

EXCALIBUR
10

Learned Exascale Computational Imaging (LEXCI)

UNLOCKING NEW HORIZONS IN RECOVERING IMAGES FROM RAW DATA BY LEVERAGING EXASCALE COMPUTATIONAL IMAGING

Jason McEwen (PI)
Mullard Space Science Laboratory (MSSL)
University College London (UCL)

ExCALIBUR workshop 2024



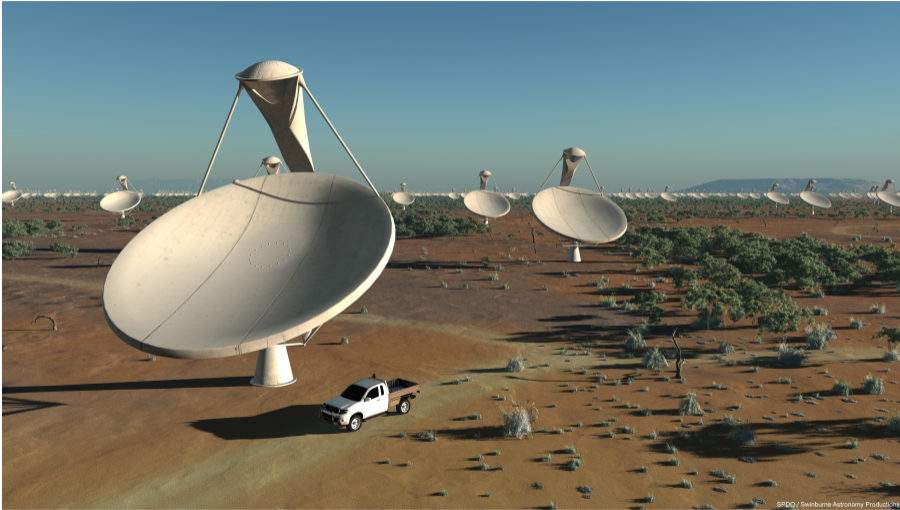
UK Research
and Innovation



UK Atomic
Energy
Authority

SKA Exascale

Square Kilometre Array (SKA): next-gen radio interferometric telescope



SKA science goals

Orders of magnitude improvement in sensitivity and resolution.

Unlock broad range of science goals.



Probing
the
cosmic
dawn

→



Challenging
Einstein

→



Cosmology
and dark
energy

→



Exploring
galaxy
evolution

→



Our
home
galaxy

→



Seeking
the
origins of
life

→



Studying
our
nearest
star

→



Understanding
cosmic
magnetism

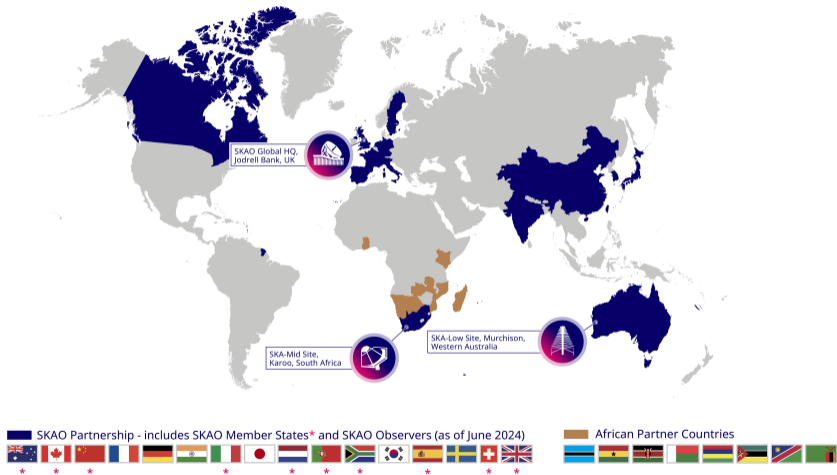
→



The
bursting
sky

→


SKA partners




SKA sites and data rates


SKA-mid – the SKA's mid-frequency instrument

The SKA Observatory (SKAO) is a next-generation radio astronomy facility that will revolutionise our understanding of the Universe. It will have a uniquely distributed character: one observatory operating two telescopes on three continents. The two telescopes, named SKA-low and SKA-mid, will be observing the Universe at different frequencies. They are also called interferometers as they each comprise a large number of individual elements working together to form a single large telescope.







