Sparsity in imaging:
Fourier measurements in compressed sensing

Clarice Poon
University of Bath

January, 2019

Recovery statement

We have so far shown:

Theorem

Let $x \in \mathbb{C}^N$ *and suppose we observe* $y = P_\Omega U x + e$ *whre*

- U *be an unitary matrix,*
- Ω *consists of* m *indices chosen uniformly at random.*
- $\|e\| \leqslant \eta$.

Suppose that

$$m \gtrsim K^2 s \ln(N) \ln(\varepsilon^{-1}),$$

where $K \stackrel{\text{def.}}{=} \sqrt{N}\mu$ *and* $\mu \stackrel{\text{def.}}{=} \max_{k,j} \left|U_{k,j}\right|$ *is the coherence of* U*.*
Then, with probability at least $1 - \varepsilon$*, any solution* $\tilde{x}$ *to*

$$\min \|z\|_1 \text{ subject to } \|P_\Omega U z - y\| \leqslant \eta$$

satisfies

$$\|\tilde{x} - x\|_2 \lesssim \sigma_s(x)_1 + \sqrt{s}\eta.$$

Recovery statement

We have so far shown:

Theorem

Let $x \in \mathbb{C}^N$ and suppose we observe $y = P_\Omega U x + e$ whre

- *U be an unitary matrix,*
- *Ω consists of m indices chosen uniformly at random.*
- *$\|e\| \leqslant \eta$.*

Suppose that

$$m \gtrsim K^2 s \ln(N) \ln(\varepsilon^{-1}),$$

where $K \stackrel{\text{def.}}{=} \sqrt{N}\mu$ and $\mu \stackrel{\text{def.}}{=} \max_{k,j} \left|U_{k,j}\right|$ is the coherence of U. Then, with probability at least $1-\varepsilon$, any solution $\tilde{x}$ to

$$\min \|z\|_1 \text{ subject to } \|P_\Omega U z - y\| \leqslant \eta$$

satisfies

$$\|\tilde{x} - x\|_2 \lesssim \sigma_s(x)_1 + \sqrt{s}\eta.$$

Accurate recovery guaranteed provided that

- U is incoherent, that is $\mu = 1$,
- x is s-sparse (or approximately s-sparse),
- We observe $m = \mathcal{O}(s \log(N))$ samples uniformly at random.

Compressed sensing in action

Applications: Magnetic Resonance Imaging (MRI), X-ray Computed Tomography, Electron Microscopy, Seismology, Radio interferometry,....

Mathematically: We observe samples of the Fourier transform, and typical images are sparse in wavelets.

CS approach: Solve

$$\min_{z \in \mathbb{C}^{N^2}} \|U_* z\|_1 \text{ subject to } \left\| P_\Omega U_{df} z - P_\Omega \hat{f} \right\|_2 \leqslant \delta$$

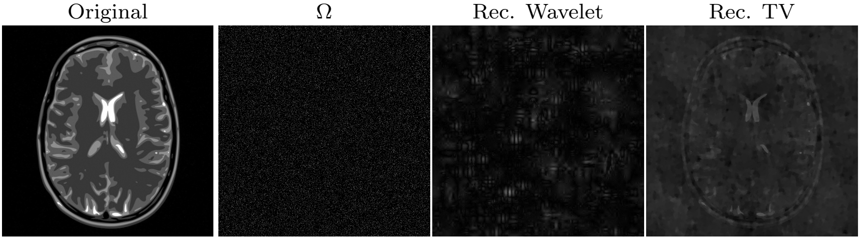
where U_{df} is the discrete Fourier transform, and U_* is some sparsifying transform (e.g. wavelets).

Compressed sensing in action

If U_* is a discrete wavelet transform, then $\mu(U_{df}U_*^{-1}) = 1$ so $K = \sqrt{N}$, so

$$K^2 \cdot s \cdot \log(N)\log(\varepsilon^{-1}+1) > N!$$

Also, uniform random sampling does not work.



