

The AIRI plug-and-play algorithm for image reconstruction in radio-interferometry: variations and robustness

Matthieu Terris,^{1†} Chao Tang,^{1,2†} Adrian Jackson² and Yves Wiaux^{1★†}

¹*Institute of Sensors, Signals, and Systems, Heriot-Watt University, Currie, Edinburgh EH14 4AS, UK*

²*EPCC, University of Edinburgh, Potterrow, Edinburgh EH8 9BT, UK*

Accepted 2024 December 12. Received 2024 December 12; in original form 2024 August 22

ABSTRACT

Plug-and-play (PnP) algorithms are appealing alternatives to proximal algorithms when solving inverse imaging problems. By learning a Deep Neural Network (DNN) denoiser behaving as a proximal operator, one waives the computational complexity of optimization algorithms induced by sophisticated image priors, and the sub-optimality of handcrafted priors compared to DNNs. Such features are highly desirable in radio-interferometric (RI) imaging, where precision and scalability of the image reconstruction process are key. In previous work, we introduced AIRI (‘AI Regularization in radio-interferometric Imaging’), PnP counterpart to the unconstrained variant of the SARA (Sparsity Averaging Reweighted Analysis) optimization algorithm, relying on a forward–backward algorithmic backbone. Here, we introduce variations of AIRI towards a more general and robust PnP paradigm in RI imaging. First, we show that the AIRI denoisers can be used without any alteration to instantiate a PnP counterpart to the constrained SARA optimization algorithm itself, relying on a primal–dual forward–backward algorithmic backbone, thus extending the remit of the AIRI paradigm. Secondly, we show that AIRI algorithms are robust to strong variations in the nature of the training data set, with denoisers trained on medical images yielding similar reconstruction quality to those trained on astronomical images. Thirdly, we develop a functionality to quantify the model uncertainty introduced by the randomness in the training process. We validate the image reconstruction and uncertainty quantification functionality of AIRI algorithms against the SARA family and CLEAN, both in simulation and on real data of the ESO 137-006 galaxy acquired with the MeerKAT telescope. AIRI code is available in the BASPLib code library[†] on GitHub.

Key words: techniques: image processing – techniques: interferometric.

1 INTRODUCTION

Synthesis imaging by radio interferometry (RI) in astronomy leverages antenna arrays to observe the sky with high resolution and sensitivity (Thompson, Moran & Swenson 2017). A new generation of RI telescopes is currently emerging, whose science goals range from studying cosmic magnetism, dark matter, and dark energy, to understanding the structure and evolution of stars and galaxies. Targeting ever increasing resolution and dynamic range to resolve highly complex structure distributions, new arrays feature ever larger antenna counts. They are consequently characterized by unprecedented volumes of observed data, and image sizes. The flagship Square Kilometre Array (SKA) will soon come online (Scaife 2020), the South African MeerKAT array, and the Australian Pathfinder ASKAP two major precursors.

In this new era, the Fourier inverse problem for image formation from visibility data is extremely challenging, with RI imaging algorithms required to deliver a new regime of joint precision and

scalability. The CLEAN algorithm holds a near-total monopoly in the field (Högbom 1974; Offringa et al. 2014; Thompson et al. 2017), owing its success to its simplicity and associated scalability. CLEAN is a highly specialized precursor to the ‘Matching Pursuit’ algorithm (Mallat & Zhang 1993), iteratively identifying image components from data residuals back-projected in the pixel domain (Högbom 1974; Schwab 1984), or some other multiscale representation (Cornwell 2008). It can loosely be categorized among the variational inference techniques (Marsh & Richardson 1987). Yet, it requires manual fine tuning, limit resolution, and dynamic range by design to compensate for a simplistic regularization approach, lacks versatility to handle complex emission and complex calibration effects, and a proper uncertainty quantification functionality.

Building from early work framed in the context of the theory of compressed sensing (Wiaux et al. 2009), advanced optimization, sampling, and deep-learning approaches have been developed (Kazemi et al. 2011; Dabbech et al. 2015; Garsden et al. 2015; Junklewitz et al. 2016; Repetti et al. 2017; Cai, Pereyra & McEwen 2018; Repetti, Pereyra & Wiaux 2019; Arras et al. 2021; Connor et al. 2022; Dabbech et al. 2022; Gheller & Vazza 2022; Aghabiglou et al. 2024; Dia et al. 2023, ; Liaudat et al. 2024; Roth et al. 2023; Terris et al. 2023; Wilber et al. 2023a, b; Dabbech et al. 2024). Few of those techniques have made their way through from

* E-mail: yves.wiaux@hw.ac.uk

† Equal contribution.

‡ Code available at: <https://basp-group.github.io/BASPLib/>.

extensive methodological study to validation on challenging real data applications. First, the variational Bayes approach underpinning the RESOLVE algorithm has shown potential to deliver superior precision and robustness to CLEAN (Arras et al. 2021; Roth et al. 2023), but at the expense of a high computational cost affecting efficiency and scalability. Secondly, modern variational inference techniques using sparsity-based regularization have been shown to deliver remarkable precision, propelled by advanced proximal optimization algorithms. These range from the original SARA, which relies on a primal-dual forward-backward (PDFB) algorithmic backbone to solve an optimization problem with a constrained data-fidelity term (Carrillo, McEwen & Wiaux 2012; Onose et al. 2016; Thouvenin et al. 2023), to uSARA, which relies on a forward-backward (FB) algorithmic backbone to solve an optimization problem with an unconstrained data-fidelity term (Repetti & Wiaux 2020; Dabbech et al. 2022; Wilber et al. 2023a). Thirdly, a ‘plug-and-play’ version of uSARA was introduced, dubbed AIRI, which consists in substituting the sparsity-promoting regularization operator with a learned Deep Neural Network (DNN) denoiser. AIRI was demonstrated to further enhance imaging precision over uSARA. It also brings a significant acceleration owing to the fast inference of its DNN denoiser (Terris et al. 2023; Wilber et al. 2023b).

In this work, we introduce variations of AIRI towards a more general and robust PnP paradigm for RI imaging, showing that (i) AIRI denoisers can be plugged without any alteration in other algorithmic backbones than FB, in particular in PDFB to instantiate a PnP counterpart to SARA itself; (ii) AIRI algorithms are robust to strong variations in the nature of the training data set; (iii) AIRI algorithms can be endowed with a functionality to quantify the uncertainty inherent to its learned regularization model.

The remainder of this paper is organized as follows. In Section 2, we detail the proposed approach unifying constrained and unconstrained approaches for PnP and optimization approaches. In Section 3, we validate the image reconstruction of AIRI algorithms against state-of-the-art algorithms on simulated measurements. In Section 4, we showcase the performance of the proposed algorithm on real measurements of the ESO 137–006 galaxy acquired with the MeerKAT telescope (Jonas 2016).

2 PROPOSED METHODOLOGY

In this section, we recall the RI inverse problem and introduce our constrained PnP algorithm, dubbed cAIRI. We next detail our methodology for curating image data sets with high dynamic range.

2.1 The RI inverse problem

Focusing on monochromatic intensity imaging on a small field of view (FoV), and in the absence of atmospheric and instrumental perturbations, each pair of antennas of an RI array acquires a noisy Fourier component of the intensity image to be formed, called a visibility. The associated Fourier mode (also called uv -point) is given by the projection of the corresponding baseline, expressed in units of the observation wavelength, on to the plane perpendicular to the line of sight (Thompson et al. 2017). Gathering measurements from all antenna pairs throughout the duration of acquisition, an RI array thus samples an incomplete coverage of the spatial Fourier domain of the image of interest. A discrete formulation of the underpinning inverse problem, aiming to restore the target image $\bar{\mathbf{x}} \in \mathbb{R}^n$ from the measured complex visibilities $\mathbf{y} \in \mathbb{C}^m$, reads

$$\mathbf{y} = \Phi \bar{\mathbf{x}} + \mathbf{e}, \quad (1)$$

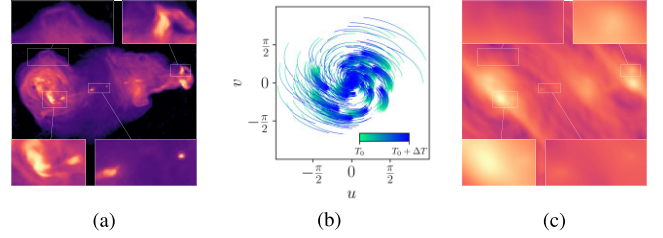


Figure 1. Illustration of a simulated RI imaging problem. Panel (a) shows a groundtruth radio galaxy, displayed in logarithmic scale. Panel (b) shows the sampled points in the spatial Fourier domain. At each time $T_0 \leq t \leq T_0 + \Delta T$, each pair of antennas samples a single point, during a total acquisition time ΔT . As the acquisition time increases, the locus of the points moves in the Fourier domain due to the Earth’s rotation. The colourbar illustrates this effect. Panel (c) shows the back-projected measurements in the image domain.

Algorithm 1 AIRI algorithm

```

1: given  $0 < \gamma < 2/L$ , denoiser  $D$ ,  $\mathbf{x}_0 \in \mathbb{R}^n$ ,  $\xi > 0$ 
2: repeat for  $k = 0, 1, \dots$ 
3:    $\mathbf{x}_{k+1} = D(\mathbf{x}_k - \gamma(\text{Re}\{\Phi^\dagger \Phi \mathbf{x}_k\} - \text{Re}\{\Phi^\dagger \mathbf{y}\}))$ 
4: until  $\|\mathbf{x}_{k+1} - \mathbf{x}_k\| / \|\mathbf{x}_{k+1}\| < \xi$ 
5: return  $\mathbf{x}_{k+1}$ 

```

where $\Phi = \mathbf{G}\mathbf{F}\mathbf{Z} \in \mathbb{C}^{m \times n}$ is the measurement operator, $\mathbf{G} \in \mathbb{C}^{m \times d}$ is a sparse interpolation matrix, encoding the non-uniform Fourier transform, $\mathbf{F} \in \mathbb{C}^{d \times d}$ is the 2D Discrete Fourier Transform, $\mathbf{Z} \in \mathbb{R}^{d \times n}$ is a zero-padding operator, incorporating the correction for the convolution performed through the operator \mathbf{G} , and $\mathbf{e} \in \mathbb{C}^m$ is a realization of some i.i.d. Gaussian random noise, with mean 0 and standard deviation $\eta > 0$. Backprojecting problem (1) in the image domain gives the alternative formulation

$$\text{Re}\{\Phi^\dagger \mathbf{y}\} = \text{Re}\{\Phi^\dagger \Phi\} \bar{\mathbf{x}} + \text{Re}\{\Phi^\dagger \mathbf{e}\}, \quad (2)$$

where $(\cdot)^\dagger$ denotes the complex conjugate transpose and $\text{Re}\{\cdot\}$ the real part. $\text{Re}\{\Phi^\dagger \mathbf{y}\}$ is known as the dirty image, $\text{Re}\{\Phi^\dagger \Phi\}$ is the operator representing the convolution of $\bar{\mathbf{x}}$ by the point spread function, also known as the dirty beam, and $\text{Re}\{\Phi^\dagger \mathbf{e}\}$ is the noise backprojected in the image domain. An illustration of these concepts is provided in Fig. 1.