Location: South Africa




Frequency range:
350 MHz to 15.4 GHz
with a goal of 24 GHz




197 dishes
(including 42 steerable dishes)



Total collecting area:
33,000m²
or
126 tennis courts



Maximum distance between dishes:
150km



Data transfer rate:
8.8 Terabits per second

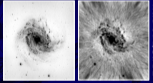




Image quality of SKA-mid (left) versus the best current facility operating in the same frequency range, the Jansky Very Large Array (JVA) in the United States (right). SKA-mid's resolution will be 4x better than JVA.



Compared to the JVA, the current best similar instrument in the world:


4x the resolution **5x** more sensitive **60x** the survey speed


www.skatelescope.org @SKAO f SKA Observatory in SKA Observatory SKA Observatory @skaoobservatory




SKA-low – the SKA's low-frequency instrument

The SKA Observatory (SKAO) is a next-generation radio astronomy facility that will revolutionise our understanding of the Universe. It will have a uniquely distributed character: one observatory operating two telescopes on three continents. The two telescopes, named SKA-low and SKA-mid, will be observing the Universe at different frequencies. They are also called interferometers as they each comprise a large number of individual elements working together to form a single large telescope.







Location: Australia



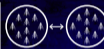
Frequency range:
50 MHz to 350 MHz




131,072 antennas spread between 512 stations



Total collecting area:
0.4km²



Maximum distance between stations:
>65km



Data transfer rate:
7.2 Terabits per second

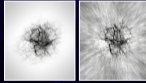




Image quality of SKA-low (left) versus the best current facility operating in the same frequency range, the LOFAR in the Netherlands (right). SKA-low's resolution will be similar to LOFAR.



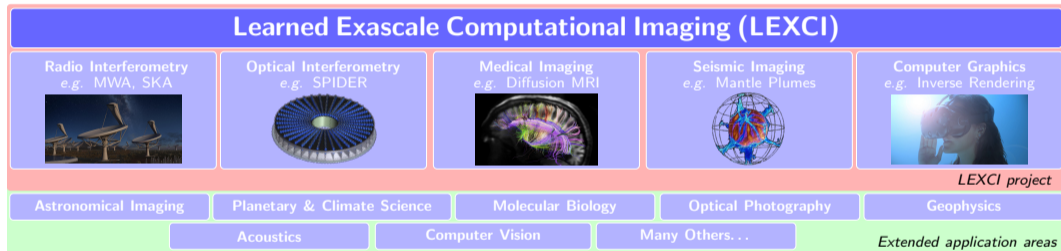
Compared to LOFAR Netherlands, the current best similar instrument in the world:

25% better resolution **8x** more sensitive **135x** the survey speed

www.skatelescope.org @SKAO f SKA Observatory in SKA Observatory SKA Observatory @skaoobservatory



Application domains more broadly

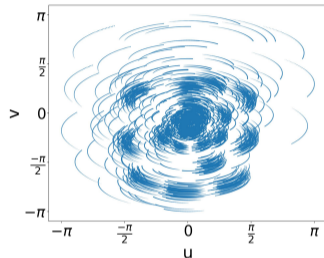


Imaging Strategy

Radio interferometric telescopes acquire “Fourier” measurements



“Fourier”
Measurements



Interferometric imaging is an **exascale computational inverse imaging problem**.

Radio interferometric inverse problem

Radio interferometric imaging ill-posed inverse problem:

$$y = \Phi(x) + n$$

$$y \xleftarrow{\text{forward model}} x$$

$$y \xrightarrow{\text{inverse inference}} x$$

for data (visibilities) y , telescope model Φ , underlying image x and noise n .

