RFI-aware and low-cost maximum likelihood imaging for high-sensitivity radio telescopes

J. Wang*, M. N. El Korso‡, *Member, IEEE*, L. Bacharach*, P. Larzabal*, *Member, IEEE*

***Abstract*—This paper addresses the challenge of interference mitigation and reduction of computational cost in the context of radio interferometric imaging. We propose a novel maximum-likelihood-based methodology based on the antenna sub-array switching technique, which strikes a refined balance between imaging accuracy and computational efficiency. In addition, we tackle robustness regarding radio interference by modeling the additive noise as t-distributed. Through simulation results, we demonstrate the superiority of the t-distributed noise model over the conventional Gaussian noise model in scenarios involving interferences. We evidence that our proposed switching approach yields similar imaging performances with far fewer visibilities compared to the full array configuration, thus, diminishing the computational complexity.**

***Index Terms*—Expectation-Maximization algorithm, antenna array processing, interferometry, radio-astronomical imaging.**

I. Introduction

The rapid advancement of new-generation radio telescopes in the last decades, exemplified by projects like the Square Kilometer Array (SKA) [1, 2] employing over a hundred thousand antennas, aims for an era of unprecedented sensitivity and resolution in astronomical measurements. However, alongside this expansion in observational capacity arises a dual challenge: 1) 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 [3–6], and 2) the computational burden imposed by the large amount of data generated by the increasing number of antennas. This challenge is pressing, with estimates indicating that the processing demand for SKA Phase I can reach up to 100 Megawatts, split between the telescope itself and the data processing centers [7]. This poses a significant challenge in terms of energy consumption and cooling requirements, especially for telescopes situated in remote semi-desert regions [8]. These demands underscore the need for innovative approaches that can effectively mitigate interference phenomena, minimize energy consumption, and preserve imaging quality during processing, which forms the focal point of this paper.

First, to address the challenge of interferences, various RFI mitigation strategies have emerged. Broadly, these strategies can be categorized into two types: collaborative RFI mitigation [9, 10], which relies on prior information about the RFI source, and blind passive RFI mitigation. In this paper, we focus on the latter. Conventional approaches often rely on identifying and flagging strong outliers in the raw data, employing techniques such as manual inspection [11] or spatial filtering on the correlation matrix [12–14]. While effective for strong interferences, such methods struggle with weaker signals, increasingly prevalent with the growing number of sensors in large-scale interferometers [15]. Unlike thermal noise, which can be appropriately modeled using a Gaussian distribution, RFI originates from diverse sources such as communication systems or high voltage transmission lines, making it challenging to model accurately [16, 17]. Theoretical and empirical investigations have demonstrated the relevance of t-distribution in modeling interference [18, 19].

Second, to tackle the computational complexity challenge, efficient techniques for partitioning and leveraging antenna sub-arrays, dedicated to different types of measurement missions, have been proposed to balance the imperatives of accuracy and computational cost. Antenna placement strategies, based on optimization of estimation lower bounds, have been a focal point in the literature. Notably, methodologies leveraging criteria such as the Cramér-Rao bound (CRB) [20–23] and the Barankin-type bound (BTB) [24–27] have been proposed. By employing these criteria, the selection of antenna sub-arrays presents itself as a promising avenue for reducing the number of antennas utilized in the imaging process, thereby curtailing computational complexity and energy consumption. However, while these approaches demonstrate efficacy under specific conditions, their performances may vary depending on the application scenario. For instance, CRB-based sub-array selection tends to yield a beampattern with a thin mainlobe, whereas BTB-based methods provide a trade-off between mainlobe width and sidelobe level. Recognizing the complementary strengths of these methodologies, we propose a switching sub-array method that combines both approaches, rather than privileging one over the other depending on application scenarios. Our proposed methodology relies on dynamic switching between sub-arrays over time, exploiting the inherent correlation between them. This approach reduces the number of measurements (visibilities) required, significantly decreasing the computational complexity of the imaging process. Furthermore, by utilizing fewer antennas simultaneously, we minimize overall power consumption and cooling needs, contributing to a more cost-effective and energy-efficient solution for next-generation radio interferometers.

In this paper, we propose an antenna sub-arrays switching imaging algorithm to address the interference mitigation and high computational demand. A t-distribution noise model is employed to model radio interferences, leveraging its heavy-tailed structure to model the outliers and thereby improve the imaging quality compared to traditional additive Gaussian noise model. Due to the intractability of the direct maximum

* *SATIE,* ‡ *L2S, Université Paris-Saclay, 91190 Gif-sur-Yvette, France*

likelihood estimation method, the proposed strategy is formulated as an expectation-conditional maximization (ECM) algorithm.