2.2 Unconstrained and constrained approaches

A common approach for solving the ill-posed inverse problem (1) consists in reformulating it as a minimization problem yielding an estimate $\hat{\mathbf{x}}$ of $\bar{\mathbf{x}}$ as

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} f(\mathbf{x}) + \lambda r(\mathbf{x}), \quad (3)$$

where $\lambda > 0$ is a regularization parameter. In turn, appropriate choices for f , r , and λ ensure the quality of the reconstructed image $\hat{\mathbf{x}}$. Given the Gaussian nature of the noise, there are primarily two possible strategies for the choice of f .

A first possible choice consists in the least squared error, *i.e.*

$$f(\mathbf{x}) = \frac{1}{2} \|\Phi \mathbf{x} - \mathbf{y}\|^2, \quad (4)$$

and in this case, algorithms for solving (3) typically take the form of Algorithm 1, where $L = \|\Phi\|_F^2$ is the spectral norm of the measurement operator, γ is the step size and $D = \text{prox}_{\gamma\lambda r}$ is

Algorithm 2 cAIRI algorithm

```

1: given  $0 < \gamma < 1/L$ , denoiser  $D$ ,  $\mathbf{x}_0 \in \mathbb{R}^n$ ,  $\xi > 0$ 
2: repeat~for  $k = 0, 1, \dots$ 
3:    $\mathbf{x}_{k+1} = D(\mathbf{x}_k - \gamma \text{Re}\{\Phi^\dagger \mathbf{u}_k\})$ 
4:    $\mathbf{v}_k = \mathbf{u}_k + \Phi(2\mathbf{x}_{k+1} - \mathbf{x}_k)$ 
5:    $\mathbf{u}_{k+1} = \mathbf{v}_k - \text{prox}_{\iota_{\mathcal{B}(\mathbf{y}, \varepsilon)}}(\mathbf{v}_k)$ 
6: until  $\|\mathbf{x}_{k+1} - \mathbf{x}_k\| / \|\mathbf{x}_{k+1}\| < \xi$ 
7: return  $\mathbf{x}_{k+1}$ 

```

the proximity operator¹ of $\gamma\lambda r$. In particular, the uSARA imaging algorithm (Repetti & Wiaux 2020; Terris et al. 2023) writes as Algorithm 1 for a handcrafted sparsity-inducing prior r . A second strategy consists in adopting a constrained approach by minimizing $r(\mathbf{x})$ subject to $\|\Phi\mathbf{x} - \mathbf{y}\| \leq \varepsilon$, which can be reformulated as (3) with

$$f(\mathbf{x}) = \iota_{\mathcal{B}(\mathbf{y}, \varepsilon)}(\Phi\mathbf{x}), \quad (5)$$

where $\iota_{\mathcal{B}(\mathbf{y}, \varepsilon)}(\cdot)$ denotes the indicator² function of $\mathcal{B}(\mathbf{y}, \varepsilon)$, the ℓ_2 ball of radius ε and centred in \mathbf{y} . The value of ε can be estimated with the bound on a χ^2 distribution with $2m$ degrees of freedom, i.e. $\varepsilon^2 = (2m + 4\sqrt{m})\eta^2$ (Carrillo et al. 2012). This problem leads to different algorithms, a classical instance being given in Algorithm 2, where $D = \text{prox}_{\gamma r}$.

In particular, the SARA algorithm (Onose et al. 2016) writes as Algorithm 2 when $D = \text{prox}_{\gamma r}$ for a well-chosen sparsity-inducing prior function r .

Although Algorithms 1 and 2 exhibit distinct algorithmic backbones, these two algorithms can be related within the unifying minimization framework (3) up to the choice of f (Combettes & Pesquet 2011; Pesquet et al. 2021). In Terris et al. (2023), we proposed to adopt a PnP approach and replaced D in Algorithm 1 with a denoising DNN and dubbed the resulting algorithm AIRI. In this work, we propose to extend this approach to the constrained framework and introduce constrained AIRI (cAIRI) Algorithm 2, the constrained variant of AIRI. Importantly, these two algorithms, as well as uSARA and SARA, fall in the same general framework propelled by the minimization framework (3). Furthermore, AIRI and cAIRI can be seen as generalizations of the state-of-the-art uSARA and SARA algorithms, but relying on denoising DNNs D with higher expressive power than sparsity inducing priors underpinning uSARA and SARA.

2.3 Generating synthetic low-dynamic range data sets

We define the dynamic range of an image as the ratio of pixel intensities between brightest and faintest features. For natural images, this ratio tends to be low. In the absence of groundtruth RI image data sets, Terris et al. (2023) proposed to generate a low-dynamic range data set from Optical Astronomical Imaging (OAI) images to serve as a basis for training PnP denoisers (as well as end-to-end DNNs for benchmarking) to solve the inverse problem (1).

In order to investigate the influence of the nature of the images used in the training of the denoiser, we also propose to generate a low-dynamic range data set from magnetic resonance imaging (MRI) images. The second data set we examine contains MRI images derived from the fastMRI single-coil knee images (Zbontar et al.

¹The proximity operator of a convex function r is defined as $\text{prox}_r(\mathbf{y}) = \arg\min_{\mathbf{x}} r(\mathbf{x}) + \frac{1}{2}\|\mathbf{x} - \mathbf{y}\|^2$.

²The indicator ι_C of a closed convex set C is the function defined as $\iota_C(\mathbf{x}) = 0$ if $\mathbf{x} \in C$ and $\iota_C(\mathbf{x}) = +\infty$ otherwise.

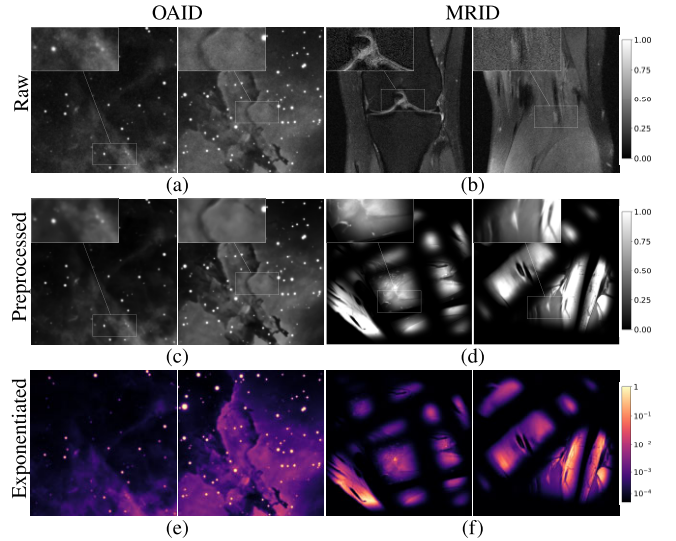


Figure 2. The different pre-processing steps of our OAID (left) and MRI data set (MRID; right). Panels (a) and (b) show raw images; (c) and (d) show images pre-processed with a SCUNet and the soft-thresholding removing the background; (e) and (f) show images after exponentiation. For images in MRID, random translations and rotations are also applied to increase image size and to break symmetries as discussed in Section 2.3.

2018). To break symmetries and enlarge the image size, random patches from each MRI slices are extracted, rotated, translated, and concatenated. An edge tapering filter is then applied to eliminate image discontinuities, resulting in new MRI images with varied content.

Both the OAI and MRI-born images were pre-processed with a blind image restoration network (SCUNet; Zhang et al. 2023) to remove noise and artefacts. Pixel-wise soft-thresholding is then applied to remove residual background. This processing strategy yields 5000 low-dynamic range OAI-born and MRI-born groundtruth images of size 512×512 , respectively.

We show in Figs 2(a)–(d) the raw and pre-processed images from both optical astronomy image data set (OAID) and MRI data set (MRID). While the semantic content of images strongly differs between the two data sets, both are normalized to a maximum intensity of 1. However, the residual noise levels differ between the two. As a result, the minimum non-zero intensity, denoted σ_0 , also varies between the data sets, leading to different dynamic ranges. For each data set, σ_0 is chosen such that the intensities of 1 per cent non-zero pixels in the corresponding data set are below this value, which are around 0.02 and 0.01 for OAID and MRID, respectively.

2.4 Training a shelf of high dynamic range denoisers

A distinctive characteristic of radio images is their high dynamic range, with important features exhibiting very faint intensity (for instance radio plumes or radio relics, see Missaglia et al. 2019; Wittor et al. 2021). Observation with modern instrument will typically target values up to 10^6 or even higher. Considering without loss of generality images with peak intensity normalized to 1, this implies faint intensities down to as low as 10^{-6} . In order to convert the available low-dynamic range images to high dynamic ranges, Terris et al. (2023) proposed a simple exponentiation procedure. Given a low-dynamic range image \mathbf{u} , we apply a pixel-wise exponentiation

transform, denoted rexp_a :

$$\text{rexp}_a : \mathbf{u} \mapsto \mathbf{u}_{\max}(a^{u/\mathbf{u}_{\max}} - 1)/a. \quad (6)$$

Indeed, it can be shown that this transform increases the dynamic range of the image for sufficiently large values of a . In this work, we assume a maximum pixel intensity of 1, i.e. $\mathbf{u}_{\max} = 1$.

