Learning-based probabilistic subarray switching for robust low-cost interferometric imaging

Jianhua Wang*, Mohammed Nabil El Korso‡, Lucien Bacharach*, Pascal Larzabal*

***Abstract*—Computational cost poses a significant challenge in next-generation interferometric imaging systems. In these systems, the large number of antennas makes it impractical to process all measurements simultaneously due to computational capacity constraints. To reduce the computational burden while preserving image reconstruction quality, we propose a subarray switching strategy that utilizes fewer antennas and different antenna configurations. To take into consideration the influence of the image reconstruction algorithm on the design of the subarray switching pattern and to fully exploit the flexibility of the switching strategy, we propose a probabilistic deep learning-based method for designing antenna switching patterns, named by Probabilistic Antenna Switcher (PAS). In addition to the computational challenge, interferometric systems are also particularly sensitive to the presence of radio frequency interferences (RFI), which heavily affects imaging quality. In order to address this issue, we show that it is possible to combine the proposed PAS with a RFI detection module. Specifically, this module is a neural network that is trained to identify and minimize the impact of RFI-affected antennas in the subarray selection process. This results in a RFI-aware PAS (RaPAS) , which balances computational efficiency, imaging quality, and robustness against RFI. The full code will be available at https://github.com/jianhua-WANG-ENS/PAS_RaPAS.**

***Index Terms*—Array processing, antenna selection, imaging, interferometric array**

I. Introduction

The emergence of data-intensive interferometers [1–3] has driven substantial advancements in sensitivity and resolution through the inclusion of a greater number of antennas. However, alongside this expansion in observational capacity arises a dual challenge: 1) the computational challenges and cost [4, 5] imposed by the increased number of antennas and the resulting data dimensionality; 2) the susceptibility to interferences, notably Radio Frequency Interference (RFI), due to heightened sensitivity, where the presence of RFI can significantly degrade the quality of reconstructed images [6–8]

First, to tackle the challenge of computational complexity, a practical approach is antenna subarray selection, which mitigates computational complexity by partitioning the full array into smaller subsets and utilizing only a fraction of the antennas for each measurement [9, 10]. This method not only reduces the data size but also lowers computational demands, making large-scale, data-intensive interferometry more manageable. Despite its advantages, optimal subarray selection remains a complex problem. Existing methods, such as deterministic approaches based on lower bounds of Mean Square Error (MSE) for source localization include the Cramér-Rao Bound (CRB) [11–13] and the Barankin-Type Bound (BTB) [14]. These methods provide structured antenna configurations, either for a single source or multiple well-separated sources. Specifically, the CRB-based method generates a beampattern with a narrow mainlobe but high sidelobes, while the BTB-based method achieves a balance between mainlobe width and sidelobe levels [10, 15, 16]. As a result, the CRB-based method provides the highest imaging resolution but may introduce artifacts due to the high sidelobes. In contrast, the BTB-based method offers a better trade-off between resolution and artifacts, though at the cost of reduced resolution compared to the CRB-based approach.

An antenna subarray switching strategy between different types of subarrays is introduced in recent work [17, 18] to reduce computational cost while preserving the quality of reconstructed images. Assuming two different switching patterns, this strategy is implemented by combining CRB-based and BTB-based subarrays to exploit the advantages of both methods. The results demonstrate the efficiency of this switching strategy compared to using a single subarray.

However, this switching strategy is optimal only when the selected imaging algorithm attains the CRB or BTB, which is not always the case. In practice, the performance of switching between different subarrays is influenced by the chosen imaging algorithm [19]. Therefore, it is essential to consider the imaging algorithm when designing antenna subarray switching patterns. Furthermore, to maximize the potential and explore the flexibility of the switching strategy, it is necessary to be able to design more than two subarray patterns.

Second, RFI poses a long-standing challenge in interferometry. It originates from various sources, such as communication systems (e.g., mobile phones) or high-voltage transmission lines [20, 21]. Conventional thresholding-based methods often focus on detecting and flagging strong outliers in the raw data. These techniques include manual inspection [22] or spatial filtering of the correlation matrix [23–25]. In contrast to thermal noise, which can be modeled using a Gaussian distribution, RFI exhibits complex statistical properties that are difficult to model accurately. Recent approaches have employed t-distributions to better capture the statistical characteristics of RFI in radio-astronomical imaging [26, 27]. However, the latter RFI mitigation approaches face significant challenges, primarily due to how the mitigation task is formulated. Conventional thresholding-based [28, 29] and data-driven [30, 31] methods treat RFI mitigation as a binary classification or segmentation problem. In other words, they focus solely on

* *SATIE, Université Paris-Saclay, 91190 Gif-sur-Yvette, France;* ‡ *L2S, Université Paris-Saclay, 91190 Gif-sur-Yvette, France*
This research was funded in part by l'Agence Nationale de la Recherche (ANR), project ANR-24-CE48-5182.

identifying and labeling parts of the visibilities corrupted by RFI, a process known as RFI flagging. Once these corrupted visibilities are identified, they are typically discarded from further analysis. This approach heavily relies on the accuracy of RFI detection. As a result, valuable information may be lost, depending on the location of the affected visibilities and the subsequent task, in our case, image reconstruction.

To address these limitations, we propose a method for designing antenna subarray switching patterns while considering the chosen imaging algorithm. Given the complexity of real-world measurements, analytical solutions to this switching pattern design problem are infeasible. Instead, our method employs deep learning to learn the optimal antenna subarray switching pattern using Gumbel reparameterization to enable backpropagation gradient estimation. Additionally, we introduce a RFI detection block to mitigate the impact of RFI during the antenna selection process. This block analyzes the measurements and estimates the probability of each antenna being contaminated by RFI. This probability is then combined with the antenna selection probability to determine the most optimal and reliable antennas for forming the subarray switching pattern. This approach provides a new perspective on mitigating RFI by addressing it at the data acquisition stage.

The content of this paper is organized as follows. In Section II, we introduce the interferometric sensing model and formulate the framework of antenna switching. In Section III, we present our proposed method, which optimizes the antenna switching pattern while considering the image reconstruction process. In Section IV, we demonstrate the effectiveness of our approach based on the Very Large Array (VLA) configuration. We also analyze the reduction in computational complexity achieved with the proposed method. Finally, the work is summarized in Section V.

II. System model and problem formulation

A. Interferometric data model

Interferometry involves measuring the spatial coherence of the electric field between pairs of antennas within the antenna array. These measurements, termed visibilities, provide information about the electromagnetic radiation emitted by celestial sources under observation. Without loss of generality and to simplify notation, we assume that each measurement contains the same number of visibilities. The interferometric imaging model for each snapshot measurement block k is formulated as

$$\mathbf{r}_k = \mathbf{H}_k\mathbf{x} + \mathbf{n}_k,\ k = 1, \cdots, K \tag{1}$$

where $\mathbf{r}_k \in \mathbb{C}^N$ denotes the observed visibility vector, $\mathbf{H}_k \in \mathbb{C}^{N\times m}$ is the forward operator, corresponding to the incomplete Fourier sampling operator, whereas $\mathbf{x} \in \mathbb{R}^m$ represents the image vector with a number of image pixels m to be restored, K is the total number of snapshot measurement block, and N represent the number of per-snapshot visibilities. The additive noise is assumed to follow a Gaussian distribution $\mathbf{n}_k \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_N)$.

B. Problem formulation

The main idea of this paper is to select the different "best" ℓ antennas (out of a L-element array) for each measurement k, forming a subarray switching pattern. We then use the visibilities corresponding to the selected antennas to achieve the best possible image reconstruction quality while reducing computational complexity.

Let $\mathbf{m}_k \in \{0,1\}^L$ be a ℓ-hot encoding vector, where the i-th element indicates whether the i-th antenna is selected (1) or not (0) for the k-th measurement. The number of selected antennas is ℓ, i.e., $\|\mathbf{m}_k\|_1 = \ell$. Based on this, we define a switching matrix $\mathbf{M}_k$ to select the visibilities corresponding to the chosen antennas

$$\mathbf{M}_k = \zeta\left(\mathrm{diag}(\mathbf{m}_k \otimes \mathbf{m}_k)\right) \tag{2}$$

where $\mathbf{M}_k \in \{0,1\}^{n\times N}$ and the operator $\zeta(\cdot)$ remove all rows containing only zeros and $\otimes$ denotes the Kronecker product. Given the relationship between the number of antennas and visibilities, we have $N = \frac{L(L-1)}{2}$ and $n = \frac{\ell(\ell-1)}{2}$ where N is the total number of visibilities and n is the number of selected visibilities. For consistency, we assume that the number of selected antennas (and consequently the number of selected visibilities) remains the same for each measurement. Under this assumption, the selected visibility vector $\mathbf{y}_k \in \mathbb{C}^n$ can be expressed as

$$\mathbf{y}_k = \mathbf{M}_k\mathbf{r}_k,\ k = 1, \cdots, K \tag{3}$$

The primary objective in the radio-telescope imaging problem is to reconstruct the sky model image with the highest possible accuracy while minimizing computational overhead. In this work, we aim to achieve this by optimizing the switching matrices $\mathbf{M}_k$ and taking into account the image reconstruction process. This optimization problem can be formulated as follows

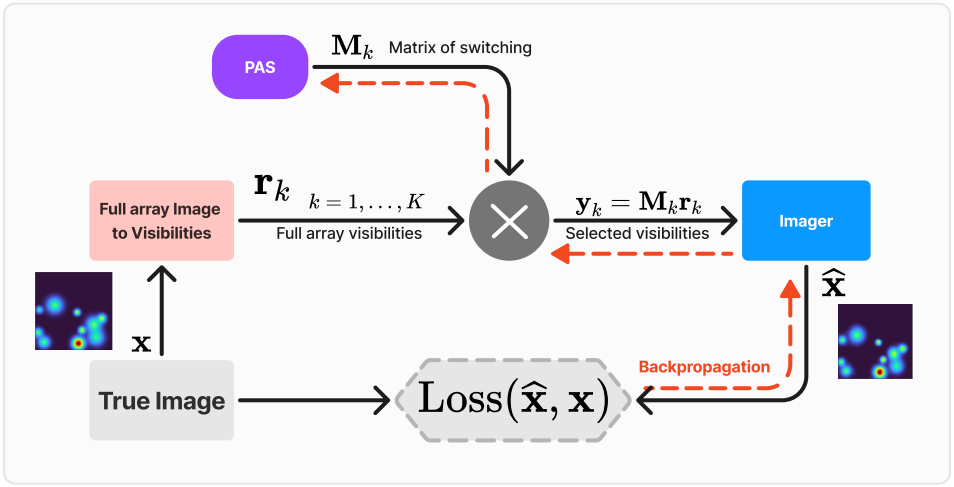
$$\begin{aligned}
\mathbf{m}_k^*, \boldsymbol{\theta}^* = \underset{\mathbf{m}_k, \boldsymbol{\theta}}{\arg\min} \quad & d\left(\hat{\mathbf{x}}, \mathbf{x}\right),\ k = 1, \ldots, K && \text{(4a)}\\
\text{subject to} \quad & \hat{\mathbf{x}} = f_{\boldsymbol{\theta}}\left(\mathbf{m}_k, \mathbf{r}_k\right), && \text{(4b)}\\
& \mathbf{m}_k \in \{0,1\}^L, && \text{(4c)}\\
& \|\mathbf{m}_k\|_1 = \ell. && \text{(4d)}
\end{aligned}$$

Here, the function $d(\cdot,\cdot)$ defines the distance between the reconstructed image $\hat{\mathbf{x}}$ and the true image $\mathbf{x}$. The image reconstruction function $f_{\boldsymbol{\theta}}(\cdot)$ is parameterized by $\boldsymbol{\theta}$, whose details are provided in subsequent sections. The second constraint (4c) ensures binary selection of antennas, while the third (4d) enforces that the total number of selected antennas equals ℓ.

