Sparse image reconstruction for the SPIDER optical interferometric telescope

Jason McEwen
www.jasonmcewen.org

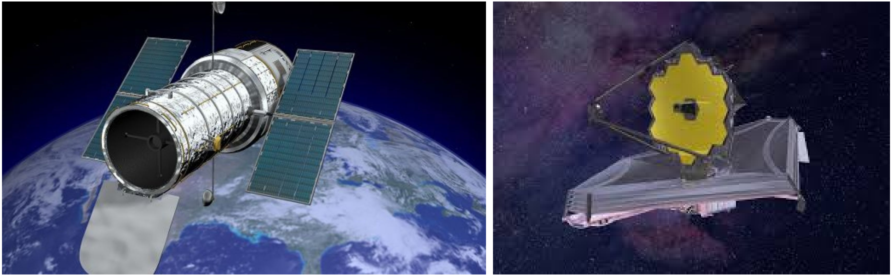
Mullard Space Science Laboratory (MSSL)
University College London (UCL)

Pratley & McEwen (2019): arXiv:1903.05638

UC Davis, June 2019

Optical astronomical telescopes

- Hubble Space Telescope (HST) has transformed our understanding of the Universe.
- Hubble's scientific successor, the James Webb Space Telescope (JWST), will lead to further scientific advances.
- But Hubble and JWST are extremely **large** and **heavy**, and **expensive in cost** and **power** consumption.



(a) Hubble Space Telescope (HST)

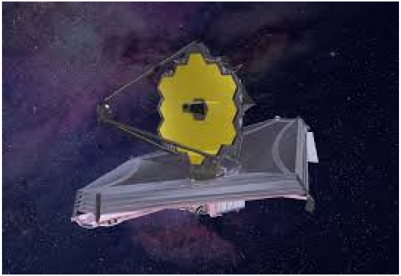
(b) James Web Space Telescope (JWST)

Figure: Optical telescopes

Optical astronomical telescopes

- Hubble Space Telescope (HST) has transformed our understanding of the Universe.
- Hubble's scientific successor, the James Webb Space Telescope (JWST), will lead to further scientific advances.
- But Hubble and JWST are extremely **large** and **heavy**, and **expensive in cost** and **power** consumption.

(a) Hubble Space Telescope (HST)



(b) James Web Space Telescope (JWST)

Figure: Optical telescopes

Segmented Planar Imaging Detector for Electro-optical Reconnaissance (SPIDER)

- SPIDER imaging device developed by Prof. Ben Yoo and colleagues at UC Davis and Lockheed Martin (Kendrick *et al.* 2013; Duncan *et al.* 2015).
- SPIDER is a small-scale **interferometric optical imaging device** that first uses a lenslet array to measure multiple interferometer baselines, then uses photonic integrated circuits (PICs) to miniaturize the measurement acquisition.

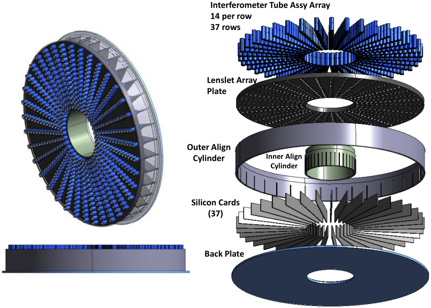


Figure: SPIDER payload design [Credit: Kendrick *et al.* 2013]

SPIDER

- SPIDER **reduces the weight, cost, and power consumption** of optical telescopes.

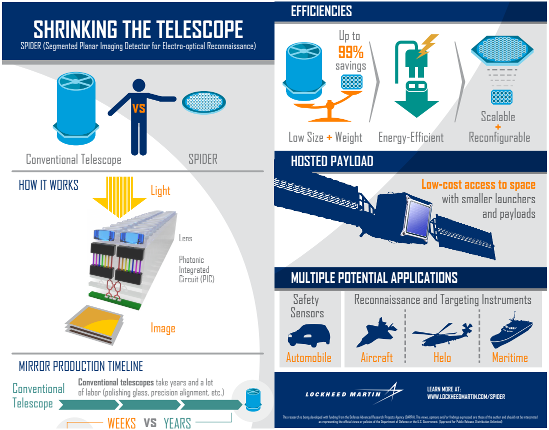


Figure: SPIDER advantages [Credit: Lockheed Martin]

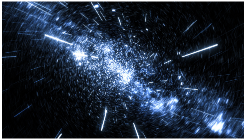
SPIDER

- Unlike traditional optical interferometry, the SPIDER telescope can accurately retrieve both **phase and amplitude** information, making the measurement process analogous to a radio interferometer.
- Accurate interferometric image reconstruction methods from radio astronomy can thus be applied to image SPIDER observations.

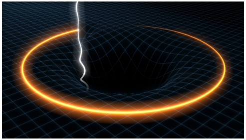
Figure: SPIDER imaging is analogous to astronomical radio interferometry

Next-generation of radio interferometry rapidly approaching

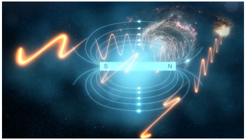
- Next-generation of radio interferometric telescopes will provide **orders of magnitude improvement in sensitivity**.
- Unlock broad range of science goals.



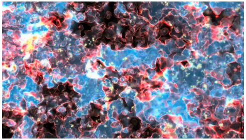
(a) **Dark energy**



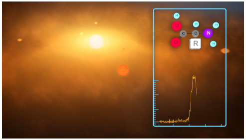
(b) **General relativity**



(c) **Cosmic magnetism**



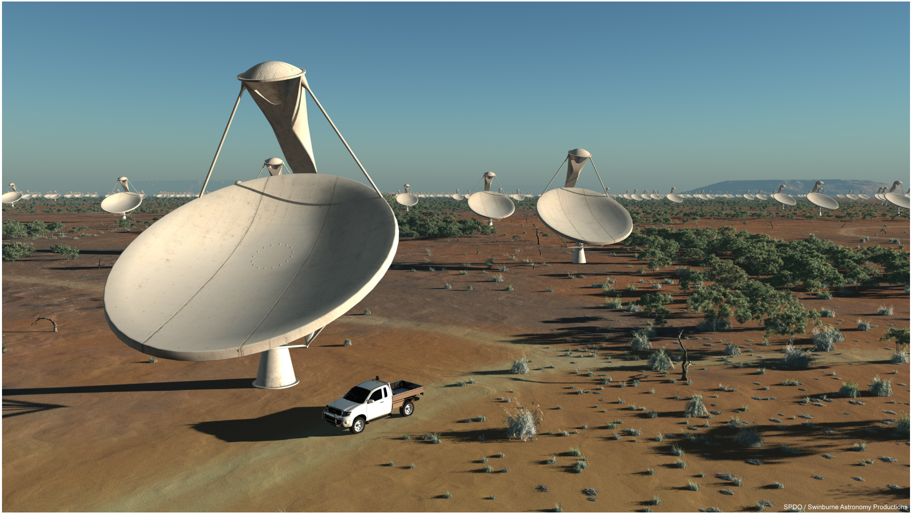
(d) **Epoch of reionization**



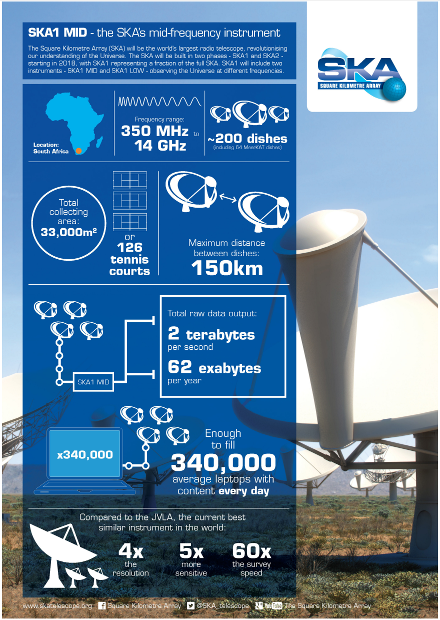
(e) **Exoplanets**

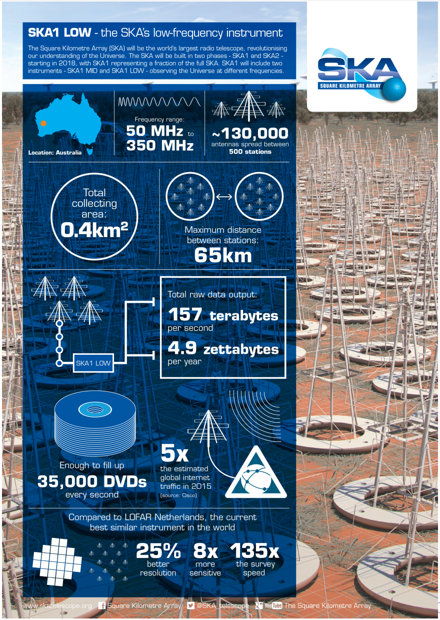
Figure: SKA science goals. **[Credit: SKA Organisation]**

Square Kilometre Array (SKA)

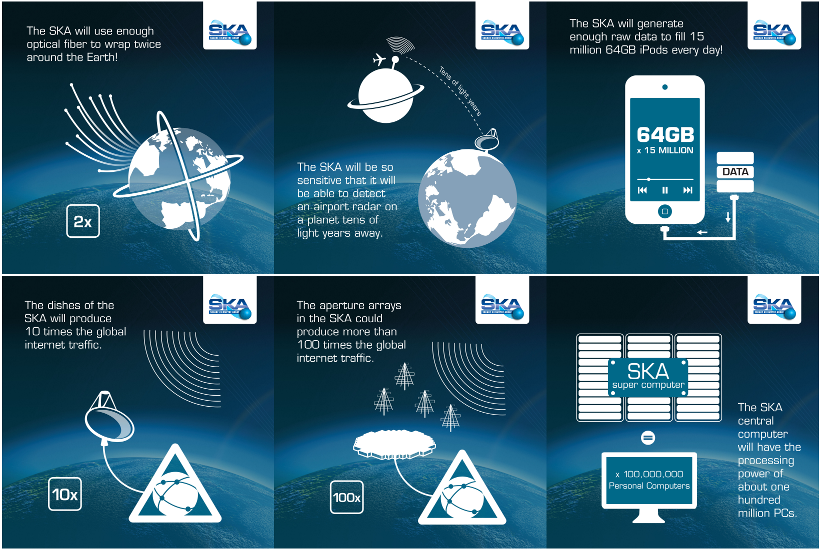


SKA sites


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

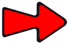


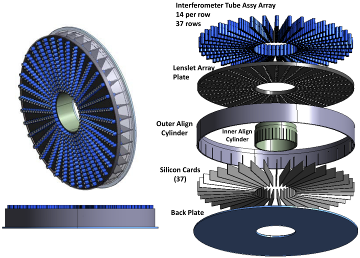
The SKA poses a considerable big-data challenge

Potential to transfer techniques from radio interferometry to SPIDER

Recent advances in radio interferometric imaging could be transferred to SPIDER imaging:

1. High-fidelity imaging
2. Efficient algorithms and implementations
3. Uncertainty quantification
4. Online imaging





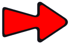
Outline

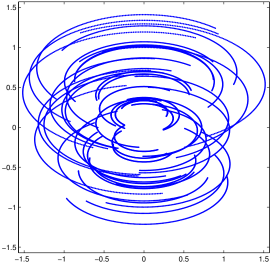
1. Radio interferometric imaging
2. Uncertainty quantification (MCMC sampling)
3. Uncertainty quantification (MAP estimation)
4. Online imaging

Outline

Radio interferometric telescopes acquire “Fourier” measurements

“Fourier”
Measurements





Radio interferometric inverse problem

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

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

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.* $\boxed{\boldsymbol{\Phi} = \mathbf{GFA}}$, 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.* $\boxed{\boldsymbol{\Phi} = \mathbf{GFA}}$, 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

Synthesis and analysis frameworks

- 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

(For orthogonal bases the two approaches are identical but otherwise very different.)

Sparse regularisation

Synthesis and analysis frameworks

- 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

(For orthogonal bases the two approaches are identical but otherwise very different.)

Sparse regularisation

Synthesis and analysis frameworks

- 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

(For orthogonal bases the two approaches are identical but otherwise very different.)

Sparse regularisation

SARA algorithm

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

- Overcomplete dictionary composed of a concatenation of orthonormal bases:

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

  with following bases: Dirac (*i.e.* pixel basis); Haar wavelets (promotes gradient sparsity); Daubechies wavelets two to eight $\Rightarrow$ concatenation of 9 bases.

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

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

Sparse regularisation

SARA algorithm

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

- Overcomplete dictionary composed of a concatenation of orthonormal bases:

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

with following bases: Dirac (*i.e.* pixel basis); Haar wavelets (promotes gradient sparsity); Daubechies wavelets two to eight $\Rightarrow$ concatenation of 9 bases.

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

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

SARA

Sparse regularisation

SARA algorithm

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

- Overcomplete dictionary composed of a concatenation of orthonormal bases:

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

  with following bases: Dirac (*i.e.* pixel basis); Haar wavelets (promotes gradient sparsity); Daubechies wavelets two to eight $\Rightarrow$ concatenation of 9 bases.

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

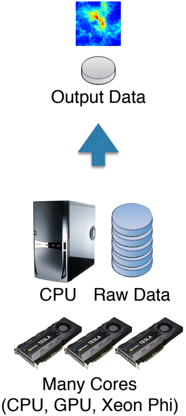
$$\boxed{\min_{\boldsymbol{x} \in \mathbb{R}^N} \|\mathbf{W}\mathbf{\Psi}^\dagger \boldsymbol{x}\|_1 \quad \text{subject to} \quad \|\boldsymbol{y} - \mathbf{\Phi}\boldsymbol{x}\|_2 \leq \epsilon \quad \text{and} \quad \boldsymbol{x} \geq 0}$$

SARA

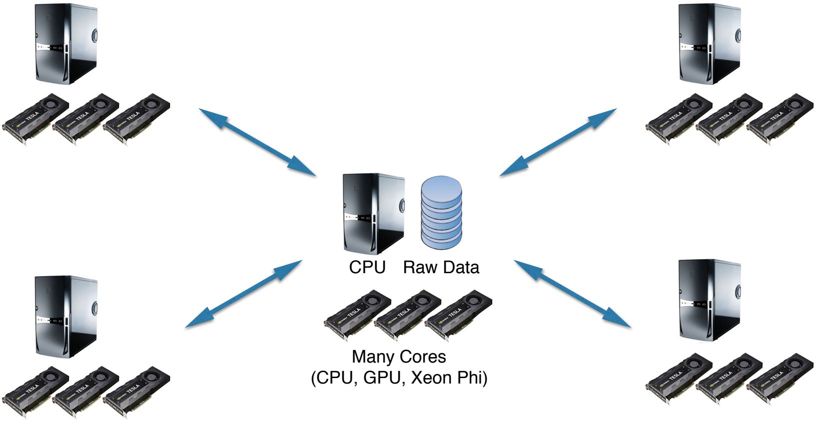
Distributed and parallelised convex optimisation

- Solve resulting convex optimisation problems by proximal splitting.
- **Distributed and parallelised** sparse convex optimization for radio interferometry with PURIFY (Pratley, McEwen, *et al.* 2019; arXiv:1903.04502)
- **Load balancing** for distributed interferometric image reconstruction (Pratley, McEwen 2019; arXiv:1903.07621)
- Image **2 billion visibilities (measurements)** on 50 nodes of HPC cluster.

