# AstroStatistics & AstroInformatics

## in the context of the SKA and LSST

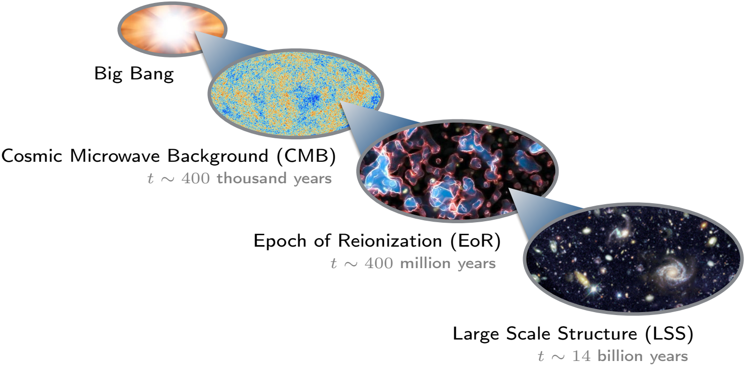
Jason McEwen
www.jasonmcewen.org
@jasonmcewen

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

AI for CERN and SKA, Alan Turing Institute

September 2018

# Unanswered fundamental questions

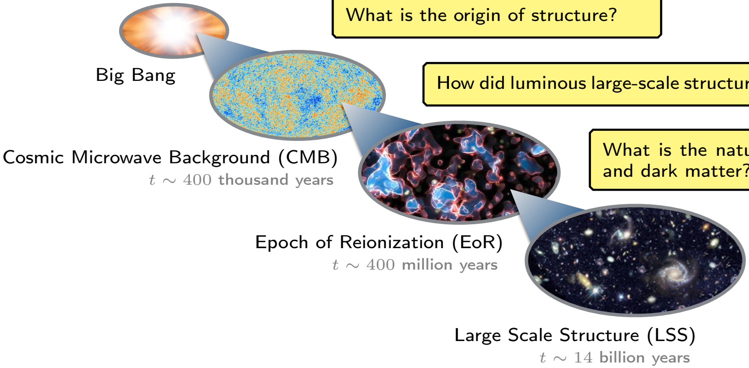


# Unanswered fundamental questions

What is the origin of structure?

How did luminous large-scale structure form?

What is the nature of dark energy and dark matter?

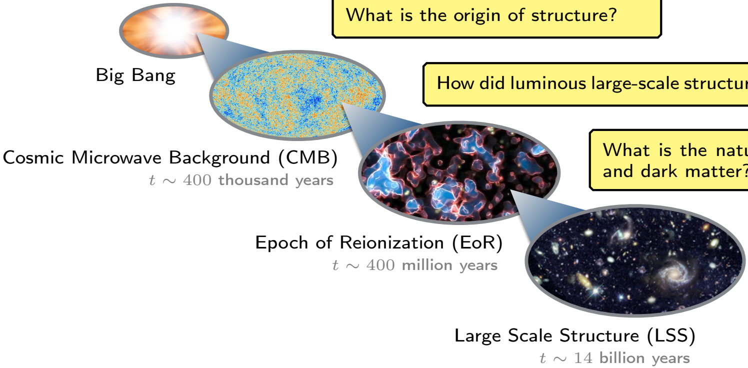


# Unanswered fundamental questions

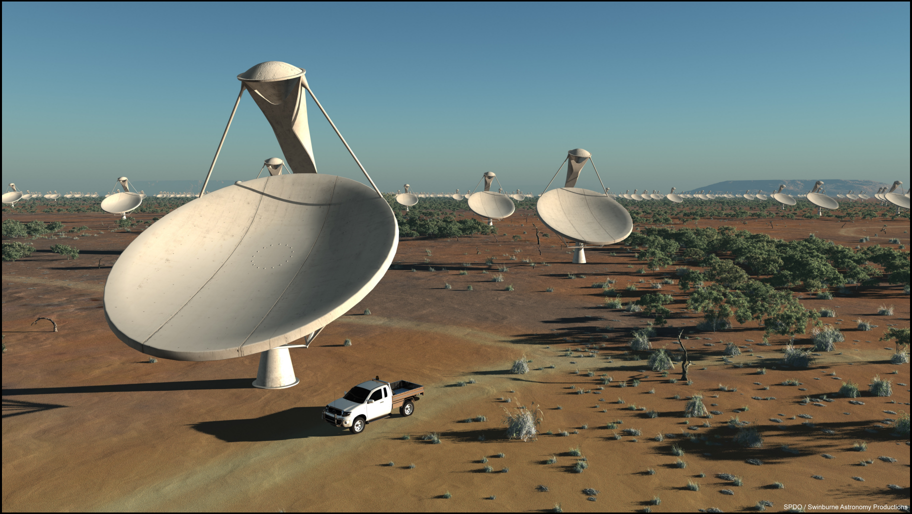
What is the origin of structure?

How did luminous large-scale structure form?

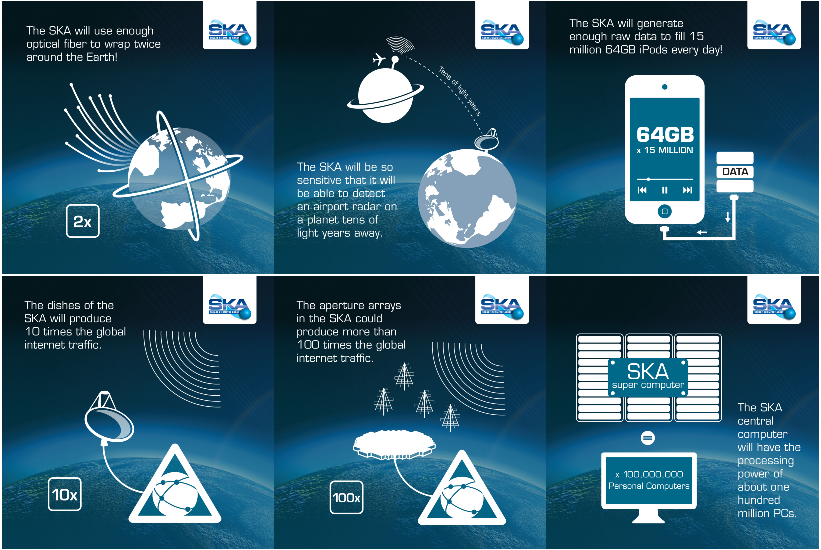
What is the nature of dark energy and dark matter?




# Square Kilometre Array (SKA)

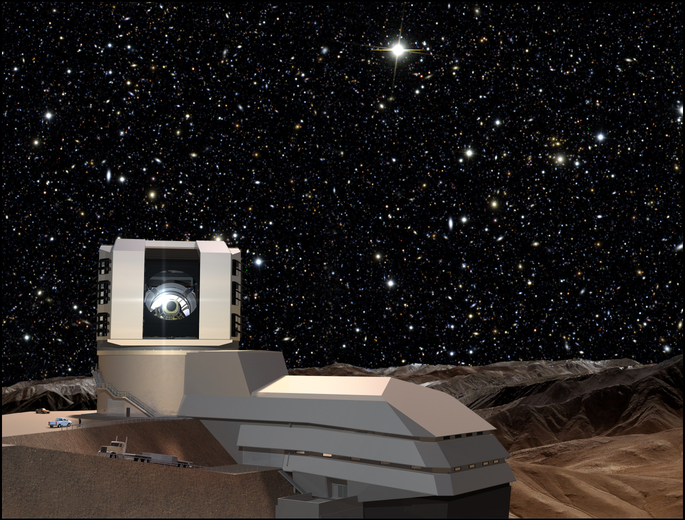


# The SKA poses a considerable big-data challenge



# The SKA poses a considerable big-data challenge

# Large Synoptic Survey Telescope (LSST)



Credit: LSST

# Large Synoptic Survey Telescope (LSST)

**Data Releases:**

Number of Data Releases = 11

Date of DR1 release = Date of Operations Start+ 12 months

Estimated numbers for DR-1 release

- Objects = 18 billion
- Sources = 350 billion (single epoch)
- Forced Sources = 0.75 trillion

Estimated numbers for DR-11

- Objects = 37 billion
- Sources = 7 trillion (single epoch)
- Forced Sources = 30 trillion
- Visits observed = 2.75 million
- Images collected = 5.5 million

**Alert Production:**

Real-time alert latency = 60 seconds

Average number of alerts per night= "about 10 million"