Big-Data \Rightarrow Big-Compute

since compute scales as $\mathcal{O}(M)$ for M data measurements.

Optimisation vs sampling

Inverse problem is ill-posed \Rightarrow **inject regularising prior information.**

MAP estimation

- + Based on optimisation so **computationally efficient.**
- **No uncertainties** (traditionally).
- **Hand-crafted priors** (traditionally).

MCMC sampling

- Based on sampling so **computationally demanding.**
- + **Uncertainties** encoded in posterior.
- **Hand-crafted priors** (traditionally).

Goals:

- + **Computationally efficient** (optimisation + distribution).
- + **Quantifies uncertainties**.
- + **Data-driven AI priors** (enhance reconstruction fidelity).

Goals:

- + **Computationally efficient** (optimisation + distribution).
- + **Quantifies uncertainties**.
- + **Data-driven AI priors** (enhance reconstruction fidelity).

Achieve by combining:

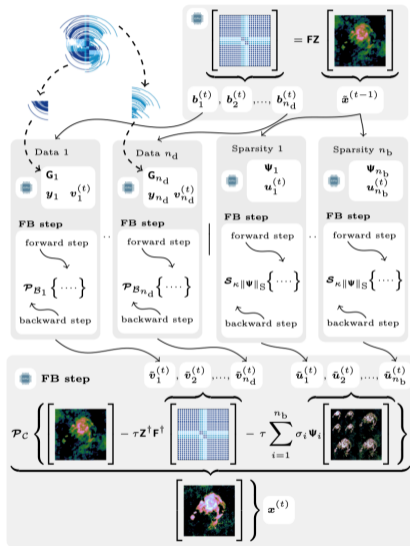
1. **Statistical framework:** Bayesian inference and MAP estimation.
2. **Mathematical theory:** probability concentration theorem for log-convex distributions.
3. **Constrained AI model:** convex AI model with explicit potential.

Exascale Algorithms

Exascale Algorithms

Blocking for Distribution

Block distributed primal dual algorithm



Block distributed primal dual algorithm with AI prior



Exascale Algorithms

Uncertainty Quantification

Convex probability concentration for uncertainty quantification

Posterior credible region:

$$p(\mathbf{x} \in C_\alpha | \mathbf{y}) = \int_{\mathbf{x} \in \mathbb{R}^N} p(\mathbf{x} | \mathbf{y}) \mathbb{1}_{C_\alpha} d\mathbf{x} = 1 - \alpha.$$

Consider the **highest posterior density (HPD)** region

$$C_\alpha^* = \{ \mathbf{x} : -\log p(\mathbf{x}) \leq \gamma_\alpha \}, \quad \text{with } \gamma_\alpha \in \mathbb{R}, \quad \text{and } p(\mathbf{x} \in C_\alpha^* | \mathbf{y}) = 1 - \alpha \text{ holds.}$$

Theorem 3.1 (Pereyra 2017)

Suppose the posterior $\log p(\mathbf{x} | \mathbf{y}) \propto \log \mathcal{L}(\mathbf{x}) + \log \pi(\mathbf{x})$ is **log-concave** on \mathbb{R}^N . Then, for any $\alpha \in (4e^{[-N/3]}, 1)$, the HPD region C_α^* is contained by

$$\hat{C}_\alpha = \left\{ \mathbf{x} : \log \mathcal{L}(\mathbf{x}) + \log \pi(\mathbf{x}) \leq \hat{\gamma}_\alpha = \log \mathcal{L}(\hat{\mathbf{x}}_{\text{MAP}}) + \log \pi(\hat{\mathbf{x}}_{\text{MAP}}) + \sqrt{N}\tau_\alpha + N \right\},$$

with a positive constant $\tau_\alpha = \sqrt{16 \log(3/\alpha)}$ independent of $p(\mathbf{x} | \mathbf{y})$.

