Radio interferometry in the big-data era of the Square Kilometre Array (SKA)

Jason McEwen
www.jasonmcewen.org
@jasonmcewen

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

UCL Mathematical & Physical Sciences (MAPS) Faculty Research Festival

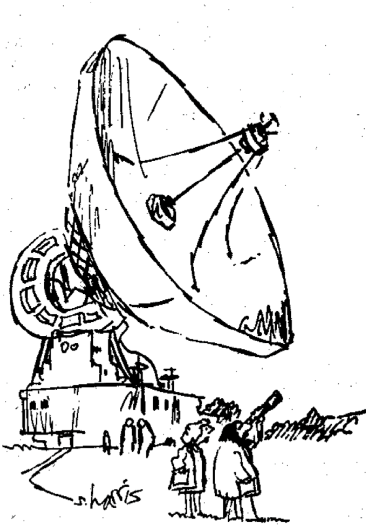
April 2016

Outline

Outline

1. Radio interferometry and the SKA
2. Interferometric imaging with compressive sensing
3. Scalable algorithms

Radio telescopes are big!



"Just checking."

Radio telescopes are big!

Radio interferometric telescopes

Next-generation of radio interferometry rapidly approaching

- Many pathfinder radio interferometric telescopes coming online, *e.g.* LOFAR, ASKAP, MeerKAT, MWA.
- Square Kilometre Array (SKA) construction scheduled to begin 2018.
- Broad range of science goals.

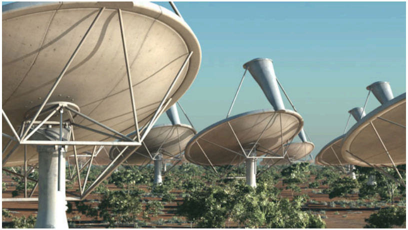


Figure: Artist impression of SKA dishes. [Credit: SKA Organisation]

Next-generation of radio interferometry rapidly approaching

- Many pathfinder radio interferometric telescopes coming online, *e.g.* LOFAR, ASKAP, MeerKAT, MWA.
- Square Kilometre Array (SKA) construction scheduled to begin 2018.
- Broad range of science goals.

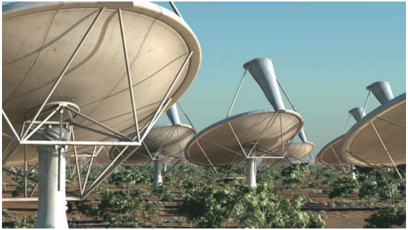
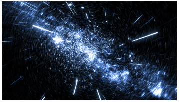
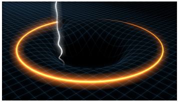


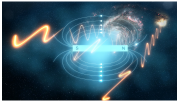
Figure: Artist impression of SKA dishes. [Credit: SKA Organisation]



