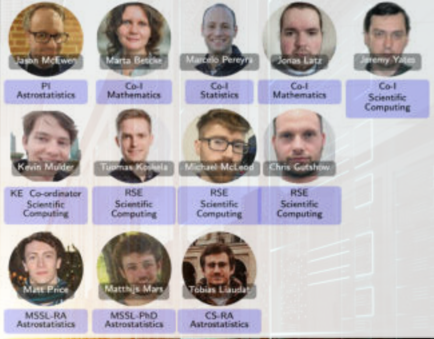
Learned Exascale Computational Imaging (LEXCI)

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

SKA Exascale

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

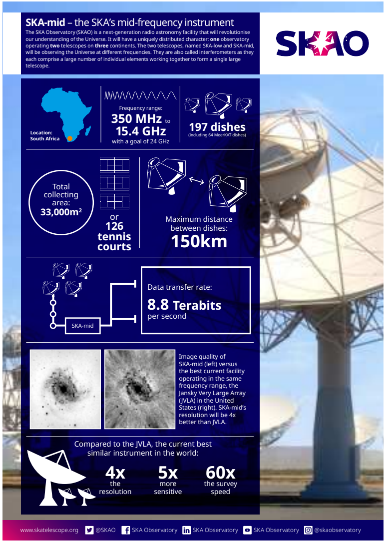
SKA science goals

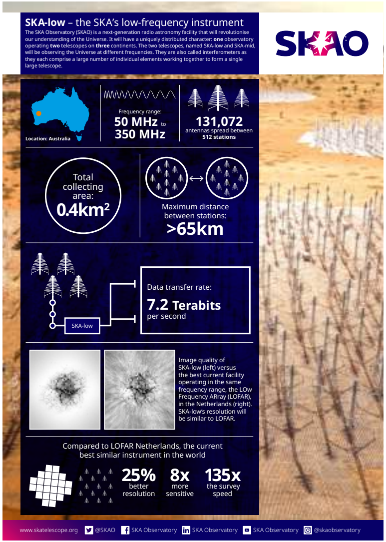
Orders of magnitude improvement in sensitivity and resolution.

Unlock broad range of science goals.

SKA partners

SKA sites and data rates



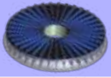


Application domains more broadly

Learned Exascale Computational Imaging (LEXCI)

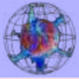
Radio Interferometry
e.g. MWA, SKA

Optical Interferometry
e.g. SPIDER



Medical Imaging
e.g. Diffusion MRI

Seismic Imaging
e.g. Mantle Plumes



Computer Graphics
e.g. Inverse Rendering

LEXCI project

Astronomical Imaging | **Planetary & Climate Science** | **Molecular Biology** | **Optical Photography** | **Geophysics**

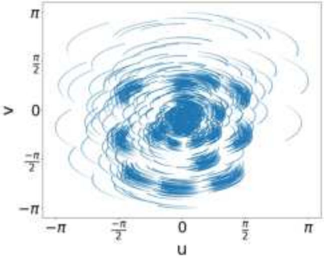
Acoustics | **Computer Vision** | **Many Others...**

Extended application areas

Imaging Strategy

Radio interferometric telescopes acquire “Fourier” measurements

“Fourier”
Measurements
$\Rightarrow$



Interferometric imaging is an exascale computational inverse imaging problem.

Radio interferometric inverse problem

Radio interferometric imaging **ill-posed inverse problem**:

$$\boxed{y = \Phi(x) + n}$$

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

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

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**.
+ Uncertatinties encoded in posterior.
- **Hand-crafted priors** (traditionally).

Computational imaging strategy

Goals:

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

Computational imaging strategy

Goals:

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

Achieve by combining:

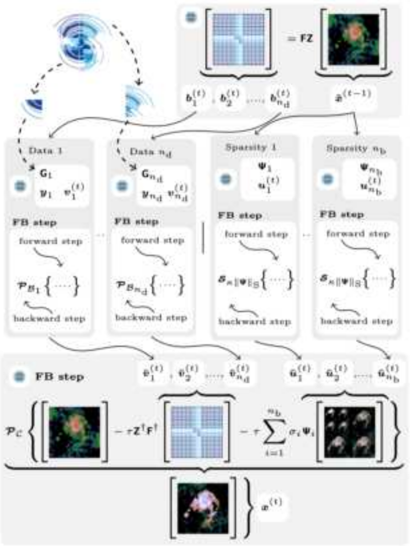
1. **Statistical framework:** Bayesian inference and MAP estimation.
2. **Mathematical theory:** probability concentration theorem for log-convex distributions.
3. **Constrained AI model:** convex AI model with explicit potential.