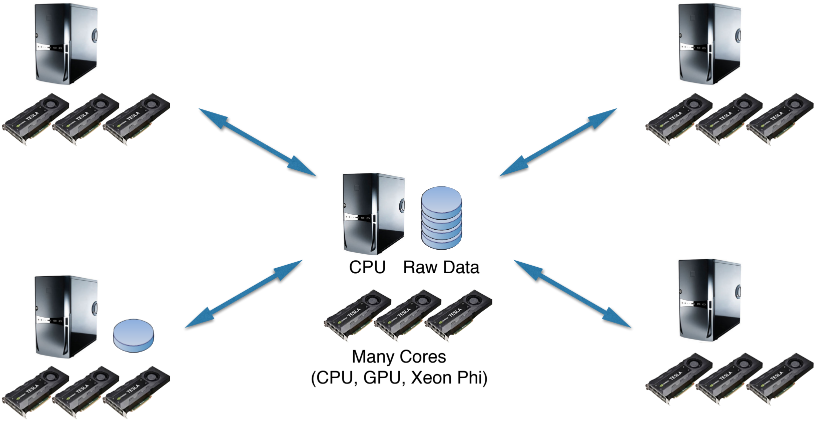
Standard algorithms



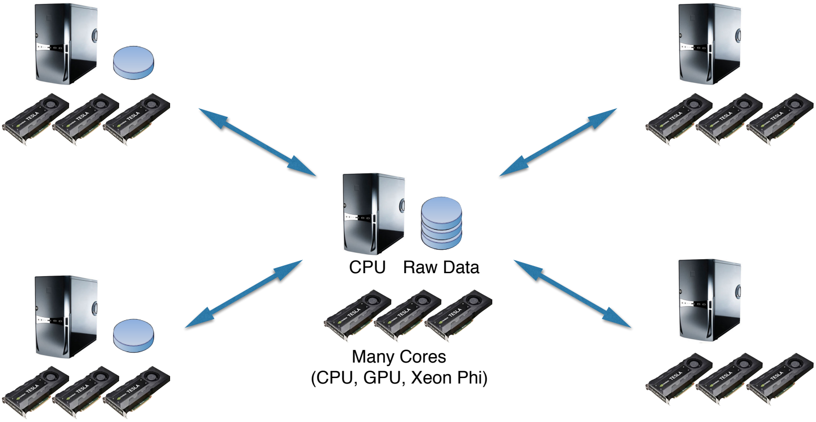
Highly distributed and parallelised algorithms



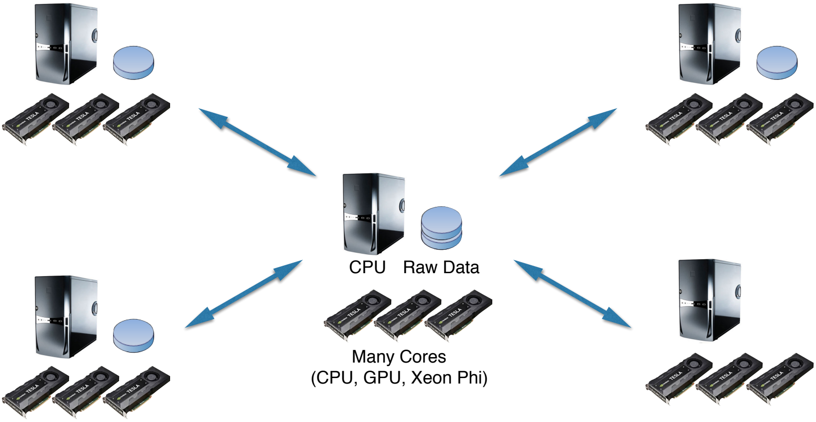
Highly distributed and parallelised algorithms



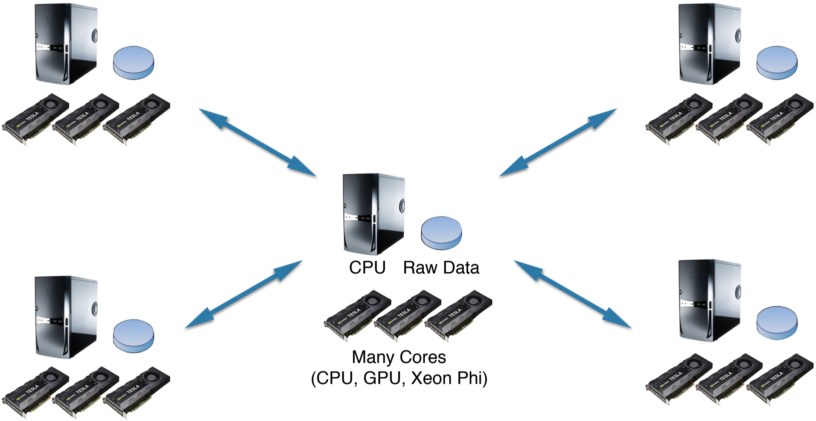
Highly distributed and parallelised algorithms



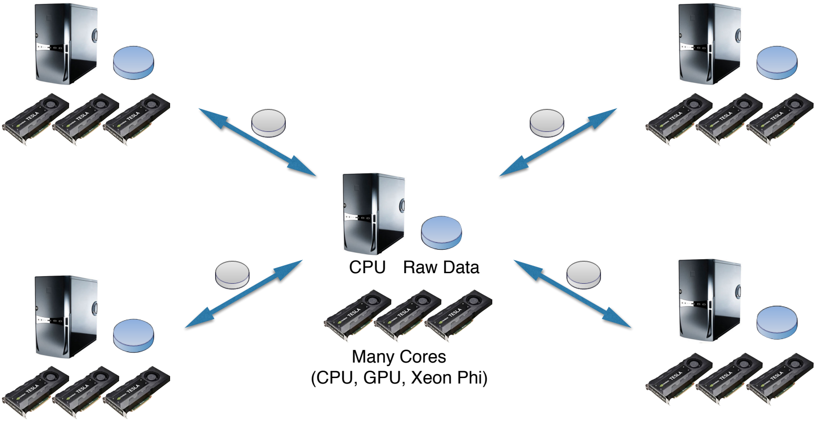
Highly distributed and parallelised algorithms



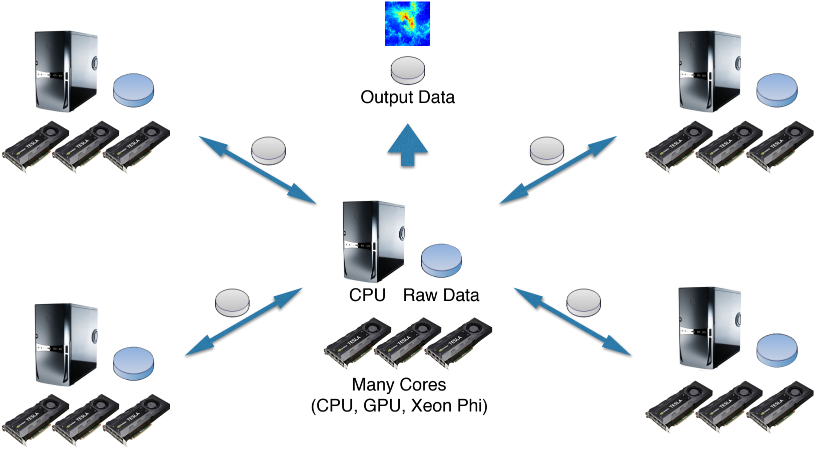
Highly distributed and parallelised algorithms



Highly distributed and parallelised algorithms



Highly distributed and parallelised algorithms



Public open-source codes

PURIFY code http://astro-informatics.github.io/purify/

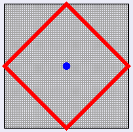
Next-generation radio interferometric imaging

d'Avezac, Carrillo, Christidi, Guichard, McEwen, Perez-Suarez, Pratley, Wiaux

Project lead: McEwen

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://astro-informatics.github.io/sopt/



Sparse OPTimisation

d'Avezac, Carrillo, Christidi, Guichard, McEwen, Perez-Suarez, Pratley, Wiaux

Project lead: McEwen

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

Imaging observations from the VLA and ATCA with PURIFY

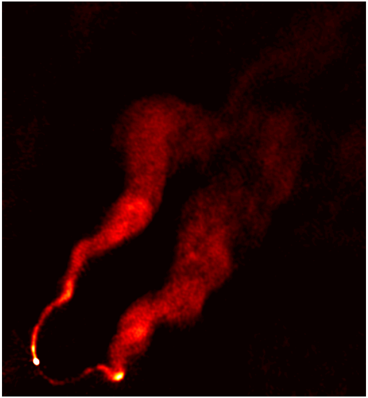
(a) NRAO Very Large Array (VLA)

(b) Australia Telescope Compact Array (ATCA)

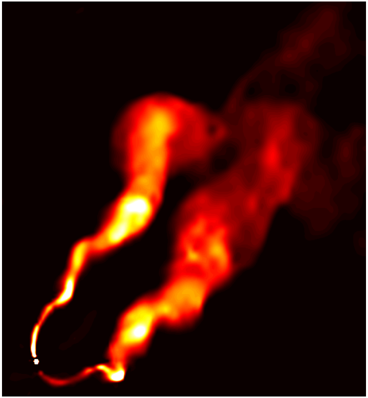
Figure: Radio interferometric telescopes considered

PURIFY reconstruction

VLA observation of 3C129



(a) CLEAN (uniform)

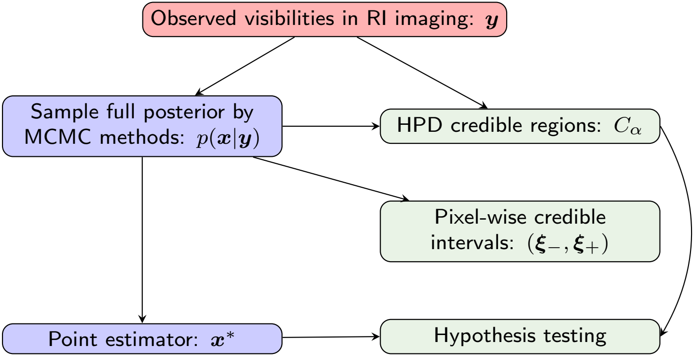


(b) PURIFY

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

Outline

MCMC sampling and uncertainty quantification



Uncertainty quantification for radio interferometric imaging: I. proximal MCMC methods (Cai, Pereyra &McEwen 2018a; arXiv:1711.04818)

MCMC sampling the full posterior distribution

- Sample full posterior distribution $\mathrm{P}(\boldsymbol{x} \,|\, \boldsymbol{y})$.
- MCMC methods for high-dimensional problems (like interferometric imaging):
  - Gibbs sampling (sample from conditional distributions)
  - Hamiltonian MC (HMC) sampling (exploit gradients)
  - Metropolis adjusted Langevin algorithm (MALA) sampling (exploit gradients)

Require MCMC approach to support sparsity priors, which shown to be highly effective.

MCMC sampling the full posterior distribution

- Sample full posterior distribution $\mathrm{P}(\boldsymbol{x} \mid \boldsymbol{y})$.
- MCMC methods for high-dimensional problems (like interferometric imaging):
  - Gibbs sampling (sample from conditional distributions)
  - Hamiltonian MC (HMC) sampling (exploit gradients)
  - Metropolis adjusted Langevin algorithm (MALA) sampling (exploit gradients)

Require MCMC approach to support sparsity priors, which shown to be highly effective.

MCMC sampling the full posterior distribution

- Sample full posterior distribution $\mathrm{P}(\boldsymbol{x} \mid \boldsymbol{y})$.
- MCMC methods for high-dimensional problems (like interferometric imaging):
  - Gibbs sampling (sample from conditional distributions)
  - Hamiltonian MC (HMC) sampling (exploit gradients)
  - Metropolis adjusted Langevin algorithm (MALA) sampling (exploit gradients)

Require MCMC approach to support sparsity priors, which shown to be highly effective.

MCMC sampling with gradients

Langevin dynamics

- Consider posteriors of the following form:

$$\mathrm{P}(\boldsymbol{x} \mid \boldsymbol{y}) = \underbrace{\boxed{\pi(\boldsymbol{x})}}_{\text{Posterior}} \propto \exp\big(-\underbrace{\boxed{g(\boldsymbol{x})}}_{\text{Smooth}}\big)$$

- If $g(x)$ differentiable can adopt MALA (Langevin dynamics).
- Based on Langevin diffusion process $\mathcal{L}(t)$, with π as stationary distribution:

$$\mathrm{d}\mathcal{L}(t) = \frac{1}{2}\nabla \log \pi\big(\mathcal{L}(t)\big)\mathrm{d}t + \mathrm{d}\mathcal{W}(t), \quad \mathcal{L}(0) = l_0$$

  where $\mathcal{W}$ is Brownian motion.
- Need gradients so cannot support sparse priors.

MCMC sampling with gradients

Langevin dynamics

- Consider posteriors of the following form:

$$\mathrm{P}(\boldsymbol{x} \mid \boldsymbol{y}) = \underbrace{\boxed{\pi(\boldsymbol{x})}}_{\text{Posterior}} \propto \exp\big(-\underbrace{\boxed{g(\boldsymbol{x})}}_{\text{Smooth}}\big)$$

- If $g(\boldsymbol{x})$ differentiable can adopt MALA (Langevin dynamics).
- Based on Langevin diffusion process $\mathcal{L}(t)$, with π as stationary distribution:

$$\mathrm{d}\mathcal{L}(t) = \frac{1}{2}\nabla \log \pi\big(\mathcal{L}(t)\big)\mathrm{d}t + \mathrm{d}\mathcal{W}(t), \quad \mathcal{L}(0) = l_0$$

where $\mathcal{W}$ is Brownian motion.

- Need gradients so cannot support sparse priors.

MCMC sampling with gradients

Langevin dynamics

- Consider posteriors of the following form:

$$\mathrm{P}(\boldsymbol{x} \mid \boldsymbol{y}) = \underbrace{\boxed{\pi(\boldsymbol{x})}}_{\text{Posterior}} \propto \exp\big(-\underbrace{\boxed{g(\boldsymbol{x})}}_{\text{Smooth}}\big)$$

- If $g(\boldsymbol{x})$ differentiable can adopt MALA (Langevin dynamics).
- Based on Langevin diffusion process $\mathcal{L}(t)$, with π as stationary distribution:

$$\mathrm{d}\mathcal{L}(t) = \frac{1}{2}\nabla \log \pi\big(\mathcal{L}(t)\big)\mathrm{d}t + \mathrm{d}\mathcal{W}(t), \quad \mathcal{L}(0) = l_0$$

  where $\mathcal{W}$ is Brownian motion.
- Need gradients so cannot support sparse priors.

MCMC sampling with gradients

Langevin dynamics

- Consider posteriors of the following form:

$$\mathrm{P}(\boldsymbol{x} \mid \boldsymbol{y}) = \underbrace{\boxed{\pi(\boldsymbol{x})}}_{\text{Posterior}} \propto \exp\big(-\underbrace{\boxed{g(\boldsymbol{x})}}_{\text{Smooth}}\big)$$

- If $g(\boldsymbol{x})$ differentiable can adopt MALA (Langevin dynamics).
- Based on Langevin diffusion process $\mathcal{L}(t)$, with π as stationary distribution:

$$\mathrm{d}\mathcal{L}(t) = \frac{1}{2} \underbrace{\boxed{\nabla \log \pi\big(\mathcal{L}(t)\big)}}_{\text{Gradient}} \mathrm{d}t + \mathrm{d}\mathcal{W}(t), \quad \mathcal{L}(0) = l_0$$

  where $\mathcal{W}$ is Brownian motion.