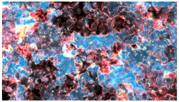
(a) Dark-energy



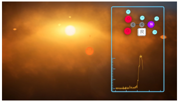
(b) GR



(c) Cosmic magnetism



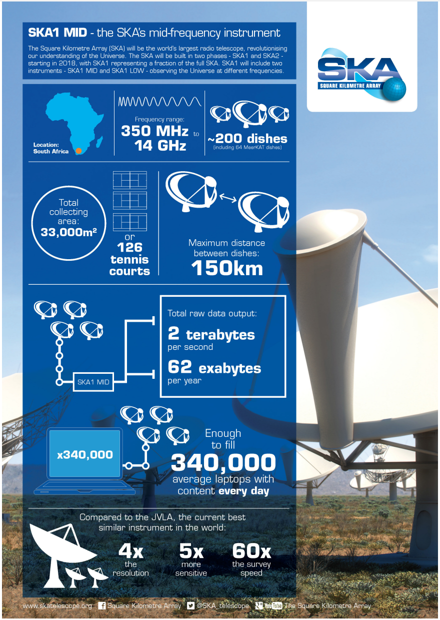
(d) Epoch of reionization

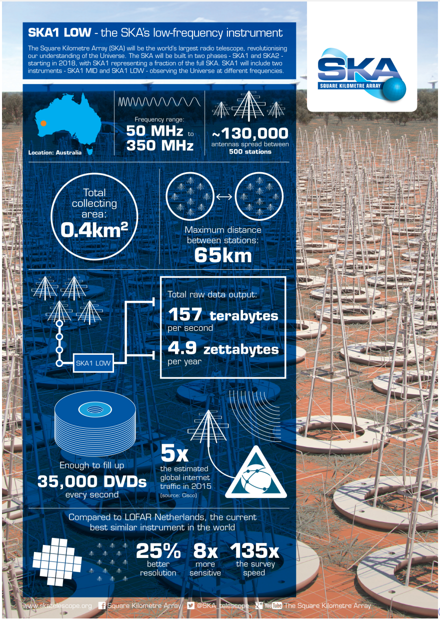


(e) Exoplanets

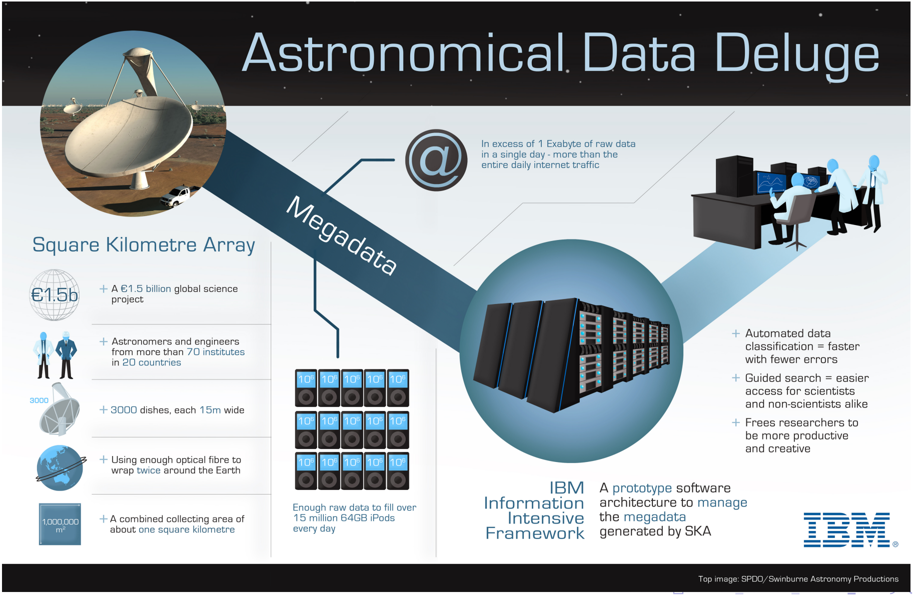
Figure: SKA science goals. [Credit: SKA Organisation]

SKA sites

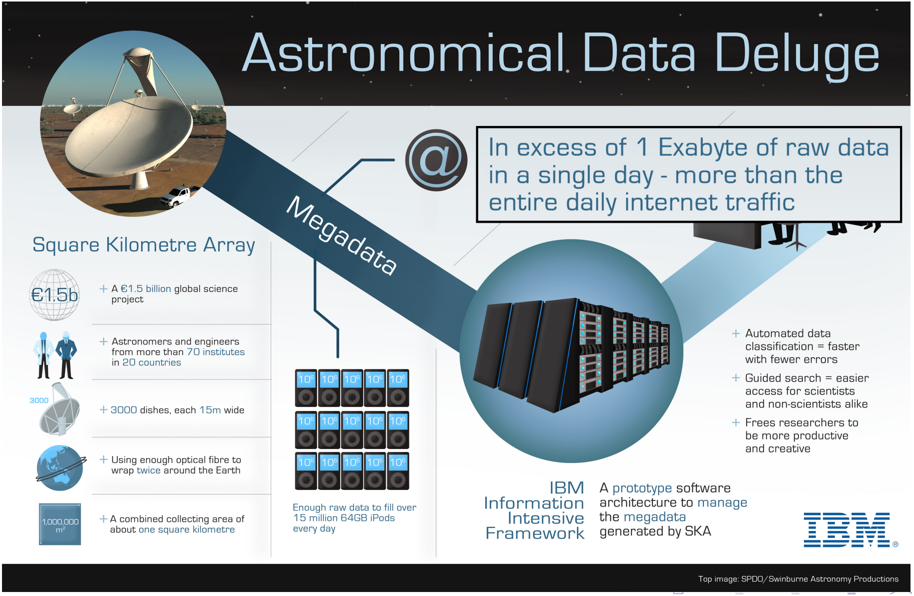

SKA1 MID - the SKA's mid-frequency instrument
The Square Kilometre Array (SKA) will be the world's largest radio telescope, revolutionising our understanding of the Universe. The SKA will be built in two phases - SKA1 and SKA2 - starting in 2018, with SKA1 representing a fraction of the full SKA. SKA1 will include two instruments - SKA1 MID and SKA1 LOW - observing the Universe at different frequencies.
SKA
SQUARE KILOMETRE ARRAY
Location: South Africa
Frequency range:
350 MHz to 14 GHz
~200 dishes
(including 64 MeerKAT dishes)
Total collecting area: 33,000m²
or 126 tennis courts
Maximum distance between dishes: 150km
SKA1 MID
Total raw data output:
2 terabytes per second
62 exabytes per year
x340,000
Enough to fill 340,000 average laptops with content every day
Compared to the JVLA, the current best similar instrument in the world:
4x the resolution
5x more sensitive
60x the survey speed
www.skatelescope.org
Square Kilometre Array
@SKA_telescope
The Square Kilometre Array