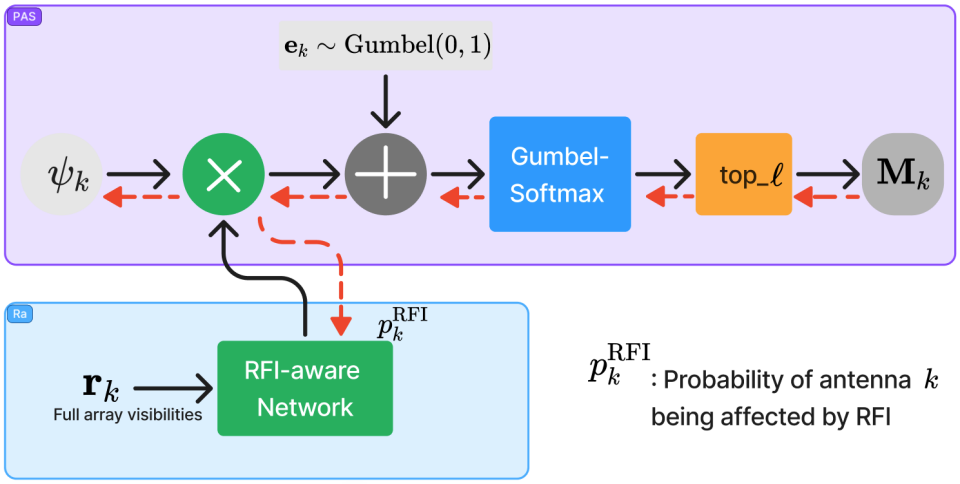
Denote K_{swit}, the number of switching patterns to design, the number of possible subarray combinations is given by

$$C = \binom{L}{\ell}^{K_{\text{swit}}} = \left(\frac{L!}{\ell!(L-\ell)!}\right)^{K_{\text{swit}}}.$$

Taking the VLA telescope as an example, as in [26], selecting $\ell = 15$ antennas from an array of $L = 27$ elements for $K_{\text{swit}} = 2$ switching patterns results in $C = 3.02 \times 10^{14}$ possible configurations. Exhaustively searching through all combinations is computationally infeasible.



(a)



(b)

Fig. 1: (a) Overview of the proposed framework, comprising a Probabilistic Antenna Switcher (PAS) and an image reconstruction block (Imager), trained to achieve an optimal antenna switching pattern. For each input true image, the corresponding visibilities of the full-array are simulated according to the antenna telescope observation model described in (1). The switching matrix $\mathbf{M}_k$ is applied to select a subset of the measurements, significantly reducing the total data size compared to the full-array measurements. The selected visibilities are processed by the image reconstruction block to produce a reconstructed image corresponding to the current switching pattern. By minimizing the distance between the reconstructed image and the true image, the PAS is optimized. (b) The illustration of the proposed PAS and RFI-aware PAS (RaPAS) blocks.

III. Probabilistic Antenna Switcher Learning Framework

Given that our antenna switching pattern design is intended for radio-astronomical imaging, we do not adopt learning approaches based on minimal bounds, such as in [32], where the CRB of the selected antennas is used as a criterion. Instead, we incorporate the imaging-related parameter $\boldsymbol{\theta}$ directly into our optimization problem (4a). Our criterion is the quality of the reconstructed images, which will be detailed in this section. This approach leads to the development of an end-to-end learning framework for designing the switching pattern, as illustrated in Fig. 1.

A. Probabilistic Antenna Switcher (PAS)

Direct optimization of the antenna selection vector $\mathbf{m}_k$ is intractable due to its discrete nature and following the constraints (4b) and (4c) within the end-to-end learning framework. To address this, we propose using a discrete sampling distribution [33, 34].

Specifically, each antenna selection vector $\mathbf{m}_k$ is modeled as the ℓ-hot encoding of an independent categorical random variable $\mathbf{c}_k \sim \mathrm{Cat}(L, \boldsymbol{\pi}_k)$. Here, $\boldsymbol{\pi}_k \in \mathbb{R}^L$ is a vector of L class probabilities, whose i-th element $\pi_{k,i}$ represents the probability of selecting the i-th antenna. We parameterize $\boldsymbol{\pi}_k$

using unnormalized logits $\boldsymbol{\psi}_k$, such that

$$\pi_{k,i} = \frac{e^{\psi_{k,i}}}{\sum_{j=1}^{L} e^{\psi_{k,j}}},\ i \in \{1, 2, \ldots, L\}. \tag{5}$$

Thus, rather than optimizing $\mathbf{m}_k$, we will optimize the distribution parameter $\boldsymbol{\psi}_k$ of $\mathbf{m}_k$.

Now the problem is how to take sample from the categorical distribution $\mathrm{Cat}\,(L, \boldsymbol{\pi}_k)$. For the forward pass, we need a ℓ-hot realization from $\mathrm{Cat}\,(L, \boldsymbol{\pi}_k)$ to reflect the selection of ℓ antenna, corresponding to the constraints (4c) and (4d). To address it, we use the Gumbel top-K trick [35–37]. A realization of $\mathbf{c}_k$ can be obtained by

$$\mathbf{m}_k = \ell\_\mathrm{hot}_L\left(\tilde{\mathbf{c}}_k\right) = \ell\_\mathrm{hot}_L\left(\mathrm{top}\_\ell\left(\boldsymbol{\psi}_k + \mathbf{e}_k\right)\right) \tag{6}$$

where the operation $\ell\_\mathrm{hot}_L(\cdot)$ returns a ℓ-hot vector of length L, $\mathrm{top}\_\ell(\cdot)$ works in similar way to the $\arg\max$ operator while it takes ℓ largest elements without replacement, and $\mathbf{e}_k \sim \mathrm{Gumbel}\,(0, 1)$ follows a standard Gumbel distribution. The Gumbel distribution is commonly used to model the maximum of a set of independent samples. It is equivalent to twice the negative log-transform of an uniform distribution, i.e., $\mathbf{e}_k = -\log(-\log(\mathbf{u}_k))$, with $\mathbf{u}_k \sim \mathcal{U}(0, 1)$, which makes it easy to generate realizations of $\mathbf{e}_k$.

However, the operation $\ell\_\mathrm{hot}_L\left(\mathrm{top}\_\ell\left(\cdot\right)\right)$ is not differentiable, and thus prevents gradient backpropagation. To circumvent this issue, we relax it by converting it into ℓ $\arg\max$ operation that together create a ℓ-hot vector by sampling without replacement. Each $\arg\max(\cdot)$ operation is relaxed by applying the Gumbel-softmax trick [38]. The Gumbel-softmax representation of $\mathbf{m}_k$ is defined as

$$\begin{aligned}\mathbf{m}_k^{\text{Gumbel}} &= \underset{\tau}{\mathrm{softmax}}\left(\boldsymbol{\psi}_k + \mathbf{e}_k\right) \\ &= \frac{\exp\left(\left(\boldsymbol{\psi}_k + \mathbf{e}_k\right)/\tau\right)}{\sum_{j=1}^{L} \exp\left(\left(\psi_{k,j} + e_{k,j}\right)/\tau\right)},\end{aligned} \tag{7}$$

where τ is the temperature parameter that controls the degree of approximation. When $\tau > 0$, the Gumbel-softmax performs soft sampling, i.e., generates a continuous version of $\mathbf{m}_k$, which enables backpropagation during training. However, for the forward pass, $\mathbf{m}_k$ must be binary, to comply with constraint (4c), and thus to perform antenna selection. Theoretically, this is achieved as τ approaches zero. In practice, we employ an approach known as the straight-through Gumbel-softmax estimator [39], which allows us to set $\tau = 0$, perform hard sampling, and to finally obtain a valid ℓ-hot $\mathbf{m}_k$ vector for the forward pass. The detailed structure of the Probabilistic Antenna Switcher (PAS) block is shown in Fig. 1(b).

B. RFI-aware Probabilistic Antenna Switcher (RaPAS)

Regarding RFI mitigation, conventional thresholding-based [28, 29] and data-driven [30, 31] methods are computationally expensive, in particular for next-generation radio interferometers, as they involve analyzing all the visibilities for potential flagging. As an alternative, we propose to estimate the probability of RFI contamination at the antenna level before the selection process.

Assuming that RFI contamination is concentrated on a subset of antennas, we introduce a RFI processing step which assesses the likelihood of RFI contamination for each antenna. This block assigns a probability score to each antenna, indicating its likelihood of being affected by RFI. This probability is then integrated into the antenna selection process to optimize the design of the subarray switching pattern.

The input to the proposed RFI-aware network consists of covariance matrices derived from the received signals, obtained by rearranging the full-array visibilities $\mathbf{r}_k$. Specifically, the input data is formatted as a $n \times n \times 3$ real-valued tensor by stacking the real part, imaginary part, and element-wise norm of the covariance matrix, where n is the number of antennas in the full array. This representation effectively captures the essential statistical properties of the received signals, providing a comprehensive and informative input to the network.

The input tensor is processed through three 2D convolutional layers, each followed by a batch normalization layer and a Rectified Linear Unit (ReLU) activation function. These layers extract spatial and statistical features from the covariance matrix. To generate the final output, the network includes two fully connected (linear) layers. The last layer employs a sigmoid activation function to transform the output into probability values, each representing the likelihood of a specific antenna being affected by RFI.