- Need gradients so cannot support sparse priors.

Proximity operators

A brief aside

- Define proximity operator:

$$\boxed{\mathrm{prox}_g^{\lambda}(\boldsymbol{x}) = \arg\min_{\boldsymbol{u}} \left[g(\boldsymbol{u}) + \|\boldsymbol{u} - \boldsymbol{x}\|^2 / 2\lambda \right]}$$

- Generalisation of projection operator:

$$\mathcal{P}_{\mathcal{C}}(\boldsymbol{x}) = \arg\min_{\boldsymbol{u}} \left[\imath_{\mathcal{C}}(\boldsymbol{u}) + \|\boldsymbol{u} - \boldsymbol{x}\|^2 / 2 \right],$$

where $\imath_{\mathcal{C}}(\boldsymbol{u}) = \infty$ if $\boldsymbol{u} \notin \mathcal{C}$ and zero otherwise.

Proximity operators

A brief aside

- Define proximity operator:

$$\boxed{\mathrm{prox}_g^{\lambda}(\boldsymbol{x}) = \arg\min_{\boldsymbol{u}} \left[g(\boldsymbol{u}) + \|\boldsymbol{u} - \boldsymbol{x}\|^2 / 2\lambda \right]}$$

- Generalisation of projection operator:

$$\mathcal{P}_{\mathcal{C}}(\boldsymbol{x}) = \arg\min_{\boldsymbol{u}} \left[\imath_{\mathcal{C}}(\boldsymbol{u}) + \|\boldsymbol{u} - \boldsymbol{x}\|^2 / 2 \right],$$

where $\imath_{\mathcal{C}}(\boldsymbol{u}) = \infty$ if $\boldsymbol{u} \notin \mathcal{C}$ and zero otherwise.

Proximity operators

A brief aside

- Define proximity operator:

$$\boxed{\mathrm{prox}_g^{\lambda}(\boldsymbol{x}) = \arg\min_{\boldsymbol{u}} \Big[g(\boldsymbol{u}) + \|\boldsymbol{u} - \boldsymbol{x}\|^2 / 2\lambda \Big]}$$

- Generalisation of projection operator:

$$\mathcal{P}_{\mathcal{C}}(\boldsymbol{x}) = \arg\min_{\boldsymbol{u}} \Big[\imath_{\mathcal{C}}(\boldsymbol{u}) + \|\boldsymbol{u} - \boldsymbol{x}\|^2 / 2 \Big] ,$$

where $\imath_{\mathcal{C}}(\boldsymbol{u}) = \infty$ if $\boldsymbol{u} \notin \mathcal{C}$ and zero otherwise.

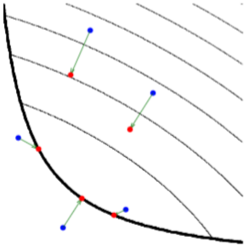


Figure: Illustration of proximity operator [Credit: Parikh & Boyd (2013)]

Proximal MCMC methods

- Exploit proximal calculus.
- "Replace gradients with sub-gradients".

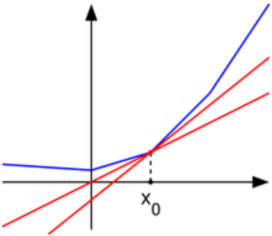


Figure: Illustration of sub-gradients [Credit: Wikipedia (Maksim)]

Proximal MALA

Moreau approximation

- Moreau approximation of $f(\boldsymbol{x}) \propto \exp(-g(\boldsymbol{x}))$:

$$\boxed{f_\lambda^{\mathsf{MA}}(\boldsymbol{x}) = \sup_{\boldsymbol{u} \in \mathbb{R}^N} f(\boldsymbol{u}) \exp\left(-\frac{\|\boldsymbol{u}-\boldsymbol{x}\|^2}{2\lambda}\right)}$$

- Important properties of $f_\lambda^{\mathsf{MA}}(\boldsymbol{x})$:

1. As $\lambda \rightarrow 0$, $f_\lambda^{\mathsf{MA}}(\boldsymbol{x}) \rightarrow f(\boldsymbol{x})$
2. $\nabla \log f_\lambda^{\mathsf{MA}}(\boldsymbol{x}) = (\mathrm{prox}_g^\lambda(\boldsymbol{x}) - \boldsymbol{x})/\lambda$

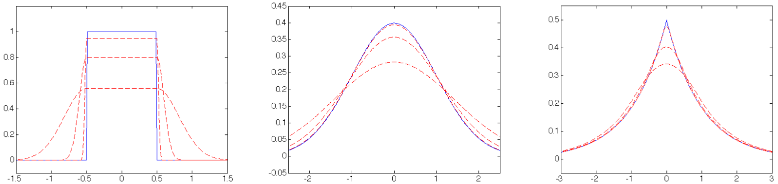


Figure: Illustration of Moreau approximations [Credit: Pereyra 2016a]

Proximal MALA

Moreau approximation

- Moreau approximation of $f(\boldsymbol{x}) \propto \exp(-g(\boldsymbol{x}))$:

$$\boxed{f_{\lambda}^{\mathsf{MA}}(\boldsymbol{x}) = \sup_{\boldsymbol{u}\in\mathbb{R}^N} f(\boldsymbol{u}) \exp\left(-\frac{\|\boldsymbol{u}-\boldsymbol{x}\|^2}{2\lambda}\right)}$$

- Important properties of $f_{\lambda}^{\mathsf{MA}}(\boldsymbol{x})$:

1. As $\lambda \rightarrow 0$, $f_{\lambda}^{\mathsf{MA}}(\boldsymbol{x}) \rightarrow f(\boldsymbol{x})$
2. $\nabla \log f_{\lambda}^{\mathsf{MA}}(\boldsymbol{x}) = (\mathrm{prox}_g^{\lambda}(\boldsymbol{x}) - \boldsymbol{x})/\lambda$

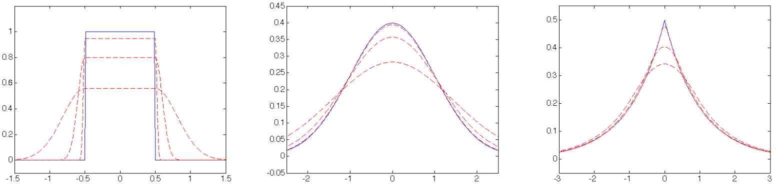


Figure: Illustration of Moreau approximations [Credit: Pereyra 2016a]

Proximal MALA

MCMC sampling

Proximal Metropolis adjusted Langevin algorithm (P-MALA)

Pereyra (2016a)

- Consider log-convex posteriors: $\mathrm{P}(\boldsymbol{x} \mid \boldsymbol{y}) = \pi(\boldsymbol{x}) \propto \exp\big(-\boxed{g(\boldsymbol{x})}\big)$. Convex
- Langevin diffusion process $\mathcal{L}(t)$, with π as stationary distribution ($\mathcal{W}$ Brownian motion):

$$\mathrm{d}\mathcal{L}(t) = \frac{1}{2}\nabla \log \pi\big(\mathcal{L}(t)\big)\mathrm{d}t + \mathrm{d}\mathcal{W}(t), \quad \mathcal{L}(0) = l_0 \ .$$

- Euler discretisation and apply Moreau approximation to π:

$$\boldsymbol{l}^{(m+1)} = \boldsymbol{l}^{(m)} + \frac{\delta}{2}\boxed{\nabla \log \pi(\boldsymbol{l}^{(m)})} + \sqrt{\delta}\boldsymbol{w}^{(m)} \ .$$

$$\nabla \log \pi_\lambda(\boldsymbol{x}) = (\mathrm{prox}_g^\lambda(\boldsymbol{x}) - \boldsymbol{x})/\lambda$$

- Metropolis-Hastings accept-reject step.

Proximal MALA

MCMC sampling

Proximal Metropolis adjusted Langevin algorithm (P-MALA)

Pereyra (2016a)

- Consider log-convex posteriors: $\mathrm{P}(\boldsymbol{x} \mid \boldsymbol{y}) = \pi(\boldsymbol{x}) \propto \exp(-\underbrace{\boxed{g(\boldsymbol{x})}}_{\text{Convex}})$.
- Langevin diffusion process $\mathcal{L}(t)$, with π as stationary distribution ($\mathcal{W}$ Brownian motion):

$$\mathrm{d}\mathcal{L}(t) = \frac{1}{2}\nabla \log \pi\big(\mathcal{L}(t)\big)\mathrm{d}t + \mathrm{d}\mathcal{W}(t), \quad \mathcal{L}(0) = l_0 \ .$$

- Euler discretisation and apply Moreau approximation to π:

$$\boldsymbol{l}^{(m+1)} = \boldsymbol{l}^{(m)} + \frac{\delta}{2}\boxed{\nabla \log \pi(\boldsymbol{l}^{(m)})} + \sqrt{\delta}\boldsymbol{w}^{(m)} \ .$$

$$\nabla \log \pi_\lambda(\boldsymbol{x}) = (\mathrm{prox}_g^\lambda(\boldsymbol{x}) - \boldsymbol{x})/\lambda$$

- Metropolis-Hastings accept-reject step.

Proximal MALA

MCMC sampling

Proximal Metropolis adjusted Langevin algorithm (P-MALA)

Pereyra (2016a)

- Consider log-convex posteriors: $\mathrm{P}(\boldsymbol{x} \mid \boldsymbol{y}) = \pi(\boldsymbol{x}) \propto \exp(-\underbrace{\boxed{g(\boldsymbol{x})}}_{\text{Convex}})$.
- Langevin diffusion process $\mathcal{L}(t)$, with π as stationary distribution ($\mathcal{W}$ Brownian motion):

$$\mathrm{d}\mathcal{L}(t) = \frac{1}{2}\nabla \log \pi\big(\mathcal{L}(t)\big)\mathrm{d}t + \mathrm{d}\mathcal{W}(t), \quad \mathcal{L}(0) = l_0 \ .$$

- Euler discretisation and apply Moreau approximation to π:

$$\boldsymbol{l}^{(m+1)} = \boldsymbol{l}^{(m)} + \frac{\delta}{2}\boxed{\nabla \log \pi(\boldsymbol{l}^{(m)})} + \sqrt{\delta}\boldsymbol{w}^{(m)} \ .$$

$$\nabla \log \pi_\lambda(\boldsymbol{x}) = (\mathrm{prox}_g^\lambda(\boldsymbol{x}) - \boldsymbol{x})/\lambda$$

- Metropolis-Hastings accept-reject step.

Proximal MALA

MCMC sampling

Proximal Metropolis adjusted Langevin algorithm (P-MALA)

Pereyra (2016a)

- Consider log-convex posteriors: $\mathrm{P}(\boldsymbol{x} \mid \boldsymbol{y}) = \pi(\boldsymbol{x}) \propto \exp(-\boxed{g(\boldsymbol{x})})$. Convex
- Langevin diffusion process $\mathcal{L}(t)$, with π as stationary distribution ($\mathcal{W}$ Brownian motion):

$$\mathrm{d}\mathcal{L}(t) = \frac{1}{2}\nabla \log \pi\big(\mathcal{L}(t)\big)\mathrm{d}t + \mathrm{d}\mathcal{W}(t), \quad \mathcal{L}(0) = l_0 \, .$$

- Euler discretisation and apply Moreau approximation to π:

$$\boldsymbol{l}^{(m+1)} = \boldsymbol{l}^{(m)} + \frac{\delta}{2}\boxed{\nabla \log \pi(\boldsymbol{l}^{(m)})} + \sqrt{\delta}\boldsymbol{w}^{(m)} \, .$$

$$\nabla \log \pi_\lambda(\boldsymbol{x}) = (\mathrm{prox}_g^\lambda(\boldsymbol{x}) - \boldsymbol{x})/\lambda$$

- Metropolis-Hastings accept-reject step.

Proximal MALA

Computing proximity operators for the analysis case

- Recall posterior: $\pi(\boldsymbol{x}) \propto \exp\big(-g(\boldsymbol{x})\big)$.
- Let $\bar{g}(\boldsymbol{x}) = \bar{f}_1(\boldsymbol{x}) + \bar{f}_2(\boldsymbol{x})$, where $\boxed{\bar{f}_1(\boldsymbol{x}) = \mu \|\boldsymbol{\Psi}^\dagger \boldsymbol{x}\|_1}$ and $\boxed{\bar{f}_2(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{x}\|_2^2 / 2\sigma^2}$.

  Prior Likelihood
- Must solve an optimisation problem for each iteration!

$$\operatorname{prox}_{\bar{g}}^{\delta/2}(\boldsymbol{x}) = \operatorname*{argmin}_{\boldsymbol{u} \in \mathbb{R}^N} \left\{ \mu \|\boldsymbol{\Psi}^\dagger \boldsymbol{u}\|_1 + \frac{\|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{u}\|_2^2}{2\sigma^2} + \frac{\|\boldsymbol{u} - \boldsymbol{x}\|_2^2}{\delta} \right\}.$$

- Taylor expansion at point $\boldsymbol{x}$: $\|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{u}\|_2^2 \approx \|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{x}\|_2^2 + 2(\boldsymbol{u} - \boldsymbol{x})^\top \boldsymbol{\Phi}^\dagger(\boldsymbol{\Phi}\boldsymbol{x} - \boldsymbol{y})$.
- Then proximity operator approximated by

$$\operatorname{prox}_{\bar{g}}^{\delta/2}(\boldsymbol{x}) \approx \operatorname{prox}_{\bar{f}_1}^{\delta/2}\left(\boldsymbol{x} - \delta \boldsymbol{\Phi}^\dagger(\boldsymbol{\Phi}\boldsymbol{x} - \boldsymbol{y})/2\sigma^2\right).$$

Single forward-backward iteration

- Analytic approximation:

$$\operatorname{prox}_{\bar{g}}^{\delta/2}(\boldsymbol{x}) \approx \bar{\boldsymbol{v}} + \boldsymbol{\Psi}\left(\operatorname{soft}_{\mu\delta/2}(\boldsymbol{\Psi}^\dagger \bar{\boldsymbol{v}}) - \boldsymbol{\Psi}^\dagger \bar{\boldsymbol{v}}\right),$$ where $\bar{\boldsymbol{v}} = \boldsymbol{x} - \delta \boldsymbol{\Phi}^\dagger(\boldsymbol{\Phi}\boldsymbol{x} - \boldsymbol{y})/2\sigma^2$.

Proximal MALA

Computing proximity operators for the analysis case

- Recall posterior: $\pi(\boldsymbol{x}) \propto \exp\big(-g(\boldsymbol{x})\big)$.
- Let $\bar{g}(\boldsymbol{x}) = \bar{f}_1(\boldsymbol{x}) + \bar{f}_2(\boldsymbol{x})$, where $\boxed{\bar{f}_1(\boldsymbol{x}) = \mu \|\boldsymbol{\Psi}^\dagger \boldsymbol{x}\|_1}$ and $\boxed{\bar{f}_2(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{x}\|_2^2 / 2\sigma^2}$.

