

Generative imaging for radio interferometry with fast uncertainty quantification

Matthijs Mars,^{1,2★} Tobías I. Liaudat,³ Jessica J. Whitney,¹ Marta M. Betcke,⁴ Jason D. McEwen^{1,5†}

¹Mullard Space Science Laboratory (MSSL), University College London (UCL), Dorking RH5 6NT, UK

²Leiden Observatory, Leiden University, Leiden 2333 CC, The Netherlands

³IRFU, CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France

⁴Department of Computer Science, University College London (UCL), London WC1E 6BT, UK

⁵Alan Turing Institute, London NW1 2DB, UK

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

With the rise of large radio interferometric telescopes, particularly the SKA, there is a growing demand for computationally efficient image reconstruction techniques. Existing reconstruction methods, such as the CLEAN algorithm or proximal optimisation approaches, are iterative in nature, necessitating a large amount of compute. These methods either provide no uncertainty quantification or require large computational overhead to do so. Learned reconstruction methods have shown promise in providing efficient and high quality reconstruction. In this article we explore the use of generative neural networks that enable efficient approximate sampling of the posterior distribution for high quality reconstructions with uncertainty quantification. Our RI-GAN framework, builds on the regularised conditional generative adversarial network (rcGAN) framework by integrating a gradient U-Net (GU-Net) architecture - a hybrid reconstruction model that embeds the measurement operator directly into the network. This framework uses Wasserstein GANs to improve training stability in combination with regularisation terms that combat mode collapse, which are typical problems for conditional GANs. This approach takes as input the dirty image and the point spread function (PSF) of the observation and provides efficient, high-quality image reconstructions that are robust to varying visibility coverages, generalises to images with an increased dynamic range, and provides informative uncertainty quantification. Our methods provide a significant step toward computationally efficient, scalable, and uncertainty-aware imaging for next-generation radio telescopes.

Key words: machine learning – image processing – interferometric imaging

1 INTRODUCTION

In order to make reliable scientific conclusions from astronomical observations, it is crucial to quantify the uncertainties in the image reconstructions. However, most current radio interferometric imaging methods do not provide uncertainty quantification, and methods that do are often computationally expensive and not scalable to the large data volumes expected from next-generation radio telescopes. Additionally, the advent of large radio interferometer telescopes such as the Square Kilometre Array (SKA, [Dewdney et al. 2009](#)) leads to an increasing need for efficient image reconstruction methods, making the provision of uncertainty quantification significantly more challenging. One of the main challenges in radio interferometric imaging is dealing with the increasingly large amount of visibilities acquired by the telescopes as this increases the computational complexity of evaluating the measurement operator of the telescope.

Current reconstruction methods, such as CLEAN ([Högbom 1974](#); [Bhatnagar & Cornwell 2004](#); [Bhatnagar et al. 2008](#); [Stewart et al. 2011](#); [Offringa et al. 2014](#)), as well as more recently developed methods based on compressed sensing techniques (e.g. [Carrillo et al. 2012](#),

[2013](#); [Pratley et al. 2018](#)) are based on iterative optimisation. These methods are computationally expensive as they have to evaluate the measurement operator at each iteration and typically provide only a point estimate without quantifying the uncertainties. While scalable and highly distributed methods have been developed to manage large data volumes (e.g. [Pratley et al. 2019](#)) as well as some methods to estimate the uncertainties in the reconstructions ([Cai et al. 2018b](#); [Liaudat et al. 2024](#); [Cherif et al. 2024](#)), these methods still struggle with the data volumes expected from the SKA.

Statistical methods like Markov chain Monte Carlo (MCMC) are capable of performing Bayesian inference by sampling the full posterior distribution and quantify the uncertainties in the reconstruction (e.g. [Sutter et al. 2014](#); [Cai et al. 2018a](#)). However, these methods are particularly computationally expensive due to the iterative nature of MCMC sampling techniques, where the measurement operator needs to be evaluated a large amount of times. Therefore, these methods do not scale well to larger data volumes.

An alternative Bayesian approach, that does not require sampling of the full posterior distribution, is the RESOLVE method ([Junklewitz et al. 2016](#); [Greiner et al. 2017](#); [Arras et al. 2018](#)). This method uses a variational approach to approximate the posterior distribution using a technique called metric Gaussian variational inference (MGVI,

★ E-mail: academic@matthijsmars.com

† E-mail: jason.mcewen@ucl.ac.uk

Knollmüller & Enßlin 2020) and uses samples from this approximate posterior to estimate the uncertainties in the reconstruction.

Recent advances in machine learning for radio interferometric image reconstruction have shown promise in terms of superior reconstruction quality and computational efficiency. Learned regularisation methods, which employ a learned prior or regulariser in a variational regularisation framework, have demonstrated high reconstruction quality, robustness to varying measurement operators, and good generalising to unseen data (e.g. Terris et al. 2022, 2023; Liaudat et al. 2024). Post-processing methods, which are computationally more efficient than iterative based approaches, remove artifacts from an initial reconstruction and provide high-quality results (e.g. Allam Jr 2016; Connor et al. 2022; Mars et al. 2023, 2025). These methods are less robust to varying visibility coverages and require full retraining for each new observation, though more efficient training strategies can be used to mitigate this somewhat (Mars et al. 2025). Unrolled methods, which are trained to perform a fixed number of learned updates in an iterative algorithm, have shown to provide high quality reconstructions and require far fewer iterations than fully iterative methods (e.g. Mars et al. 2025; Aghabiglou et al. 2024; Dabbech et al. 2023). These methods explicitly include the measurement operator in the reconstruction network, significantly improving reconstruction quality compared to post-processing methods and providing robustness to varying visibility coverages (Mars et al. 2025).

While most learned methods only provide a single point estimate of the image, recent approaches consider the inclusion of uncertainty quantification. One strategy is to train multiple versions of the learned reconstruction models to represent different realisations of the prior distribution, which can then estimate the model uncertainty in the reconstruction by providing multiple reconstructions and examining their variance (Terris et al. 2023). Another strategy uses the theory of probability concentration (extending the work of Cai et al. (2018b) to the learned setting) to provide approximate uncertainty quantification by using a convex learned prior without the need for sampling or generating several realisations (Liaudat et al. 2024).

Conditional generative imaging methods create a generating function that maps a latent code vector (typically a vector of random noise) to an image sample of the posterior distribution, conditioned on the observed data. Popular generative models in astronomical inverse problems have been either score-based (e.g. Remy et al. 2023; Dia et al. 2025) or diffusion-based (e.g. Wang et al. 2023; Drozdova et al. 2024) methods. These methods, despite providing high quality reconstructions and diverse posterior samples, are orders of magnitude slower in generating samples than methods successfully applied in other imaging problems based on conditional generative adversarial networks (cGANs, e.g., Isola et al. 2017; Zhao et al. 2018, 2021; Adler & Öktem 2025), conditional variational autoencoders (cVAEs, e.g., Edupuganti et al. 2021; Tonolini et al. 2020; Sohn et al. 2015), or conditional normalizing flows (cNFs, e.g., Ardizzone et al. 2019; Winkler et al. 2023; Sun & Bouman 2021).

In this work we focus on using cGANs for radio interferometric image reconstruction with approximate posterior samples to build our radio interferometry GAN (RI-GAN) framework. While cGANs have been known to suffer from mode-collapse, resulting in reduced sample diversity, recent work by Bendel et al. (2023) has introduced a regularised cGAN (rcGAN). This approach uses regularisation to prevent mode-collapse and promote sample diversity, with successful applications in inpainting and magnetic resonance imaging (MRI) reconstruction problems. Extensions to astronomical imaging, e.g., by Whitney et al. (2024a,b) in mass mapping, show promise in providing both efficient and diverse posterior samples.

A significant challenge in constructing learned methods for radio interferometric imaging is the variability of the measurement operator, which depends on the telescope’s uv -coverage and varies for each observation based on the telescope’s pointing, the duration of the observation, and the rotation of the Earth. In order to encode this information into the reconstruction, we additionally condition the generator of the rcGAN with the observation’s point spread function (PSF). Previous work has shown that conditioning learned reconstruction methods on the measurement operator through explicit embedding of the operator in the reconstruction network significantly enhances reconstruction quality, robustness to varying visibility coverages, and generalisation to out-of-distribution data (Mars et al. 2025). Building on this, we also present an rcGAN generator that explicitly embeds the operator within the network. This additional conditioning enhances the reconstruction, robustness, and generalisation capabilities of our method.

