Radio interferometric imaging with compressive sensing

Jason McEwen

http://www.jasonmcewen.org/

Department of Physics and Astronomy, University College London (UCL)

London Cosmology Discussion Meeting (LCDM) :: January 2013

Outline

Outline

1. Radio interferometry
2. An introduction to compressive sensing
3. Compressed sensing for radio interferometric imaging
4. Spread spectrum
5. Continuous visibilities

Next-generation of radio interferometry rapidly approaching

- Square Kilometre Array (SKA) first observations planned for 2019.
- Many other pathfinder telescopes under construction, *e.g.* LOFAR, ASKAP, MeerKAT, MWA.
- New modelling and imaging techniques required to ensure the next-generation of interferometric telescopes reach their full potential.

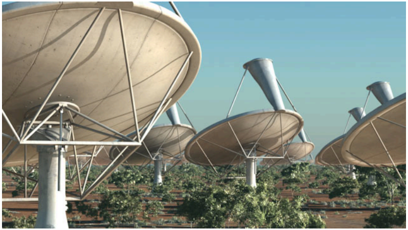
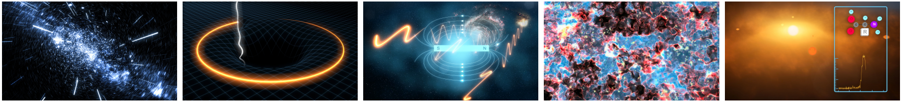


Figure: Artist impression of SKA dishes. [Credit: SKA Organisation]



(a) Dark-energy (b) GR (c) Cosmic magnetism (d) EoR (e) Exoplanets

Figure: SKA science goals. [Credit: SKA Organisation]

Radio interferometry

- The complex visibility measured by an interferometer is given by

$$y(\boldsymbol{u}, w) = \int_{D^2} A(\boldsymbol{l})\, x_{\rm p}(\boldsymbol{l})\, \mathrm{e}^{-\mathrm{i}2\pi[\boldsymbol{u}\cdot\boldsymbol{l} + w\,(n(\boldsymbol{l})-1)]}\, \frac{\mathrm{d}^2\boldsymbol{l}}{n(\boldsymbol{l})}$$
$$= \int_{D^2} A(\boldsymbol{l})\, x_{\rm p}(\boldsymbol{l})\, C(\|\boldsymbol{l}\|_2)\, \mathrm{e}^{-\mathrm{i}2\pi\boldsymbol{u}\cdot\boldsymbol{l}}\, \frac{\mathrm{d}^2\boldsymbol{l}}{n(\boldsymbol{l})} ,$$

  where $\boldsymbol{l} = (l, m)$, $\|\boldsymbol{l}\|^2 + n^2(\boldsymbol{l}) = 1$ and the w-component $C(\|\boldsymbol{l}\|_2)$ is given by

$$C(\|\boldsymbol{l}\|_2) \equiv \mathrm{e}^{\mathrm{i}2\pi w\left(1-\sqrt{1-\|\boldsymbol{l}\|^2}\right)} .$$

- Various assumptions are often made regarding the size of the field-of-view (FoV):
  - Small-field with $\|\boldsymbol{l}\|^2\, w \ll 1 \quad \Rightarrow \quad C(\|\boldsymbol{l}\|_2) \simeq 1$
  - Small-field with $\|\boldsymbol{l}\|^4\, w \ll 1 \quad \Rightarrow \quad C(\|\boldsymbol{l}\|_2) \simeq \mathrm{e}^{\mathrm{i}\pi w \|\boldsymbol{l}\|^2}$
  - Wide-field $\quad \Rightarrow \quad C(\|\boldsymbol{l}\|_2) = \mathrm{e}^{\mathrm{i}2\pi w\left(1-\sqrt{1-\|\boldsymbol{l}\|^2}\right)}$

- Interferometric imaging: recover an image from noisy and incomplete Fourier measurements.

Radio interferometric inverse problem

- Consider the ill-posed inverse problem of radio interferometric imaging:

$$\boldsymbol{y} = \Phi \boldsymbol{x} + \boldsymbol{n} ,$$

where $\boldsymbol{y}$ are the measured visibilities, Φ is the linear measurement operator, $\boldsymbol{x}$ is the underlying image and $\boldsymbol{n}$ is instrumental noise.

- Measurement operator $\Phi = \mathbf{M}\,\mathbf{F}\,\mathbf{C}\,\mathbf{A}$ may incorporate:
  - primary beam $\mathbf{A}$ of the telescope;
  - w-component modulation $\mathbf{C}$ (responsible for the spread spectrum phenomenon);
  - Fourier transform $\mathbf{F}$;
  - masking $\mathbf{M}$ which encodes the incomplete measurements taken by the interferometer.

Radio interferometric inverse problem

- Consider the ill-posed inverse problem of radio interferometric imaging:

$$\boxed{\boldsymbol{y} = \Phi \boldsymbol{x} + \boldsymbol{n}} \,,$$

  where $\boldsymbol{y}$ are the measured visibilities, Φ is the linear measurement operator, $\boldsymbol{x}$ is the underlying image and $\boldsymbol{n}$ is instrumental noise.

- Measurement operator $\boxed{\Phi = \mathbf{M}\,\mathbf{F}\,\mathbf{C}\,\mathbf{A}}$ may incorporate:
  - primary beam $\mathbf{A}$ of the telescope;
  - w-component modulation $\mathbf{C}$ (responsible for the spread spectrum phenomenon);
  - Fourier transform $\mathbf{F}$;
  - masking $\mathbf{M}$ which encodes the incomplete measurements taken by the interferometer.

Outline

Compressive sensing (CS)

- **"Nothing short of revolutionary."**
  – National Science Foundation
- Developed by Emmanuel Candes and David Donoho (and others).
- Next evolution of wavelet analysis.
- The mystery of JPEG compression (discrete cosine transform; wavelet transform).
- Acquisition versus imaging.