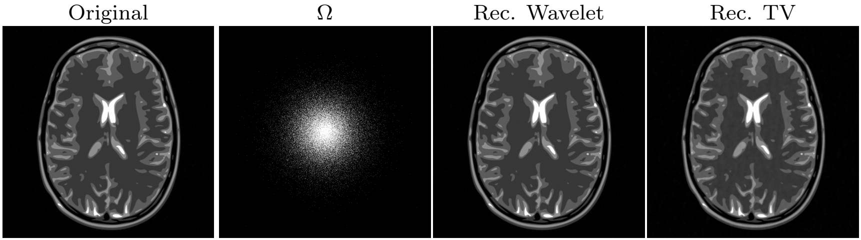
Test phantom constructed by Guerquin-Kern, Lejeune, Pruessmann, Unser, 2012

Compressed sensing in action

If U_* is a discrete wavelet transform, then $\mu(U_{df} U_*^{-1}) = 1$ so $K = \sqrt{N}$, so

$$K^2 \cdot s \cdot \log(N) \log(\varepsilon^{-1} + 1) > N!$$

Also, uniform random sampling does not work.



Lustig, Donoho & Pauli '07, Lustig et al. '08: Sample more densely at low Fourier frequencies and less at higher Fourier frequencies. *Why?*

Test phantom constructed by Guerquin-Kern, Lejeune, Pruessmann, Unser, 2012

Outline

The coherence barrier

For essentially any wavelet basis $\{\varphi_j\}_{j\in\mathbb{N}}$, if $U_{dw,N}$ is its discrete wavelet transform and $U_{df,N}$ is the discrete Fourier transform, then

$$U_{df,N}U_{dw,N}^{-1} \xrightarrow{\text{WOT}} U, \qquad N \to \infty$$

where $\{\psi_j\}_{j\in\mathbb{N}} = \{e^{2\pi ik\cdot}\}_{k\in\mathbb{Z}}$,

$$U = \begin{pmatrix} \langle\varphi_1,\psi_1\rangle & \langle\varphi_2,\psi_1\rangle & \cdots \\ \langle\varphi_1,\psi_2\rangle & \langle\varphi_2,\psi_2\rangle & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}, \quad \mu(U) \geqslant c.$$

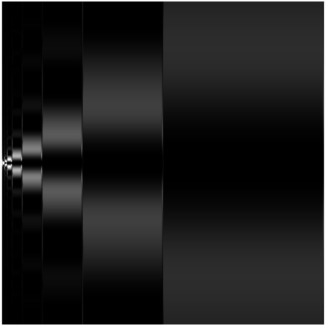
Any systems arising from the discretization of continuous problems will always run into the coherence barrier.

Asymptotic incoherence

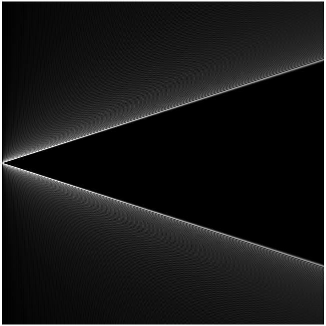
If U is the Fourier-wavelets matrix, then

$$\mu(P_N^{\perp}U), \mu(UP_N^{\perp}) = \mathcal{O}\left(N^{-1}\right).$$

Fourier to DB4



Fourier to Legendre polynomials



Notation: $P_N x = (x_1, \cdots, x_N, 0, 0, \cdots)$ and $P_N^{\perp} x = (0, \cdots, 0, x_{N+1}, x_{N+2}, \cdots)$

Local coherence

Instead of coherence, divide U into rectangular blocks using $\mathbf{N} = (N_j)_{j=1}^r, \mathbf{M} = (M_j)_{j=1}^r$, and consider local coherence

$$\mu_{\mathbf{N},\mathbf{M}}(k,l) = \mu(P_{N_{k-1}}^{N_k} U P_{M_{l-1}}^{M_l}).$$

where $P_n^m \alpha = (\ldots, 0, \alpha_{n+1}, \alpha_{n+2}, \ldots, \alpha_m, 0 \ldots)$.

Implication of asymptotic incoherence: sample more at low Fourier frequencies where the local coherence is high and less at higher Fourier frequencies.