The remainder of this article is structured as follows. Section 2 provides an overview of the radio interferometric imaging problem, its challenges, and a discussion of established reconstruction methods. In Section 3 we introduce the concept of generative imaging and discuss the rcGAN framework. Section 4 describes our two proposed RI-GAN models, one based on a U-Net and the other based on a gradient U-Net (GU-Net; Mars et al. 2023, 2025) generator. It also details the implementation of the rcGAN for radio interferometric imaging and the training procedure. In Section 5 we evaluate the performance of the rcGAN on simulated data and compare the two models. We also investigate the number of samples needed for the reconstructions and uncertainty estimates, and the correlation of the uncertainties with the errors of the reconstructions. Additionally, we evaluate the models on simulated measurements from the 30 Doradus region. Finally, in Section 6 we summarise our findings and discuss future work.

2 RADIO INTERFEROMETRIC IMAGING

Aperture synthesis in radio interferometric imaging is an essential technique in radio astronomy to obtain high-resolution images of the radio sky. The process involves measuring visibilities, which are Fourier components of the sky brightness distribution, using an array of radio telescopes. However, reconstructing high-fidelity images from these visibilities is a challenging problem due to the non-uniform sampling of the Fourier domain, the presence of noise, and the wide-field effects that occur when the baselines are not co-planar.

The measurement process for a radio telescope can be described as (see, e.g., Thompson et al. 2017)

$$\mathcal{V}(u, v, w) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{\sqrt{1-l^2-m^2}} A_N(l, m) I(l, m) \times \exp \left\{ -i2\pi \left[ul + vm + w \left(\sqrt{1-l^2-m^2} - 1 \right) \right] \right\} dl dm, \quad (1)$$

where $\mathcal{V}(u, v, w)$ represents the acquired visibility, $A_N(l, m)$ is the primary beam pattern, and $I(l, m)$ the sky brightness distribution. If the telescope array is small, the baselines can be considered to be co-planar and the w -term can be ignored, reducing the equation to a 2D Fourier transform. This results in the sky brightness distribution convolved with the point spread function (PSF) of the telescope. If the baselines are not co-planar, the w -term cannot be ignored and needs to be included in the measurement operator and for large fields of view direction dependent effects (DDEs) need to be accounted for (Smirnov 2011).

The measurement acquisition process can be summarised as

$$\mathbf{y} = \Phi \mathbf{x} + \mathbf{n}, \quad (2)$$

where $\mathbf{y} \in \mathbb{C}^M$ are the measured visibilities, $\mathbf{x} \in \mathbb{R}^N$ is the image to be reconstructed, $\mathbf{n} \in \mathbb{C}^M$ the noise in the measurements, and $\Phi : \mathbb{R}^N \rightarrow \mathbb{C}^M$ is the measurement operator of the telescope. The measurement operator is typically modelled using a non-uniform fast Fourier transform (NUFFT, [Duijndam & Schonewille 1997](#); [Fessler & Sutton 2003](#)), which grids the visibilities to a uniform grid in order to use the fast Fourier transform (FFT, [Dutt & Rokhlin 1993](#)) to speed up the computation. More details on implementing this measurement operator in the context of radio interferometry can be found in e.g., [Pratley et al. \(2018\)](#); [Mars et al. \(2023\)](#).

For the co-planar baseline case, the response of the telescope can be described as a convolution with the PSF of the telescope

$$\mathbf{x}_{\text{dirty}} = \Phi^* \mathbf{y} = \Phi^* \Phi \mathbf{x} = I_{\text{PSF}} * \mathbf{x}, \quad (3)$$

where $\mathbf{x}_{\text{dirty}} \in \mathbb{R}^N$ is the naturally weighted dirty image, $\Phi^* : \mathbb{C}^M \rightarrow \mathbb{R}^N$ is the adjoint of the measurement operator and where the PSF of the telescope is described as $I_{\text{PSF}} \triangleq \Phi^* \Phi \delta$, with $\delta \in \mathbb{R}^N$ being an image with a value of 1 in the centre pixel and 0 elsewhere. However, the wide-field effects and DDEs that occur from large fields of view and non-coplanar baselines are not modelled by this convolution.

3 GENERATIVE IMAGING

In this work we focus on using cGANs to provide our generative framework. Where unconditional GANs map from a latent code vector to an image, conditional GANs are also conditioned on the observed data. cGANs consist of two parts: a generator network G_θ that generates samples from the posterior distribution and a discriminator network D_ϕ that tries to distinguish between true samples and generated samples. The generator is trained to produce samples that are close to the true posterior distribution, while the discriminator is trained to distinguish between true and generated samples. These two networks are trained in an adversarial fashion, where the generator tries to fool the discriminator and the discriminator tries to distinguish between true and generated samples. This adversarial process pushes both networks to improve, leading to generators that can provide high quality samples.

Generative conditional modelling, such as cGANs, offer a powerful framework for radio interferometric imaging, enabling uncertainty quantification and generation of diverse image samples that are consistent with observed data. The goal is to train a generating function $G_\theta : \mathbf{Z} \times \mathbf{Y} \rightarrow \mathbf{X}$ that maps from the observed data $\mathbf{y} \in \mathbf{Y} \in \mathbb{C}^M$ and a latent code vector $\mathbf{z} \in \mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ to an image $\hat{\mathbf{x}} \in \mathbf{X} \in \mathbb{R}^N$ such that $\hat{\mathbf{x}} = G_\theta(\mathbf{z}, \mathbf{y}) \sim p_{\hat{\mathbf{x}}|\mathbf{y}}(\cdot | \mathbf{y})$, with the generated posterior distribution $p_{\hat{\mathbf{x}}|\mathbf{y}}(\cdot | \mathbf{y})$ being as close as possible to the true posterior distribution of the image given the observed data $p_{\mathbf{x}|\mathbf{y}}(\cdot | \mathbf{y})$. Additionally, one also trains a discriminator function $D_\phi : \mathbf{X} \rightarrow [0, 1]$ that tries to distinguish between true samples $\mathbf{x} \sim p_{\mathbf{x}|\mathbf{y}}(\cdot | \mathbf{y})$ and generated samples $\hat{\mathbf{x}} \sim p_{\hat{\mathbf{x}}|\mathbf{y}}(\cdot | \mathbf{y})$.

However, GANs, and subsequently cGANs, suffer from training instability and mode collapse ([Arjovsky & Bottou 2017](#)), where the generator produces samples with low diversity. Using the original GAN loss function [Goodfellow et al. \(2014\)](#), when the discriminator gets too good at distinguishing between the distributions of true and generated samples, the gradient of the generator vanishes and the training becomes unstable. In order to combat this, Wasserstein GANs (WGANs) have been introduced ([Arjovsky et al. 2017](#)), which use the Wasserstein-1 distance to measure the distance between the

true and generated samples. The dual formulation of the Wasserstein-1 distance is given by

$$W_1(p_{\mathbf{x}|\mathbf{y}}(\cdot | \mathbf{y}), p_{\hat{\mathbf{x}}|\mathbf{y}}(\cdot | \mathbf{y})) = \sup_{D_\phi \in \text{Lip}_1} \mathbb{E}_{\mathbf{x}|\mathbf{y}}\{D_\phi(\mathbf{x}, \mathbf{y})\} - \mathbb{E}_{\hat{\mathbf{x}}|\mathbf{y}}\{D_\phi(\hat{\mathbf{x}}, \mathbf{y})\}, \quad (4)$$

where we have a discriminator that is a 1-Lipschitz function ($D_\phi \in \text{Lip}_1$) and $\hat{\mathbf{x}} = G_\theta(\mathbf{z}, \mathbf{y})$ is the generated image. The loss function of the WGAN generator can be found by taking the expectation over several measurements $\mathbf{y} \in \mathbf{Y}$ and latent code vectors $\mathbf{z} \in \mathcal{N}(\mathbf{0}, \mathbf{I})$ and is given by

$$\mathcal{L}_{\text{adv}}(\theta, \phi) \triangleq \mathbb{E}_{\mathbf{x}, \mathbf{z}, \mathbf{y}} \{D_\phi(\mathbf{x}, \mathbf{y}) - D_\phi(G_\theta(\mathbf{z}, \mathbf{y}), \mathbf{y})\}, \quad (5)$$

which approximates the Wasserstein-1 distance. Through training the discriminator, the quality of the Wasserstein approximation improves. The loss function for the discriminator is thus given by

$$\mathcal{L}_{\text{discriminator}}(\theta, \phi) = -\mathcal{L}_{\text{adv}}(\theta, \phi) + \mathcal{L}_{\text{gp}}(\phi), \quad (6)$$

where $\mathcal{L}_{\text{gp}}(\phi)$ is a gradient penalty term that ensures the discriminator remains 1-Lipschitz ([Gulrajani et al. 2017](#)) and the loss is minimised with respect to the discriminator parameters ϕ . The expectation value over $\mathbf{x}, \mathbf{y}$ can be expressed as an average over the training pairs $\{\mathbf{x}_i, \mathbf{y}_i\}$ and the weights of the generator and discriminator networks are optimised alternately by minimising the loss functions.