To effectively utilize the predictions provided by the RFI-aware block, we combine this information with the PAS. Instead of entirely excluding antennas identified as RFI-contaminated, we incorporate the RFI probabilities into the selection process, weighting the antenna selection probabilities accordingly. For the sake of simplicity, we assume that the final selection probability for each antenna is computed as

$$\boldsymbol{p}_k = \boldsymbol{\psi}_{\mathrm{k}} \odot (\mathbf{1} - \boldsymbol{p}_k^{\mathrm{RFI}}), \tag{8}$$

where $\boldsymbol{p}_k \in \mathbb{R}^L$ represents the probability of antennas being selected by the PAS, $\boldsymbol{p}_k^{\mathrm{RFI}} \in \mathbb{R}^L$ denotes the probability of antennas being affected by RFI, and it should be noted that the Hadamard product $\odot$ is applied element-wise across the probability vectors of the antennas. After that step, we add the top-ℓ block to pick the ℓ antennas with highest selection probability in the antenna selection matrix $\mathbf{M}_k$. Figure. 1(b) illustrates the procedures of the proposed RaPAS.

By adopting this approach, we avoid the overly conservative strategy of outright discarding RFI-affected antennas. In some cases, an antenna may only be mildly affected by RFI, and excluding it completely could degrade the overall quality of the reconstructed image. By balancing the selection probabilities using both the PAS and the RFI-aware block, we enable a more nuanced compromise, allowing the PAS to optimize the final image quality while accounting for RFI presence. A pseudo-code implementation of the proposed PAS and RaPAS is provided in Algorithm 1, assuming a software framework capable of automatic differentiation.

C. Image reconstruction

The image reconstruction process generates an image from the measured visibilities. The quality of the reconstructed image is crucial in radio-astronomical imaging. Since image reconstruction algorithms require significant computational

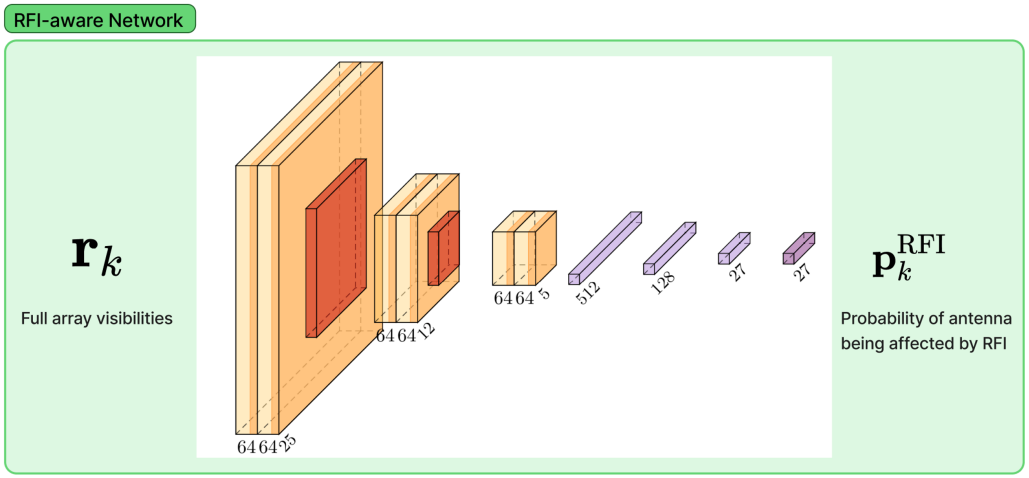


Fig. 2: Architecture of the RFI-aware network.

resources, reducing the total number of visibilities directly decreases the computational cost, highlighting the importance of designing an efficient antenna switching pattern.

Modern radio-interferometric imaging requires advanced reconstruction techniques to achieve the high dynamic range, angular resolution, and sensitivity demanded by contemporary instruments such as LOFAR, MeerKAT, and SKA. Beyond classical CLEAN-based approaches [40, 41], state-of-the-art methods leverage optimisation- and inference-driven approaches, including compressed-sensing formulations (e.g., SARA [42], MORESANE [43], FISTA-accelerated variants [44]), Bayesian imaging frameworks such as RESOLVE [45], and more recently deep-learning-based methods such as Polish [46] and R2D2 [47]. These techniques offer improved fidelity for extended emission, robustness to noise, and faster convergence.

In this work, we adopt ISTA as a baseline optimisation method to isolate and evaluate the contribution of the proposed PAS and switching strategy. While ISTA is a simple and well-known proximal-gradient algorithm, it is not representative of the most advanced imaging techniques currently used in radio astronomy. Its use here is therefore intended solely for proof-of-concept validation of the proposed strategy, rather than to achieve state-of-the-art reconstruction performance.

Let $\mathbf{M}$ denote the vertical concatenation of all $\mathbf{M}_k$, for $k = 1, 2, \ldots, K$,

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_1 \\ \mathbf{M}_2 \\ \vdots \\ \mathbf{M}_K \end{bmatrix} \in \{0,1\}^{nK \times N}.$$

with the same definition applied to $\mathbf{y}$ and $\mathbf{H}$. In radio-astronomical imaging, a basic reconstruction method applies the inverse Fourier transform to the selected visibilities, producing a preliminary result known as the dirty image

$$\mathbf{x}_{\text{dirty}} = \text{Re}\left((\mathbf{MH})^H \mathbf{y}\right). \tag{9}$$

It is important to stress that the dirty image is not used for scientific interpretation, as it does not provide an accurate reconstruction. By construction, it represents the convolution of the true image with the point spread function of the antenna array. It provides only an initial assessment of the quality of the selected antenna switching pattern.

Several optimization-based algorithms have been developed to refine the dirty image. To implement a more general imaging approach, we formulate the image reconstruction problem as a convex minimization problem [48]

$$\min_{\mathbf{x}} g(\mathbf{x}) + \alpha \mathcal{R}(\mathbf{x}), \tag{10}$$

where $g(\mathbf{x})$ enforces data fidelity, $\mathcal{R}(\mathbf{x})$ introduces sparsity in the reconstructed image, and α is a positive tuning parameter. A higher value of α promotes a sparser solution, leading to more elements of $\mathbf{x}$ being zero.

Considering the i.i.d. Gaussian random noise assumption in model (1), and choosing the data-fidelity term $g(\mathbf{x})$ as the MSE loss, which corresponds to the negative log-likelihood of the measurements, we obtain

$$g(\mathbf{x}) = \frac{1}{2}\|\mathbf{y} - \mathbf{MHx}\|_2^2. \tag{11}$$

For the regularization term $\mathcal{R}(\mathbf{x})$, following conventional practices in radio astronomy [49, 50], we use the ℓ_1-norm as the regularization function $\mathcal{R}(\mathbf{x}) = \|\mathbf{x}\|_1$. Thus, the image reconstruction problem (10) can be written as

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{MHx}\|_2^2 + \alpha\|\mathbf{x}\|_1. \tag{12}$$

A widely-used class of algorithms for solving the above image reconstruction problem are gradient-based methods. For example, the Iterative Soft-Thresholding Algorithm (ISTA) [51] which uses an iterative proximal scheme. The updating step for the reconstructed image is

$$\mathbf{x}^{(t+1)} = \underset{\alpha\gamma||\mathbf{x}||_1}{\text{prox}} \left(\mathbf{x}^{(t)} + \gamma\left(\mathbf{H}^H\mathbf{M}^H\boldsymbol{\varepsilon}^{(t)}\right)\right) \tag{13}$$

where $0 < \gamma < 2/\sigma_{\text{H}}$, and σ_{H} is the spectral radius of $\mathbf{H}^H\mathbf{M}^H\mathbf{MH}$. The residual at iteration (t) is defined as $\boldsymbol{\varepsilon}^{(t)} = \mathbf{y} - \mathbf{MHx}^{(t)}$. Here, the proximal operator is defined as

$$\underset{\alpha\gamma||\mathbf{x}||_1}{\text{prox}} (\mathbf{v})_i = \text{sgn}(\mathbf{v}_i)\left(|\mathbf{v}_i| - \alpha\gamma\right)_+ \tag{14}$$

with $(|\mathbf{v}_i| - \alpha\gamma)_+ = \max(0, |\mathbf{v}_i| - \alpha\gamma)$. This procedure efficiently updates $\mathbf{x}$ in an iterative manner, incorporating sparsity-inducing regularization with parameter α.

Regarding the complexity of the chosen image reconstruction method, we evaluate the computational complexity by measuring the number of floating-point operations (FLOPs). For the ISTA algorithm, the update rule at each iteration is given by (13). By summing the computational costs of the two steps for computing $\mathbf{e}^{(t)}$ and $\mathbf{x}^{(t)}$, the total number of FLOPs required is given by

$$n_{\text{ite,swit}}(K^2Nmn + KNm + K^2Nmn + K^2Nn)$$

where $n_{\text{ite,swit}}$ represents the number of iterations in the ISTA algorithm with the switching strategy. Given that the number of selected visibilities $n \gg 1$ and the number of image pixels $m \gg 1$, this expression can be approximated as $2n_{\text{ite,swit}}K^2Nmn$.

Following the same statement, the complexity of ISTA with full array (without switching strategy) will be $2n_{\text{ite,full}}K^2N^2m$, where $n_{\text{ite,full}}$ represents the number of iterations in the ISTA algorithm with full array. Thus, the ratio of computational complexity is $\frac{n_{\text{ite,swit}}n}{n_{\text{ite,full}}N}$. The computational savings achieved with the switching strategy will be given in details in the next section.

IV. Numerical Evaluation

In this section, we present numerical results to assess the performance of the proposed approach using PAS and RaPAS with synthetic and real radio astronomical images.

A. Implementation details

The simulations are based on the Very Large Array (VLA) [52], a well-known radio-astronomical telescope comprising 27 antennas arranged in three branches, each containing nine antennas. The VLA supports four configurations (A, B, C, and D) that provide varying resolution levels across different frequency bands. For this study, we employ the A configuration and conduct narrowband observations at a frequency channel of 300 MHz over a 4 h observation period.

The training datasets consist of true images discretized into 64×64 pixels, corresponding to a field of view (FOV) $3.16''$ (arc-second). These images are then vectorized, resulting in $\mathbf{x} \in \mathbb{R}^{4096}$ in the model (1). For each discretized true image $\mathbf{x}$, the corresponding true visibilities $\mathbf{r}_k$, where $k = 1, \ldots, 4$, are generated to represent four observation blocks ($K = 4$). To simulate real-world conditions, additive white Gaussian noise with zero mean and variance σ^2 is introduced into the visibilities. The noise variance σ^2 is kept consistent across all observation blocks.

The architecture of the RFI-aware block is illustrated in Fig. 2. Specifically, the model begins with three 2-D convolutional layers: the first maps the 2 input features to 64 channels using a 3×3 kernel, followed by batch normalization, ReLU activation, and 2×2 max pooling. A second 64-channel convolution block (3×3 kernel, batch normalization, ReLU, and max pooling) further reduces the spatial dimension. A third 64-channel convolution is applied with padding of 1 to preserve the 5×5 resolution, followed again by batch normalization and ReLU. The resulting feature maps are flattened into a 64×5×5 dimensional vector, which is passed through two fully connected layers: the first with 512 units and ReLU activation, and the second with 128 units and ReLU activation. A final linear layer projects the features to a vector of size 27, providing the model's output, per-antenna contamination probabilities.

