CLEANing Cygnus A deep and fast with R2D2

Arwa Dabbech [ID],[1] Amir Aghabiglou [ID],[1] Chung San Chu [ID],[1] and Yves Wiaux [ID][1]

[1]Institute of Sensors, Signals and Systems, Heriot-Watt University, Edinburgh EH14 4AS, United Kingdom

ABSTRACT

A novel deep learning paradigm for synthesis imaging by radio interferometry in astronomy was recently proposed, dubbed "**R**esidual-to-**R**esidual **D**NN series for high-**D**ynamic range imaging" (**R2D2**). In this work, we start by shedding light on R2D2's algorithmic structure, interpreting it as a learned version of CLEAN with minor cycles substituted with a deep neural network (DNN) whose training is iteration-specific. We then proceed with R2D2's first demonstration on real data, for monochromatic intensity imaging of the radio galaxy Cygnus A from S band observations with the Very Large Array (VLA). We show that the modeling power of R2D2's learning approach enables delivering high-precision imaging, superseding the resolution of CLEAN, and matching the precision of modern optimization and plug-and-play algorithms, respectively uSARA and AIRI. Requiring few major-cycle iterations only, R2D2 provides a much faster reconstruction than uSARA and AIRI, known to be highly iterative, and is at least as fast as CLEAN.

1. INTRODUCTION

Modern radio telescopes are able to map the radio sky over large fields of view and wide frequency bandwidths with unprecedented depth and detail, owing to the sheer amounts of the data they acquire. Leveraging these capabilities in image formation raises significant data-processing challenges. The underlying radio-interferometric (RI) inverse problem calls for efficient imaging algorithms able to both deliver high-precision reconstruction via tailored regularization models, and scale to large data volumes and image dimensions. Since its inception by Högbom (1974), the CLEAN paradigm has been the standard for RI imaging owing to its simplicity and computational efficiency. In essence, CLEAN is a Matching Pursuit (MP) algorithm (Lannes et al. 1997), iteratively identifying model components by projecting the residual dirty image onto a sparsity dictionary, which in Högbom's version (hereafter Hö-CLEAN) is the identity basis. Its reconstruction is obtained by smoothing the identified model components via a restoring beam, and is complemented by the addition of the

Corresponding author: Yves Wiaux
y.wiaux@hw.ac.uk

residual dirty image. Variants of CLEAN have been devised over five decades to overcome its shortfalls in terms of precision, stability, and scalability (e.g., Clark 1980; Wakker & Schwarz 1988; Bhatnagar & Cornwell 2004; Cornwell 2008). The Cotton-Schwab approach (CS-CLEAN; Schwab 1984) introduced a nested iterative structure, with an inner loop (minor cycles) identifying model components, whose exact contributions are removed from the RI data all at once at the next iteration of an outer loop (major cycle). The major cycle would typically exhibit few iterations only, conferring to CLEAN its scalability. Multi-scale CLEAN (MS-CLEAN; Cornwell 2008) introduces a bespoke multi-scale basis to substitute the identity basis, improving reconstruction quality. However, CLEAN's restored image is by design restricted to instrumental resolution due to the restoring beam, and limited in dynamic range by the addition of the final residual dirty image.

Numerous algorithms for RI imaging emerged from optimization theory over the past two decades (e.g., Wiaux et al. 2009; Li et al. 2011; Carrillo et al. 2012; Garsden et al. 2015; Dabbech et al. 2015). Backed by compressive sensing theory, these methods inject hand-crafted sparsity-based image models into the RI data, leveraging convergent optimization algorithmic structures. Their iterative structure is different to CLEAN's,

but shares a two-step iteration process, alternating a data-fidelity step involving the computation of data residuals, and a regularization step enforcing the chosen image model. Their latest evolution, uSARA, is underpinned by the Forward-Backward structure, which promotes an advanced handcrafted sparsity-based image model via a so-called "proximal regularization operator" (Repetti & Wiaux 2020; Terris et al. 2022). uSARA has been demonstrated on real gigabyte-scale RI data (Dabbech et al. 2022; Wilber et al. 2023a). Nonetheless, the complexity and the highly-iterative nature of these methods have hindered their adoption by the wider radio astronomy community.

Novel algorithms for computational imaging empowered by deep learning have emerged across a wide range of applications, including radio astronomy. Purely data-driven end-to-end DNNs can provide real-time solutions to the RI inverse problem (Connor et al. 2022; Schmidt et al. 2022). However, they raise important interpretability and reliability concerns as standard architectures cannot ensure consistency of the recovered image with the observed data. Alternative Plug-and-Play (PnP) approaches, interfacing optimization theory and deep learning, can circumvent these concerns by implicitly injecting a learned image model into the data via pre-trained denoising DNNs (see Kamilov et al. 2023, for a review). One such example is AIRI (Terris et al. 2022), which is underpinned by the same algorithmic structure as uSARA, with its regularization operator substituted with a denoising DNN. AIRI has been demonstrated to deliver slightly superior imaging precision than uSARA, while exhibiting a lower computational cost thanks to its DNN-encapsulated regularization process (Dabbech et al. 2022; Wilber et al. 2023b). However, PnP algorithms still share the highly iterative nature and associated limited scalability of their optimization counterparts. Unrolled DNNs represent another deep learning approach, consisting in learning at once a finite number of iterations of an optimization structure (see Monga et al. 2021, for a review). They offer more interpretability than purely data-driven DNNs thanks to the integration of data consistency layers. However, embedding measurement operators in the network architecture also becomes impractical at large scale, primarily for training, but also for image reconstruction.

Most recently, we have proposed a novel deep learning approach dubbed R2D2, standing for **R**esidual-to-**R**esidual **D**NN series for high-**D**ynamic range imaging (Aghabiglou et al. 2024). R2D2's reconstruction is formed as a series of residual images, iteratively estimated as outputs of DNNs taking the previous iter-

ation's image estimate and associated back-projected data residual as inputs.

In this work, we start by shedding light on R2D2's algorithmic structure, interpreting it as a learned version of CLEAN. Three R2D2 variants have been introduced and extensively studied in simulation, namely R2D2, R2D2-Net, and R3D3. We analyze their positions in the landscape of CLEAN variants, with specific comparison to Hö-CLEAN, CS-CLEAN, and MS-CLEAN. We then demonstrate the R2D2 paradigm on high-sensitivity VLA observations of the radio galaxy Cygnus A at S band, both in terms of precision and computational speed. Firstly, the reconstruction quality of all R2D2 variants at least equates that of AIRI and uSARA, outperforming all CLEAN variants. Secondly, requiring few major-cycle iterations only (on par with CLEAN), all R2D2 variants provide a much faster reconstruction than uSARA and AIRI, and at least as fast as CLEAN.

The remainder of this paper is structured as follows. Section 2 introduces the RI inverse problem. Section 3 provides a summary of the R2D2 paradigm in its three variants. Section 4 presents a CLEAN perspective on the R2D2 paradigm. Application of telescope-specific incarnations of the R2D2 variants is provided in Section 5, for the formation of Cygnus A images from VLA observations. Conclusions are presented in Section 6.

