

A scale determination method for MSMFS CLEAN based on gradient descent optimizer Postprint

Authors: Xueying He, Lei Tan, Ying Mei, Hui Deng, Please provide the academic text you would like translated.

Date: 2025-07-25T09:22:19+00:00

Abstract

The performance of the deconvolution algorithm plays a crucial role in data processing of radio interferometers. The multi-scale multi-frequency synthesis (MSMFS) CLEAN is a widely used deconvolution algorithm for radio interferometric imaging, which combines the advantages of both wide-band synthesis imaging and multi-scale imaging and can substantially improve performance. However, how best to effectively determine the optimal scale is an important problem when implementing the MSMFS CLEAN algorithm. In this study, we proposed a Gaussian fitting method for multiple sources based on the gradient descent algorithm, with consideration of the influence of the point spread function (PSF). After fitting, we analyzed the fitting components using statistical analysis to derive reasonable scale information through the model parameters. A series of simulation validations demonstrated that the scales extracted by our proposed algorithm are accurate and reasonable. The proposed method can be applied to the deconvolution algorithm and provide modeling analysis for Gaussian sources, offering data support for source extraction algorithms.

Full Text

A Scale Determination Method for MSMFS CLEAN Based on Gradient Descent Optimizer

Xueying He¹², Lei Tan¹², Ying Mei^{12*}, Hui Deng^{12}

¹Center for Astrophysics, Guangzhou University, Guangzhou 510006, China

²Great Bay Center of National Astronomical Data Center, Guangzhou 510006, China

Correspondence: meiying@gzhu.edu.cn

Received: February 19, 2025; **Accepted:** May 9, 2025; **Published Online:** May 11, 2025

DOI: <https://doi.org/10.61977/ati2025021>

CSTR: <https://cstr.cn/32083.14.ati2025021>

© 2025 Editorial Office of Astronomical Techniques and Instruments, Yunnan Observatories, Chinese Academy of Sciences. This is an open access article under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>)

Citation: He, X. Y., Tan, L., Mei, Y., et al. 2025. A scale determination method for MSMFS CLEAN based on gradient descent optimizer. *Astronomical Techniques and Instruments*, 2(4): 219–225. <https://doi.org/10.61977/ati2025021>.

Abstract

The performance of the deconvolution algorithm plays a crucial role in data processing of radio interferometers. The multi-scale multi-frequency synthesis (MSMFS) CLEAN is a widely used deconvolution algorithm for radio interferometric imaging, which combines the advantages of both wide-band synthesis imaging and multi-scale imaging and can substantially improve performance. However, how best to effectively determine the optimal scale is an important problem when implementing the MSMFS CLEAN algorithm. In this study, we proposed a Gaussian fitting method for multiple sources based on the gradient descent algorithm, with consideration of the influence of the point spread function (PSF). After fitting, we analyzed the fitting components using statistical analysis to derive reasonable scale information through the model parameters. A series of simulation validations demonstrated that the scales extracted by our proposed algorithm are accurate and reasonable. The proposed method can be applied to the deconvolution algorithm and provide modeling analysis for Gaussian sources, offering data support for source extraction algorithms.

Keywords: Radio astronomy; Deconvolution; Synthesis imaging

1. Introduction

SKA1-Low, a low-frequency component of the Square Kilometre Array (SKA) project¹, is located at the Murchison Radio Astronomy Observatory (MRO) in Western Australia, operating over a frequency range of 50–350 MHz. SKA1-Low comprises 512 stations, each equipped with 256 dual-polarized antennas. It has 131,072 antennas in total, and the core area of SKA1-Low extends over approximately 1 km². It is designed to probe the early universe with unprecedented sensitivity and resolution. Its scientific objectives encompass a broad spectrum of astrophysical inquiries, including investigation of the Cosmic Dawn and the

Epoch of Reionization, during which the first stars and galaxies formed and ionized the neutral hydrogen in the interstellar medium. Additionally, it aims to explore galaxy evolution, cosmic magnetism, transient phenomena, pulsars, and perform tests of general relativity. All of those studies could potentially unveil unforeseen discoveries that will reshape our understanding of cosmology. The array design, featuring a dense core and three logarithmic spiral arms extending to a maximum baseline of 74 km, optimizes its ability to map large-scale structures and to detect faint signals in the distant universe².

