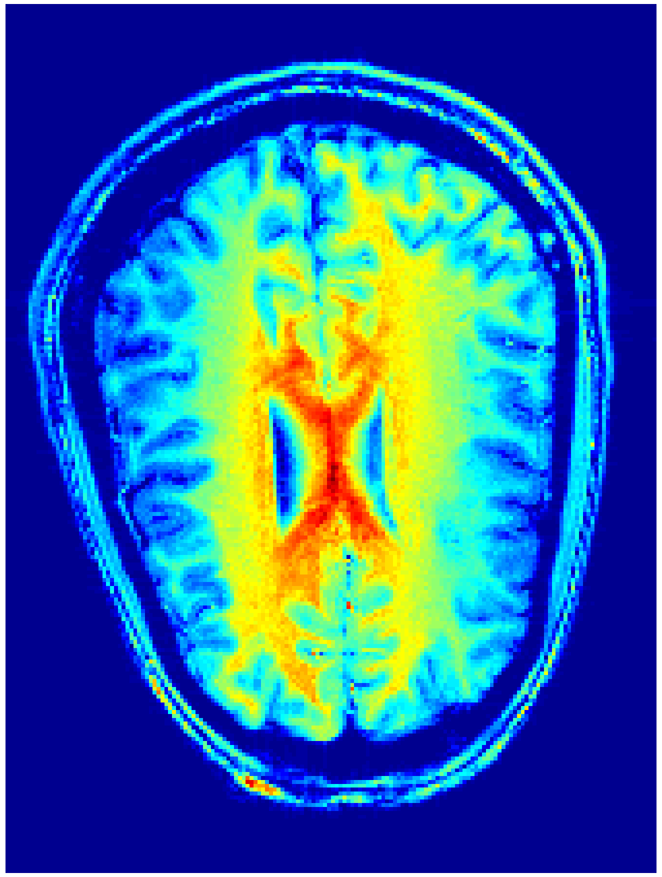
Original



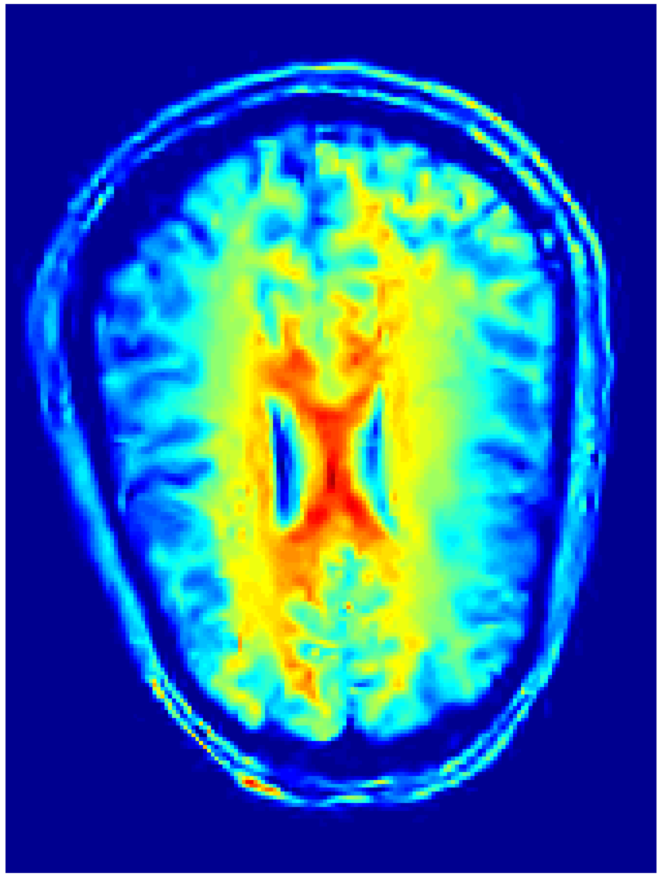
Backprojected image

Impact for MR imaging

* Superiority of SARA for Fourier acquisition on a brain image, for 5% under-sampling, i.e. 20-fold acceleration, with 30 dB input noise: visual assessment.



