Radio interferometric imaging with compressive sensing

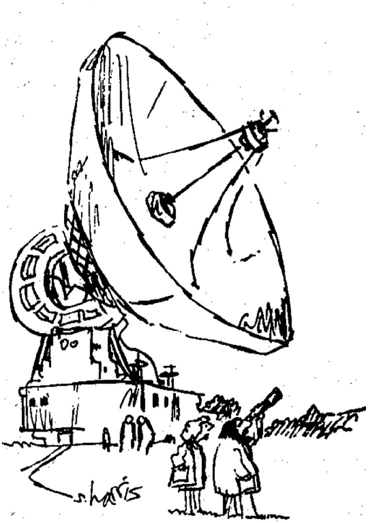
Jason McEwen
www.jasonmcewen.org
@jasonmcewen

Mullard Space Science Laboratory (MSSL)
University College London (UCL)

In collaboration with Laura Wolz, Filipe Abdalla, Rafael Carrillo & Yves Wiaux

Inverse Problems – from Theory to Application, Bristol, August 2013

Radio telescopes are big!



"Just checking."

Radio telescopes are big!

Radio interferometric telescopes

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.
- Broad range of science goals.
- New modelling and imaging techniques required to ensure the next-generation of interferometric telescopes reach their full potential.

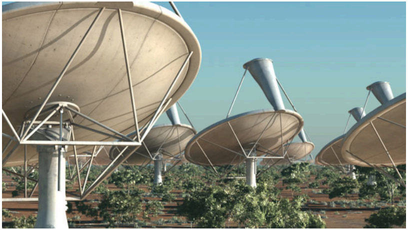


Figure: Artist impression of SKA dishes. [Credit: SKA Organisation]

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.
- Broad range of science goals.
- New modelling and imaging techniques required to ensure the next-generation of interferometric telescopes reach their full potential.

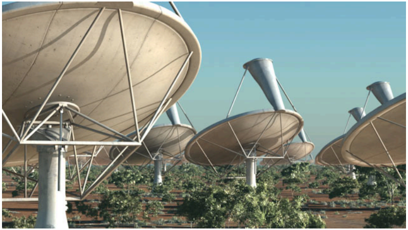
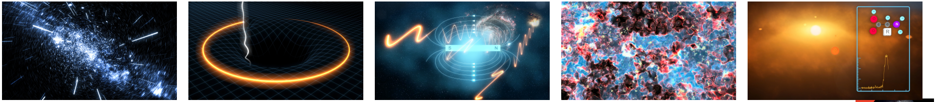


Figure: Artist impression of SKA dishes. [Credit: SKA Organisation]



(a) Dark-energy (b) GR (c) Cosmic magnetism (d) Epoch of reionization (e) Exoplanets

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

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.
- Broad range of science goals.
- New modelling and imaging techniques required to ensure the next-generation of interferometric telescopes reach their full potential.

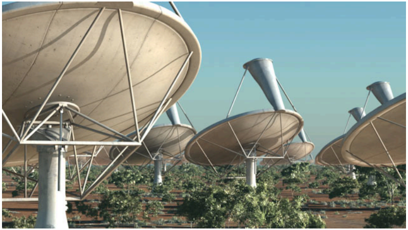
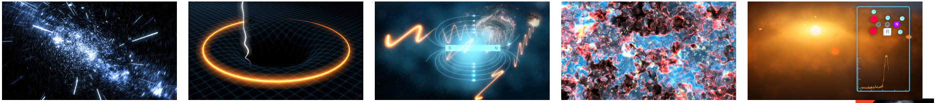


Figure: Artist impression of SKA dishes. [Credit: SKA Organisation]



(a) Dark-energy (b) GR (c) Cosmic magnetism (d) Epoch of reionization (e) Exoplanets

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

Outline

Outline

1. Inverse Problem
2. Interferometric Imaging with Compressive Sensing
3. Spread Spectrum Effect
4. Continuous Visibilities

Radio interferometric inverse problem

- The complex visibility measured by an interferometer is given by

$$\boxed{y(\boldsymbol{u}, w) = \int_{D^2} A(\boldsymbol{l})\, x(\boldsymbol{l})\, C(\|\boldsymbol{l}\|_2)\, \mathrm{e}^{-\mathrm{i}2\pi \boldsymbol{u}\cdot\boldsymbol{l}}\, \frac{\mathrm{d}^2\boldsymbol{l}}{n(\boldsymbol{l})}}\,,$$

visibilities

where the w-modulation $C(\|\boldsymbol{l}\|_2)$ is given by

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

w-modulation

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

Radio interferometric inverse problem

- The complex visibility measured by an interferometer is given by

$$\boxed{y(\boldsymbol{u}, w) = \int_{D^2} A(\boldsymbol{l})\, x(\boldsymbol{l})\, C(\|\boldsymbol{l}\|_2)\, \mathrm{e}^{-\mathrm{i}2\pi \boldsymbol{u}\cdot\boldsymbol{l}} \, \frac{\mathrm{d}^2\boldsymbol{l}}{n(\boldsymbol{l})}} \,,$$

visibilities

where the w-modulation $C(\|\boldsymbol{l}\|_2)$ is given by

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

w-modulation

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

Radio interferometric inverse problem

- The complex visibility measured by an interferometer is given by

$$\boxed{y(\boldsymbol{u}, w) = \int_{D^2} A(\boldsymbol{l})\, x(\boldsymbol{l})\, C(\|\boldsymbol{l}\|_2)\, \mathrm{e}^{-\mathrm{i}2\pi \boldsymbol{u}\cdot\boldsymbol{l}}\, \frac{\mathrm{d}^2\boldsymbol{l}}{n(\boldsymbol{l})}},$$

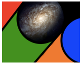
visibilities

where the w-modulation $C(\|\boldsymbol{l}\|_2)$ is given by

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

w-modulation

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



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-modulation modulation $\mathbf{C}$;
  - 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-modulation modulation $\mathbf{C}$;
  - 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-modulation modulation $\mathbf{C}$;
    - Fourier transform $\mathbf{F}$;
    - masking $\mathbf{M}$ which encodes the incomplete measurements taken by the interferometer.