  Prior Likelihood
- Must solve an optimisation problem for each iteration!

$$\boxed{\operatorname{prox}_{\bar{g}}^{\delta/2}(\boldsymbol{x}) = \underset{\boldsymbol{u} \in \mathbb{R}^N}{\operatorname{argmin}} \left\{ \mu \|\boldsymbol{\Psi}^\dagger \boldsymbol{u}\|_1 + \frac{\|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{u}\|_2^2}{2\sigma^2} + \frac{\|\boldsymbol{u} - \boldsymbol{x}\|_2^2}{\delta} \right\}}.$$

- Taylor expansion at point $\boldsymbol{x}$: $\|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{u}\|_2^2 \approx \|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{x}\|_2^2 + 2(\boldsymbol{u} - \boldsymbol{x})^\top \boldsymbol{\Phi}^\dagger (\boldsymbol{\Phi}\boldsymbol{x} - \boldsymbol{y})$.
- Then proximity operator approximated by

$$\boxed{\operatorname{prox}_{\bar{g}}^{\delta/2}(\boldsymbol{x}) \approx \operatorname{prox}_{\bar{f}_1}^{\delta/2}\left(\boldsymbol{x} - \delta \boldsymbol{\Phi}^\dagger(\boldsymbol{\Phi}\boldsymbol{x} - \boldsymbol{y})/2\sigma^2\right)}.$$

Single forward-backward iteration

- Analytic approximation:

$$\boxed{\operatorname{prox}_{\bar{g}}^{\delta/2}(\boldsymbol{x}) \approx \bar{\boldsymbol{v}} + \boldsymbol{\Psi}\left(\operatorname{soft}_{\mu\delta/2}(\boldsymbol{\Psi}^\dagger \bar{\boldsymbol{v}}) - \boldsymbol{\Psi}^\dagger \bar{\boldsymbol{v}}\right)}, \quad \text{where } \bar{\boldsymbol{v}} = \boldsymbol{x} - \delta \boldsymbol{\Phi}^\dagger(\boldsymbol{\Phi}\boldsymbol{x} - \boldsymbol{y})/2\sigma^2.$$

Proximal MALA

Computing proximity operators for the analysis case

- Recall posterior: $\pi(\boldsymbol{x}) \propto \exp\big(-g(\boldsymbol{x})\big)$.
- Let $\bar{g}(\boldsymbol{x}) = \bar{f}_1(\boldsymbol{x}) + \bar{f}_2(\boldsymbol{x})$, where $\boxed{\bar{f}_1(\boldsymbol{x}) = \mu \|\boldsymbol{\Psi}^\dagger \boldsymbol{x}\|_1}$ and $\boxed{\bar{f}_2(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{x}\|_2^2 / 2\sigma^2}$.

  Prior Likelihood
- Must solve an optimisation problem for each iteration!

$$\boxed{\operatorname{prox}_{\bar{g}}^{\delta/2}(\boldsymbol{x}) = \operatorname*{argmin}_{\boldsymbol{u} \in \mathbb{R}^N} \left\{ \mu \|\boldsymbol{\Psi}^\dagger \boldsymbol{u}\|_1 + \frac{\|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{u}\|_2^2}{2\sigma^2} + \frac{\|\boldsymbol{u} - \boldsymbol{x}\|_2^2}{\delta} \right\}}.$$

- Taylor expansion at point $\boldsymbol{x}$: $\|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{u}\|_2^2 \approx \|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{x}\|_2^2 + 2(\boldsymbol{u} - \boldsymbol{x})^\top \boldsymbol{\Phi}^\dagger (\boldsymbol{\Phi}\boldsymbol{x} - \boldsymbol{y})$.
- Then proximity operator approximated by

$$\boxed{\operatorname{prox}_{\bar{g}}^{\delta/2}(\boldsymbol{x}) \approx \operatorname{prox}_{\bar{f}_1}^{\delta/2}\left(\boldsymbol{x} - \delta \boldsymbol{\Phi}^\dagger (\boldsymbol{\Phi}\boldsymbol{x} - \boldsymbol{y}) / 2\sigma^2\right)}.$$

Single forward-backward iteration

- Analytic approximation:

$$\boxed{\operatorname{prox}_{\bar{g}}^{\delta/2}(\boldsymbol{x}) \approx \bar{\boldsymbol{v}} + \boldsymbol{\Psi}\left(\operatorname{soft}_{\mu\delta/2}(\boldsymbol{\Psi}^\dagger \bar{\boldsymbol{v}}) - \boldsymbol{\Psi}^\dagger \bar{\boldsymbol{v}}\right)}, \quad \text{where } \bar{\boldsymbol{v}} = \boldsymbol{x} - \delta \boldsymbol{\Phi}^\dagger (\boldsymbol{\Phi}\boldsymbol{x} - \boldsymbol{y}) / 2\sigma^2.$$

Proximal MALA

Computing proximity operators for the analysis case

- Recall posterior: $\pi(\boldsymbol{x}) \propto \exp\big(-g(\boldsymbol{x})\big)$.
- Let $\bar{g}(\boldsymbol{x}) = \bar{f}_1(\boldsymbol{x}) + \bar{f}_2(\boldsymbol{x})$, where $\boxed{\bar{f}_1(\boldsymbol{x}) = \mu \|\boldsymbol{\Psi}^\dagger \boldsymbol{x}\|_1}$ and $\boxed{\bar{f}_2(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{x}\|_2^2 / 2\sigma^2}$.

  Prior Likelihood

- Must solve an optimisation problem for each iteration!

$$\boxed{\operatorname{prox}_{\bar{g}}^{\delta/2}(\boldsymbol{x}) = \underset{\boldsymbol{u} \in \mathbb{R}^N}{\operatorname{argmin}} \left\{ \mu \|\boldsymbol{\Psi}^\dagger \boldsymbol{u}\|_1 + \frac{\|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{u}\|_2^2}{2\sigma^2} + \frac{\|\boldsymbol{u} - \boldsymbol{x}\|_2^2}{\delta} \right\}} .$$

- Taylor expansion at point $\boldsymbol{x}$: $\|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{u}\|_2^2 \approx \|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{x}\|_2^2 + 2(\boldsymbol{u} - \boldsymbol{x})^\top \boldsymbol{\Phi}^\dagger (\boldsymbol{\Phi}\boldsymbol{x} - \boldsymbol{y})$.
- Then proximity operator approximated by

$$\boxed{\operatorname{prox}_{\bar{g}}^{\delta/2}(\boldsymbol{x}) \approx \operatorname{prox}_{\bar{f}_1}^{\delta/2} \left(\boldsymbol{x} - \delta \boldsymbol{\Phi}^\dagger (\boldsymbol{\Phi}\boldsymbol{x} - \boldsymbol{y}) / 2\sigma^2 \right)} .$$

Single forward-backward iteration

- Analytic approximation:

$$\boxed{\operatorname{prox}_{\bar{g}}^{\delta/2}(\boldsymbol{x}) \approx \bar{\boldsymbol{v}} + \boldsymbol{\Psi} \left(\operatorname{soft}_{\mu\delta/2}(\boldsymbol{\Psi}^\dagger \bar{\boldsymbol{v}}) - \boldsymbol{\Psi}^\dagger \bar{\boldsymbol{v}} \right)} , \text{ where } \bar{\boldsymbol{v}} = \boldsymbol{x} - \delta \boldsymbol{\Phi}^\dagger (\boldsymbol{\Phi}\boldsymbol{x} - \boldsymbol{y}) / 2\sigma^2 .$$

Proximal MALA

Computing proximity operators for the synthesis case

- Recall posterior: $\pi(\boldsymbol{x}) \propto \exp\big(-g(\boldsymbol{x})\big)$.
- Let $\hat{g}(\boldsymbol{x}(\boldsymbol{a})) = \hat{f}_1(\boldsymbol{a}) + \hat{f}_2(\boldsymbol{a})$, where $\boxed{\hat{f}_1(\boldsymbol{a}) = \mu\|\boldsymbol{a}\|_1}$ and $\boxed{\hat{f}_2(\boldsymbol{a}) = \|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{a}\|_2^2/2\sigma^2}$.

  Prior Likelihood

- Must solve an optimisation problem for each iteration!

$$\mathrm{prox}_{\hat{g}}^{\delta/2}(\boldsymbol{a}) = \underset{\boldsymbol{u}\in\mathbb{R}^L}{\mathrm{argmin}} \left\{ \mu\|\boldsymbol{u}\|_1 + \frac{\|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{u}\|_2^2}{2\sigma^2} + \frac{\|\boldsymbol{u} - \boldsymbol{a}\|_2^2}{\delta} \right\}.$$

- Taylor expansion at point $\boldsymbol{a}$: $\|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{u}\|_2^2 \approx \|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{a}\|_2^2 + 2(\boldsymbol{u} - \boldsymbol{a})^\top \boldsymbol{\Psi}^\dagger\boldsymbol{\Phi}^\dagger(\boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{a} - \boldsymbol{y})$.
- Then proximity operator approximated by

$$\mathrm{prox}_{\hat{g}}^{\delta/2}(\boldsymbol{a}) \approx \mathrm{prox}_{\hat{f}_1}^{\delta/2}\left(\boldsymbol{a} - \delta\boldsymbol{\Psi}^\dagger\boldsymbol{\Phi}^\dagger(\boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{a} - \boldsymbol{y})/2\sigma^2\right).$$

Single forward-backward iteration

- Analytic approximation:

$$\mathrm{prox}_{\hat{g}}^{\delta/2}(\boldsymbol{a}) \approx \mathrm{soft}_{\mu\delta/2}\left(\boldsymbol{a} - \delta\boldsymbol{\Psi}^\dagger\boldsymbol{\Phi}^\dagger(\boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{a} - \boldsymbol{y})/2\sigma^2\right).$$

Proximal MALA

Computing proximity operators for the synthesis case

- Recall posterior: $\pi(\boldsymbol{x}) \propto \exp\bigl(-g(\boldsymbol{x})\bigr)$.
- Let $\hat{g}(\boldsymbol{x}(\boldsymbol{a})) = \hat{f}_1(\boldsymbol{a}) + \hat{f}_2(\boldsymbol{a})$, where $\boxed{\hat{f}_1(\boldsymbol{a}) = \mu\|\boldsymbol{a}\|_1}$ and $\boxed{\hat{f}_2(\boldsymbol{a}) = \|\boldsymbol{y} - \boldsymbol{\Phi\Psi a}\|_2^2/2\sigma^2}$.

  Prior Likelihood
- Must solve an optimisation problem for each iteration!

$$\boxed{\operatorname{prox}_{\hat{g}}^{\delta/2}(\boldsymbol{a}) = \underset{\boldsymbol{u}\in\mathbb{R}^L}{\operatorname{argmin}} \left\{ \mu\|\boldsymbol{u}\|_1 + \frac{\|\boldsymbol{y} - \boldsymbol{\Phi\Psi u}\|_2^2}{2\sigma^2} + \frac{\|\boldsymbol{u} - \boldsymbol{a}\|_2^2}{\delta} \right\}}.$$

- Taylor expansion at point $\boldsymbol{a}$: $\|\boldsymbol{y} - \boldsymbol{\Phi\Psi u}\|_2^2 \approx \|\boldsymbol{y} - \boldsymbol{\Phi\Psi a}\|_2^2 + 2(\boldsymbol{u} - \boldsymbol{a})^\top \boldsymbol{\Psi}^\dagger \boldsymbol{\Phi}^\dagger (\boldsymbol{\Phi\Psi a} - \boldsymbol{y})$.
- Then proximity operator approximated by

$$\boxed{\operatorname{prox}_{\hat{g}}^{\delta/2}(\boldsymbol{a}) \approx \operatorname{prox}_{\hat{f}_1}^{\delta/2}\left(\boldsymbol{a} - \delta\boldsymbol{\Psi}^\dagger\boldsymbol{\Phi}^\dagger(\boldsymbol{\Phi\Psi a} - \boldsymbol{y})/2\sigma^2\right)}.$$

Single forward-backward iteration

- Analytic approximation:

$$\boxed{\operatorname{prox}_{\hat{g}}^{\delta/2}(\boldsymbol{a}) \approx \operatorname{soft}_{\mu\delta/2}\left(\boldsymbol{a} - \delta\boldsymbol{\Psi}^\dagger\boldsymbol{\Phi}^\dagger(\boldsymbol{\Phi\Psi a} - \boldsymbol{y})/2\sigma^2\right)}.$$

Proximal MALA

Computing proximity operators for the synthesis case

- Recall posterior: $\pi(\boldsymbol{x}) \propto \exp\big(-g(\boldsymbol{x})\big)$.
- Let $\hat{g}(\boldsymbol{x}(\boldsymbol{a})) = \hat{f}_1(\boldsymbol{a}) + \hat{f}_2(\boldsymbol{a})$, where $\boxed{\hat{f}_1(\boldsymbol{a}) = \mu \|\boldsymbol{a}\|_1}$ and $\boxed{\hat{f}_2(\boldsymbol{a}) = \|\boldsymbol{y} - \boldsymbol{\Phi\Psi a}\|_2^2 / 2\sigma^2}$.

  Prior Likelihood
- Must solve an optimisation problem for each iteration!

$$\boxed{\mathrm{prox}_{\hat{g}}^{\delta/2}(\boldsymbol{a}) = \operatorname*{argmin}_{\boldsymbol{u}\in\mathbb{R}^L} \left\{ \mu\|\boldsymbol{u}\|_1 + \frac{\|\boldsymbol{y} - \boldsymbol{\Phi\Psi u}\|_2^2}{2\sigma^2} + \frac{\|\boldsymbol{u} - \boldsymbol{a}\|_2^2}{\delta} \right\}}\,.$$

- Taylor expansion at point $\boldsymbol{a}$: $\|\boldsymbol{y} - \boldsymbol{\Phi\Psi u}\|_2^2 \approx \|\boldsymbol{y} - \boldsymbol{\Phi\Psi a}\|_2^2 + 2(\boldsymbol{u} - \boldsymbol{a})^\top \boldsymbol{\Psi}^\dagger \boldsymbol{\Phi}^\dagger (\boldsymbol{\Phi\Psi a} - \boldsymbol{y})$.
- Then proximity operator approximated by

$$\boxed{\mathrm{prox}_{\hat{g}}^{\delta/2}(\boldsymbol{a}) \approx \mathrm{prox}_{\hat{f}_1}^{\delta/2}\left(\boldsymbol{a} - \delta \boldsymbol{\Psi}^\dagger \boldsymbol{\Phi}^\dagger (\boldsymbol{\Phi\Psi a} - \boldsymbol{y})/2\sigma^2\right)}\,.$$

  Single forward-backward iteration
- Analytic approximation:

$$\boxed{\mathrm{prox}_{\hat{g}}^{\delta/2}(\boldsymbol{a}) \approx \mathrm{soft}_{\mu\delta/2}\left(\boldsymbol{a} - \delta \boldsymbol{\Psi}^\dagger \boldsymbol{\Phi}^\dagger (\boldsymbol{\Phi\Psi a} - \boldsymbol{y})/2\sigma^2\right)}\,.$$

Proximal MALA

Computing proximity operators for the synthesis case

- Recall posterior: $\pi(\boldsymbol{x}) \propto \exp\big(-g(\boldsymbol{x})\big)$.
- Let $\hat{g}(\boldsymbol{x}(\boldsymbol{a})) = \hat{f}_1(\boldsymbol{a}) + \hat{f}_2(\boldsymbol{a})$, where $\boxed{\hat{f}_1(\boldsymbol{a}) = \mu\|\boldsymbol{a}\|_1}$ and $\boxed{\hat{f}_2(\boldsymbol{a}) = \|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{a}\|_2^2/2\sigma^2}$.

