# A Robust EM Algorithm for Radio Interferometric Imaging in The Presence of Outliers

Yassine Mhiri*, Mohammed Nabil El Korso†, Arnaud Breloy† and Pascal Larzabal*
* SATIE, Université Paris Saclay, 91190 Gif-Sur-Yvette, France
† Paris-Nanterre University/LEME, 92410 Ville d'Avray, France
{yassine.mhiri}{pascal.larzabal}@universite-paris-saclay.fr,
{m.elkorso}{abreloy}@parisnanterre.fr

***Abstract*—Image synthesis in the context of radio interferometric data can be expressed as a signal reconstruction from incomplete Fourier measurements. Most imaging techniques for radio interferometry lie in minimizing the least square error between the reconstructed image and the observed data assuming an additive white gaussian noise. In this paper, we derive an expectation-maximization based imaging algorithm that handles the presence of outliers in the observed data. Subsequently, we propose a new generic image synthesis algorithm based on the expectation-maximization algorithm, leading to a computationally efficient method.**

***Index Terms*—Radioastronomy, expectation-maximization, image restauration**

## I. Introduction

Radio interferometers produce images of the sky from the correlation between multiple sensors, bringing new insights into various scientific domains such as solar monitoring, planetology, and astrophysics. A new generation of large-scale radio interferometers, such as the LOFAR or the SKA, are designed with the promise to improve further the resolution and sensitivity of radio astronomical images over large bandwidth using multiple sensors, resulting in a large collecting area and high-resolution imaging. Nevertheless, these technological advances bring several signal processing challenges to exploit the scientific potential of such instruments fully. As an example, the SKA will gather a large amount of data which will have to be processed in parallel. It will thus be necessary to process and store the useful data with tight computational time and storing constraints.

Image synthesis can be expressed as a signal reconstruction from incomplete Fourier measurements and leads to an ill-posed inverse problem [1], [2]. The particularity of radio interferometric data lie in the coverage of the Fourier space (UV-coverage) that depends on the sensors coordinates as well as the observation time. In particular, large-scale radio interferometers will improve the UV coverage, increasing the amount of data to process. The resulting deconvolution problem can be solved using greedy algorithms and has led to numerous methods based on the CLEAN algorithm [3]. On the other hand, methods based on convex optimization and the theory of compressed sensing are developed, assuming a sparse representation of the image in the wavelet domain [2], [4], [5]. Both methods are closely related to the maximum likelihood estimate of the image assuming visibilities probed under Gaussian noise. Sparse representation of the signal has been at the heart of radio interferometric imaging synthesis algorithm. In fact, from the very early days of interferometry and the Högbom CLEAN algorithm [6] to the actual multiscale CLEAN [3] used in current radio interferometers, CLEAN methods that assume sparsity in the pixels coefficients have been of undoubtful efficiency. Furthermore, the rise of compressed sensing (CS) in the early 2000 has brought a new paradigm, considering explicit sparse prior in the wavelet domain of the sky image [2], [5], [7]. The Expectation-Maximization (EM) algorithm is a versatile tool for maximum likelihood estimation (MLE) in latent data models from which a wide range of signal processing methods have been derived [8]. Furthermore, the potentiality of online and federated variants of the EM algorithm makes it a suitable choice for large-scale radio interferometric data. Notably, a computationally efficient EM algorithm had been proposed in the context of image restoration for sparse signal in the wavelet domain corrupted by a white Gaussian noise [9].

However, Radio-interferometric data cannot be appropriately modeled using white Gaussian noise. In fact, the new generation of large-scale radio interferometers, such as the SKA, will be able to detect weaker signals thanks to sensors being more and more sensitive, increasing the chances of having data corrupted by outliers. Namely, despite the efforts made to build such instruments in radio-quiet zones (in which radio transmissions are rare or restricted), man-made radio waves still significantly impact observation in radioastronomy [10]. Moreover, the generalization of large-scale telecommunication networks and the growing use of satellites makes RFI modeling and mitigation a significant challenge in array signal processing. More generally, the presence of outliers in the observed data can significantly affect the performances of imaging algorithms, leading to poor reconstructions. Likewise, image synthesis and calibration are intertwined, and calibration errors propagate and degrade the reconstructed images [11]. Theoretical and experimental analyses have been conducted to demonstrate the suitability of non-Gaussian heavy-tailed distribution to model the presence of outliers in the radioastronomy context [12], [13]. From this observation, multiple calibration algorithms that account for the possible presence of outliers have been derived. An iterative hard thresholding algorithm for wavelet coefficient estimation based on the Lorentzian cost

function to tackle the presence of outliers has been proposed in [14] but has not been applied to radio interferometric data.