2. RI INVERSE PROBLEM

Under the assumptions of non-polarized monochromatic radio emission spanning a narrow field of view, RI data, also termed visibilities, are noisy Fourier measurements of the intensity image of interest. The sampled Fourier coverage is defined by the array antenna configuration and the specifications of the observation (e.g., direction of sight, total observation duration). Let $\boldsymbol{x}^{\star} \in \mathbb{R}_{+}^{N}$ denote a discrete representation of the unknown radio image. A discrete model of the RI data, denoted by $\boldsymbol{y} \in \mathbb{C}^{M}$, reads $\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{x}^{\star} + \boldsymbol{n}$, where $\boldsymbol{n} \in \mathbb{C}^{M}$ is an additive white Gaussian noise, with mean zero and standard deviation $\tau > 0$. The RI measurement operator $\boldsymbol{\Phi} \in \mathbb{C}^{M \times N}$ is a Fourier sampling operator, often incorporating a data-weighting scheme to mitigate the non-uniform Fourier sampling and enhance the effective resolution of the data (e.g., Briggs weighting; Briggs 1995). In high-sensitivity regimes, the measurement operator model is more complex, encompassing direction dependent effects (DDEs) induced by atmospheric perturbations and instrumental errors (Smirnov 2011).

The image-domain formulation of the RI inverse problem can be obtained by back-projecting the data $\boldsymbol{y}$ via

$\boldsymbol{\Phi}^\dagger$, the adjoint of the measurement operator, as follows:

$$\boldsymbol{x}_{\mathrm{d}} = \mathbf{D}\boldsymbol{x}^\star + \mathbf{b}, \qquad (1)$$

where $\boldsymbol{x}_{\mathrm{d}} = \kappa\mathrm{Re}\{\boldsymbol{\Phi}^\dagger\boldsymbol{y}\} \in \mathbb{R}^N$ is the normalized back-projected data, often termed the *dirty* image. The normalization factor $\kappa > 0$ ensures that the point spread function (PSF; $\mathbf{h} = \kappa\mathrm{Re}\{\boldsymbol{\Phi}^\dagger\boldsymbol{\Phi}\}\boldsymbol{\delta} \in \mathbb{R}^N$, with $\boldsymbol{\delta} \in \mathbb{R}^N$ standing for the image with value one at its center and zero otherwise) has a peak value equal to one. The linear operator $\mathbf{D} \triangleq \kappa\mathrm{Re}\{\boldsymbol{\Phi}^\dagger\boldsymbol{\Phi}\} \in \mathbb{R}^{N\times N}$ encodes Fourier de-gridding and gridding operations, and maps the image of interest to the dirty image space. Finally, the normalized back-projected noise reads $\mathbf{b} = \kappa\mathrm{Re}\{\boldsymbol{\Phi}^\dagger\boldsymbol{n}\} \in \mathbb{R}^N$.

3. THE R2D2 PARADIGM

3.1. *Algorithmic structure*

The R2D2 algorithm (Aghabiglou et al. 2024) is underpinned by a sequence of DNN modules denoted by $\{\mathbf{N}_{\widehat{\boldsymbol{\theta}}^{(i)}}\}_{1\leq i\leq I}$, which are described by their learned parameters $\{\widehat{\boldsymbol{\theta}}^{(i)} \in \mathbb{R}^Q\}_{1\leq i\leq I}$. With the image estimate being initialized as $\boldsymbol{x}^{(0)} = \mathbf{0} \in \mathbb{R}^N$, at any iteration $i \in \{1,\ldots,I\}$, the network $\mathbf{N}_{\widehat{\boldsymbol{\theta}}^{(i)}}$ takes as input both the previous image estimate $\boldsymbol{x}^{(i-1)}$ and associated residual dirty image $\boldsymbol{r}^{(i-1)}$. The latter is updated from the dirty image by removing the contribution of the previous image estimate: $\boldsymbol{r}^{(i-1)} = \boldsymbol{x}_{\mathrm{d}} - \mathbf{D}\boldsymbol{x}^{(i-1)}$. The current image estimate is updated from the output of the network as:

$$\boldsymbol{x}^{(i)} = \boldsymbol{x}^{(i-1)} + \mathbf{N}_{\widehat{\boldsymbol{\theta}}^{(i)}}(\boldsymbol{r}^{(i-1)}, \boldsymbol{x}^{(i-1)}), \qquad (2)$$

where the predicted residual image $\mathbf{N}_{\widehat{\boldsymbol{\theta}}^{(i)}}(\boldsymbol{r}^{(i-1)}, \boldsymbol{x}^{(i-1)})$ captures emission from the input residual dirty image $\boldsymbol{r}^{(i-1)}$, and corrects for estimation errors in $\boldsymbol{x}^{(i-1)}$, thus progressively enhancing the resolution and dynamic range of the reconstruction. Assuming I DNN modules in the R2D2 sequence, the final reconstruction $\widehat{\boldsymbol{x}}$ takes a simple series expression, as the sum of output residual images from all DNN modules. The iteration structure of R2D2 is reminiscent of MP, but its model components at each iteration are identified in a learned manifold of basis functions encoded by a DNN, rather than a handcrafted sparsity basis fixed across iterations.

3.2. *Training procedure*

R2D2 DNN modules are trained in a sequential manner from a dataset consisting of K image pairs of the ground truth images and their associated dirty images $\{\boldsymbol{x}_k^\star, \boldsymbol{x}_{\mathrm{d}k}\}_{1\leq k\leq K}$. At any iteration $i \geq 1$, the network $\mathbf{N}_{\widehat{\boldsymbol{\theta}}^{(i)}}$ is trained taking the previous image estimates and associated residual dirty images $\{\boldsymbol{x}_k^{(i-1)}, \boldsymbol{r}_k^{(i-1)}\}_{1\leq k\leq K}$ as input, with $\{\boldsymbol{x}_k^{(0)} = \mathbf{0}\}_{1\leq k\leq K}$

and $\{\boldsymbol{r}_k^{(0)} = \boldsymbol{x}_{\mathrm{d}k}\}_{1\leq k\leq K}$. Its learned parameter vector $\widehat{\boldsymbol{\theta}}^{(i)}$ is minimizer of the loss function:

$$\min_{\boldsymbol{\theta}^{(i)}\in\mathbb{R}^Q} \sum_{k=1}^K \|\boldsymbol{x}_k^\star - [\boldsymbol{x}_k^{(i-1)} + \mathbf{N}_{\boldsymbol{\theta}^{(i)}}(\boldsymbol{r}_k^{(i-1)}, \boldsymbol{x}_k^{(i-1)})]_+\|_1, \quad (3)$$

where $\|.\|_1$ stands for the ℓ_1-norm, and $[.]_+$ denotes the projection onto the positive orthant $\mathbb{R}_+^N$, ensuring non-negativity of the image estimate at any point in the iterative sequence. The image estimates and associated residual dirty images $\{\boldsymbol{x}_k^{(i)}, \boldsymbol{r}_k^{(i)}\}_{1\leq k\leq K}$ are updated from the outputs of the trained network. Alongside their corresponding ground truth images, they serve as the training dataset for the subsequent DNN module. Training of the DNN series concludes when the reconstruction metrics on the validation dataset reach a point of saturation.