**Data and compute sizes:**

Final image collection (DR11) = 0.5 Exabytes

Final database size (DR11) = 15 PB

Final disk storage = 0.4 Exabytes

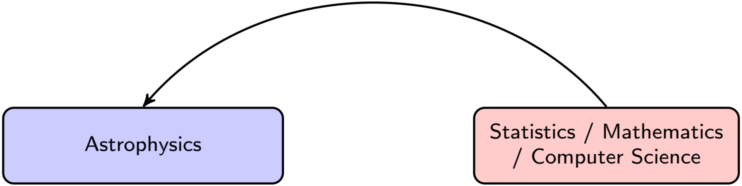
Peak number of nodes = 1750 nodes

Peak compute power in LSST data centers = 1.8 PFLOPS

# Astrostatistics & Astroinformatics

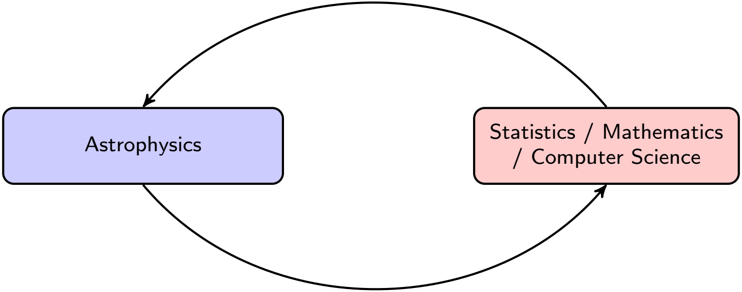
## Closing the loop

*Extracting weak observational signatures of fundamental physics from complex data-sets requires sensitive, robust and principled analysis techniques.*



# Astrostatistics & Astroinformatics
## Closing the loop

*Extracting weak observational signatures of fundamental physics from complex data-sets requires sensitive, robust and principled analysis techniques.*




*Constructing appropriate analysis techniques requires a deep understanding of cosmological problems and methodological foundations.*

# UCL Centre for Doctoral Training (CDT) in Data Intensive Science (DIS)

- UCL STFC CDT focused on Data Intensive Science (DIS), *i.e.* Data Science for Science.
  https://www.hep.ucl.ac.uk/cdt-dis/
- Aims:
  - Train next generation of leaders in the field of DIS (in both academic and industry).
  - Promote development and application of novel DIS techniques.
  - Promote knowledge transfer:
    - between academic fields;
    - between non-academic and academic organisations.
- Unique opportunity to bring together DIS research from perspective of applications, methodologies, and theoretical foundations.

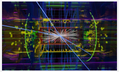
UCL

Science & Technology
Facilities Council

# UCL Centre for Doctoral Training (CDT) in Data Intensive Science (DIS)

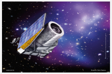
## Who we are

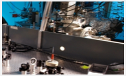
**Particle Physics**
**Dpt. of Physics and Astronomy**
**(20 CDT Staff Members)**



**Astrophysics**
**Dpt. of Physics and Astronomy**
**(20 CDT Staff Members)**

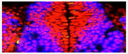
**Department of**
**Space and Climate Science**
**(20 CDT Staff Members)**

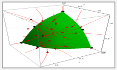




**Atomic & Molecular Physics**
**Dpt. of Physics and Astronomy**
**(2 CDT Staff Members)**

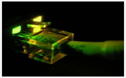
**Department of**
**Computer Science**
**(8 CDT Staff Members)**

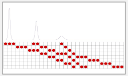




**Department of**
**Mathematics**
**(9 CDT Staff Members)**

**Department of**
**Electrical Engineering**
**(3 CDT Staff Members)**





**Department of**
**Statistical Science**
**(5 CDT Staff Members)**

Aim to foster closer collaboration between these areas to aid the development of novel DIS techniques or applications to new areas.

# UCL Centre for Doctoral Training (CDT) in Data Intensive Science (DIS)

## Industrial partners

**IT Companies**

**Data Intensive Companies**

**Public Sector Organisations**

**Non-Academic Research**

**Public-Private Partnerships**

- Students will undertake 6 month internships with partners on a DIS project
- Promote knowledge transfer between academic and non-academic organisations.
- More organisations joining since winning the bid.

# Outline

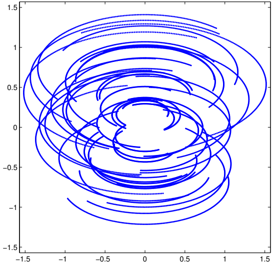
1. Distributed and parallelised algorithms
2. Online algorithms
3. Uncertainty quantification
4. Machine learning

# Outline

1. Distributed and parallelised algorithms
2. Online algorithms
3. Uncertainty quantification
4. Machine learning

# Radio interferometric telescopes acquire "Fourier" measurements

"Fourier"
Measurements



# Radio interferometric inverse problem

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

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

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

- Measurement operator, *e.g.* $\boxed{\boldsymbol{\Phi} = \mathbf{GFA}}$, may incorporate:
  - primary beam $\mathbf{A}$ of the telescope;
  - Fourier transform $\mathbf{F}$;
  - convolutional de-gridding $\mathbf{G}$ to interpolate to continuous $uv$-coordinates;
  - direction-dependent effects (DDEs)...

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:

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

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

- Measurement operator, *e.g.* $\boxed{\mathbf{\Phi} = \mathbf{GFA}}$, may incorporate:
  - primary beam $\mathbf{A}$ of the telescope;
  - Fourier transform $\mathbf{F}$;
  - convolutional de-gridding $\mathbf{G}$ to interpolate to continuous $uv$-coordinates;
  - direction-dependent effects (DDEs)...

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:

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

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

- Measurement operator, *e.g.* $\boxed{\boldsymbol{\Phi} = \mathbf{GFA}}$, may incorporate:
  - primary beam **A** of the telescope;
  - Fourier transform **F**;
  - convolutional de-gridding **G** to interpolate to continuous $uv$-coordinates;
  - direction-dependent effects (DDEs)...

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

# Sparse regularisation

## Motivated by compressive sensing

- Sparse synthesis regularisation problem:

$$\boxed{x_{\text{synthesis}} = \boldsymbol{\Psi} \times \underset{\boldsymbol{\alpha}}{\arg\min} \left[ \left\| \boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{\alpha} \right\|_2^2 + \lambda \left\| \boldsymbol{\alpha} \right\|_1 \right]}$$

Synthesis framework

where consider sparsifying (*e.g.* wavelet) representation of image: $\boxed{\boldsymbol{x} = \boldsymbol{\Psi}\boldsymbol{\alpha}}$ .

- Sparse analysis regularisation problem (Elad *et al.* 2007, Nam *et al.* 2012):

$$\boxed{x_{\text{analysis}} = \underset{\boldsymbol{x}}{\arg\min} \left[ \left\| \boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{x} \right\|_2^2 + \lambda \left\| \boldsymbol{\Psi}^\dagger \boldsymbol{x} \right\|_1 \right]}$$

Analysis framework

- Sparsity averaging reweighted analysis (SARA) (Carrillo, McEwen & Wiaux 2012; Carrillo, McEwen, Van De Ville, Thiran & Wiaux 2013) with overcomplete dictionary:

$$\boxed{\boldsymbol{\Psi} = \left[ \boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2, \ldots, \boldsymbol{\Psi}_q \right]}$$

# Sparse regularisation

## Motivated by compressive sensing

- Sparse synthesis regularisation problem:

$$\boxed{\boldsymbol{x}_{\text{synthesis}} = \boldsymbol{\Psi} \times \arg\min_{\boldsymbol{\alpha}} \left[ \left\| \boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{\alpha} \right\|_2^2 + \lambda \left\| \boldsymbol{\alpha} \right\|_1 \right]}$$

Synthesis framework

where consider sparsifying (*e.g.* wavelet) representation of image: $\boxed{\boldsymbol{x} = \boldsymbol{\Psi}\boldsymbol{\alpha}}$ .

- Sparse analysis regularisation problem (Elad *et al.* 2007, Nam *et al.* 2012):

$$\boxed{\boldsymbol{x}_{\text{analysis}} = \arg\min_{\boldsymbol{x}} \left[ \left\| \boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{x} \right\|_2^2 + \lambda \left\| \boldsymbol{\Psi}^\dagger \boldsymbol{x} \right\|_1 \right]}$$

Analysis framework

- Sparsity averaging reweighted analysis (SARA) (Carrillo, McEwen & Wiaux 2012; Carrillo, McEwen, Van De Ville, Thiran & Wiaux 2013) with overcomplete dictionary:

$$\boxed{\boldsymbol{\Psi} = \left[\boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2, \ldots, \boldsymbol{\Psi}_q\right]}$$

# Sparse regularisation

## Motivated by compressive sensing

- Sparse synthesis regularisation problem:

$$\boxed{\boldsymbol{x}_{\text{synthesis}} = \boldsymbol{\Psi} \times \underset{\boldsymbol{\alpha}}{\arg\min} \left[ \left\| \boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{\alpha} \right\|_2^2 + \lambda \left\| \boldsymbol{\alpha} \right\|_1 \right]}$$

Synthesis framework

where consider sparsifying (*e.g.* wavelet) representation of image: $\boxed{\boldsymbol{x} = \boldsymbol{\Psi}\boldsymbol{\alpha}}$ .