Compressive sensing (CS)

- **"Nothing short of revolutionary."**
  – National Science Foundation
- Developed by Emmanuel Candes and David Donoho (and others).
- Next evolution of wavelet analysis.
- The mystery of JPEG compression (discrete cosine transform; wavelet transform).
- Acquisition versus imaging.

(a) Emmanuel Candes

(b) David Donoho

Compressive sensing (CS)

- **"Nothing short of revolutionary."**
  – National Science Foundation
- Developed by Emmanuel Candes and David Donoho (and others).
- Next evolution of wavelet analysis.
- The mystery of JPEG compression (discrete cosine transform; wavelet transform).
- Acquisition versus imaging.

(a) Emmanuel Candes

(b) David Donoho

An introduction to compressive sensing

- Linear operator (linear algebra) representation of signal decomposition:

$$x(t) = \sum_i \alpha_i \Psi_i(t) \quad \rightarrow \quad \boldsymbol{x} = \sum_i \Psi_i \alpha_i = \begin{pmatrix} | \\ \Psi_0 \\ | \end{pmatrix} \alpha_0 + \begin{pmatrix} | \\ \Psi_1 \\ | \end{pmatrix} \alpha_1 + \cdots \quad \rightarrow \quad \boxed{\boldsymbol{x} = \Psi \boldsymbol{\alpha}}$$

- Linear operator (linear algebra) representation of measurement:

$$y_i = \langle x, \Phi_j \rangle \quad \rightarrow \quad \boldsymbol{y} = \begin{pmatrix} - \ \Phi_0 \ - \\ - \ \Phi_1 \ - \\ \vdots \end{pmatrix} \boldsymbol{x} \quad \rightarrow \quad \boxed{\boldsymbol{y} = \Phi \boldsymbol{x}}$$

- Putting it together: $\boxed{\boldsymbol{y} = \Phi \boldsymbol{x} = \Phi \Psi \boldsymbol{\alpha}}$

An introduction to compressive sensing

- Linear operator (linear algebra) representation of signal decomposition:

$$x(t) = \sum_i \alpha_i \Psi_i(t) \quad \rightarrow \quad \boldsymbol{x} = \sum_i \Psi_i \alpha_i = \begin{pmatrix} | \\ \Psi_0 \\ | \end{pmatrix} \alpha_0 + \begin{pmatrix} | \\ \Psi_1 \\ | \end{pmatrix} \alpha_1 + \cdots \quad \rightarrow \quad \boxed{\boldsymbol{x} = \Psi \boldsymbol{\alpha}}$$

- Linear operator (linear algebra) representation of measurement:

$$y_i = \langle x, \Phi_j \rangle \quad \rightarrow \quad \boldsymbol{y} = \begin{pmatrix} - & \Phi_0 & - \\ - & \Phi_1 & - \\ & \vdots & \end{pmatrix} \boldsymbol{x} \quad \rightarrow \quad \boxed{\boldsymbol{y} = \Phi \boldsymbol{x}}$$

- Putting it together: $\boxed{\boldsymbol{y} = \Phi \boldsymbol{x} = \Phi \Psi \boldsymbol{\alpha}}$

An introduction to compressive sensing

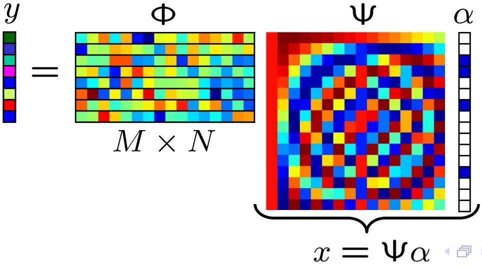
- Linear operator (linear algebra) representation of signal decomposition:

$$x(t) = \sum_i \alpha_i \Psi_i(t) \quad \rightarrow \quad \boldsymbol{x} = \sum_i \Psi_i \alpha_i = \begin{pmatrix} | \\ \Psi_0 \\ | \end{pmatrix} \alpha_0 + \begin{pmatrix} | \\ \Psi_1 \\ | \end{pmatrix} \alpha_1 + \cdots \quad \rightarrow \quad \boxed{\boldsymbol{x} = \Psi \boldsymbol{\alpha}}$$

- Linear operator (linear algebra) representation of measurement:

$$y_i = \langle x, \Phi_j \rangle \quad \rightarrow \quad \boldsymbol{y} = \begin{pmatrix} - & \Phi_0 & - \\ - & \Phi_1 & - \\ & \vdots & \end{pmatrix} \boldsymbol{x} \quad \rightarrow \quad \boxed{\boldsymbol{y} = \Phi \boldsymbol{x}}$$

- Putting it together: $\boxed{\boldsymbol{y} = \Phi \boldsymbol{x} = \Phi \Psi \boldsymbol{\alpha}}$



An introduction to compressive sensing

- Ill-posed inverse problem:

$$y = \Phi x + n = \Phi \Psi \alpha + n .$$

- Solve by imposing a regularising prior that the signal to be recovered is sparse in Ψ, *i.e.* solve the following ℓ_0 optimisation problem:

$$\alpha^\star = \arg\min_{\alpha} \|\alpha\|_0 \ \text{ such that } \ \|y - \Phi\Psi\alpha\|_2 \leq \epsilon \, ,$$

where the signal is synthesising by $x^\star = \Psi \alpha^\star$.

- Recall norms given by:

$$\|\alpha\|_0 = \text{no. non-zero elements} \qquad \|\alpha\|_1 = \sum_i |\alpha_i| \qquad \|\alpha\|_2 = \Big(\sum_i |\alpha_i|^2\Big)^{1/2}$$

- Solving this problem is difficult (combinatorial).
- Instead, solve the ℓ_1 optimisation problem (convex):

$$\alpha^\star = \arg\min_{\alpha} \|\alpha\|_1 \ \text{ such that } \ \|y - \Phi\Psi\alpha\|_2 \leq \epsilon \, .$$

An introduction to compressive sensing

- Ill-posed inverse problem:

$$\boldsymbol{y} = \Phi \boldsymbol{x} + \boldsymbol{n} = \Phi \Psi \boldsymbol{\alpha} + \boldsymbol{n} .$$

- Solve by imposing a regularising prior that the signal to be recovered is sparse in Ψ, *i.e.* solve the following ℓ_0 optimisation problem:

$$\boldsymbol{\alpha}^\star = \arg\min_{\boldsymbol{\alpha}} \|\alpha\|_0 \;\text{ such that }\; \|\boldsymbol{y} - \Phi\Psi\boldsymbol{\alpha}\|_2 \leq \epsilon ,$$

where the signal is synthesising by $\boldsymbol{x}^\star = \Psi \boldsymbol{\alpha}^\star$.