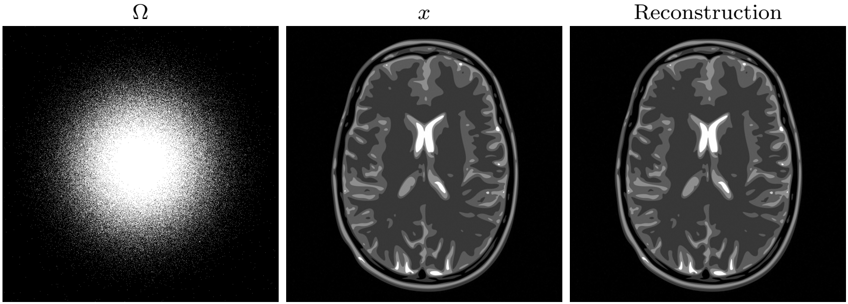
Outline

Sparsity and the flip test

In standard CS, the only signal structure considered is sparsity and RIP based results consider the recovery of all s-sparse signals using one Ω. In contrast, the flip test will demonstrate that we must look beyond sparsity.

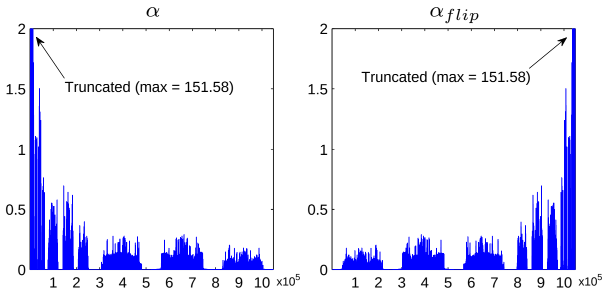
Consider the reconstruction of x from $P_\Omega U_{df} x$ by solving

$$\min \|z\|_1 \text{ subject to } P_\Omega U_{df} U_{dw}^{-1} z = P_\Omega U_{df} x.$$



The flip test

Let α be the wavelet coefficients of x.



Let $\alpha^{flip} = (\alpha_N, \ldots, \alpha_1)$ and $x^{flip} = U_{dw}^{-1}\alpha^{flip}$.
If it is enough to consider sparsity when choosing Ω, then for the same Ω,

$$\tilde{\alpha}_f \in \arg\min_z \|z\|_1 \text{ subject to } P_\Omega U_{df} U_{dw}^{-1} z = P_\Omega U_{df} x^{flip},$$

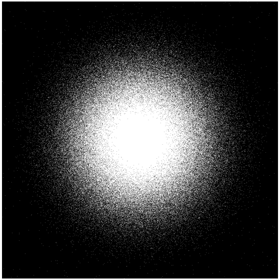
would yield

$$\tilde{\alpha}_f \approx \alpha^{flip} \implies \tilde{\alpha}_f^{flip} \approx \alpha$$

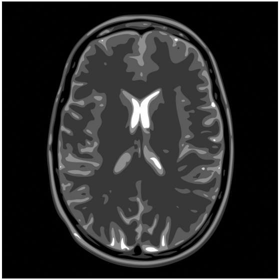
$$\hat{x} = U_{dw}^{-1}\tilde{\alpha}_f^{flip} \approx x = U_{dw}^{-1}\alpha$$