- Sparse analysis regularisation problem (Elad *et al.* 2007, Nam *et al.* 2012):

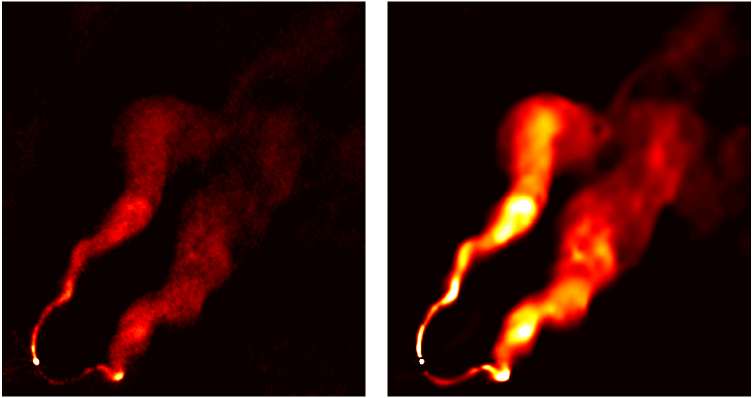
$$\boxed{\boldsymbol{x}_{\text{analysis}} = \underset{\boldsymbol{x}}{\arg\min} \left[ \left\| \boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{x} \right\|_2^2 + \lambda \left\| \boldsymbol{\Psi}^\dagger \boldsymbol{x} \right\|_1 \right]}$$

Analysis framework

- Sparsity averaging reweighted analysis (SARA) (Carrillo, McEwen & Wiaux 2012; Carrillo, McEwen, Van De Ville, Thiran & Wiaux 2013) with overcomplete dictionary:

$$\boxed{\boldsymbol{\Psi} = \left[ \boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2, \ldots, \boldsymbol{\Psi}_q \right]}$$

# Reconstruction
## VLA observation of 3C129



(a) CLEAN (uniform) (b) PURIFY

Figure: 3C129 recovered images (Pratley, McEwen, *et al.* 2016)

# Distributed and parallelised algorithms

- Solve resulting convex optimisation problems by proximal splitting.
- Block inexact ADMM algorithm to split data and measurement operator:
  (Carrillo, McEwen & Wiaux 2014; Onose, Carrillo, Repetti, McEwen, Thiran, Pesquet, & Wiaux 2016; Pratley, Johnston-Hollitt & McEwen 2018; Pratley, McEwen *et al.* in prep.)

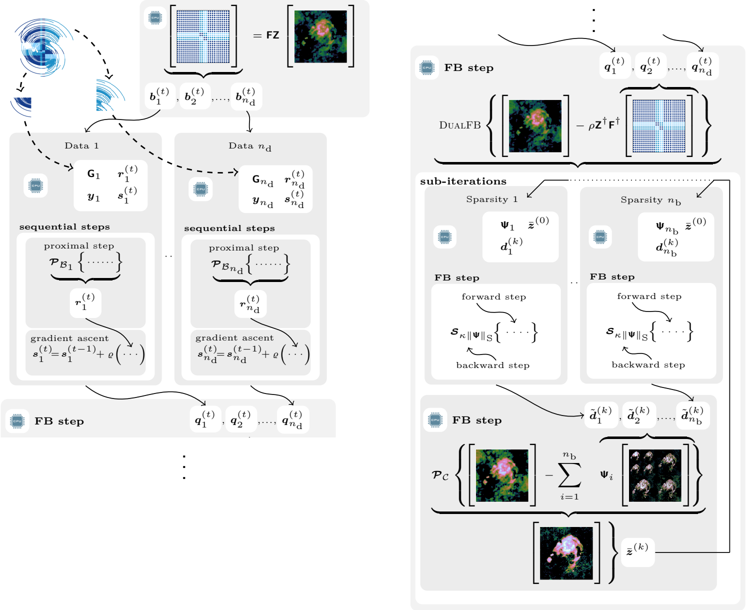
$$\boxed{\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_{n_\mathrm{d}} \end{bmatrix}}, \quad \boxed{\boldsymbol{\Phi} = \begin{bmatrix} \boldsymbol{\Phi}_1 \\ \vdots \\ \boldsymbol{\Phi}_{n_\mathrm{d}} \end{bmatrix} = \begin{bmatrix} \mathbf{G}_1\mathbf{M}_1 \\ \vdots \\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

# Distributed and parallelised algorithms

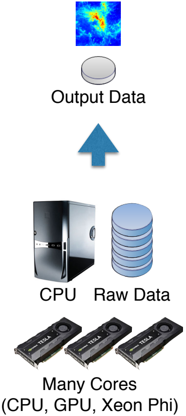
- Solve resulting convex optimisation problems by proximal splitting.
- Block inexact ADMM algorithm to split data and measurement operator:
  (Carrillo, McEwen & Wiaux 2014; Onose, Carrillo, Repetti, McEwen, Thiran, Pesquet, & Wiaux 2016; Pratley, Johnston-Hollitt & McEwen 2018; Pratley, McEwen *et al.* in prep.)

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

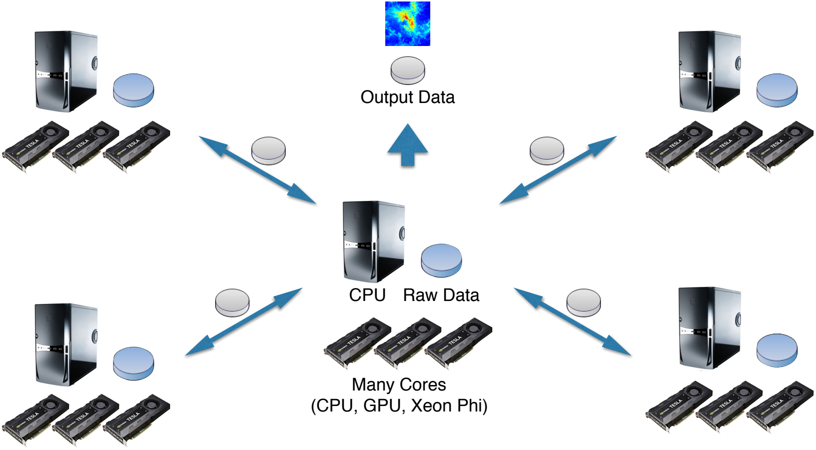
# Distributed and parallelised convex optimisation



# Standard algorithms



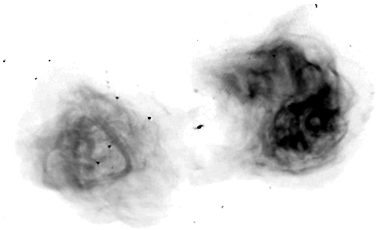
# Highly distributed and parallelised algorithms



# Highly distributed and parallelised algorithms
## Reconstruction

- Hybrid $w$-stacking and $w$-projection distributed and parallelised reconstruction (Pratley, Johnston-Hollitt & McEwen 2018)
  - 100 millions visibilities (measurements)
  - $4096 \times 4096$ pixel image ($\sim$17 million pixels)
  - $17^\circ$ field of view
  - $w$-terms of $\pm 300$ wavelengths (to account for wide fields)

Imaging with exact wide-field corrections for 100 million visibilities in 30 minutes.



# Public open-source codes

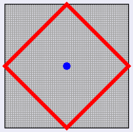
## PURIFY code http://basp-group.github.io/purify/

*Next-generation radio interferometric imaging*

Carrillo, McEwen, Wiaux, Pratley, d'Avezac

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.

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



*Sparse OPTimisation*

Carrillo, McEwen, Wiaux, Kartik, d'Avezac, Pratley, Perez-Suarez

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

# Outline

# Online algorithms

- Many standard astrophysical data analyses are performed offline.
- Data are acquired. . . and then analysed.
- Will not necessarily be possible in future.

# Online radio interferometric imaging

- Online radio interferometric imaging: assimilating and discarding visibilities on arrival (Cai, Pratley, McEwen 2018)

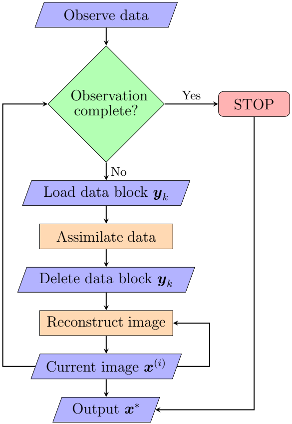



Figure: Schematic of online imaging algorithm.

# Online radio interferometric imaging

- Data storage requirements reduced dramatically.
- Computational costs can also be reduced.