Next, we link the exponentiation factor a to the target dynamic range $1/\sigma$ by following the approach of Terris et al. (2023), where the faintest intensity in the exponentiated image, $(a^{\sigma_0} - 1)/a$, is equated to the inverse of the dynamic range σ . This yields the following equation to be solved for a , given the faintest intensity σ_0 in the clean, low-dynamic range training data set, and the inverse of the target dynamic range σ :

$$(a\sigma + 1)^{1/\sigma_0} - a = 0. \quad (7)$$

We solve equation (7) numerically for the exponentiation parameter a at different training noise levels of interest to our experiments.

Applying this transform to images in OAID and MRID online during training yields synthetic high dynamic range radio-astronomical training images, as shown in Figs 2(e)–(f).

Given these high dynamic range data sets, the common approach for training the denoiser D consists in minimizing a denoising loss with Lipschitz regularization (Terris et al. 2023). Denoting θ the learnable parameters of D , for a synthetic high dynamic range sample \mathbf{u} , the training loss writes

$$\mathcal{L}(\theta, \mathbf{u}) = \underbrace{\|D_\theta(\mathbf{u} + \sigma \mathbf{w}) - \mathbf{u}\|_1}_{\text{denoising}} + \underbrace{\kappa \max\{\|\nabla Q_\theta(\tilde{\mathbf{u}})\|_S, 1 - \delta\}}_{\text{Lipschitz regularization}}, \quad (8)$$

where \mathbf{w} is the realization of Gaussian random noise with standard deviation 1, $\sigma > 0$ denotes the training noise level, Q is defined as $Q = 2D - \text{Id}$, $\|\nabla Q(\cdot)\|_S$ denotes the spectral norm of the Jacobian of Q , κ and δ are positive constants, and $\tilde{\mathbf{u}}$ is a point sampled at random on the segment $[\mathbf{u}, \mathbf{u} + \sigma \mathbf{w}]$.

Finally, to avoid the necessity of training denoisers for each measurement with different target dynamic ranges, we build a shelf of denoisers trained for different noise levels covering the full range of possible dynamic ranges of interest for a class of radio observations.

2.5 Selecting denoiser from shelf during reconstruction

During reconstruction, both the maximum image intensity and the effective noise level in the image domain need to be estimated from the data to provide an estimate of the target dynamic range, necessary for choosing the appropriate trained denoiser based on its training noise level.

First, the maximum intensity of the back-projected image, denoted as $\tilde{\alpha}$, provides a rough estimation of the sought peak value α of the target image $\bar{\mathbf{x}}$. Then, following Terris et al. (2023), the inverse of the signal-to-noise ratio of the observed data, i.e. the inverse of the target dynamic range, can be estimated from the properties of Φ and e in equation (1) as

$$\sigma_{\text{heu}} = \frac{\eta}{\tilde{\alpha} \sqrt{2\|\Phi\|_S^2}}, \quad (9)$$

where we recall that η is the standard deviation of e in equation (1).

Next, the network D_σ will be chosen where σ is the closest training noise level below σ_{heu} . The activation of the denoiser is thus rescaled with a factor $\beta = \tilde{\alpha}\sigma_{\text{heu}}/\sigma$ as

$$D(\cdot) = \beta D_\sigma(\cdot/\beta), \quad (10)$$

ensuring both that the input to the network has maximum intensity below 1 and that the effective standard deviation of the residual noise

in the input to the denoiser matches the noise level σ for which it was trained.

In practice, the peak value of the back-projected image is a loose upper bound estimation of the real peak value. To address the inaccuracy of the σ_{heu} estimate, we update $\tilde{\alpha}$ as the maximum intensity of the reconstructed image at each iteration of the PnP algorithm. The value of σ_{heu} is monitored at each iteration to check whether D_σ and the scaling factor β need to be updated. With the combination of the denoiser shelf and the selection strategy, we fully decouple the training and reconstruction.

2.6 Model uncertainty quantification

If obtaining an estimate of the solution to a specific model is desirable, it is also of interest to be able to quantify the uncertainty in the recovered solution. In practice, this uncertainty quantification is often performed by sampling from the posterior linked to r in equation (3), e.g. with MCMC methods (Laumont et al. 2022; Mukherjee et al. 2023), but other strategies have been proposed (Repetti et al. 2019; Zhang & Jin 2019; Liaudat et al. 2024). In a nutshell, these methods aim at quantifying the aleatoric uncertainty induced by the measurement procedure (1) in the reconstructed image.

Similarly to Narnhofer et al. (2021), we instead propose in this work to investigate the epistemic uncertainty in the reconstruction, induced by variations in the denoiser D . In fact, while the choice of r in equation (3) is a cornerstone of the reconstruction quality, its relationship with the denoiser D in the case of PnP algorithms remains unclear since no closed form for r is available in general,³ not to mention the high number of parameters and the (by nature) imperfect training procedure of D .

Training the same architecture D for the same task but in different settings (i.e. with different initialization, Gaussian noise realization or patch selection) leads to networks D with different weights, but consistent denoising performance. In turn, these different denoiser realizations yield distinct prior realizations. More precisely, assume that we train $K \in \mathbb{N}$ denoisers $(D_k)_{1 \leq k \leq K}$ with the same training procedure (loss and data sets) and that the sole difference between these denoisers arises from distinct random realizations influencing their training process. Intuitively, these denoisers yield unique regularizers $(r_k)_{1 \leq k \leq K}$, and plugging D_k within the PnP Algorithms 1 and 2 returns $\hat{\mathbf{x}}_k$ satisfying (3). This procedure gives us a set of solutions $\text{Sol} = \{\hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_K\}$ to each problem (1). In turn, the (pixel-wise) empirical standard deviation of Sol yields an estimator of the epistemic uncertainty (Lahlou et al. 2021).

We note that previous works have studied aleatoric and epistemic uncertainties in the context of end-to-end supervised learning with Bayesian neural networks (Kendall & Gal 2017; Kwon et al. 2018; Baumgartner et al. 2019; Xue et al. 2019; Edupuganti et al. 2020; Gao & Zhuang 2022).

3 EXPERIMENTAL RESULTS ON SIMULATED DATA

In this section, we show the performance of different algorithms for reconstructing RI images from simulated observations.

³Only a maximally monotone operator can be associated to D . These operators can be seen as a generalization of subdifferentials for convex functionals. We refer the reader to Bauschke, Combettes et al. (2011) and Pesquet et al. (2021) for more details.

3.1 Test data set and evaluation metrics

We use 3 radio images [3c353, Hercules A, and Centaurus A, see Terris et al. (2023) for more details] of size $N = 512 \times 512$ as groundtruth images to simulate testing measurements. Each image has maximum intensity value normalized to 1, and a dynamic range slightly above 10^4 . For each image, we generate RI measurements following (1) with various observation settings simulating the behaviours of the MeerKAT telescope (Jonas 2016). Namely, each simulated operator Φ arises from different pointing direction and observation time. More specifically, we consider five distinct pointing directions chosen at random, and for each direction, we vary the observation time with ΔT selected from the set {1 h, 2 h, 4 h, 8 h}. This leads to a total of 4×5 inverse problems for each groundtruth image, i.e. a total of 4×15 inverse problems with $\mathbf{y} \in \mathbb{C}^m$ such that $m = 4 \times 10^5$ for $\Delta T = 1$ h, $m = 8 \times 10^5$ for $\Delta T = 2$ h, $m = 1.6 \times 10^6$ for $\Delta T = 4$ h, $m = 3.2 \times 10^6$ for $\Delta T = 8$ h. Fig. 1(b) shows one of the test sampling patterns for a given pointing direction and observation time $\Delta T = 4$ h. The dynamic ranges of the observations fall in the range $[2.5 \times 10^3, 1.4 \times 10^4]$, with the average dynamic range for each ΔT as follows: 2.9×10^3 for $\Delta T = 1$ h, 3.9×10^3 for $\Delta T = 2$ h, 5.6×10^3 for $\Delta T = 4$ h, and 1.1×10^4 for $\Delta T = 8$ h.

To emphasize low-intensity features, radio images are usually displayed in logarithmic scale with the following explicit transform $\text{rlog}_a(\mathbf{x}) = \mathbf{x}_{\max} \log_a(a\mathbf{x}/\mathbf{x}_{\max} + 1)$, which maps an image \mathbf{x} with dynamic range $a \gg 1$ and pixel values in $[\mathbf{x}_{\max}/a; \mathbf{x}_{\max}]$ to a low-dynamic range image $\text{rlog}_a(\mathbf{x})$, with dynamic range $\log_a(2)$ and pixel values in the range $[\mathbf{x}_{\max} \log_a(2); \mathbf{x}_{\max}]$. Therefore, we evaluate the reconstruction accuracy of the high and low intensities with the SNR and logSNR metrics. Given a groundtruth image $\bar{\mathbf{x}}$ and an estimated image $\hat{\mathbf{x}}$, the reconstruction SNR is defined as $\text{SNR}(\bar{\mathbf{x}}, \hat{\mathbf{x}}) = 20 \log(\|\bar{\mathbf{x}}\|/\|\bar{\mathbf{x}} - \hat{\mathbf{x}}\|)$. Given the high DR set-up, in order to evaluate the performance of the algorithms at faint intensities, we consider the logSNR metric (Terris et al. 2023) writing $\log\text{SNR}(\bar{\mathbf{x}}, \hat{\mathbf{x}}) = \text{SNR}(\text{rlog}_a(\bar{\mathbf{x}}), \text{rlog}_a(\hat{\mathbf{x}}))$. The images in this article are presented in a logarithmic scale, i.e. for an image \mathbf{x} , visualizations show $\text{rlog}_a(\mathbf{x})$. The values of a in $\text{rlog}_a(\cdot)$ should ideally be set to the corresponding dynamic range of each measurement. For the logSNR metric, we set $a = 2.5 \times 10^3$, the lowest dynamic range of all the synthetic measurements. Choosing a single value for a preserves the consistency of the logSNR metric across different reconstructions and ΔT 's. For the visualization of images, a value of $a = 10^4$ is chosen, more representative of the dynamic range values for the $\Delta T = 4$ h, 8 h cases displayed. Because the groundtruth images are normalized, \mathbf{x}_{\max} in $\text{rlog}_a(\cdot)$ is set to 1 for consistency. Given K trained models, the model uncertainty is estimated using the empirical unbiased standard deviation (std) metric, defined as $\text{std}(\text{Sol}) = \sqrt{\sum_{k=1}^K (\hat{\mathbf{x}}_k - \text{mean}(\text{Sol}))^2} / \sqrt{K-1}$. Additionally, we calculate the relative standard deviation, defined as $\text{std}_{\text{rel}}(\text{Sol}) = \text{std}(\text{Sol}) / \text{mean}(\text{Sol})$, to compare the uncertainty for features with different intensities.