3.3. *R2D2 variants*

The R2D2 algorithm takes three variants. The first, simply referred to as R2D2 henceforth, adopts the well-established U-Net as the architecture of its DNN modules. The second, called R2D2-Net, is an unrolled version of R2D2 itself, trained end-to-end on GPUs. Its architecture consists of J U-Net layers interlaced with $J-1$ data consistency layers. The latter layers leverage a fast and memory-efficient PSF-based approximation of the operator $\mathbf{D}$ for the computation of residual images, which is instrumental to alleviate the scalability challenges encountered by unrolled deep learning approaches. The third is a nested R2D2 architecture, where the U-Net modules of R2D2 are substituted with R2D2-Net. In reference to its nested structure, this variant is dubbed "**R**ussian **D**olls" **R2D2**, in short **R3D3**. In this landscape, the R2D2-Net variant, as DNN module to R3D3, also represents R3D3's first iteration.

4. WHEN R2D2 MEETS CLEAN

In exploring the joint landscape of R2D2 and CLEAN paradigms, we rely on three main features: (i) the iterative structure, either nested or non-nested; (ii) the data model used to update the residual dirty images, relying either on the exact measurement operator and non-gridded visibilities, or a PSF-based approximation operating on gridded data; (iii) the model component basis, either handcrafted (identity or multi-scale) or learned. In what follows, CLEAN and R2D2 variants are characterized and compared via the lens of this three-fold categorization. Table 1 summarizes this landscape.

In the CLEAN paradigm, Hö-CLEAN exhibits a non-nested iterative structure, taking the dirty image as input. At each iteration, model components are identified in the identity basis from the residual dirty image.

Table 1. Joint landscape of R2D2 and CLEAN paradigms

	Hö-CLEAN	R2D2-Net	R2D2	CS-CLEAN	MS-CLEAN	R3D3
Iterative structure	non-nested	non-nested	non-nested	nested	nested	nested
Data model	gridded	gridded	non-gridded	non-gridded	non-gridded	non-gridded
Component basis	identity	learned	learned	identity	multi-scale	learned

The latter is modeled as a convolution of the sought image with the PSF, representing an approximation to the data model. CS-CLEAN exhibits a nested iterative structure, also taking the dirty image as input. The minor cycle serves as a regularization step, with the image estimate updated with multiple model components in the identity basis. The major cycle ensures data fidelity through accurate updates of the residual dirty image, whereby the contribution of the current image estimate is removed from the non-gridded visibilities using the exact measurement operator. Finally, MS-CLEAN can benefit from the same "major-minor" cycle structure as CS-CLEAN, while featuring a bespoke multi-scale component basis in lieu of the identity basis.

In the R2D2 paradigm, all variants also take the dirty image as input. The unrolled R2D2-Net alternates between DNN modules, serving as regularization layers, and approximate PSF-based data-consistency layers. It thus closely mirrors the iteration structure of Hö-CLEAN, sharing a non-nested structure and a PSF-based data model. But, its model components are identified at each regularization layer in a learned manifold of basis functions encoded by the corresponding U-Net, rather than the identity basis. R2D2 itself can be interpreted as a learned version of a hybrid CLEAN at the interface of Hö-CLEAN, CS-CLEAN, and MS-CLEAN. It indeed features the exact data model of CS-CLEAN and MS-CLEAN for the update of the residual dirty images, but is underpinned by a non-nested structure, as is Hö-CLEAN. Its model components are identified at each iteration in a learned manifold of basis functions encoded by the corresponding U-Net module, rather than a handcrafted (identity or multi-scale) basis. Finally, R3D3 exhibits a nested structure matching the major-minor cycle structure of CS-CLEAN and MS-CLEAN. Its R2D2-Net modules serve as its minor cycles, just as Hö-CLEAN (resp. a multi-scale variant) serves as CS-CLEAN's (resp. MS-CLEAN) minor cycle structure. Its model components are identified at each iteration in the same learned manifold of basis functions as R2D2.

5. APPLICATION TO CYGNUS A

5.1. *Observation and imaging settings*

The Cygnus A data considered were acquired at S band (2.052 GHz) with the following observation setting. VLA configurations A and C were combined, with a single pointing centered at the inner core of Cygnus A, given by the coordinates RA = 19h59m28.356s (J2000) and DEC = +40º4′2.07″. The total observation duration is about 19 hours conducted over multiple scans, respectively combining 7 and 12 hours with configurations A and C. The data are single-channel, acquired with integration time-step of 2 seconds and channel-width of 2 MHz. Careful calibration of the direction independent effects (DIEs) was performed in AIPS (Sebokolodi et al. 2020), following which, the data were averaged over 10 seconds totaling $M = 9.6 \times 10^5$ points. Dabbech et al. (2021) showed that these data could also benefit from the calibration of DDEs, likely attributed to pointing errors at the hotspots. However, DDE solutions from the underpinning joint calibration and imaging framework (Repetti et al. 2017) are not considered in this study as none of the benchmark algorithms aside from uSARA and AIRI could benefit from them.

As far as the imaging setting is considered, we target forming an image of size $N = 512 \times 512$, with a pixel-size of about 0.29 arcsec, corresponding to a super-resolution factor of 1.5, from Briggs-weighted data with Briggs parameter set to 0 (weights are generated in WSClean; Offringa & Smirnov 2017). The target dynamic range of the reconstruction is about 1.7×10^5, inferred from the data as the ratio between the peak pixel value of the sought image and the image-domain noise level estimated as $\tau/\sqrt{2\|\mathbf{\Phi}^\dagger\mathbf{\Phi}\|_S}$ (Terris et al. 2022), with $\|.\|_S$ denoting the spectral norm of its argument operator.

5.2. *R2D2 incarnations and implementations*

We have borrowed the first incarnations of the R2D2 variants from Aghabiglou et al. (2024). These were trained in a telescope-specific approach, more precisely for VLA. R2D2 and R3D3 were respectively trained with $I = 15$ U-Net modules, and $I = 8$ R2D2-Net modules, all sharing the same architecture composed of $J = 6$ U-Net layers. The R2D2-Net incarnation considered is exactly the first DNN module of the R3D3 incarnation.

The training dataset was composed of 20,000 pairs of ground truth images and associated dirty images of size $N = 512 \times 512$. Synthetic ground truth images, endowed with high dynamic ranges in the interval $[10^3, 5 \times 10^5]$, were created from a low-dynamic range database of optical astronomy and medical images via denoising and exponentiation procedures (Terris et al. 2022, 2023). Tar-

geting a robust incarnation, usable across a wide landscape of VLA observation settings, data were created from a variety of Fourier sampling patterns generated using MeqTrees (Noordam & Smirnov 2010), combining configurations A and C. The sampling patterns were generated by sampling uniformly at random: (i) the pointing direction; (ii) the temporal parameters (total observation duration assuming a single scan and a constant time-step of 36 seconds for a combined duration at both configurations ranging between 6 hours and 13 hours); (iii) the spectral parameters (number of consecutive frequency channels and frequency channel bandwidth) under the assumption of flat spectra of the radio emission. Additionally, random flagging of up to 20 % of the Fourier samples was applied. The simulated visibilities of size $M \in [0.2, 2] \times 10^6$ were corrupted with noise levels commensurate with the dynamic ranges of the corresponding ground truth images. In order to contain the number of varying parameters in the training, the imaging setting was fixed, with a $N = 512 \times 512$ image size, and dirty images obtained: (i) using the same weighting scheme across all data, that is Briggs weighting whose parameter is set to 0; (ii) with a pixel size ensuring a super-resolution factor of 1.5.