- Recall norms given by:

$$\|\alpha\|_0 = \text{no. non-zero elements} \qquad \|\alpha\|_1 = \sum_i |\alpha_i| \qquad \|\alpha\|_2 = \Big(\sum_i |\alpha_i|^2\Big)^{1/2}$$

- Solving this problem is difficult (combinatorial).
- Instead, solve the ℓ_1 optimisation problem (convex):

$$\boldsymbol{\alpha}^\star = \arg\min_{\boldsymbol{\alpha}} \|\alpha\|_1 \;\text{ such that }\; \|\boldsymbol{y} - \Phi\Psi\boldsymbol{\alpha}\|_2 \leq \epsilon .$$

An introduction to compressive sensing

- Ill-posed inverse problem:

$$\boldsymbol{y} = \Phi \boldsymbol{x} + \boldsymbol{n} = \Phi \Psi \boldsymbol{\alpha} + \boldsymbol{n}.$$

- Solve by imposing a regularising prior that the signal to be recovered is sparse in Ψ, *i.e.* solve the following ℓ_0 optimisation problem:

$$\boldsymbol{\alpha}^\star = \arg\min_{\boldsymbol{\alpha}} \|\alpha\|_0 \ \text{such that} \ \|\boldsymbol{y} - \Phi\Psi\boldsymbol{\alpha}\|_2 \leq \epsilon \, ,$$

  where the signal is synthesising by $\boldsymbol{x}^\star = \Psi\boldsymbol{\alpha}^\star$.

- Recall norms given by:

$$\|\alpha\|_0 = \text{no. non-zero elements} \qquad \|\alpha\|_1 = \sum_i |\alpha_i| \qquad \|\alpha\|_2 = \Big(\sum_i |\alpha_i|^2\Big)^{1/2}$$

- Solving this problem is difficult (combinatorial).

- Instead, solve the ℓ_1 optimisation problem (convex):

$$\boldsymbol{\alpha}^\star = \arg\min_{\boldsymbol{\alpha}} \|\alpha\|_1 \ \text{such that} \ \|\boldsymbol{y} - \Phi\Psi\boldsymbol{\alpha}\|_2 \leq \epsilon \, .$$

An introduction to compressive sensing

- Ill-posed inverse problem:

$$\boldsymbol{y} = \Phi \boldsymbol{x} + \boldsymbol{n} = \Phi \Psi \boldsymbol{\alpha} + \boldsymbol{n}.$$

- Solve by imposing a regularising prior that the signal to be recovered is sparse in Ψ, *i.e.* solve the following ℓ_0 optimisation problem:

$$\boldsymbol{\alpha}^\star = \arg\min_{\boldsymbol{\alpha}} \|\alpha\|_0 \text{ such that } \|\boldsymbol{y} - \Phi\Psi\boldsymbol{\alpha}\|_2 \leq \epsilon \,,$$

  where the signal is synthesising by $\boldsymbol{x}^\star = \Psi\boldsymbol{\alpha}^\star$.

- Recall norms given by:

$$\|\alpha\|_0 = \text{no. non-zero elements} \qquad \|\alpha\|_1 = \sum_i |\alpha_i| \qquad \|\alpha\|_2 = \Big(\sum_i |\alpha_i|^2\Big)^{1/2}$$

- Solving this problem is difficult (combinatorial).
- Instead, solve the ℓ_1 optimisation problem (convex):

$$\boldsymbol{\alpha}^\star = \arg\min_{\boldsymbol{\alpha}} \|\alpha\|_1 \text{ such that } \|\boldsymbol{y} - \Phi\Psi\boldsymbol{\alpha}\|_2 \leq \epsilon \,.$$

An introduction to compressive sensing

- The solutions of the ℓ_0 and ℓ_1 problems are often the same.
- Restricted isometry property (RIP):

$$(1-\delta_K)\|\boldsymbol{\alpha}\|_2^2 \leq \|\Theta\boldsymbol{\alpha}\|_2^2 \leq (1+\delta_K)\|\boldsymbol{\alpha}\|_2^2 \ ,$$

for K-sparse $\boldsymbol{\alpha}$, where $\Theta = \Phi\Psi$.

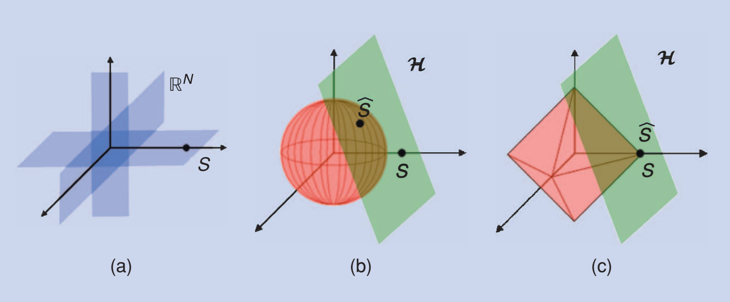


Figure: Geometry of (a) ℓ_0 (b) ℓ_2 and (c) ℓ_1 problems. [Credit: Baraniuk (2007)]

An introduction to compressive sensing

- In the absence of noise, compressed sensing is **exact!**
- Number of measurements required to achieve exact reconstruction is given by

$$M \geq c\mu^2 K \log N \ ,$$

  where K is the sparsity and N the dimensionality.
- The coherence between the measurement and sparsity basis is given by

$$\mu = \sqrt{N} \max_{i,j} |\langle \Psi_i, \Phi_j \rangle| \ .$$

- Robust to noise.
- Many new developments (*e.g.* analysis vs synthesis, cosparsity, structured sparsity) and new applications.