To assess the data-fidelity quality, we also provide residual dirty images, defined as $\text{Re}\{\Phi^\dagger(\mathbf{y} - \Phi\mathbf{x})\}$ (up to a normalization constant, see Terris et al. 2023).

3.2 AIRI variants and benchmark algorithms

In order to investigate the influence of the training data set on the denoising prior, we propose two versions for both AIRI and cAIRI underpinned by two sets of denoisers trained with different data sets. We denote by AIRI_{OAID} (resp. cAIRI_{OAID}) the AIRI (resp. cAIRI)

algorithm relying on denoisers trained on the OAID data set, while AIRI_{MRID} (resp. cAIRI_{MRID}) relies on denoisers trained on the MRID data set (see Fig. 2). Following (Terris et al. 2023), random rotations are applied before applying the denoising network, which improves stability of the algorithms (Terris et al. 2024). The algorithms are implemented with MATLAB.

We compare the performance of the proposed AIRI and cAIRI algorithms with several state-of-the-art RI imaging methods: (i) multi-scale CLEAN (MS-CLEAN) implemented in the C++ WS-Clean software package⁴ (Offringa & Smirnov 2017); (ii) the UNet from (Zhang et al. 2022) that was trained on the proposed synthetic data sets to solve the RI imaging problem in a supervised end-to-end fashion for each ΔT ; (iii) uSARA and SARA, the MATLAB implementations of which are available in the BASPLib code library. We further test our algorithms with the BM3D denoiser (Dabov et al. 2007) in Algorithms 1 and 2 instead of the proposed trained denoiser. BM3D is a non-trained denoiser and we tuned its denoising noise level to achieve the best possible image quality. However, it does not yield convergent PnP algorithms and suffers from slow inference time. The algorithm with a BM3D denoiser plugged into Algorithm 1 is named uPnP_{BM3D} and the one with BM3D plugged into Algorithm 2 is named cPnP_{BM3D}.

3.3 Training details

We use the DnCNN network architecture (Zhang et al. 2017) with the proposed modifications in Terris et al. (2023): all batch normalization layers are removed, ReLU layers are replaced by leaky ReLU layers, and an additional ReLU is added to ensure positivity of the reconstructed image.

We train a shelf of denoisers for eight noise levels covering wide possible dynamic ranges of interest, i.e. $\{2.5 \times 10^{-6}, 5.0 \times 10^{-6}, 1.0 \times 10^{-5}, 2.0 \times 10^{-5}, 4.0 \times 10^{-5}, 8.0 \times 10^{-5}, 1.6 \times 10^{-4}, 3.2 \times 10^{-4}\}$. This choice of noise levels ensures that $\sigma_{\text{heu}}/\sigma \in [1, 2]$ in equation (10). Each denoiser is trained on randomly cropped patches of size 46×46 of the OAID or MRID data sets, with the training loss (equation 8). As detailed in Terris et al. (2023), for each image of the synthetic data set, a set of randomly cropped patches \mathbf{x}_s is extracted and an associated high-dynamic range patch \mathbf{u}_s is computed with $\mathbf{u}_s = \text{rexp}_a(\mathbf{x}_s)$ where a is chosen as per equation (equation 7). The Lipschitz regularization parameter κ in equation (8) is fine-tuned for each denoiser and other training hyperparameters are the same. We refer to the denoisers trained with the non-expansiveness term for RI imaging as AIRI denoisers.

3.4 Results for image estimation

To illustrate clearly even the minor differences among different methods, we choose to show the reconstructions of one measurement with $\Delta T = 8$ h in Fig. 3, which contain more details compared to others recovered from measurements with smaller observation time. The reconstructions are organized in the ascending order of logSNR values and we only display the error maps and residual dirty images for the six best results. We first observe that MS-CLEAN successfully recovers structures with higher intensities, but not the faint and extended emissions. The central black hole is also less compact compared to the ones of other algorithms, though there is no clear gap around the central source as in reconstructions from algorithms.

⁴<https://gitlab.com/aroffringa/ws-clean/>

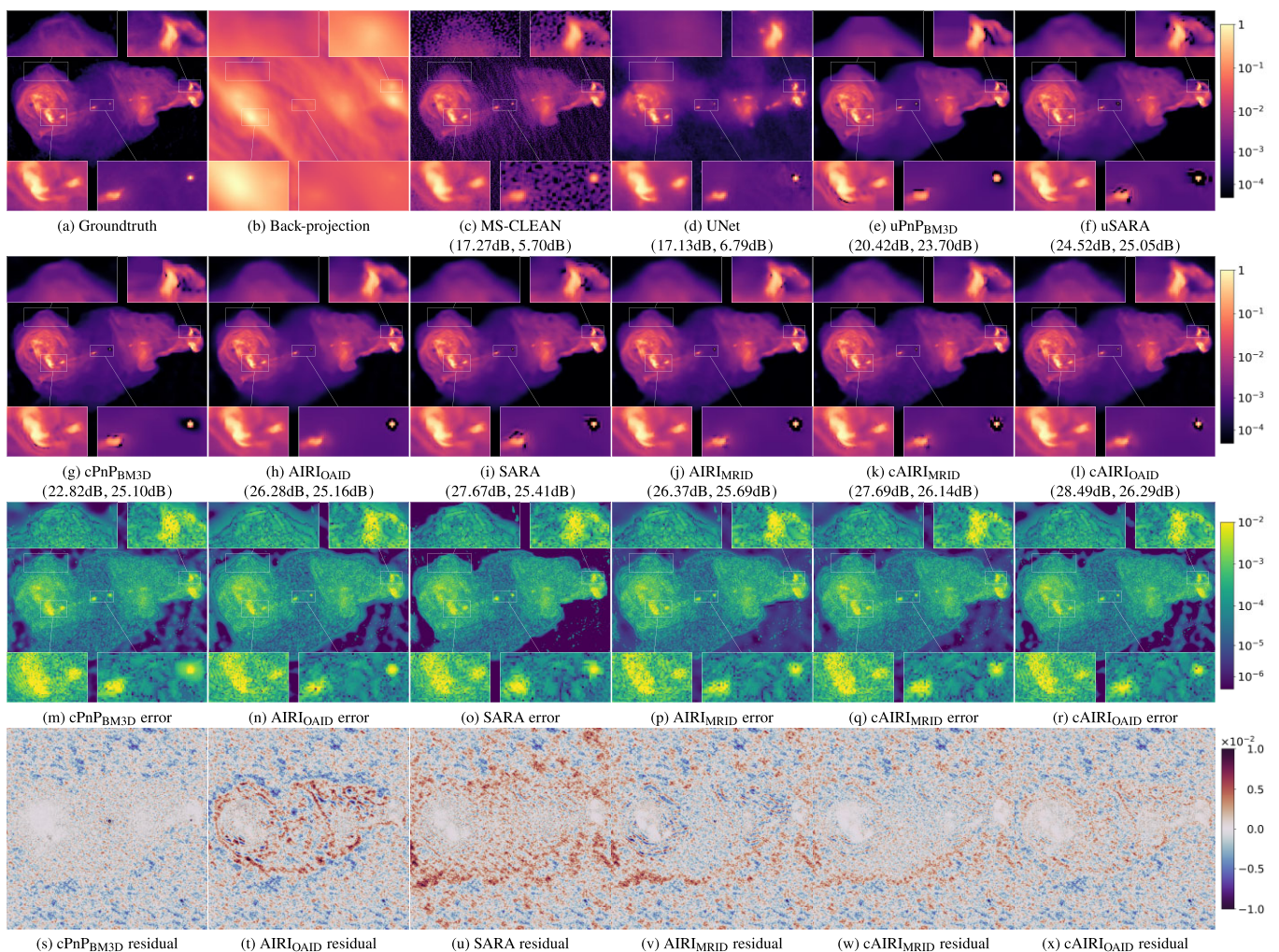


Figure 3. Results of various algorithms for the RI imaging problem. Top left image (a) shows the groundtruth image, (b) shows the simulated measurements back-projected in the real image domain. Other images on rows 1 and 2 show reconstructions with different imaging algorithms in increasing order of logSNR values. The metrics (SNR, logSNR) associated with each reconstruction are shown beneath each image. Row 3 shows error maps for different algorithms in row 2. Row 4 shows residual dirty images of associated algorithms. All images in rows 1, 2, and 3 are displayed in logarithmic scale (see colourbar on the right), while the residuals in row 4 are shown in linear scale. Insets show magnified areas of the image.

This is as expected due to the convolution with the CLEAN beam and the addition of the residual dirty image. As was observed in Terris et al. (2023), UNet recovers the high intensity emissions better than MS-CLEAN, but struggles at lower intensities (see the extended purple areas).

For the other algorithms, the constrained ones generally yield better reconstructions than their unconstrained counterpart. uSARA and SARA show faithful reconstructions, despite noticeable wavelet artefacts around the central point source, which is challenging to recover for all algorithms. While no significant difference between the two methods can be noticed in logarithmic scale, the higher SNR of the SARA reconstruction implies that it recovers high intensity values better than its unconstrained variant. BM3D-based PnP algorithms (uPnP_{BM3D} and cPnP_{BM3D}) exhibit mild stripe-like artefacts [see the top right zoom regions of panels (e) and (g)], and the discontinuities around the central black hole form significant rectangular shapes. These artefacts could be induced by the square patches selection and hard thresholding strategies within the BM3D algorithm. We recall that these algorithms were fine-tuned to achieve best performance. PnP algorithms with the proposed denoisers AIRI

and cAIRI yield better reconstructions than other methods, with little noticeable visual difference between each method despite variations in the reconstruction metrics. However, AIRI_{OAIID} and AIRI_{MRID} do not recover high intensities as well as their constrained variants cAIRI_{OAIID} and cAIRI_{MRID}. This is consistent with the results obtained with uSARA and SARA. Interestingly, the visual results of the OAIID and MRID AIRI algorithms are very similar, with the MRID variants exhibiting slightly sharper discontinuities around the bright structures. This may be due to the lack of abrupt intensities changes in the MRI knee images compared to the astronomical images in the training data sets, as shown in Fig. 2.