We note that, while the imaging setting for the pretrained R2D2 variants and the target image reconstruction of Cygnus A are the same, the VLA temporal parameters of the observation settings (total duration, integration time-step, number of scans) belong to quite different categories. Our working hypothesis in using these pre-trained variants is that the wide variety of observation settings visited at training stage endows the DNN modules with the necessary robustness, when interlaced with the update of the residual dirty image at each iteration, to generalize across different categories of observation settings. The same question in fact formally holds with regards to the target category of images, as the training dataset contains no radio image whatsoever. The capability to train on synthetic data and reconstruct radio images was extensively validated in simulation (Aghabiglou et al. 2024), but remains to be demonstrated on real data.

Python and MATLAB implementations of the R2D2 variants which can be executed on flexible hardware are available as part of the BASPLib code library. Extensive validation of image reconstruction in simulation, as well as a detailed analysis of their computational performance at both training and imaging stages are provided in Aghabiglou et al. (2024). In this study, image reconstruction with R2D2 variants was conducted using their Python implementation, executed on a single GPU.

5.3. Benchmark algorithms

The performance of the three R2D2 variants is studied in comparison with the Hö-CLEAN, CS-CLEAN, MS-CLEAN, uSARA, and AIRI. Parameters of CLEAN variants were adjusted to ensure a good compromise between speed and data fidelity, while achieving a good reconstruction quality. Both uSARA and AIRI are shipped with automated noise-driven heuristics for the choice of their regularization parameter (Terris et al. 2022; Wilber et al. 2023a). However, an adjustment within one order of magnitude from their heuristic values appeared to be necessary for best results. For CLEAN variants, we have considered the widely-used C++ software WSClean (Offringa & Smirnov 2017, see Appendix A for full commands). Both uSARA and AIRI are implemented in MATLAB as part of the BASPLib code library. All CLEAN variants and uSARA were executed on a single CPU core. AIRI was executed on a single CPU core for its data-fidelity steps and single GPU for its regularization steps (denoising DNNs).

5.4. Imaging results

Cygnus A images obtained by the different imaging methods are displayed in $\log_{10}$ scale in Figs. 1–4. Reconstructions are overlaid with additional panels consisting of (a) the associated residual dirty images displayed in linear scale to visually assess the fidelity to back-projected data, and zooms on selected regions of the radio galaxy, all displayed in $\log_{10}$ scale: (b) the inner core, (c) the West hotspots, and (d) the East hotspots. The overall visual inspection of Cygnus A reconstructions shows that R2D2 variants exhibit higher resolution than CLEAN variants, while generally corroborating the achieved depictions by AIRI and uSARA. They provide *deep* reconstructions, whose pixel values span nearly five orders of magnitude, which is in line with the target dynamic range estimate. A close-up inspection indicates that both R3D3 (Fig. 3, bottom) and AIRI (Fig. 4, bottom) stand out, owing to their high levels of detail and their limited amount of patterns that could be construed as artifacts. R2D2 (Fig. 2, bottom) and R2D2-Net (Fig. 3, top) seem to lack details in the faint extended emission. uSARA (Fig. 4, top) depicts spurious ringing and wavelet-like patterns. As expected, Hö-CLEAN (Fig. 1, top) delivers a poor reconstruction with severely limited dynamic range, due to its inherent approximate data model. Both CS-CLEAN (Fig. 1, bottom) and MS-CLEAN (Fig. 2, top) provide much improved reconstructions, with the former exhibiting grid-like artifacts due to its inadequate sparsity (identity) basis for the complex target radio source.

The inner core—The inner core consists of the point-like active galactic nucleus (AGN) of Cygnus A, from which two jets emanate (panels (b) of all figures). The reconstructions of R2D2 variants, uSARA, and AIRI show a super-resolved depiction of the region. In particular, R2D2 variants exhibit continuous emission between the AGN and both jets. uSARA exhibits wavelet-like artifacts around the AGN. CLEAN variants provide unresolved depiction of the source due to the restoring beam.

The lobes—Examination of the West and East lobes of Cygnus A highlights the ability of R2D2 variants to provide a more physical depiction of their filamentary structure than the benchmark algorithms. On the one hand, CLEAN variants deliver a smooth reconstruction. On the other hand, uSARA, and to a much lesser extent AIRI, exhibit ringing artifacts in the West lobe (pointed at with a green arrows in Fig. 4). These artifacts are likely induced by pointing errors at the hotspots, resulting in the over-fitting of the high-spatial frequency content of the data by both uSARA and AIRI. Joint DDE calibration and imaging (using either AIRI or uSARA as the imaging module) can drastically reduce (if not remove) these artifacts (see their corresponding reconstructions provided in Dabbech et al. 2024). These findings suggest that R2D2 variants may be less prone to calibration errors than AIRI and uSARA.

Focusing on the faint diffuse emission at the tails of the lobes, both R3D3 and AIRI provide consistent depiction. When flipping between their associated figures (Figs. 3–4, bottom), the R3D3 reconstruction appears sharper. One such example is the faint filamentary structure at the tail of the East lobe. However, one can not ascertain whether this sharpness is physical, particularly in absence of DDE calibration. Both R2D2 (Fig. 2, bottom) and R2D2-Net (Fig. 3, top) provide a rather smooth structure. uSARA introduces wavelet-like patterns (pointed at with red arrows in Fig. 4, top). The faint emission under scrutiny is completely buried in the noise in the Hö-CLEAN reconstruction (Fig. 1, top). A blurry and noisy depiction emerges in the reconstructions of both CS-CLEAN (Fig. 1, bottom) and MS-CLEAN (Fig. 2, top).

The hotspots—As recovered by R2D2 variants, AIRI and uSARA, the hotspots highlight the ability of these algorithms to resolve physical structure beyond instrumental resolution, in contrast with CLEAN variants (see panels (c) and (d) of all figures). Interestingly, where AIRI and uSARA exhibit artificial zero-valued pixels around the hotspots, all R2D2 variants depict continuous emission. This observation suggests the ability of R2D2 variants to achieve a more physical reconstruction.

Table 2. Imaging computational details: the number of iterations I, the allocated computing resources in CPU core (n_{cpu}) and GPU (n_{gpu}), the total imaging time ($t_{\mathrm{tot.}}$), the average time of the data-fidelity step per iteration ($t_{\mathrm{dat.}}$), and the average time of the regularization step per iteration ($t_{\mathrm{reg.}}$), all in seconds (sec). The time to compute the dirty image is excluded. For CLEAN variants, I is the number of iterations in the major cycle, $t_{\mathrm{dat.}}$ is the average time to update the residual dirty image, and $t_{\mathrm{reg.}}$ is the average time taken by a minor cycle.