As a way of assessing the fidelity of the reconstructed image and the performance of our proposed antenna switcher, we use two metrics. The first is the normalized mean square error (NMSE) defined as $\|\widehat{\mathbf{x}} - \mathbf{x}\|^2/\|\mathbf{x}\|^2$, where $\widehat{\mathbf{x}}$ denotes the reconstructed image and $\mathbf{x}$ represents the ground truth image.

Algorithm 1: Probabilistic Antenna Switcher (PAS) and RFI-aware PAS (RaPAS)

Input: Training dataset $\mathcal{D}$; number of images in the dataset $n_{\mathcal{D}}$; temperature parameter τ; initial parameters $\boldsymbol{\psi}_k$; number of observational snapshots K; number of ISTA iterations n_{iter}; Trained RFI-detection network $\mathcal{N}^{\text{RFI}}$.

Output: Trained logits vectors $\{\boldsymbol{\psi}_k\}_{k=1}^K$.

for $i \leftarrow 1$ **to** $n_{\mathcal{D}}$ **do**
  Sample a mini-batch $\mathbf{x}_i$ from $\mathcal{D}$;
  for $k \leftarrow 1$ **to** K **do**
    `// Forward model: simulate snapshot measurements`
    Simulate measurement: $\mathbf{r}_k = \mathbf{H}_k\mathbf{x}_i + \mathbf{n}_k$;
    `// RFI-aware or standard logits`
    Draw Gumbel noise $\mathbf{e}_k \sim \text{Gumbel}(0, 1)$;
    if *RaPAS is enabled* **then**
      Obtain RFI probabilities from the RFI-detect module $\mathbf{p}_k^{\text{RFI}} = \mathcal{N}^{\text{RFI}}(\mathbf{r}_k)$;
      Compute RFI-aware logits: $\mathbf{p}_k = \boldsymbol{\psi}_k \odot (\mathbf{1} - \mathbf{p}_k^{\text{RFI}})$;
    else
      Set $\mathbf{p}_k = \boldsymbol{\psi}_k$;
    end
    `// Gumbel-Softmax relaxation`
    $\mathbf{m}_k^{\text{G}} = \text{softmax}((\mathbf{p}_k + \mathbf{e}_k)/\tau)$;
    `// Hard top-`ℓ` selection (non-differentiable)`
    $\mathbf{m}_k^{\text{hard}} = \text{top\_}\ell(\mathbf{m}_k^{\text{G}})$;
    `// Straight-through estimator`
    $\mathbf{m}_k = \mathbf{m}_k^{\text{hard}}.\text{detach}() + \mathbf{m}_k^{\text{G}} - \mathbf{m}_k^{\text{G}}.\text{detach}()$;
    `// Construct switching matrices`
    $\mathbf{M}_k = \zeta(\text{diag}(\mathbf{m}_k \otimes \mathbf{m}_k))$;
    Apply switching: $\mathbf{y}_k = \mathbf{M}_k\mathbf{r}_k$;
  end
  `// Image reconstruction using ISTA`
  for $t \leftarrow 1$ **to** n_{iter} **do**
    $\mathbf{x}^{(t+1)} = \text{prox}_{\alpha\gamma\|\cdot\|_1}\left(\mathbf{x}^{(t)} + \gamma\mathbf{H}^H\mathbf{M}^H\boldsymbol{\varepsilon}^{(t)}\right)$;
  end
  Compute MSE loss: $L_i = \|\mathbf{x}_i - \widehat{\mathbf{x}}_i\|_2^2$;
  Update PAS logits $\{\boldsymbol{\psi}_k\}$ using Adam optimizer;
end

NMSE measures pixel-wise error, making it particularly sensitive to misalignments or translations of the image. Since we are concerned here with small distortions to the galaxy shapes, we additionally employ a more perception-based metric known as the Structural Similarity Index Measure (SSIM) [53]. Unlike NMSE, SSIM compares various windows within an image to assess the statistics of those regions, rather than directly computing absolute differences in pixels. The SSIM is defined as:

$$\text{SSIM}\left(\mathbf{x}, \widehat{\mathbf{x}}\right) = \frac{(2\mu_{\mathbf{x}}\mu_{\widehat{\mathbf{x}}} + c_1)(2\sigma_{\mathbf{x}\widehat{\mathbf{x}}} + c_2)}{(\mu_{\mathbf{x}}^2 + \mu_{\widehat{\mathbf{x}}}^2 + c_1)(\sigma_{\mathbf{x}}^2 + \sigma_{\widehat{\mathbf{x}}}^2 + c_2)}, \tag{15}$$

where $\mu_{\mathbf{x}}$ and $\mu_{\widehat{\mathbf{x}}}$ are the local mean values of the ground truth and reconstructed images, respectively, computed over a $l_s \times l_s$ neighborhood. $\sigma_{\mathbf{x}}^2$ and $\sigma_{\widehat{\mathbf{x}}}^2$ are the local variances, and $\sigma_{\mathbf{x}\widehat{\mathbf{x}}}$ is the local covariance. $c_1 = (k_1 D)^2$ and $c_2 = (k_2 D)^2$ are constants to stabilize the division, where D is the dynamic range of pixel values, and k_1 and k_2 are user-defined parameters.

For our implementation, we use the default parameters provided by the PIQA Python library: $k_1 = 0.01$, $k_2 = 0.03$, and a window size of 7×7. These settings ensure stable and reliable SSIM computation. We use both NMSE and SSIM to evaluate the performance of PAS.

B. Training

We train the proposed PAS using synthetic sky images. The training images are composed of Gaussian ellipsoids, which corresponds to the prior we have in the images. The number of sources in each sky image is sampled from a Poisson distribution with a mean of 11 sources per image. The fluxes of the sources are drawn from a uniform distribution in the range 0 to $10\,\text{Jansky}$ (Jy). The additive Gaussian noise level was fixed to $\sigma = 1$, corresponding to an input SNR of approximately 30 dB. Each image used $K = 4$ snapshot blocks. The dynamic range of the ground-truth images ranged between 200 and 250, which is the contrast between the brightest and faintest pixels.

During training, the dataset with 5000 images is shuffled at the beginning of each epoch to ensure robust learning. The generated dataset is split into $80\,\%$ for training and $20\,\%$ for cross validation. The function $d(\cdot, \cdot)$ in (4a), used to evaluate the distance between the reconstructed and true images, is the mean square error loss.

To demonstrate the superiority of the proposed PAS over the recent CRB-BTB switching pattern [26], we conduct simulations with a total of $L = 27$ antennas. We evaluate the performance for different numbers of selected antennas, specifically $\ell = 9, 12, 15, 18$. To ensure a fair comparison, we also design $K_{\text{swit}} = 2$ different subarray configurations, denoted as $\{\mathbf{M}_{i,K_{\text{swit}}}\}, i = 1, \ldots, K_{\text{swit}}$. It is important to note that, in practice, up to K switching patterns can be designed, meaning that each observation can have a corresponding (often different) switching pattern. In our setup, for the sake of simplicity and in order to compare the proposed scheme with the CRB-BTB switching scheme, the sequence of subarray configurations corresponding to $\mathbf{M}_1 \to \mathbf{M}_2 \to \mathbf{M}_3 \to \mathbf{M}_4$ is given by

$$\underbrace{\mathbf{M}_{1,2}}_{k=1} \to \underbrace{\mathbf{M}_{2,2}}_{k=2} \to \underbrace{\mathbf{M}_{1,2}}_{k=3} \to \underbrace{\mathbf{M}_{2,2}}_{k=4}$$

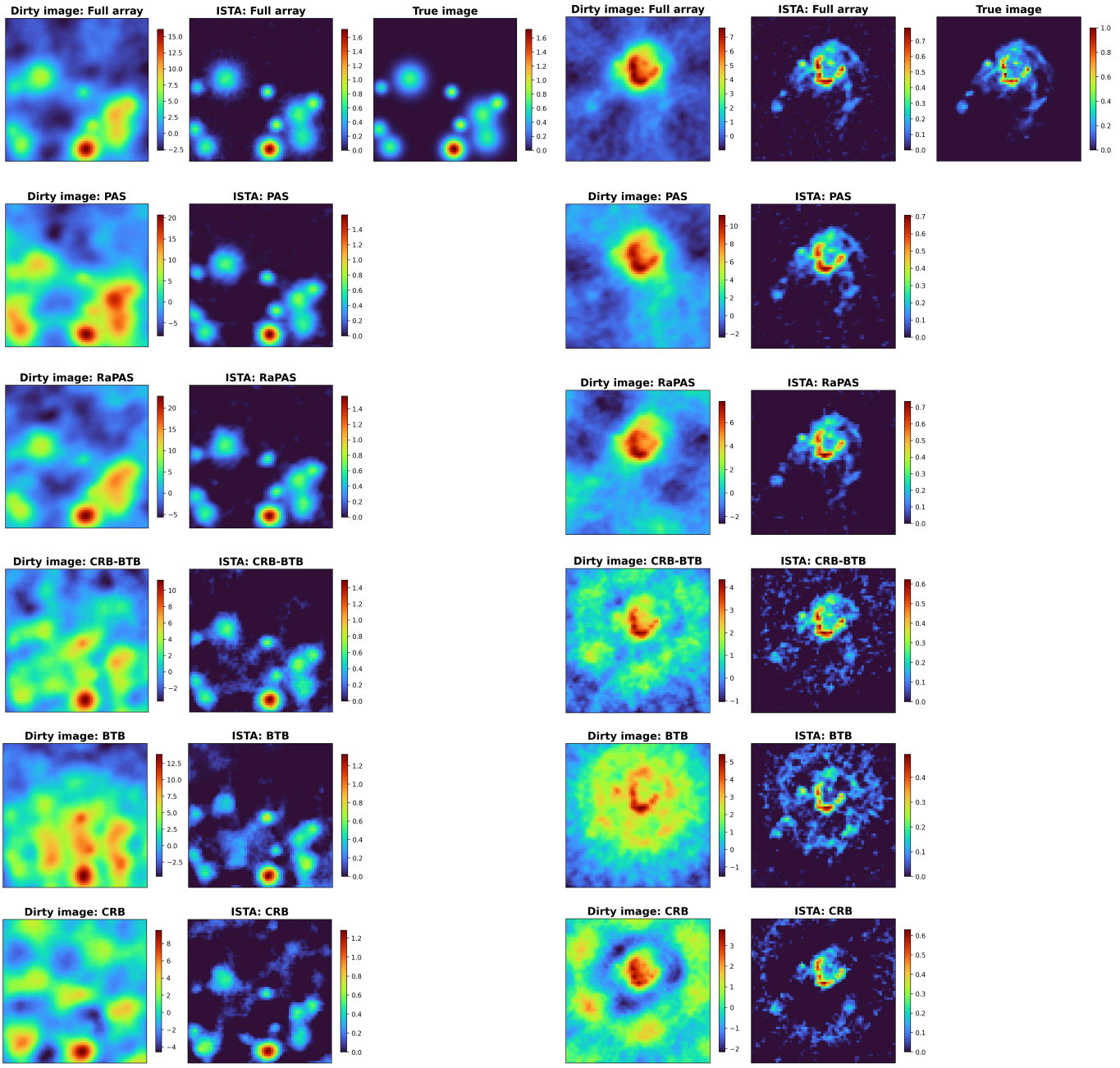
Each switching pattern $\{\mathbf{M}_{i,K_{\text{swit}}}\}, i = 1, \ldots, K_{\text{swit}}$ has a PAS trained within the framework depicted in Fig. 1. As a common choice in radio astronomical imaging problems [48], the step size in the ISTA image reconstruction algorithm (13) is set to $\gamma = 1.98/\sigma_{\text{H}}$, where σ_{H} is the largest eigenvalue of $\mathbf{H}^H\mathbf{M}^H\mathbf{M}\mathbf{H}$, as previously stated. It should be noted that σ_{H} depends on the current switching pattern $\{\mathbf{M}_{i,K_{\text{swit}}}\}, i = 1, \ldots, K_{\text{swit}}$. The regularisation parameter α was selected empirically using a small validation set of synthetic images. The chosen value provided a good trade-off between noise suppression and structural fidelity. Once selected, α was kept fixed during both PAS training and evaluation to ensure consistency across all experiments. We used $\alpha = 100$. The number of iteration of ISTA is fixed to $n_{\text{ite}} = 10$. The entire training process is conducted using the Adam optimizer implemented in PyTorch, with a learning rate of 1×10^{-3}. Training converges within 100 epochs.