Image processing for SKA1-Low is a key task in achieving the scientific goals of the project that focuses on reconstructing the sky brightness distribution from visibility data. The MSMFS CLEAN algorithm³ is one of the most widely used CLEAN algorithms and is the preferred algorithm for SKA1-Low imagery. MSMFS CLEAN integrates the advantages of both the multi-scale CLEAN⁴ and the multi-frequency CLEAN⁵ algorithms. In terms of the multi-scale element, MSMFS employs the same matched filtering technique as used in MS-CLEAN, whereby minor loops determine the position, amplitude, and scale of the main flux components in each iteration. The algorithm represents the image as a combination of truncated spherical wavelet bases that serve as scale bases. To optimize performance, these scale bases are typically preselected and carefully defined empirical values that meet the requirements of most signal sources.

Different scale settings in the MSMFS CLEAN algorithm have considerable impact on the final results of the deconvolution and operational efficiency. Determining the optimal scale present in the dirty image to be processed is an important task in application of the MSMFS CLEAN algorithm. The adaptive scale process (ASP) CLEAN algorithm⁶ introduces a more refined scale selection method, determining the optimal scale iteratively within minor loops. Although this approach enables more accurate fitting of certain sky components, it also incurs higher computational cost. In recent years, several improved algorithms based on the ASP algorithm have been developed⁷⁻⁹ to effectively enhance the efficiency of the ASP algorithm. Additionally, improvements have been made to the scale selection strategy of the ASP algorithm. For example, Zhang et al.⁹ proposed the Random Multi-scale Estimator algorithm, which employs a random perturbation mechanism to generate scale basis functions. This approach avoids scale uncertainty when using fixed scales, but does not fully resolve the issue of scale mismatch, and it still relies to some extent on the preset scale list.

For SKA1-Low, the determination of the MSMFS CLEAN scales is of great importance for the final imaging. Limited by the low observation frequency and short baseline length, SKA1-Low cannot distinguish the details of many extended sources, and a large number of extended sources might end up being observed as Gaussian sources in the final SKA1-Low imaging. In this study, we focused on improving the imaging efficiency and accuracy of MSMFS CLEAN by automatically determining the scale strategy from these Gaussian sources.

The remainder of this paper is organized as follows. We provide a detailed explanation of our scale selection strategy in Section 2. In Section 3, we de-

scribe performing simulation experiments, applying the proposed method to the MSMFS CLEAN deconvolution algorithm to validate its reliability. Finally, we summarize our work in Section 4.

2. Method

To systematically determine optimal scales for MSMFS CLEAN, we proposed a Gaussian fitting method based on gradient descent for multiple sources. This approach comprises the following three key components, and the workflow is illustrated in [Figure 1: see original paper].

1. **Source detection:** Identify potential Gaussian sources in the dirty image using local peak detection.
2. **Parameter optimization:** Fit detected sources with a multi-Gaussian model via gradient descent, incorporating the PSF to suppress sidelobe artifacts.
3. **Scale selection:** Derive a scale list from the fitted Gaussian parameters using statistical analysis of the full-width-half-maximum (FWHM) distribution.

2.1 Local Peak Detection

Before source detection, we normalize the dirty image using the Min-Max Normalization algorithm that is expressed as follows:

$$\tilde{I}_{\text{dirty}}(x, y) = \frac{I_{\text{dirty}}(x, y) - I_{\text{min}}}{I_{\text{max}} - I_{\text{min}}}$$

where \tilde{I}_{dirty} denotes the min-max normalized dirty image, I_{max} and I_{min} denote the global maximum and minimum pixel intensities in the dirty image, respectively.

We then use the local peak detection method of scikit-image¹⁰ to identify potential sources in the normalized dirty image \tilde{I}_{dirty} . This method controls the detection of reasonable peaks using two key parameters: the minimum distance between peaks (d_{min}) and the intensity threshold (T_{bright}). The minimum inter-peak distance is defined as:

$$d_{\text{min}} = \left\lceil \frac{B_{\text{maj}}}{2} \right\rceil$$

where B_{maj} denotes the synthesized Gaussian beam in pixels, $\lceil \cdot \rceil$ denotes the ceiling function, which rounds a number up to the nearest integer. This ensures that detected peaks are separated by at least half of the beam width. Additionally, T_{bright} is determined by the noise level in the dirty image. Generally,

T_{bright} could be set as 3–5 times the root mean square (RMS). Peaks below T_{bright} are discarded.