In this paper, we propose an EM-based imaging algorithm that profits from the efficiency of the EM algorithm presented in [9] requiring $O(N \log N)$ per iteration, while considering a robust cost function based on the Student-t distribution. Numerical simulations display the impact of such modelization and the performances of the proposed algorithms.

## II. Radio interferometric data model

Radio interferometers measure the spatial coherence of the electric field for all the pairs of antennas that compose the sensor array. Such measurements are called *visibilities* and can be linked to the electromagnetic radiation of the observed celestial sources. The van Cittert-Zernike theorem connects the measured visibilities with the spatial Fourier transform of the emitted radiation across the celestial sphere. More precisely, it states that, considering unpolarized, non-coherent monochromatic sources in the far-field of the radio interferometer, the measured visibilities can be identified as samples of the 2D Fourier transform of the image [1],

$$\mathcal{V}(u,v,w) = \int\int \boldsymbol{\mathcal{I}}(l,m) e^{-j2\pi(ul+vm+wn)} \frac{dldm}{\sqrt{1-l^2-m^2}}. \quad (1)$$

The Fourier plane is refered to as the *uv plane* and its coordinates are expressed using the antennas coordinates and the observation wavelength. Given two antennas positionned at $\boldsymbol{r}_p$ and $\boldsymbol{r}_q$ and an observation wavelength, $\lambda$, the associated uv coordinate is written $(u,v,w) = \frac{\boldsymbol{r}_p - \boldsymbol{r}_q}{\lambda}$. Radio-interferometric data are samples of the 2D Fourier transform of the celestial image at continuous coordinates. To be able to leverage the power of the FFT, most radio interferometric imaging algorithms consider a gridding step to interpolate the measured visibilities in a uniform grid [2], [3], [5]. Subsequently, a discrete data model for gridded radio interferometric data can be expressed from the van Cittert-Zernike theorem,

$$\boldsymbol{V}_{pq} = \sum_{i,j} \boldsymbol{I}_{ij} e^{-j2\pi(u_p l_i + v_q l_j)} \quad (2)$$

## III. An EM Algorithm for radio interferometric imaging

The general inverse problem related to the imaging problem in radioastronomy can be written as a linear model by representing the image in lexicographic order [2], [4], [5],

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}, \\ \boldsymbol{H} &= \boldsymbol{S}\boldsymbol{F}, \end{aligned} \quad (3)$$

where $\boldsymbol{H} \in \mathbb{C}^{N\times M}$ represent the forward operator and $\boldsymbol{n} \sim \mathcal{CN}(0, \sigma^2\mathbf{I})$ is an additive noise. The visibility vector, $\boldsymbol{y} \in \mathbb{C}^{N\times 1}$, is expressed as a masked Fourier transform of the vectorized image of the sky in the field of view, $\boldsymbol{x} \in \mathbb{C}^{M\times 1}$. The forward operator, $\boldsymbol{H}$ is expressed as a product of a mask matrix, $\boldsymbol{S} \in \mathbb{C}^{N\times M}$ and a Fourier basis $\boldsymbol{F} \in \mathbb{C}^{M\times M}$. We choose to add a regularization of the form of an $L_1$ norm that promotes sparse solutions in the image domain.