SKA1 LOW - the SKA's low-frequency instrument
The Square Kilometre Array (SKA) will be the world's largest radio telescope, revolutionising our understanding of the Universe. The SKA will be built in two phases - SKA1 and SKA2 - starting in 2018, with SKA1 representing a fraction of the full SKA. SKA1 will include two instruments - SKA1 MID and SKA1 LOW - observing the Universe at different frequencies.
SKA
SQUARE KILOMETRE ARRAY
Location: Australia
Frequency range:
50 MHz to 350 MHz
~130,000 antennas spread between 500 stations
Total collecting area: 0.4km²
Maximum distance between stations: 65km
SKA1 LOW
Total raw data output:
157 terabytes per second
4.9 zettabytes per year
Enough to fill up 35,000 DVDs every second
5x the estimated global internet traffic in 2015
(source: Cisco)
Compared to LOFAR Netherlands, the current best similar instrument in the world
25% better resolution
8x more sensitive
135x the survey speed
www.skatelescope.org
Square Kilometre Array
@SKA_telescope
The Square Kilometre Array

The SKA poses a considerable big-data challenge



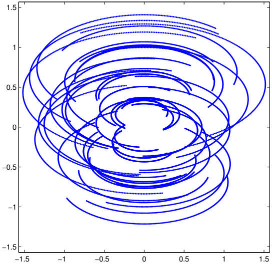
The SKA poses a considerable big-data challenge



Outline

Radio interferometric telescopes acquire “Fourier” measurements

“Fourier” Measurements



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 **A** of the telescope;
  - w-modulation modulation **C**;
  - Fourier transform **F**;
  - masking **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.

Compressive sensing

- Developed by Candes *et al.* 2006 and Donoho 2006 (and others).
- Although many underlying ideas around for a long time.
- Exploits the sparsity of natural signals.
- Acquisition versus imaging.

(a) Emmanuel Candes

(b) David Donoho

Compressive sensing

- Developed by Candes *et al.* 2006 and Donoho 2006 (and others).
- Although many underlying ideas around for a long time.
- Exploits the sparsity of natural signals.
- Acquisition versus imaging.

(a) Emmanuel Candes

(b) David Donoho

Interferometric imaging with compressed sensing

- Solve the interferometric imaging problem

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

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

- Basis Pursuit (BP) denoising problem

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

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

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

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

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

SARA algorithm 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:

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

SARA

where $\mathbf{W} \in \mathbb{R}^{D \times D}$ 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 algorithm 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:

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

SARA

  where $\mathbf{W} \in \mathbb{R}^{D \times D}$ 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 algorithm 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:

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

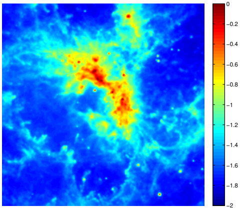
SARA

  where $\mathbf{W} \in \mathbb{R}^{D \times D}$ 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

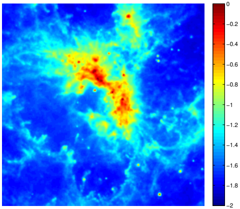
Results on simulations



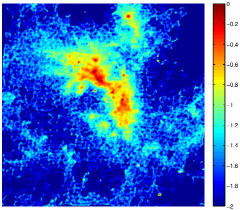
(a) Original

SARA for radio interferometric imaging

Results on simulations



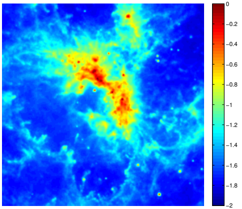
(a) Original



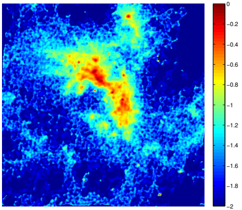
(b) “CLEAN” (SNR=16.67 dB)

SARA for radio interferometric imaging

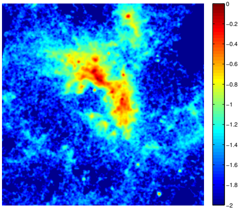
Results on simulations



(a) Original



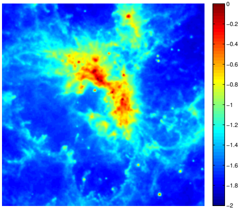
(b) “CLEAN” (SNR=16.67 dB)