Exascale Algorithms

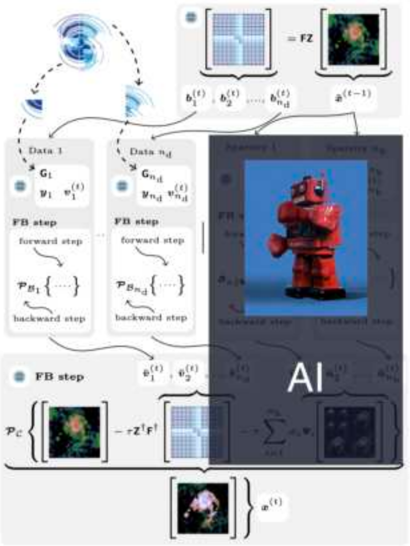
Exascale Algorithms

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:

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

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

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

Theorem 3.1 (Pereyra 2017)

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

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

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

Need **only evaluate** $\log \mathcal{L} + \log \pi$ **for the MAP estimate** 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):

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

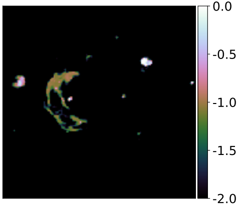
- ψ_n are learned convex profile functions with Lipschitz continuous derivative;
- N_C learned convolutional filters $\boldsymbol{h}_n$.

Properties:

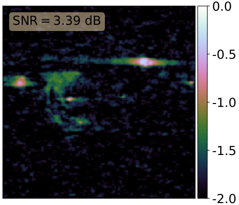
1. **Convex** + **explicit** $\Rightarrow$ leverage convex UQ theory.
2. **Smooth regulariser with known Lipschitz constant** $\Rightarrow$ theoretical convergence guarantees.

Demonstrations

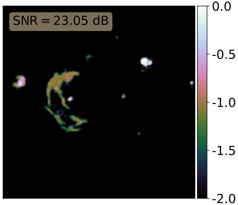
Reconstructed images



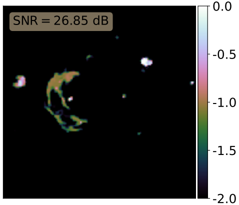
Ground truth



Dirty image
SNR=3.39 dB

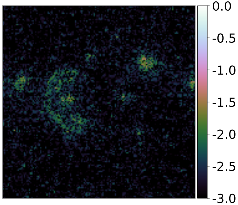


Reconstruction (classical)
SNR=23.05 dB

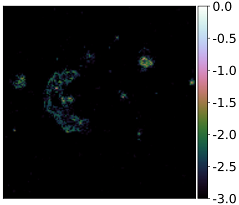


Reconstruction (learned)
SNR= 26.85 dB

(Liaudat *et al.* McEwen 2024)

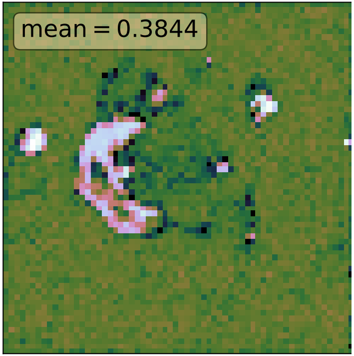


Error (classical)

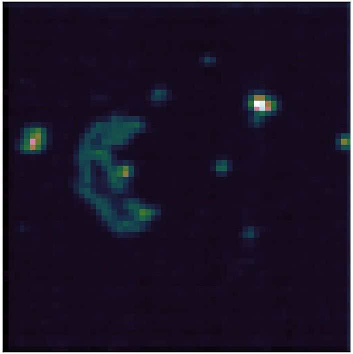


Error (learned)

Approximate local Bayesian credible intervals



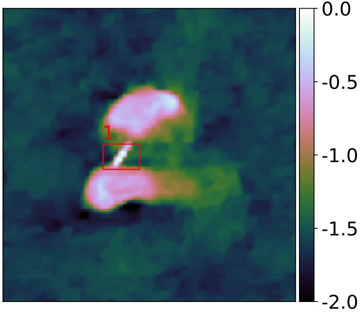
LCI
(super-pixel size 4×4)



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

(Liaudat *et al.* McEwen 2024)

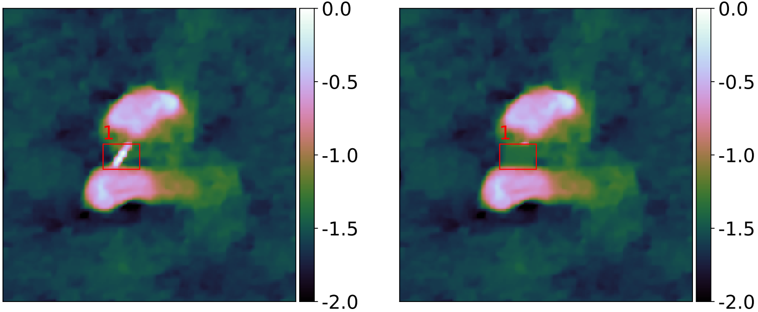
Hypothesis testing of structure



Reconstructed image

(Liaudat *et al.* McEwen 2024)