Besides the training instability and mode collapse found in unconditional GANs, conditional GANs can also suffer an additional cause of mode collapse. Since there is only one true image $\mathbf{x}_i$ for each observed measurement $\mathbf{y}_i$, there is no incentive for the generator to produce diverse samples. To prevent this, [Bendel et al. \(2023\)](#) introduced a regularised cGAN (rcGAN) that uses several regularisation terms to ensure diversity in the generated samples. Instead of minimising Equation 5, the generator is trained by minimising

$$\arg \min_{\theta} \{\beta_{\text{adv}} \mathcal{L}_{\text{adv}}(\theta, \phi) + \mathcal{L}_{1, N_{\text{train}}}(\theta) - \beta_{\text{SD}} \mathcal{L}_{\text{SD}, N_{\text{train}}}(\theta)\}, \quad (7)$$

where $\beta_{\text{adv}} > 0$ and $\beta_{\text{SD}} > 0$ are hyperparameters and $N_{\text{train}} > 2$ is the number of samples used in the regularisation terms. The $\mathcal{L}_{1, N_{\text{train}}}(\theta)$ regularisation term is an N_{train} supervised ℓ_1 loss

$$\mathcal{L}_{1, N_{\text{train}}}(\theta) \triangleq \mathbb{E}_{\mathbf{x}, \mathbf{z}_1, \dots, \mathbf{z}_{N_{\text{train}}}, \mathbf{y}} \{\|\mathbf{x} - \hat{\mathbf{x}}_{(N_{\text{train}})}\|_1\}, \quad (8)$$

with $\hat{\mathbf{x}}_{(N_{\text{train}})} = \sum_{i=1}^{N_{\text{train}}} \frac{G_\theta(\mathbf{z}_i, \mathbf{y})}{N_{\text{train}}}$ the N_{train} -sample average and $\hat{\mathbf{x}}_i \triangleq G_\theta(\mathbf{z}_i, \mathbf{y})$ the generated samples.

The $\mathcal{L}_{\text{SD}, N_{\text{train}}}(\theta)$ term is a standard deviation reward that encourages the samples to deviate by looking at the ℓ_1 distance between the individual samples and the N_{train} -sample average

$$\mathcal{L}_{\text{SD}, N_{\text{train}}}(\theta) \triangleq \sqrt{\frac{\pi}{2N_{\text{train}}(N_{\text{train}} - 1)}} \sum_{i=1}^{N_{\text{train}}} \mathbb{E}_{\mathbf{z}_1, \dots, \mathbf{z}_{N_{\text{train}}}, \mathbf{y}} \{\|\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_{(N_{\text{train}})}\|_1\}. \quad (9)$$

In the work of [Bendel et al. \(2023\)](#), the authors show that this loss function results in generated samples that follow the true posterior distribution in both mean and covariance if the elements of $\hat{\mathbf{x}}_i$ and $\mathbf{x}$ are independent Gaussian distributions conditioned on $\mathbf{y}$ ([Bendel et al. 2023](#), §Prop. 3.1):

$$\mathbb{E}_{\mathbf{z}_i|\mathbf{y}}\{\hat{\mathbf{x}}_i(\theta^*)|\mathbf{y}\} = \mathbb{E}_{\mathbf{x}|\mathbf{y}}\{\mathbf{x}|\mathbf{y}\} = \hat{\mathbf{x}}_{\text{mmse}} \quad (10)$$

$$\text{Cov}_{\mathbf{z}_i|\mathbf{y}}\{\hat{\mathbf{x}}_i(\theta^*)|\mathbf{y}\} = \text{Cov}_{\mathbf{x}|\mathbf{y}}\{\mathbf{x}|\mathbf{y}\} \quad (11)$$

In order to ensure that this standard deviation reward term results in

generated samples following the true posterior distribution in both mean and covariance, the hyperparameter β_{SD} needs to be tuned carefully, as choosing a larger value of β_{SD} tends to result in samples with larger variance.

To automatically tune the β_{SD} reward weight, Bendel et al. (2023) propose a method that relates the error of a single generated sample to that of the error on the N -average of the samples. The authors show that for independent samples of the true posterior these errors should follow the relation

$$\frac{\mathcal{E}_1}{\mathcal{E}_N} = \frac{2N}{N+1}, \quad (12)$$

where the errors are calculated over the samples of the validation set as

$$\begin{aligned} \widehat{\mathcal{E}}_N &= \frac{1}{N_{\text{val}}} \sum_{i=1}^{N_{\text{val}}} \left\| \mathbf{x}_i - \frac{1}{N_{\text{val}}} \sum_{j=1}^{N_{\text{val}}} \mathbf{G}_{\theta}(\mathbf{z}_j, \mathbf{y}_i) \right\|_1, \\ \widehat{\mathcal{E}}_1 &= \frac{1}{N_{\text{val}}} \sum_{i=1}^{N_{\text{val}}} \left\| \mathbf{x}_i - \mathbf{G}_{\theta}(\mathbf{z}_1, \mathbf{y}_i) \right\|_1. \end{aligned} \quad (13)$$

The authors then use a gradient descent method to update the β_{SD} hyperparameter using

$$\beta_{SD,t+1} = \beta_{SD,t} - \mu_{SD} \cdot \left(\log_{10} \left[\frac{\widehat{\mathcal{E}}_1}{\widehat{\mathcal{E}}_N} \right] - \log_{10} \left[\frac{2N}{N+1} \right] \right) \cdot \beta_{SD,0}, \quad (14)$$

with some $\mu_{SD} > 0$ as the learning rate. This update is performed after every epoch using the validation set to ensure that the generated samples follow the true posterior in covariance. For a more in-depth discussion on the tuning of the hyperparameters and proof of this proposition we refer the reader to Bendel et al. (2023).

4 GENERATIVE RADIO INTERFEROMETRIC IMAGING

In this work we build on the rcGAN framework of Bendel et al. (2023), adapt it to the radio interferometry problem and extend it by explicitly including the measurement operator in the generator network. A schematic of our RI-GAN framework is shown in Figure 1.

In order to provide a model that addresses the scalability problem for observations with large telescopes like the SKA, the models work using just the dirty image and PSF of the observation in order to reduce the number of evaluations of the measurement operator. These need to only be computed once per observation, which is significantly less than the number of evaluations needed for iterative optimisation approaches.

4.1 Generator

The generator network is based on the U-Net architecture (Ronneberger et al. 2015) which has been shown to be effective in learned post-processing settings in radio interferometric imaging (e.g. Al-lam Jr 2016; Terris et al. 2019; Mars et al. 2023, 2025) and has been used extensively for GANs. The input of the generator consists of several image channels:

- (i) The dirty image, $\mathbf{x}_{\text{dirty}} = \Phi^\dagger \mathbf{y}$, allowing the generator to condition on the observed data.
- (ii) The PSF of the telescope, which is used to condition the generator on the measurement operator that has a different PSF for each uv -coverage.
- (iii) A latent code image of random noise. By varying this latent code image, we generate different posterior samples.

The output of the generator is a sample from the approximate posterior distribution. The U-Net architecture consists of four layers with strided convolutions to downsample from each encoder layer to the next and transposed convolutions to upsample from each decoder layer to the next. Each encoder and decoder layer consists of residual convolutional blocks, which are based on the implementation in Adler & Öktem (2025), with skip connections from the encoder to the decoder layers. A more detailed description of the architecture can be found in the paper of the rcGAN framework (Bendel et al. 2023).