Notable differences emerge at the level of residual dirty images where the SARA algorithm exhibits higher intensities in the back-ground. With the same set of denoisers plugged in, the constrained [panels (t) and (x)] and unconstrained [panels (v) and (w)] algorithms yield similar patterns in residual images, but the ones generated with constrained algorithms tend to have lower intensities. Meanwhile, the residual dirty image of cPnP_{BM3D} [panel (s)] is the closest to pure noise in large, smooth regions compared to others, thus maintaining better data fidelity. However, its absolute error at relatively brighter

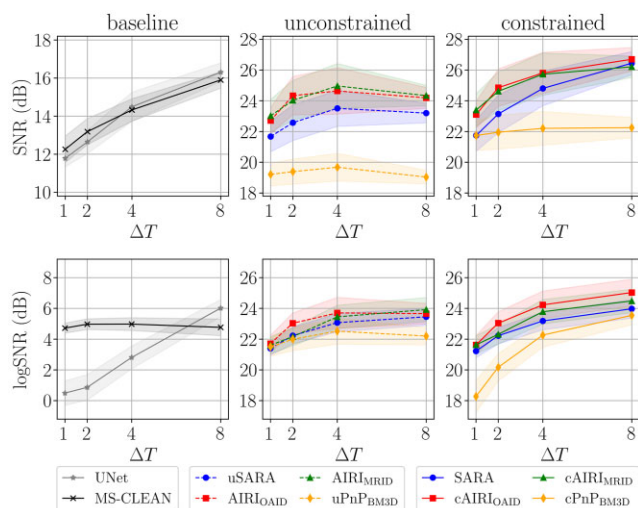


Figure 4. Reconstruction metrics as a function of the observation time. From left to right: (i) MS-CLEAN and UNet as baseline algorithms; (ii) unconstrained algorithms based on Algorithm 1; (iii) constrained algorithms based on Algorithm 2. Top row gives SNR metrics and bottom row gives logSNR metric. Each point is an average value over 15 inverse problems. The shaded areas represent the 95 per cent confidence intervals.

Table 1. Average and standard deviation of total and per iteration time (in seconds) for solving 60 simulated inverse problems with different denoisers D when plugged in Algorithms 1 and 2.

Method	Iterations	Total time (s.)		Iteration time (s.)	
		Avg.	Std.	Avg.	Std.
AIRI	6k	2846	1939	0.47	0.64
uSARA	5k	7176	3086	1.38	1.10
uPnP _{BM3D}	0.6k	10 565	2979	16.41	1.95
cAIRI	4k	2443	1723	0.61	0.59
cPnP _{BM3D}	0.4k	6311	2625	15.67	1.33
SARA	6k	5229	677	0.57	0.16

and compact structures is also higher than for other algorithms. This observation also indicates that faint residual images do not necessarily correlate with accurate reconstruction.

Fig. 4 gives the average SNR and logSNR estimated with various algorithms. On these plots, each point is an average over the 15 inverse problems considered for each observation time, error bars representing the 95 per cent confidence intervals. Both uSARA and SARA optimization algorithms significantly outperform the baseline MS-CLEAN algorithm and the UNet for the two considered metrics. AIRI and cAIRI improve by 1 to 2 dB over their pure optimization counterparts uSARA and SARA. The constrained variants perform better than the unconstrained ones, especially on the SNR where cAIRI outperforms AIRI. We explain this phenomenon by the good ability of constrained methods to recover high intensity values after fewer iterations than their unconstrained variants.

In Table 1, we provide the computational times of each algorithm for various denoisers. All algorithms were implemented in MATLAB and executed on Cirrus, an EPSRC Tier2 national HPC service. For fair comparison, AIRI algorithms were allocated 10 CPU cores and 1 NVIDIA V100 GPU, while other algorithms used 10 CPU cores. Consistent with Terris et al. (2023), both AIRI and cAIRI demonstrate faster execution than their variational counterparts. At the algorithmic level, the timings are driven by two steps: data-

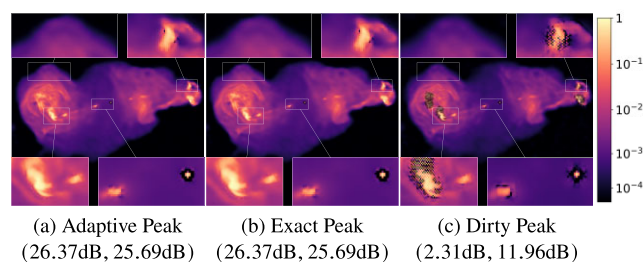


Figure 5. AIRI_{MRID} reconstructions of the same simulated measurement as in Fig. 3 with different image peak value α estimation schemes: (a) adaptive peak value estimation; (b) using the exact peak value of the groundtruth (1.00); and (c) using the maximum intensity of the back-projected image (61.18). The metrics (SNR, logSNR) associated with each reconstruction are shown beneath each image.

consistency steps (related to f) and prior related terms (related to D). For low-dimensional problems, D is the computational bottleneck. Notably, for both variational and PnP approaches, the use of a constraint in equation (5) does not translate into slower algorithms. This experiment shows that deep-learning-based models, benefiting from GPU acceleration, perform slightly faster than the SARA proximal operator. In contrast, BM3D's slow inference time raises concerns about its practical applicability. For more detailed insights into this comparison, see Terris et al. (2023), which uses the same implementation. We emphasize that these timings are strongly dependent on the specific implementation frameworks (e.g. MATLAB, CPU versus GPU) that are subject to change, and should be interpreted accordingly. Consistent comparison in a fully homogeneous PYTHON GPU framework is the subject of ongoing work (Aghabiglou et al. 2024).

3.5 Effectiveness of the adaptive peak estimation approach

In our simulated experiments, the true peak value in the sought reconstructed image is known (it is equal to 1) and the peak value of the back-projected images vary from 30 to 68 across measurement settings. To validate the adaptive peak estimation scheme proposed in Section 2.5, we ran the simulated experiments where this approach is compared to two alternatives, where $\tilde{\alpha}$ is set to either (i) the exact maximum intensity of the groundtruth image, or (ii) the maximum intensity of the back-projected image. Results in Fig. 5 show that the proposed adaptive peak estimation scheme yields reconstructions with the same quality as if the true peak value were known. If the estimated peak value differs too much from the real peak value, like in the situation from Fig. 5(c), the reconstruction quality can degrade significantly. We attribute this suboptimality to the fact that the denoisers are applied to images with peak value far below 1, a range for which they have not been trained.

3.6 Robustness to denoiser realization and model uncertainty

To investigate the robustness of the proposed approach, and following the discussion in Section 2.6, we trained $K = 15$ denoisers per training data set (i.e. OAID and MRID) under identical experimental conditions, but with different random seeds for the (pseudo) random processes during training. Each denoiser was then used in Algorithm 1 and Algorithm 2, resulting in two sets of 15 solutions, one for each training data set. For more details, see Section 2.6.

Fig. 6 shows visual results for Algorithm 1 and Algorithm 2 with different denoisers realizations. We notice that the reconstructions

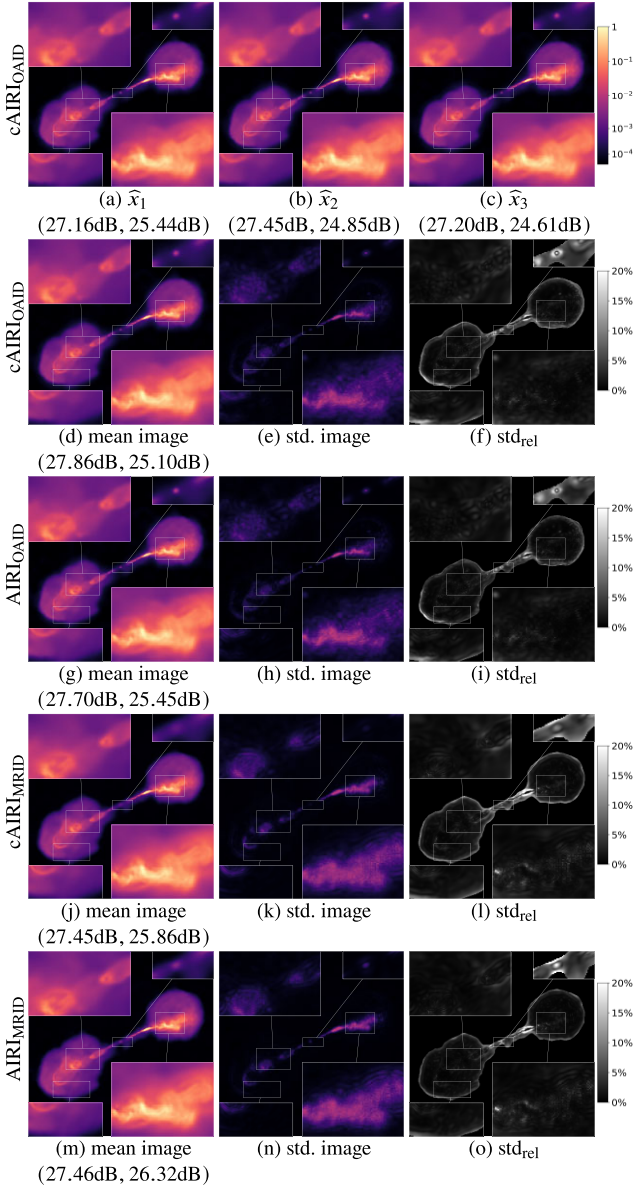


Figure 6. Sensitivity of the imaging result to the training conditions of the denoiser realization. Panels (a)–(c) show image reconstructions with cAIRI_{OAID} for three denoisers, each trained with a different random seed. Panels (d)–(f) show the mean image, standard deviation and relative standard deviation maps for reconstructions obtained with 15 networks trained on the astronomical data set and plugged in the cAIRI_{OAID} algorithm. Next rows show the same results but obtained with other AIRI variants. Insets show magnified areas of the image. Images (a)–(c), the mean images and standard deviation images share the same colour bar.

from each random seed, as well as the mean image, show no clear difference, except around the point source at the centre of the image.