A high value of temperature τ in eq (7) produces smooth probability distributions and stable gradients, whereas a low τ yields sharper, near-discrete selections. In our implementation, we employ an exponential annealing schedule starting from $\tau_0 = 1$ and decaying to a minimum value of $\tau_{\min} = 0.01$ over the course of training. This strategy allows PAS to first explore a wide range of candidate subarrays before progressively converging toward harder selections. We also conducted a sensitivity study and observed that PAS remains robust to moderate changes in the schedule: excessively high temperatures slow convergence, while prematurely low temperatures can destabilise the gradient flow. The adopted schedule therefore offers a well-balanced trade-off between stability and selection sharpness.

C. Simulation results

1) Simulation 1: performance without RFI: This simulation evaluates the performance of the proposed PAS under Gaussian white noise, which simulates thermal noise added to the visibilities. The reconstruction performance is compared across the methods: (1) **Full array**: Using all antennas without selection. (2) **Random switching**: Randomly selecting two subarray configurations. (3) **CRB subarray**: Without switching. The CRB is computed following [10], and the resulting subarray design favours antennas located at the extremities of the array branches. (4) **BTB subarray**: Without switching. (5) **CRB-BTB switching**: Recent subarray selection techniques [17]. (6) **PAS switching**: The proposed deep learning-based probabilistic antenna switcher. (7) **RaPAS switching**: The proposed RFI-aware PAS.

An example comparison of reconstructed synthetic images under thermal noise selecting $\ell = 15$ antennas for different methods is shown in Fig. 3(a). Quantitative reconstruction performance is evaluated using NMSE and SSIM. The results, averaged over 500 Monte Carlo simulations, are summarized



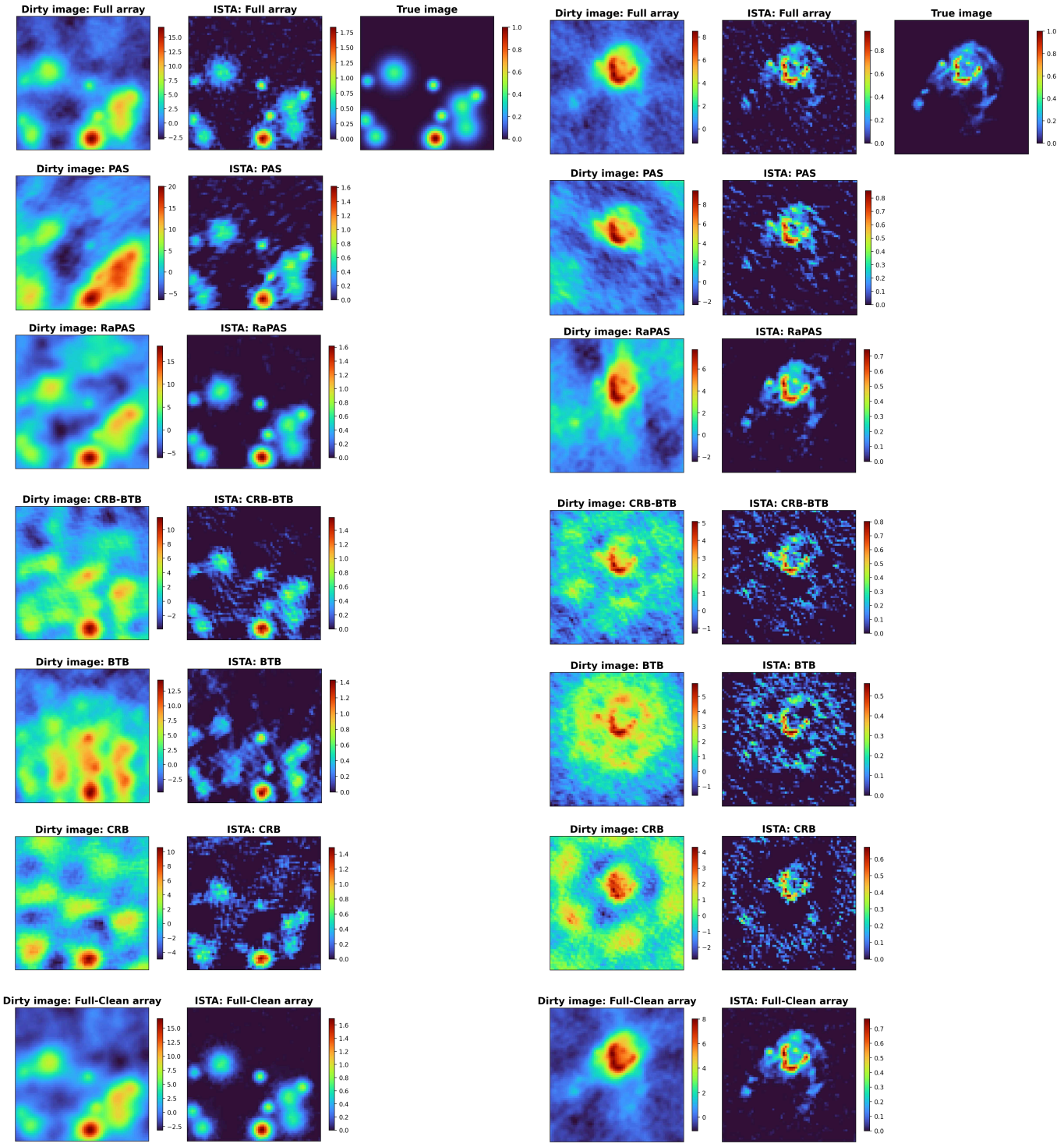
(a) Synthetic sky images

(b) Galaxy M31

Fig. 3: Different reconstructed images without RFI using the ISTA algorithm after 10 iterations with different methods selecting $\ell = 15$ antennas: **Full array (without switching)**: using all antennas without selection, **PAS switching (proposed)**, **RaPAS switching (proposed)**, **CRB-BTB switching**, **BTB subarray without switching**, and **CRB subarray without switching**. As expected, the RaPAS method does not affect the quality of antenna subarray switching or image reconstruction when only thermal noise is present.

in Fig. 5. We further validate the proposed PAS on a real radio-astronomical image of M31, focusing on the H2 region [43]. This image has been previously used to investigate radio-astronomy imaging algorithms [49, 50]. The reconstruction results with $\ell = 15$ antennas are presented in Fig. 3(b). The proposed PAS consistently outperforms CRB-BTB-based switching patterns, achieving significantly lower NMSE and higher SSIM values. These results demonstrate the superior reconstruction capability of the PAS approach. Notably, for RaPAS, the additional RFI-aware block does not degrade the performance of PAS in the presence of thermal noise alone.

2) Simulation 2: Performance under strong RFI: To simulate the impact of strong RFI noise, we add extreme noise to two randomly selected antennas in the full 27-antenna VLA array, in addition to Gaussian white noise. Statistically, the amplitude of the RFI noise exceeds three standard deviations



(a) Synthetic sky images

(b) Galaxy M31

Fig. 4: Different reconstructed images with RFI contamination using the ISTA algorithm after 10 iterations with different methods selecting $\ell = 15$ antennas: **Full array** (utilizing all antennas, including RFI-affected ones), **PAS switching** (proposed), **RaPAS switching** (proposed), **CRB-BTB switching**, **BTB subarray without switching**, **CRB subarray without switching** and **Full-Clean array** (excluding all RFI-affected antennas as a reference). As expected, the PAS method alone does not perform well in terms of image reconstruction quality when RFI is present.