To enhance the robustness and quality of the reconstruction, we incorporate the explicit use of the measurement operator into the generator network. For this, we consider a GU-Net (Mars et al. 2023, 2025) architecture, which integrates the gradient of the data-fidelity likelihood, $\mathcal{L} = \|\mathbf{y} - \Phi \mathbf{x}\|_{\ell_2}^2$, into the network. At several layers of the U-Net architecture, we compute the gradient of the data-fidelity likelihood with respect to the image $\mathbf{x}$ and combine it with the rest of the channels using convolutional layers. This gradient of the data-fidelity likelihood is given by

$$\nabla_{\mathbf{x}} \mathcal{L} \propto \Phi^*(\mathbf{y} - \Phi \mathbf{x}). \quad (15)$$

Given the computational cost of the full measurement operator with (de)gridding operations (as well as w-stacking and w-correction operations in case of non-coplanar baselines), we approximate the $\Phi^* \Phi$ operation by a convolution with the PSF of the telescope. This approximation is accurate in the case of coplanar baselines, though it neglects wide-field effects caused by non-coplanar baselines. This approximate gradient is computed at full resolution scale at the first level of the U-Net and lower resolution gradients are applied on the lower levels of the network:

$$\widetilde{\nabla}_{\mathbf{x},i} \mathcal{L} \propto \mathbf{x}_{\text{dirty},i} - \mathbf{x}_i * \mathbf{I}_{\text{PSF},i}, \quad (16)$$

where the subscript i denotes the i -th level in the network. The dirty image $\mathbf{x}_{\text{dirty},i}$ and the PSF $\mathbf{I}_{\text{PSF},i}$ are sub-sampled versions of the full resolution dirty image and PSF, sub-sampled by a factor of 2^i along each dimension to match the resolution of the channels on the i -th level. These (sub-scale) gradient operations are applied at the beginning of each encoder and decoder layer in the network.

4.2 Discriminator

The discriminator network is a standard convolutional neural network (CNN) consisting of one convolutional layer followed by six convolutional downsampling blocks. Each block consists of a 2×2 average pool layer and a convolutional layer with instance normalisation and leaky ReLU activation functions (slope 0.2). The input of the discriminator is either an image from the true distribution or a generated image from the generator network. The output of the discriminator is a single scalar that represents the estimated Wasserstein score for the image.

4.3 Training data and training

To train the generative networks, we compiled a large dataset of 13,000 simulated radio interferometric measurements of 360×360 galaxy images created from the IllustrisTNG simulations (Nelson et al. 2019). The images are created by binning and in-painting the image from the simulated particles, as described in more detail in Mars et al. (2023). The measurements are simulated for different uv -coverages of the MeerKAT telescope with a total of 241,920 visibilities, where we vary the pointing of the telescope. However, we

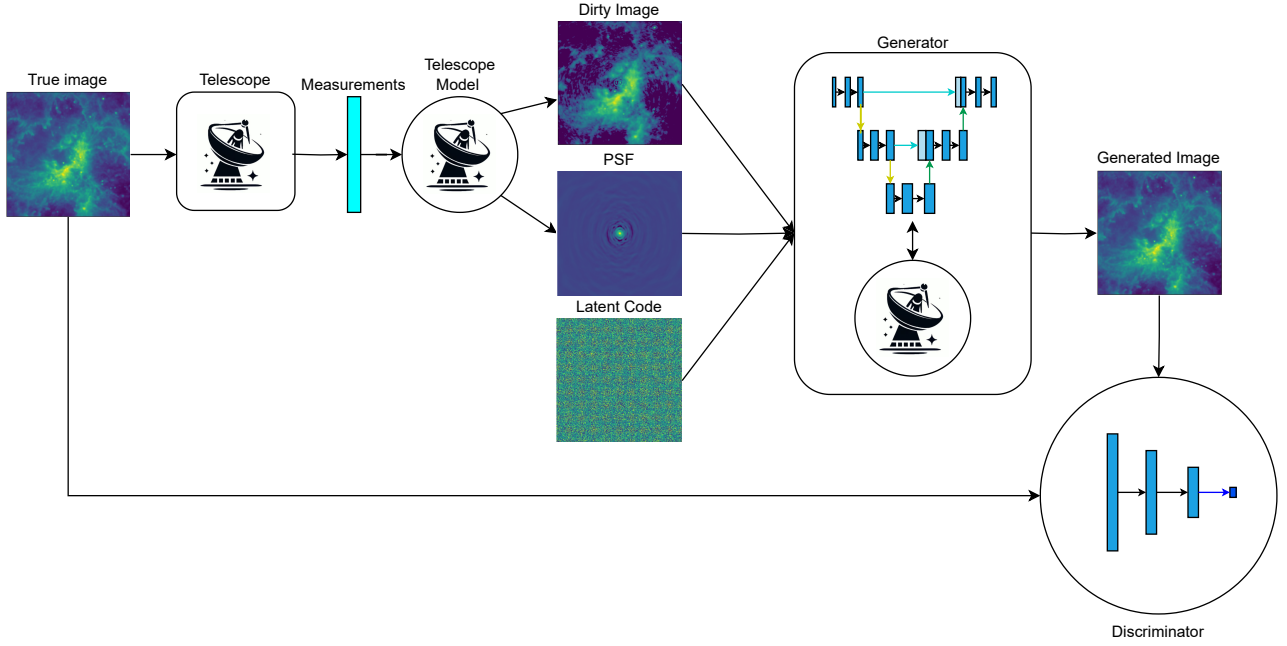


Figure 1. Our RI-GAN framework consisting of the generator and discriminator networks. From the true image we simulate measurements from which we can create the dirty image and PSF for an observation. These two images, combined with a latent code image, are used as input to the generator network. The generator network outputs an approximate sample from the posterior distribution. The discriminator gets passed either a generated image or a true image and outputs a scalar.

assume co-planar baselines and neglect the w -components of the baselines. The added noise is calculated according to the Brown Equation (Brown 1998) as implemented in the casatools simulation module that is part of CASA (Team et al. 2022). From the contaminated measurements, we create the dirty images and PSFs using Briggs weighting (using a robustness parameter of 0.5). For training, we use 10,000 images, with 1,500 images for validation and testing each.

The RI-GANs are trained for 100 epochs on 4 NVIDIA A100 GPUs with 80Gb VRAM each, using an effective batch size of 16; training takes about 48 hours.

4.4 Point estimate and uncertainty estimation

The trained generator network can be used to generate approximate posterior samples when given the dirty image and PSF of an observation. From these samples, we compute a point estimate by taking the mean over N samples, where N is the number of samples generated. In order to estimate the uncertainty of the point estimate, we compute the standard deviation of the samples. These choices are motivated by proposition 3.1 in Bendel et al. (2023), summarised in Equations 10 and 11. In order to reduce the training and validation time we use $N_{\text{train}} = 2$ and $N_{\text{val}} = 8$ samples for training and validation respectively and found no significant drop in training performance compared to training using a larger number of samples. During evaluations we want to generate a larger number of samples in order to get a better estimate of the reconstruction and its uncertainty, so we use $N_{\text{eval}} = 32$ samples for the evaluation of the model. Note that for creating a reconstruction of a single observation, these samples can be generated in parallel on one or a few GPUs, which makes the generation of samples for a single observation very fast.

5 IMAGE RECONSTRUCTION WITH RI-GAN

The performance of the U-Net and GU-Net RI-GAN models is assessed based on reconstruction quality, uncertainty estimation, and computational efficiency. We evaluate these models using the same simulated data as described in the previous section and further validate them using simulated radio interferometric measurements from the 30 Doradus region. On the 30 Doradus region we also compare our reconstructions to the CLEAN algorithm (Offringa et al. 2014).

The quality of the image reconstructions is quantified using the signal-to-noise ratio:

$$\text{SNR}(\mathbf{x}_{\text{true}}, \mathbf{x}_{\text{pred}}) = 20 \cdot \log_{10} \left(\frac{\|\mathbf{x}_{\text{true}}\|_2}{\|\mathbf{x}_{\text{true}} - \mathbf{x}_{\text{pred}}\|_2} \right). \quad (17)$$

We also quantify the correlation of the uncertainty estimate with the absolute error of the reconstruction using the Pearson correlation coefficient:

$$\text{corr}(\mathbf{x}_{\text{uncertainty}}, |\mathbf{x}_{\text{true}} - \mathbf{x}_{\text{pred}}|) = \frac{\text{cov}(\mathbf{x}_{\text{uncertainty}}, |\mathbf{x}_{\text{true}} - \mathbf{x}_{\text{pred}}|)}{\sigma_{\mathbf{x}_{\text{uncertainty}}} \sigma_{|\mathbf{x}_{\text{true}} - \mathbf{x}_{\text{pred}}|}}, \quad (18)$$

where cov is the covariance, σ is the standard deviation and $\mathbf{x}_{\text{uncertainty}}$ is the uncertainty estimate of the reconstruction.