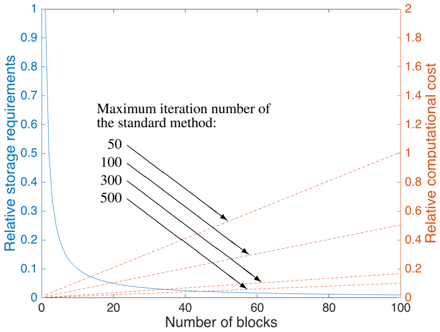



Figure: Storage and computational costs.

# Online radio interferometric imaging

- Theoretical guarantees that recover images of same fidelity as offline approach.

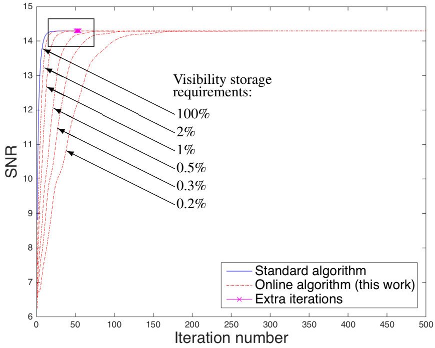
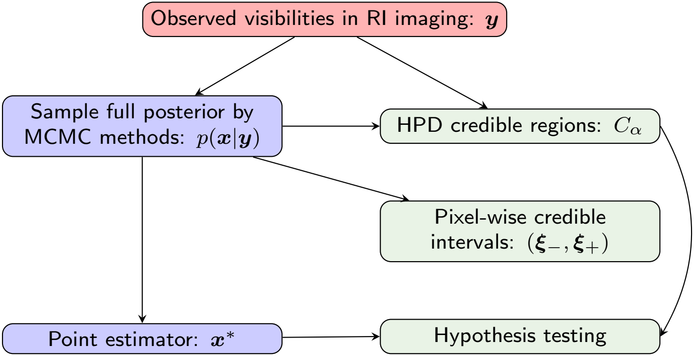


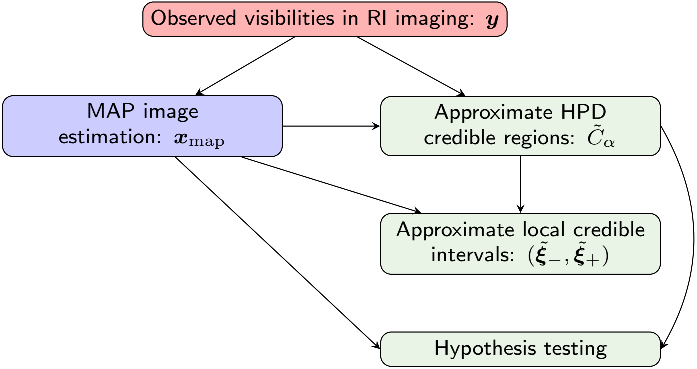

Figure: Reconstruction fidelity vs iteration number.

# Outline

# *Proximal MCMC sampling* and uncertainty quantification



## *MAP estimation* and uncertainty quantification



# Approximate Bayesian credible regions for MAP estimation

- Combine uncertainty quantification with fast sparse regularisation to scale to big-data.
- Recall $C_\alpha$ denotes the highest posterior density (HPD) Bayesian credible region with confidence level $(1-\alpha)\%$ defined by posterior iso-contour: $C_\alpha = \{\boldsymbol{x} : g(\boldsymbol{x}) \leq \gamma_\alpha\}$.
- Analytic approximation of $\gamma_\alpha$:

$$\boxed{\tilde{\gamma}_\alpha = g(\boldsymbol{x}^\star) + N(\tau_\alpha + 1)}$$

where $\tau_\alpha = \sqrt{16\log(3/\alpha)/N}$ and $\alpha \in (4\exp(-N/3), 1)$ (Pereyra 2016b).

- Define approximate HPD regions by $\tilde{C}_\alpha = \{\boldsymbol{x} : g(\boldsymbol{x}) \leq \tilde{\gamma}_\alpha\}$.
- Compute $\boldsymbol{x}^\star$ by sparse regularisation, then estimate local Bayesian credible intervals and perform hypothesis testing using approximate HPD regions.

# Approximate Bayesian credible regions for MAP estimation

- Combine uncertainty quantification with fast sparse regularisation to scale to big-data.
- Recall $C_\alpha$ denotes the highest posterior density (HPD) Bayesian credible region with confidence level $(1-\alpha)\%$ defined by posterior iso-contour: $C_\alpha = \{\boldsymbol{x} : g(\boldsymbol{x}) \leq \gamma_\alpha\}$.
- Analytic approximation of $\gamma_\alpha$:

$$\boxed{\tilde{\gamma}_\alpha = g(\boldsymbol{x}^\star) + N(\tau_\alpha + 1)}$$

where $\tau_\alpha = \sqrt{16 \log(3/\alpha)/N}$ and $\alpha \in (4\exp(-N/3), 1)$ (Pereyra 2016b).

- Define approximate HPD regions by $\tilde{C}_\alpha = \{\boldsymbol{x} : g(\boldsymbol{x}) \leq \tilde{\gamma}_\alpha\}$.
- Compute $x^\star$ by sparse regularisation, then estimate local Bayesian credible intervals and perform hypothesis testing using approximate HPD regions.

# Approximate Bayesian credible regions for MAP estimation

- Combine uncertainty quantification with fast sparse regularisation to scale to big-data.
- Recall $C_\alpha$ denotes the highest posterior density (HPD) Bayesian credible region with confidence level $(1-\alpha)\%$ defined by posterior iso-contour: $C_\alpha = \{\boldsymbol{x} : g(\boldsymbol{x}) \leq \gamma_\alpha\}$.
- Analytic approximation of $\gamma_\alpha$:

$$\boxed{\tilde{\gamma}_\alpha = g(\boldsymbol{x}^\star) + N(\tau_\alpha + 1)}$$

where $\tau_\alpha = \sqrt{16\log(3/\alpha)/N}$ and $\alpha \in (4\exp(-N/3), 1)$ (Pereyra 2016b).

- Define approximate HPD regions by $\tilde{C}_\alpha = \{\boldsymbol{x} : g(\boldsymbol{x}) \leq \tilde{\gamma}_\alpha\}$.
- Compute $\boldsymbol{x}^\star$ by sparse regularisation, then estimate local Bayesian credible intervals and perform hypothesis testing using approximate HPD regions.

# Local Bayesian credible intervals for MAP estimation

## Local Bayesian credible intervals for sparse reconstruction
(Cai, Pereyra & McEwen 2017b)

Let $\Omega$ define the area (or pixel) over which to compute the credible interval $(\tilde{\xi}_-, \tilde{\xi}_+)$ and $\boldsymbol{\zeta}$ be an index vector describing $\Omega$ (*i.e.* $\zeta_i = 1$ if $i \in \Omega$ and 0 otherwise).

Consider the test image with the $\Omega$ region replaced by constant value $\xi$:

$$\boxed{\boldsymbol{x}' = \boldsymbol{x}^\star(\mathcal{I} - \boldsymbol{\zeta}) + \xi\boldsymbol{\zeta}} \, .$$

Given $\tilde{\gamma}_\alpha$ and $\boldsymbol{x}^\star$, compute the credible interval by

$$\boxed{\begin{aligned} \tilde{\xi}_- &= \min_{\xi} \left\{ \xi \mid g_{\boldsymbol{y}}(\boldsymbol{x}') \leq \tilde{\gamma}_\alpha, \ \forall \xi \in [-\infty, +\infty) \right\}, \\ \tilde{\xi}_+ &= \max_{\xi} \left\{ \xi \mid g_{\boldsymbol{y}}(\boldsymbol{x}') \leq \tilde{\gamma}_\alpha, \ \forall \xi \in [-\infty, +\infty) \right\}. \end{aligned}}$$

# Local Bayesian credible intervals for MAP estimation

## Local Bayesian credible intervals for sparse reconstruction

(Cai, Pereyra & McEwen 2017b)

Let $\Omega$ define the area (or pixel) over which to compute the credible interval $(\tilde{\xi}_-, \tilde{\xi}_+)$ and $\boldsymbol{\zeta}$ be an index vector describing $\Omega$ (*i.e.* $\zeta_i = 1$ if $i \in \Omega$ and 0 otherwise).