Hypothesis testing of structure

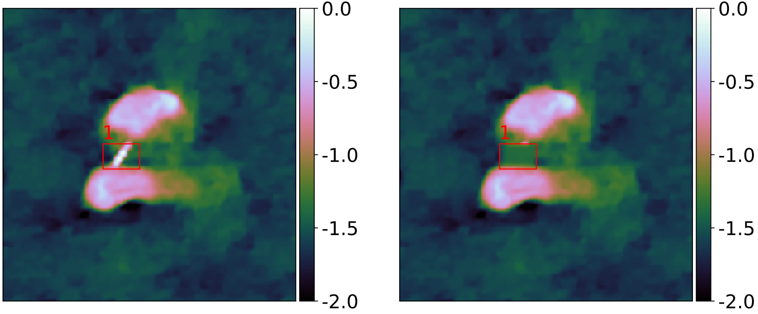


Reconstructed image

Surrogate test image (region removed)

(Liaudat *et al.* McEwen 2024)

Hypothesis testing of structure



Reconstructed image

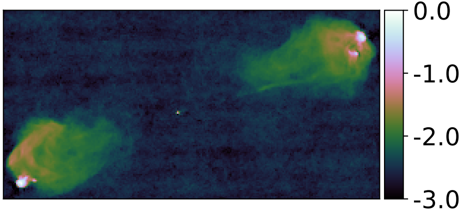
Surrogate test image (region removed)

Reject null hypothesis

⇒ **structure physical**

(Liaudat *et al.* McEwen 2024)

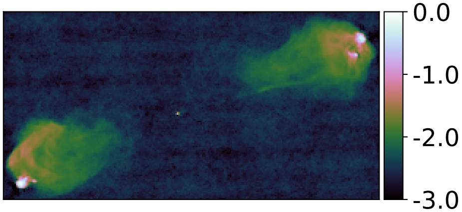
Hypothesis testing of substructure



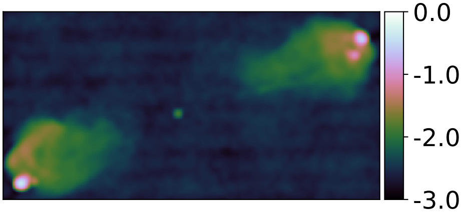
Reconstructed image

(Liaudat *et al.* McEwen 2024)

Hypothesis testing of substructure



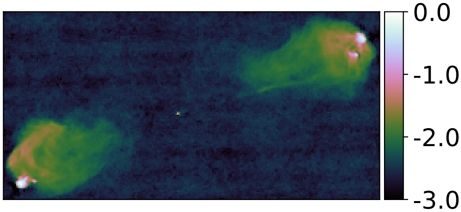
Reconstructed image



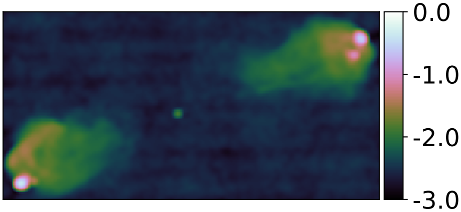
Surrogate test image (blurred)

(Liaudat *et al.* McEwen 2024)

Hypothesis testing of substructure



Reconstructed image

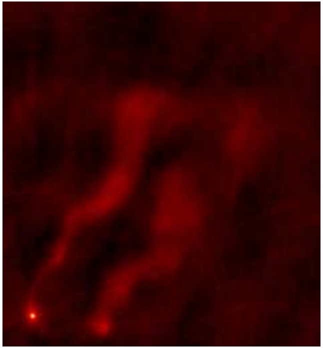


Surrogate test image (blurred)

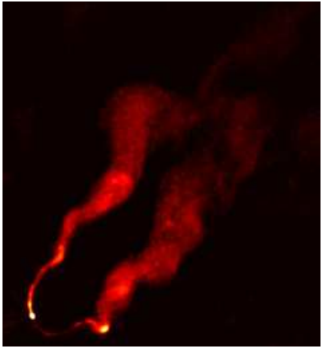
Reject null hypothesis $\Rightarrow$ **substructure physical**

(Liaudat *et al.* McEwen 2024)

Imaging 3C128 with VLA



Dirty image

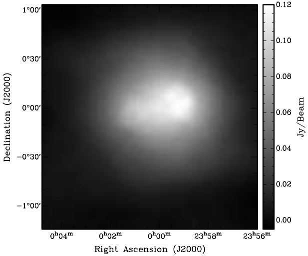


CLEAN (Traditional approach)

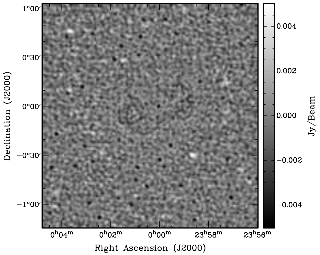
PURIFY (Ours)

(Pratley, McEwen *et al.* 2018)

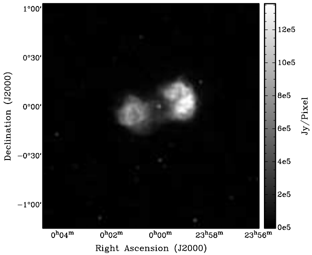
Imaging Fornax A with MWA



Dirty image



Residuals



Reconstruction

(Pratley, Johnston-Hollitt & McEwen 2020)

Code

Open-source codes

PURIFY code

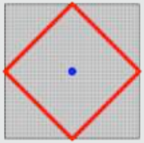
https://github.com/astro-informatics/purify

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

Summary

- 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...