5.1 Reconstruction quality

Figure 2 shows reconstructions using both the U-Net and GU-Net RI-GANs. The reconstructions are obtained by generating 32 posterior samples, with the mean representing the reconstruction and the standard deviation of these samples representing the reconstruction uncertainty. Both models reconstruct the images at high quality and

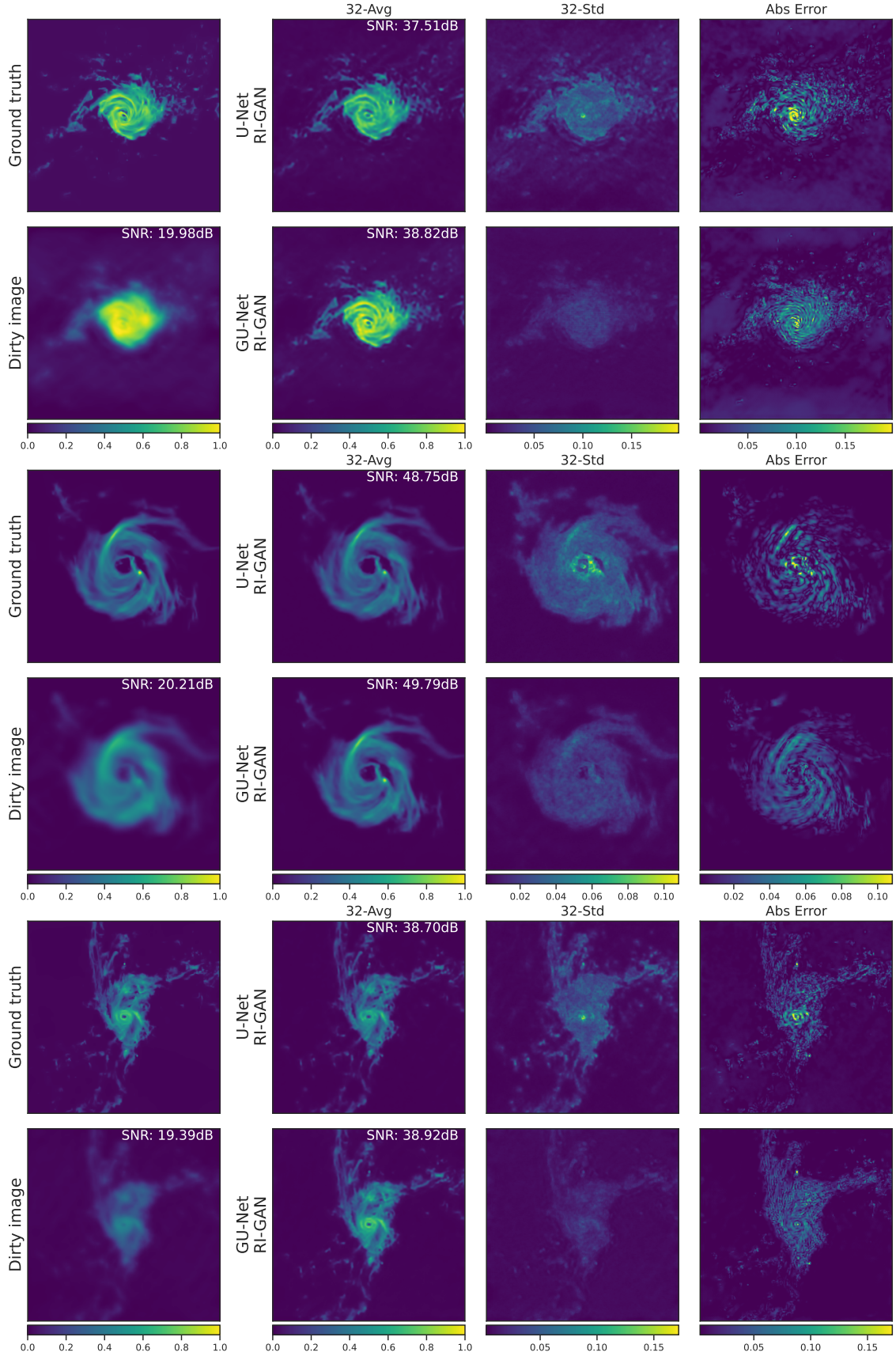


Figure 2. Reconstructions using U-Net and GU-Net RI-GANs from simulated measurements of galaxies in test set. The reconstructions are calculated by averaging over 32 generated approximate posterior samples and the uncertainties are obtained by taking the standard deviation over these samples. Both models provide a large improvement over the dirty image. The GU-Net, which includes updates using an approximate gradient of the data fidelity term, provides higher quality reconstructions and lower uncertainties overall compared to the U-Net.

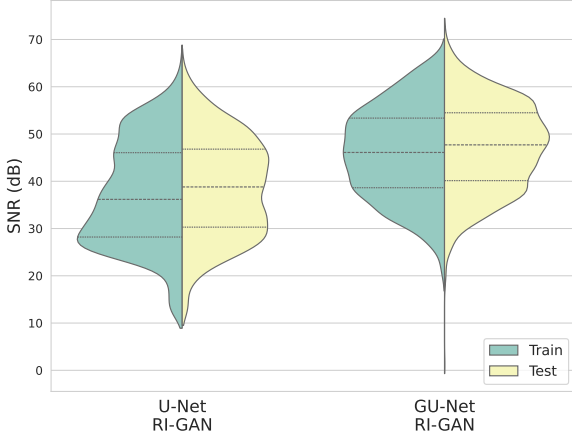


Figure 3. Distribution of the SNR for reconstructions using the U-Net and GU-Net RI-GANs for both the training and test set. The reconstructions are made by averaging over 32 generated posterior samples. Over both the train and test set, the GU-Net provides reconstructions with higher SNR compared to the U-Net. Besides that, the performance between the training and test set is similar for the GU-Net RI-GAN indicating that there is no overfitting to the train set. The dashed and dotted lines indicate the mean and quartiles of the distributions respectively.

the absolute error of the reconstructions visually correlates with the estimated uncertainty. The GU-Net consistently provides a higher average reconstruction quality and lower uncertainty compared to the U-Net. This trend is also evident in the SNR distribution across the dataset, as shown in Figure 3.

5.2 Number of samples

The number of samples used to calculate the mean of the reconstructions affects the quality of the reconstructions and the uncertainties. During training and validation, we are able to use a smaller number of samples to speed up the training process, yet for the final reconstructions we want to use a larger number of samples to get the best estimate of the reconstruction and its uncertainty. Figure 4 shows the effect of increasing the number of samples for a particular reconstruction. As expected, increasing the number of samples used for the average improves the SNR of the reconstruction. This effect plateaus after 32 samples, indicating that the quality of the reconstruction does not significantly improve beyond this point. Additionally, the correlation between the estimated uncertainty also increases as a function of the number of samples, saturating around 32 samples at 0.58 for the U-Net and 0.69 for the GU-Net generators demonstrating that the uncertainty is a reasonable indicator of the error of the reconstruction. The GU-Net reconstructions outperform the U-Net in both the SNR and the uncertainty correlation for any number of samples.

The observation that for a lower number of samples both the SNR and the uncertainty correlation are significantly lower than for higher number of samples indicates that there is significant diversity in the reconstructions, suggesting that we do not suffer from mode collapse.