$$\mathcal{R}(\boldsymbol{x}) = \alpha\|\boldsymbol{x}\|_1. \quad (4)$$

Although, it can be noticed that the transposition of the proposed algorithms to sparse priors in the wavelet domain is straightforward, leading to a range of imaging algorithms. Thus, the imaging problem consists at estimating the image, $\boldsymbol{x}$, under a Gaussian noise with a sparsity constraint. The associated maximum likelihood estimation problem can be solved using least square methods. Such methods might lack of computational efficiency for large scale data. We derive in this section an EM algorithm based on [9] to solve the associated regularized maximum likelihood estimation problem in $O(N\log(N))$ cost. To do so, we consider as complete data space $\boldsymbol{\xi} = (\boldsymbol{z}, \boldsymbol{y})$, where $\boldsymbol{z}$ is a noised complete Fourier transform of the image as presented in the following. We consider a decomposition of the global noise affecting the visibilities,

$$\boldsymbol{n} = \boldsymbol{S}\boldsymbol{e} + \boldsymbol{b}, \quad (5)$$

where $\boldsymbol{e}, \sim \mathcal{CN}(0, \sigma_e^2\mathbf{I})$ and $\boldsymbol{b}, \sim \mathcal{CN}(0, \boldsymbol{\Sigma}_b)$. Thus the observed data can be written as,

$$\boldsymbol{y} = \boldsymbol{S}\boldsymbol{z} + \boldsymbol{b}. \quad (6)$$

where,

$$\boldsymbol{z} = \boldsymbol{F}\boldsymbol{x} + \boldsymbol{e}. \quad (7)$$

The condition on the covariance matrices for the noise decomposition to be valid reads,

$$\boldsymbol{\Sigma}_b = \sigma^2\mathbf{I} - \sigma_e^2\boldsymbol{S}\boldsymbol{S}^T. \quad (8)$$

For $\boldsymbol{\Sigma}_b$ to be positive semidefinite, its eigenvalues must be greater or equal to zero, leading to the condition $\sigma_e^2 \leq \sigma^2/\lambda_1$ on $\sigma_e^2$ where $\lambda_1$ is the highest eigenvalues of $\boldsymbol{S}$. Since $\boldsymbol{S}$ is a mask matrix, its highest eigenvalue is $\lambda_1 = 1$. The condition on $\sigma_e^2$ thus reads $\sigma_e^2 \leq \sigma^2$.

The surrogate function associated to the expectation step of the EM algorithm is written as,

$$\begin{aligned} Q(\boldsymbol{x}|\boldsymbol{x}^{(m)}) &= \mathbb{E}_{\boldsymbol{z}|\boldsymbol{y},\boldsymbol{x}^{(m)})}\left[\log\ p(\boldsymbol{z},\boldsymbol{y};\boldsymbol{x})\right] \\ &\propto \|\hat{\boldsymbol{z}}^{(m)} - \boldsymbol{F}\boldsymbol{x}\|^2_{\boldsymbol{\Sigma}_e^{-1}}, \end{aligned} \quad (9)$$

where $\hat{\boldsymbol{z}}^{(m)} = \mathbb{E}_{\boldsymbol{z}|\boldsymbol{y},\boldsymbol{x}^{(m)})}[\boldsymbol{z}]$. Since both $\boldsymbol{z}$ and $\boldsymbol{y}$ are considerered complex multivariate gaussian, the conditional distribution of $\boldsymbol{z}$ given $\boldsymbol{y}$ is also a complex multivariate gaussian, leading to the following expression,

$$\hat{\boldsymbol{z}}^{(m)} = \boldsymbol{F}\boldsymbol{x}^{(m)} + \frac{\sigma_e^2}{\sigma^2}\boldsymbol{S}^T(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}^{(m)}). \quad (10)$$

The $M-step$ of the imaging *EM* algorithm reads,

$$\boldsymbol{x}^{(m+1)} = \underset{\boldsymbol{x}}{\text{argmax}}\ Q(\boldsymbol{x}|\boldsymbol{x}^{(m)}) + \mathcal{R}(\boldsymbol{x}), \quad (11)$$

leading to the following update,

$$\boldsymbol{x}^{(m+1)} = \mathcal{T}_\alpha\left(\boldsymbol{F}^H\boldsymbol{z}^{(m)}\right), \quad (12)$$

where $\mathcal{T}_\alpha$ is a soft-thresholding function that promotes sparsity in the solution based on iterative soft thresholding methods [15],

$$\mathcal{T}_\alpha(\boldsymbol{x}) = \text{sign}(\boldsymbol{x}) \times \max(0, |\boldsymbol{x} - \alpha|) \quad (13)$$

The computational load of the E-Step and M-step defined by equations (10) and (12) are dominated by the multiplication by a Fourier matrix, thus leading to efficient implementations of $O(N\log(N))$ cost using an FFT.