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

Outline

1. Inverse Problem
2. Interferometric Imaging with Compressive Sensing
3. Spread Spectrum Effect
4. Continuous Visibilities

Interferometric imaging with compressed sensing

- Solve the interferometric imaging problem

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

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

- Basis pursuit (BP) denoising problem

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

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

- Total Variation (TV) denoising problem

$$\boxed{\boldsymbol{x}^\star = \underset{\boldsymbol{x}}{\arg\min} \|\boldsymbol{x}\|_{\text{TV}} \;\; \text{such that} \;\; \|\boldsymbol{y} - \Phi\boldsymbol{x}\|_2 \leq \epsilon \,.} \;\text{TVDN}$$

Interferometric imaging with compressed sensing

- Solve the interferometric imaging problem

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

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

- Basis pursuit (BP) denoising problem

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

BPDN

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

- Total Variation (TV) denoising problem

$$\boxed{\boldsymbol{x}^\star = \underset{\boldsymbol{x}}{\arg\min} \|\boldsymbol{x}\|_{\text{TV}} \;\text{ such that }\; \|\boldsymbol{y} - \Phi\boldsymbol{x}\|_2 \leq \epsilon \,.}$$

TVDN

Interferometric imaging with compressed sensing

- Solve the interferometric imaging problem

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

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

- Basis pursuit (BP) denoising problem

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

BPDN

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

- Total Variation (TV) denoising problem

$$\boxed{\boldsymbol{x}^\star = \underset{\boldsymbol{x}}{\arg\min} \|\boldsymbol{x}\|_{\text{TV}} \ \text{ such that } \ \|\boldsymbol{y} - \Phi\boldsymbol{x}\|_2 \leq \epsilon \,.}$$

TVDN

SARA for radio interferometric imaging

Algorithm

- Sparsity averaging reweighted analysis (SARA) for RI imaging (Carrillo, McEwen & 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:

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

  SARA

  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

Algorithm

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

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

  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:

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

  SARA

  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 → approximate the ℓ_0 problem.

SARA for radio interferometric imaging

Algorithm

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

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

  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:

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

  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

Algorithm

- Sparsity averaging reweighted analysis (SARA) for RI imaging (Carrillo, McEwen & Wiaux 2012)

- Consider a dictionary composed of a concatenation of orthonormal bases, i.e.

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

  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

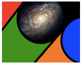
- Promote average sparsity by solving the reweighted ℓ_1 analysis problem:

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

SARA

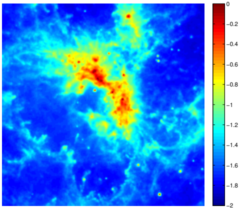
  where $W \in \mathbb{R}^{D \times D}$ is a diagonal matrix with positive weights.

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



SARA for radio interferometric imaging

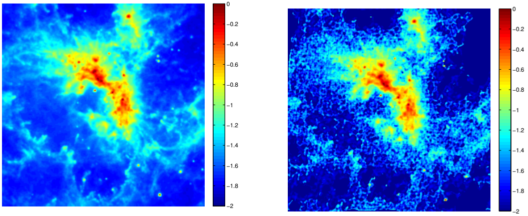
Results on simulations



(a) Original

SARA for radio interferometric imaging

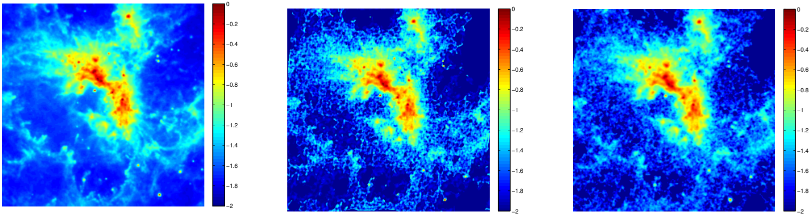
Results on simulations



(a) Original (b) BP (SNR=16.67 dB)

SARA for radio interferometric imaging

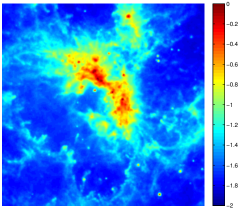
Results on simulations



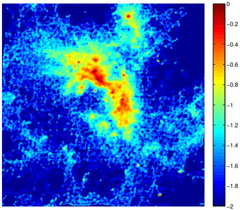
(a) Original (b) BP (SNR=16.67 dB) (c) IUWT (SNR=17.87 dB)

SARA for radio interferometric imaging

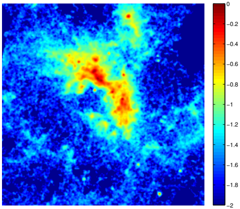
Results on simulations



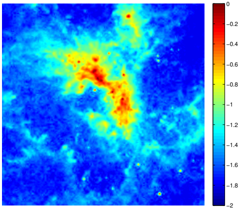
(a) Original



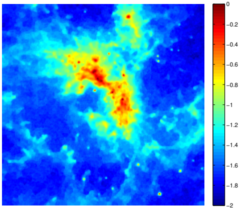
(b) BP (SNR=16.67 dB)



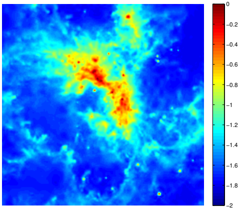
(c) IUWT (SNR=17.87 dB)



(d) BPDb8 (SNR=24.53 dB)



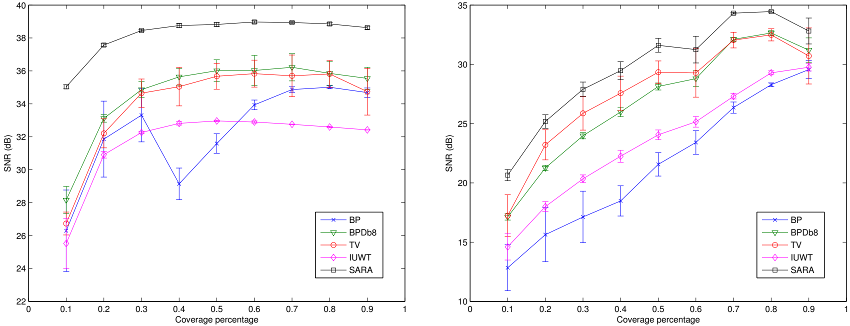
(e) TV (SNR=26.47 dB)



(f) SARA (SNR=29.08 dB)

SARA for radio interferometric imaging

Results on simulations



(a) M31

(b) 30Dor

Figure: Reconstruction fidelity vs visibility coverage.