Need only evaluate $\log \mathcal{L} + \log \pi$ for the MAP estimate \mathbf{x}_{MAP} !

Exascale Algorithms

AI Data-Driven Prior

Convex AI prior

Adopt **neural-network-based convex regulariser** R

(Goujon *et al.* 2022; Liaudat *et al.* McEwen 2024):

$$R(\mathbf{x}) = \sum_{n=1}^{N_C} \sum_k \psi_n((\mathbf{h}_n * \mathbf{x})[k]),$$

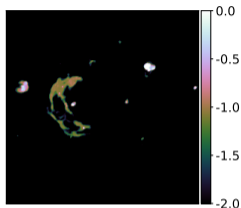
- ▷ ψ_n are learned convex profile functions with Lipschitz continuous derivative;
- ▷ N_C learned convolutional filters \mathbf{h}_n .

Properties:

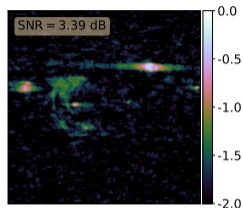
1. **Convex + explicit** \Rightarrow leverage convex UQ theory.
2. **Smooth regulariser with known Lipschitz constant** \Rightarrow theoretical convergence guarantees.

Demonstrations

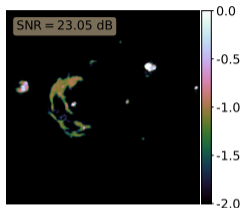
Reconstructed images



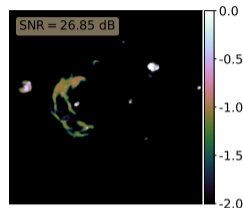
Ground truth



Dirty image
SNR=3.39 dB

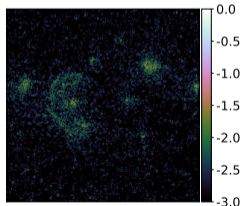


Reconstruction (classical)
SNR=23.05 dB

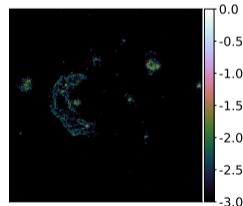


Reconstruction (learned)
SNR= 26.85 dB

(Liaudat *et al.* McEwen 2024)



Error (classical)

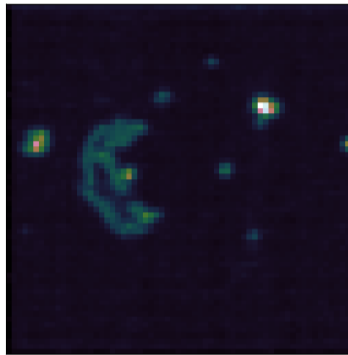


Error (learned)

Approximate local Bayesian credible intervals



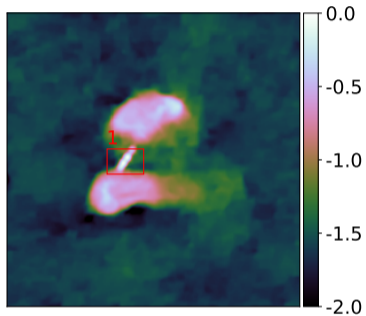
LCI
(super-pixel size 4×4)



MCMC standard deviation
(super-pixel size 4×4)

(Liaudat *et al.* McEwen 2024)

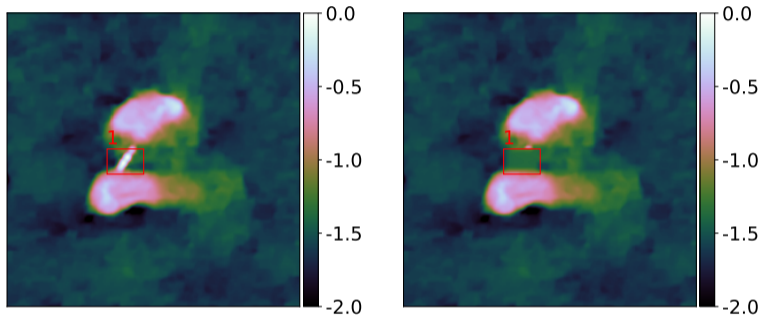
Hypothesis testing of structure



Reconstructed image

(Liaudat *et al.* McEwen 2024)