  Prior Likelihood
- Must solve an optimisation problem for each iteration!

$$\boxed{\mathrm{prox}_{\hat{g}}^{\delta/2}(\boldsymbol{a}) = \underset{\boldsymbol{u}\in\mathbb{R}^L}{\mathrm{argmin}} \left\{ \mu\|\boldsymbol{u}\|_1 + \frac{\|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{u}\|_2^2}{2\sigma^2} + \frac{\|\boldsymbol{u} - \boldsymbol{a}\|_2^2}{\delta} \right\}}\ .$$

- Taylor expansion at point $\boldsymbol{a}$: $\|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{u}\|_2^2 \approx \|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{a}\|_2^2 + 2(\boldsymbol{u} - \boldsymbol{a})^\top \boldsymbol{\Psi}^\dagger \boldsymbol{\Phi}^\dagger (\boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{a} - \boldsymbol{y})$.
- Then proximity operator approximated by

$$\boxed{\mathrm{prox}_{\hat{g}}^{\delta/2}(\boldsymbol{a}) \approx \mathrm{prox}_{\hat{f}_1}^{\delta/2}\left(\boldsymbol{a} - \delta\boldsymbol{\Psi}^\dagger\boldsymbol{\Phi}^\dagger(\boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{a} - \boldsymbol{y})/2\sigma^2\right)}\ .$$

Single forward-backward iteration

- Analytic approximation:

$$\boxed{\mathrm{prox}_{\hat{g}}^{\delta/2}(\boldsymbol{a}) \approx \mathrm{soft}_{\mu\delta/2}\left(\boldsymbol{a} - \delta\boldsymbol{\Psi}^\dagger\boldsymbol{\Phi}^\dagger(\boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{a} - \boldsymbol{y})/2\sigma^2\right)}\ .$$

MYULA
Moreau-Yosida approximation

- Moreau-Yosida approximation (Moreau envelope) of f:

$$\boxed{f_\lambda^{\mathsf{MY}}(\boldsymbol{x}) = \inf_{\boldsymbol{u}\in\mathbb{R}^N} f(\boldsymbol{u}) + \frac{\|\boldsymbol{u}-\boldsymbol{x}\|^2}{2\lambda}}$$

- Important properties of $f_\lambda^{\mathsf{MY}}(\boldsymbol{x})$:
  1. As $\lambda \rightarrow 0$, $f_\lambda^{\mathsf{MY}}(\boldsymbol{x}) \rightarrow f(\boldsymbol{x})$
  2. $\nabla f_\lambda^{\mathsf{MY}}(\boldsymbol{x}) = (\boldsymbol{x} - \mathrm{prox}_f^\lambda(\boldsymbol{x}))/\lambda$

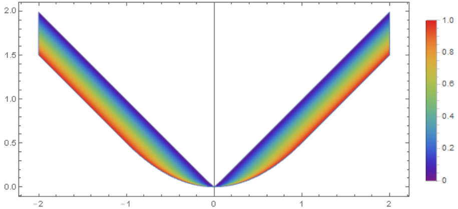


Figure: Illustration of Moreau-Yosida envelope of $|x|$ for varying λ [Credit: Stack exchange (ubpdqn)]

MYULA

Moreau-Yosida approximation

- Moreau-Yosida approximation (Moreau envelope) of f:

$$\boxed{f_\lambda^{\mathsf{MY}}(\boldsymbol{x}) = \inf_{\boldsymbol{u}\in\mathbb{R}^N} f(\boldsymbol{u}) + \frac{\|\boldsymbol{u}-\boldsymbol{x}\|^2}{2\lambda}}$$

- Important properties of $f_\lambda^{\mathsf{MY}}(\boldsymbol{x})$:

  1. As $\lambda \rightarrow 0$, $f_\lambda^{\mathsf{MY}}(\boldsymbol{x}) \rightarrow f(\boldsymbol{x})$
  2. $\nabla f_\lambda^{\mathsf{MY}}(\boldsymbol{x}) = (\boldsymbol{x} - \mathrm{prox}_f^\lambda(\boldsymbol{x}))/\lambda$

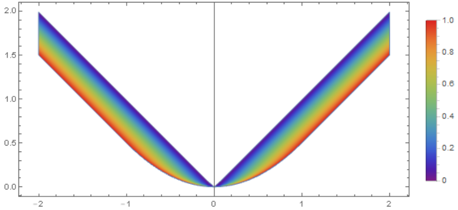


Figure: Illustration of Moreau-Yosida envelope of $|x|$ for varying λ [Credit: Stack exchange (ubpdqn)]

MYULA

MCMC sampling

Moreau-Yosida unadjusted Langevin algorithm (MYULA)
Durmus, Moulines & Pereyra (2016)

- Consider log-convex posteriors: $\mathrm{P}(\boldsymbol{x} \mid \boldsymbol{y}) = \pi(\boldsymbol{x}) \propto \exp\big(-g(\boldsymbol{x})\big)$, where

$$g(\boldsymbol{x}) = \underbrace{\boxed{f_1(\boldsymbol{x})}}_{\text{Convex}} + \underbrace{\boxed{f_2(\boldsymbol{x})}}_{\text{Smooth}} .$$

- Langevin diffusion process $\mathcal{L}(t)$, with π as stationary distribution ($\mathcal{W}$ Brownian motion):

$$\mathrm{d}\mathcal{L}(t) = \frac{1}{2}\nabla \log \pi\big(\mathcal{L}(t)\big)\mathrm{d}t + \mathrm{d}\mathcal{W}(t), \quad \mathcal{L}(0) = l_0 .$$

- Euler discretisation and apply Moreau-Yosida approximation to f_1:

$$\boldsymbol{l}^{(m+1)} = \boldsymbol{l}^{(m)} + \frac{\delta}{2}\boxed{\nabla \log \pi(\boldsymbol{l}^{(m)})} + \sqrt{\delta}\boldsymbol{w}^{(m)} .$$

$$\nabla \log \pi(\boldsymbol{x}) \approx \big(\mathrm{prox}_{f_1}^{\lambda}(\boldsymbol{x}) - \boldsymbol{x}\big)/\lambda - \nabla f_2(\boldsymbol{x})$$

- No Metropolis-Hastings accept-reject step. Converges geometrically fast, where bias can be made arbitrarily small. To achieve precision target ϵ requires:
  - Worst case: order $N^5 \log^2(\epsilon^{-1})\epsilon^{-2}$ iterations.
  - Strong convexity worst case: order $N \log(N) \log^2(\epsilon^{-1})\epsilon^{-2}$ iterations.

MYULA

MCMC sampling

Moreau-Yosida unadjusted Langevin algorithm (MYULA)

Durmus, Moulines & Pereyra (2016)

- Consider log-convex posteriors: $\mathrm{P}(\boldsymbol{x} \mid \boldsymbol{y}) = \pi(\boldsymbol{x}) \propto \exp\big(-g(\boldsymbol{x})\big)$, where

$$g(\boldsymbol{x}) = \underbrace{\boxed{f_1(\boldsymbol{x})}}_{\text{Convex}} + \underbrace{\boxed{f_2(\boldsymbol{x})}}_{\text{Smooth}} .$$

- Langevin diffusion process $\mathcal{L}(t)$, with π as stationary distribution ($\mathcal{W}$ Brownian motion):

$$\mathrm{d}\mathcal{L}(t) = \frac{1}{2}\nabla \log \pi\big(\mathcal{L}(t)\big)\mathrm{d}t + \mathrm{d}\mathcal{W}(t), \quad \mathcal{L}(0) = l_0 .$$

- Euler discretisation and apply Moreau-Yosida approximation to f_1:

$$\boldsymbol{l}^{(m+1)} = \boldsymbol{l}^{(m)} + \frac{\delta}{2}\boxed{\nabla \log \pi(\boldsymbol{l}^{(m)})} + \sqrt{\delta}\boldsymbol{w}^{(m)} .$$

$$\nabla \log \pi(\boldsymbol{x}) \approx \big(\mathrm{prox}_{f_1}^{\lambda}(\boldsymbol{x}) - \boldsymbol{x}\big)/\lambda - \nabla f_2(\boldsymbol{x})$$

- No Metropolis-Hastings accept-reject step. Converges geometrically fast, where bias can be made arbitrarily small. To achieve precision target ϵ requires:
  - Worst case: order $N^5 \log^2(\epsilon^{-1})\epsilon^{-2}$ iterations.
  - Strong convexity worst case: order $N \log(N) \log^2(\epsilon^{-1})\epsilon^{-2}$ iterations.

MYULA

MCMC sampling

Moreau-Yosida unadjusted Langevin algorithm (MYULA)
Durmus, Moulines & Pereyra (2016)

- Consider log-convex posteriors: $\mathrm{P}(\boldsymbol{x} \mid \boldsymbol{y}) = \pi(\boldsymbol{x}) \propto \exp\bigl(-g(\boldsymbol{x})\bigr)$, where

$$g(\boldsymbol{x}) = \underbrace{\boxed{f_1(\boldsymbol{x})}}_{\text{Convex}} + \underbrace{\boxed{f_2(\boldsymbol{x})}}_{\text{Smooth}} .$$

- Langevin diffusion process $\mathcal{L}(t)$, with π as stationary distribution ($\mathcal{W}$ Brownian motion):

$$\mathrm{d}\mathcal{L}(t) = \frac{1}{2} \nabla \log \pi\bigl(\mathcal{L}(t)\bigr)\mathrm{d}t + \mathrm{d}\mathcal{W}(t), \quad \mathcal{L}(0) = l_0 .$$

- Euler discretisation and apply Moreau-Yosida approximation to f_1:

$$\boldsymbol{l}^{(m+1)} = \boldsymbol{l}^{(m)} + \frac{\delta}{2} \boxed{\nabla \log \pi(\boldsymbol{l}^{(m)})} + \sqrt{\delta}\boldsymbol{w}^{(m)} .$$

$$\nabla \log \pi(\boldsymbol{x}) \approx \bigl(\mathrm{prox}_{f_1}^{\lambda}(\boldsymbol{x}) - \boldsymbol{x}\bigr)/\lambda - \nabla f_2(\boldsymbol{x})$$

- No Metropolis-Hastings accept-reject step. Converges geometrically fast, where bias can be made arbitrarily small. To achieve precision target ϵ requires:
  - Worst case: order $N^5 \log^2(\epsilon^{-1})\epsilon^{-2}$ iterations.
  - Strong convexity worst case: order $N \log(N) \log^2(\epsilon^{-1})\epsilon^{-2}$ iterations.

MYULA

MCMC sampling

Moreau-Yosida unadjusted Langevin algorithm (MYULA)

Durmus, Moulines & Pereyra (2016)

- Consider log-convex posteriors: $\mathrm{P}(\boldsymbol{x} \mid \boldsymbol{y}) = \pi(\boldsymbol{x}) \propto \exp\big(-g(\boldsymbol{x})\big)$, where

$$g(\boldsymbol{x}) = \underbrace{\boxed{f_1(\boldsymbol{x})}}_{\text{Convex}} + \underbrace{\boxed{f_2(\boldsymbol{x})}}_{\text{Smooth}} .$$

- Langevin diffusion process $\mathcal{L}(t)$, with π as stationary distribution ($\mathcal{W}$ Brownian motion):

$$\mathrm{d}\mathcal{L}(t) = \frac{1}{2}\nabla \log \pi\big(\mathcal{L}(t)\big)\mathrm{d}t + \mathrm{d}\mathcal{W}(t), \quad \mathcal{L}(0) = l_0 .$$

- Euler discretisation and apply Moreau-Yosida approximation to f_1:

$$\boldsymbol{l}^{(m+1)} = \boldsymbol{l}^{(m)} + \frac{\delta}{2}\boxed{\nabla \log \pi(\boldsymbol{l}^{(m)})} + \sqrt{\delta}\boldsymbol{w}^{(m)} .$$

$$\nabla \log \pi(\boldsymbol{x}) \approx \big(\mathrm{prox}_{f_1}^{\lambda}(\boldsymbol{x}) - \boldsymbol{x}\big)/\lambda - \nabla f_2(\boldsymbol{x})$$

- No Metropolis-Hastings accept-reject step. Converges geometrically fast, where bias can be made arbitrarily small. To achieve precision target ϵ requires:
  - Worst case: order $N^5 \log^2(\epsilon^{-1})\epsilon^{-2}$ iterations.
  - Strong convexity worst case: order $N \log(N) \log^2(\epsilon^{-1})\epsilon^{-2}$ iterations.

MYULA

Computing proximity operators for the analysis case

- Recall posterior: $\pi(\boldsymbol{x}) \propto \exp\big(-g(\boldsymbol{x})\big)$.
- Let $\bar{g}(\boldsymbol{x}) = \bar{f}_1(\boldsymbol{x}) + \bar{f}_2(\boldsymbol{x})$, where $\underbrace{\boxed{\bar{f}_1(\boldsymbol{x}) = \mu\|\boldsymbol{\Psi}^\dagger \boldsymbol{x}\|_1}}_{\text{Prior}}$ and $\underbrace{\boxed{\bar{f}_2(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{x}\|_2^2/2\sigma^2}}_{\text{Likelihood}}$.
- Only need to compute proximity operator of $\bar{f}_1$, which can be computed analytically without any approximation:

$$\boxed{\mathrm{prox}_{\bar{f}_1}^{\delta/2}(\boldsymbol{x}) = \boldsymbol{x} + \boldsymbol{\Psi}\left(\mathrm{soft}_{\mu\delta/2}(\boldsymbol{\Psi}^\dagger \boldsymbol{x}) - \boldsymbol{\Psi}^\dagger \boldsymbol{x}\right)}.$$

MYULA

Computing proximity operators for the analysis case

- Recall posterior: $\pi(\boldsymbol{x}) \propto \exp\bigl(-g(\boldsymbol{x})\bigr)$.
- Let $\bar{g}(\boldsymbol{x}) = \bar{f}_1(\boldsymbol{x}) + \bar{f}_2(\boldsymbol{x})$, where $\underbrace{\boxed{\bar{f}_1(\boldsymbol{x}) = \mu \|\boldsymbol{\Psi}^\dagger \boldsymbol{x}\|_1}}_{\text{Prior}}$ and $\underbrace{\boxed{\bar{f}_2(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{x}\|_2^2 / 2\sigma^2}}_{\text{Likelihood}}$.
- Only need to compute proximity operator of f_1, which can be computed analytically without any approximation:

$$\boxed{\mathrm{prox}_{\bar{f}_1}^{\delta/2}(\boldsymbol{x}) = \boldsymbol{x} + \boldsymbol{\Psi}\left(\mathrm{soft}_{\mu\delta/2}(\boldsymbol{\Psi}^\dagger \boldsymbol{x}) - \boldsymbol{\Psi}^\dagger \boldsymbol{x}\right)} \, .$$

MYULA

Computing proximity operators for the synthesis case

- Recall posterior: $\pi(\boldsymbol{x}) \propto \exp\big(-g(\boldsymbol{x})\big)$.
- Let $\hat{g}(\boldsymbol{x}(\boldsymbol{a})) = \hat{f}_1(\boldsymbol{a}) + \hat{f}_2(\boldsymbol{a})$, where $\boxed{\hat{f}_1(\boldsymbol{a}) = \mu \|\boldsymbol{a}\|_1}$ and $\boxed{\hat{f}_2(\boldsymbol{a}) = \|\boldsymbol{y} - \boldsymbol{\Phi\Psi a}\|_2^2 / 2\sigma^2}$.
  Prior Likelihood
- Only need to compute proximity operator of f_1, which can be computed analytically without any approximation:

$$\boxed{\mathrm{prox}_{\hat{f}_1}^{\delta/2}(\boldsymbol{a}) = \mathrm{soft}_{\mu\delta/2}(\boldsymbol{a})}\,.$$

MYULA

Computing proximity operators for the synthesis case

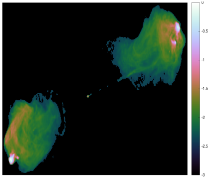
- Recall posterior: $\pi(\boldsymbol{x}) \propto \exp\big(-g(\boldsymbol{x})\big)$.
- Let $\hat{g}(\boldsymbol{x}(\boldsymbol{a})) = \hat{f}_1(\boldsymbol{a}) + \hat{f}_2(\boldsymbol{a})$, where $\boxed{\hat{f}_1(\boldsymbol{a}) = \mu \|\boldsymbol{a}\|_1}$ and $\boxed{\hat{f}_2(\boldsymbol{a}) = \|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{a}\|_2^2 / 2\sigma^2}$.