Outline

1. Inverse Problem
2. Interferometric Imaging with Compressive Sensing
3. Spread Spectrum Effect
4. Continuous Visibilities

Spread spectrum effect

Preliminaries: sparsity and coherence

- What drives the quality of compressive sensing reconstruction?
- Number of measurements M required to achieve exact reconstruction given by

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

where K is the sparsity and N the dimensionality.

- Coherence between the measurement vectors and atoms of sparsity dictionary given by

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

Spread spectrum effect

Preliminaries: sparsity and coherence

- What drives the quality of compressive sensing reconstruction?
- Number of measurements M required to achieve exact reconstruction given by

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

where K is the sparsity and N the dimensionality.

- Coherence between the measurement vectors and atoms of sparsity dictionary given by

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

Spread spectrum effect

Preliminaries: sparsity and coherence

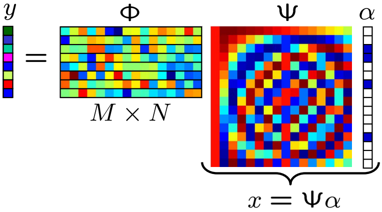
- What drives the quality of compressive sensing reconstruction?
- Number of measurements M required to achieve exact reconstruction given by

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

where K is the sparsity and N the dimensionality.

- Coherence between the measurement vectors and atoms of sparsity dictionary given by

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



Spread spectrum effect

Review

- Non-coplanar baselines and wide fields → w-modulation → spread spectrum effect (first considered by Wiaux *et al.* 2009b).
- The w-modulation operator $\mathbf{C}$ has elements defined by

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

giving rise to to a linear chirp.

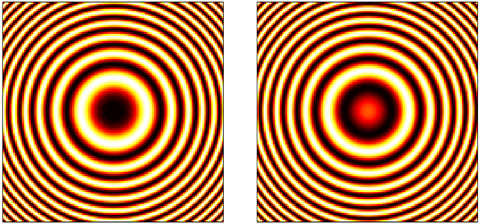
Spread spectrum effect

Review

- Non-coplanar baselines and wide fields → w-modulation → spread spectrum effect (first considered by Wiaux *et al.* 2009b).
- The w-modulation operator $\mathbf{C}$ has elements defined by

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

giving rise to to a linear chirp.



(a) Real part (b) Imaginary part

Figure: Chirp modulation.

Spread spectrum effect

Review

Spread spectrum effect in a nutshell

1. Radio interferometers take (essentially) Fourier measurements.
2. Recall, the coherence is the maximum inner product between measurement vectors and sparsifying atoms.
3. Thus, coherence is (essentially) the maximum of the Fourier coefficients of the atoms of the sparsifying dictionary.
4. w-modulation spreads the spectrum of the atoms of the sparsifying dictionary, reducing the maximum Fourier coefficient.
5. Spreading the spectrum reduces coherence, thus improving reconstruction fidelity.

- Consistent with findings of Carozzi et al. (2013) from information theoretic approach.
- Studied for constant w (for simplicity) by Wiaux *et al.* (2009b).
- Studied for varying w (with realistic images and various sparse representations) by Wolz *et al.* (2013).

Spread spectrum effect

Review

Spread spectrum effect in a nutshell

1. Radio interferometers take (essentially) Fourier measurements.
2. Recall, the coherence is the maximum inner product between measurement vectors and sparsifying atoms.
3. Thus, coherence is (essentially) the maximum of the Fourier coefficients of the atoms of the sparsifying dictionary.
4. w-modulation spreads the spectrum of the atoms of the sparsifying dictionary, reducing the maximum Fourier coefficient.
5. Spreading the spectrum reduces coherence, thus improving reconstruction fidelity.

- Consistent with findings of Carozzi et al. (2013) from information theoretic approach.
- Studied for constant w (for simplicity) by Wiaux *et al.* (2009b).
- Studied for varying w (with realistic images and various sparse representations) by Wolz *et al.* (2013).

Spread spectrum effect

Review

Spread spectrum effect in a nutshell

1. Radio interferometers take (essentially) Fourier measurements.
2. Recall, the coherence is the maximum inner product between measurement vectors and sparsifying atoms.
3. Thus, coherence is (essentially) the maximum of the Fourier coefficients of the atoms of the sparsifying dictionary.
4. w-modulation spreads the spectrum of the atoms of the sparsifying dictionary, reducing the maximum Fourier coefficient.
5. Spreading the spectrum reduces coherence, thus improving reconstruction fidelity.

- Consistent with findings of Carozzi et al. (2013) from information theoretic approach.
- Studied for constant w (for simplicity) by Wiaux *et al.* (2009b).
- Studied for varying w (with realistic images and various sparse representations) by Wolz *et al.* (2013).

Spread spectrum effect

Review

Spread spectrum effect in a nutshell

1. Radio interferometers take (essentially) Fourier measurements.
2. Recall, the coherence is the maximum inner product between measurement vectors and sparsifying atoms.
3. Thus, coherence is (essentially) the maximum of the Fourier coefficients of the atoms of the sparsifying dictionary.
4. w-modulation spreads the spectrum of the atoms of the sparsifying dictionary, reducing the maximum Fourier coefficient.
5. Spreading the spectrum reduces coherence, thus improving reconstruction fidelity.

- Consistent with findings of Carozzi et al. (2013) from information theoretic approach.
- Studied for constant w (for simplicity) by Wiaux *et al.* (2009b).
- Studied for varying w (with realistic images and various sparse representations) by Wolz *et al.* (2013).