An introduction to compressive sensing

- In the absence of noise, compressed sensing is **exact!**
- Number of measurements required to achieve exact reconstruction is given by

$$M \geq c\mu^2 K \log N \ ,$$

where K is the sparsity and N the dimensionality.

- The coherence between the measurement and sparsity basis is given by

$$\mu = \sqrt{N} \max_{i,j} |\langle \Psi_i, \Phi_j \rangle| \ .$$

- Robust to noise.
- Many new developments (*e.g.* analysis vs synthesis, cosparsity, structured sparsity) and new applications.

An introduction to compressive sensing

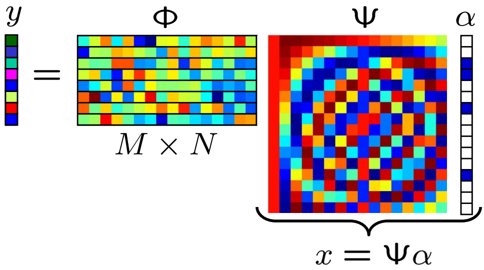
- In the absence of noise, compressed sensing is **exact!**
- Number of measurements required to achieve exact reconstruction is given by

$$M \geq c\mu^2 K \log N \,,$$

where K is the sparsity and N the dimensionality.

- The coherence between the measurement and sparsity basis is given by

$$\mu = \sqrt{N} \max_{i,j} |\langle \Psi_i, \Phi_j \rangle| \,.$$



- Robust to noise.
- Many new developments (*e.g.* analysis vs synthesis, cosparsity, structured sparsity) and new applications.

An introduction to compressive sensing

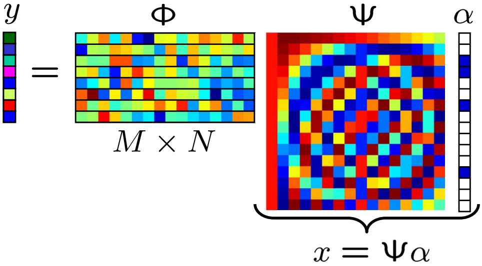
- In the absence of noise, compressed sensing is **exact!**
- Number of measurements required to achieve exact reconstruction is given by

$$M \geq c\mu^2 K \log N \ ,$$

where K is the sparsity and N the dimensionality.

- The coherence between the measurement and sparsity basis is given by

$$\mu = \sqrt{N} \max_{i,j} |\langle \Psi_i, \Phi_j \rangle| \ .$$



- Robust to noise.
- Many new developments (*e.g.* analysis vs synthesis, cosparsity, structured sparsity) and new applications.

Outline

Interferometric imaging with compressed sensing

- Solve the interferometric imaging problem

$$\boldsymbol{y} = \Phi \boldsymbol{x} + \boldsymbol{n} \quad \text{with} \quad \Phi = \mathbf{M}\,\mathbf{F}\,\mathbf{C}\,\mathbf{A}\,,$$

  by applying a prior on sparsity of the signal in a sparsifying dictionary Ψ.

- Solve basis pursuit denoising problem

$$\boldsymbol{\alpha}^\star = \arg\min_{\boldsymbol{\alpha}} \|\alpha\|_1 \ \text{such that} \ \|\boldsymbol{y} - \Phi\Psi\boldsymbol{\alpha}\|_2 \leq \epsilon\,,$$

  where the image is synthesised by $\boldsymbol{x}^\star = \Psi\boldsymbol{\alpha}^\star$.

- Various choices for sparsifying dictionary Ψ, *e.g.* Dirac basis, Daubechies wavelets.
- Analysis versus synthesis problems, *e.g.* SARA algorithm.
- Recall the potential trade-off between sparsity and coherence.

Interferometric imaging with compressed sensing

- Solve the interferometric imaging problem

$$\boldsymbol{y} = \Phi \boldsymbol{x} + \boldsymbol{n} \quad \text{with} \quad \Phi = \mathbf{M}\,\mathbf{F}\,\mathbf{C}\,\mathbf{A}\,,$$

  by applying a prior on sparsity of the signal in a sparsifying dictionary Ψ.

- Solve basis pursuit denoising problem

$$\boldsymbol{\alpha}^\star = \underset{\boldsymbol{\alpha}}{\arg\min} \|\alpha\|_1 \ \text{such that} \ \|\boldsymbol{y} - \Phi\Psi\boldsymbol{\alpha}\|_2 \leq \epsilon\,,$$

  where the image is synthesised by $\boldsymbol{x}^\star = \Psi\boldsymbol{\alpha}^\star$.

- Various choices for sparsifying dictionary Ψ, *e.g.* Dirac basis, Daubechies wavelets.
- Analysis versus synthesis problems, *e.g.* SARA algorithm.
- Recall the potential trade-off between sparsity and coherence.

SARA for radio interferometric imaging

- Sparsity averaging reweighted analysis (SARA) for RI imaging (Carrillo, JDM & Wiaux 2012)
- Consider a dictionary composed of a concatenation of orthonormal bases, i.e.

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

  thus $\Psi \in \mathbb{R}^{N \times D}$ with $D = qN$.

- We consider the following bases:
  - Dirac, *i.e.* pixel basis
  - Haar wavelets (promotes gradient sparsity)
  - Daubechies wavelet bases two to eight.

  $\Rightarrow$ concatenation of 9 bases

- Promote average sparsity by solving the reweighted ℓ_1 analysis problem:

$$\min_{\bar{x} \in \mathbb{R}^N} \| W \Psi^T \bar{x} \|_1 \quad \text{subject to} \quad \| y - \Phi \bar{x} \|_2 \leq \epsilon \quad \text{and} \quad \bar{x} \geq 0 \,,$$

  where $W \in \mathbb{R}^{D \times D}$ is a diagonal matrix with positive weights.