II. Radio interferometric data model

Radio interferometers involve measuring the spatial coherence of the electric field between pairs of antennas in the antenna array. These measurements, known as visibilities, provide information about the electromagnetic radiation emitted by celestial sources under observation. The main idea of this paper is to dynamically switch the antenna sub-arrays during the observation process, which leads to a selection of q_k sensors from the whole array composed of q sensors for the k-th snapshot block. This results in a visibility vector is denoted by $\mathbf{r}_k^{\text{obs}} \in \mathbb{C}^{N_k=\frac{q_k(q_k-1)}{2}}$ (in contrast to the full visibility vector denoted by $\mathbf{r}_k \in \mathbb{C}^{N=\frac{q(q-1)}{2}}$). Specifically, in radio-interferometry the visibility $\mathbf{r}_k^{\text{obs}}$ is obtained by the vectorization of the upper triangular part of the sample covariance matrix of the observed signal at the sensor level, without the diagonal terms [1]. The (unobserved) full visibiliy vector reads

$$\mathbf{r}_k = \mathbf{H}\mathbf{x} + \mathbf{b}_k, \qquad k = 1, \ldots, K \tag{1}$$

with K denoting the total number of snapshot blocks and where $\mathbf{H} \in \mathbb{C}^{N\times n}$ is the forward operator, corresponding to a spatial discrete Fourier transform, symbol $\mathbf{x} \in \mathbb{R}^n$ represents the image vector of size n (number of image pixels) to be reconstructed, and $\mathbf{b}_k \in \mathbb{C}^N$ is the noise vector corresponding to the k-th observation block. The noise $\mathbf{b}_k$ is modeled by a t-distribution to take into account the presence of interferences, i.e., $[\mathbf{b}_k]_i = \frac{1}{\sqrt{\tau_{ki}}}[\mathbf{t}_k]_i, \ i = 1, \cdots, N$ with the speckles $\mathbf{t}_k \in \mathbb{C}^N$ being independent and identically distributed (i.i.d.) complex Gaussian random variables with zero mean and variance σ^2, i.e., $\mathbf{t}_k \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_N)$. The texture τ_{ki} follows a Gamma distribution with the shape and the inverse-scale parameters both $\nu/2$, i.e., $\tau_{ki} \sim \Gamma\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$, where the degree of freedom ν is a positive tuning parameter, which models the non-Gaussianity of the noise.

To assess the effectiveness of modeling visibility noise, we conducted a comparative analysis between fitting the noise with a t-distribution and a Gaussian distribution. Both distributions were applied to fit the simulated visibility noise including RFI. It is worth mentioning that the t-distribution offers a superior fit to the visibility noise, particularly in the presence of RFI. This observation highlights the relevance of the t-distribution in modeling visibility noise, especially when dealing with interferences (cf. supplementary materials, Fig.3).

Our objective is to estimate the image vector $\mathbf{x}$, the noise power σ^2, and the degree of freedom ν from the noisy observed visibilities $\mathbf{r}_k^{\text{obs}}$ with $k = 1, \cdots, K$. The log-likelihood function of the observed visibilities can be expressed as

$$\begin{aligned} L(\mathbf{x}, \sigma^2, \nu|\mathbf{r}^{\text{obs}}) &= \ln \prod_{k=1}^{K} p(\mathbf{r}_k^{\text{obs}}|\mathbf{x}, \sigma^2, \nu) \\ &= \sum_{k=1}^{K} \int \ln p(\mathbf{r}_k|\mathbf{x}, \sigma^2, \nu) d\mathbf{r}_k^{\text{miss}} \end{aligned} \tag{2}$$

where $\mathbf{r}^{\text{obs}}$ is the collection set of $\left\{\mathbf{r}_1^{\text{obs}}, \ldots, \mathbf{r}_K^{\text{obs}}\right\}$, and $\mathbf{r}_k^{\text{miss}}$ denoted the missing part of $\mathbf{r}_k$.

A regularization function $\mathcal{R}(\mathbf{x})$ is added to take into account the sparsity of image vector $\mathbf{x}$, leading to the penalized maximum likelihood estimation problem. Finally, The interferometric imaging problem can be formulated

$$(\hat{\mathbf{x}}, \hat{\sigma}^2, \hat{\nu}) = \underset{\mathbf{x}, \sigma^2, \nu}{\arg\max} \sum_{k=1}^{K} \int \ln p(\mathbf{r}_k|\mathbf{x}, \sigma^2, \nu) d\mathbf{r}_k^{\text{miss}} + \mathcal{R}(\mathbf{x}) \tag{3}$$

III. ECM-based imaging algorithm

The aforementioned regularized maximum likelihood problem is intractable due to the form of t-distribution and the missing part of visibilities, i.e, $\mathbf{r}_k^{\text{miss}}$. To overcome these obstacles, we propose employing a variant of the EM algorithm known as the ECM algorithm [28]. The latter consists in performing several E-step, calculating the conditional expectation of the complete-data log-likelihood, and CM-step, conditionally maximizing that expectation with respect to the model parameters, one unknown parameter at a time, while holding the others fixed. Given that we have three parameters to estimate, namely $\{\sigma^2, \mathbf{x}, \nu\}$ (which will be denoted in the sequel by $\boldsymbol{\theta}$), and the likelihood function is not separable with respect to these parameters, we adopt a sequential estimation approach. Specifically, the estimation of each parameter occurs with the other two parameters held fixed, leading to the use of the CM-step rather than the traditional M-step.