Hypothesis testing of structure

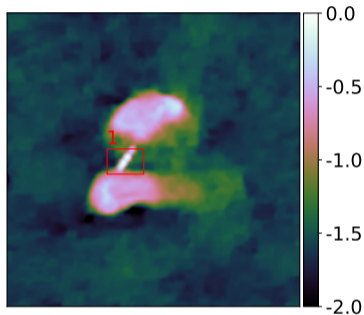


Reconstructed image

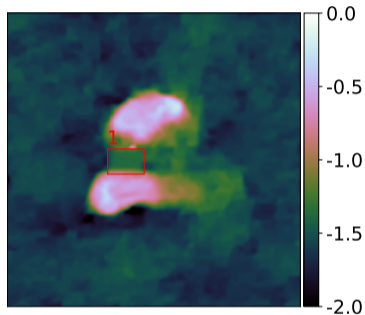
Surrogate test image (region removed)

(Liaudat *et al.* McEwen 2024)

Hypothesis testing of structure



Reconstructed image

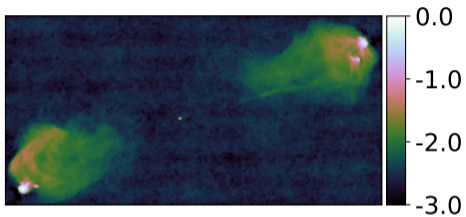


Surrogate test image (region removed)

Reject null hypothesis
⇒ structure physical

(Liaudat *et al.* McEwen 2024)

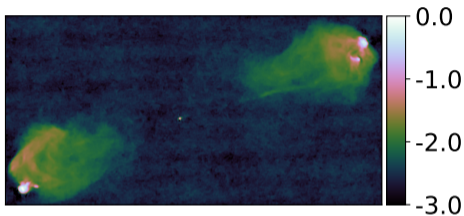
Hypothesis testing of substructure



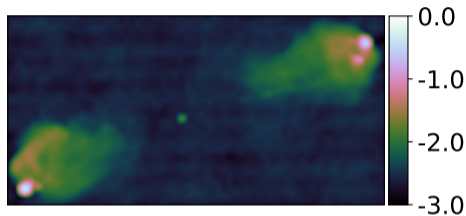
Reconstructed image

(Liaudat *et al.* McEwen 2024)

Hypothesis testing of substructure



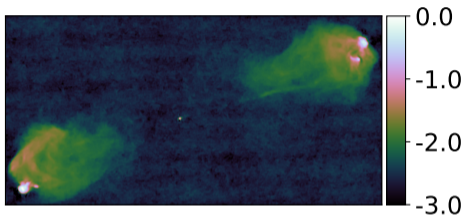
Reconstructed image



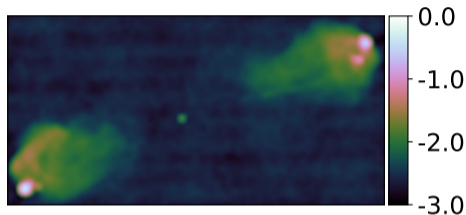
Surrogate test image (blurred)

(Liaudat *et al.* McEwen 2024)

Hypothesis testing of substructure



Reconstructed image

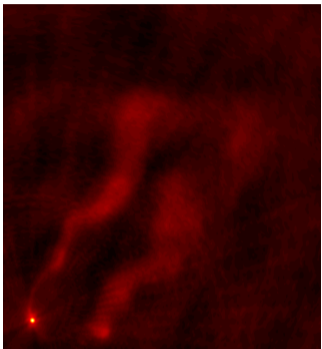


Surrogate test image (blurred)