**Algorithm 1** EM Imager

**input:** $\boldsymbol{y}, \boldsymbol{S}, \boldsymbol{F}, \sigma_e^2, \sigma^2$
**output:** $\hat{\boldsymbol{x}}$
**initialize:** $\hat{\boldsymbol{x}} \leftarrow \boldsymbol{x}_0$

1: **while** stop criterion not reached **do**
2: E-Step : $\hat{\boldsymbol{z}}^{(m)}$ obtained from (10)
3: M-Step : $\hat{\boldsymbol{x}}^{(m+1)}$ obtained from (12)
4: **end while**

## IV. A Robust Imaging Algorithm

We present in this section an EM algorithm for imaging in the context of radio interferometric data affected by outliers. Robust noise distribution can efficiently model the presence of outliers. Amongst them, the Student-t is a distribution of choice, leading to robust estimates. We therefore consider a Student-$t$ distributed noise for the measured visibilities,

$$\boldsymbol{n}_n = \frac{1}{\sqrt{\tau}_n}\boldsymbol{t}_n, \quad \tau_n \sim \Gamma(\frac{\nu}{2},\frac{\nu}{2}), \quad \boldsymbol{t} \sim \mathcal{CN}(0,\sigma^2\mathbf{I}). \tag{14}$$

The number of degree of freedom (DoF) of the Student-$t$ distribution, $\nu$, is assumed known, and the parameters to estimate are $\boldsymbol{\theta} = (\boldsymbol{x}, \sigma^2)$. Notice however that data-driven methods to get an estimate the DoF of the Student-t can be implemented [16]. The log-likelihood of the model can then be expressed,

$$\begin{aligned}\mathcal{L}(\boldsymbol{y},\boldsymbol{\theta}) = C + \sum_{n=1}^{N}\frac{1}{2}\log\sigma^2 - \\ \frac{\nu+1}{2}\log\left(1+\frac{1}{\nu\sigma^2}|\boldsymbol{y}_n-(\boldsymbol{H}\boldsymbol{x})_n|^2\right).\end{aligned} \tag{15}$$

Direct maximization of the likelihood is not tractable and we propose to develop a SAGE algorithm to perform maximum likelihood estimation of the image and noise parameters. In practice, the SAGE consists in defining multiple spaces of parameters and their corresponding complete data spaces in which EM procedures are performed. Thereby, we define two sets of complete data from which we derive a SAGE procedure [17]. For each hidden data space we compute the expectation of the log-likelihood of the complete data given the observed incomplete data in the *E-step* as a surrogate function that is maximized in the *M-step*. Thus the set of complete data is defined so that the likelihood of its associated parameters leads to closed form maximizations. The first set of complete data is analogous to the one presented in section III and is used to solve the imaging problem,

$$\boldsymbol{\xi}^{[1]} = (\boldsymbol{y}, \boldsymbol{z}, \tau_1, \ldots, \tau_N). \tag{16}$$

Notice that the the global noise can be written $\boldsymbol{n} = \boldsymbol{\Omega}^{-1/2}\boldsymbol{t}$ with $\boldsymbol{\Omega} = \text{diag}(\tau_1,\ldots,\tau_N)$ and $\boldsymbol{n}|\boldsymbol{\Omega} \sim \mathcal{CN}(0,\sigma^2\mathbf{I})$ . Thus, for a given $\boldsymbol{\Omega} = \text{diag}(\tau_1,\ldots,\tau_N)$, the global noise can be decomposed as the sum of two Gaussian components,

$$\boldsymbol{n}_{|\boldsymbol{\Omega}} = \boldsymbol{S}\boldsymbol{e} + \boldsymbol{n}'. \tag{17}$$

The measured visibilities, given $\boldsymbol{\Omega}$, can thus be written,

$$\boldsymbol{y}_{|\boldsymbol{\Omega}} = \boldsymbol{S}\boldsymbol{z} + \boldsymbol{n}' \tag{18}$$

$$\boldsymbol{z} = \boldsymbol{F}\boldsymbol{x} + \boldsymbol{e}, \tag{19}$$

where $\boldsymbol{e} \sim \mathcal{CN}(0,\sigma_e^2\mathbf{I})$.