In our case, we consider $\{\mathbf{r}_k^{\text{obs}}, \mathbf{r}_k^{\text{miss}}, \tau_{ki}\}$ with $k = 1, \cdots, K$, $i = 1, 2, \ldots, N$ as the complete data space. The complete log-likelihood function can be written as

$$\begin{aligned} L(\boldsymbol{\theta}|\mathbf{r}^{\text{obs}}, \mathbf{r}^{\text{miss}}, \boldsymbol{\tau}) &= \sum_{k=1}^{K}\sum_{i=1}^{N_k} \ln \Gamma\left(\tau_{ki}|\frac{\nu}{2}, \frac{\nu}{2}\right) \\ &+ \sum_{k=1}^{K} \ln \mathcal{CN}\left(\mathbf{r}_k^{\text{obs}}, \mathbf{r}^{\text{miss}}|\mathbf{M}_k\mathbf{H}\mathbf{x}, \sigma^2\mathbf{M}_k\boldsymbol{\Omega}_k^{-1}\mathbf{M}_k^H\right) \end{aligned} \tag{4}$$

where $\boldsymbol{\Omega}_k = \text{diag}(\tau_{k1}, \cdots, \tau_{kN})$, and $\mathbf{M}_k = \mathcal{T}\left(\mathbf{P}_\text{k} \otimes \mathbf{P}_\text{k}\right) \in \mathbb{R}^{N_k\times N}$ with $\mathbf{P}_k$ the selection matrix sensors for the k-th snapshot block, constructed by selecting q_k sensors over q, and $\mathcal{T}(\cdot)$ is the operator that retains only the upper triangular terms during the vectorization of the sample covariance matrix.

A. E-step

Proposition 1. *Denoting $Q_1(\mathbf{x}, \sigma^2|\boldsymbol{\theta}^{(m)})$ the part of $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(m)})$ relative to σ^2 and $\mathbf{x}$, and $Q_2(\nu|\boldsymbol{\theta}^{(m)})$ the part of $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(m)})$ relative to ν, the E-step reads*

$$Q_1(\mathbf{x}, \sigma^2|\boldsymbol{\theta}^{(m)}) \propto -\sum_{k=1}^{K} N_k \ln(\sigma^2) - \frac{1}{\sigma^2}\sum_{k=1}^{K}\left(\mathbf{e}_k^H\mathbf{M}_k^T\widetilde{\boldsymbol{\Omega}}_k^{(m)}\mathbf{M}_k\mathbf{e}_k\right) \tag{5}$$

where $\widetilde{\boldsymbol{\Omega}}_k^{(m)} \in \mathbb{R}^{N_k\times N_k}$ is a diagonal matrix defined as $\widetilde{\boldsymbol{\Omega}}_k^{(m)} = \text{diag}\left(\widetilde{\tau_{k1}}^{(m)}, \ldots, \widetilde{\tau_{kN_k}}^{(m)}\right)$, $\mathbf{e}_k$ indicates the residual vector defined as $\mathbf{e}_k \triangleq \mathbf{r}_k - \mathbf{H}\mathbf{x}$, and

$$\begin{aligned} Q_2(\nu|\boldsymbol{\theta}^{(m)}) = &\sum_{k=1}^{K}\sum_{i=1}^{N_k} \frac{\nu}{2}\left(\widetilde{\ln(\tau_{ki})}^{(m)} - \widetilde{\tau_{ki}}^{(m)}\right) \\ &+ \left(\sum_{k=1}^{K} N_k\right)\left(\frac{\nu}{2}\ln\left(\frac{\nu}{2}\right) - \ln\left(\Gamma\left(\frac{\nu}{2}\right)\right)\right) \end{aligned} \tag{6}$$

$$\text{with} \quad \widetilde{\tau_{ki}}^{(m)} = \frac{\nu^{(m)}+1}{\nu^{(m)}+\sigma^{-2(m)}\left([\mathbf{M}_k\mathbf{e}_k^{(m)}]_i\right)^2} \tag{7}$$

$$\widetilde{\ln(\tau_{ki})}^{(m)} = \phi\left(\frac{\nu^{(m)}+1}{2}\right) - \ln\left(\frac{\nu^{(m)}+\sigma^{-2(m)}\left([\mathbf{M}_k\mathbf{e}_k^{(m)}]_i\right)^2}{2}\right) \tag{8}$$

where $\phi(\cdot)$ is the digamma function defined as the logarithmic derivative of the gamma function.

The proof can be found in the supplementary material.

B. CM-step

1) Update of σ^2 and $\mathbf{x}$: The maximization step consists in solving the optimization problem

$$\mathbf{x}^{(m+1)}, \sigma^{2(m+1)} = \arg\max_{\mathbf{x},\sigma^2}\left(Q_1(\mathbf{x},\sigma^2|\boldsymbol{\theta}^{(m)}) + \mathcal{R}(\mathbf{x})\right) \tag{9}$$

The resulting optimization problem can be solved alternatively to produce the updated estimated sky image $\mathbf{x}^{(m+1)}$ and the noise power $\sigma^{2(m+1)}$, which results in

$$\sigma^{2(m+1)} = \frac{1}{\sum_{k=1}^K N_k}\sum_{k=1}^K\left(\mathbf{e}_k^{(m)H}\mathbf{M}_k^T\widetilde{\boldsymbol{\Omega}}_k^{(m)}\mathbf{M}_k\mathbf{e}_k^{(m)}\right) \tag{10}$$

$$\mathbf{x}^{(m+1)} = \arg\min_{\mathbf{x}}\sum_{k=1}^K \mathbf{e}_k^H\mathbf{M}_k^T\widetilde{\boldsymbol{\Omega}}_k^{(m)}\mathbf{M}_k\mathbf{e}_k + \mathcal{R}(\mathbf{x}) \tag{11}$$