The relative standard deviation, which highlights regions with mean pixel intensities surpassing the estimated noise levels as defined by equation (9), generally demonstrates that the empirical standard deviation remains modest, typically less than 10 per cent of the mean intensity across most of the image. Notably, this deviates around the central source where the standard deviation increases to about 20 per cent of the mean intensity. This observation is particularly relevant as the central source estimation is often unsatisfying, most algorithms hallucinating a black ring around it. Our cAIRI

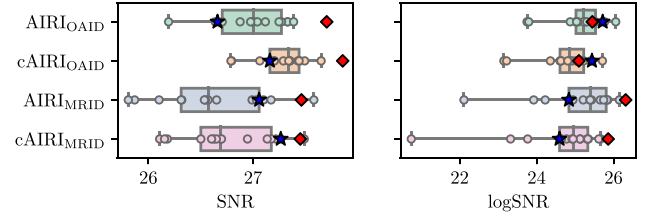


Figure 7. Results for 15 different runs of the AIRI and cAIRI algorithms for simulated measurements with $\Delta T = 4$ h. The red diamond indicates the reconstruction metric of the mean image of the 15 reconstructions. The blue star indicates the models that are used in our deterministic experiments (in Figs 3 and 4).

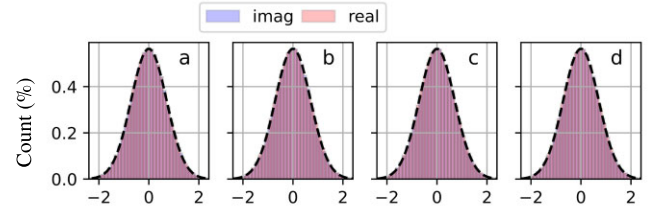


Figure 8. Histograms (real and imaginary parts) of the whitened residuals of the visibilities, i.e. $(\Phi \mathbf{x} - \mathbf{y})/\tau$ associated with the reconstructions from Fig. 7. (a) AIRI_{OAID}, (b) cAIRI_{OAID}, (c) AIRI_{MRID}, (d) cAIRI_{MRID}. The black dashed line represents the p.d.f. of Gaussian distribution with 0 mean and a variance of 0.5.

algorithm demonstrates better performance in estimating central sources, especially when combined with OAID AIRI denoisers, which gives a smaller relative standard deviation around the central sources (the top right zoom in Fig. 6l).

Fig. 7 gives the reconstruction metrics for 15 denoiser realizations in the context of a single inverse problem at $\Delta T = 4$ h, with selected samples depicted in Fig. 6. The reconstruction quality offered by different denoiser realizations spans about 1.5 dB in SNR and 4 dB in logSNR. As illustrated in Figs 6 and 7, the variations in SNR and logSNR metrics among the reconstructed images are not readily discernible at the visual image quality.

In general, AIRI tends to provide better logSNR metrics than cAIRI, while the opposite holds for SNR, and reconstructions produced with cAIRI exhibit less variation in SNR compared to those generated with AIRI for the selected measurement. Interestingly, the SNR of both the mean images is slightly higher than the mean SNR, indicating that the mean image offers enhanced fidelity. Yet, computing this image necessitates multiple denoiser realizations and algorithm runs.

Furthermore, we notice in Fig. 6 that while algorithms relying on OAID-based denoisers tend to provide reconstructions with higher SNR, those relying on MRID-based denoisers tend to provide higher logSNR. Visually, we observe a similar trend as in the previous sections where MRID-based algorithms tend to yield less precise point sources, but these differences are mild.

Lastly, we show in Fig. 8 the histogram of the whitened visibility residuals, i.e. $(\Phi \mathbf{x} - \mathbf{y})/\tau$, on all inverse problems from Section 3. We notice that these residuals closely follow the expected probability density function of the additive measurement noise distribution, supporting the statistical consistency of our reconstructions.

As the overall variation in reconstruction quality remains small, these results suggest that our AIRI PnP algorithms are resilient to variations inherent to the training process.

4 APPLICATION TO REAL ASTRONOMICAL DATA

In this section, we validate our robust PnP approach on real observations of the radio galaxy ESO 137–006, the loudest radio galaxy in the Norma cluster, acquired with the MeerKAT telescope. This FoV has recently sparked interest in the radio astronomy community for the observation of collimated synchrotron threads (Ramatsoku et al. 2020; Dabbech et al. 2022). We stress that we rely on the same shelf of denoisers that was used in the experiments on simulated data, meaning that, in line with the PnP motivation, no denoisers had to be trained specifically for the experiments in this section.

4.1 Data acquisition

The considered measurements correspond to a subset of the data imaged in Ramatsoku et al. (2020) and Dabbech et al. (2022) acquired by MeerKAT. More specifically, we selected observations at 10 consecutive frequency channels, spanning the frequency range 963 to 971 MHz for the formation a monochromatic intensity image. The measurement is self-calibrated for phase using the WSClean imager (Offringa & Smirnov 2017) and the CubiCal calibration suite (Kenyon et al. 2018). Full details are provided in Ramatsoku et al. (2020). We maintain the same image pixel size as in Dabbech et al. (2022) ($1.68 \text{ arcsec pixel}^{-1}$) corresponding to a super-resolution factor around 2, but with an image size of 2560×2560 and a FoV of 1.19×1.19 square degrees. These imaging considerations minimizes the computational load and reduces the w-effect, and is also motivated by the limited presence of objects of interest outside this window.

4.2 Imaging results

We show the full imaged FoV in Fig. 9 for both the baseline MS-CLEAN algorithm⁵ and cAIRI_{OAID}. More reconstruction results are displayed in Fig. 10, focusing on the radio galaxy ESO 137–006. The MS-CLEAN hyperparameter configuration largely follows Dabbech et al. (2022), with two modifications: we enabled the w-gridder and used MS-CLEAN’s automatic scale selection instead of a pre-defined set of scales, as we found this did not significantly impact the reconstruction quality. Due to the visual similarities of the reconstructions between AIRI and cAIRI with same set of denoisers, only cAIRI results are illustrated. The maximum intensities of the reconstructions given by various methods are around $0.05 \text{ Jy pixel}^{-1}$ with heuristic noise level around 7.6×10^{-6} , which yields a dynamic range of approximately 10^4 . Therefore, we set $x_{\max} = 0.05$ and $a = 10^4$ in $\text{rlog}_a(\cdot)$ for visualization. Overall, the imaging quality for all algorithms aligns with the observations from experiments using simulated data (see Section 3). Both the pure optimization algorithms (SARA and uSARA) and the PnP algorithms significantly outperform MS-CLEAN. The cAIRI_{OAID} reconstruction shown in Fig. 9 achieves higher spatial resolution compared to MS-CLEAN. This is particularly evident in zoom regions iii and

iv, where cAIRI_{OAID} recovers more compact point sources while preserving similar flux values. Specifically, for the point source in region iii, MS-CLEAN and cAIRI_{OAID} recover fluxes of 0.12 and 0.13 Jy, respectively, while for region iv, the recovered fluxes are 0.19 and 0.21 Jy. However, SARA exhibits significant wavelet artefacts, which could be explained by calibration errors in estimating Φ and inadequate estimation of the noise standard deviation for e in equation (1), on which SARA’s hyperparameters rely (Thouvenin et al. 2023).

The artefacts induced by residual calibration errors can be seen in all algorithms, presenting as patterns distributed on arches and some filaments around the core of the galaxy. Both cAIRI_{OAID} and cAIRI_{MRID} produce images with fewer artefacts than SARA and better resolution (see bottom-left zoom). We however underline that both reconstructions show algorithm specific artefacts: the cAIRI_{OAID} yields reconstructions with dotted artefacts (see the filaments in Fig. 10h), while the cAIRI_{MRID} yields light ringing artefacts in the reconstruction (see the filaments in Fig. 10j). The reconstruction of MS-CLEAN is smoother than others due to the convolution with the clean beam, though it gives fainter residual image (see panel c). Meanwhile, residuals in panels (e), (g), (c), and (k) of Fig. 10 suggest that the reconstructions with uSARA, SARA, cAIRI_{OAID}, and cAIRI_{MRID} show less accurate data-fidelity constraints. However, as was observed on simulated data, lower residuals do not necessarily imply more accurate reconstructions (see Fig. 3), especially with the presence of calibration errors.

As in Section 3.6, we run cAIRI with 15 different denoisers and investigate the statistics of the reconstructed images. The results on denoisers trained on OAID show that the mean Fig. 11(a) is consistent with a single sample Fig. 10(h), and that only very mild variations occur between the different solutions to the problem, as suggested by the empirical standard deviation image Fig. 11(c). Furthermore, the small ratio between the standard deviation and the mean Fig. 11(e) confirms that the variations in the reconstructions related to the model uncertainty are small compared to the image itself.

We observe different results for denoisers trained on MRID in Fig. 11(b), (d), and (f). The mean image Fig. 11(b) shows a different background value than on the sample Fig. 10(j), along with stronger localized artefacts. The standard deviation map Fig. 11(d) suggests that different denoiser realizations can yield highly different background values in the reconstruction. The ratio map Fig. 11(f) suggests that these variations dominate the signal at very faint intensities.

This study suggests that models trained on the proposed OAID data set are more robust than those trained on the MRID data set. In practice, we observed that some of the networks trained on the MRID data set showed unstable results when plugged in the PnP algorithm for real imaging, despite the non-expansive constraint of the denoisers. We did not observe such phenomenon in the simulated experiments, and presume this might result from calibration errors in the measurement operator.