Spread spectrum effect

Review

Spread spectrum effect in a nutshell

1. Radio interferometers take (essentially) Fourier measurements.
2. Recall, the coherence is the maximum inner product between measurement vectors and sparsifying atoms.
3. Thus, coherence is (essentially) the maximum of the Fourier coefficients of the atoms of the sparsifying dictionary.
4. w-modulation spreads the spectrum of the atoms of the sparsifying dictionary, reducing the maximum Fourier coefficient.
5. Spreading the spectrum reduces coherence, thus improving reconstruction fidelity.

- Consistent with findings of Carozzi et al. (2013) from information theoretic approach.
- Studied for constant w (for simplicity) by Wiaux *et al.* (2009b).
- Studied for varying w (with realistic images and various sparse representations) by Wolz *et al.* (2013).

Spread spectrum effect

Review

Spread spectrum effect in a nutshell

1. Radio interferometers take (essentially) Fourier measurements.
2. Recall, the coherence is the maximum inner product between measurement vectors and sparsifying atoms.
3. Thus, coherence is (essentially) the maximum of the Fourier coefficients of the atoms of the sparsifying dictionary.
4. w-modulation spreads the spectrum of the atoms of the sparsifying dictionary, reducing the maximum Fourier coefficient.
5. Spreading the spectrum reduces coherence, thus improving reconstruction fidelity.

- Consistent with findings of Carozzi et al. (2013) from information theoretic approach.
- Studied for constant w (for simplicity) by Wiaux *et al.* (2009b).
- Studied for varying w (with realistic images and various sparse representations) by Wolz *et al.* (2013).

Spread spectrum effect

Review

Spread spectrum effect in a nutshell

1. Radio interferometers take (essentially) Fourier measurements.
2. Recall, the coherence is the maximum inner product between measurement vectors and sparsifying atoms.
3. Thus, coherence is (essentially) the maximum of the Fourier coefficients of the atoms of the sparsifying dictionary.
4. w-modulation spreads the spectrum of the atoms of the sparsifying dictionary, reducing the maximum Fourier coefficient.
5. Spreading the spectrum reduces coherence, thus improving reconstruction fidelity.

- Consistent with findings of Carozzi et al. (2013) from information theoretic approach.
- Studied for constant w (for simplicity) by Wiaux *et al.* (2009b).
- Studied for varying w (with realistic images and various sparse representations) by Wolz *et al.* (2013).

Spread spectrum effect

Review

Spread spectrum effect in a nutshell

1. Radio interferometers take (essentially) Fourier measurements.
2. Recall, the coherence is the maximum inner product between measurement vectors and sparsifying atoms.
3. Thus, coherence is (essentially) the maximum of the Fourier coefficients of the atoms of the sparsifying dictionary.
4. w-modulation spreads the spectrum of the atoms of the sparsifying dictionary, reducing the maximum Fourier coefficient.
5. Spreading the spectrum reduces coherence, thus improving reconstruction fidelity.

- Consistent with findings of Carozzi et al. (2013) from information theoretic approach.
- Studied for constant w (for simplicity) by Wiaux *et al.* (2009b).
- Studied for varying w (with realistic images and various sparse representations) by Wolz *et al.* (2013).

Spread spectrum effect

Sparse w-projection

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

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

- Naively, expressing the application of the w-modulation in this manner is computationally less efficient that the original formulation but it has two important advantages.
- Different w for each (u, v), while still exploiting FFT.
- Many of the elements of $\hat{\mathbf{C}}$ will be close to zero.
- Support of w-modulation in Fourier space determined dynamically.

Spread spectrum effect

Sparse w-projection

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

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

- Naively, expressing the application of the w-modulation in this manner is computationally less efficient that the original formulation but it has two important advantages.

- Different w for each (u, v), while still exploiting FFT.

- Many of the elements of $\hat{\mathbf{C}}$ will be close to zero.

- Support of w-modulation in Fourier space determined dynamically.

Spread spectrum effect

Sparse w-projection

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

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

- Naively, expressing the application of the w-modulation in this manner is computationally less efficient that the original formulation but it has two important advantages.
- Different w for each (u, v), while still exploiting FFT.
- Many of the elements of $\hat{\mathbf{C}}$ will be close to zero.
- Support of w-modulation in Fourier space determined dynamically.

Spread spectrum effect

Sparse w-projection

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

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

- Naively, expressing the application of the w-modulation in this manner is computationally less efficient that the original formulation but it has two important advantages.
- Different w for each (u, v), while still exploiting FFT.
- Many of the elements of $\hat{\mathbf{C}}$ will be close to zero.
- Support of w-modulation in Fourier space determined dynamically.

Spread spectrum effect

Sparse w-projection

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

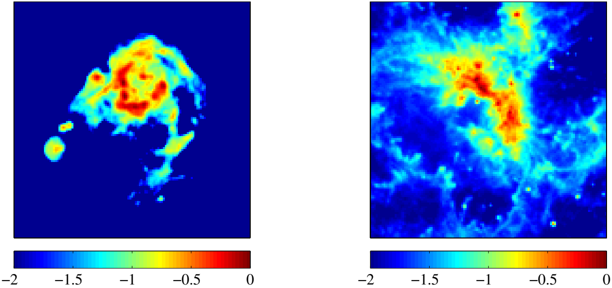
$$\boxed{\Phi = \mathbf{M}\,\mathbf{F}\,\mathbf{C}\,\mathbf{A} \quad \Rightarrow \quad \Phi = \hat{\mathbf{C}}\,\mathbf{F}\,\mathbf{A}} \, .$$

- Naively, expressing the application of the w-modulation in this manner is computationally less efficient that the original formulation but it has two important advantages.
- Different w for each (u, v), while still exploiting FFT.
- Many of the elements of $\hat{\mathbf{C}}$ will be close to zero.
- Support of w-modulation in Fourier space determined dynamically.