The ℓ_1-norm is used as the regularization function to promote the sparsity in the image vector $\mathbf{x}$, i.e. $\mathcal{R}(\mathbf{x}) = \alpha\|\mathbf{x}\|_1$, where α is a positive tuning parameter. A higher value of α promotes a sparser image vector $\mathbf{x}$, resulting in more elements of $\mathbf{x}$ being zero. We employ the ISTA algorithm [29] to solve (11), which gives

$$\mathbf{x}^{(m+1)} = \operatorname*{prox}_{\alpha||\mathbf{x}||_1}\left(\mathbf{x}^{(m)} + \sum_{k=1}^K \mathbf{H}^H\mathbf{M}_k^T\widetilde{\boldsymbol{\Omega}}_k^{(m)}\mathbf{M}_k\mathbf{e}_k\right) \tag{12}$$

and the proximal operator is defined as $\text{prox}_{\alpha||\mathbf{u}||_1}(\mathbf{u})_i = \text{sgn}(\mathbf{u}_i)\left(|\mathbf{u}_i|-\alpha\right)_+$ with $\left(|\mathbf{u}_i|-\alpha\right)_+ = \max(0, |\mathbf{u}_i|-\alpha)$.

2) Update of ν: The maximization step for the update of ν consists in solving the optimization problem, $\nu^{(m+1)} = \arg\max_\nu Q_2(\nu|\boldsymbol{\theta}^{(m)})$. It can be addressed by finding the critical points of the objective function. To do so, we differentiate $Q_2(\nu|\boldsymbol{\theta}^{(m)})$ with respect to ν, setting the derivative to zero, and then solving for ν. The derivative takes the form

$$\begin{aligned}\frac{\partial Q_2(\nu|\boldsymbol{\theta}^{(m)})}{\partial\nu} =& \frac{1}{2}\sum_{k=1}^K\sum_{i=1}^{N_k}\left(\widetilde{\ln(\tau_{ki})}^{(m)} - \widetilde{\tau_{ki}}^{(m)}\right)\\ &+\frac{1}{2}\left(\sum_{k=1}^K N_k\right)\left(\ln\left(\frac{\nu}{2}\right)+1-\phi\left(\frac{\nu}{2}\right)\right)\end{aligned} \tag{13}$$

By setting this derivative to zero and after some algebraic manipulations, the update of $\nu^{(m+1)}$ becomes the solution of

$$\begin{aligned}&\frac{1}{\sum_{k=1}^K N_k}\sum_{k=1}^K\sum_{i=1}^{N_k}\left[\phi\left(\frac{\nu^{(m)}+1}{2}\right) - \ln\left(\frac{\nu^{(m)}+1}{2}\right)\right]\\ &-\phi(\frac{\nu}{2}) + \ln\left(\frac{\nu}{2}\right) = \sum_{k=1}^K\sum_{i=1}^{N_k}\left[\frac{-\ln\left(\widetilde{\tau_{ki}}^{(m)}\right)+\widetilde{\tau_{ki}}^{(m)}}{\sum_{k=1}^K N_k}\right] - 1\end{aligned} \tag{14}$$

By noting that $\tau - \ln\tau \geq 1$, $\forall\tau > 0$, we obtain that the right hand side of the above equation is non-negative. In addition, the function $-\phi(\nu/2) + \ln(\nu/2)$ monotonically decreases in the interval $(0, +\infty)$, ensuring the uniqueness of the solution. We employ a one-dimensional search method to solve (14) for ν. In our context, the search interval for ν is restricted to $2 < \nu < 10$ because the complex t-distribution exhibits finite second-order moments for $\nu > 2$ and tends to a Gaussian for $\nu \approx 10$ as established by [30] (a summary of the proposed algorithm is provided in the supplementary materials Algo. 1).

IV. Numerical Evaluation

A. Simulation Setup

In this section, we employ the proposed algorithm and validate its effectiveness by conducting simulations in the context of radio-astronomical imaging. The simulations are based on the existing radio telescope VLA [31], which is composed of three branches with 9 antennas on each branch, resulting into a total of 27 antennas. The M31 galaxy serves as the true image, which is based on the H2 region in the M31 galaxy [32] and has been previously utilized for investigating radio-astronomical imaging algorithms [33]. The true image is discretized into 84×84 pixels, and then vectorized, resulting in $\mathbf{x} \in \mathbb{R}^{7056}$ in the system model (1). Then the true visibility is calculated from the discretized true image $\mathbf{x}$. To simulate thermal noise, additive white Gaussian noise with zero mean and a variance of σ^2 is incorporated into the visibilities. To introduce RFI, we model 100 RFI events as additional point sources with same power into the scene. 1% of the visibilities are affected by the RFI, reflecting the selective impact on sensors. The power of each RFI event is quantified relatively to the source power, as defined by $P_{\text{RFIdB}} = 10\log_{10}(P_{\text{RFI}}/P_0)$, where $P_0 = \|\mathbf{x}\|^2$ is the source power. The number of observation blocks K is set to 24. The performance of different algorithms is assessed using the normalized mean square error (NMSE) defined as $\|\widehat{\mathbf{x}} - \mathbf{x}\|^2/\|\mathbf{x}\|^2$ with $\widehat{\mathbf{x}}$ the reconstructed image and the structural similarity index (SSIM) [34].

B. Simulation Results

First, to validate the effectiveness of the t-distribution noise model, we conducted a comparative analysis between the non-switching version (obtained by setting $\mathbf{M}_k = \mathbf{I}_N$, $\forall k$) of the imaging algorithm proposed in this paper and a Gaussian-based approach, which assumes white Gaussian-distributed noise. The performance of the two imagers with full array VLA, in terms of the NMSE of the reconstructed images against different RFI noise level with 1000 Monte Carlo is shown in Table. I. The results indicate that the imaging algorithm based on the t-distribution noise model consistently outperforms that based on a Gaussian noise assumption.