5 CONCLUSION

We have introduced variations of the AIRI PnP algorithm, towards a more general and robust PnP paradigm for RI imaging. First, we have shown that the AIRI denoisers can be plugged without any alteration into a PDFB optimization backbone, leading to cAIRI, PnP variant to the constrained SARA optimization algorithm itself. This extends the remit of the AIRI paradigm, with the original AIRI and the new cAIRI respectively representing the PnP counterparts to uSARA and SARA. Results from simulated and real data confirm that cAIRI performs better than AIRI, while each respectively improved

⁵The full command is: `wsclean -size 2560 2560 -scale 1.68asec -channel-range 3 12 -channels-out 1 -mem 80 -weight briggs 0 -super-weight 1.0 -weighting-rank-filter-size 16 -taper-gaussian 0 -use-wgridder -kernel-size 7 -oversampling 63 -pol I -intervals-out 1 -auto-threshold 0.5 -auto-mask 2 -gain 0.1 -mgain 0.9 -multiscale -multiscale-scale-bias 0.6 -niter 5000000.`

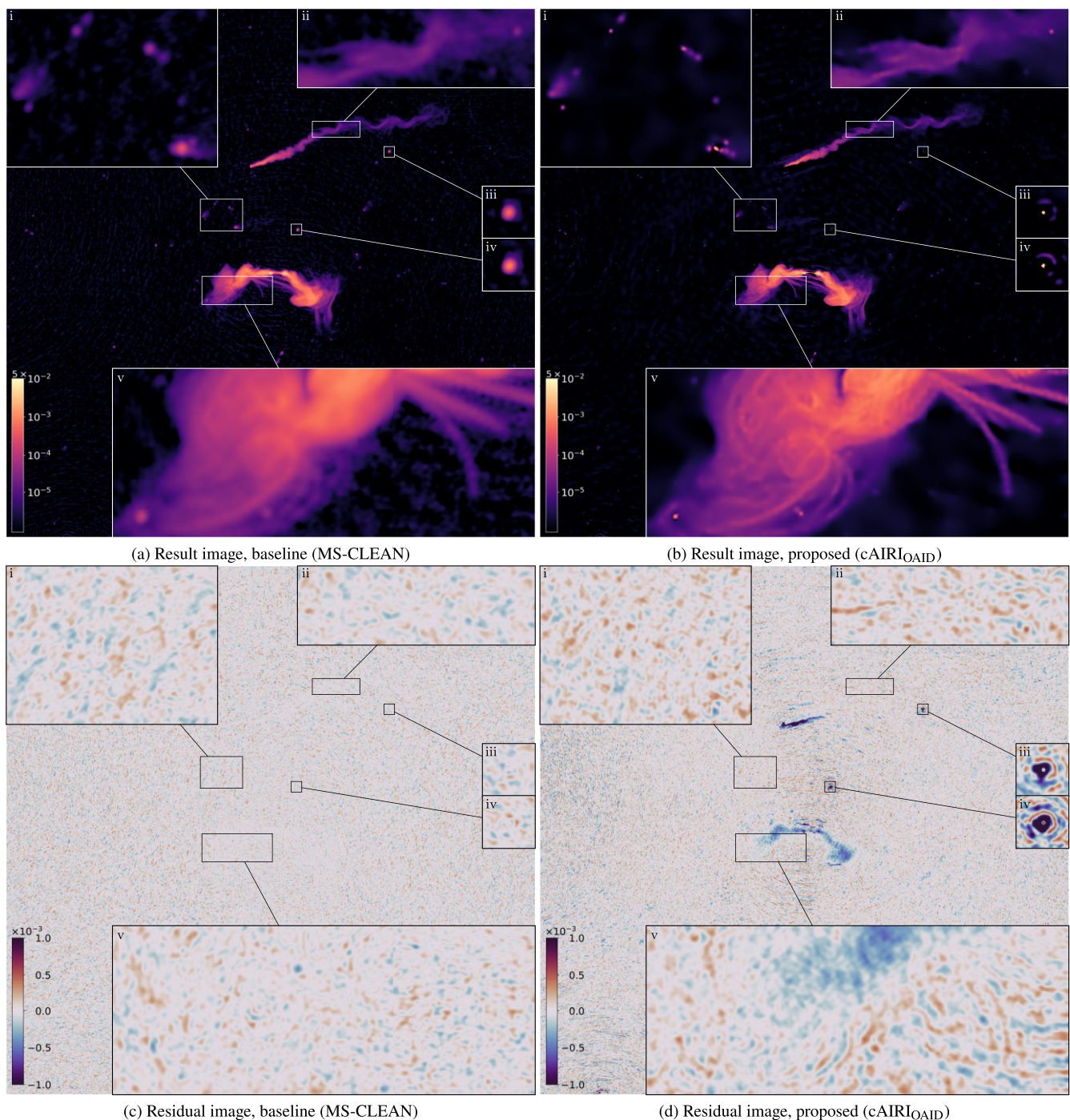


Figure 9. Imaging of ESO 137–006 galaxy and other radio sources of interest from data acquired with the MeerKAT telescope with an image size of 2560×2560 and a FoV of 1.19×1.19 square degree: (a) imaging result with the baseline MS-CLEAN algorithm; (b) imaging result with the proposed cAIRI_{OAID} algorithm; (c) residual dirty image of MS-CLEAN; (d) residual dirty image of cAIRI_{OAID}.

on their pure optimization counterparts. All algorithms deliver better reconstruction quality than MS-CLEAN. We interpret the reported higher performance of the constrained data-fidelity approach in cAIRI over AIRI as being related to the characteristic high-dynamic ranges of interest in RI imaging, with some high intensity point sources better resolved with constrained algorithms.

Secondly, we have studied the robustness of the AIRI paradigm to strong variations in the nature of the training data set, with denoisers trained on MRI-born images (MRID) yielding similar reconstruction

quality to those trained on optical astronomy images (OAID). The capability to adapt the dynamic range of the data set to a specific target dynamic range of observation via bespoke exponentiation procedures enables training a shelf of denoisers tailored to high dynamic ranges. In this context, we have proposed to refine the estimated target dynamic range of observation at each iteration during reconstruction, and switch denoisers accordingly on the fly. This adaptive set-up was shown to improve and stabilise the reconstruction process.

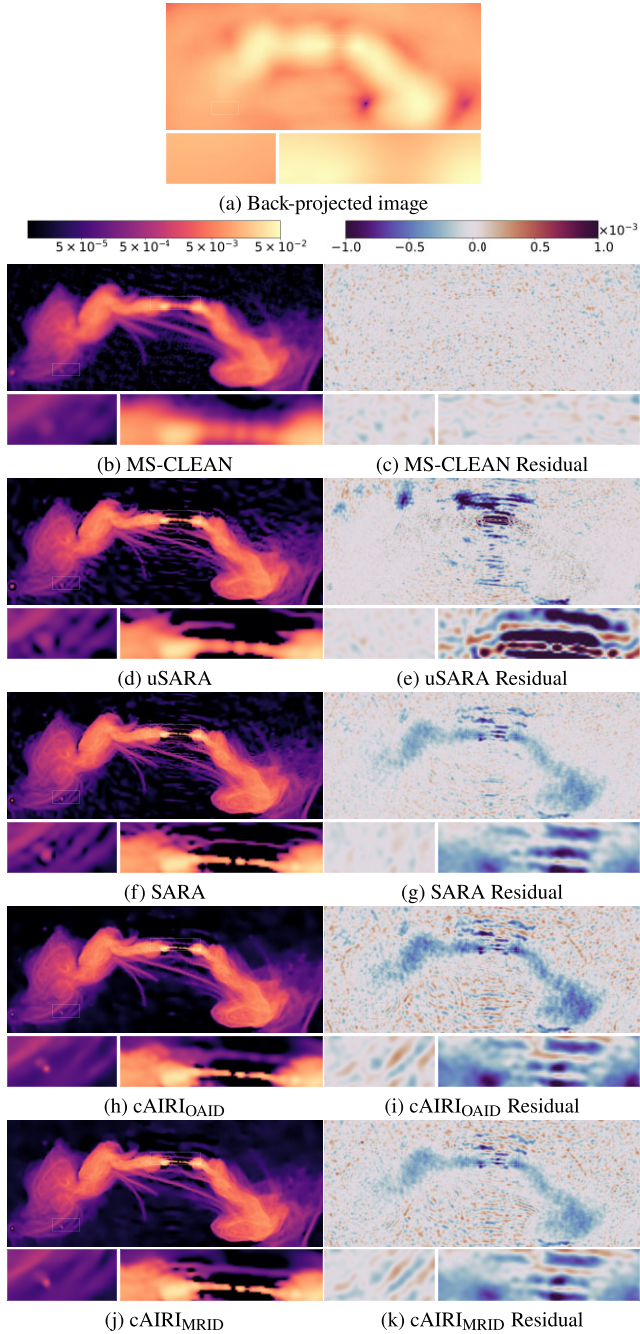


Figure 10. Imaging of a wide FoV (same as in Fig. 9) containing the ESO 137–006 galaxy. Due to the large size (2560×2560), we only display a centre-crop of size 640×256 over the ESO 137–006 galaxy. Below each image, we provide zooms on areas of interest corresponding to white rectangles. Panel (a) shows the back-projected dirty image. For the rest panels, reconstructions of various methods are shown in the left column. The corresponding residual dirty images are shown in the right column.