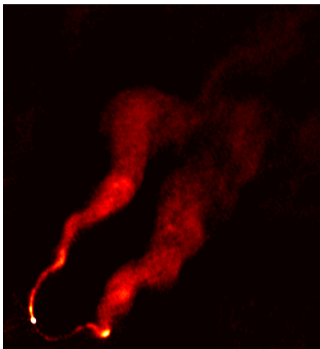
Reject null hypothesis \Rightarrow **substructure physical**

(Liaudat *et al.* McEwen 2024)

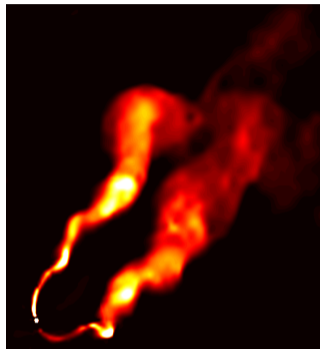
Imaging 3C128 with VLA



Dirty image



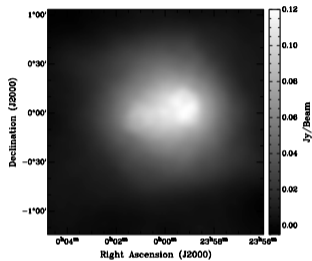
CLEAN (Traditional approach)



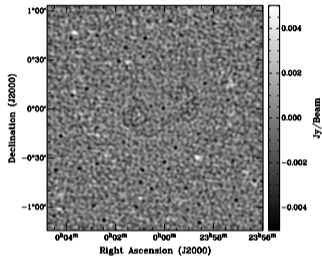
PURIFY (Ours)

(Pratley, McEwen *et al.* 2018)

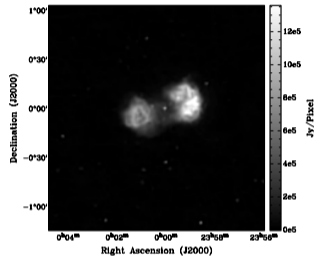
Imaging Fornax A with MWA



Dirty image



Residuals



Reconstruction

(Pratley, Johnston-Hollitt & McEwen 2020)

Code

PURIFY code

<https://github.com/astro-informatics/purify>

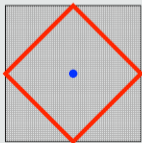


Next-generation radio interferometric imaging

PURIFY is a highly distributed and parallelized open-source C++ code for radio interferometric imaging, leveraging recent developments in the field of variational regularization, convex optimisation, and learned imaging.

SOPT code

<https://github.com/astro-informatics/sopt>



Sparse OPTimisation

SOPT is a highly distributed and parallelized open-source C++ code for variational regularization and convex optimisation, with learned data-driven priors.

Computational strategy

- ▷ Big data and big compute BUT small AI models (big sims to generate training data)
- ▷ Training and prototyping in **Python** on current-generation hardware
- ▷ Imaging (production) in **C++** on **exascale** hardware
- ▷ **Spack** package manager
- ▷ **Benchmarking**
 - ▷ Integrated in ExCALIBUR *Benchmarking for Performance Portable ExCALIBUR Applications* (see talk by Tuomas Koskela)
 - ▷ Tested on NVIDIA Grace Hopper on UCL Contender
 - ▷ Tested on Intel with OmniPath network on UCL Kathleen
 - ▷ Isambard 3 Technical Preparatory Access



ONNX



Spack

- ▷ **Learned exascale computational inverse imaging (LEXCI)** framework for the SKA and beyond
 1. **Highly distributed and parallelised**
 2. **Highly realistic telescope modelling**
 3. **Superior reconstruction quality** by using learned AI data-driven priors
 4. **Uncertainty quantification for exascale imaging** with learned priors for the first time
 5. **Validated** by MCMC sampling (for low-dimensional setting)
- ▷ **Benchmarking** underway...