The associated EM algorithm is then expressed,

$$\boldsymbol{x}^{(m+1)} = \underset{\boldsymbol{x}}{\text{argmin}}\mathbb{E}_{\boldsymbol{\xi}^{[1]}|\boldsymbol{y},\boldsymbol{\theta}^{(m)}}\left[-\log p(\boldsymbol{\xi}^{[1]};\boldsymbol{x})\right]. \tag{20}$$

The surrogate function for the first set of complete data reads,

$$\begin{aligned}Q(\boldsymbol{\theta}_1|\boldsymbol{\theta}_1^{(m)}) &\propto \mathbb{E}_{\boldsymbol{\xi}^{[1]}|\boldsymbol{y},\boldsymbol{\theta}^{(m)}}\left[\log p(\boldsymbol{z}|(\tau_n)_{n\leq N};\boldsymbol{\theta})\right] \\ &\propto (\hat{\boldsymbol{z}}^{(m)}-\boldsymbol{F}\boldsymbol{x})^H(\hat{\boldsymbol{z}}^{(m)}-\boldsymbol{F}\boldsymbol{x})\end{aligned} \tag{21}$$

where $\hat{\boldsymbol{z}} = \mathbb{E}_{\boldsymbol{\xi}^{[1]}|\boldsymbol{y},\boldsymbol{\theta}^{(m)}}[\boldsymbol{z}]$ and is computed during the *E-Step* using the rules of conditional expectations, given that $p(\boldsymbol{z}|(\tau_n)_{n\leq N},\boldsymbol{y})$ is a multivariate complex Gaussian distribution. The latter reads,

$$\hat{\boldsymbol{z}} = \boldsymbol{F}\boldsymbol{x} + \frac{\sigma_e^2}{\sigma^2}\boldsymbol{S}^T\hat{\boldsymbol{\Omega}}^{(m)}(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}), \tag{22}$$

where $\hat{\boldsymbol{\Omega}}^{(m)} = \text{diag}(\hat{\tau}_1^{(m)},\ldots,\hat{\tau}_N^{(m)})$ and,

$$\hat{\tau}_n^{(m)} = \frac{\nu+1}{\nu+\sigma^{-2}|\boldsymbol{y}_n-(\boldsymbol{H}\boldsymbol{x}^{(m)})_n|^2}. \tag{23}$$

Considering a sparse regularization on the image, the $M-Step$ is similar to the one presented in section III, leading to the following update for the image,

$$\boldsymbol{x}^{(m+1)} = \mathcal{T}_\alpha\left(\boldsymbol{F}^H\hat{\boldsymbol{z}}\right). \tag{24}$$

We choose as the second set of complete data, used to derive an estimate of $\sigma^2$,

$$\boldsymbol{\xi}^{[2]} = (\boldsymbol{y}, \tau_1, \ldots, \tau_N). \tag{25}$$

The $E-Step$ of the associated EM procedure reads,

$$\begin{aligned}Q(\boldsymbol{\theta}_2|\boldsymbol{\theta}_2^{(m)}) &\propto \mathbb{E}_{\boldsymbol{\xi}^{[2]}|\boldsymbol{y},\boldsymbol{\theta}^{(m)}}\left[\log p(\boldsymbol{y}|(\tau_n)_{n\leq N};\boldsymbol{\theta})\right] \\ &\propto \frac{1}{\sigma^2}(\hat{\boldsymbol{y}}-\boldsymbol{H}\boldsymbol{x})^H\hat{\boldsymbol{\Omega}}^{(m)}(\hat{\boldsymbol{y}}-\boldsymbol{H}\boldsymbol{x}) \\ &\quad + \log(\sigma^2)\end{aligned} \tag{26}$$