	I	n_{cpu}	n_{gpu}	$t_{\mathrm{tot.}}$ (sec)	$t_{\mathrm{dat.}}$ (sec)	$t_{\mathrm{reg.}}$ (sec)
Hö-CLEAN	1	1	-	21.87	-	21.87
CS-CLEAN	7	1	-	33.69	2.38	2.43
MS-CLEAN	9	1	-	39.01	2.50	1.83
R2D2	15	-	1	3.31	0.13	0.10
R2D2-Net	1	-	1	0.97	-	0.97
R3D3	8	-	1	1.83	0.06	0.18
uSARA	477	1	-	1197	0.64	1.83
AIRI	1783	1	1	672	0.32	0.05

Image-domain data fidelity—We evaluate data fidelity delivered by the imaging algorithms by scrutinizing their residual dirty images (i.e., back-projected data residual) displayed in panels (a) of all figures, whose standard deviation values are reported in the corresponding captions. AIRI, uSARA, and CS-CLEAN obtain residual dirty images with the lowest standard deviation values, reflecting their high data fidelity. Both MS-CLEAN and R2D2 variants provide comparable values, slightly above the best performing algorithms. Among R2D2 variants, R3D3 performs better than both R2D2 and R2D2-Net, owing to its underpinning model-informed DNN modules on the one hand, and its iterative structure on the other hand. Hö-CLEAN delivers the lowest fidelity with its standard deviation value being more than one order of magnitude higher than the rest. This numerical analysis aligns with the overall visual examination. A closer inspection of the R2D2 variants reveals discernible structure at the pixel positions of the Western and Eastern hotspots, where the highest emission is concentrated. While similar behavior has been observed in high-dynamic acquisition regimes in simulation (Aghabiglou et al. 2024), these patterns are possibly amplified by the pointing errors at the hotspots.

It is worth noting that the quest for noise-like data residuals is only really meaningful in combination with the recovery of a model satisfying physical constraints. The fact that CLEAN variants accept negative components eases the reduction of their data residuals. On the contrary, R2D2 variants, uSARA, and AIRI explicitly enforce the positivity of the recovered intensity images.

Computational cost—Computational details of the imaging algorithms are provided in Table 2. R2D2 variants are able to deliver reconstructions in a few seconds only. We note that reported times include a computational overhead for initializing DNN modules at the first iteration, inducing a large difference between the average time of the regularization step between R2D2-Net on one hand, and R2D2 and R3D3 on the other. Both CS-CLEAN and MS-CLEAN require around half a minute, whereas the highly iterative algorithms AIRI and uSARA take about 10 and 20 minutes, respectively. When compared to CS-CLEAN and MS-CLEAN, both R2D2 and R3D3 perform comparable number of passes through the non-gridded visibilities. Their DNN modules provide a boost to their speed in comparison with the minor cycles of the advanced CLEAN variants. R2D2-Net is also at least one order of magnitude faster than its CLEAN counterpart Hö-CLEAN. However, one must recognize the differences in implementation and hardware between the imaging algorithms, reflected in the varying average time of their data fidelity step (all calling for the computation of the residual dirty images).

Of note, the training of R2D2 variants required the order of thousands of CPU core hours and GPU hours (Aghabiglou et al. 2024). This is however a one-off cost accommodating any number of subsequent reconstructions from VLA data, as exemplified by the fact that the present analysis did not require any training.

6. CONCLUSIONS

We have shed light on R2D2's algorithmic structure, interpreting it as a learned version of CLEAN. The R2D2, R2D2-Net, and R3D3 variants have been cautiously positioned in the landscape of Hö-CLEAN, CS-CLEAN, and MS-CLEAN. In this context, we have provided the first demonstration of R2D2 on real data, for monochromatic intensity imaging of the radio galaxy Cygnus A from S band observations with the VLA. R2D2 variants deliver superior imaging quality to CLEAN's, while offering some acceleration potential owing to the substitution of minor cycles with DNNs. In comparison to uSARA and AIRI, R2D2 variants at least equate their imaging precision, at a fraction of the cost. Not to mention that the regularization parameter of uSARA and AIRI has been adjusted manually, while R2D2 runs in an automated manner. These results confirm that R2D2's learned approach offers an immense potential for fast precision RI imaging, already generalizing across categories of images and observations settings.

Future research includes investigating the robustness of the R2D2 paradigm to accommodate a much wider variety of observation and imaging settings, ranging from varying data-weighting schemes, to enabling flexible super-resolution factors and image dimensions, and ultimately an all-instrument-encompassing incarnation. Further developments towards improving its capability to deliver new regimes of precision are warranted, e.g., by leveraging novel DNN architectures (including diffusion models, Ho et al. 2020) and training loss functions for more efficient and physics-informed training and reconstruction. Finally, endowing the R2D2 paradigm with an image-splitting procedure, as implemented in DDFacet (Tasse et al. 2018), WSClean (Offringa & Smirnov 2017), and Faceted HyperSARA (Thouvenin et al. 2022), is a necessary evolution to efficiently handle large image sizes. In a nutshell, not only has CLEAN been leading the way for decades, but its algorithmic structure might well form the backbone of the next-generation deep learned-based imaging algorithms for radio astronomy.
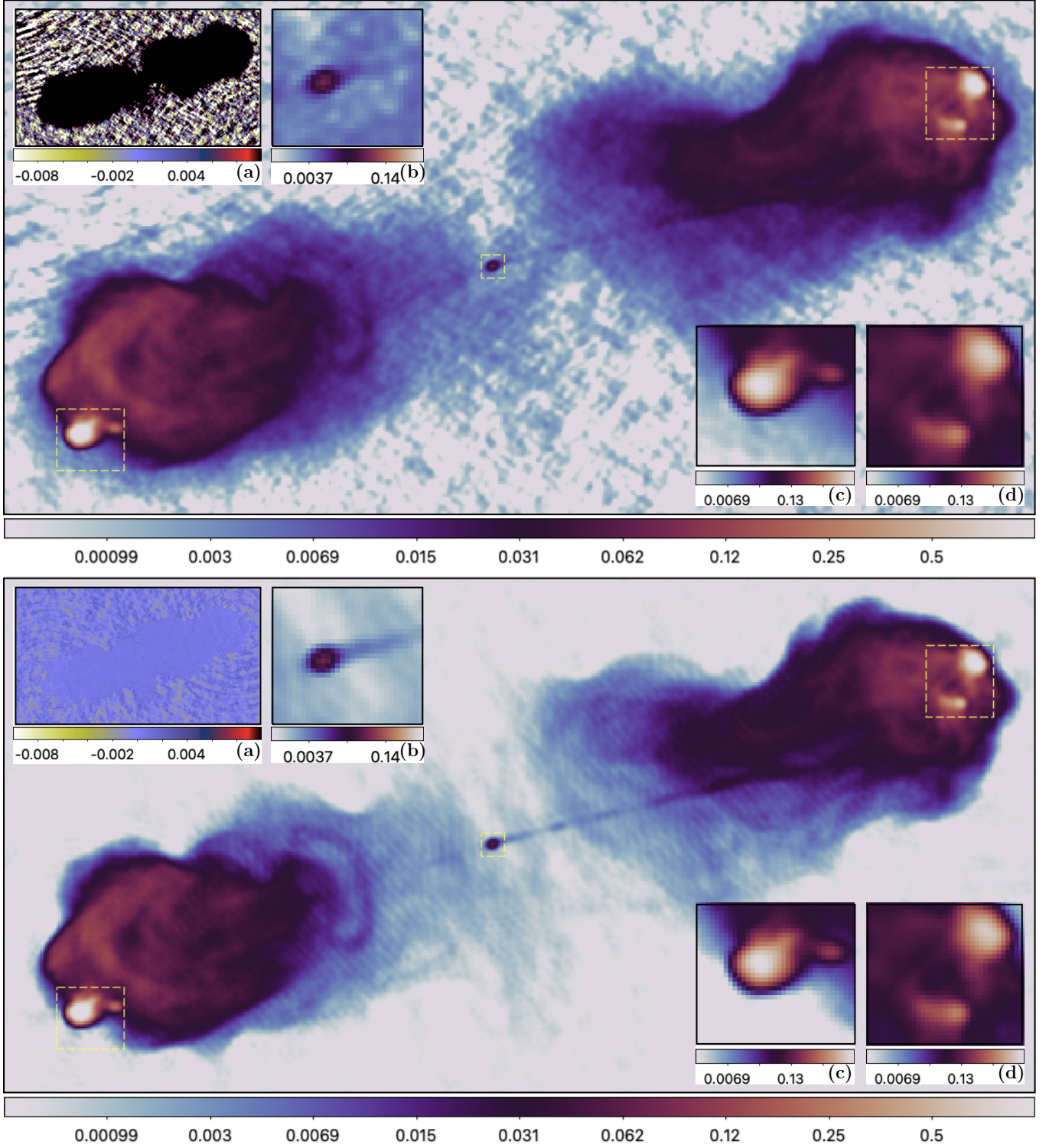
7. DATA AVAILABILITY