Table I: NMSE of non-switching imagers for various RFI.

RFI level (in dB)	1	2	3	4	5
Full-VLA Gaussain	0.032	0.033	0.037	0.043	0.052
Full-VLA t-student	**0.029**	**0.029**	**0.029**	**0.029**	**0.029**

Next, to implement the proposed switching robust imaging algorithm, two specific sub-arrays will be used. Thus, in order to avoid poor selections and/or potential ambiguities in sub-array selection arising from random switching, we advocate for a method that toggles between two types of sub-arrays based on the minimization of mean square error (MSE) lower bounds, namely, the CRB and the BTB, building upon our previous work [25]. The CRB-based antenna selection method yields a beampattern with a thin mainlobe, while the BTB-based method provides a relevant trade-off between mainlobe width and sidelobe level. The switching approach aims at leveraging the benefits of both methods. These antenna sub-arrays (AS) are derived from the full VLA configuration comprising 27 antennas, and are designated as AS-CRB and AS-BTB based on the CRB and BTB criteria, respectively.

We evaluated the performance of the proposed switching approach alongside CRB-based and BTB-based sub-arrays, as well as the full VLA configuration. The NMSE and SSIM of reconstructed images obtained from 1000 Monte Carlo simulations under 5 dB RFI noise powers is given in Table II. Reconstructed images under 5 dB RFI noise power are showcased in Fig. 1. It should be noted that we set the convergence criteria for all methods to the same standard, defined as the stability of the reconstructed image ($||\mathbf{x}^{(m+1)} - \mathbf{x}^{(m)}||^2 \leq \epsilon = 10^{-4}$). As shown from the NMSE and SSIM, and by comparing directly the reconstructed images, the proposed switching method performs similarly to the full VLA and consistently outperforms both AS-CRB and AS-BTB, particularly under strong RFI conditions.

The gain of the computational complexity (for a large N) is $(\beta^2 + 3\beta)/4$ in which $\beta = N_k/N$ is ratio of visibility selection defined as number of selected visibilities divided by full visibilities. Table II details the number of antennas and visibilities required by each method. Notably, the switching method requires approximately one-third of the visibilities compared to the full array, and the actual runtime for 10 iterations is consistent with the ratio of selected visibilities, further highlighting the efficiency of the switching algorithm. Fig. 2(a) illustrates the evolution of the difference between the reconstructed images at each iteration until convergence condition is met ($\epsilon = 10^{-4}$). The switching method converges much faster than the CRB and BTB sub-arrays, while importantly maintaining a convergence speed close to that of the full VLA. To compare image quality after a fixed number of iterations, Fig. 2(b) showcases the reconstructed images after 10 iterations. Here, the switching method achieves image quality comparable to the full VLA, while surpassing both AS-CRB and AS-BTB. Importantly, a comparison of the images in Fig. 2(b) and Fig. 1 reveals that the image quality of the switching method after 10 iterations is similar to its converged state, indicating a significant reduction in computational time. In summary, the proposed switching sub-array approach effectively balances image quality with computational efficiency. It achieves image fidelity comparable to the full VLA while utilizing significantly fewer antennas and visibilities. Additionally, it converges faster and yields superior image quality than CRB and BTB-based sub-arrays.

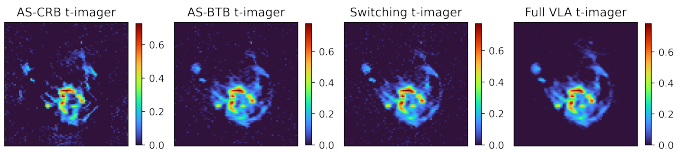
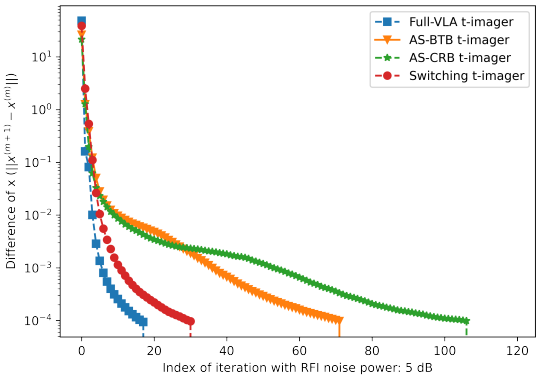
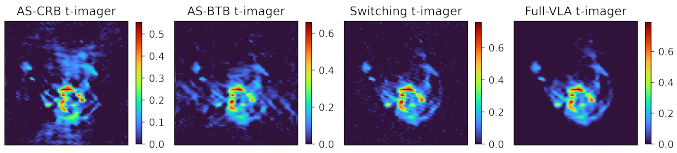


Fig. 1: Reconstructed image for RFI 5dB after satisfying the convergence condition.



(a)



(b)

Fig. 2: (a) Convergence of $\mathbf{x}^{(m)}$. (b) Reconstructed images after 10 iterations (RFI 5dB).

Table II: Number of antennas/visibilities required by each method, NMSE, SSIM, and the run time of 10 iterations of each method for RFI 5dB.