where $\hat{\boldsymbol{\Omega}}^{(m)} = \text{diag}(\hat{\tau}_1^{(m)},\ldots,\hat{\tau}_N^{(m)})$ and,

$$\hat{\tau}_n^{(m)} = \frac{\nu+1}{\nu+\sigma^{-2^{(m)}}|\boldsymbol{y}_n-(\boldsymbol{H}\boldsymbol{x})_n|^2}. \tag{27}$$

The $M-Step$ leads to direct maximization,

$$\sigma^{2^{(m+1)}} = \frac{1}{N}\left((\hat{\boldsymbol{y}}-\boldsymbol{H}\boldsymbol{x})^H\hat{\boldsymbol{\Omega}}^{(m)}(\hat{\boldsymbol{y}}-\boldsymbol{H}\boldsymbol{x})\right). \tag{28}$$

Here again, the computational complexity of an iteration lie in a multiplication by a Fourier matrix and is implemented

in $O(N\log(N))$. In fact, the added normalisation step represented by the diagonal matrix, $\hat{\boldsymbol{\Omega}}$, can be expressed as a point-wise multiplication with a vector of $O(N)$ cost.

**Algorithm 2** Robust EM Imager

**input:** $\boldsymbol{y}, \boldsymbol{S}, \boldsymbol{F}, \sigma_e^2$
**output:** $\hat{\boldsymbol{x}}, \hat{\sigma}^2$
**initialize:** $\hat{\boldsymbol{x}} \leftarrow \boldsymbol{x}_0$, $\hat{\sigma}^2 \leftarrow \sigma_0^2$

**while** stop criterion not reached **do**

*Phase 1 – SAGE step 1 : Image estimation*

E-Step:
  $\hat{\boldsymbol{\tau}}_n^{(m)}$ obtained from (23) for $n \in [1, N]$
  $\hat{\boldsymbol{z}}^{(m)}$ obtained from (22)
M-Step:
  $\hat{\boldsymbol{x}}^{(m+1)}$ obtained from (24)

*Phase 2 – SAGE step 2 : noise variance estimation*

E-Step:
  $\hat{\boldsymbol{\tau}}_n^{(m)}$ obtained from (27) for $n \in [1, N]$
M-Step:
  $\hat{\sigma}^{2^{(m+1)}}$ obtained from (28)

**end while**

## V. Numerical results

In this section, we present several numerical results to illustrate the robustness of the proposed model and the performances of the implemented imaging algorithms. The test image used is *M31*, based on a $H2$ region in the M31 galaxy. We chose this image for its previous use in the study of imaging algorithm for radio interferometer [4], [18]. Likewise, its ability to efficiently model compact and extended structures makes of the M31 model image a good candidate to test the sparsity prior considered in the proposed imaging algorithms. The ground truth $256 \times 256$ discrete model of M31 is shown in figure 1.

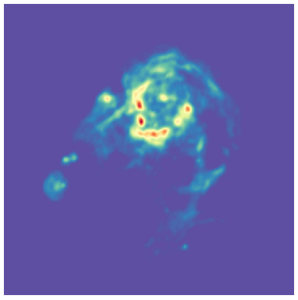

Fig. 1. M31 model image

The UV coverage is simulated using the rascil python library [19] and is plotted in figure 2. Subsequently, the uv plane of the LOWBD2-CORE configuration of the SKA is simulated and used to generate the visibilities from the ground truth image.