R2D2 codes are available alongside the AIRI and uSARA codes in the BASPLib code library on GitHub. BASPLib is developed and maintained by the Biomedical and Astronomical Signal Processing Laboratory (BASP). The VLA-trained R2D2 and R3D3 DNN series are available in the dataset Aghabiglou et al. (2024). Cygnus A reconstructions are available in FITS format in the dataset Dabbech et al. (2024). Observations of Cygnus A were provided by the National Radio Astronomy Observatory (NRAO; Program code: 14B-336). The self-calibrated data can be shared upon request to R. A. Perley (NRAO).
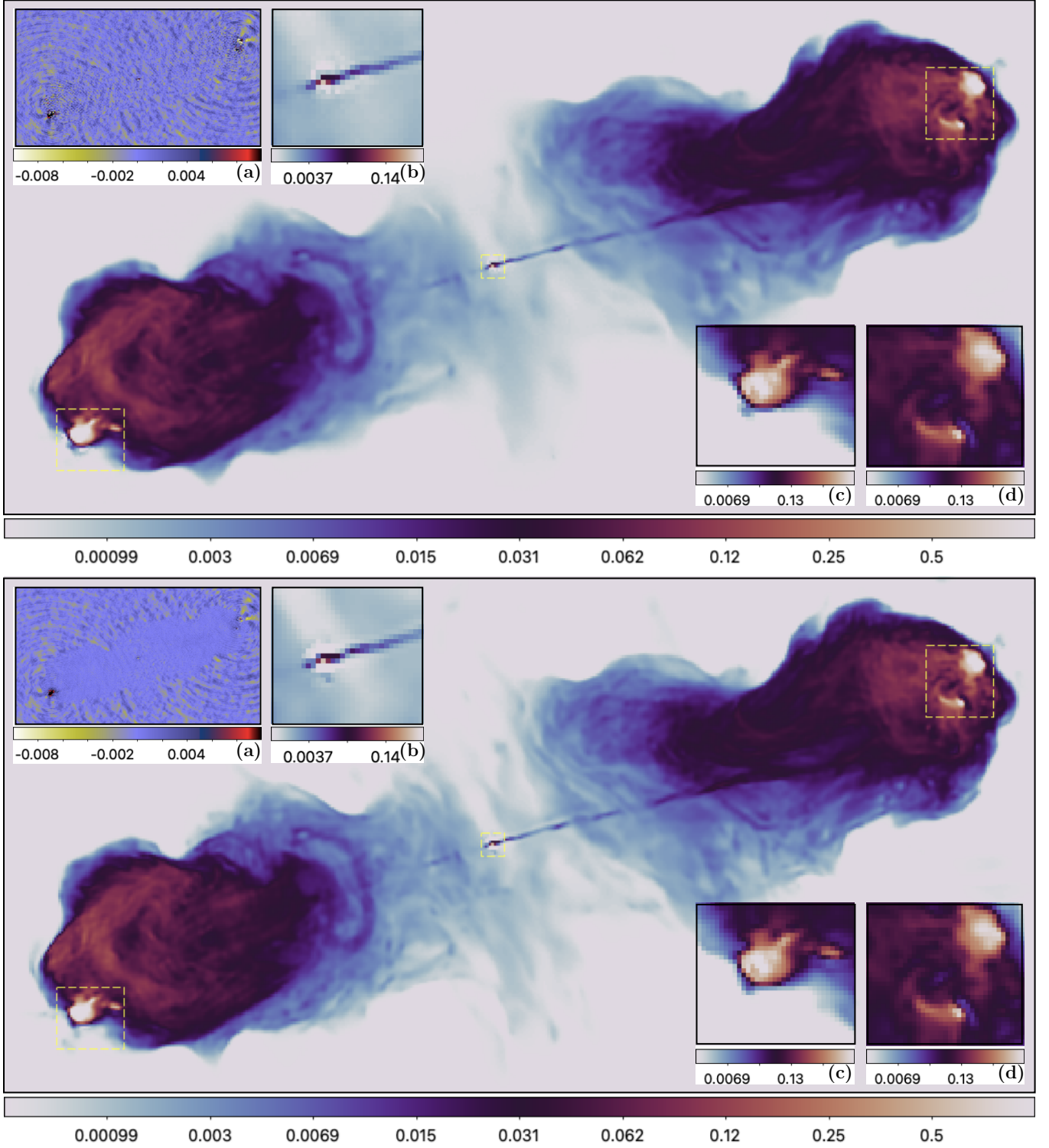
Figure 1. Cygnus A: reconstructions of Hö-CLEAN (top) and CS-CLEAN (bottom), displayed in $\log_{10}$ scale (negative pixels set to zero for visualization purposes). Their associated residual dirty images are provided in linear scale in panels (a), with standard deviation values 359×10^{-4}, and 8.6×10^{-4}, respectively. Reconstructions are overlaid by