  Prior Likelihood
- Only need to compute proximity operator of f_1, which can be computed analytically without any approximation:

$$\boxed{\mathrm{prox}_{\hat{f}_1}^{\delta/2}(\boldsymbol{a}) = \mathrm{soft}_{\mu\delta/2}(\boldsymbol{a})}\,.$$

Numerical experiments

MYULA with analysis model

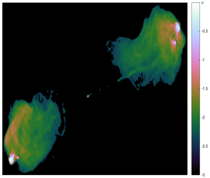


(a) Ground truth

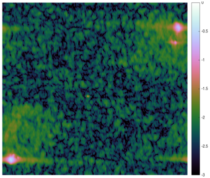
Figure: Cygnus A

Numerical experiments

MYULA with analysis model



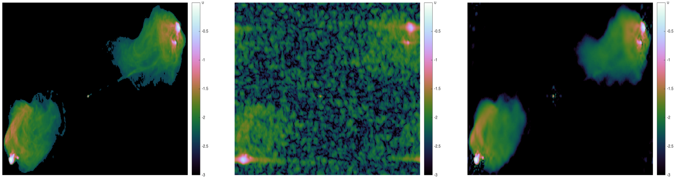
(a) Ground truth



(b) Dirty image

Figure: Cygnus A

Numerical experiments
MYULA with analysis model

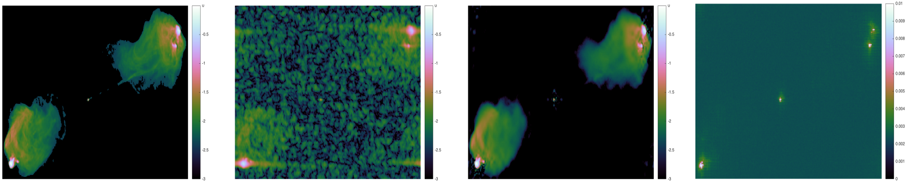


(a) Ground truth (b) Dirty image (c) Mean recovered image

Figure: Cygnus A

Numerical experiments

MYULA with analysis model

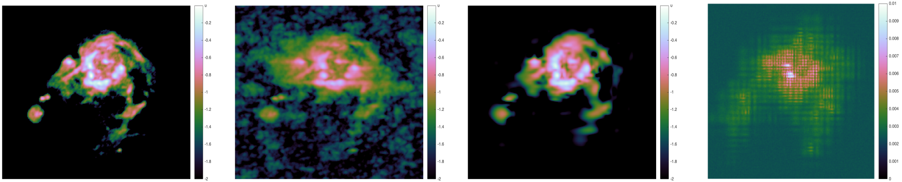


(a) Ground truth (b) Dirty image (c) Mean recovered image (d) Credible interval length

Figure: Cygnus A

Numerical experiments

MYULA with analysis model

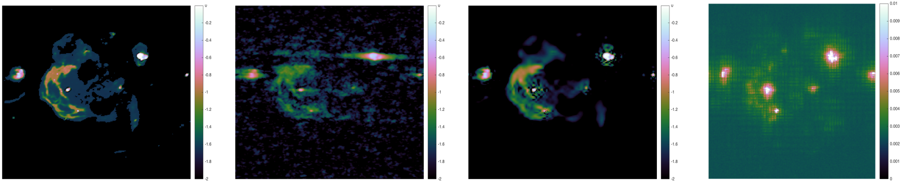


(a) Ground truth (b) Dirty image (c) Mean recovered image (d) Credible interval length

Figure: HII region of M31

Numerical experiments

MYULA with analysis model

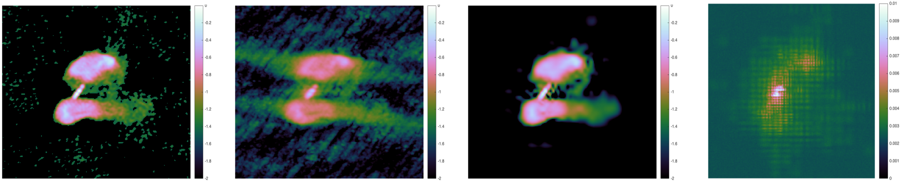


(a) Ground truth (b) Dirty image (c) Mean recovered image (d) Credible interval length

Figure: W28 Supernova remnant

Numerical experiments

MYULA with analysis model



(a) Ground truth (b) Dirty image (c) Mean recovered image (d) Credible interval length

Figure: 3C288

Numerical experiments

Computation time

Table: CPU time in minutes for Proximal MCMC sampling

Image	Method	CPU time (min) Analysis	CPU time (min) Synthesis
Cygnus A	P-MALA	2274	1762
	MYULA	1056	942
M31	P-MALA	1307	944
	MYULA	618	581
W28	P-MALA	1122	879
	MYULA	646	598
3C288	P-MALA	1144	881
	MYULA	607	538

Hypothesis testing
Method

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

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 artefact with confidence $(1-\alpha)\%$, i.e.
   - If $x' \in C_\alpha$ 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_α 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_α 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_α 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_α 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 $x' \notin C_\alpha$ then reject hypothesis that structure is an artifact with confidence $(1-\alpha)\%$, *i.e.* structure most likely physical.
   - If $x' \in C_\alpha$ 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_α 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

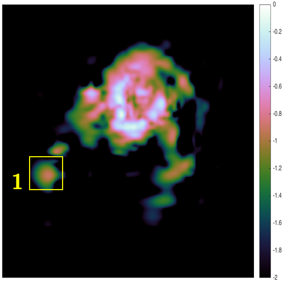
- Perform hypothesis tests of image structure using Bayesian credible regions (Pereyra 2016b).
- Let C_α denote the highest posterior density (HPD) Bayesian credible region with confidence level $(1-\alpha)\%$ defined by posterior iso-contour: $C_\alpha = \{\boldsymbol{x} : g(\boldsymbol{x}) \leq \gamma_\alpha\}$.

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

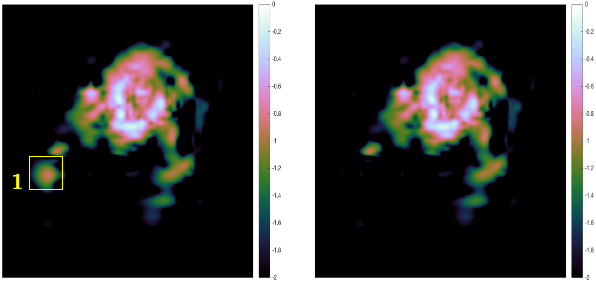


(a) Recovered image

Figure: HII region of M31

Hypothesis testing

Numerical experiments



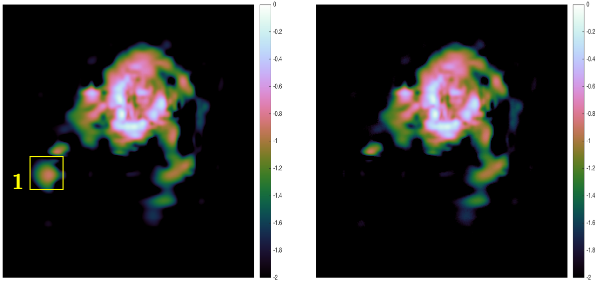
(a) Recovered image

(b) Surrogate with region removed

Figure: HII region of M31

Hypothesis testing

Numerical experiments



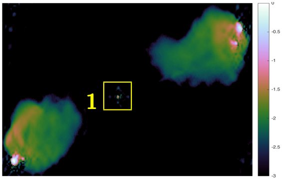
(a) Recovered image (b) Surrogate with region removed

Figure: HII region of M31

1. Reject null hypothesis

$\Rightarrow$ structure physical

Hypothesis testing
Numerical experiments

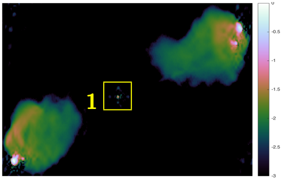


(a) Recovered image

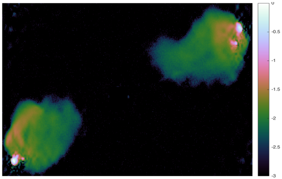
Figure: Cygnus A

Hypothesis testing

Numerical experiments



(a) Recovered image

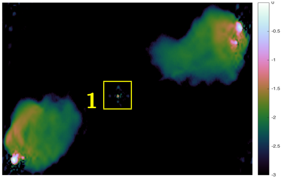


(b) Surrogate with region removed

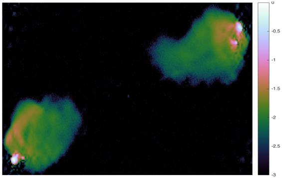
Figure: Cygnus A

Hypothesis testing

Numerical experiments



(a) Recovered image



(b) Surrogate with region removed

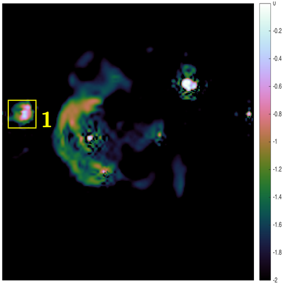
Figure: Cygnus A

1. Cannot reject null hypothesis

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

Hypothesis testing

Numerical experiments

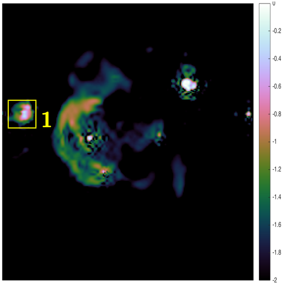


(a) Recovered image

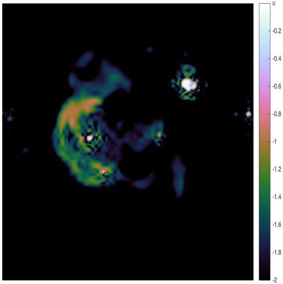
Figure: Supernova remnant W28

Hypothesis testing

Numerical experiments



(a) Recovered image

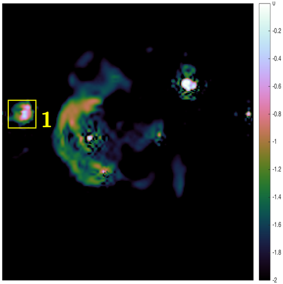


(b) Surrogate with region removed

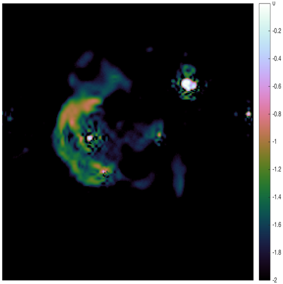
Figure: Supernova remnant W28

Hypothesis testing

Numerical experiments



(a) Recovered image



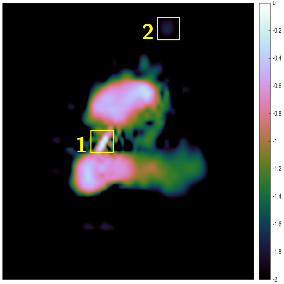
(b) Surrogate with region removed

Figure: Supernova remnant W28

1. Reject null hypothesis

$\Rightarrow$ structure physical

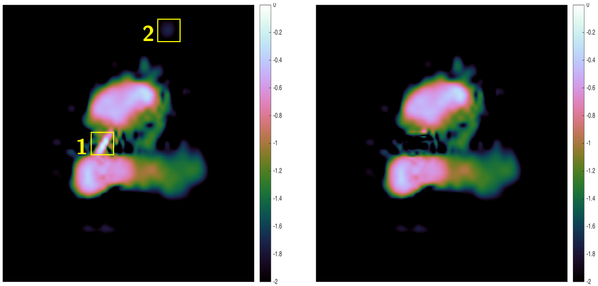
Hypothesis testing
Numerical experiments



(a) Recovered image

Figure: 3C288

Hypothesis testing
Numerical experiments

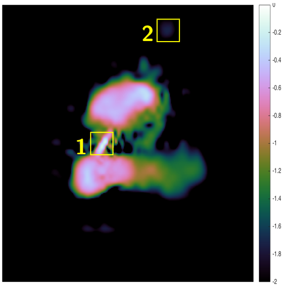


(a) Recovered image (b) Surrogate with region removed

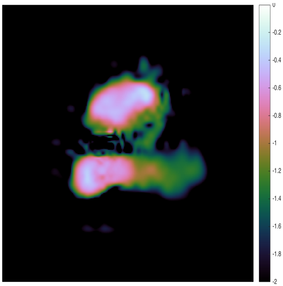
Figure: 3C288

Hypothesis testing

Numerical experiments



(a) Recovered image



(b) Surrogate with region removed

Figure: 3C288

1. Reject null hypothesis

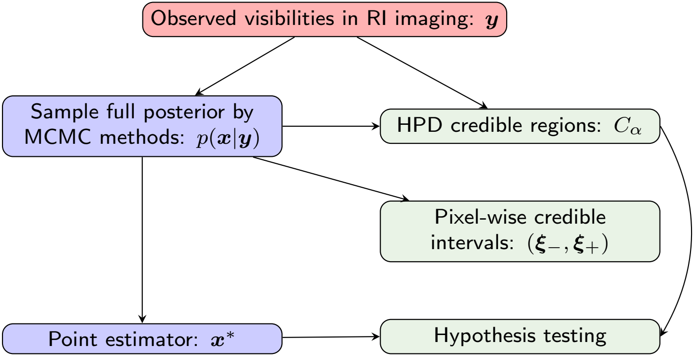
   ⇒ structure physical

2. Cannot reject null hypothesis

   ⇒ cannot make strong statistical statement about origin of structure

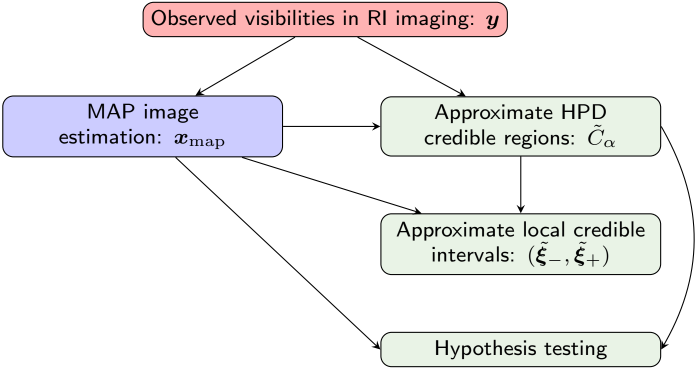
Outline

Proximal MCMC sampling and uncertainty quantification



Uncertainty quantification for radio interferometric imaging: I. proximal MCMC methods (Cai, Pereyra &McEwen 2018a; arXiv:1711.04818)

MAP estimation and uncertainty quantification



Uncertainty quantification for radio interferometric imaging: II. MAP estimation
(Cai, Pereyra &McEwen 2018b; arXiv:1711.04819)