- Solve a sequence of reweighted ℓ_1 problems using the solution of the previous problem as the inverse weights $\rightarrow$ approximate the ℓ_0 problem.

SARA for radio interferometric imaging

- Sparsity averaging reweighted analysis (SARA) for RI imaging (Carrillo, JDM & Wiaux 2012)
- Consider a dictionary composed of a concatenation of orthonormal bases, i.e.

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

thus $\Psi \in \mathbb{R}^{N \times D}$ with $D = qN$.

- We consider the following bases:
  - Dirac, *i.e.* pixel basis
  - Haar wavelets (promotes gradient sparsity)
  - Daubechies wavelet bases two to eight.

$\Rightarrow$ concatenation of 9 bases

- Promote average sparsity by solving the reweighted ℓ_1 analysis problem:

$$\min_{\bar{\boldsymbol{x}} \in \mathbb{R}^N} \|W \Psi^T \bar{\boldsymbol{x}}\|_1 \quad \text{subject to} \quad \|\boldsymbol{y} - \Phi \bar{\boldsymbol{x}}\|_2 \le \epsilon \quad \text{and} \quad \bar{\boldsymbol{x}} \ge 0 \,,$$

where $W \in \mathbb{R}^{D \times D}$ is a diagonal matrix with positive weights.

- Solve a sequence of reweighted ℓ_1 problems using the solution of the previous problem as the inverse weights $\rightarrow$ approximate the ℓ_0 problem.

SARA for radio interferometric imaging

- Sparsity averaging reweighted analysis (SARA) for RI imaging (Carrillo, JDM & Wiaux 2012)
- Consider a dictionary composed of a concatenation of orthonormal bases, i.e.

$$\Psi = \frac{1}{\sqrt{q}}[\Psi_1, \Psi_2, \dots, \Psi_q],$$

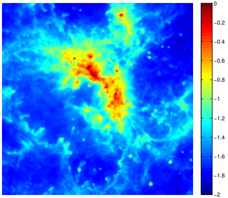
  thus $\Psi \in \mathbb{R}^{N \times D}$ with $D = qN$.
- We consider the following bases:
  - Dirac, *i.e.* pixel basis
  - Haar wavelets (promotes gradient sparsity)
  - Daubechies wavelet bases two to eight.

  $\Rightarrow$ concatenation of 9 bases
- Promote average sparsity by solving the reweighted ℓ_1 analysis problem:

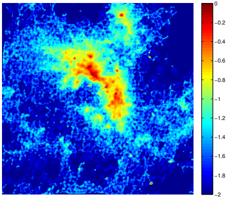
$$\min_{\bar{\boldsymbol{x}} \in \mathbb{R}^N} \|W\Psi^T \bar{\boldsymbol{x}}\|_1 \quad \text{subject to} \quad \|\boldsymbol{y} - \Phi\bar{\boldsymbol{x}}\|_2 \leq \epsilon \quad \text{and} \quad \bar{\boldsymbol{x}} \geq 0 \,,$$

  where $W \in \mathbb{R}^{D \times D}$ is a diagonal matrix with positive weights.
- Solve a sequence of reweighted ℓ_1 problems using the solution of the previous problem as the inverse weights $\rightarrow$ approximate the ℓ_0 problem.

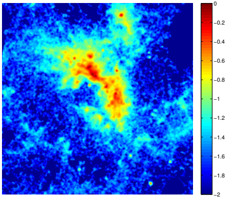
SARA for radio interferometric imaging



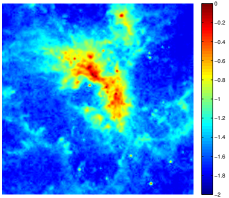
(a) Original



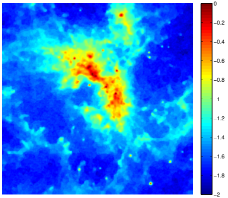
(b) BP (SNR=16.67 dB)



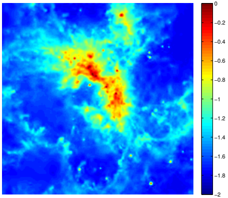
(c) IUWT (SNR=17.87 dB)



(d) BPDb8 (SNR=24.53 dB)



(e) TV (SNR=26.47 dB)



(f) SARA (SNR=29.08 dB)

Figure: Reconstruction example of 30Dor from 30% of visibilities.

SARA for radio interferometric imaging

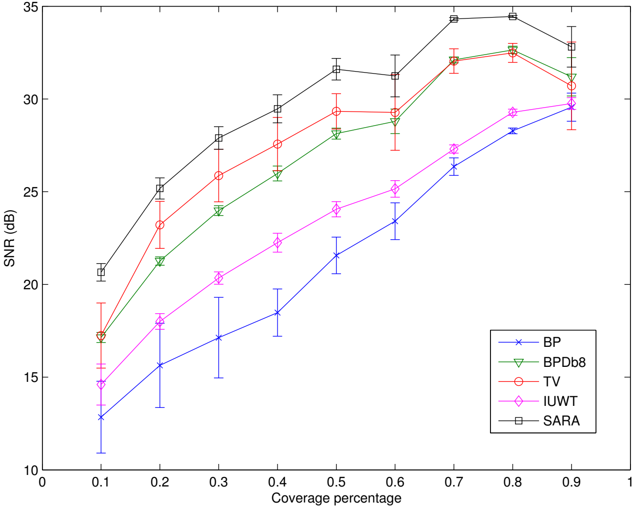


Figure: Reconstruction fidelity vs visibility coverage for 30Dor.

Outline

Review of the spread spectrum phenomenon

- The w-component modulation gives rise to the spread spectrum phenomenon first considered by Wiaux *et al.* (2009b).
- The w-component operator $\mathbf{C}$ has elements defined by

$$C(l, m) \equiv e^{i2\pi w\left(1-\sqrt{1-l^2-m^2}\right)} \simeq e^{i\pi w \|\boldsymbol{l}\|^2} \quad \text{for} \quad \|\boldsymbol{l}\|^4 \, w \ll 1 \, ,$$

giving rise to to a linear chirp.