zooms on key regions of the radio galaxy: the inner core of Cygnus A in panels (b), the East hotspots in panels (c), and the West hotspots in panels (d), all displayed in $\log_{10}$ scale. Reconstructions are available in FITS format in the dataset Dabbech et al. (2024).
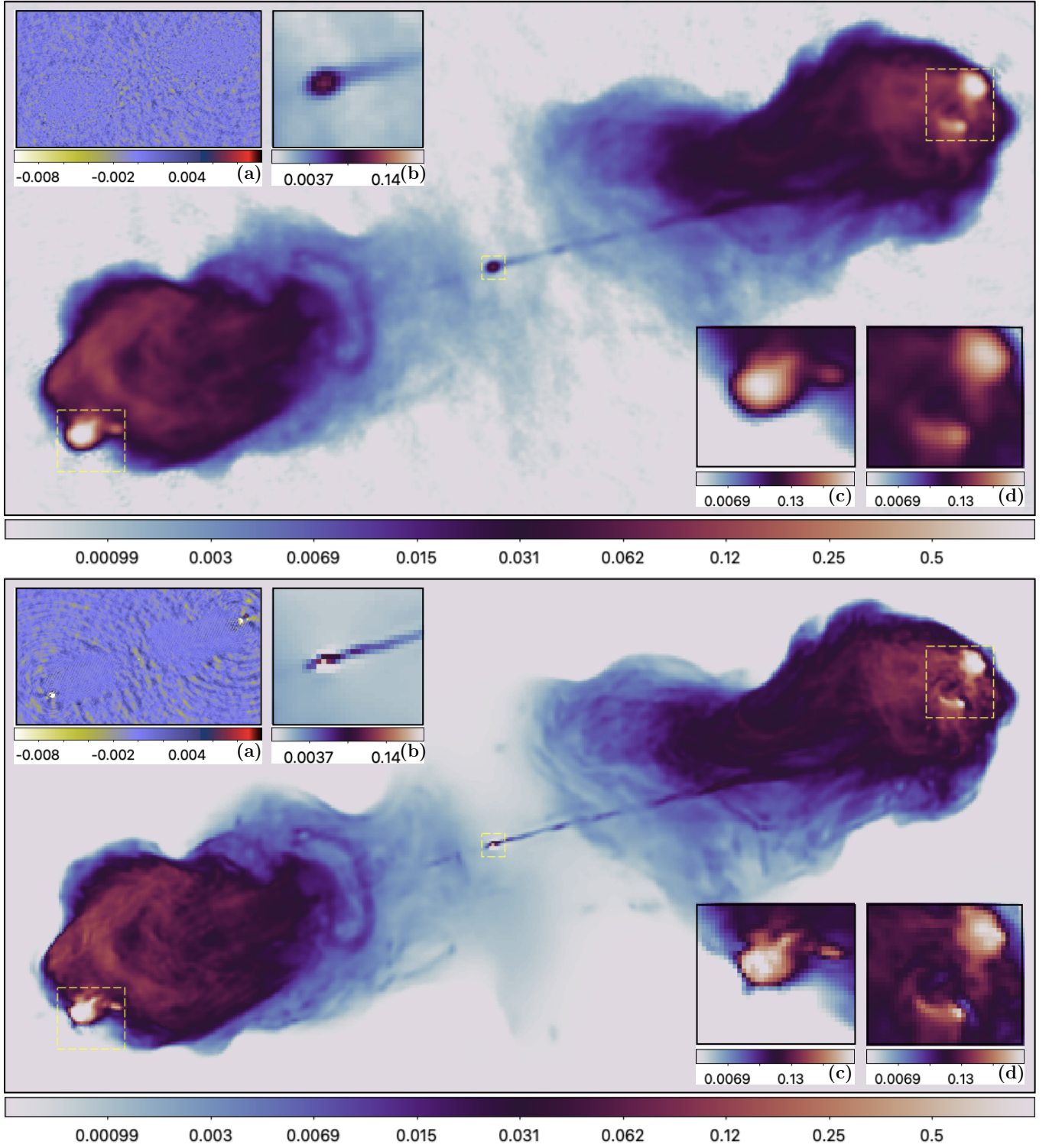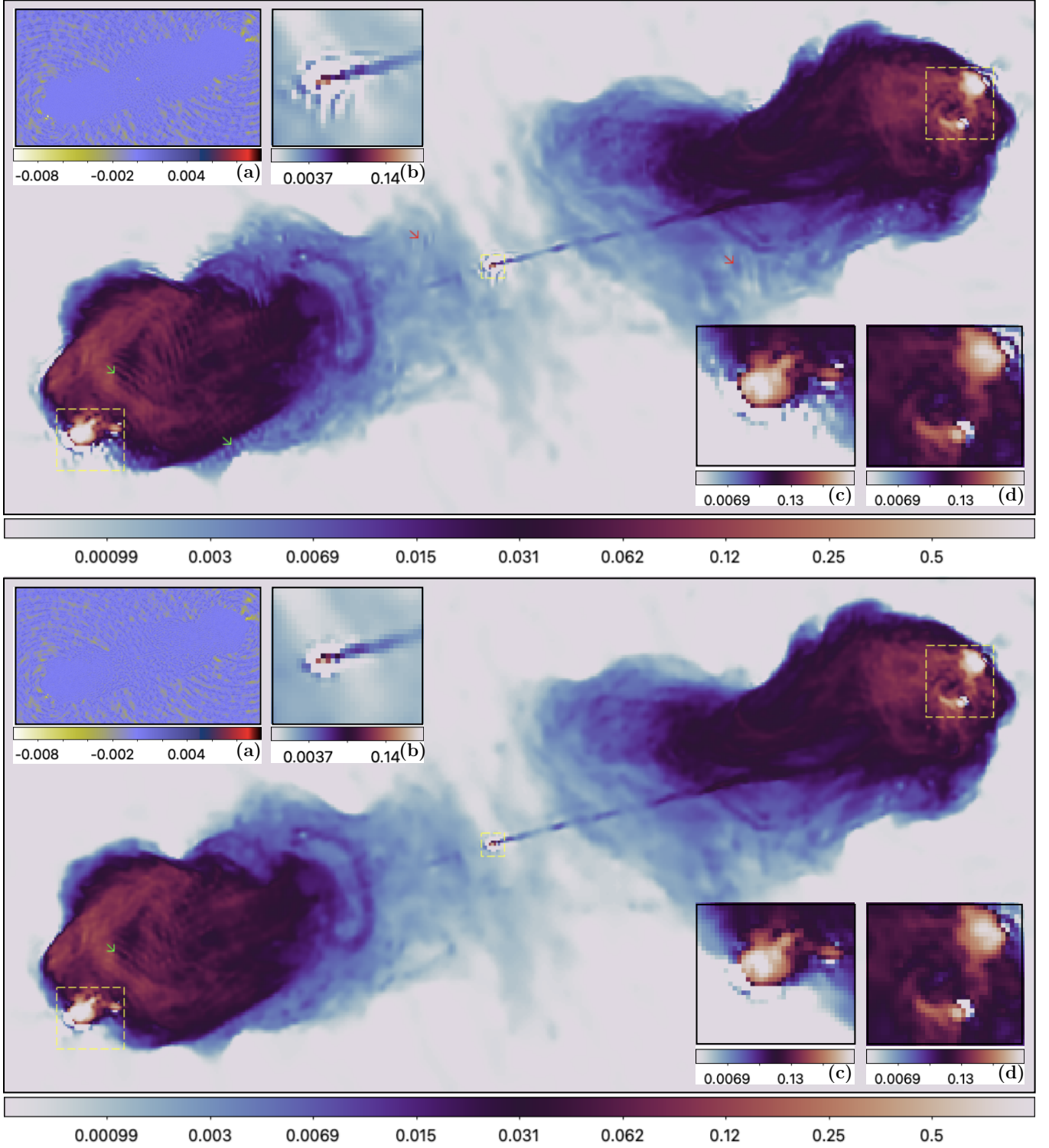
Figure 2. Cygnus A: reconstructions of MS-CLEAN (negative pixels set to zero for visualization purposes; top) and R2D2 (bottom), both displayed in $\log_{10}$ scale. Their associated residual dirty images are provided in linear scale in panels (a), with standard deviation values 10.4×10^{-4}, and 11.7×10^{-4}, respectively. Reconstructions are overlaid by zooms on key regions of the radio galaxy: the inner core of Cygnus A in panels (b), the East hotspots in panels (c), and the West hotspots in panels (d), all displayed in $\log_{10}$ scale. Reconstructions are available in FITS format in the dataset Dabbech et al. (2024).