		Ant. No.	Vis. No.	NMSE	SSIM	Run-time
Non-switching	Array-Full	27	8424	0.0279	0.965	1.6098s
	Array-CRB	15	2520	0.2058	0.703	0.5465s
	Array-BTB	15	2520	0.0446	0.858	0.5502s
Switching method		**15**	**2520**	**0.0364**	**0.908**	**0.5553**s

V. Conclusions

This paper presents an antenna switching imaging algorithm for radio astronomy aiming to tackle the challenges of interference mitigation at reduced computational cost. We formulated the imaging as a maximum likelihood estimation problem, incorporating a t-distribution noise model within a sub-array switching framework. Leveraging this model, we derived an ECM algorithm for efficient image reconstruction. Simulations validated the superiority of the t-distribution model in the presence of RFI. The proposed switching algorithm achieved image quality comparable to the full-array VLA configuration while using significantly fewer visibilities and maintaining similar convergence times. These results demonstrate the efficacy of the proposed approach in achieving low-cost and robust radio-astronomical imaging.

References

[1] M. G. Labate, M. Waterson, B. Alachkar, A. Hendre, P. Lewis, M. Bartolini, P. Dewdney, Highlights of the square kilometre array low frequency (SKA-low) telescope, Journal of Astronomical Telescopes, Instruments, and Systems 8 (1) (2022) 011024–011024.

[2] G. P. Swart, P. E. Dewdney, A. Cremonini, Highlights of the SKA1-Mid telescope architecture, Journal of Astronomical Telescopes, Instruments, and Systems 8 (1) (2022) 011021–011021.

[3] A.-J. Boonstra, S. J. Wijnholds, S. van der Tol, 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.

[4] S. J. Wijnholds, S. van der Tol, R. Nijboer, A.-J. van der Veen, Calibration challenges for future radio telescopes, IEEE Signal Processing Magazine 27 (1) (2009) 30–42.

[5] W. A. Baan, Implementing RFI mitigation in radio science, Journal of Astronomical Instrumentation 8 (01) (2019) 1940010.

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

[7] P. J. Hall, Power considerations for the Square Kilometre Array (SKA) radio telescope, in: 2011 XXXth URSI General Assembly and Scientific Symposium, IEEE, 2011, pp. 1–4.

[8] A. Hotele, S. Winberg, L. Schwardt, Immersion cooling monitoring and prediction system for Meerkat radio astronomy imager, in: 2021 International Conference on Electrical, Communication, and Computer Engineering (ICECCE), IEEE, 2021, pp. 1–6.

[9] S. Chakraborty, D. Saha, A. Dutta, G. Hellbourg, LOCI: Learning Low Overhead Collaborative Interference Cancellation for Radio Astronomy, in: ICC 2023-IEEE International Conference on Communications, IEEE, 2023, pp. 6391–6396.

[10] S. Chakraborty, G. Hellbourg, M. Careem, D. Saha, A. Dutta, Collaboration with Cellular Networks for RFI Cancellation at Radio Telescope, IEEE Transactions on Cognitive Communications and Networking 9 (3) (2023) 765–778.

[11] A. Offringa, A. De Bruyn, S. Zaroubi, M. Biehl, A LOFAR RFI detection pipeline and its first results, Proceedings of Science 107 (2010).

[12] G. Hellbourg, R. Weber, K. Abed-Meraim, 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.

[13] R. Levanda, A. Leshem, Adaptive selective sidelobe canceller beamformer with applications to interference mitigation in radio astronomy, IEEE Transactions on Signal Processing 61 (20) (2013) 5063–5074.

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

[15] S. Tingay, M. Sokolowski, R. Wayth, D. Ung, A survey of spatially and temporally resolved radio frequency interference in the FM band at the Murchison Radio-astronomy Observatory, Publications of the Astronomical Society of Australia 37 (2020) e039.

[16] C. Tasse, Nonlinear Kalman filters for calibration in radio interferometry, Astronomy & Astrophysics 566 (2014) A127.

[17] S. Salvini, S. J. Wijnholds, Fast gain calibration in radio astronomy using alternating direction implicit methods: Analysis and applications, Astronomy & Astrophysics 571 (2014) A97.

[18] V. Ollier, M. N. El Korso, A. Ferrari, R. Boyer, P. Larzabal, Robust distributed calibration of radio interferometers with direction dependent distortions, Signal Processing 153 (2018) 348–354.

[19] E. Ollila, D. E. Tyler, V. Koivunen, H. V. Poor, Compound-Gaussian clutter modeling with an inverse Gaussian texture distribution, IEEE Signal Processing Letters 19 (12) (2012) 876–879.

[20] M. Juhlin, A. Jakobsson, Optimal sensor placement for localizing structured signal sources, Signal Processing (2022) 108679.

[21] H. Gazzah, J. P. Delmas, CRB-based design of linear antenna arrays for near-field source localization, IEEE Transactions on Antennas and Propagation 62 (4) (2014) 1965–1974.

[22] J. P. Delmas, M. N. El Korso, H. Gazzah, M. Castella, CRB analysis of planar antenna arrays for optimizing near-field source localization, Signal Processing 127 (2016) 117–134.

[23] M. N. El Korso, A. Renaux, R. Boyer, S. Marcos, Deterministic performance bounds on the mean square error for near field source localization, IEEE transactions on signal processing 61 (4) (2012) 871–877.

[24] J. Tabrikian, O. Isaacs, I. Bilik, Cognitive antenna selection for automotive radar using Bobrovsky-Zakai bound, IEEE Journal of Selected Topics in Signal Processing 15 (4) (2021) 892–903.