Spread spectrum effect for varying w

Results on simulations

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

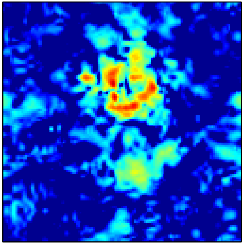


(a) M31 (b) 30Dor

Figure: Ground truth images in logarithmic scale.

Spread spectrum effect for varying w

Results on simulations



(a) $w_{\mathrm{d}} = 0 \rightarrow$ SNR$=5$dB

Figure: Reconstructed images of M31 for 10% coverage.

Spread spectrum effect for varying w

Results on simulations

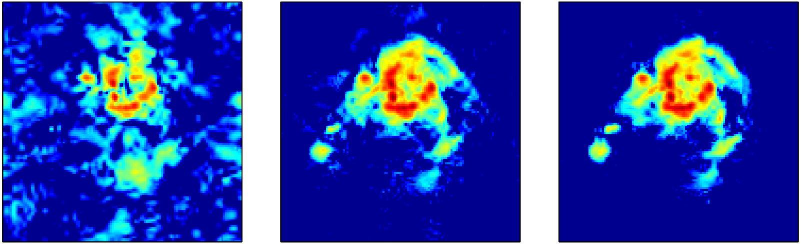
(a) $w_{\mathrm{d}} = 0 \rightarrow$ SNR$=$ 5dB

(c) $w_{\mathrm{d}} = 1 \rightarrow$ SNR$=$ 19dB

Figure: Reconstructed images of M31 for 10% coverage.

Spread spectrum effect for varying w

Results on simulations

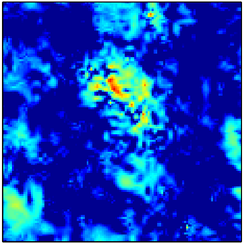


(a) $w_\mathrm{d} = 0 \rightarrow$ SNR$=$ 5dB (b) $w_\mathrm{d} \sim \mathcal{U}(0, 1) \rightarrow$ SNR$=$ 16dB (c) $w_\mathrm{d} = 1 \rightarrow$ SNR$=$ 19dB

Figure: Reconstructed images of M31 for 10% coverage.

Spread spectrum effect for varying w

Results on simulations



(a) $w_{\rm d} = 0 \rightarrow$ SNR$=2$dB

Figure: Reconstructed images of 30Dor for 10% coverage.

Spread spectrum effect for varying w

Results on simulations

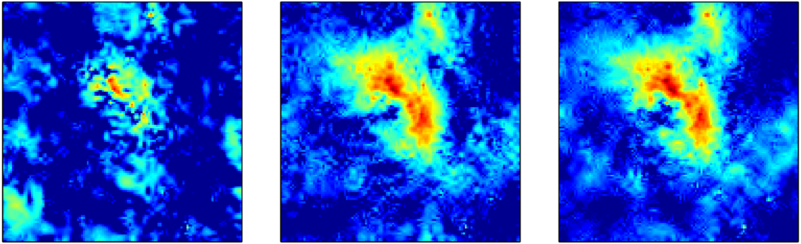
(a) $w_{\rm d} = 0 \rightarrow$ SNR$=$ 2dB

(c) $w_{\rm d} = 1 \rightarrow$ SNR$=$ 15dB

Figure: Reconstructed images of 30Dor for 10% coverage.

Spread spectrum effect for varying w

Results on simulations

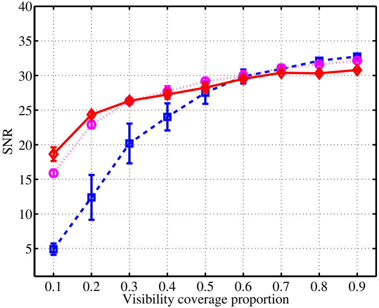


(a) $w_\mathrm{d} = 0 \rightarrow$ SNR$=$ 2dB (b) $w_\mathrm{d} \sim \mathcal{U}(0, 1) \rightarrow$ SNR$=$ 12dB (c) $w_\mathrm{d} = 1 \rightarrow$ SNR$=$ 15dB

Figure: Reconstructed images of 30Dor for 10% coverage.

Spread spectrum effect for varying w

Results on simulations



(a) Daubechies 8 (Db8) wavelets

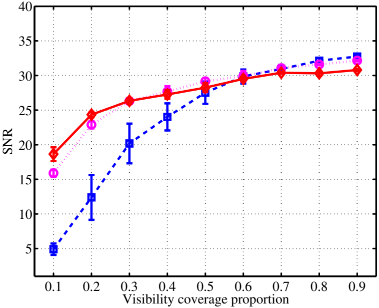
Figure: Reconstruction fidelity for M31.

Improvement in reconstruction fidelity due to the spread spectrum effect 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 effect for varying w

Results on simulations



(a) Daubechies 8 (Db8) wavelets

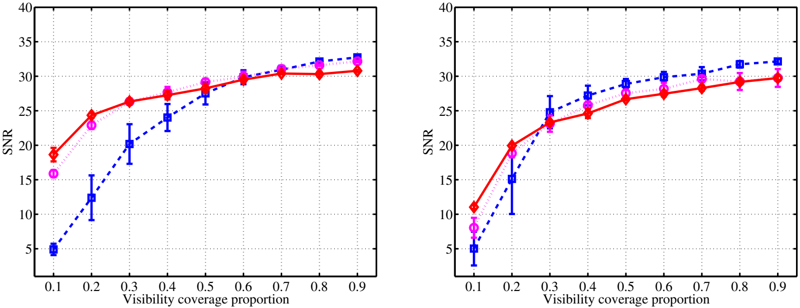
Figure: Reconstruction fidelity for M31.

Improvement in reconstruction fidelity due to the spread spectrum effect 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 effect for varying w

Results on simulations



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

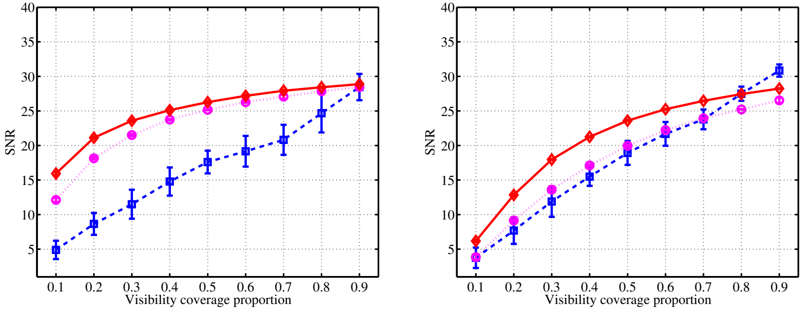
Figure: Reconstruction fidelity for M31.