Figure 3. Cygnus A: reconstructions of R2D2-Net (also the first iteration of R3D3; top), and R3D3 (bottom), both displayed in $\log_{10}$ scale. Their associated residual dirty images are provided in linear scale in panels (a), with standard deviation values 13.4×10^{-4}, and 9.7×10^{-4}, respectively. Reconstructions are overlaid by zooms on key regions of the radio galaxy: the inner core of Cygnus A in panels (b), the East hotspots in panels (c), and the West hotspots in panels (d), all displayed in $\log_{10}$ scale. Reconstructions are available in FITS format in the dataset Dabbech et al. (2024).

Figure 4. Cygnus A: reconstructions of uSARA (top) and AIRI (bottom), both displayed in $\log_{10}$ scale. Their associated residual dirty images are provided in linear scale in panels (a), with standard deviation values 7.2×10^{-4}, and 7.4×10^{-4}, respectively. Reconstructions are overlaid by zooms on key regions of the radio galaxy: the inner core of Cygnus A in panels (b), the East hotspots in panels (c), and the West hotspots in panels (d), all displayed in $\log_{10}$ scale. Reconstructions are available in FITS format in the dataset Dabbech et al. (2024).

APPENDIX

A. WSCLEAN COMMANDS

MS-CLEAN: `wsclean -size 512 512 -scale 0.29170266asec -weight briggs 0 -mgain 0.8 -gain 0.1 -auto-threshold .5 -auto-mask 1.5 -multiscale -padding 2 -niter 2000000 -j 1`

CS-CLEAN: `wsclean -size 512 512 -scale 0.29170266asec -weight briggs 0 -mgain 0.8 -gain 0.1 -auto-threshold .5 -auto-mask 1.5 -padding 2 -niter 2000000 -j 1`

Hö-CLEAN: `wsclean -size 1024 1024 -scale 0.29170266asec -weight briggs 0 -mgain 1 -gain 0.1 -threshold 0.001Jy -auto-mask 1 -padding 2 -nmiter 1 -niter 500000 -j 1`

REFERENCES

Aghabiglou, A., Chu, C. S., A., Dabbech, A., & Wiaux, Y. 2024, R2D2 deep neural network series for radio-interferometric imaging, doi: 10.17861/99cbe654-5071-4625-b59d-a26c790cbeb4

Aghabiglou, A., San Chu, C., Dabbech, A., & Wiaux, Y. 2024, preprint arXiv:2403.05452

Bhatnagar, S., & Cornwell, T. J. 2004, A&A, 426, 747

Briggs, D. 1995, in AAS Meeting Abstracts, Vol. 187, 112–02

Carrillo, R. E., McEwen, J. D., & Wiaux, Y. 2012, MNRAS, 426, 1223

Clark, B. G. 1980, A&A, 89, 377

Connor, L., Bouman, K. L., Ravi, V., & Hallinan, G. 2022, MNRAS, 514, 2614

Cornwell, T. J. 2008, IEEE J. Sel. Top. Signal Process., 2, 793

Dabbech, A., Aghabiglou, A., Chu, C. S., , & Wiaux, Y. 2024, Cygnus A reconstructions at S band with R2D2, doi: 10.17861/76034345-9e4d-488f-9558-3bd7674f41ee

Dabbech, A., Ferrari, C., Mary, D., et al. 2015, A&A, 576, A7

Dabbech, A., Repetti, A., Perley, R. A., Smirnov, O. M., & Wiaux, Y. 2021, MNRAS, 506, 4855

Dabbech, A., Terris, M., Jackson, A., et al. 2022, ApJL, 939

Garsden, H., Girard, J., Starck, J.-L., et al. 2015, A&A, 575, A90

Ho, J., Jain, A., & Abbeel, P. 2020, Adv. Neural Inf. Process. Syst., 33, 6840

Högbom, J. 1974, A&AS, 15, 417

Kamilov, U. S., Bouman, C. A., Buzzard, G. T., & Wohlberg, B. 2023, IEEE Signal Process. Mag., 40, 85

Lannes, A., Anterrieu, E., & Marechal, P. 1997, A&AS, 123, 183

Li, F., Cornwell, T. J., & de Hoog, F. 2011, A&A, 528, A31

Monga, V., Li, Y., & Eldar, Y. C. 2021, IEEE Signal Process. Mag., 38, 18

Noordam, J. E., & Smirnov, O. M. 2010, A&A, 524, A61

Offringa, A. R., & Smirnov, O. 2017, MNRAS, 471, 301

Repetti, A., Birdi, J., Dabbech, A., & Wiaux, Y. 2017, MNRAS, 470, 3981

Repetti, A., & Wiaux, Y. 2020, in Proc. of ICASSP 2020 IEEE, 1434–1438

Schmidt, K., Geyer, F., Fröse, S., et al. 2022, A&A, 664, A134

Schwab, F. R. 1984, AJ, 89, 1076

Sebokolodi, M. L. L., Perley, R., Eilek, J., et al. 2020, ApJ, 903, 36

Smirnov, O. M. 2011, A&A, 527, A106

Tasse, C., Hugo, B., Mirmont, M., et al. 2018, A&A, 611, A87

Terris, M., Dabbech, A., Tang, C., & Wiaux, Y. 2022, MNRAS, 518, 604

Terris, M., Tang, C., Jackson, A., & Wiaux, Y. 2023, preprint arXiv:2312.07137

Thouvenin, P.-A., Abdulaziz, A., Dabbech, A., Repetti, A., & Wiaux, Y. 2022, MNRAS, 521, 1

Wakker, B. P., & Schwarz, U. 1988, A&A, 200, 312

Wiaux, Y., Jacques, L., Puy, G., Scaife, A. M., & Vandergheynst, 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