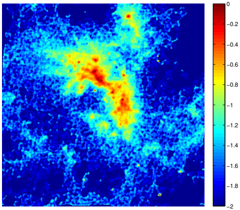
(c) “MS-CLEAN” (SNR=17.87 dB)

SARA for radio interferometric imaging

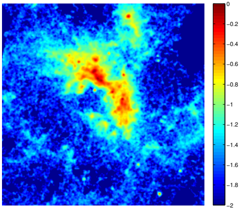
Results on simulations



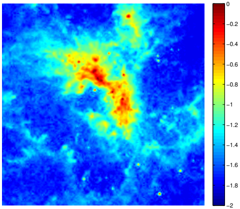
(a) Original



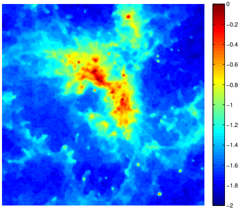
(b) “CLEAN” (SNR=16.67 dB)



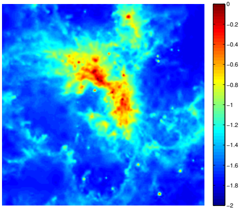
(c) “MS-CLEAN” (SNR=17.87 dB)



(d) BPDb8 (SNR=24.53 dB)



(e) TV (SNR=26.47 dB)



(f) SARA (SNR=29.08 dB)

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.* 2014).
- 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

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

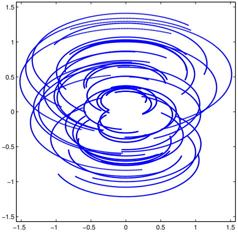
- But this is impractically slow!
- Incorporated gridding into our CS interferometric imaging framework (Carrillo *et al.* 2014).
- Model with measurement operator

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

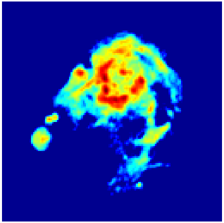
where we incorporate:
  - convolutional gridding operator $\mathbf{G}$;
  - fast Fourier transform $\mathbf{F}$;
  - normalisation operator $\mathbf{D}$ to undo the convolution gridding;
  - zero-padding operator $\mathbf{Z}$ to upsample the discrete visibility space.

Supporting continuous visibilities

Results on simulations



(a) Coverage

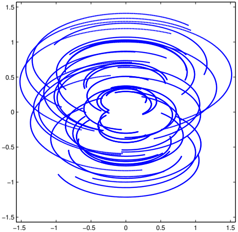


(b) M31 (ground truth)

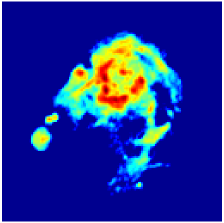
Figure: Reconstructed images from continuous visibilities.

Supporting continuous visibilities

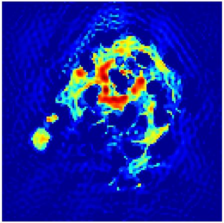
Results on simulations



(a) Coverage



(b) M31 (ground truth)

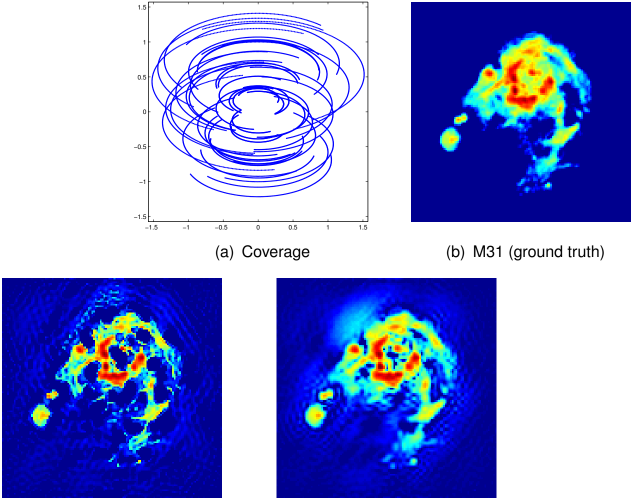


(c) “CLEAN” (SNR$=8.2$dB)

Figure: Reconstructed images from continuous visibilities.

Supporting continuous visibilities

Results on simulations



(a) Coverage

(b) M31 (ground truth)

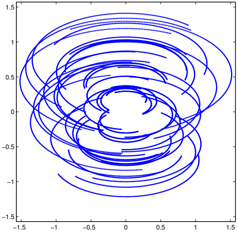
(c) “CLEAN” (SNR$=8.2$dB)

(d) “MS-CLEAN” (SNR$=11.1$dB)

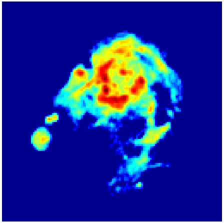
Figure: Reconstructed images from continuous visibilities.