Approximate Bayesian credible regions for MAP estimation

- Combine uncertainty quantification with fast sparse regularisation to scale to big-data.
- Recall C_α 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 γ_α:

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

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_α 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 γ_α:

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

  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_α 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 γ_α:

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

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 2018b)

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

Consider the test image with the Ω region replaced by constant value ξ:

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

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 2018b)

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

Consider the test image with the Ω region replaced by constant value ξ:

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

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

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

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

Local Bayesian credible intervals for MAP estimation

Local Bayesian credible intervals for sparse reconstruction
(Cai, Pereyra & McEwen 2018b)

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

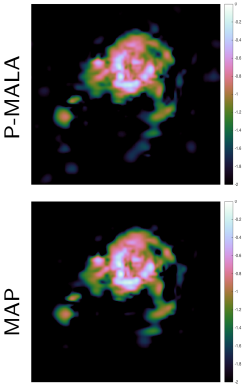
Consider the test image with the Ω region replaced by constant value ξ:

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

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

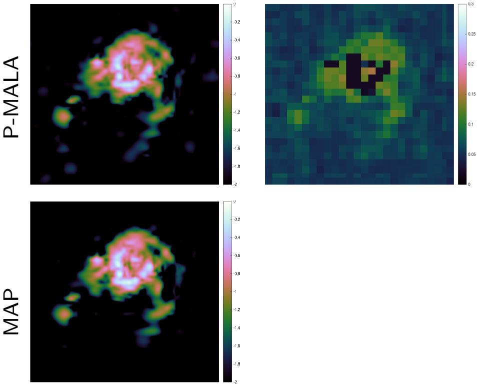
(b) local credible interval (grid size 10×10 pixels)

(c) local credible interval (grid size 20×20 pixels)

(d) local credible interval (grid size 30×30 pixels)

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

Numerical experiments



(a) point estimators

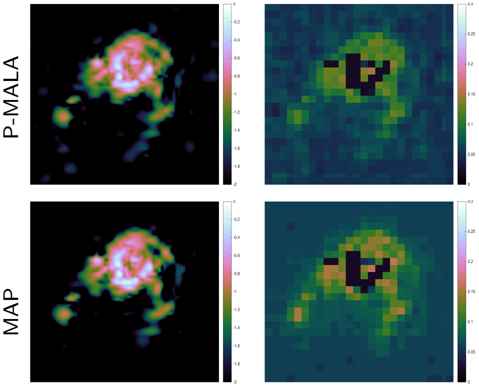
(b) local credible interval (grid size 10×10 pixels)

(c) local credible interval (grid size 20×20 pixels)

(d) local credible interval (grid size 30×30 pixels)

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

Numerical experiments



(a) point estimators

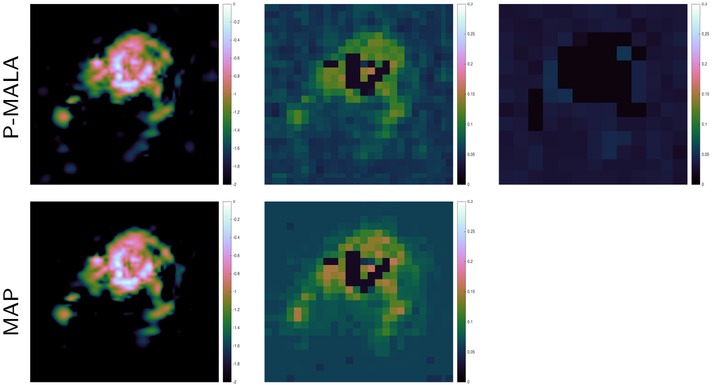
(b) local credible interval (grid size 10×10 pixels)

(c) local credible interval (grid size 20×20 pixels)

(d) local credible interval (grid size 30×30 pixels)

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

Numerical experiments



(a) point estimators

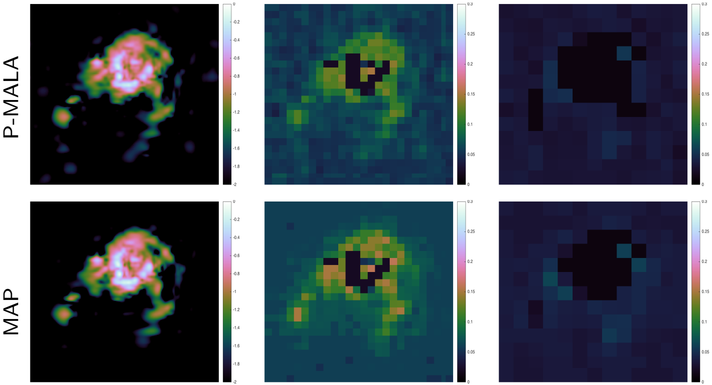
(b) local credible interval (grid size 10×10 pixels)

(c) local credible interval (grid size 20×20 pixels)

(d) local credible interval (grid size 30×30 pixels)

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

Numerical experiments



(a) point estimators

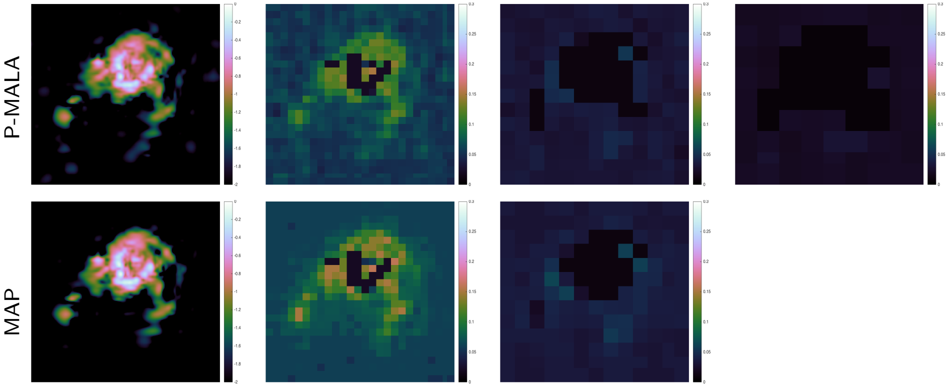
(b) local credible interval (grid size 10×10 pixels)

(c) local credible interval (grid size 20×20 pixels)

(d) local credible interval (grid size 30×30 pixels)

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

Numerical experiments



(a) point estimators

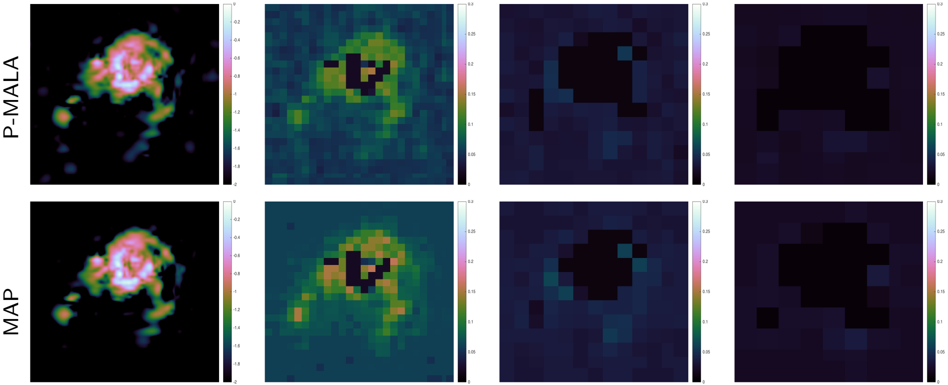
(b) local credible interval (grid size 10×10 pixels)

(c) local credible interval (grid size 20×20 pixels)

(d) local credible interval (grid size 30×30 pixels)

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

Numerical experiments



(a) point estimators

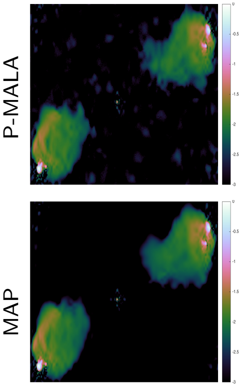
(b) local credible interval (grid size 10×10 pixels)

(c) local credible interval (grid size 20×20 pixels)

(d) local credible interval (grid size 30×30 pixels)

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

Numerical experiments



(a) point estimators

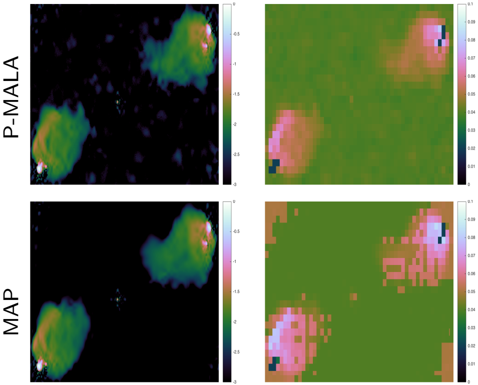
(b) local credible interval (grid size 10×10 pixels)

(c) local credible interval (grid size 20×20 pixels)

(d) local credible interval (grid size 30×30 pixels)

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

Numerical experiments



(a) point estimators

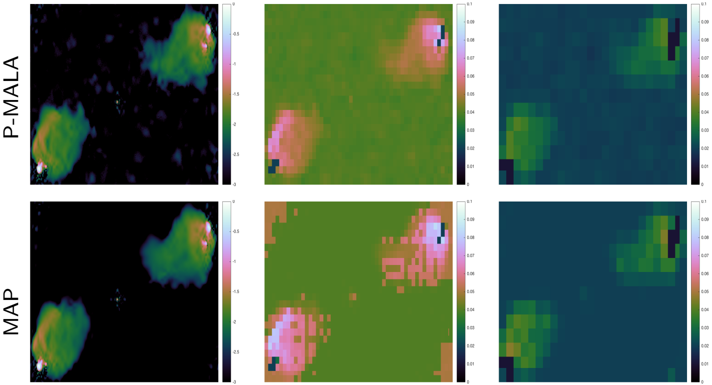
(b) local credible interval (grid size 10×10 pixels)

(c) local credible interval (grid size 20×20 pixels)

(d) local credible interval (grid size 30×30 pixels)

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

Numerical experiments



(a) point estimators

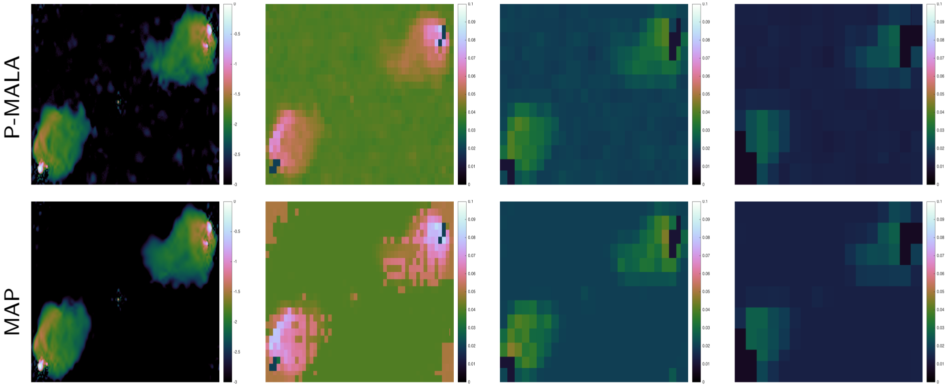
(b) local credible interval (grid size 10×10 pixels)

(c) local credible interval (grid size 20×20 pixels)

(d) local credible interval (grid size 30×30 pixels)

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

Numerical experiments



(a) point estimators

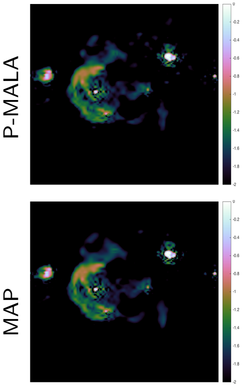
(b) local credible interval (grid size 10×10 pixels)

(c) local credible interval (grid size 20×20 pixels)

(d) local credible interval (grid size 30×30 pixels)

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

Numerical experiments



(a) point estimators

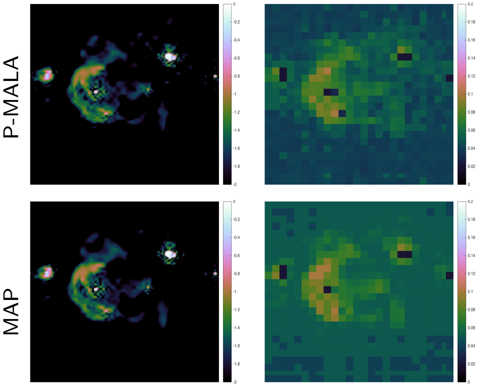
(b) local credible interval (grid size 10×10 pixels)

(c) local credible interval (grid size 20×20 pixels)

(d) local credible interval (grid size 30×30 pixels)

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

Numerical experiments



(a) point estimators

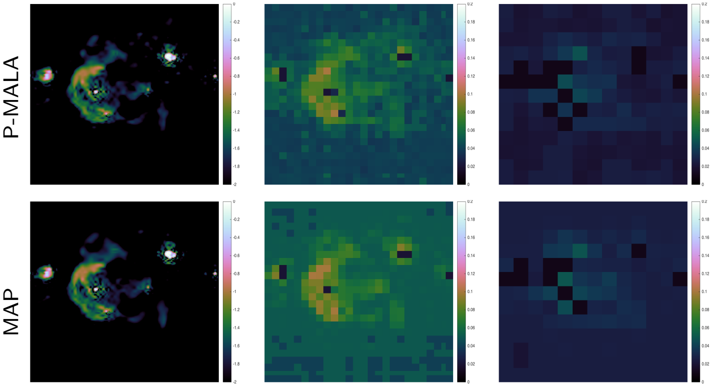
(b) local credible interval (grid size 10×10 pixels)

(c) local credible interval (grid size 20×20 pixels)

(d) local credible interval (grid size 30×30 pixels)

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

Numerical experiments



(a) point estimators

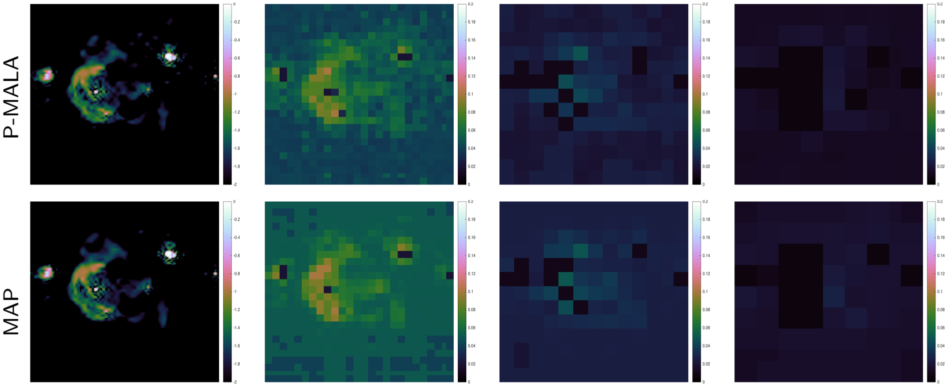
(b) local credible interval (grid size 10×10 pixels)

(c) local credible interval (grid size 20×20 pixels)

(d) local credible interval (grid size 30×30 pixels)

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

Numerical experiments



(a) point estimators

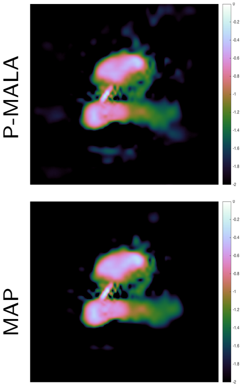
(b) local credible interval (grid size 10×10 pixels)