Consider the test image with the $\Omega$ region replaced by constant value $\xi$:

$$\boxed{\boldsymbol{x}' = \boldsymbol{x}^\star(\mathcal{I} - \boldsymbol{\zeta}) + \xi\boldsymbol{\zeta}} \ .$$

Given $\tilde{\gamma}_\alpha$ and $\boldsymbol{x}^\star$, compute the credible interval by

$$\boxed{\begin{aligned} \tilde{\xi}_- &= \min_\xi \left\{ \xi \mid g_{\boldsymbol{y}}(\boldsymbol{x}') \leq \tilde{\gamma}_\alpha, \ \forall \xi \in [-\infty, +\infty) \right\}, \\ \tilde{\xi}_+ &= \max_\xi \left\{ \xi \mid g_{\boldsymbol{y}}(\boldsymbol{x}') \leq \tilde{\gamma}_\alpha, \ \forall \xi \in [-\infty, +\infty) \right\}. \end{aligned}}$$

# Local Bayesian credible intervals for MAP estimation

## Local Bayesian credible intervals for sparse reconstruction

(Cai, Pereyra & McEwen 2017b)

Let $\Omega$ define the area (or pixel) over which to compute the credible interval $(\tilde{\xi}_-, \tilde{\xi}_+)$ and $\boldsymbol{\zeta}$ be an index vector describing $\Omega$ (*i.e.* $\zeta_i = 1$ if $i \in \Omega$ and 0 otherwise).

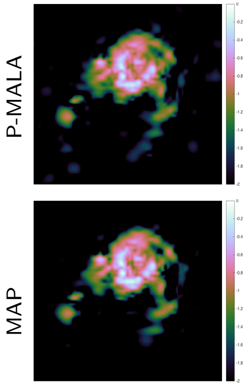
Consider the test image with the $\Omega$ region replaced by constant value $\xi$:

$$\boxed{\boldsymbol{x}' = \boldsymbol{x}^\star(\mathcal{I} - \boldsymbol{\zeta}) + \xi\boldsymbol{\zeta}}\ .$$

Given $\tilde{\gamma}_\alpha$ and $\boldsymbol{x}^\star$, compute the credible interval by

$$\boxed{\begin{aligned}\tilde{\xi}_- &= \min_{\xi}\left\{\xi \mid g_{\boldsymbol{y}}(\boldsymbol{x}') \leq \tilde{\gamma}_\alpha,\ \forall \xi \in [-\infty, +\infty)\right\},\\ \tilde{\xi}_+ &= \max_{\xi}\left\{\xi \mid g_{\boldsymbol{y}}(\boldsymbol{x}') \leq \tilde{\gamma}_\alpha,\ \forall \xi \in [-\infty, +\infty)\right\}.\end{aligned}}$$

# Numerical experiments



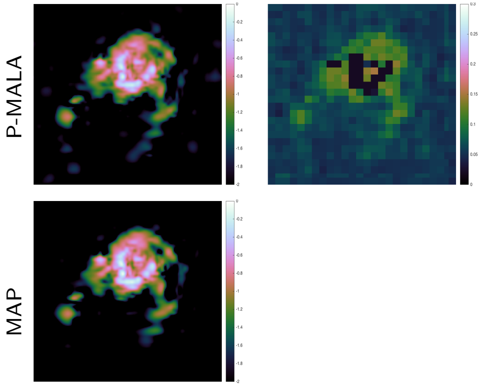

(a) point estimators

(b) local credible interval (grid size $10 \times 10$ pixels)

(c) local credible interval (grid size $20 \times 20$ pixels)

Figure: Length of local credible intervals for M31 for the analysis model.

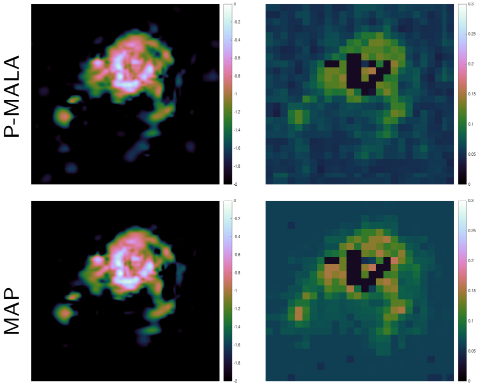
# Numerical experiments




(a) point estimators (b) local credible interval (grid size $10 \times 10$ pixels) (c) local credible interval (grid size $20 \times 20$ pixels)

Figure: Length of local credible intervals for M31 for the analysis model.

# Numerical experiments



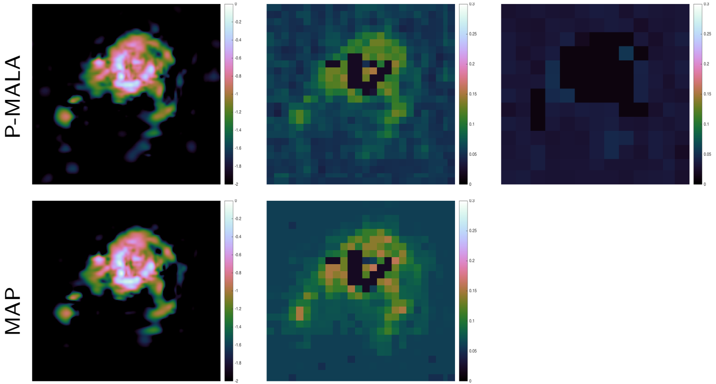

(a) point estimators

(b) local credible interval (grid size $10 \times 10$ pixels)

(c) local credible interval (grid size $20 \times 20$ pixels)

Figure: Length of local credible intervals for M31 for the analysis model.

# Numerical experiments



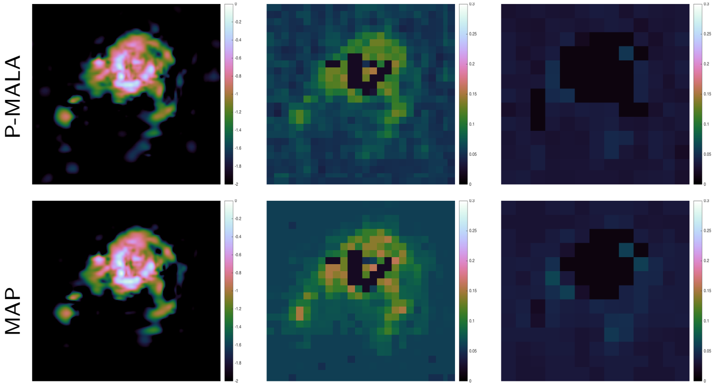

(a) point estimators

(b) local credible interval (grid size $10 \times 10$ pixels)

(c) local credible interval (grid size $20 \times 20$ pixels)

Figure: Length of local credible intervals for M31 for the analysis model.

# Numerical experiments



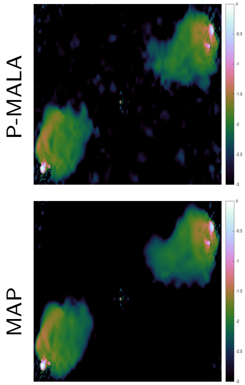

(a) point estimators

(b) local credible interval (grid size $10 \times 10$ pixels)

(c) local credible interval (grid size $20 \times 20$ pixels)

Figure: Length of local credible intervals for M31 for the analysis model.

# Numerical experiments



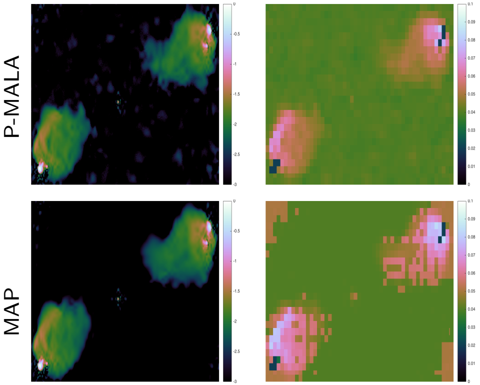

(a) point estimators

(b) local credible interval (grid size $10 \times 10$ pixels)

(c) local credible interval (grid size $20 \times 20$ pixels)

Figure: Length of local credible intervals for Cygnus A for the analysis model.

# Numerical experiments



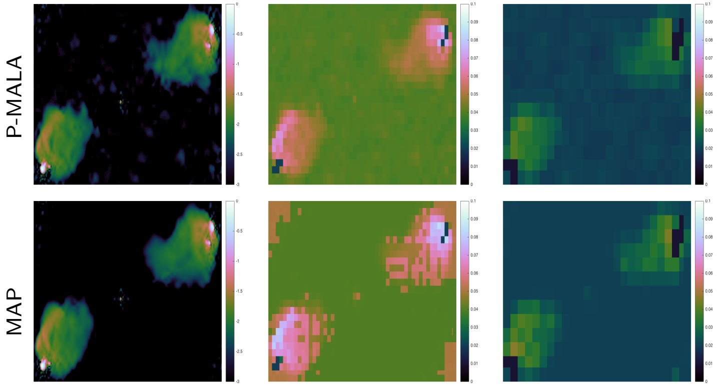

(a) point estimators

(b) local credible interval (grid size $10 \times 10$ pixels)

(c) local credible interval (grid size $20 \times 20$ pixels)

Figure: Length of local credible intervals for Cygnus A for the analysis model.