Improvement in reconstruction fidelity due to the spread spectrum effect 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 effect for varying w

Results on simulations



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

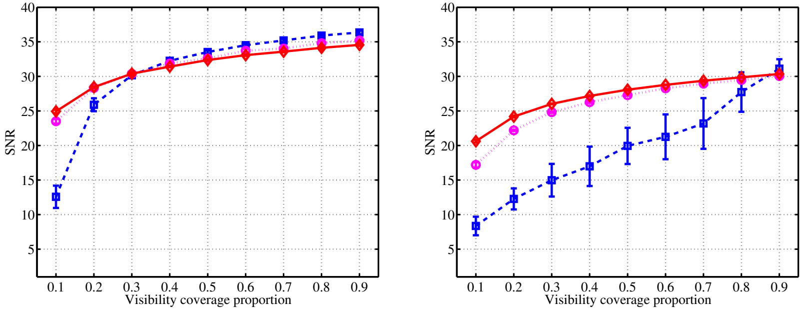
Figure: Reconstruction fidelity for 30Dor.

Improvement in reconstruction fidelity due to the spread spectrum effect 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 effect for varying w

Results on simulations



(a) M31 (b) 30 Dor

Figure: Reconstruction fidelity using SARA.

Improvement in reconstruction fidelity due to the spread spectrum effect 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.

Outline

1. Inverse Problem
2. Interferometric Imaging with Compressive Sensing
3. Spread Spectrum Effect
4. Continuous Visibilities

Supporting continuous visibilities

Algorithm

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

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

- But this is impractically slow!
- Incorporated gridding into our CS interferometric imaging framework (Carrillo *et al.* 2013).
- Model with measurement operator

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

where we incorporate:

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

Supporting continuous visibilities

Algorithm

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

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

- But this is impractically slow!
- Incorporated gridding into our CS interferometric imaging framework (Carrillo *et al.* 2013).
- Model with measurement operator

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

where we incorporate:

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

Supporting continuous visibilities

Algorithm

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

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

- But this is impractically slow!
- Incorporated gridding into our CS interferometric imaging framework (Carrillo *et al.* 2013).
- Model with measurement operator

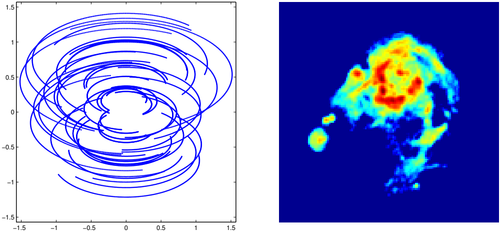
$$\boxed{\Phi = \mathbf{G}\,\mathbf{F}\,\mathbf{D}\,\mathbf{Z}} \ ,$$

where we incorporate:

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

Supporting continuous visibilities

Results on simulations



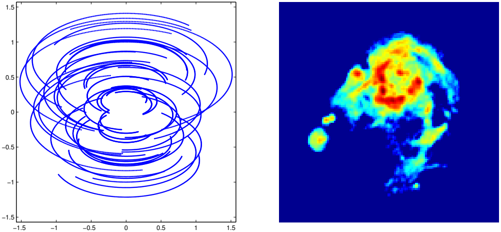
(a) Coverage

(b) M31 (ground truth)

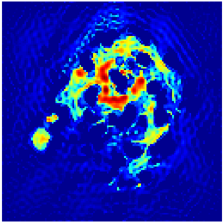
Figure: Reconstructed images from continuous visibilities.

Supporting continuous visibilities

Results on simulations



(a) Coverage (b) M31 (ground truth)

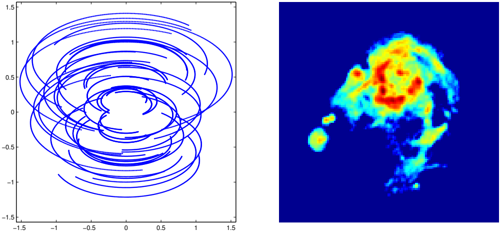


(c) Dirac basis → SNR= 8.2dB

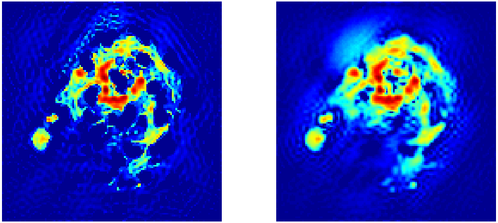
Figure: Reconstructed images from continuous visibilities.

Supporting continuous visibilities

Results on simulations



(a) Coverage (b) M31 (ground truth)

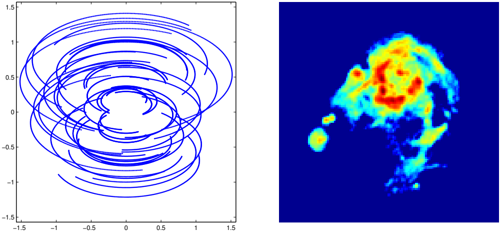


(c) Dirac basis → SNR= 8.2dB (d) Db8 wavelets → SNR= 11.1dB

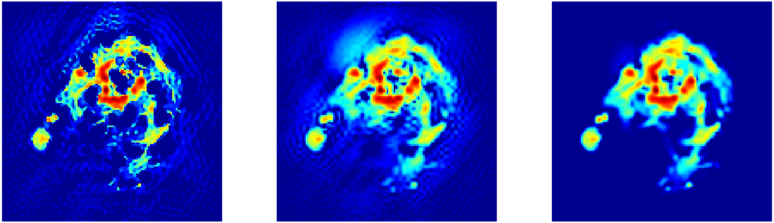
Figure: Reconstructed images from continuous visibilities.