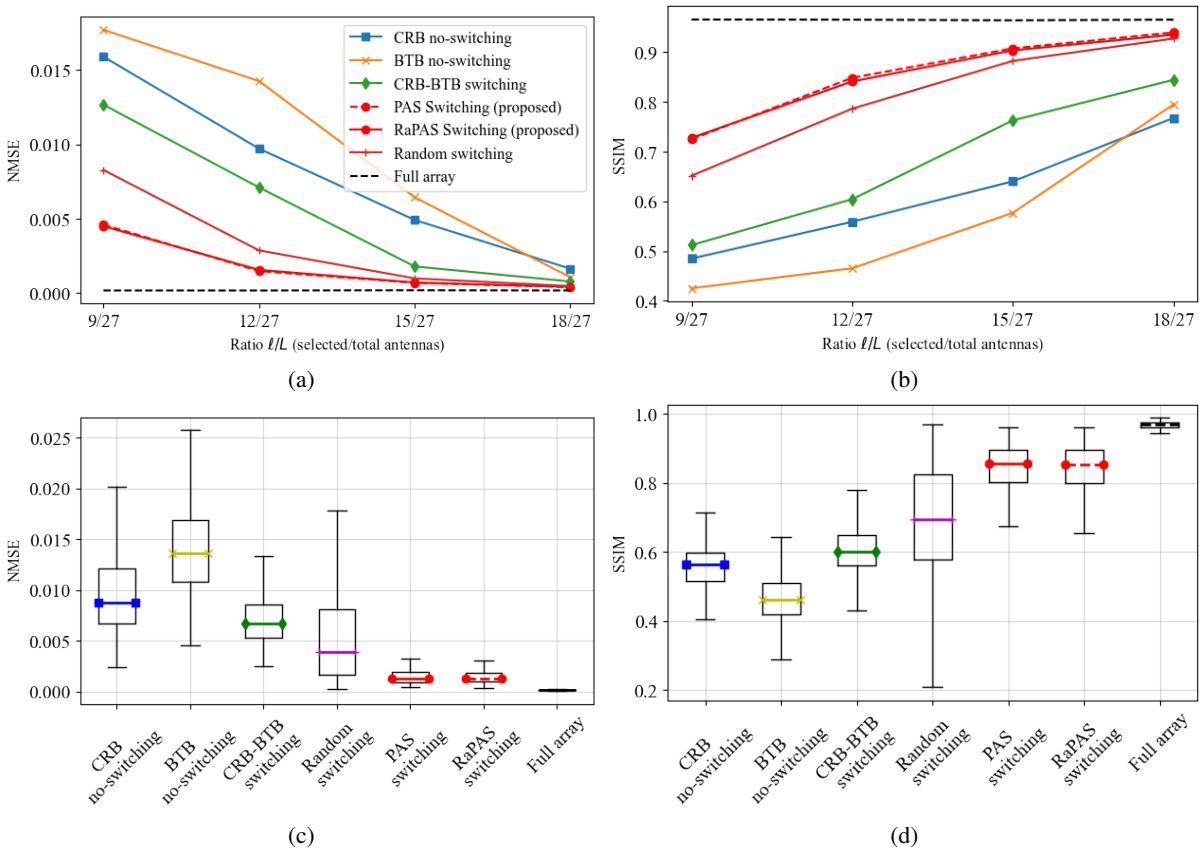


Fig. 5: Reconstruction performance: (a) NMSE and (b) SSIM of synthetic images in the absence of RFI, evaluated for different numbers of selected antennas: $\ell = 9, 12, 15, 18$ out of $L = 27$ total antennas. The ISTA algorithm was run for 10 iterations, and results were averaged over 500 Monte Carlo simulations. (c) and (d) show the boxplots of NMSE and SSIM, respectively, comparing different methods for $\ell = 12$ selected antennas. As expected, the RaPAS method does not affect the quality of antenna subarray switching or image reconstruction when only thermal noise is present.

of the Gaussian noise, ensuring it remains distinguishable from the Gaussian noise distribution. This amplitude, referred to as the RFI factor, is set to 10. To evaluate the effectiveness of the RFI exclusion functionality, the two RFI-affected antennas are randomized for each switching block to test the robustness of the RFI exclusion functionality.

The training dataset for the RFI-aware block is derived from the previously described Gaussian dataset, with RFI noise added using the same RFI factor of 10. The RFI-aware block is trained in a supervised manner using the Binary Cross Entropy (BCE) loss, comparing the predicted probabilities of contamination with the ground-truth labels (1 for the contaminated antennas and 0 for the others). Training is performed in PyTorch with the Adam optimizer and a learning rate of 1×10^{-5}, and typically converges within 50 epochs. Once trained, the RFI-aware block is incorporated into the RaPAS framework. In addition to the previously described methods, a new case is included: **Full-Clean array**: Using all antennas without any RFI contamination.

An example comparison of reconstructed synthetic images corrupted by RFI selecting $\ell = 15$ antennas is shown in Fig. 4(a). Figure 6 summarized the NMSE and SSIM metrics, averaged over 500 Monte Carlo trials, for different reconstruction methods with synthetic images. The reconstructed M31 image using the proposed RaPAS method with $\ell = 15$ antennas is presented in Fig. 4(b). The proposed RaPAS method effectively mitigates the impact of RFI, resulting in significantly fewer artifacts in the reconstructed images. It achieves substantially lower NMSE and higher SSIM values compared to the Full array case (which includes RFI-affected antennas) and performs comparably with the Full-Clean array case (excluding RFI-affected antennas). These results demonstrate the superior RFI-mitigation capabilities of the RaPAS approach.

3) Summary: In the absence of RFI, the results in Fig. 5 show that for all configurations with different numbers of selected antennas, the proposed PAS strategy consistently outperforms all other tested methods. This advantage is particularly significant when the number of selected antennas is small compared to the total number of available antennas. In this regime, PAS demonstrates its potential, compared to the recently proposed CRB-BTB switching strategy in the literature.

In the presence of RFI contamination, as shown in Fig. 6, as expected, the proposed PAS alone is not sufficient to mitigate the interference. Its performance is limited by the use of

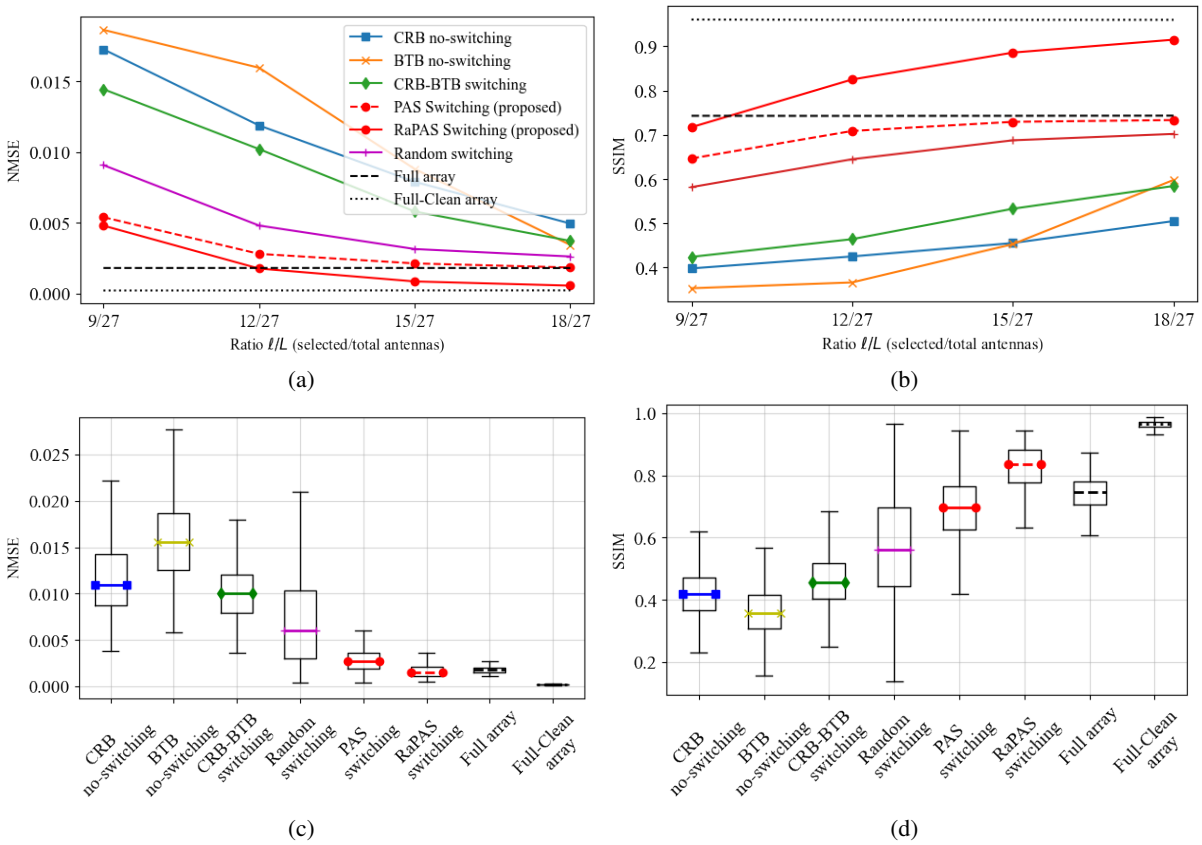


Fig. 6: Reconstruction performance: (a) NMSE and (b) SSIM of synthetic images under RFI contamination, evaluated for different numbers of selected antennas: $\ell = 9, 12, 15, 18$ out of $L = 27$ total antennas. The ISTA algorithm was run for 10 iterations, and results were averaged over 500 Monte Carlo simulations. (c) and (d) show the boxplots of NMSE and SSIM, respectively, comparing different methods for $\ell = 12$ selected antennas. As expected, the PAS method alone does not perform well in terms of image reconstruction quality when RFI is present.

the full array, which includes the contaminated antennas. In contrast, the proposed RaPAS method effectively handles RFI. From a moderate number of selected antennas (e.g., 12 out of 27 in our simulation), RaPAS shows strong performance. Remarkably, from 12 selected antennas onward, RaPAS can outperform the full-array configuration (27 antennas, including 2 contaminated ones).

Regarding the random selection method, although its average performance is acceptable, it suffers from high variance, as shown in Fig. 5(c), Fig. 5(d), Fig. 6(c) and Fig. 6(d). This large variability means that, in practice, it cannot be considered a reliable or practical strategy.

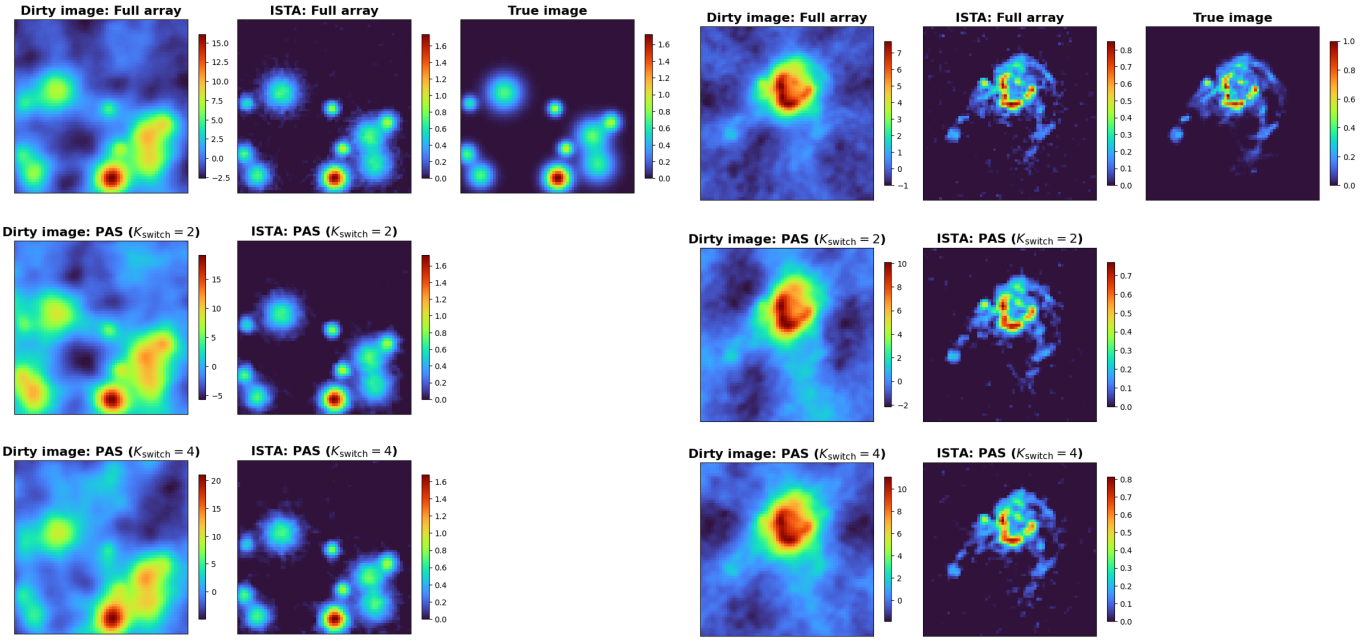
D. Energy reduction

To analyze the reduction in energy consumption achieved by our proposed method, we consider two aspects: the antenna level and the Science Processing Center (SPC) level. By reducing the number of active antennas, power consumption can be directly decreased, either by deactivating antennas or switching them to a low-power mode (reducing power consumption to less than 5% of their nominal consumption) [54]. Since fewer antennas are in use, fewer measurements (visibilities) are generated, leading to reduced computational requirements for imaging algorithms. This, in turn, lowers the energy consumption at the SPC level.