The flip test

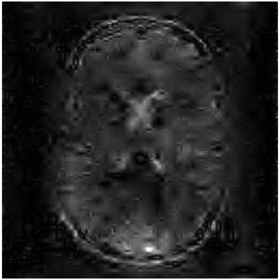
Ω



Standard reconstruction

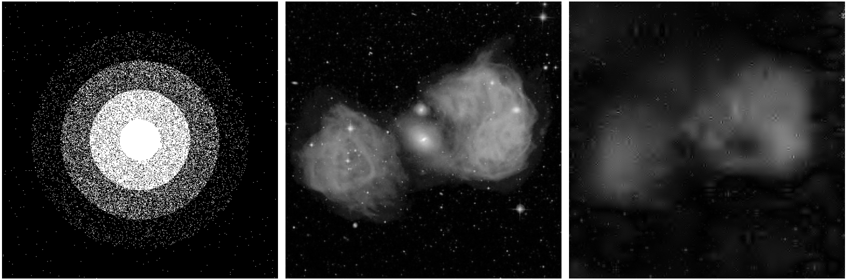


Reconstruction from flipped coefficients



The flip test

Ω | Standard reconstruction | Reconstruction from flipped coefficients



We can repeat this test for different images, different sampling patterns,

The flip test

Ω

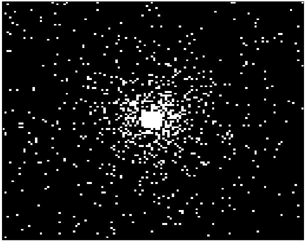


Image 1

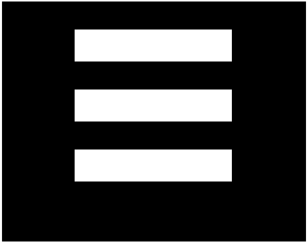
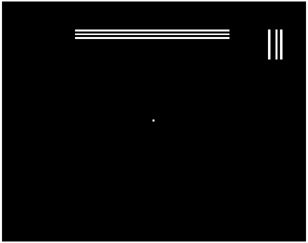


Image 2

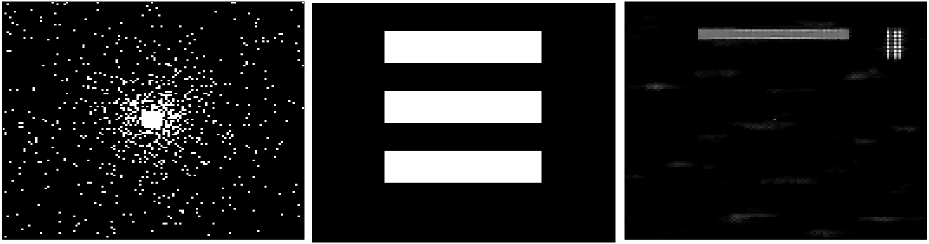


We can repeat this test for different images, different sampling patterns, or even different sparsifying transforms such as *-lets or total variation,

$$\min_{z \in \mathbb{C}^{N^2}} \|z\|_{TV} \text{ subject to } P_\Omega U_{df} z = P_\Omega U_{df} x$$

The flip test

Ω | Image 1 reconstruction | Image 2 reconstruction



We can repeat this test for different images, different sampling patterns, or even the sparsifying transform to that of *-lets or total variation,

$$\min_{z \in \mathbb{C}^{N^2}} \|z\|_{TV} \text{ subject to } P_\Omega U_{df} z = P_\Omega U_{df} x$$

The optimal choice of Ω cannot depend on sparsity alone.

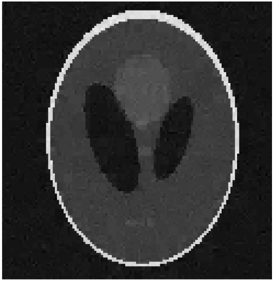
Remark on the RIP

The flip test demonstrated that Ω cannot depend on sparsity alone, and in fact, the RIP is absent.

This is in contrast to random Gaussian measurements are insensitive to sparsity structure:

Standard reconstruction

Reconstruction from flipped coefficients



Asymptotic sparsity

Natural images are not just sparse, but asymptotically sparse.

Given $\varepsilon \in (0,1)$ and

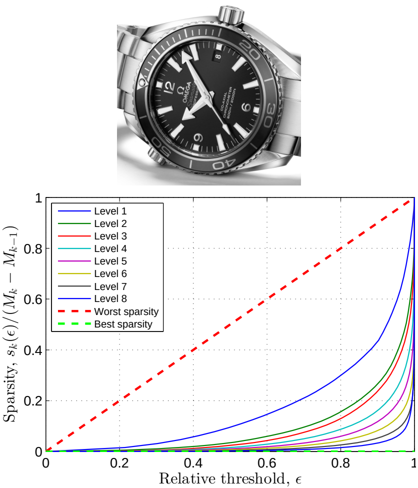
$$f = \sum_{j \in \mathbb{N}} \alpha_j \varphi_j.$$