Supporting continuous visibilities

Results on simulations



(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.

Conclusions & outlook

- Effectiveness of compressive sensing for radio interferometric imaging demonstrated.
- Theory of compressive sensing can guide telescope design.
- We have just released the PURIFY code to scale to the realistic setting.
- Includes state-of-the-art convex optimisation algorithms that support parallelisation.
- Application of (unconstrained) BPDN to LOFAR by Garsden *et al.* 2014.

Apply to observations made by real interferometric telescopes.

Conclusions & outlook

- Effectiveness of compressive sensing for radio interferometric imaging demonstrated.
- Theory of compressive sensing can guide telescope design.
- We have just released the PURIFY code to scale to the realistic setting.
- Includes state-of-the-art convex optimisation algorithms that support parallelisation.
- Application of (unconstrained) BPDN to LOFAR by Garsden *et al.* 2014.

Apply to observations made by real interferometric telescopes.

Conclusions & outlook

- Effectiveness of compressive sensing for radio interferometric imaging demonstrated.
- Theory of compressive sensing can guide telescope design.
- We have just released the PURIFY code to scale to the realistic setting.
- Includes state-of-the-art convex optimisation algorithms that support parallelisation.
- Application of (unconstrained) BPDN to LOFAR by Garsden *et al.* 2014.

Apply to observations made by real interferometric telescopes.

PURIFY code `http://basp-group.github.io/purify/`

Next-generation radio interferometric imaging
Carrillo, McEwen, Wiaux

PURIFY is an open-source code that provides functionality to perform radio interferometric imaging, leveraging recent developments in the field of compressive sensing and convex optimisation.

Conclusions & outlook

- Effectiveness of compressive sensing for radio interferometric imaging demonstrated.
- Theory of compressive sensing can guide telescope design.
- We have just released the PURIFY code to scale to the realistic setting.
- Includes state-of-the-art convex optimisation algorithms that support parallelisation.
- Application of (unconstrained) BPDN to LOFAR by Garsden *et al.* 2014.

Apply to observations made by real interferometric telescopes.

PURIFY code `http://basp-group.github.io/purify/`

Next-generation radio interferometric imaging
Carrillo, McEwen, Wiaux

PURIFY is an open-source code that provides functionality to perform radio interferometric imaging, leveraging recent developments in the field of compressive sensing and convex optimisation.

Extra Slides

Sparse w-projection

Sparsity of w-modulation kernel in Fourier space

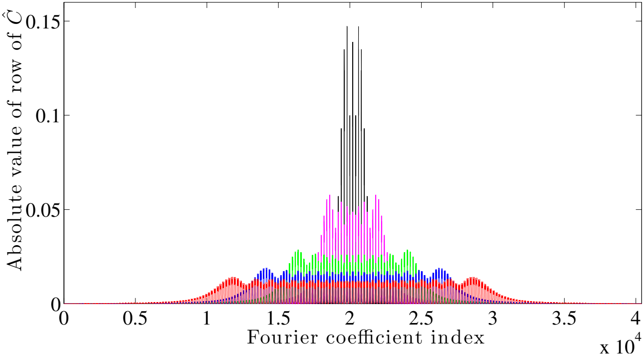


Figure: Rows of Fourier representation of w-modulation operator $\hat{\mathbf{C}}$.

Sparse w-projection

Dynamic sparsification

- We make a sparse matrix approximation of $\hat{\mathsf{C}}$ to speed up its computational application and reduce memory requirements.
- Sparsify $\hat{\mathsf{C}}$ by dynamic thresholding.
- Retain $E\%$ of the energy content for each visibility measurement.
- Support of w-modulation kernel in Fourier space determined dynamically, so don't require any prior information about structure.
- Generic procedure applicable for any direction-dependent effect (DDE).

Sparse w-projection

Dynamic sparsification

- We make a sparse matrix approximation of $\hat{\mathsf{C}}$ to speed up its computational application and reduce memory requirements.
- Sparsify $\hat{\mathsf{C}}$ by dynamic thresholding.
- Retain E% of the energy content for each visibility measurement.
- Support of w-modulation kernel in Fourier space determined dynamically, so don't require any prior information about structure.
- Generic procedure applicable for any direction-dependent effect (DDE).

Sparse w-projection

Dynamic sparsification

- We make a sparse matrix approximation of $\hat{\mathsf{C}}$ to speed up its computational application and reduce memory requirements.
- Sparsify $\hat{\mathsf{C}}$ by dynamic thresholding.
- Retain E% of the energy content for each visibility measurement.
- Support of w-modulation kernel in Fourier space determined dynamically, so don't require any prior information about structure.
- Generic procedure applicable for any direction-dependent effect (DDE).

Sparse w-projection

Sparsified w-modulation kernels

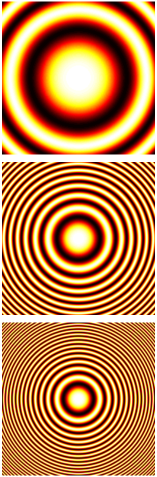


Figure: w-modulation kernel.

Sparse w-projection

Sparsified w-modulation kernels

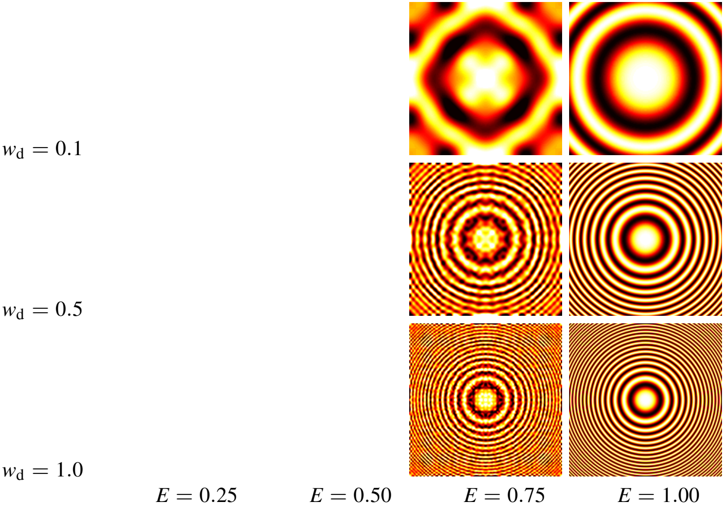


Figure: w-modulation kernel.