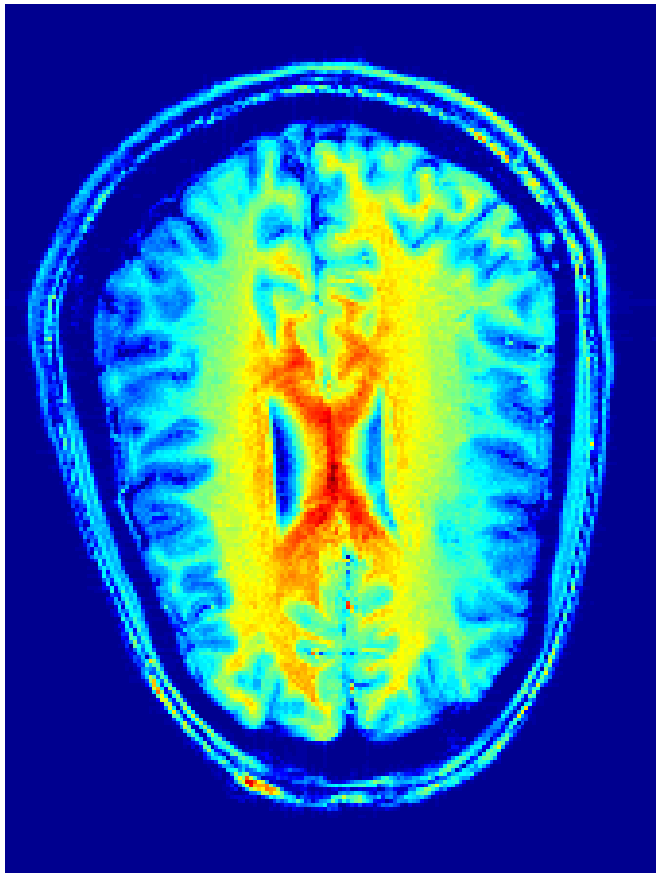
Original



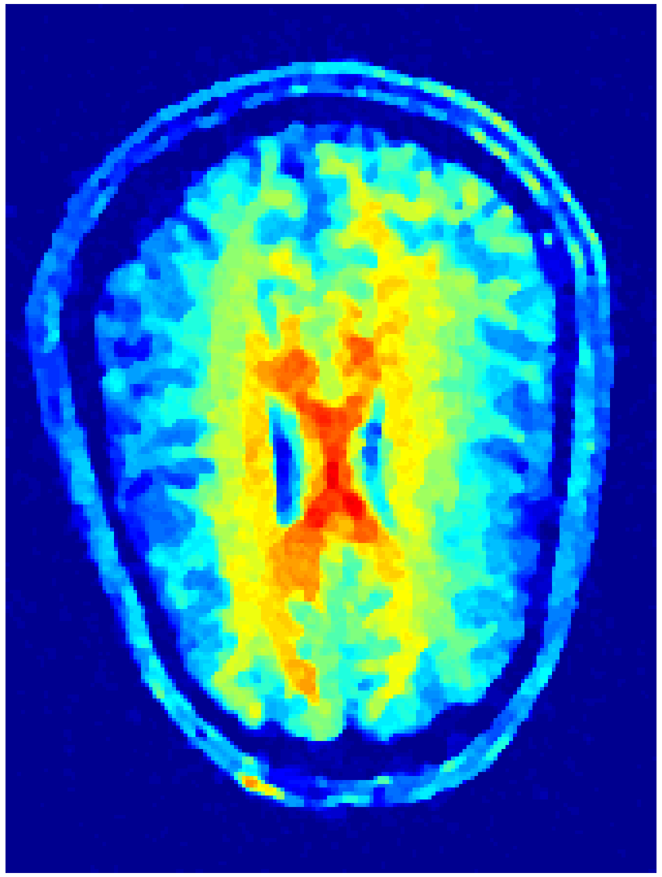
SARA (18.8 dB)

Impact for MR imaging

* Superiority of SARA for Fourier acquisition on a brain image, for 5% under-sampling, i.e. 20-fold acceleration, with 30 dB input noise: visual assessment.



Original



TV (17.3 dB)

TAKE-HOME MESSAGE

- **Compressive sampling** enables accurate sparse signal recovery from **drastic under-sampling**, relying on randomness, incoherence, and nonlinear recovery.

- Beyond the first steps of the theory, **new acquisition approaches** as spread spectrum and **enhanced signal models** as sparsity averaging are important evolutions.

- **The theory brings important evolutions for Fourier imaging applications**, in particular for radio astronomy and medical imaging.