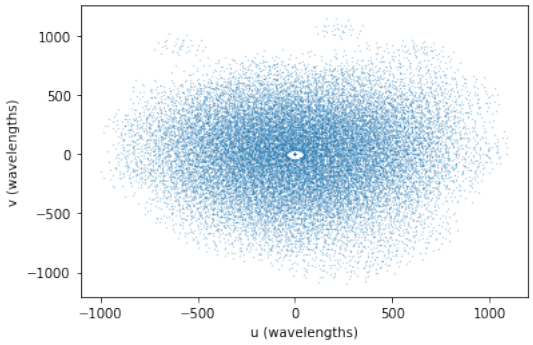


Fig. 2. UV coverage

We simulate an observation with Gaussian thermal noise added to the visibilities. To model the possible presence of outliers, $10\%$ of the visibilities are corrupted with an additional Gaussian noise. The power of the outliers are defined by the variance of the noise, $P_o = 10\log\frac{\sigma_o^2}{v_o}$. The performances of the algorithms are evaluated using the Normalized Mean Square Error between the estimated image and the ground truth, $\text{NMSE} = \frac{1}{100}\sum_{i=1}^{100}\frac{\|\hat{\boldsymbol{x}}_i - \boldsymbol{x}\|_2^2}{\|\boldsymbol{x}\|_2^2}$ for 100 monte carlo realizations. There has been sufficient work on the convergence properties of the EM algorithm, though it is known that an EM procedure increases the likelihood at each iteration converging to at least a local maxima [8], [20]. Yet, the EM algorithm is known for its slow convergence and dependance to initialization when the space of parameters is of high dimensions [21]. Thus we choose to initialize the proposed EM algorithms with the dirty image to avoid convergence to local maxima of the likelihood.

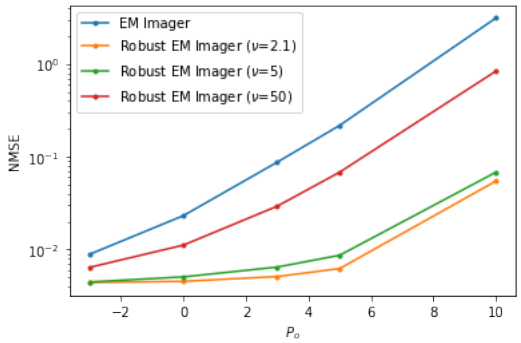


Fig. 3. Evolution of the NMSE vs the power of the outliers

Figure 3 displays the evolution of the NMSE with the power of the outliers for various values of the DoF of the Student-t distribution as well as for the EM imaging algorithm presented in section 1 that does not consider the presence of outliers in the modelization. A Student-t based imaging algorithm bring robustness to the estimation process, leading to less bias in the resulting image in the presence of outliers. The Student-t based imaging algorithm gets closer to a Gaussian-based imaging algorithm as the DoF. The choice of the Dof of the Student-t distribution is crucial and can be understood intuitively. In fact, as the DoF grows, it can be stated from the

central limit theorem that the Student-t distribution converges towards a Gaussian distribution, thus implicitely assuming few outliers in the observed data. Conversely low values for the DoF implicitely assume a strong presence of outliers in the data leading to more robuste estimate.