5.3 Reconstruction of 30 Doradus

To test the generators on out-of-distribution and more realistic data, we use an image of the 30 Doradus region, which has a larger dynamic range (~ 600) than the simulated galaxies on which the models are trained (though not as high a dynamic range as real radio images which can be in the order of 10^6). The measurements are simulated

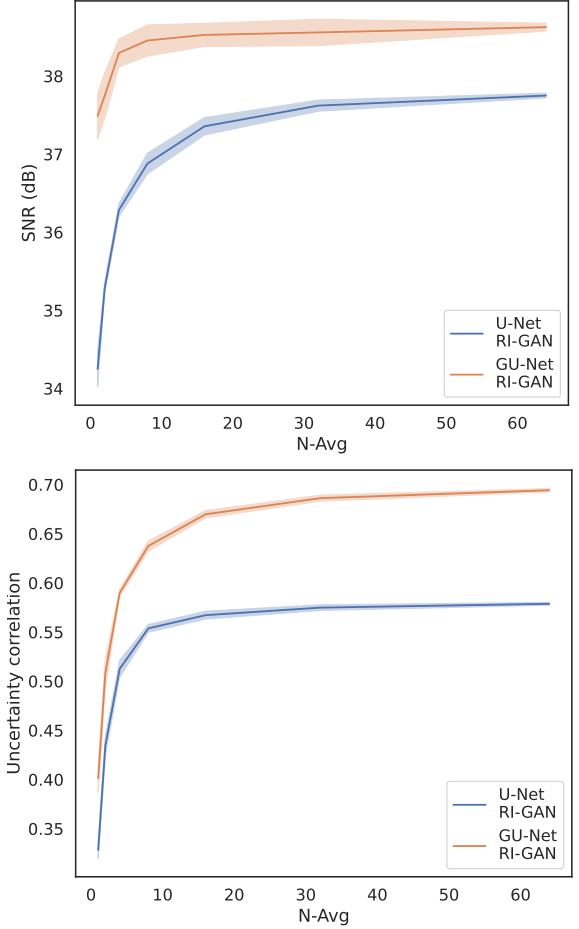


Figure 4. Effect of the number of samples on reconstruction quality and the correlation between the uncertainty and the absolute error for both the U-Net and the GU-Net RI-GANs for an example image. The shaded area represents the standard deviation of repeating the generation 100 times. Both the quality of the reconstruction as quantified by the SNR and the correlation of the uncertainty with the absolute error increase as a function of the number of samples. The GU-Net consistently outperforms the U-Net on both metrics.

similarly to the training data but with a different, unseen uv -coverage. Figures 5 and 6 show the reconstructions on both a linear scale and a logarithmic scale to better showcase the performance at a higher dynamic range. The GU-Net, which includes explicit evaluations of the measurement operator, improves reconstruction quality (SNR: 46.03dB) compared to the U-Net (SNR: 30.80dB), and the CLEAN algorithm (SNR: 32.55dB). The increased performance of the GU-Net is due to the explicit inclusion of the measurement operator in the generator, making it more robust to changes in the uv -coverage and out-of-distribution data of a larger dynamic range. Also, note that the neural networks are trained on simulated data of a lower dynamic range and we expect that the performance of the GU-Net would improve further if it were trained on data more representative of the 30 Doradus region.

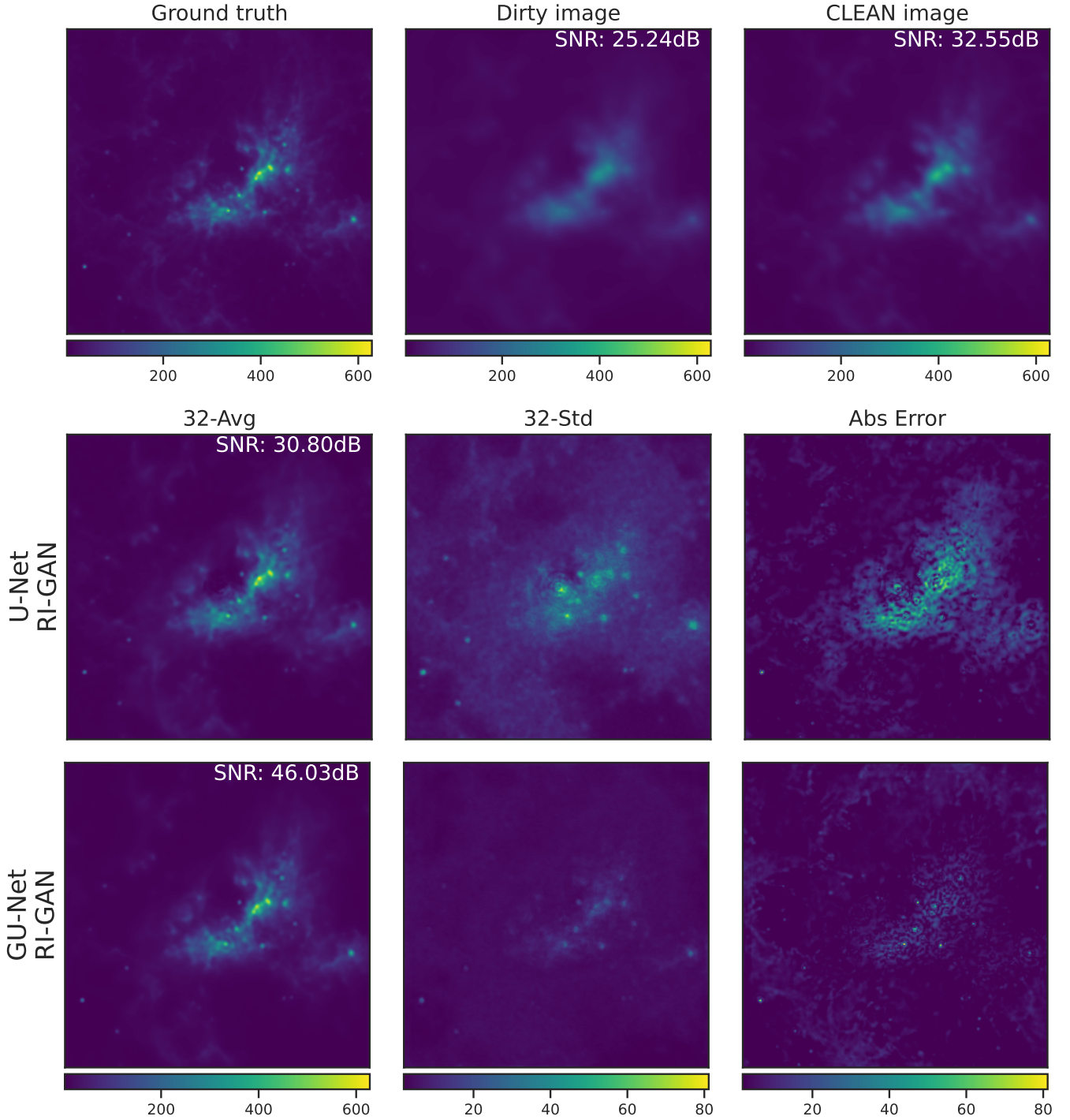


Figure 5. Reconstructions using U-Net and GU-Net RI-GANs from simulated measurements of an out-of-distribution image of the 30 Doradus region using an unseen uv -coverage. The reconstructions are calculated by averaging over 32 generated approximate posterior samples and the uncertainties are obtained by taking the standard deviation over these samples. The top row shows the ground truth, dirty image and the reconstruction using the CLEAN algorithm. The bottom rows show the reconstructions, uncertainties and absolute error using the U-Net and GU-Net RI-GANs respectively. All images are shown on a linear intensity scale. The GU-Net provides higher quality reconstructions and lower uncertainties compared to the U-Net and the CLEAN algorithm.