Thirdly, we analysed the epistemic uncertainty quantification method in PnP algorithms. Our experiments show that AIRI denoisers, despite variations in parameter initialization, noise realization, and patch selection, consistently yield similar results when integrated into AIRI or cAIRI algorithms. This ensemble of solutions provides an epistemic uncertainty metric for the AIRI paradigm. It is worth pointing out that, while results show minimal relative standard

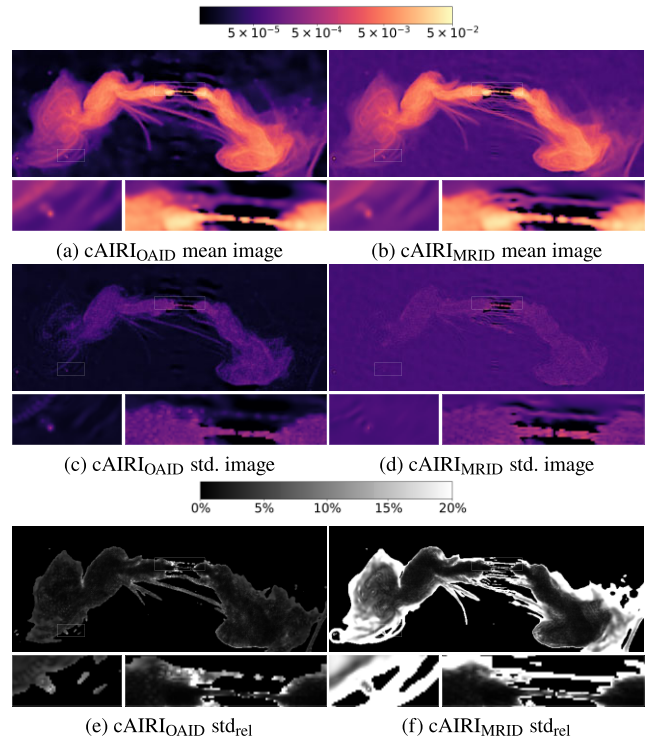


Figure 11. Sensitivity of the MeerKAT real measurement imaging results to the training conditions of the denoiser realization for cAIRI_{OAID} and cAIRI_{MRID}. Same as Fig. 10, we only display a centre-crop of size 640×256 over the ESO 137–006 galaxy. The left column shows the statistical imaging results over 15 different reconstructions of cAIRI_{OAID} with denoisers trained with OAID and different random status. The right column shows cAIRI_{OAID} results with various denoisers trained with MRID. Each row shows the pixel-wise mean images, standard deviation images and relative standard deviation maps, respectively.

deviations in reconstructed pixel intensities for OAID denoisers, higher values are reported with MRID denoisers, suggesting greater imaging robustness with the former over the latter, in particular in larger-dimensional image setting and in the presence of calibration errors.

Finally, future research directions include developing wideband and calibration functionality for the AIRI paradigm.

ACKNOWLEDGEMENTS

The authors thank O. Smirnov for providing the real observations with MeerKAT and A. Dabbech for helpful discussions. This work was supported by the Engineering and Physical Sciences Research Council (EPSRC), with work funded under the EPSRC grants EP/T028270/1 and EP/T028351/1, and the STFC grant ST/W000970/1. The research used Cirrus, a UK National Tier-2 HPC Service at EPCC funded by the University of Edinburgh and EPSRC (EP/P020267/1). The MeerKAT telescope is operated by SARAO, which is a facility of DST/NRF.

DATA AVAILABILITY

AIRI and cAIRI codes are available alongside the SARA codes in the [BASPLib](#) code library on GitHub. BASPLib is developed and maintained by the Biomedical and Astronomical Signal Processing

Laboratory (BASP). The denoisers trained with OAID and MRID are available on the [Heriot-Watt University research portal](#).

Images used to generate training, validation, and testing data sets are sourced as follows. Optical astronomy images are gathered from NOIRLab/NSF/AURA/H.Schweiker/WIYN/T.A.Rector (University of Alaska Anchorage). Medical images are obtained from the NYU fastMRI Initiative data base (Zbontar et al. 2018).

The utilized real data are observations with the MeerKAT telescope (Project ID SCI-20190418-SM-01).

REFERENCES

- Aghabiglou A., Chu C. S., Dabbech A., Wiaux Y., 2024, *ApJS*, 273, 3
- Arras P., Bester H. L., Perley R. A., Leike R., Smirnov O., Westermann R., Enßlin T. A., 2021, *A&A*, 646, A84
- Baumgartner C. F. et al., 2019, in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part II 22*. Springer Nature, p. 119
- Bauschke H. H. et al., 2011, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. CMS Books in Mathematics. Springer
- Cai X., Pereyra M., McEwen J. D., 2018, *MNRAS*, 480, 4154
- Carrillo R. E., McEwen J. D., Wiaux Y., 2012, *MNRAS*, 426, 1223
- Combettes P. L., Pesquet J.-C., 2011, *Fixed-point Algorithms for Inverse Problems in Science and Engineering*. p. 185
- Connor L., Bouman K. L., Ravi V., Hallinan G., 2022, *MNRAS*, 514, 2614
- Cornwell T. J., 2008, *IEEE J. Select. Topics Signal Process.*, 2, 793
- Dabbech A., Aghabiglou A., Chu C. S., Wiaux Y., 2024, *ApJL*, 966, L34
- Dabbech A., Ferrari C., Mary D., Slezak E., Smirnov O., Kenyon J. S., 2015, *A&A*, 576, A7
- Dabbech A., Terris M., Jackson A., Ramatsoku M., Smirnov O. M., Wiaux Y., 2022, *ApJ*, 939, L4
- Dabov K., Foi A., Katkovnik V., Egiazarian K., 2007, *IEEE Trans. Image Process.*, 16, 2080
- Dia N., Yantovski-Barth M., Adam A., Bowles M., Lemos P., Scaife A. M., Hezaveh Y., Perreault-Levasseur L., 2023, *Machine Learning and the Physical Sciences Workshop, NeurIPS 2023*
- Edupuganti V., Mardani M., Vasanawala S., Pauly J., 2020, *IEEE Trans. Med. Imaging*, 40, 239
- Gao S., Zhuang X., 2022, *IEEE Trans. Pattern Anal. Machine Intell.*, 45, 1405
- Garsden H. et al., 2015, *A&A*, 575, A90
- Gheller C., Vazza F., 2022, *MNRAS*, 509, 990
- Högbom J., 1974, *A&AS*, 15, 417
- Jonas J., 2016, *MeerKAT Science: On the Pathway to the SKA*, 1
- Junklewitz H., Bell M., Selig M., Enßlin T., 2016, *A&A*, 586, A76
- Kazemi S., Yatawatta S., Zaroubi S., Lampropoulos P., De Bruyn A., Koopmans L., Noordam J., 2011, *MNRAS*, 414, 1656
- Kendall A., Gal Y., 2017, *Advances in Neural Information Processing Systems*. Vol. 30
- Kenyon J., Smirnov O., Grobler T., Perkins S., 2018, *MNRAS*, 478, 2399
- Kwon Y., Won J.-H., Kim B. J., Paik M. C., 2018, *Med. Imaging Deep Learn.*, 4, 2
- Lahlou S., Jain M., Nekoei H., Butoi V. I., Bertin P., Rector-Brooks J., Korablyov M., Bengio Y., 2021, *Transactions on Machine Learning Research*
- Laumont R., Bortoli V. D., Almansa A., Delon J., Durmus A., Pereyra M., 2022, *SIAM J. Imaging Sci.*, 15, 701
- Liaudat T. I., Mars M., Price M. A., Pereyra M., Betcke M. M., McEwen J. D., 2024, *RAS Techniques and Instruments*, 3, 505
- Mallat S. G., Zhang Z., 1993, *IEEE Trans. Sig. Process.*, 41, 3397
- Marsh K., Richardson J., 1987, *A&A*, 182, 174
- Missaglia V., Massaro F., Capetti A., Paolillo M., Kraft R., Baldi R. D., Paggi A., 2019, *A&A*, 626, A8
- Mukherjee S., Hauptmann A., Öktem O., Pereyra M., Schönlieb C.-B., 2023, *IEEE Signal Process. Magaz.*, 40, 164
- Narnhofer D., Effland A., Kobler E., Hammernik K., Knoll F., Pock T., 2021, *IEEE Trans. Med. Imaging*, 41, 279
- Offringa A. et al., 2014, *MNRAS*, 444, 606
- Offringa A., Smirnov O., 2017, *MNRAS*, 471, 301
- Onose A., Carrillo R. E., Repetti A., McEwen J. D., Thiran J.-P., Pesquet J.-C., Wiaux Y., 2016, *MNRAS*, 462, 4314
- Pesquet J.-C., Repetti A., Terris M., Wiaux Y., 2021, *SIAM J. Imaging Sci.*, 14, 1206
- Ramatsoku M. et al., 2020, *A&A*, 636, L1
- Repetti A., Birdi J., Dabbech A., Wiaux Y., 2017, *MNRAS*, 470, 3981
- Repetti A., Pereyra M., Wiaux Y., 2019, *SIAM J. Imaging Sci.*, 12, 87
- Repetti A., Wiaux Y., 2020, in *ICASSP 2020–2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. p. 1434
- Roth J., Arras P., Reinecke M., Perley R. A., Westermann R., Enßlin T. A., 2023, *A&A*, 678, A177
- Scaife A., 2020, *Phil. Trans. R. Soc. A*, 378, 20190060
- Schwab F., 1984, *AJ*, 89, 1076
- Terris M., Dabbech A., Tang C., Wiaux Y., 2023, *MNRAS*, 518, 604
- Terris M., Moreau T., Pustelnik N., Tachella J., 2024, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. p. 25255
- Thompson A. R., Moran J. M., Swenson G. W., 2017, *Interferometry and Synthesis in Radio Astronomy*. Springer Nature
- Thouvenin P.-A., Abdulaziz A., Dabbech A., Repetti A., Wiaux Y., 2023, *MNRAS*, 521, 1
- Wiaux Y., Jacques L., Puy G., Scaife A. M., Vanderghenst P., 2009, *MNRAS*, 395, 1733
- Wilber A. G., Dabbech A., Jackson A., Wiaux Y., 2023a, *MNRAS*, 522, 5558
- Wilber A. G., Dabbech A., Terris M., Jackson A., Wiaux Y., 2023b, *MNRAS*, 522, 5576
- Wittor D., Ettori S., Vazza F., Rajpurohit K., Hoeft M., Domínguez-Fernández P., 2021, *MNRAS*, 506, 396
- Xue Y., Cheng S., Li Y., Tian L., 2019, *Optica*, 6, 618
- Zbontar J. et al., 2018, preprint ([arXiv:1811.08839](#))
- Zhang C., Jin B., 2019, preprint ([arXiv:1908.01010](#))
- Zhang K., Li Y., Liang J., Cao J., Zhang Y., Tang H., Timofte R., Van Gool L., 2023, *Mach. Intell. Res.*, 20, 822
- Zhang K., Li Y., Zuo W., Zhang L., Van Gool L., Timofte R., 2022, *IEEE Trans. Pattern Anal. Machine Intell.*, 44, 6360
- Zhang K., Zuo W., Chen Y., Meng D., Zhang L., 2017, *IEEE Trans. Image Process.*, 26, 3142

This paper has been typeset from a \LaTeX file prepared by the author.