Supporting continuous visibilities

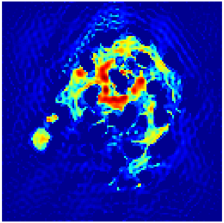
Results on simulations



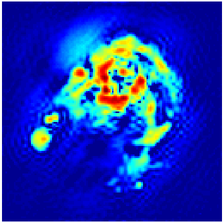
(a) Coverage



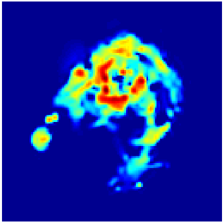
(b) M31 (ground truth)



(c) "CLEAN" (SNR= 8.2dB)



(d) "MS-CLEAN" (SNR= 11.1dB)

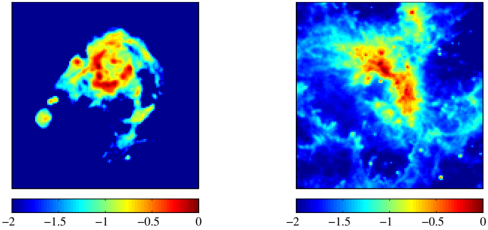


(e) SARA (SNR= 13.4dB)

Figure: Reconstructed images from continuous visibilities.

Optimising telescope configurations
Spread spectrum effect

- Use theory of compressive sensing to optimise telescope configurations.
- Non-coplanar baselines and wide fields $\rightarrow$ w-modulation $\rightarrow$ spread spectrum effect which reduces coherence $\rightarrow$ improves reconstruction quality (first considered by Wiaux *et al.* 2009b).
- Perform simulations to assess the effectiveness of the spread spectrum effect in the presence of varying w (Wolz, McEwen *et al.* 2013).

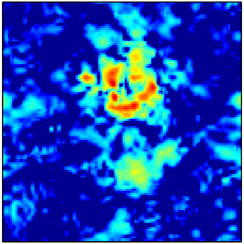


(a) M31 (b) 30Dor

Figure: Ground truth images in logarithmic scale.

Optimising telescope configurations

Results on simulations

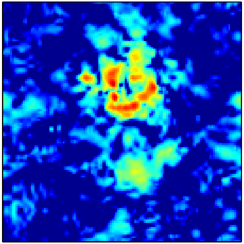


(a) $w_{\mathrm{d}} = 0 \rightarrow$ SNR$= 5$dB

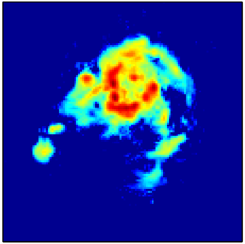
Figure: Reconstructed images of M31 for 10% coverage.

Optimising telescope configurations

Results on simulations



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

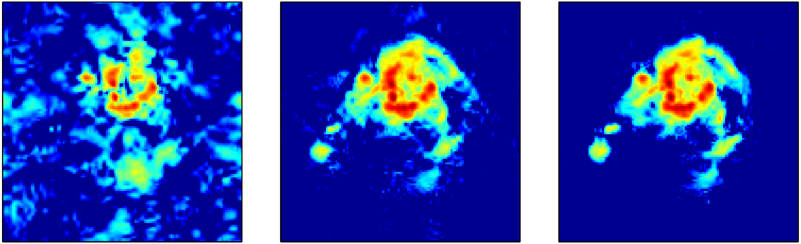


(c) $w_{\mathrm{d}} = 1 \rightarrow$ SNR$=$ 19dB

Figure: Reconstructed images of M31 for 10% coverage.

Optimising telescope configurations

Results on simulations

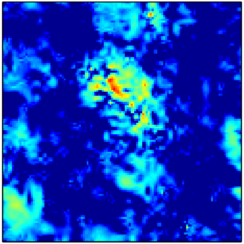


(a) $w_{\mathrm{d}} = 0 \rightarrow$ SNR= 5dB (b) $w_{\mathrm{d}} \sim \mathcal{U}(0, 1) \rightarrow$ SNR= 16dB (c) $w_{\mathrm{d}} = 1 \rightarrow$ SNR= 19dB

Figure: Reconstructed images of M31 for 10% coverage.