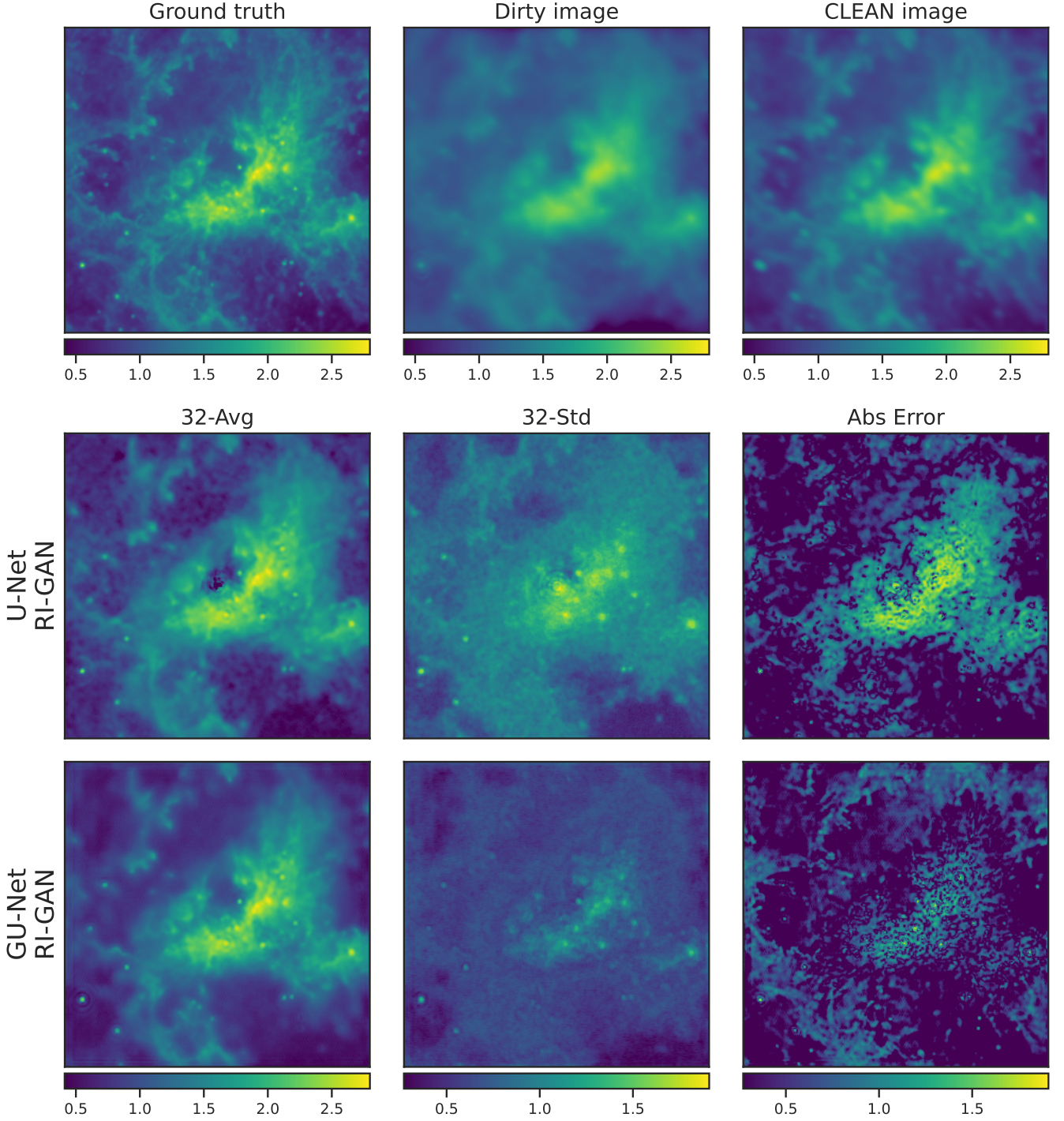


Figure 6. Reconstructions using U-Net and GU-Net RI-GANs from simulated measurements of an out-of-distribution image of the 30 Doradus region using an unseen uv -coverage. The reconstructions are calculated by averaging over 32 generated approximate posterior samples and the uncertainties are obtained by taking the standard deviation over these samples. The top row shows the ground truth, dirty image and the reconstruction using the CLEAN algorithm. The bottom rows show the reconstructions, uncertainties and absolute error using the U-Net and GU-Net RI-GANs respectively. All images are shown on a logarithmic intensity scale. The GU-Net provides higher quality reconstructions and lower uncertainties compared to the U-Net and higher resolution compared to the CLEAN algorithm.

6 CONCLUSION

The use of generative models for interferometric imaging is a promising approach to provide fast image reconstructions of high quality as well as approximate uncertainty quantification, while being robust to different uv -coverages. In this work we proposed our RI-GAN framework which builds on the rcGAN framework (Bendel et al. 2023). These generative frameworks use a standard deviation regularisation term and automatic tuning of the regularisation strength to have the generated posterior samples follow the true posterior distribution in both mean and covariance. For our framework we explore both the use of a U-Net and a GU-Net generator (Mars et al. 2023, 2025). The GU-Net generator is a modified U-Net generator, which includes explicit use of the measurement operator, by including at several points in the network an approximate gradient of an L_2 -norm data fidelity term and using a convolution of the images with the point spread function (PSF). Adding the measurement operator explicitly in the reconstruction network has been shown to improve the quality of the reconstructions as well as provide robustness to varying visibility coverages (Mars et al. 2025).

In our evaluation of the two models on measurements simulated from galaxies in the IllustrisTNG simulations, we find that the GU-Net RI-GAN consistently outperforms its U-Net counterpart in terms of image quality and shows a lower estimated uncertainty. When looking at the number of generated posterior samples needed for reconstruction and uncertainty estimation, we find that the quality of the reconstructions and the correlation of the uncertainties with the errors of the reconstructions increases with the number of samples, but saturates at approximately 32 samples. We also find that the inclusion of the measurement information explicitly in the generator increases the quality of the reconstruction and the uncertainties. For the test example we see that the correlation of the estimated uncertainty and the absolute error of the reconstruction is 0.69 for the GU-Net and 0.58 for the U-Net RI-GANs, showing that the uncertainties are representative of the actual errors.

The models are also evaluated on an out-of-distribution image from the 30 Doradus region which has a larger dynamic range (~ 600). In this example we see that the GU-Net RI-GAN provides significantly better reconstructions and is better at reconstructing the larger dynamic range than the U-Net variant. Additionally, the GU-Net provides lower estimated uncertainties. Including the measurement information explicitly in the generator clearly also improves the ability of the model to generalise to out-of-distribution data. The GU-Net RI-GAN also produces higher quality reconstructions than the CLEAN algorithm, which is used as a baseline for comparison. We expect that additional improvements in the reconstruction quality and uncertainty estimates can be achieved by training on more representative data, such as images with a larger dynamic range, and by including bright point sources in the training data.

In conclusion, we presented two generative models built on our RI-GAN framework for interferometric image reconstruction. Overall the GU-Net RI-GAN, which includes the measurement operator explicitly in the generator, provides better reconstructions and lower uncertainties than the U-Net variant and generalises better to out-of-distribution data. Using this approach we can provide fast and high-quality image reconstructions of interferometric measurements, while also providing uncertainty estimates that correlate well with the actual errors of the reconstructions.

Further work is needed to perform a more in depth analysis of the statistical significance of the uncertainties to verify whether the generated samples follow the true posterior distribution. One method

to do this would be to use estimate the coverage probabilities using the generated samples (Hermans et al. 2022; Lemos et al. 2023).

Additionally, more work is needed to evaluate and adapt these models to real radio observations. This includes the inclusion of non-coplanar baselines in simulation, addition of bright point sources in the training data, application to larger image sizes and evaluation on images that have a dynamic range more representative of true radio observations.

ACKNOWLEDGEMENTS

We thank Tariq Blecher for help generating MeerKAT visibility coverages. The authors would like to acknowledge the use of OpenAI's ChatGPT for providing stylistic suggestions but take responsibility for the accuracy of this manuscript. Matthijs Mars is supported by the UCL Centre for Doctoral Training in Data Intensive Science (STFC Training grant ST/P006736/1). This work is also partially supported by EPSRC and STFC (grant numbers EP/T517793/1; EP/W007673/1; ST/W001136/1). This work also used computing equipment funded by the Research Capital Investment Fund (RCIF) provided by UKRI, and partially funded by the UCL Cosmoparticle Initiative.

DATA AVAILABILITY

The code used to generate the results in this paper can be found at <https://github.com/astro-informatics/rcGAN>. The trained final models and all necessary data to run them are stored at <https://zenodo.org/records/16529320>. The data used in this paper is available upon reasonable request to the corresponding author.

REFERENCES

- Adler J., Öktem O., 2025, in Bubba T. A., ed., , Data-Driven Models in Inverse Problems. De Gruyter, Berlin, Boston, pp 359–412, doi:doi:10.1515/9783111251233-011
- Aghabiglou A., Chu C. S., Dabbech A., Wiaux Y., 2024, *The Astrophysical Journal Supplement Series*, 273, 3
- Allam Jr T., 2016, MSc. Thesis, University College London
- Arduzzone L., Lüth C., Kruse J., Rother C., Köthe U., 2019, Guided Image Generation with Conditional Invertible Neural Networks ([arXiv:1907.02392](https://arxiv.org/abs/1907.02392)), doi:10.48550/arXiv.1907.02392
- Arjovsky M., Bottou L., 2017, in International Conference on Learning Representations.
- Arjovsky M., Chintala S., Bottou L., 2017, in Proceedings of the 34th International Conference on Machine Learning. PMLR, pp 214–223
- Arras P., Knollrüller J., Junklewitz H., Enßlin T. A., 2018, in 2018 26th European Signal Processing Conference (EUSIPCO). pp 2683–2687, doi:10.23919/EUSIPCO.2018.8553533
- Bendel M., Ahmad R., Schniter P., 2023, *Advances in Neural Information Processing Systems*, 36, 68673
- Bhatnagar S., Cornwell T. J., 2004, *Astronomy & Astrophysics*, 426, 747
- Bhatnagar S., Cornwell T. J., Golap K., Uson J. M., 2008, *Astronomy & Astrophysics*, 487, 419
- Brown R. L., 1998, in Advanced Technology MMW, Radio, and Terahertz Telescopes. SPIE, pp 231–237, doi:10.1117/12.317357
- Cai X., Pereyra M., McEwen J. D., 2018a, *Monthly Notices of the Royal Astronomical Society*, 480, 4154
- Cai X., Pereyra M., McEwen J. D., 2018b, *Monthly Notices of the Royal Astronomical Society*, 480, 4170
- Carrillo R. E., McEwen J. D., Wiaux Y., 2012, *Monthly Notices of the Royal Astronomical Society*, 426, 1223