# Numerical experiments



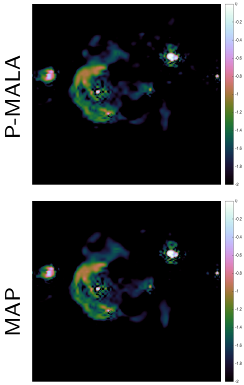

(a) point estimators

(b) local credible interval (grid size $10 \times 10$ pixels)

(c) local credible interval (grid size $20 \times 20$ pixels)

Figure: Length of local credible intervals for Cygnus A for the analysis model.

# Numerical experiments



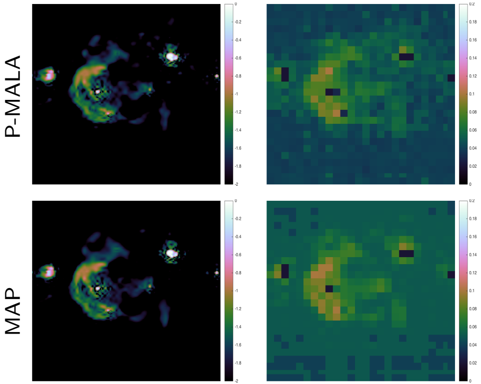

(a) point estimators

(b) local credible interval (grid size $10 \times 10$ pixels)

(c) local credible interval (grid size $20 \times 20$ pixels)

Figure: Length of local credible intervals for W28 for the analysis model.

# Numerical experiments



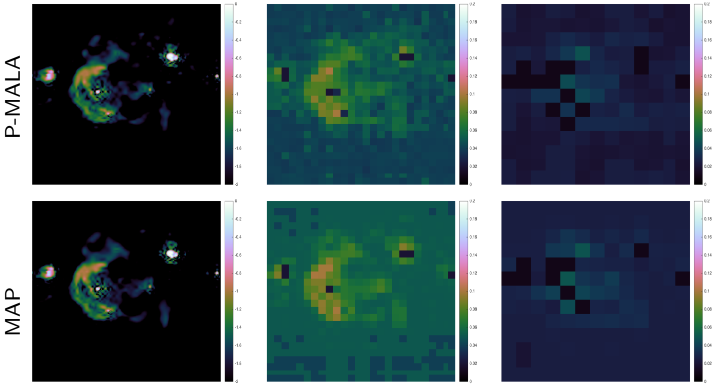

(a) point estimators

(b) local credible interval (grid size $10 \times 10$ pixels)

(c) local credible interval (grid size $20 \times 20$ pixels)

Figure: Length of local credible intervals for W28 for the analysis model.

# Numerical experiments



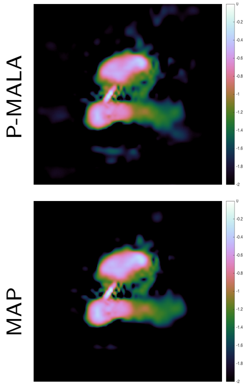

(a) point estimators

(b) local credible interval (grid size $10 \times 10$ pixels)

(c) local credible interval (grid size $20 \times 20$ pixels)

Figure: Length of local credible intervals for W28 for the analysis model.

# Numerical experiments



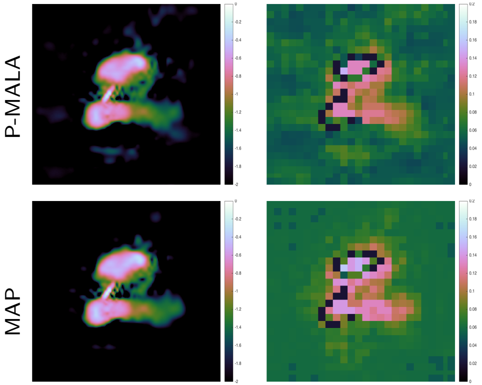

(a) point estimators

(b) local credible interval (grid size $10 \times 10$ pixels)

(c) local credible interval (grid size $20 \times 20$ pixels)

Figure: Length of local credible intervals for 3C288 for the analysis model.

# Numerical experiments



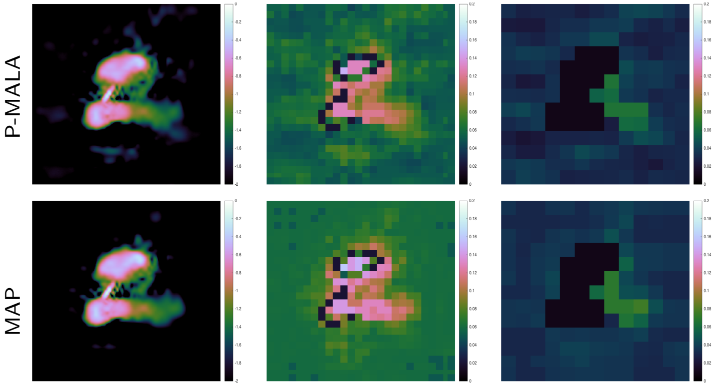

(a) point estimators

(b) local credible interval (grid size $10 \times 10$ pixels)

(c) local credible interval (grid size $20 \times 20$ pixels)

Figure: Length of local credible intervals for 3C288 for the analysis model.

# Numerical experiments




(a) point estimators

(b) local credible interval (grid size $10 \times 10$ pixels)

(c) local credible interval (grid size $20 \times 20$ pixels)

Figure: Length of local credible intervals for 3C288 for the analysis model.

# Computation time

Table: CPU time in minutes for Proximal MCMC sampling and MAP estimation

| Image | Method | CPU time Analysis | CPU time Synthesis |
|---|---|---|---|
| Cygnus A | P-MALA | 2274 | 1762 |
| | MYULA | 1056 | 942 |
| | MAP | .07 | .04 |
| M31 | P-MALA | 1307 | 944 |
| | MYULA | 618 | 581 |
| | MAP | .03 | .02 |
| W28 | P-MALA | 1122 | 879 |
| | MYULA | 646 | 598 |
| | MAP | .06 | .04 |
| 3C288 | P-MALA | 1144 | 881 |
| | MYULA | 607 | 538 |
| | MAP | .03 | .02 |

# Hypothesis testing
## Method

- Perform hypothesis tests of image structure using Bayesian credible regions (Pereyra 2016b).
- Let $C_\alpha$ denote the highest posterior density (HPD) Bayesian credible region with confidence level $(1-\alpha)\%$ defined by posterior iso-contour: $C_\alpha = \{\boldsymbol{x} : g(\boldsymbol{x}) \leq \gamma_\alpha\}$.

**Hypothesis testing of physical structure**

1. Remove structure of interest from recovered image $\boldsymbol{x}^\star$.
2. Inpaint background (noise) into region, yielding surrogate image $\boldsymbol{x}'$.
3. Test whether $\boldsymbol{x}' \in C_\alpha$:
   - If $\boldsymbol{x}' \notin C_\alpha$ then reject hypothesis that structure is an artifact with confidence $(1-\alpha)\%$, i.e.
   - If $\boldsymbol{x}' \in C_\alpha$ uncertainly too high to draw strong conclusions about the physical nature of the structure.

# Hypothesis testing

## Method

- Perform hypothesis tests of image structure using Bayesian credible regions (Pereyra 2016b).
- Let $C_\alpha$ denote the highest posterior density (HPD) Bayesian credible region with confidence level $(1-\alpha)\%$ defined by posterior iso-contour: $C_\alpha = \{\boldsymbol{x} : g(\boldsymbol{x}) \leq \gamma_\alpha\}$.

**Hypothesis testing of physical structure**

1. Remove structure of interest from recovered image $x^\star$.
2. Inpaint background (noise) into region, yielding surrogate image $x'$.
3. Test whether $x' \in C_\alpha$:
   - If $x' \notin C_\alpha$ then reject hypothesis that structure is an artifact with confidence $(1-\alpha)\%$, i.e.
   - If $x' \in C_\alpha$ uncertainty too high to draw strong conclusions about the physical nature of the structure.

# Hypothesis testing
## Method

- Perform hypothesis tests of image structure using Bayesian credible regions (Pereyra 2016b).
- Let $C_\alpha$ denote the highest posterior density (HPD) Bayesian credible region with confidence level $(1-\alpha)\%$ defined by posterior iso-contour: $C_\alpha = \{\boldsymbol{x} : g(\boldsymbol{x}) \leq \gamma_\alpha\}$.

### Hypothesis testing of physical structure

1. Remove structure of interest from recovered image $\boldsymbol{x}^\star$.
2. Inpaint background (noise) into region, yielding surrogate image $\boldsymbol{x}'$.
3. Test whether $\boldsymbol{x}' \in C_\alpha$:
   - If $\boldsymbol{x}' \notin C_\alpha$ then reject hypothesis that structure is an artifact with confidence $(1-\alpha)\%$, *i.e.* structure most likely physical.
   - If $\boldsymbol{x}' \in C_\alpha$ uncertainly too high to draw strong conclusions about the physical nature of the structure.