Optimising telescope configurations

Results on simulations

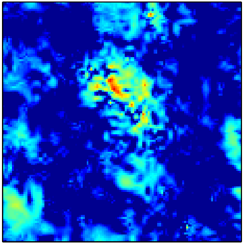


(a) $w_{\mathrm{d}} = 0 \rightarrow$ SNR$= 2$dB

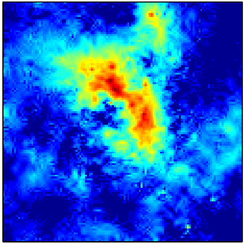
Figure: Reconstructed images of 30Dor for 10% coverage.

Optimising telescope configurations

Results on simulations



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

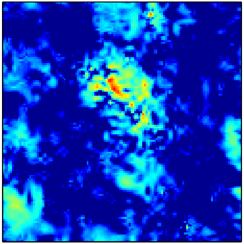


(c) $w_{\mathrm{d}} = 1 \rightarrow$ SNR$=$ 15dB

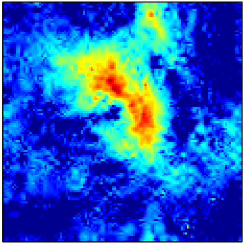
Figure: Reconstructed images of 30Dor for 10% coverage.

Optimising telescope configurations

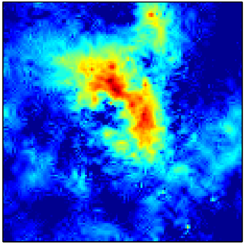
Results on simulations



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



(b) $w_{\rm d} \sim \mathcal{U}(0,1) \rightarrow$ SNR$=$ 12dB

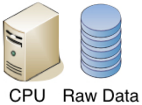


(c) $w_{\rm d} = 1 \rightarrow$ SNR$=$ 15dB

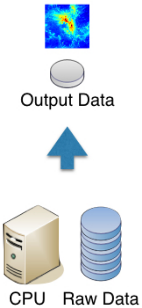
Figure: Reconstructed images of 30Dor for 10% coverage.

Outline

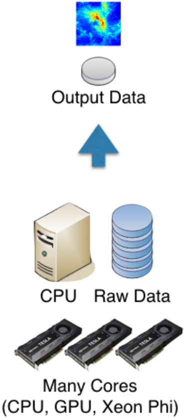
Standard algorithms



Standard algorithms



Standard algorithms



Block algorithm

- Block algorithm to split data and measurement operator
  (Carrillo, McEwen & Wiaux 2014; Onose, Carrillo, Repetti, McEwen, Thiran, Pesquet & Wiaux 2016)

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_{n_\mathrm{d}} \end{bmatrix}, \quad \Phi = \begin{bmatrix} \Phi_1 \\ \vdots \\ \Phi_{n_\mathrm{d}} \end{bmatrix} = \begin{bmatrix} \mathbf{G}_1 \mathbf{M}_1 \\ \vdots \\ \mathbf{G}_{n_\mathrm{d}} \mathbf{M}_{n_\mathrm{d}} \end{bmatrix} \mathbf{F}\mathbf{Z}.$$

- For SARA, sparsifying operator can also be naturally split into constituent dictionaries:

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

- Leads to a highly distributed and parallelised algorithmic structure.

Block algorithm

- Block algorithm to split data and measurement operator
(Carrillo, McEwen & Wiaux 2014; Onose, Carrillo, Repetti, McEwen, Thiran, Pesquet & Wiaux 2016)

$$\boxed{\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_{n_\mathrm{d}} \end{bmatrix}, \quad \Phi = \begin{bmatrix} \Phi_1 \\ \vdots \\ \Phi_{n_\mathrm{d}} \end{bmatrix} = \begin{bmatrix} \mathbf{G}_1 \mathbf{M}_1 \\ \vdots \\ \mathbf{G}_{n_\mathrm{d}} \mathbf{M}_{n_\mathrm{d}} \end{bmatrix} \mathbf{FZ}.}$$

- For SARA, sparsifying operator can also be naturally split into constituent dictionaries:

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

- Leads to a highly distributed and parallelised algorithmic structure.

Block algorithm

- Block algorithm to split data and measurement operator
(Carrillo, McEwen & Wiaux 2014; Onose, Carrillo, Repetti, McEwen, Thiran, Pesquet & Wiaux 2016)