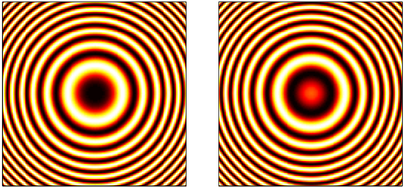
- For the (essentially) Fourier measurements of interferometric telescopes the coherence is the maximum modulus of the Fourier transform of the atoms of the sparsifying dictionary.
- Modulation by the chirp spreads the spectrum of the atoms of the sparsifying dictionary.
- Consequently, spreading the spectrum increases the incoherence between the sensing and sparsity bases, thus improving reconstruction fidelity.

Review of the spread spectrum phenomenon

- The w-component modulation gives rise to the spread spectrum phenomenon first considered by Wiaux *et al.* (2009b).
- The w-component operator $\mathbf{C}$ has elements defined by

$$C(l,m) \equiv e^{i2\pi w\left(1-\sqrt{1-l^2-m^2}\right)} \simeq e^{i\pi w\|\boldsymbol{l}\|^2} \quad \text{for} \quad \|\boldsymbol{l}\|^4\, w \ll 1 ,$$

giving rise to to a linear chirp.



(a) Real part (b) Imaginary part

Figure: Chirp modulation.

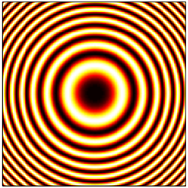
- For the (essentially) Fourier measurements of interferometric telescopes the coherence is the maximum modulus of the Fourier transform of the atoms of the sparsifying dictionary.
- Modulation by the chirp spreads the spectrum of the atoms of the sparsifying dictionary.
- Consequently, spreading the spectrum increases the incoherence between the sensing and sparsity bases, thus improving reconstruction fidelity.

Review of the spread spectrum phenomenon

- The w-component modulation gives rise to the spread spectrum phenomenon first considered by Wiaux *et al.* (2009b).
- The w-component operator $\mathbf{C}$ has elements defined by

$$C(l,m) \equiv e^{i2\pi w\left(1-\sqrt{1-l^2-m^2}\right)} \simeq e^{i\pi w\|\boldsymbol{l}\|^2} \quad \text{for} \quad \|\boldsymbol{l}\|^4\, w \ll 1\,,$$

giving rise to to a linear chirp.


(a) Real part

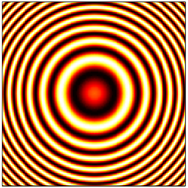

(b) Imaginary part

Figure: Chirp modulation.

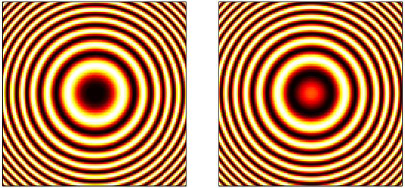
- For the (essentially) Fourier measurements of interferometric telescopes the coherence is the maximum modulus of the Fourier transform of the atoms of the sparsifying dictionary.
- Modulation by the chirp spreads the spectrum of the atoms of the sparsifying dictionary.
- Consequently, spreading the spectrum increases the incoherence between the sensing and sparsity bases, thus improving reconstruction fidelity.

Review of the spread spectrum phenomenon

- The w-component modulation gives rise to the spread spectrum phenomenon first considered by Wiaux *et al.* (2009b).
- The w-component operator $\mathbf{C}$ has elements defined by

$$C(l,m) \equiv e^{i2\pi w\left(1-\sqrt{1-l^2-m^2}\right)} \simeq e^{i\pi w \|\boldsymbol{l}\|^2} \quad \text{for} \quad \|\boldsymbol{l}\|^4 \, w \ll 1 \, ,$$

giving rise to to a linear chirp.



(a) Real part (b) Imaginary part

Figure: Chirp modulation.

- For the (essentially) Fourier measurements of interferometric telescopes the coherence is the maximum modulus of the Fourier transform of the atoms of the sparsifying dictionary.
- Modulation by the chirp spreads the spectrum of the atoms of the sparsifying dictionary.
- Consequently, spreading the spectrum increases the incoherence between the sensing and sparsity bases, thus improving reconstruction fidelity.

Spread spectrum phenomenon for varying w

- Improved reconstruction fidelity of the spread spectrum phenomenon demonstrated with simulations by Wiaux *et al.* (2009b).
- However, previous analysis was restricted to fixed w for simplicity.
- Recently, we have examined the spread spectrum phenomenon for varying w.
- Work of Laura Wolz, in collaboration with JDM, Filipe Abdalla, Rafael Carrillo and Yves Wiaux.

- Apply the w-projection algorithm (Cornwell *et al.* 2008) to shift the chirp modulation through the Fourier transform:

$$\boxed{\Phi = \mathbf{M}\,\mathbf{F}\,\mathbf{C}\,\mathbf{A} \quad \Rightarrow \quad \Phi = \mathbf{M}\,\tilde{\mathbf{C}}\,\mathbf{F}\,\mathbf{A}}\,.$$

- Consider different w for each (u, v) and threshold each Fourier transformed chirp (each row of $\tilde{\mathbf{C}}$) to approximate $\tilde{\mathbf{C}}$ accurately by a sparse matrix.

Spread spectrum phenomenon for varying w

- Improved reconstruction fidelity of the spread spectrum phenomenon demonstrated with simulations by Wiaux *et al.* (2009b).
- However, previous analysis was restricted to fixed w for simplicity.
- Recently, we have examined the spread spectrum phenomenon for varying w.
- Work of Laura Wolz, in collaboration with JDM, Filipe Abdalla, Rafael Carrillo and Yves Wiaux.

- Apply the w-projection algorithm (Cornwell *et al.* 2008) to shift the chirp modulation through the Fourier transform:

$$\boxed{\Phi = \mathbf{M}\,\mathbf{F}\,\mathbf{C}\,\mathbf{A} \quad \Rightarrow \quad \Phi = \mathbf{M}\,\tilde{\mathbf{C}}\,\mathbf{F}\,\mathbf{A}}\ .$$

- Consider different w for each (u, v) and threshold each Fourier transformed chirp (each row of $\tilde{\mathbf{C}}$) to approximate $\tilde{\mathbf{C}}$ accurately by a sparse matrix.