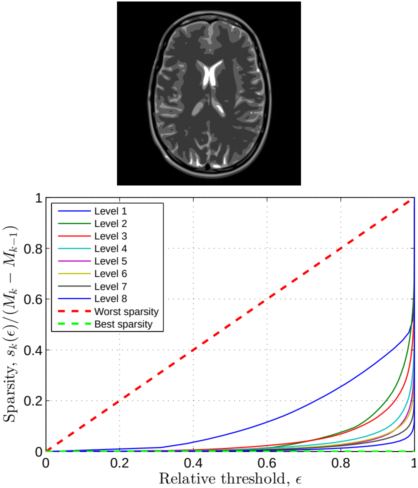
we define the number of significant wavelet coefficients of f in the k^{th} scale as

$$s_k(\varepsilon) = \min\{n : \left\| \sum_{j=1}^{n} \alpha_{\pi(j)} \varphi_{\pi(j)} \right\|_2 \geqslant \varepsilon \left\| \sum_{j=1+M_{k-1}}^{M_k} \alpha_j \varphi_j \right\|_2 \}$$

where the $\{M_{k-1}+1, \ldots, M_k\}$ be indices corresponding to the k^{th} scale and π is a permutation of the indices in $\{M_{k-1}+1, \ldots, M_k\}$ such that $\left|\alpha_{\pi(1)}\right| \geqslant \left|\alpha_{\pi(2)}\right| \geqslant \left|\alpha_{\pi(3)}\right| \geqslant \ldots$.

Asymptotic sparsity





$$\frac{s_k(\varepsilon)}{M_k - M_{k-1}} \to 0, \qquad k \to \infty$$

Variable density sampling patterns work because they exploit this additional structure.

Sparsity in levels

For $\mathbf{M} = (M_j)_{j=1}^r \in \mathbb{N}^r, s = (s_j)_{j=1}^r \in \mathbb{N}^r$ with $0 = M_0 < M_1 < \ldots < M_r = N$, $\alpha \in \mathbb{C}^N$ is $(\mathbf{s}, \mathbf{M})$-sparse if

$$|\{j : \alpha_j \neq 0\} \cap \{M_{k-1}+1, \ldots, M_k\}| = s_k.$$

Outline

Multi-level sampling scheme

Let $r \in \mathbb{N}$, $\mathbf{N} = \{N_k\}_{k=1}^{r} \in \mathbb{N}^r$, $\mathbf{m} = \{m_k\}_{k=1}^{r} \in \mathbb{N}^r$ be such that

$$0 = N_0 < N_1 < \cdots < N_r = N, \qquad m_k \leqslant N_k - N_{k-1}.$$

$\Omega = \Omega_1 \cup \cdots \cup \Omega_r$ is an (N, m)-sampling scheme if Ω_k consists of m_k indices drawn uniformly at random from $\{N_{k-1}+1, \ldots, N_k\}$.

Multi-level sampling scheme

Let $r \in \mathbb{N}$, $\mathbf{N} = \{N_k\}_{k=1}^r \in \mathbb{N}^r$, $\mathbf{m} = \{m_k\}_{k=1}^r \in \mathbb{N}^r$ be such that

$$0 = N_0 < N_1 < \cdots < N_r = N, \qquad m_k \leqslant N_k - N_{k-1}.$$

$\Omega = \Omega_1 \cup \cdots \cup \Omega_r$ is an (N, m)-sampling scheme if Ω_k consists of m_k indices drawn uniformly at random from $\{N_{k-1} + 1, \ldots, N_k\}$.

Q: *Let $U \in \mathcal{B}(\ell^2(\mathbb{N}))$ be an isometry. If x is approximately $(\mathbf{s}, \mathbf{M})$-sparse*

$$\sigma_{\mathbf{s},\mathbf{M}}(\alpha) = \inf_{z \text{ is } (\mathbf{s},\mathbf{M})\text{-sparse}} \|z - \alpha\|_1 \ll 1.$$

then how should $\mathbf{N}$ and $\mathbf{m}$ be chosen so as to guarantee robust and stable recovery of x by solving

$$\inf_{z \in \mathbb{C}^N} \|z\|_1 \text{ subject to } \|P_\Omega U x - P_\Omega U z\|_2 \leqslant \delta?$$

Recovery result

Let $\varepsilon \in (0, e^{-1}]$ and x be approximately $(\mathbf{s}, \mathbf{M})$-sparse. Suppose that $\Omega = \Omega_{\mathbf{N},\mathbf{m}}$ satisfies the following.