For the case of $\ell = 15$ selected antennas, the computational load at the antenna level reduces to $\frac{\ell}{L} = \frac{15}{27} = 55.5\%$ of the initial consumption, yielding a 44.5% energy saving. At the scientific computing level, which concerns imaging algorithms, assuming the same number of ISTA iterations for both the switching and full array cases, the computational load ratio becomes $\frac{n}{N} = \frac{\ell(\ell-1)}{L(L-1)} = 29.9\%$, implying a 70% energy saving. Furthermore, this computational reduction is accompanied by a decrease in hardware requirements due to the reduction in data flow, proportional to the visibility ratio, i.e., $\frac{n}{N} = 29.9\%$, which represents a 70% reduction in data flow.

E. Extension to the case $K_{swit} > 2$

The proposed PAS framework is compatible with an arbitrary number of switching configurations per snapshot, denoted K_{swit}. Each configuration is learned independently through a dedicated logits vector $\boldsymbol{\psi}_k$, and the switching masks $\{\mathbf{m}_k\}_{k=1}^{K_{\text{swit}}}$ are applied sequentially to the corresponding measurements. In the above part, we primarily consider $K_{\text{swit}} = 2$ as a representative setting, to do a direct comparison with the bound-based deterministic switching patterns.



(a) Synthetic images (b) Galaxy M31

Fig. 7: Reconstructed images in the absence of RFI using the ISTA algorithm after 10 iterations with different methods selecting $\ell = 15$ antennas on (a) Synthetic images and (b) Galaxy M31: **Full array**, **PAS switching** ($K_{\text{swit}} = 2$), and **PAS switching** ($K_{\text{swit}} = 4$).

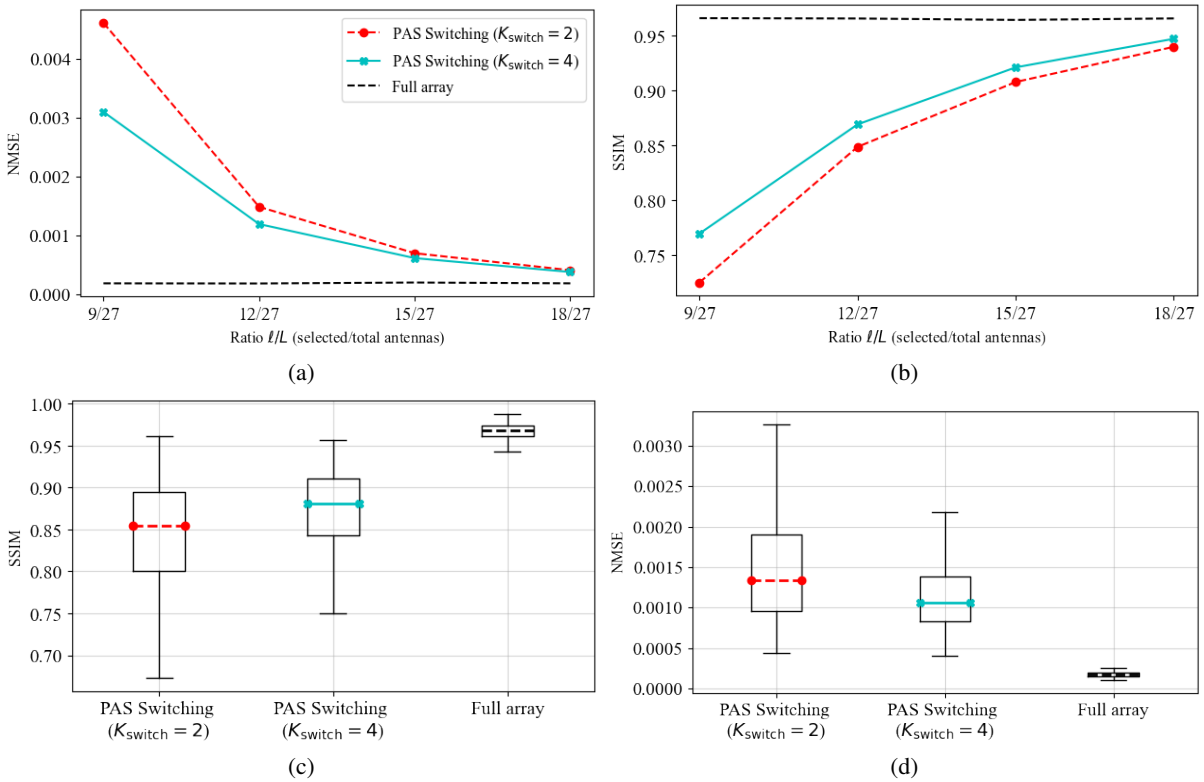


Fig. 8: Reconstruction performance of the PAS switching method for different values of K_{swit} : (a) NMSE and (b) SSIM of synthetic images in the absence of RFI, evaluated for different numbers of selected antennas: $\ell = 9, 12, 15, 18$ out of $L = 27$ total antennas. The ISTA algorithm was run for 10 iterations, and results were averaged over 500 Monte Carlo simulations. (c) and (d) show the boxplots of NMSE and SSIM, respectively, comparing different methods for $\ell = 12$ selected antennas.

To assess the generalisability of the PAS framework, we additionally evaluated PAS with $K_{\text{swit}} = 4$ in the absence of RFI, while keeping all training conditions identical to those used for $K_{\text{swit}} = 2$, as illustrated in Fig. 5. To facilitate a clear comparison, we do not replicate the results of all other methods in Fig. 5; instead, we retain only the PAS results for $K_{\text{swit}} = 2$ and $K_{\text{swit}} = 4$, together with the full-array baseline.

An example of reconstructed synthetic images and M31 for the PAS $K_{\text{swit}} = 2$ and $K_{\text{swit}} = 4$ is shown in Fig. 7. Quantitative reconstruction performance on synthetic data is evaluated using NMSE and SSIM. The results, averaged over 500 Monte Carlo simulations and summarized in Fig. 8, show that the proposed PAS switching strategy continues to yield consistent improvements in both NMSE and SSIM as the number of switching configurations increases (from $K_{\text{swit}} = 2$ to 4). This behaviour is consistent with the design of PAS, as additional switching patterns increase the diversity of sampled baselines without requiring modifications to the underlying optimisation pipeline.

Overall, the proposed method scales naturally to $K_{\text{swit}} > 2$, and these results further support the generalisability of PAS to more complex acquisition settings. It should also be noted that as the number of selected antennas ℓ increases, particularly when it exceeds half of the total number of antennas, the performance gain achieved by switching progressively diminishes.

V. Conclusions

This paper presents a deep learning framework for optimizing antenna subarray switching patterns while considering the image reconstruction process in radio-astronomical imaging. The proposed probabilistic antenna switcher (PAS) is enhanced with a RFI processing network, which assesses the likelihood of RFI contamination for each antenna and incorporates this information into the selection process. To evaluate the performance of the proposed PAS and RFI-aware PAS (RaPAS), we considered a switching pattern between two subarrays. We assessed their effectiveness under both thermal noise and RFI noise. The results demonstrate that PAS achieves superior reconstruction quality compared to the recent CRB-BTB-based switching patterns. Additionally, RaPAS effectively mitigates the impact of RFI noise. The findings also highlight the ability of our approach to reduce computational resources compared to using the full array.

Future work will focus on relaxing the assumption of perfectly calibration measurement model. A natural extension is to replace the imaging block with a joint calibration and imaging method [55]. In addition, future studies will expose PAS to more diverse and realistic simulation conditions to ensure reliable performance across varying uv-coverage patterns and noise levels. Finally, replacing the RFI detection module in RaPAS with predictive inference models of interference could enable online antenna selection, thereby achieving acquisition-stage reduction, without requiring full visibility computation.

References

[1] A. Schutz, A. Ferrari, D. Mary, É. Thiébaut, and F. Soulez, "Large scale 3D image reconstruction in optical interferometry," in *2015 23rd European Signal Processing Conference (EUSIPCO)*. IEEE, 2015, pp. 474–478.

[2] S. Vijay Kartik, R. E. Carrillo, J.-P. Thiran, and Y. Wiaux, "A Fourier dimensionality reduction model for big data interferometric imaging," *Monthly Notices of the Royal Astronomical Society*, vol. 468, no. 2, pp. 2382–2400, 2017.

[3] D. H. T. Minh and Y.-N. Ngo, "ComSAR: A new algorithm for processing Big Data SAR Interferometry," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*. IEEE, 2021, pp. 820–823.

[4] X. Wang, Y. Guo, F. Wen, J. He, and T.-K. Truong, "EMVS-MIMO radar with sparse Rx geometry: Tensor modeling and 2D direction finding," *IEEE Transactions on Aerospace and Electronic Systems*, 2023.

[5] H. Wang, L. Xu, Z. Yan, and T. A. Gulliver, "Low-complexity MIMO-FBMC sparse channel parameter estimation for industrial big data communications," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 5, pp. 3422–3430, 2020.

[6] A.-J. Boonstra, S. J. Wijnholds, S. van der Tol, and B. Jeffs, "Calibration, sensitivity and RFI mitigation requirements for LOFAR," in *Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, vol. 5. IEEE, 2005, pp. v–869.

[7] W. A. Baan, "Implementing rfi mitigation in radio science," *Journal of Astronomical Instrumentation*, vol. 8, no. 01, p. 1940010, 2019.

[8] S. Wei, H. Zhang, X. Zeng, Z. Zhou, J. Shi, and X. Zhang, "CARNet: An effective method for SAR image interference suppression," *International Journal of Applied Earth Observation and Geoinformation*, vol. 114, p. 103019, 2022.