# Hypothesis testing

## Method

- Perform hypothesis tests of image structure using Bayesian credible regions (Pereyra 2016b).
- Let $C_\alpha$ denote the highest posterior density (HPD) Bayesian credible region with confidence level $(1-\alpha)\%$ defined by posterior iso-contour: $C_\alpha = \{\boldsymbol{x} : g(\boldsymbol{x}) \leq \gamma_\alpha\}$.

**Hypothesis testing of physical structure**

1. Remove structure of interest from recovered image $\boldsymbol{x}^\star$.
2. Inpaint background (noise) into region, yielding surrogate image $\boldsymbol{x}'$.
3. Test whether $\boldsymbol{x}' \in C_\alpha$:
    - If $\boldsymbol{x}' \notin C_\alpha$ then reject hypothesis that structure is an artifact with confidence $(1-\alpha)\%$, *i.e.* structure most likely physical.
    - If $\boldsymbol{x}' \in C_\alpha$ uncertainly too high to draw strong conclusions about the physical nature of the structure.

# Hypothesis testing
## Method

- Perform hypothesis tests of image structure using Bayesian credible regions (Pereyra 2016b).
- Let $C_\alpha$ denote the highest posterior density (HPD) Bayesian credible region with confidence level $(1-\alpha)\%$ defined by posterior iso-contour: $C_\alpha = \{\boldsymbol{x} : g(\boldsymbol{x}) \leq \gamma_\alpha\}$.

### Hypothesis testing of physical structure

1. Remove structure of interest from recovered image $\boldsymbol{x}^\star$.
2. Inpaint background (noise) into region, yielding surrogate image $\boldsymbol{x}'$.
3. Test whether $\boldsymbol{x}' \in C_\alpha$:
   - If $\boldsymbol{x}' \notin C_\alpha$ then reject hypothesis that structure is an artifact with confidence $(1-\alpha)\%$, *i.e.* structure most likely physical.
   - If $\boldsymbol{x}' \in C_\alpha$ uncertainly too high to draw strong conclusions about the physical nature of the structure.

# Hypothesis testing
## Method

- Perform hypothesis tests of image structure using Bayesian credible regions (Pereyra 2016b).
- Let $C_\alpha$ denote the highest posterior density (HPD) Bayesian credible region with confidence level $(1-\alpha)\%$ defined by posterior iso-contour: $C_\alpha = \{\boldsymbol{x} : g(\boldsymbol{x}) \leq \gamma_\alpha\}$.

### Hypothesis testing of physical structure

1. Remove structure of interest from recovered image $\boldsymbol{x}^\star$.
2. Inpaint background (noise) into region, yielding surrogate image $\boldsymbol{x}'$.
3. Test whether $\boldsymbol{x}' \in C_\alpha$:
   - If $\boldsymbol{x}' \notin C_\alpha$ then reject hypothesis that structure is an artifact with confidence $(1-\alpha)\%$, *i.e.* structure most likely physical.
   - If $\boldsymbol{x}' \in C_\alpha$ uncertainly too high to draw strong conclusions about the physical nature of the structure.

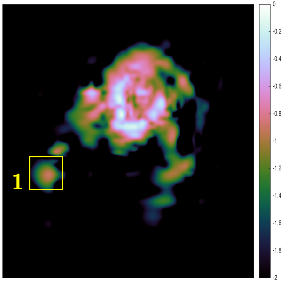
# Hypothesis testing
## Method

- Perform hypothesis tests of image structure using Bayesian credible regions (Pereyra 2016b).
- Let $C_\alpha$ denote the highest posterior density (HPD) Bayesian credible region with confidence level $(1-\alpha)\%$ defined by posterior iso-contour: $C_\alpha = \{\boldsymbol{x} : g(\boldsymbol{x}) \leq \gamma_\alpha\}$.

### Hypothesis testing of physical structure

1. Remove structure of interest from recovered image $\boldsymbol{x}^\star$.
2. Inpaint background (noise) into region, yielding surrogate image $\boldsymbol{x}'$.
3. Test whether $\boldsymbol{x}' \in C_\alpha$:
   - If $\boldsymbol{x}' \notin C_\alpha$ then reject hypothesis that structure is an artifact with confidence $(1-\alpha)\%$, *i.e.* structure most likely physical.
   - If $\boldsymbol{x}' \in C_\alpha$ uncertainly too high to draw strong conclusions about the physical nature of the structure.

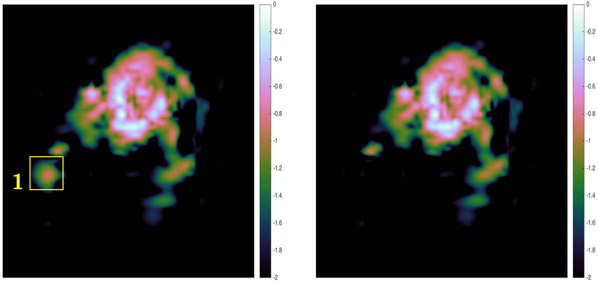
# Hypothesis testing
## Numerical experiments




(a) Recovered image

Figure: HII region of M31

# Hypothesis testing
## Numerical experiments

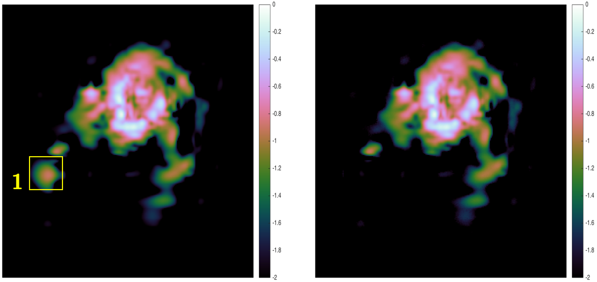



(a) Recovered image (b) Surrogate with region removed

Figure: HII region of M31

# Hypothesis testing

## Numerical experiments




(a) Recovered image

(b) Surrogate with region removed

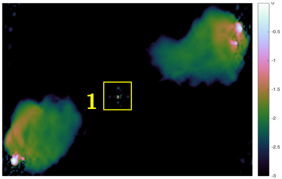
Figure: HII region of M31

1. Reject null hypothesis

   $\Rightarrow$ structure physical

# Hypothesis testing

## Numerical experiments

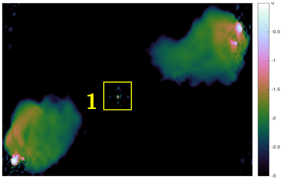



(a) Recovered image

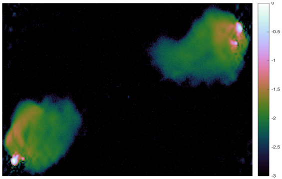
Figure: Cygnus A

# Hypothesis testing

## Numerical experiments




(a) Recovered image

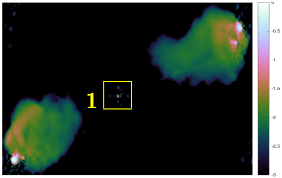



(b) Surrogate with region removed

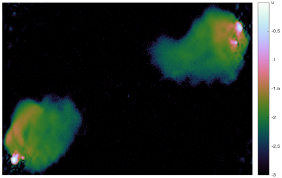
Figure: Cygnus A

# Hypothesis testing

## Numerical experiments




(a) Recovered image



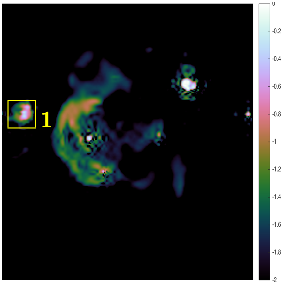

(b) Surrogate with region removed

Figure: Cygnus A

1. Cannot reject null hypothesis

$\Rightarrow$ cannot make strong statistical statement about origin of structure

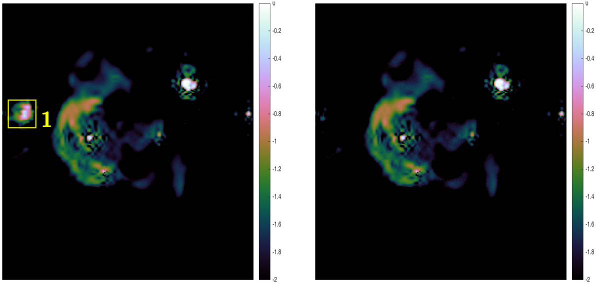
# Hypothesis testing
## Numerical experiments