(c) local credible interval (grid size 20×20 pixels)

(d) local credible interval (grid size 30×30 pixels)

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

Numerical experiments



(a) point estimators

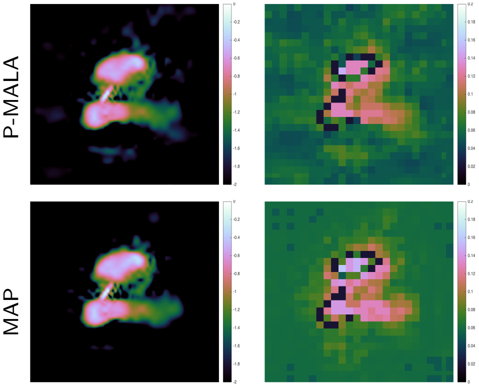
(b) local credible interval (grid size 10×10 pixels)

(c) local credible interval (grid size 20×20 pixels)

(d) local credible interval (grid size 30×30 pixels)

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

Numerical experiments



(a) point estimators

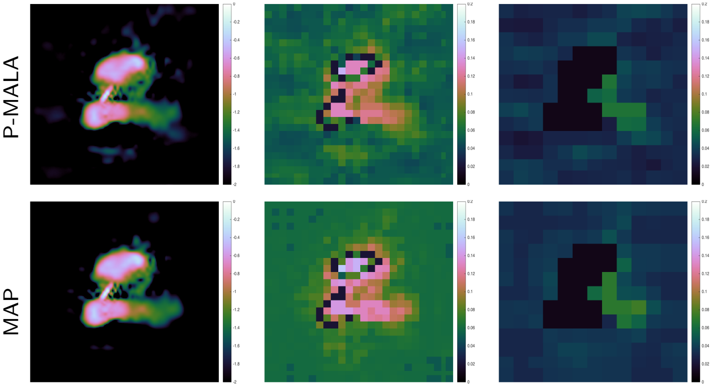
(b) local credible interval (grid size 10×10 pixels)

(c) local credible interval (grid size 20×20 pixels)

(d) local credible interval (grid size 30×30 pixels)

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

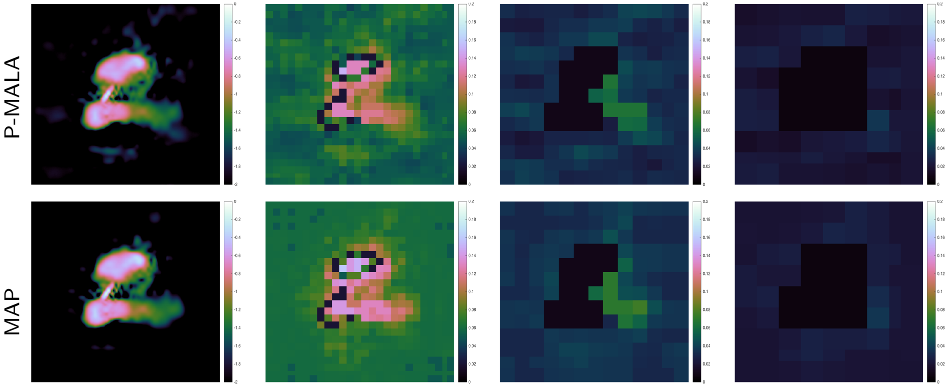
Numerical experiments



(a) point estimators (b) local credible interval (grid size 10×10 pixels) (c) local credible interval (grid size 20×20 pixels) (d) local credible interval (grid size 30×30 pixels)

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

Numerical experiments



(a) point estimators

(b) local credible interval (grid size 10×10 pixels)

(c) local credible interval (grid size 20×20 pixels)

(d) local credible interval (grid size 30×30 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

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

1. Radio interferometric imaging
2. Uncertainty quantification (MCMC sampling)
3. Uncertainty quantification (MAP estimation)
4. Online imaging

Online imaging

- Online radio interferometric imaging (Cai, Pratley & McEwen 2019; arXiv:1712.04462)
- Perform **image reconstruction simultaneously with data acquisition**.
  - Assimilate data on arrival and then discard.
  - Dramatically **reduces data storage requirements**.
  - Additional **computational savings**.
  - **Theoretical guarantee** that recover same fidelity as offline approach.

Online imaging

- Online radio interferometric imaging (Cai, Pratley & McEwen 2019; arXiv:1712.04462)
- Perform **image reconstruction simultaneously with data acquisition**.
  - Assimilate data on arrival and then discard.
  - Dramatically **reduces data storage requirements**.
  - Additional **computational savings**.
  - **Theoretical guarantee** that recover same fidelity as offline approach.

Online imaging
Algorithm overview

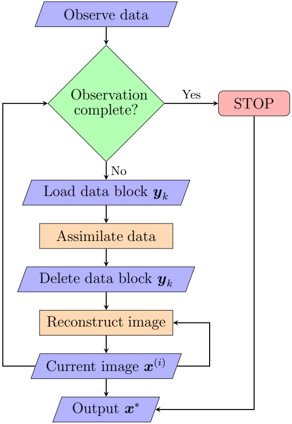


Figure: Online radio interferometric imaging.

Online imaging

Storage and computational savings

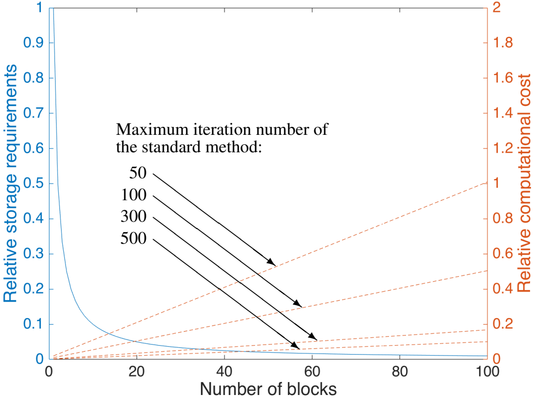
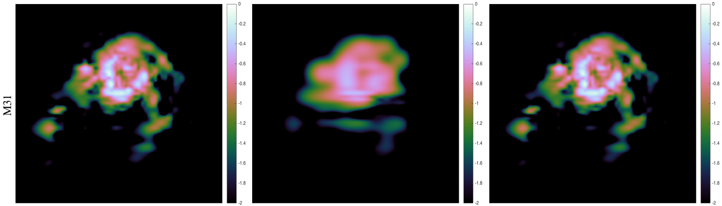


Figure: Storage and computational savings.

Online imaging
Image reconstruction



Offline algo
(storage 100% visibilities)

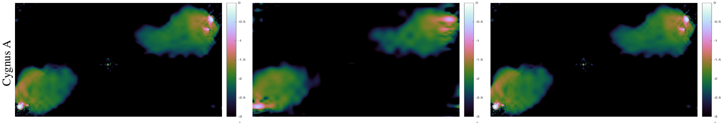
Offline algo
(storage 2% visibilities)

Online algo
(storage 2% visibilities)

Figure: Comparison between images reconstructed by the offline and online algorithms for M31.

Online imaging

Image reconstruction



Offline algo
(storage 100% visibilities)

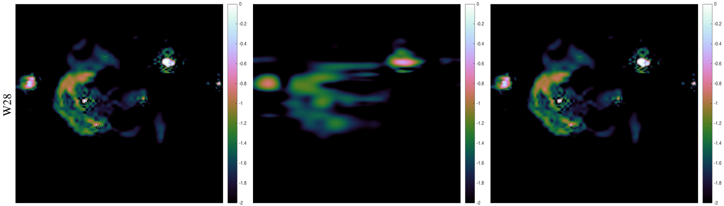
Offline algo
(storage 2% visibilities)

Online algo
(storage 2% visibilities)

Figure: Comparison between images reconstructed by the offline and online algorithms for Cygnus A.

Online imaging

Image reconstruction

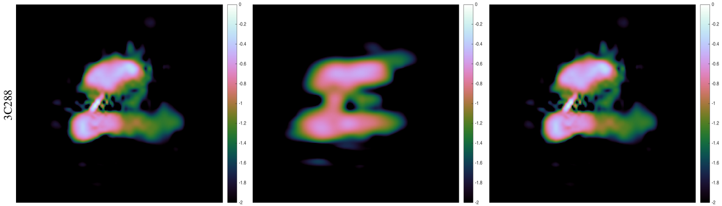


Offline algo (storage 100% visibilities) | Offline algo (storage 2% visibilities) | Online algo (storage 2% visibilities)

Figure: Comparison between images reconstructed by the offline and online algorithms for W28.

Online imaging

Image reconstruction



Offline algo
(storage 100% visibilities)

Offline algo
(storage 2% visibilities)

Online algo
(storage 2% visibilities)

Figure: Comparison between images reconstructed by the offline and online algorithms for 3C288.

Online imaging

Reconstruction fidelity

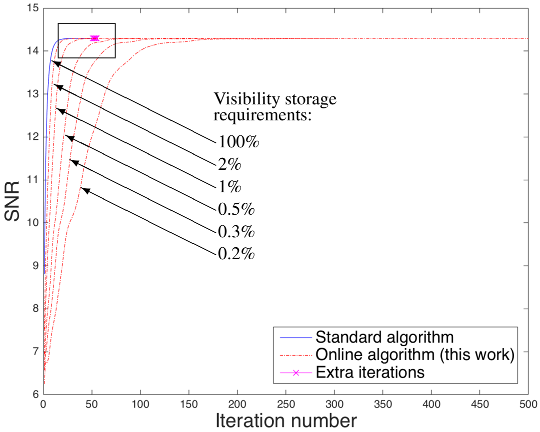


Figure: SNR vs iteration number for M31.

Online imaging
Reconstruction fidelity

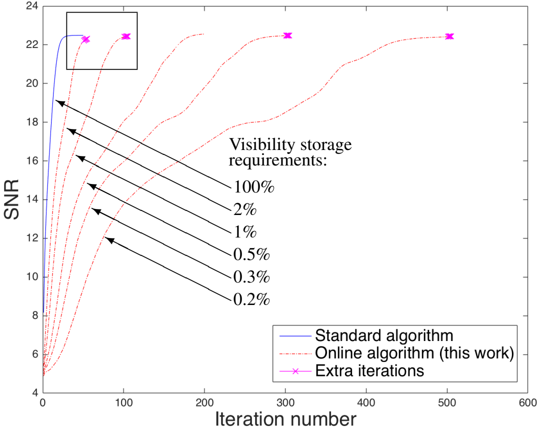


Figure: SNR vs iteration number for Cygnus A.

Online imaging
Reconstruction fidelity

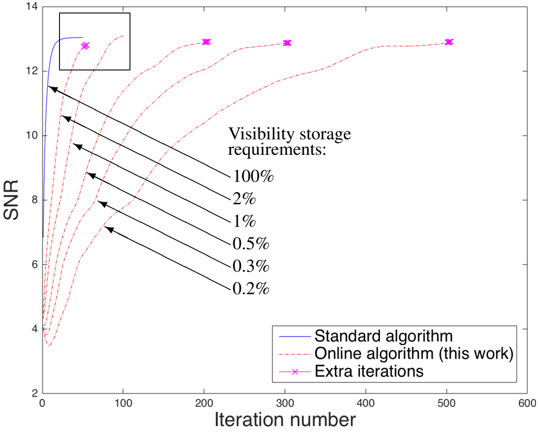


Figure: SNR vs iteration number for W28.

Online imaging
Reconstruction fidelity

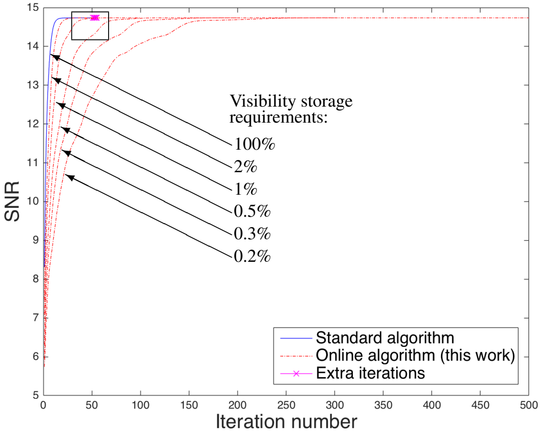


Figure: SNR vs iteration number for 3C288.

Conclusions

1. **Sparse priors** for radio interferometry **highly effective**, with **efficient implementations**.
   - PURIFY code provides robust framework for imaging interferometric observations (`http://astro-informatics.github.io/purify/`).
   - SOPT code for efficient and distributed sparse regularisation (`http://astro-informatics.github.io/sopt/`).
2. **Uncertainty quantification** to support sparse priors efficiently in full Bayesian framework:
   - Recover **Bayesian credible intervals**.
   - Perform **hypothesis testing** to test whether structure physical.
3. **Online imaging** to perform imaging simultaneously with data acquisition:
   - Dramatically **reduce storage requirements**.
   - Additional **computational savings**.

Potential to apply to SPIDER imaging.

Supported by:

EPSRC Engineering and Physical Sciences Research Council

The Leverhulme Trust

Science & Technology Facilities Council

Conclusions

1. **Sparse priors** for radio interferometry **highly effective**, with **efficient implementations**.
   - PURIFY code provides robust framework for imaging interferometric observations (`http://astro-informatics.github.io/purify/`).
   - SOPT code for efficient and distributed sparse regularisation (`http://astro-informatics.github.io/sopt/`).
2. **Uncertainty quantification** to support sparse priors efficiently in full Bayesian framework:
   - Recover **Bayesian credible intervals**.
   - Perform **hypothesis testing** to test whether structure physical.
3. **Online imaging** to perform imaging simultaneously with data acquisition:
   - Dramatically **reduce storage requirements**.
   - Additional **computational savings**.

Potential to apply to SPIDER imaging.

Supported by:

EPSRC Engineering and Physical Sciences Research Council

The Leverhulme Trust

Science & Technology Facilities Council

Conclusions

1. **Sparse priors** for radio interferometry **highly effective**, with **efficient implementations**.
   - PURIFY code provides robust framework for imaging interferometric observations (`http://astro-informatics.github.io/purify/`).
   - SOPT code for efficient and distributed sparse regularisation (`http://astro-informatics.github.io/sopt/`).
2. **Uncertainty quantification** to support sparse priors efficiently in full Bayesian framework:
   - Recover **Bayesian credible intervals**.
   - Perform **hypothesis testing** to test whether structure physical.
3. **Online imaging** to perform imaging simultaneously with data acquisition:
   - Dramatically **reduce storage requirements**.
   - Additional **computational savings**.

Potential to apply to SPIDER imaging.

Supported by:

EPSRC Engineering and Physical Sciences Research Council

The Leverhulme Trust

Science & Technology Facilities Council

Conclusions

1. **Sparse priors** for radio interferometry **highly effective**, with **efficient implementations**.
   - PURIFY code provides robust framework for imaging interferometric observations (`http://astro-informatics.github.io/purify/`).
   - SOPT code for efficient and distributed sparse regularisation (`http://astro-informatics.github.io/sopt/`).
2. **Uncertainty quantification** to support sparse priors efficiently in full Bayesian framework:
   - Recover **Bayesian credible intervals**.
   - Perform **hypothesis testing** to test whether structure physical.
3. **Online imaging** to perform imaging simultaneously with data acquisition:
   - Dramatically **reduce storage requirements**.
   - Additional **computational savings**.

Potential to apply to SPIDER imaging.

Supported by:

EPSRC Engineering and Physical Sciences Research Council

The Leverhulme Trust

Science & Technology Facilities Council