Sparse w-projection

Sparsified w-modulation kernels

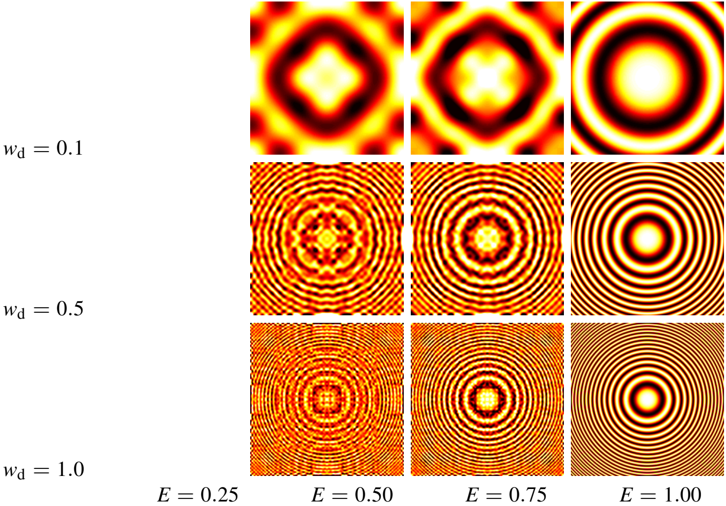


Figure: w-modulation kernel.

Sparse w-projection

Sparsified w-modulation kernels

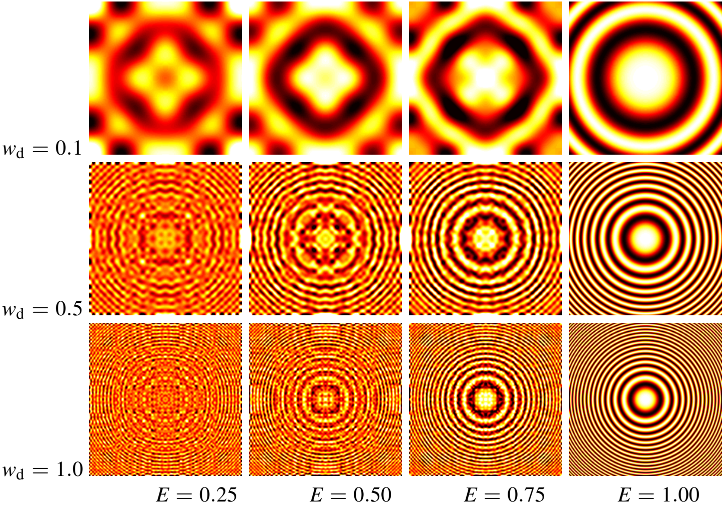


Figure: w-modulation kernel.

Sparse w-projection

Proportion of non-zero entries

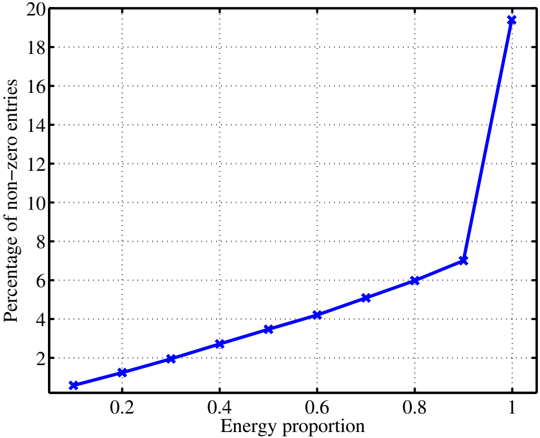


Figure: Percentage of non-zero entries as a function of preserved energy proportion.

Sparse w-projection

Runtime improvements

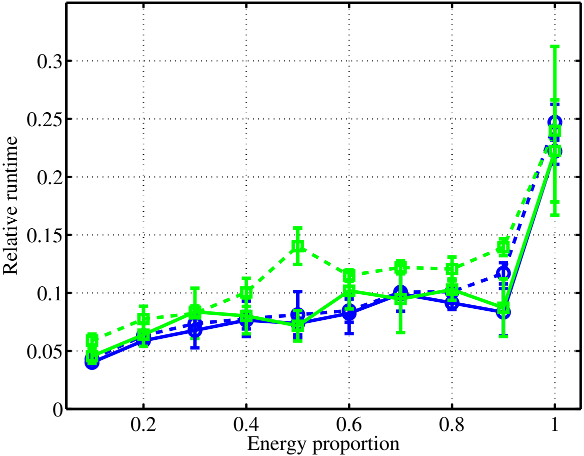


Figure: Relative runtime as a function of preserved energy proportion for 10% (dashed) and 50% (solid) visibility coverages.

Sparse w-projection

Impact on reconstruction quality

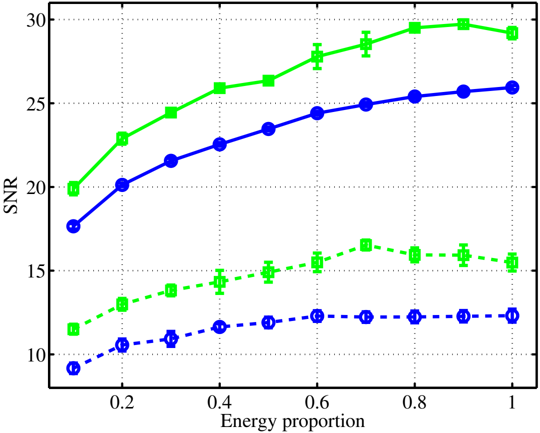


Figure: Reconstruction quality as a function of preserved energy proportion for 10% (dashed) and 50% (solid) visibility coverages.