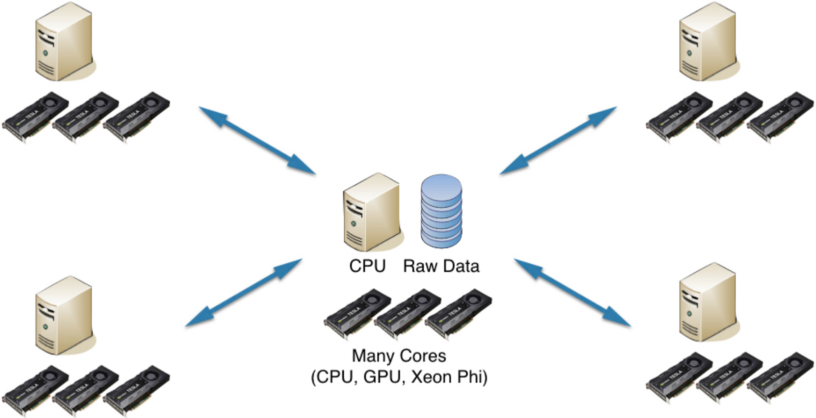
$$\boxed{\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_{n_\mathrm{d}} \end{bmatrix}, \quad \Phi = \begin{bmatrix} \Phi_1 \\ \vdots \\ \Phi_{n_\mathrm{d}} \end{bmatrix} = \begin{bmatrix} \mathbf{G}_1 \mathbf{M}_1 \\ \vdots \\ \mathbf{G}_{n_\mathrm{d}} \mathbf{M}_{n_\mathrm{d}} \end{bmatrix} \mathbf{FZ}.}$$

- For SARA, sparsifying operator can also be naturally split into constituent dictionaries:

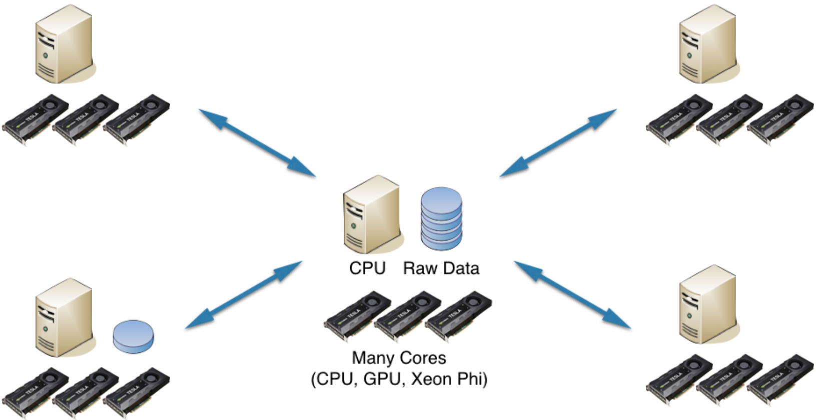
$$\boxed{\Psi = \frac{1}{\sqrt{q}} [\Psi_1, \Psi_2, \ldots, \Psi_q].}$$

- Leads to a highly distributed and parallelised algorithmic structure.

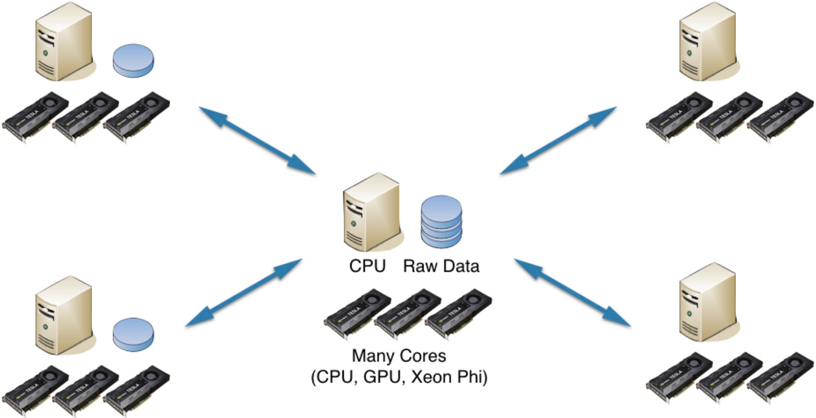
Highly parallelised and distributed algorithms



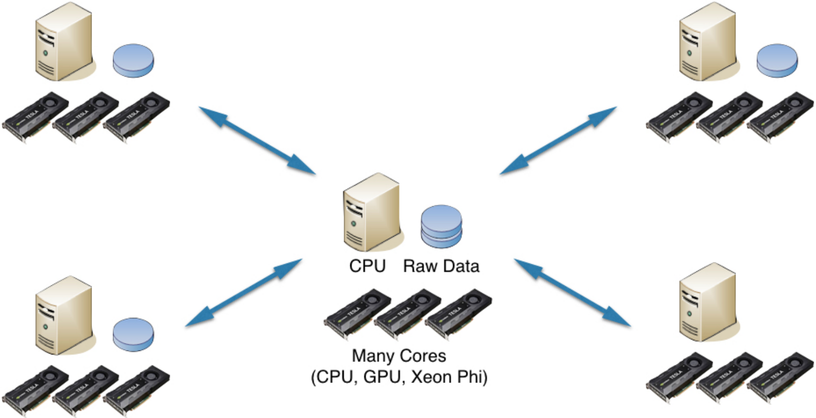
Highly parallelised and distributed algorithms