Eventually, we obtained all possible peak information, including the coordinates (x_i, y_i) of each detected peak. These coordinates represent the potential locations of the sources in the image. The list of detected peaks is denoted as $\{(x_i, y_i) | i = 1, \dots, N_g\}$, where N_g is the number of peaks.

2.2 Gradient Descent Based Multi-Gaussian Fitting

We employ gradient descent with the Adaptive Moment Estimation (Adam) optimizer to fit the multi-Gaussian model, which dynamically adjusts the Gaussian parameters to improve convergence and accuracy. Generally, a two-dimensional Gaussian source $G(x, y)$ can be defined as follows:

$$G(x, y) = I \exp \left\{ -\frac{[(x - \mu_x) \cos \theta + (y - \mu_y) \sin \theta]^2}{2\sigma_x^2} - \frac{[-(x - \mu_x) \sin \theta + (y - \mu_y) \cos \theta]^2}{2\sigma_y^2} \right\}$$

where (μ_x, μ_y) is the center position, σ_x and σ_y (expected values) are the standard deviations along the major and minor axes, $\theta \in [0, \pi]$ is the position angle (PA) relative to the x-axis, and I is the peak intensity.

For matching the symmetric scale kernel (tapered quadratic kernel) used in MSMFS CLEAN, we simplified the Gaussian model by setting both $\sigma_x = \sigma_y = \sigma$, and the PA as 0. Therefore, we modeled each Gaussian source $G_k(x, y)$ with these parameters:

The positions of the center $(\mu_{k;x}, \mu_{k;y})$ obtained from the peak position in Section 2.1 are defined as follows:

$$(\mu_{\text{init},k;x}, \mu_{\text{init},k;y}) = (x_k, y_k), \quad k = 1, \dots, N$$

The standard deviation $\sigma_{\text{init},k}$ was set uniformly with the angular resolution of the instrument in pixels.

The intensity $I_{\text{init},k}$ was uniformly assigned to all components, using the median brightness of the detected peaks to avoid bias toward bright sources during the first iterations.

We generate a model image I_{model} using these Gaussian sources to reconstruct the sky model. To optimize computational performance, we restricted the range of each Gaussian source to a certain region of sigma as follows:

$$I_{\text{model}}(x, y) = \sum_{k=1}^N G_k(x', y'), \quad (x', y') \in \Omega_k$$

where Ω_k represents the effective region for the k -th Gaussian source. Ω_k is defined as a square centered at $(\mu_{k;x}, \mu_{k;y})$, which is determined by the standard deviation of the Gaussian model. The side length W of the Gaussian model is defined as 6 times the standard deviation:

$$W = 6\sigma_k$$

Considering the sidelobe effect, we convolve the image of the sky model with the PSF to effectively suppress the sidelobe effect caused by the instrument response. The mean squared error (MSE) loss function is defined as follows:

$$L = \| [I_{\text{model}}(x, y) \otimes P(x, y)] - \tilde{I}_{\text{dirty}}(x, y) \|^2 = \| \mathcal{F}^{-1}[\mathcal{F}(I_{\text{model}}) \cdot \mathcal{F}(P)] - \tilde{I}_{\text{dirty}} \|^2$$

where $P(x, y)$ represents the PSF and \otimes denotes the convolution operation. We move the convolution operation to the spatial frequency domain for computation and then transform it back to the spatial domain through the inverse Fourier transform \mathcal{F}^{-1} . Then, we calculate the MSE.

During the fitting, the intensity I_k is constrained to be non-negative by applying the ReLU activation function, which is defined as follows:

$$I_k = \text{ReLU}(I_k) = \begin{cases} I_k & \text{if } I_k > 0 \\ 0 & \text{otherwise} \end{cases}$$

which enforces physical plausibility by suppressing non-physical negative values.

2.3 Scale Determination Strategy

During the calculation process of the MSMFS CLEAN, the scale (in pixels) can be derived from the FWHM of the Gaussian source. The relationship between the FWHM and the standard deviation σ is expressed as follows:

$$\text{FWHM} = 2\sqrt{2 \ln 2} \sigma$$

After the model converges, we calculate the FWHM based on the Gaussian parameter σ and then determine the list of scale sizes. The specific steps are as follows:

1. Produce a histogram of the FWHM distribution with a bin width of one pixel.
2. Extract quantile-based scales such as the minimum, the first quartile (Q1), the median, the third quartile (Q3), and the maximum from the scale distribution. If adjacent quantiles are close in scale, they are merged to avoid redundancy.