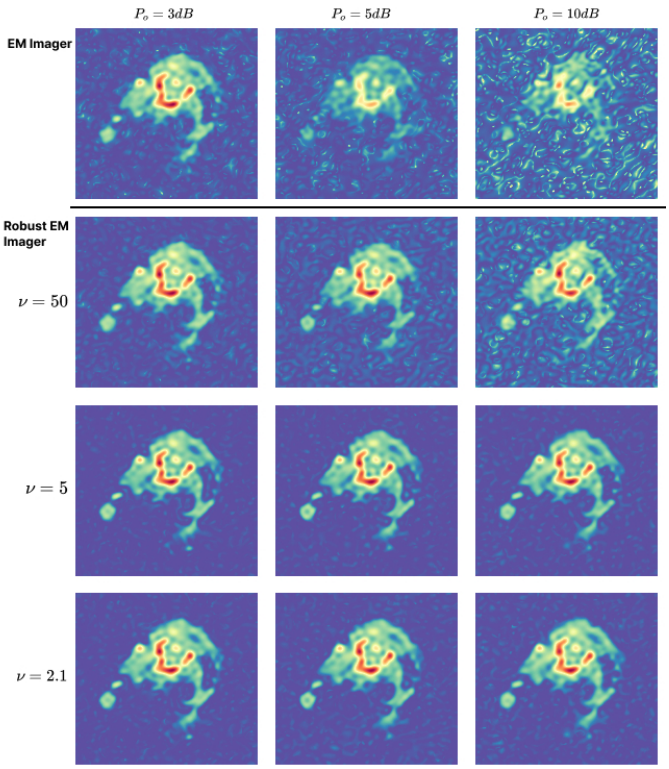


Fig. 4. Estimated images

## VI. Conclusion

In this paper, we present a new generic method for radio interferometric imaging using EM based algorithms. We propose a robust image synthesis method by modeling the additive noise with a Student-t distribution and designing an EM algorithm to perform maximum likelihood estimation. Numerical simulations illustrate the robustness of the proposed method to the presence of outliers in the observed data. The proposed methods lead to reasonable computational costs in the context of large scale radio interferometers and can be derived in an online fashion or in a federated learning framework by leveraging the versatility of EM based algorithms.

## References


[1] A. Thompson, J. Moran, and G. Swenson, *Interferometry and Synthesis in Radio Astronomy. Third edition*. Springer, 2017.

[2] Y. Wiaux, L. Jacques, G. Puy, A. M. Scaife, and P. Vandergheynst, "Compressed sensing imaging techniques for radio interferometry," *Monthly Notices of the Royal Astronomical Society*, vol. 395, no. 3, pp. 1733–1742, 2009.

[3] A. Offringa and O. Smirnov, "An optimized algorithm for multiscale wideband deconvolution of radio astronomical images," *Monthly Notices of the Royal Astronomical Society*, vol. 471, no. 1, pp. 301–316, 2017.

[4] 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.

[5] L. Pratley, J. D. McEwen, M. d'Avezac, R. E. Carrillo, A. Onose, and Y. Wiaux, "Robust sparse image reconstruction of radio interferometric observations with purify," *Monthly Notices of the Royal Astronomical Society*, vol. 473, no. 1, pp. 1038–1058, 2018.

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

[7] 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.

[8] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, no. 1, pp. 1–22, 1977.

[9] M. A. Figueiredo and R. D. Nowak, "An em algorithm for wavelet-based image restoration," *IEEE Transactions on Image Processing*, vol. 12, no. 8, pp. 906–916, 2003.

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

[11] S. J. Wijnholds and A.-J. van der Veen, "Fundamental imaging limits of radio telescope arrays," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 5, pp. 613–623, 2008.

[12] V. Ollier, M. N. El Korso, R. Boyer, P. Larzabal, and M. Pesavento, "Robust calibration of radio interferometers in non-gaussian environment," *IEEE Transactions on Signal Processing*, vol. 65, no. 21, pp. 5649–5660, 2017.

[13] S. Kazemi and S. Yatawatta, "Robust radio interferometric calibration using the t-distribution," *Monthly Notices of the Royal Astronomical Society*, vol. 435, no. 1, pp. 597–605, 2013.

[14] R. E. Carrillo and K. E. Barner, "Lorentzian iterative hard thresholding: Robust compressed sensing with prior information," *IEEE Transactions on Signal Processing*, vol. 61, no. 19, pp. 4822–4833, 2013.

[15] 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.

[16] F. Pascal, E. Ollila, and D. P. Palomar, "Improved estimation of the degree of freedom parameter of multivariate $t$-distribution," in *2021 29th European Signal Processing Conference (EUSIPCO)*. IEEE, 2021, pp. 860–864.

[17] J. A. Fessler and A. O. Hero, "Space-alternating generalized expectation-maximization algorithm," *IEEE Transactions on signal processing*, vol. 42, no. 10, pp. 2664–2677, 1994.

[18] 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.

[19] SKAO, "Radio astronomy simulation, calibration, and imaging library," https://gitlab.com/ska-telescope/external/rascil, May 2022.

[20] C. J. Wu, "On the convergence properties of the em algorithm," *The Annals of statistics*, pp. 95–103, 1983.

[21] Z. Wang, Q. Gu, Y. Ning, and H. Liu, "High dimensional expectation-maximization algorithm: Statistical optimization and asymptotic normality," *arXiv preprint arXiv:1412.8729*, 2014.