(ii) $m_k \gtrsim (N_k - N_{k-1}) \cdot \left(\sum_{l=1}^{r} \mu_{\mathbf{N},\mathbf{M}}(k,l) \cdot s_l\right) \cdot \log(s\varepsilon^{-1}) \cdot \log(N)$,

(iii) $m_k \gtrsim \hat{m}_k \cdot \log(s\varepsilon^{-1}) \cdot \log(N)$, where $\hat{m}_k$ satisfies

$$1 \gtrsim \sum_{k=1}^{r} \left(\frac{N_k - N_{k-1}}{\hat{m}_k} - 1 \right) \cdot \mu_{\mathbf{N},\mathbf{M}}(k,l) \cdot \tilde{s}_k.$$

Then, with probability exceeding $1 - \varepsilon$, any minimizer $\hat{x}$ satisfies

$$\|\hat{x} - x\|_2 \lesssim \delta \cdot \left(1 + \frac{\sqrt{\log_2(\varepsilon^{-1})}}{\log_2(N)} \right) \cdot \sqrt{s} + \sigma_{\mathbf{s},\mathbf{M}}(x).$$

Recovery of wavelet coefficients from partial Fourier data

- $\{N_k\}_{k=1}^r$ and $\{M_k\}_{k=1}^r$ correspond to wavelet scales.
- The mother wavelet Ψ has v vanishing moments.
- There exists $\alpha \geqslant 1$, $C > 0$ such that $\left|\hat{\Psi}(\xi)\right| \leqslant \frac{C}{(1+|\xi|)^\alpha}$ for all $\xi \in \mathbb{R}$.

It suffices that

$$m_k \gtrsim \mathcal{L} \cdot \left(\hat{s}_k + \sum_{l=1}^{k-2} s_j \cdot 2^{-(\alpha - \frac{1}{2})(k-l)} + \sum_{l=k+2}^{r} s_l \cdot 2^{-v(l-k)} \right)$$

where $\hat{s}_k = \max\{s_{k-1}, s_k, s_{k+1}\}$ and $\mathcal{L} = \log(s\varepsilon^{-1}) \cdot \log(N)$

NB: $m_1 + \ldots + m_r \gtrsim \mathcal{L} \cdot (s_1 + \ldots + s_r)$.

Resolution Dependence (5% samples, varying resolution)

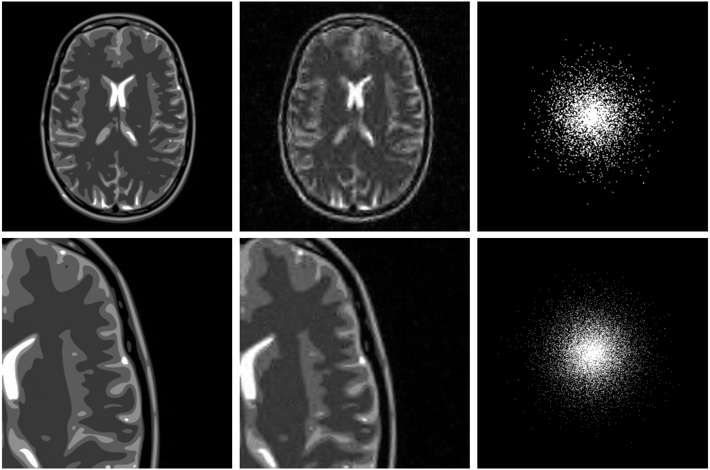
Asymptotic sparsity and asymptotic incoherence are only witnessed when N is large. Thus, V. D. sampling only reaps their benefits for large values of N and the success of compressed sensing is resolution dependent.

256x256

Error:
19.86%

512x512

Error:
10.69%



Resolution Dependence (5% samples, varying resolution)