3. Construct scales of geometric progression. Generally, such a list of scale sizes can be 0, 2, 4, 8, 16 or 0, 3, 10, 30, or 0, 1.5, 3, 6, 12 pixels⁴. The final entry in the list should reach the expected maximum scale.

The final scale is determined using the quantile-based scales (Step 2) or the geometric progression scales (Step 3). To avoid occupying an excessive resource space and to control the efficiency of the algorithm to within a reasonable range, we restrict the list size to eight. This approach balances the computational load and the flexibility of the model.

3. Simulation and Validation

3.1 Simulated Data for Validation

To validate our method, we used the Radio Astronomy Simulation, Calibration, and Imaging Library 2 (RASCIL2)¹ to generate simulated observation data based on snapshot observations using the configuration given in the SKA1-LOW settings for LOWR3. The observation frequency is 150 MHz. In the simulation, we set the phase center as (RA, Dec) = (+50°, -40°), and set the field of view (FoV) as 2.75°, imaging size 4096 × 4096 pixels, and the cell size was 2.42"/pixel. For simplicity, primary beam effects were ignored in this simulation. [Figure 2: see original paper] illustrates the UV-plane sampling distribution.

The simulation data were taken from the GaLactic and Extragalactic All-sky Murchison Widefield Array catalog^{11,12}. Referring to the ring distribution of sources mentioned in Zhang et al.¹³ and the extended source parameters suggested by the tiered radio extragalactic continuum simulation method^{14,15}, we injected eight Gaussian sources into the UV data. The parameters of these eight extended sources, including the peak intensity, flux, FWHM of the major and minor axes, and the PA, are listed in . [Figure 3: see original paper] shows a diagram of the sky image and the dirty image. It should be noted that we cannot use the RMS of the dirty image (3.6×10^{-2} Jy/beam) directly based on the full dirty image. In practice, we need to select a small region of the image without sources and then calculate the RMS, which is 3.25×10^{-3} Jy/beam.

3.2 Experimental Parameters and Results

As described in Section 2.1, the local peak detection algorithm is used to detect Gaussian sources. A total of 523 potential sources were detected in the dirty image. Among these detected sources, in addition to the correct source positions, a large number of noise sources were falsely detected, as illustrated in [Figure 4: see original paper].

With a learning rate of 0.01 and the Adam optimizer running for 4,000 iterations, the generated loss curve is shown in [Figure 5: see original paper].

After the fitting process, we finally obtained 74 sources with their Gaussian

parameters. It is evident that the false detections caused by sidelobes are effectively reduced by the ReLU activation function in intensity parameters. The final fitted model image is shown in [Figure 6: see original paper].

The FWHM histogram for the remaining sources is shown in [Figure 7: see original paper]. The minimum, first quartile (Q1), median, third quartile (Q3), and maximum of the FWHM distribution are 16, 19, 22, 25, and 41, respectively. According to the method described in Section 2.3, we choose 0, 6, 12, 22, and 40, approximately representing a geometric progression.

3.3 Scale Validations

We further applied the “tclean” task in CASA 6.6.3¹⁶ to deconvolute the dirty image to validate the availability of the scales determined in this study. During deconvolution, we set ‘niter’ as 10,000, ‘gain’ as 0.6, ‘threshold’ as 0.025 Jy, and ‘nsigma’ as 1.5, without applying a mask. We set ‘deconvolver’ as “asp” and “multiscale” to invoke the ASP algorithm and the MSMFS CLEAN algorithm, respectively. For MSMFS CLEAN, we set the parameter ‘scales’ as [0, 6, 12, 22, 40]. Meanwhile, some geometric progression scales were used to compare the deconvolution performance. For ASP CLEAN, we chose the largest scale from the scales used in MSMFS CLEAN and assigned its value to the parameter ‘largestscale’.

To evaluate image deconvolution quality, we used several numerical metrics, including the RMS of the residual image, the dynamic range of the restored image, the structural similarity index (SSIM) evaluated against the true smooth image, and the total run time. Among these, the SSIM measures similarity by considering three aspects of the images: luminance, contrast, and structure, where a value close to 1 indicates a high degree of similarity. The results are listed in .