(a) Recovered image

Figure: Supernova remnant W28

# Hypothesis testing
## Numerical experiments

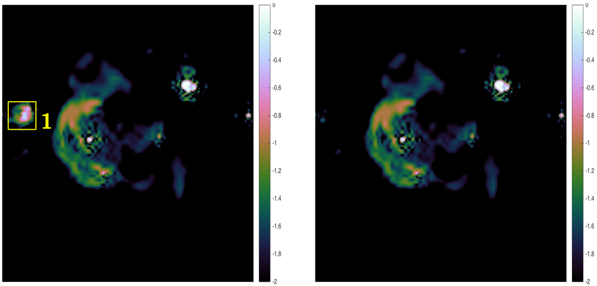



(a) Recovered image

(b) Surrogate with region removed

Figure: Supernova remnant W28

# Hypothesis testing
## Numerical experiments




(a) Recovered image

(b) Surrogate with region removed

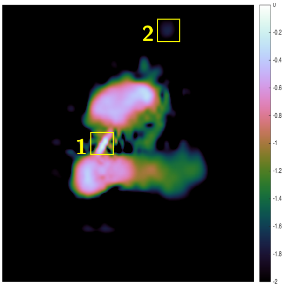
Figure: Supernova remnant W28

1. Reject null hypothesis

$\Rightarrow$ structure physical

# Hypothesis testing

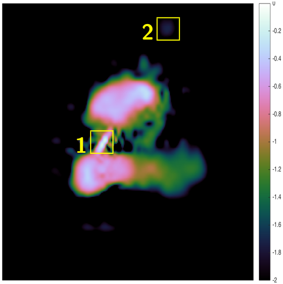
## Numerical experiments



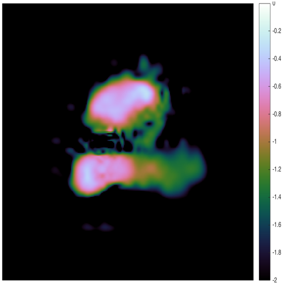

(a) Recovered image

Figure: 3C288

# Hypothesis testing
## Numerical experiments




(a) Recovered image

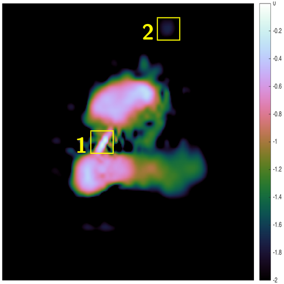



(b) Surrogate with region removed

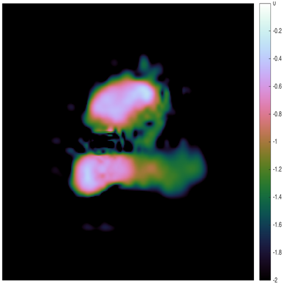
Figure: 3C288

# Hypothesis testing

## Numerical experiments




(a) Recovered image




(b) Surrogate with region removed

Figure: 3C288

1. Reject null hypothesis

   ⇒ structure physical

2. Cannot reject null hypothesis

   ⇒ cannot make strong statistical statement about origin of structure

# Hypothesis testing

## Comparison of numerical experiments

Table: Comparison of hypothesis tests for different methods for the analysis model.

| Image | Test area | Ground truth | Method | Hypothesis test |
|---|---|---|---|---|
| M31 | 1 | ✓ | P-MALA | ✓ |
| | | | MYULA | ✓ |
| | | | MAP | ✓ |
| Cygnus A | 1 | ✓ | P-MALA | ✗ |
| | | | MYULA* | ✗ |
| | | | MAP | ✗ |
| W28 | 1 | ✓ | P-MALA | ✓ |
| | | | MYULA | ✓ |
| | | | MAP | ✓ |
| 3C288 | 1 | ✓ | P-MALA | ✓ |
| | | | MYULA | ✓ |
| | | | MAP | ✓ |
| | 2 | ✗ | P-MALA | ✗ |
| | | | MYULA | ✗ |
| | | | MAP | ✗ |

(* Can correctly detect physical structure if use median point estimator.)

# Outline

# Deep learning methods for radio interferometric imaging

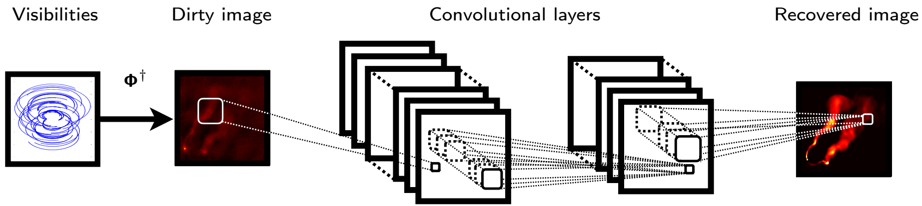



Figure: Deep learning architecture for interferometric imaging (Allam & McEwen, in prep.)

# Deep learning methods for radio interferometric imaging

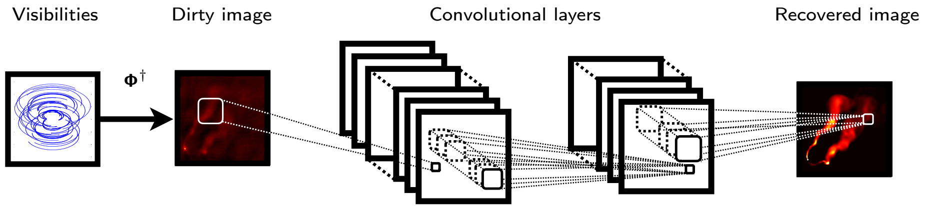



Figure: Deep learning architecture for interferometric imaging (Allam & McEwen, in prep.)

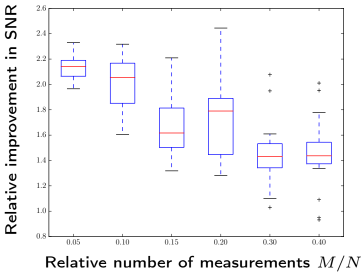



Figure: Improvement in signal-to-noise-ratio (SNR)

# Artist impression of Supernova explosion

## Thermonuclear explosion or core collapse

# Supernova classification

## Spectroscopic classification

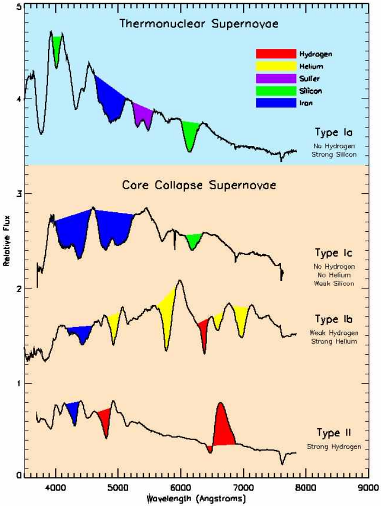



Figure: Spectroscopic observations

# Supernova classification

## Photometric classification

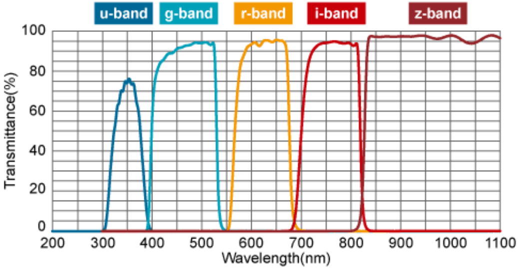
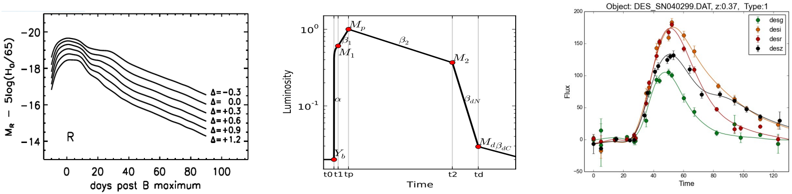



Figure: Photometric observations.

# Supernova classification
## Photometric classification

- Photometric Supernova classification by machine learning (Lochner, McEwen, Peiris, Lahav & Winter 2016)
- Limited training data.
- Go beyond single techniques to study classes.




(a) Templates (b) Generic parameterisations (c) Wavelets (non-parametric)

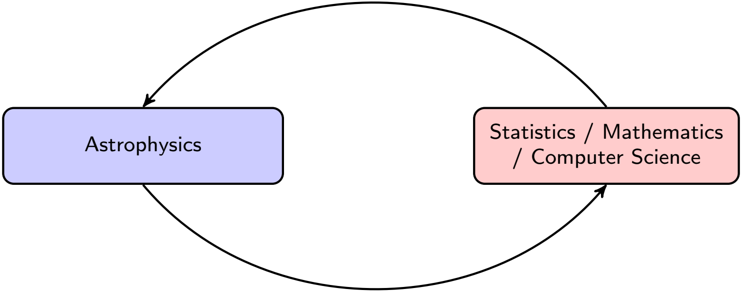
Figure: Feature selection classes (in order of increasing model independence)

- Integrate physics into machine learning (scale and dilation invariance).
- Understand physical requirements: representative training, redshift.

# Astrostatistics & Astroinformatics

## Closing the loop

*Extracting weak observational signatures of fundamental physics from complex data-sets requires sensitive, robust and principled analysis techniques.*




*Constructing appropriate analysis techniques requires a deep understanding of cosmological problems and methodological foundations.*