1024x1024

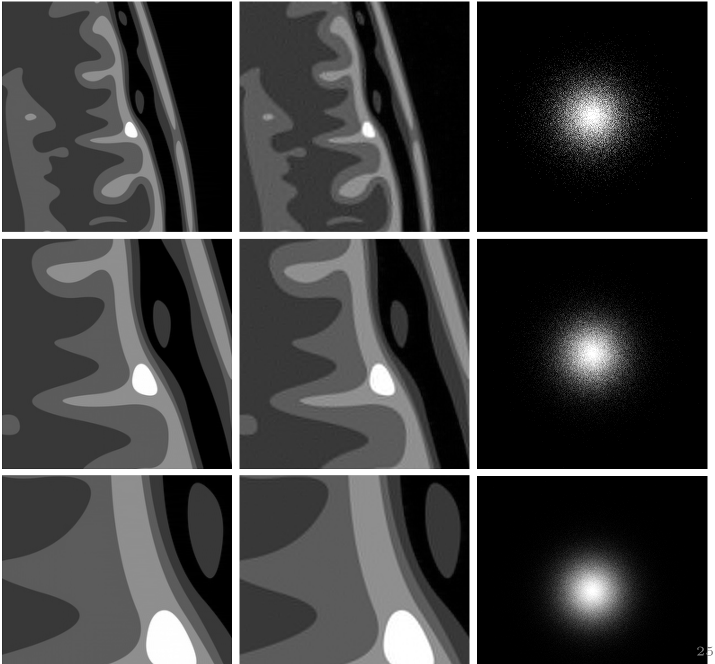
Error:
7.35%

2048x2048

Error:
4.87%

4096x4096

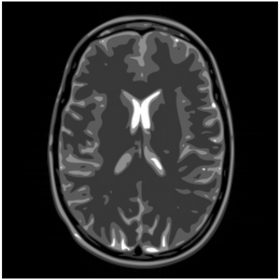
Error:
3.06%



Recovering Fine Details

At finer wavelet scales, the presence of sparsity and incoherence with Fourier samples allows us to subsample. Thus, compressed sensing allows one to enhance fine details without increasing the number of samples.

In the next example, consider the reconstruction of a 2048×2048 test phantom with details added at the finest wavelet scale.



Recovering Fine Details

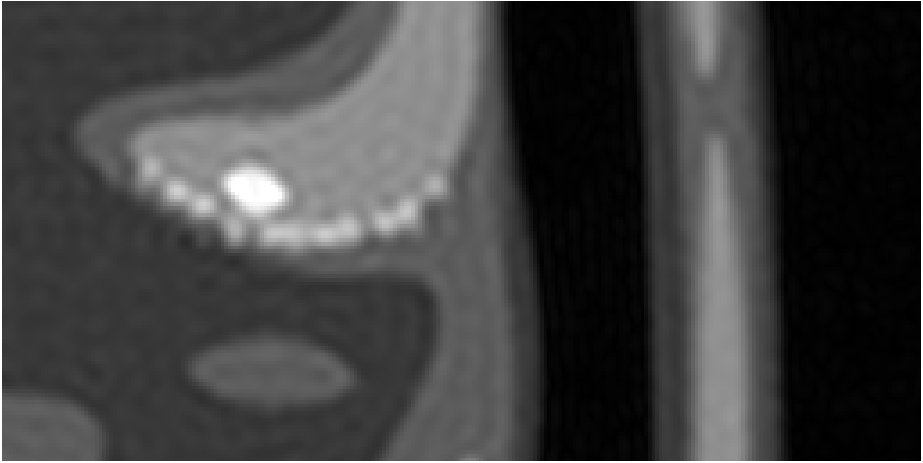


Figure: 2048×2048 linear reconstruction from the first 512×512 Fourier samples (6.25%)

Recovering Fine Details

Figure: 2048 × 2048 reconstruction from a multilevel scheme using 512 × 512 Fourier samples (6.25%)

Summary

- There are many real world problems where there is no incoherence or RIP.
- The case of wavelet sparsity and Fourier measurements in interesting for many imaging applications and here, instead of incoherence and sparsity, we have asymptotic incoherence and asymptotic sparsity.
- Two key consequences:
  1. (1) CS is resolution dependent.
  2. (2) Successful recovery is signal dependent, thus, an understanding of the structure imposed by the sparsifying transform can lead to optimal sampling patterns.
- On a practical note, one should see compressed sensing in these situations as a means of enhancing resolution ...

Sources

- **Breaking the coherence barrier: A new theory for compressed sensing.** Adcock, Hansen, Poon & Roman, *Forum of Mathematics, Sigma. Vol. 5. Cambridge University Press (2017).*
- **Stable and robust sampling strategies for compressive imaging.** Krahmer & Ward, *IEEE transactions on image processing 23.2 (2014): 612-622*
- **Compressed sensing with structured sparsity and structured acquisition.** Boyer, Bigot & Weiss. *Applied and Computational Harmonic Analysis (2017)*