[25] J. Wang, L. Bacharach, P. Larzabal, M. N. El Korso, A comparison of antenna placement criteria based on the Cramér-Rao and Barankin bounds for radio interferometer arrays, Signal Processing (2024) 109404.

[26] R. McAulay, E. Hofstetter, Barankin bounds on parameter estimation, IEEE Transactions on Information Theory 17 (6) (1971) 669–676.

[27] M. Morelande, B. Ristic, Signal-to-noise ratio threshold effect in track before detect, IET radar, sonar & navigation 3 (6) (2009) 601–608.

[28] C. Liu, D. B. Rubin, ML estimation of the t distribution using EM and its extensions, ECM and ECME, Statistica Sinica (1995) 19–39.

[29] A. Beck, M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, SIAM journal on imaging sciences 2 (1) (2009) 183–202.

[30] E. Ollila, D. E. Tyler, V. Koivunen, H. V. Poor, Complex elliptically symmetric distributions: Survey, new results and applications, IEEE Transactions on signal processing 60 (11) (2012) 5597–5625.

[31] P. J. Napier, A. R. Thompson, R. D. Ekers, The very large array: Design and performance of a modern synthesis radio telescope, Proceedings of the IEEE 71 (11) (1983) 1295–1320.

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

[33] R. E. Carrillo, J. D. McEwen, Y. Wiaux, PURIFY: a new approach to radio-interferometric imaging, Monthly Notices of the Royal Astronomical Society 439 (4) (2014) 3591–3604.

[34] R. Dosselmann, X. D. Yang, A comprehensive assessment of the structural similarity index, Signal, Image and Video Processing 5 (2011) 81–91.

Supplementary materials

A. Comparison of t-distribution fitting and Gaussian distribution fitting

The following figure displays a comparison of t-distribution fitting and Gaussian distribution fitting. Initially, clean visibilities were derived from a specific sky image. Subsequently, additive white Gaussian noise with zero mean and a variance of σ^2 was introduced to simulate thermal noise. Additionally, several RFI events were simulated as additional point sources with equal power to represent RFI noise, affecting 1% of the visibilities to reflect the selective impact on sensors. This process resulted in the simulated visibility noise.

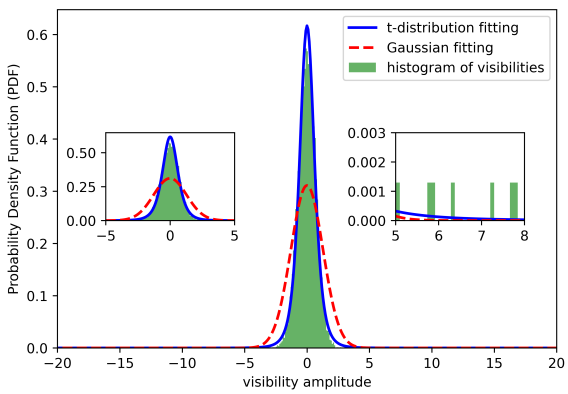


Fig. 3: Comparison of t-distribution fitting and Gaussian distribution fitting.

B. Proof of Proposition 1

Proof. Considering the data model defined in (1), the complete data log-likelihood is expressed as:

$$Q_1\left(\mathbf{x},\sigma^2|\boldsymbol{\theta}^{(m)}\right)=\mathbb{E}_{\tau|\mathbf{r}^{\text{obs}},\boldsymbol{\theta}^{(m)}}\left(L(\boldsymbol{\theta}|\mathbf{r}^{\text{obs}},\mathbf{r}^{\text{miss}},\boldsymbol{\tau})\right) \quad (15)$$

By substituting the complete data log-likelihood function with (4) and disregarding terms not dependent on σ^2, we derive

$$\begin{aligned}Q_1(\mathbf{x},\sigma^2|\boldsymbol{\theta}^{(m)}) &\propto \sum_{k=1}^{K}\mathbb{E}_{\tau|\mathbf{r}_k^{\text{obs}},\boldsymbol{\theta}^{(m)}}(-\ln|\sigma^2\mathbf{M}_k\boldsymbol{\Omega}_k^{-1}\mathbf{M}_k^H| \\ &-\mathbf{e}_k^H\mathbf{M}_k^T\left(\sigma^2\mathbf{M}_k\boldsymbol{\Omega}_k^{-1}\mathbf{M}_k^H\right)^{-1}\mathbf{M}_k\mathbf{e}_k) \\ &\propto \sum_{k=1}^{K}\left(-\ln|\sigma^2\mathbf{I}_{N_k}|+\ln|\widetilde{\boldsymbol{\Omega}}_k^{(m)}|-\mathbf{e}_k^H\mathbf{M}_k^T\frac{\widetilde{\boldsymbol{\Omega}}_k^{(m)}}{\sigma^2}\mathbf{M}_k\mathbf{e}_k\right) \\ &\propto \sum_{k=1}^{K}\left(-N_k\ln(\sigma^2)-\mathbf{e}_k^H\mathbf{M}_k^T\frac{\widetilde{\boldsymbol{\Omega}}_k^{(m)}}{\sigma^2}\mathbf{M}_k\mathbf{e}_k\right)\end{aligned}$$

where

$$\widetilde{\boldsymbol{\Omega}}_k^{(m)}=\mathbb{E}_{\tau|\mathbf{r}_k^{\text{obs}},\boldsymbol{\theta}^{(m)}}\left(\left(\mathbf{M}_k\boldsymbol{\Omega}_k^{-1}\mathbf{M}_k^H\right)^{-1}\right) \quad (16)$$