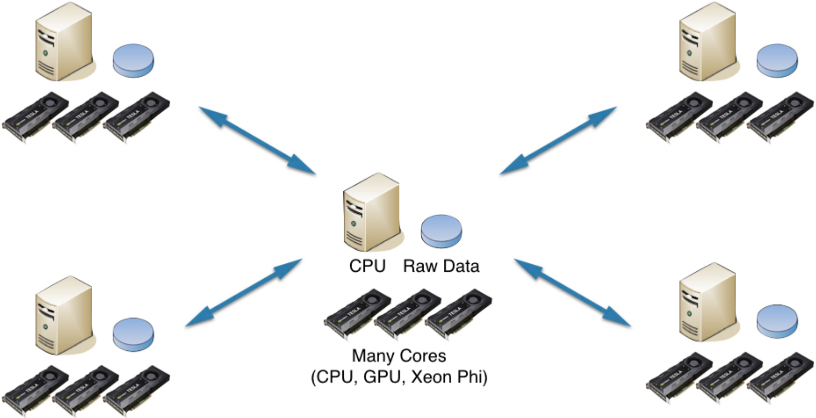
Highly parallelised and distributed algorithms



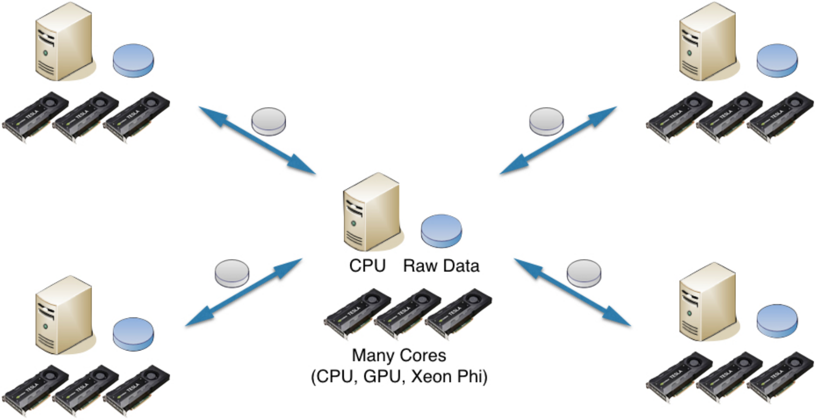
Highly parallelised and distributed algorithms



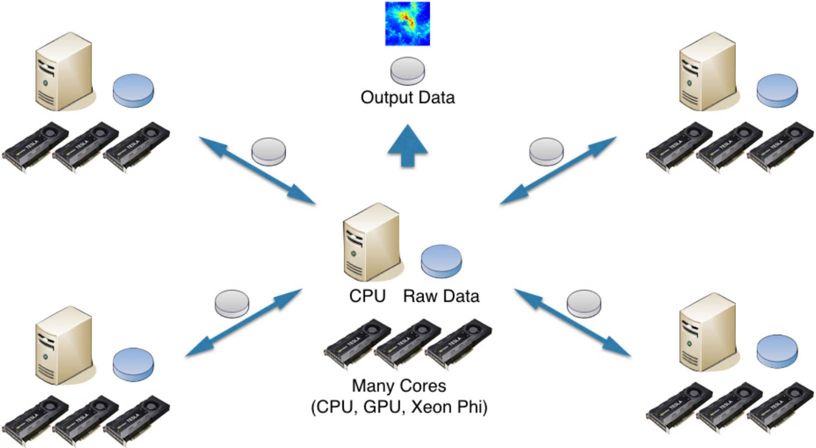
Highly parallelised and distributed algorithms



Highly parallelised and distributed algorithms

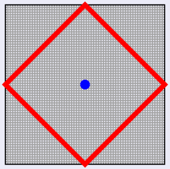


Highly parallelised and distributed algorithms



Public codes

SOPT code http://basp-group.github.io/sopt/



Sparse OPTimisation
Carrillo, McEwen, Wiaux

SOPT is an open-source code that provides functionality to perform sparse optimisation using state-of-the-art convex optimisation algorithms.

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 be used to optimise telescope configuration.
- State-of-the-art convex optimisation algorithms that support distribution.

Applying to observations made by real interferometric telescopes.

Developing fast convex optimisation algorithms that are parallelised and distributed to scale to big-data.

Supported by:

EPSRC Engineering and Physical Sciences Research Council

Science & Technology Facilities Council