Spread spectrum phenomenon for varying w

- Improved reconstruction fidelity of the spread spectrum phenomenon demonstrated with simulations by Wiaux *et al.* (2009b).
- However, previous analysis was restricted to fixed w for simplicity.
- Recently, we have examined the spread spectrum phenomenon for varying w.
- Work of Laura Wolz, in collaboration with JDM, Filipe Abdalla, Rafael Carrillo and Yves Wiaux.

- Apply the w-projection algorithm (Cornwell *et al.* 2008) to shift the chirp modulation through the Fourier transform:

$$\boxed{\Phi = \mathbf{M}\,\mathbf{F}\,\mathbf{C}\,\mathbf{A} \quad \Rightarrow \quad \Phi = \mathbf{M}\,\tilde{\mathbf{C}}\,\mathbf{F}\,\mathbf{A}}\,.$$

- Consider different w for each (u, v) and threshold each Fourier transformed chirp (each row of $\tilde{\mathbf{C}}$) to approximate $\tilde{\mathbf{C}}$ accurately by a sparse matrix.

Spread spectrum phenomenon for varying w

- Perform simulations to assess the effectiveness of the spread spectrum phenomenon in the presence of varying w.
- Consider idealised simulations with uniformly random visibility sampling.

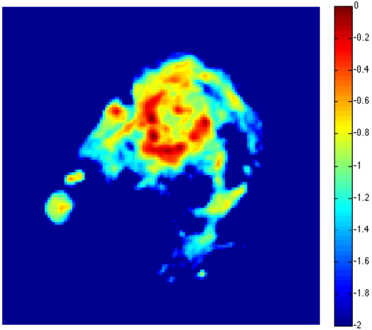
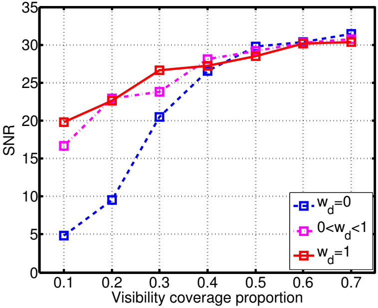


Figure: M31 (ground truth).

Spread spectrum phenomenon for varying w



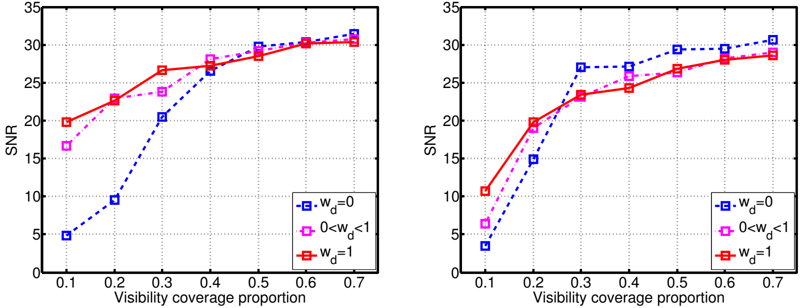
(a) Daubechies 8 (Db8) wavelets

Figure: Reconstruction fidelity.

The improvement in reconstruction fidelity due to the spread spectrum phenomenon for varying w is almost as large as the case of constant maximum w!

- As expected, for the case where coherence is already optimal, there is little improvement.

Spread spectrum phenomenon for varying w



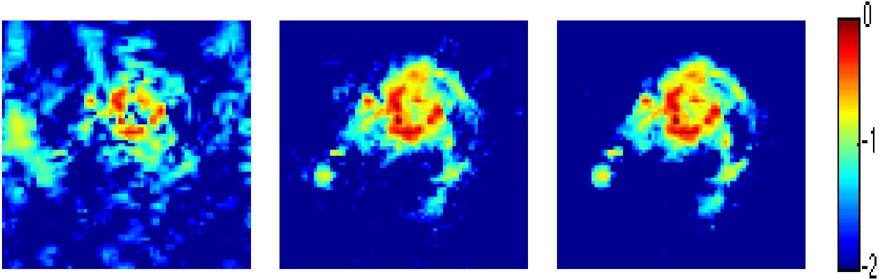
(a) Daubechies 8 (Db8) wavelets (b) Dirac basis

Figure: Reconstruction fidelity.

The improvement in reconstruction fidelity due to the spread spectrum phenomenon for varying w is almost as large as the case of constant maximum w!

- As expected, for the case where coherence is already optimal, there is little improvement.

Spread spectrum phenomenon for varying w



(a) $w_{\mathrm{d}} = 0 \rightarrow$ SNR= 4.8dB (b) $w_{\mathrm{d}} \sim \mathcal{U}(0,1) \rightarrow$ SNR= 16.7dB (c) $w_{\mathrm{d}} = 1 \rightarrow$ SNR= 19.8dB

Figure: Reconstructed images for 10% coverage.

Outline

Supporting continuous visibilities

- Ideally we would like to model the continuous Fourier transform operator

$$\Phi = \mathbf{F}^{c} .$$

- But this is slow!
- We have incorporated gridding into our CS interferometric imaging framework.
- Work of Rafael Carrillo, in collaboration with Yves Wiaux and JDM.
- Model with the measurement operator

$$\boxed{\Phi = \mathbf{G}\,\mathbf{F}\,\mathbf{Z}\,\mathbf{D}} ,$$

where we incorporate:

  - convolutional gridding operator $\mathbf{G}$;
  - fast Fourier transform $\mathbf{F}$;
  - zero-padding operator $\mathbf{Z}$ to upsample the discrete visibility space;
  - normalisation operator $\mathbf{D}$ to undo the convolution gridding (reciprocal of the inverse Fourier transform of the gridding kernel).