The calculation of $\widetilde{\boldsymbol{\Omega}}_k^{(m)}$ involves determining the conditional distribution of τ_{ki} given $\mathbf{r}_k$, that is a Gamma distribution [28]

$$\tau_{ki}|\mathbf{r}_k\sim\text{Gamma}\left(\frac{\nu+1}{2},\frac{\nu+\sigma^{-2}[\mathbf{r}_k-\mathbf{H}\mathbf{x}]_i^2}{2}\right) \quad (17)$$

with expectation

$$\mathbb{E}_{\tau_{ki}|\mathbf{r}_k}(\tau_{ki})=\frac{\nu+1}{\nu+\sigma^{-2}[\mathbf{r}_k-\mathbf{H}\mathbf{x}]_i^2},\ i=1,\dots,N \quad (18)$$

Considering the fact that $\boldsymbol{\Omega}_k$ is a diagonal matrix and incorporating the effect of the selection matrix $\mathbf{M}_k$, we have $\left(\mathbf{M}_k\boldsymbol{\Omega}_k^{-1}\mathbf{M}_k^H\right)^{-1}=\left(\mathbf{M}_k\boldsymbol{\Omega}_k\mathbf{M}_k^H\right)$ and thus we obtain

$$[\widetilde{\boldsymbol{\Omega}}_k^{(m)}]_{ii}=\widetilde{\tau_{ki}}^{(m)}=\frac{\nu^{(m)}+1}{\nu^{(m)}+\sigma^{-2(m)}([\mathbf{M}_k\mathbf{e}_k^{(m)}]_i)^2} \quad (19)$$

As for the part of ν, based on the complete-data log-likelihood function (4) and keeping only the terms dependent on ν, we derive

$$Q_2(\nu|\boldsymbol{\theta}^{(m)})=\sum_{k=1}^{K}\mathbb{E}_{\tau|\mathbf{r}^{\text{obs}},\sigma^{2(m)}}\left(\sum_{i=1}^{N_k}\left(\ln\Gamma\left(\tau_{ki}|\frac{\nu}{2},\frac{\nu}{2}\right)\right)\right) \quad (20)$$

Considering the definition of the Gamma function $p(\tau_{ki}|\nu)=\frac{\left(\frac{\nu}{2}\right)^{\frac{\nu}{2}}}{\Gamma(\frac{\nu}{2})}\tau_{ki}^{\frac{\nu}{2}-1}e^{-\frac{\nu}{2}\tau_{ki}}$, the surrogate function can be developed as

$$\begin{aligned}Q_2(\nu|\boldsymbol{\theta}^{(m)}) &= \sum_{k=1}^{K}\sum_{i=1}^{N_k}\frac{\nu}{2}\ln\left(\frac{\nu}{2}\right)-\widetilde{\ln(\tau_{ki})}^{(m)} \\ &\quad -\ln(\Gamma\left(\frac{\nu}{2}\right))+\frac{\nu}{2}\left(\widetilde{\ln(\tau_{ki})}^{(m)}-\widetilde{\tau_{ki}}^{(m)}\right) \\ &\propto \sum_{k=1}^{K}\sum_{i=1}^{N_k}\frac{\nu}{2}\ln\left(\frac{\nu}{2}\right)-\ln(\Gamma\left(\frac{\nu}{2}\right)) \\ &\quad +\frac{\nu}{2}\left(\widetilde{\ln(\tau_{ki})}^{(m)}-\widetilde{\tau_{ki}}^{(m)}\right)\end{aligned}$$

with

$$\widetilde{\tau_{ki}}^{(m)}\triangleq\mathbb{E}_{\tau_{ki}|\mathbf{r}_k^{\text{obs}},\sigma^{2(m)}}(\tau_{ki})$$

and

$$\widetilde{\ln(\tau_{ki})}^{(m)}\triangleq\mathbb{E}_{\tau_{ki}|\mathbf{r}_k^{\text{obs}},\sigma^{2(m)}}(\ln(\tau_{ki}))$$

defined as in (7) and (8).

This ends the proof. □

C. ECM antenna switching imager

In the following table we outline the proposed ECM antenna switching imager

Algorithm 1: ECM antenna switching imager

Input: $\mathbf{H}$, α, and $\mathbf{r}_k^{\text{obs}}$, $k=1,\dots,K$
Output: Estimates of $\hat{\mathbf{x}}$, $\hat{\sigma}^2$ and $\hat{\nu}$
Initialisation: $\mathbf{x}^{(m)}\leftarrow\mathbf{x}^{(0)}$, $(\sigma^2)^{(m)}\leftarrow(\sigma^2)^{(0)}$, and $\nu^{(m)}\leftarrow\nu^{(0)}$

while *stop criterion not met* **do**
  Calculate $\mathbf{e}_k^{(m+1)}$ and $\widetilde{\boldsymbol{\Omega}}_k^{(m)}$ using (7) for $k\in[1,K]$ `// E-step`
  Update $\sigma^{2(m+1)}$ using (10) `// CM-step`
  Update $\mathbf{x}^{(m+1)}$ using (12) `// CM-step`
  Update $\nu^{(m+1)}$ using (14) `// CM-step`