- Carrillo R. E., McEwen J. D., Van De Ville D., Thiran J.-P., Wiaux Y., 2013, *IEEE Signal Processing Letters*, 20, 591
- Cherif M., Liaudat T. I., Kern J., Kervazo C., Bobin J., 2024, Uncertainty Quantification for Fast Reconstruction Methods Using Augmented Equivariant Bootstrap: Application to Radio Interferometry ([arXiv:2410.23178](#)), doi:10.48550/arXiv.2410.23178
- Connor L., Bouman K. L., Ravi V., Hallinan G., 2022, *Monthly Notices of the Royal Astronomical Society*, 514, 2614
- Dabbech A., Aghabiglou A., Chu C. S., Wiaux Y., 2023, CLEANing Cygnus A Deep and Fast with R2D2, <https://arxiv.org/abs/2309.03291v3>
- Dewdney P., Hall P., Schilizzi R., Lazio T., 2009, *Proceedings of the IEEE*, 97, 1482
- Dia N., Yantovski-Barth M. J., Adam A., Bowles M., Perreault-Levasseur L., Hezaveh Y., Scaife A., 2025, IRIS: A Bayesian Approach for Image Reconstruction in Radio Interferometry with Expressive Score-Based Priors ([arXiv:2501.02473](#)), doi:10.48550/arXiv.2501.02473
- Drozдова M., et al., 2024, *Astronomy & Astrophysics*, 683, A105
- Duijndam A., Schonewille M., 1997, in SEG Technical Program Expanded Abstracts 1997. Society of Exploration Geophysicists, pp 1135–1138, doi:10.1190/1.1885591
- Dutt A., Rokhlin V., 1993, *SIAM Journal on Scientific Computing*, 14, 1368
- Edupuganti V., Mardani M., Vasanawala S., Pauly J., 2021, *IEEE Transactions on Medical Imaging*, 40, 239
- Fessler J., Sutton B., 2003, *IEEE Transactions on Signal Processing*, 51, 560
- Goodfellow I. J., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y., 2014, Generative Adversarial Networks ([arXiv:1406.2661](#)), doi:10.48550/arXiv.1406.2661
- Greiner M., Vacca V., Junklewitz H., Enßlin T. A., 2017, fastRESOLVE: Fast Bayesian Imaging for Aperture Synthesis in Radio Astronomy ([arXiv:1605.04317](#)), doi:10.48550/arXiv.1605.04317
- Gulrajani I., Ahmed F., Arjovsky M., Dumoulin V., Courville A. C., 2017, in Advances in Neural Information Processing Systems. Curran Associates, Inc.
- Hermans J., Delaunoy A., Rozet F., Wehenkel A., Begy V., Louppe G., 2022, Transactions on Machine Learning Research
- Högbom J. A., 1974, *Astronomy and Astrophysics*, 15, 55
- Isola P., Zhu J.-Y., Zhou T., Efros A. A., 2017, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp 1125–1134
- Junklewitz H., Bell M. R., Selig M., Enßlin T. A., 2016, *Astronomy & Astrophysics*, 586, A76
- Knollmüller J., Enßlin T. A., 2020, Metric Gaussian Variational Inference ([arXiv:1901.11033](#)), doi:10.48550/arXiv.1901.11033
- Lemos P., Coogan A., Hezaveh Y., Perreault-Levasseur L., 2023, in Proceedings of the 40th International Conference on Machine Learning. PMLR, pp 19256–19273
- Liaudat T. I., Mars M., Price M. A., Pereyra M., Betcke M. M., McEwen J. D., 2024, RAS Techniques and Instruments, 3, 505
- Mars M., Betcke M. M., McEwen J. D., 2023, *RAS Techniques and Instruments*, 2, 760
- Mars M., Betcke M. M., McEwen J. D., 2025, *RAS Techniques and Instruments*, 4, rzaf025
- Nelson D., et al., 2019, *Computational Astrophysics and Cosmology*, 6, 2
- Offringa A. R., et al., 2014, *Monthly Notices of the Royal Astronomical Society*, 444, 606
- Pratley L., McEwen J. D., d’Avezac M., Carrillo R. E., Onose A., Wiaux Y., 2018, *Monthly Notices of the Royal Astronomical Society*, 473, 1038
- Pratley L., McEwen J. D., d’Avezac M., Cai X., Perez-Suarez D., Christidi I., Guichard R., 2019, Distributed and Parallel Sparse Convex Optimization for Radio Interferometry with PURIFY ([arXiv:1903.04502](#)), doi:10.48550/arXiv.1903.04502
- Remy B., Lanasse F., Jeffrey N., Liu J., Starck J.-L., Osato K., Schrabback T., 2023, *Astronomy & Astrophysics*, 672, A51
- Ronneberger O., Fischer P., Brox T., 2015, in Navab N., Hornegger J., Wells W. M., Frangi A. F., eds., Vol. 9351, Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Springer International Publishing, Cham, pp 234–241, doi:10.1007/978-3-319-24574-4_28
- Smirnov O. M., 2011, *Astronomy & Astrophysics*, 527, A107
- Sohn K., Lee H., Yan X., 2015, in Advances in Neural Information Processing Systems. Curran Associates, Inc.
- Stewart I. M., Fenech D. M., Muxlow T. W. B., 2011, *Astronomy & Astrophysics*, 535, A81
- Sun H., Bouman K. L., 2021, *Proceedings of the AAAI Conference on Artificial Intelligence*, 35, 2628
- Sutter P. M., et al., 2014, *Monthly Notices of the Royal Astronomical Society*, 438, 768
- Team T. C., et al., 2022, *Publications of the Astronomical Society of the Pacific*, 134, 114501
- Terris M., Abdulaziz A., Dabbech A., Jiang M., Repetti A., Pesquet J.-C., Wiaux Y., 2019, Signal Processing with Adaptive Sparse Structured Representations (SPARS) workshop, p. 3
- Terris M., Dabbech A., Tang C., Wiaux Y., 2022, *Monthly Notices of the Royal Astronomical Society*, 518, 604
- Terris M., Tang C., Jackson A., Wiaux Y., 2023, Plug-and-Play Imaging with Model Uncertainty Quantification in Radio Astronomy ([arXiv:2312.07137](#)), doi:10.48550/arXiv.2312.07137
- Thompson A. R., Moran J. M., Swenson G. W., 2017, Interferometry and Synthesis in Radio Astronomy. Astronomy and Astrophysics Library, Springer International Publishing, Cham, doi:10.1007/978-3-319-44431-4
- Tonolini F., Radford J., Turpin A., Faccio D., Murray-Smith R., 2020, Journal of Machine Learning Research, 21, 1
- Wang R., Chen Z., Luo Q., Wang F., 2023, in , ECAI 2023. IOS Press, pp 2499–2506, doi:10.3233/FAIA230554
- Whitney J., Liaudat T., Price M., Mars M., McEwen J. D., 2024a, Using Conditional GANs for Convergence Map Reconstruction with Uncertainties ([arXiv:2406.15424](#)), doi:10.48550/arXiv.2406.15424
- Whitney J. J., Liaudat T. I., Price M. A., Mars M., McEwen J. D., 2024b, Generative Modelling for Mass-Mapping with Fast Uncertainty Quantification ([arXiv:2410.24197](#)), doi:10.48550/arXiv.2410.24197
- Winkler C., Worrall D., Hoozeboom E., Welling M., 2023, Learning Likelihoods with Conditional Normalizing Flows ([arXiv:1912.00042](#)), doi:10.48550/arXiv.1912.00042
- Zhao H., Li H., Maurer-Stroh S., Cheng L., 2018, *Medical Image Analysis*, 49, 14
- Zhao S., Cui J., Sheng Y., Dong Y., Liang X., Chang E. I.-C., Xu Y., 2021, in International Conference on Learning Representations.

This paper has been typeset from a $\text{\LaTeX}$ file prepared by the author.