[9] M. Juhlin and A. Jakobsson, "Optimal sensor placement for localizing structured signal sources," *Signal Processing*, p. 108679, 2022.

[10] J. Wang, L. Bacharach, P. Larzabal, and M. N. El Korso, "A comparison of antenna placement criteria based on the cramér-rao and barankin bounds for radio interferometer arrays," *Signal Processing*, p. 109404, 2024.

[11] H. Gazzah and J. P. Delmas, "CRB-based design of linear antenna arrays for near-field source localization," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 4, pp. 1965–1974, 2014.

[12] J. P. Delmas, M. N. El Korso, H. Gazzah, and M. Castella, "CRB analysis of planar antenna arrays for optimizing near-field source localization," *Signal Processing*, vol. 127, pp. 117–134, 2016.

[13] M. N. El Korso, A. Renaux, R. Boyer, and S. Marcos, "Deterministic performance bounds on the mean square error for near field source localization," *IEEE transactions on signal processing*, vol. 61, no. 4, pp. 871–877, 2012.

[14] J. Tabrikian, O. Isaacs, and I. Bilik, "Cognitive antenna selection for automotive radar using bobrovsky-zakai bound," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 4, pp. 892–903, 2021.

[15] R. McAulay and E. Hofstetter, "Barankin bounds on parameter estimation," *IEEE Transactions on Information Theory*, vol. 17, no. 6, pp. 669–676, 1971.

[16] M. Morelande and B. Ristic, "Signal-to-noise ratio threshold effect in track before detect," *IET radar, sonar & navigation*, vol. 3, no. 6, pp. 601–608, 2009.

[17] J. Wang, L. Bacharach, M. N. El Korso, and P. Larzabal, "Antenna sub-array switching strategy for low-cost & high-resolution radio-astronomical imaging," in *2023 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*. IEEE, 2023, pp. 1–6.

[18] J. Wang, M. N. El Korso, L. Bacharach, and P. Larzabal, "Low-rank EM-based imaging for large-scale switched interferometric arrays," *IEEE Signal Processing Letters*, 2024.

[19] H. Sun, A. V. Dalca, and K. L. Bouman, "Learning a probabilistic strategy for computational imaging sensor selection," in *2020 IEEE International Conference on Computational Photography (ICCP)*. IEEE, 2020, pp. 1–12.

[20] C. Tasse, "Nonlinear Kalman filters for calibration in radio interferometry," *Astronomy & Astrophysics*, vol. 566, p. A127, 2014.

[21] S. Salvini and S. J. Wijnholds, "Fast gain calibration in radio astronomy using alternating direction implicit methods: Analysis and applications," *Astronomy & Astrophysics*, vol. 571, p. A97, 2014.

[22] A. Offringa, A. De Bruyn, S. Zaroubi, and M. Biehl, "A LOFAR RFI detection pipeline and its first results," *arXiv preprint arXiv:1007.2089*, 2010.

[23] G. Hellbourg, R. Weber, K. Abed-Meraim, and A.-J. Boonstra, "RFI spatial processing at Nancay observatory: Approaches and experiments," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2014, pp. 5387–5391.

[24] R. Levanda and A. Leshem, "Adaptive selective sidelobe canceller beam-

former with applications to interference mitigation in radio astronomy," *IEEE Transactions on Signal Processing*, vol. 61, no. 20, pp. 5063–5074, 2013.

[25] G. Sharma and A. A. B. Raj, "Low complexity interference mitigation technique in IRCI-free SAR imaging algorithms," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.

[26] J. Wang, M. El Korso, L. Bacharach, and P. Larzabal, "RFI-aware and low-cost maximum likelihood imaging for high-sensitivity radio telescopes," *IEEE Signal Processing Letters*, 2024.

[27] Y. Mhiri, M. N. El Korso, A. Breloy, and P. Larzabal, "Regularized maximum likelihood estimation for radio interferometric imaging in the presence of radiofrequency interferences," *Signal Processing*, vol. 220, p. 109430, 2024.

[28] A. Offringa, J. Van De Gronde, and J. Roerdink, "A morphological algorithm for improving radio-frequency interference detection," *Astronomy & astrophysics*, vol. 539, p. A95, 2012.

[29] W. A. Baan, P. Fridman, and R. Millenaar, "Radio frequency interference mitigation at the westerbork synthesis radio telescope: Algorithms, test observations, and system implementation," *The Astronomical Journal*, vol. 128, no. 2, p. 933, 2004.

[30] J. Akeret, C. Chang, A. Lucchi, and A. Refregier, "Radio frequency interference mitigation using deep convolutional neural networks," *Astronomy and computing*, vol. 18, pp. 35–39, 2017.

[31] R.-Q. Yan, C. Dai, W. Liu, J.-X. Li, S.-Y. Chen, X.-C. Yu, S.-F. Zuo, and X.-L. Chen, "Radio frequency interference detection based on the ac-unet model," *Research in Astronomy and Astrophysics*, vol. 21, no. 5, p. 119, 2021.

[32] A. M. Elbir and K. V. Mishra, "Sparse array selection across arbitrary sensor geometries with deep transfer learning," *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, no. 1, pp. 255–264, 2020.

[33] I. A. Huijben, B. S. Veeling, K. Janse, M. Mischi, and R. J. van Sloun, "Learning sub-sampling and signal recovery with applications in ultrasound imaging," *IEEE Transactions on Medical Imaging*, vol. 39, no. 12, pp. 3955–3966, 2020.

[34] J. Sauder, M. Genzel, and P. Jung, "Gradient-based learning of discrete structured measurement operators for signal recovery," *IEEE Journal on Selected Areas in Information Theory*, vol. 3, no. 3, pp. 481–492, 2022.

[35] S. M. Xie and S. Ermon, "Reparameterizable subset sampling via continuous relaxations," in *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, 2019, pp. 3919–3925.

[36] W. Kool, H. Van Hoof, and M. Welling, "Stochastic beams and where to find them: The Gumbel-Top-K trick for sampling sequences without replacement," in *International Conference on Machine Learning*. PMLR, 2019, pp. 3499–3508.

[37] T. Plötz and S. Roth, "Neural nearest neighbors networks," *Advances in Neural information processing systems*, vol. 31, 2018.

[38] E. Jang, S. Gu, and B. Poole, "Categorical Reparametrization with Gumble-Softmax," in *International Conference on Learning Representations (ICLR 2017)*. OpenReview. net, 2017.

[39] Y. Bengio, N. Léonard, and A. Courville, "Estimating or propagating gradients through stochastic neurons for conditional computation," *arXiv preprint arXiv:1308.3432*, 2013.

[40] J. Högbom, "Aperture synthesis with a non-regular distribution of interferometer baselines," *Astronomy and Astrophysics Supplement, Vol. 15, p. 417*, vol. 15, p. 417, 1974.

[41] A. Offringa, B. McKinley, N. Hurley-Walker, F. Briggs, R. Wayth, D. Kaplan, M. Bell, L. Feng, A. Neben, J. Hughes *et al.*, "Wsclean: an implementation of a fast, generic wide-field imager for radio astronomy," *Monthly Notices of the Royal Astronomical Society*, vol. 444, no. 1, pp. 606–619, 2014.

[42] R. E. Carrillo, J. D. McEwen, and Y. Wiaux, "Sparsity averaging reweighted analysis (sara): a novel algorithm for radio-interferometric imaging," *Monthly Notices of the Royal Astronomical Society*, vol. 426, no. 2, pp. 1223–1234, 2012.

[43] A. Dabbech, C. Ferrari, D. Mary, E. Slezak, O. Smirnov, and J. S. Kenyon, "Moresane: Model reconstruction by synthesis-analysis estimators-a sparse deconvolution algorithm for radio interferometric imaging," *Astronomy & Astrophysics*, vol. 576, p. A7, 2015.

[44] H. Garsden, J. Girard, J.-L. Starck, S. Corbel, C. Tasse, A. Woiselle, J. McKean, A. S. Van Amesfoort, J. Anderson, I. Avruch *et al.*, "Lofar sparse image reconstruction," *Astronomy & astrophysics*, vol. 575, p. A90, 2015.

[45] P. Arras, H. L. Bester, R. A. Perley, R. Leike, O. Smirnov, R. Westermann, and T. A. Enßlin, "Comparison of classical and bayesian imaging in radio interferometry-cygnus a with clean and resolve," *Astronomy & Astrophysics*, vol. 646, p. A84, 2021.

[46] L. Connor, K. L. Bouman, V. Ravi, and G. Hallinan, "Deep radio-interferometric imaging with polish: Dsa-2000 and weak lensing," *Monthly Notices of the Royal Astronomical Society*, vol. 514, no. 2, pp. 2614–2626, 2022.

[47] A. Aghabiglou, C. San Chu, A. Dabbech, and Y. Wiaux, "The r2d2 deep neural network series paradigm for fast precision imaging in radio astronomy," *The Astrophysical Journal Supplement Series*, vol. 273, no. 1, p. 3, 2024.

[48] M. Terris, A. Dabbech, C. Tang, and Y. Wiaux, "Image reconstruction algorithms in radio interferometry: From handcrafted to learned regularization denoisers," *Monthly Notices of the Royal Astronomical Society*, vol. 518, no. 1, pp. 604–622, 2023.

[49] R. E. Carrillo, J. D. McEwen, and Y. Wiaux, "PURIFY: a new approach to radio-interferometric imaging," *Monthly Notices of the Royal Astronomical Society*, vol. 439, no. 4, pp. 3591–3604, 2014.

[50] X. Cai, M. Pereyra, and J. D. McEwen, "Uncertainty quantification for radio interferometric imaging–I. Proximal MCMC methods," *Monthly Notices of the Royal Astronomical Society*, vol. 480, no. 3, pp. 4154–4169, 2018.

[51] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM journal on imaging sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[52] P. J. Napier, A. R. Thompson, and R. D. Ekers, "The very large array: Design and performance of a modern synthesis radio telescope," *Proceedings of the IEEE*, vol. 71, no. 11, pp. 1295–1320, 1983.

[53] R. Dosselmann and X. D. Yang, "A comprehensive assessment of the structural similarity index," *Signal, Image and Video Processing*, vol. 5, pp. 81–91, 2011.

[54] "SKA phase 1 system requirements specification," 2017. [Online]. Available: https://www.skao.int/sites/default/files/documents/d3-SKA-TEL-SKO-0000008-Rev11_SKA1SystemRequirementSpecification.pdf

[55] A. Dabbech, A. Repetti, R. A. Perley, O. M. Smirnov, and Y. Wiaux, "Cygnus a jointly calibrated and imaged via non-convex optimization from vla data," *Monthly Notices of the Royal Astronomical Society*, vol. 506, no. 4, pp. 4855–4876, 2021.