Supporting continuous visibilities

- Ideally we would like to model the continuous Fourier transform operator

$$\Phi = \mathbf{F}^{\mathrm{c}} .$$

- But this is slow!
- We have incorporated gridding into our CS interferometric imaging framework.
- Work of Rafael Carrillo, in collaboration with Yves Wiaux and JDM.
- Model with the measurement operator

$$\boxed{\Phi = \mathbf{G}\,\mathbf{F}\,\mathbf{Z}\,\mathbf{D}} ,$$

  where we incorporate:
  - convolutional gridding operator $\mathbf{G}$;
  - fast Fourier transform $\mathbf{F}$;
  - zero-padding operator $\mathbf{Z}$ to upsample the discrete visibility space;
  - normalisation operator $\mathbf{D}$ to undo the convolution gridding (reciprocal of the inverse Fourier transform of the gridding kernel).

Supporting continuous visibilities

- Ideally we would like to model the continuous Fourier transform operator

$$\Phi = \mathbf{F}^{\mathrm{c}} \, .$$

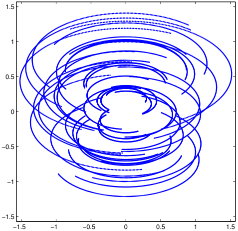
- But this is slow!
- We have incorporated gridding into our CS interferometric imaging framework.
- Work of Rafael Carrillo, in collaboration with Yves Wiaux and JDM.
- Model with the measurement operator

$$\boxed{\Phi = \mathbf{G}\,\mathbf{F}\,\mathbf{Z}\,\mathbf{D}} \, ,$$

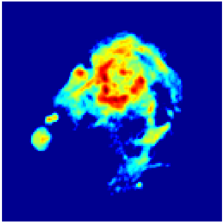
where we incorporate:

  - convolutional gridding operator $\mathbf{G}$;
  - fast Fourier transform $\mathbf{F}$;
  - zero-padding operator $\mathbf{Z}$ to upsample the discrete visibility space;
  - normalisation operator $\mathbf{D}$ to undo the convolution gridding (reciprocal of the inverse Fourier transform of the gridding kernel).

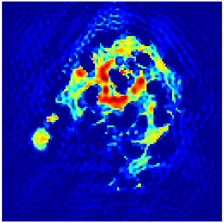
Reconstruction with continuous visibilities



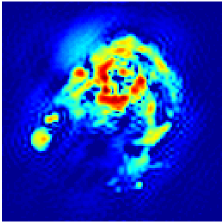
(a) Coverage



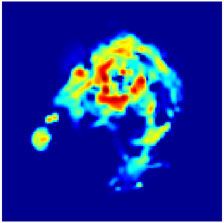
(b) M31 (ground truth)



(c) Dirac basis → SNR= 8.2dB



(d) Db8 wavelets → SNR= 11.1dB



(e) SARA → SNR= 13.4dB

Figure: Reconstructed images from continuous visibilities.

Summary

- Effectiveness of compressive sensing for radio interferometric imaging demonstrated already (Wiaux *et al.* 2009a, Wiaux *et al.*2009b, Wiaux *et al.* 2009c, McEwen & Wiaux 2011, Carrillo *et al.* 2012).
- Provide improvements in reconstruction fidelity, flexibility and computation time.
- Important to take these methods to the realistic setting so that their advantages can be realised on observations made by real radio interferometric telescopes.

- Taken first steps toward more realistic setting.
- Studied the spread spectrum phenomenon for varying w.
- The improvement in reconstruction fidelity due to the spread spectrum phenomenon for varying w is almost as large as the case of constant maximum w!
- Incorporated a gridding operator into our framework to support continuous visibilities.

Summary

- Effectiveness of compressive sensing for radio interferometric imaging demonstrated already (Wiaux *et al.* 2009a, Wiaux *et al.*2009b, Wiaux *et al.* 2009c, McEwen & Wiaux 2011, Carrillo *et al.* 2012).
- Provide improvements in reconstruction fidelity, flexibility and computation time.
- Important to take these methods to the realistic setting so that their advantages can be realised on observations made by real radio interferometric telescopes.

- Taken first steps toward more realistic setting.
- Studied the spread spectrum phenomenon for varying w.
- The improvement in reconstruction fidelity due to the spread spectrum phenomenon for varying w is almost as large as the case of constant maximum w!
- Incorporated a gridding operator into our framework to support continuous visibilities.

Summary

- Effectiveness of compressive sensing for radio interferometric imaging demonstrated already (Wiaux *et al.* 2009a, Wiaux *et al.*2009b, Wiaux *et al.* 2009c, McEwen & Wiaux 2011, Carrillo *et al.* 2012).
- Provide improvements in reconstruction fidelity, flexibility and computation time.
- Important to take these methods to the realistic setting so that their advantages can be realised on observations made by real radio interferometric telescopes.

- Taken first steps toward more realistic setting.
- Studied the spread spectrum phenomenon for varying w.
- The improvement in reconstruction fidelity due to the spread spectrum phenomenon for varying w is almost as large as the case of constant maximum w!
- Incorporated a gridding operator into our framework to support continuous visibilities.

Summary

- Effectiveness of compressive sensing for radio interferometric imaging demonstrated already (Wiaux *et al.* 2009a, Wiaux *et al.*2009b, Wiaux *et al.* 2009c, McEwen & Wiaux 2011, Carrillo *et al.* 2012).
- Provide improvements in reconstruction fidelity, flexibility and computation time.
- Important to take these methods to the realistic setting so that their advantages can be realised on observations made by real radio interferometric telescopes.

- Taken first steps toward more realistic setting.
- Studied the spread spectrum phenomenon for varying w.
- The improvement in reconstruction fidelity due to the spread spectrum phenomenon for varying w is almost as large as the case of constant maximum w!
- Incorporated a gridding operator into our framework to support continuous visibilities.

Outlook

- BUT... so far we remain idealised.
- We (Rafael Carrillo, JDM and Yves Wiaux) are developing an optimised C code (`PURIFY`) to scale to the realistic setting.
- Preliminary tests indicate that this code provides in excess of an order of magnitude speed improvement and supports scaling to very large data-sets.