From the experimental results, it is apparent that the scale [0, 6, 12, 22, 40], adopted in this study, is reasonable. From both the evaluation metrics and the final imaging results, it can be seen that the MSMFS results based on this scale of deconvolution have a lower RMS residual, a higher dynamic range, and reliable image fidelity. At the same time, the deconvolution computation is less time-consuming. This suggests that the maximum scale calculated in this study is optimal.

4. Discussions and Conclusions

In this study, we applied a Gaussian fitting method based on gradient descent to analyze dirty imagery in radio interferometry. We used CUDA parallel computing within the PyTorch framework for parameter fitting. This approach leverages the parallel processing capabilities of GPUs, substantially enhancing

computational efficiency, with full account of the impact of the point spread function.

This analytical strategy was shown to successfully generate a scale distribution map that clearly shows the scale distribution characteristics. By examining this distribution, we can gain insight into the Gaussian source scales and subsequently select a scale suitable for the MSMFS CLEAN algorithm. When applied to the imaging deconvolution process, this approach effectively reduces the reliance on the manually defined scale input list. This algorithm is not only applicable to image deconvolution, but also provides a reference for the precise extraction of radio sources.

Generally, the scale detection algorithm proposed in this study provides reasonable scale parameters that can improve the quality of an image while reducing the time required to perform the deconvolution. This represents an important reference for future SKA1-Low imaging data processing.

Despite the results achieved in this study, certain limitations remain. The study was based on a simulation approach, and thus all the experiments were based on simulation observations. This clearly has an impact on the validation of the overall algorithm, including how to inject more rationality and realism with the simulation extended sources, which is the focus of our future work.

Another limitation is that the current algorithm has room for further improvement. In the local peak detection algorithm, the design of the threshold parameter can weaken the detection effect on weak sources, especially when they are affected by higher margins or lower signal-to-noise ratios. Moreover, setting the distance parameter might result in more closely spaced sources being masked by larger distributions of sidelobe structures and thus remaining undetected.

Acknowledgements

This work is supported by the China National SKA Program (2020SKA0110300), the Natural Science Foundation of China (12433012, 12373097), and the Guangdong Province Basic and Applied Basic Research Foundation Project of Guangdong Province (2024A1515011503).

AI Disclosure Statement

Kimi was employed for language and grammar checks within the article. The authors carefully reviewed, edited, and revised the Kimi-generated texts to their own preferences, assuming ultimate responsibility for the content of the publication.

Author Contributions

Xueying He and Ying Mei conceived the ideas and designed the study. Lei Tan designed and applied the local peak detection method for candidate source identification. Xueying He and Ying Mei conducted the Gradient Descent Based Multi-Gaussian Fitting, scale determination experiments, and method validation (simulation and scale analysis). Hui Deng supervised the project, validated the results, and provided critical revisions. All authors read and approved the final manuscript.

Declaration of Interests

The authors declare no competing interests.

References

1. Pupillo, Dewdney, P., Hall, P., Schilizzi, R., et al. 2009. The Square Kilometre Array. *Proceedings of the IEEE*, 97(8): 1482–1496.
2. Macario, et al. 2021. Characterization of the SKA1-low prototype station aperture array verification System 2. *Journal of Astronomical Telescopes, Instruments, and Systems*, 8(1): 11–14.
3. Rau, U., Cornwell, T. J. 2011. A multi-scale multi-frequency deconvolution algorithm for synthesis imaging in radio interferometry. *Astronomy & Astrophysics*, 532: A71.
4. Bernardi, Wieringa, Cornwell, T. J. 2008. Multiscale CLEAN deconvolution of radio synthesis images. *IEEE Journal of Selected Topics in Signal Processing*, 2(5): 793–801.
5. Sault, H. 1994. Multi-frequency synthesis techniques in radio interferometric imaging. *Astronomy & Astrophysics Supplement*, 108: 585–594.
6. Bhatnagar, J. 2004. Scale sensitive deconvolution of interferometric images-I. Adaptive Scale Pixel (Asp) decomposition. *Astronomy & Astrophysics*, 426(2): 747–754.
7. Zhang, L., Bhatnagar, S., Rau, U., et al. 2016. Efficient implementation of the adaptive scale pixel decomposition algorithm. *Astronomy & Astrophysics*, 592: A128.
8. Zhang, L. 2018. Fused CLEAN deconvolution for compact and diffuse emission. *Astronomy & Astrophysics*, 618: A117.
9. Zhang, L., Mi, L. G., Zhang, M., et al. 2021. Parameterized reconstruction with random scales for radio synthesis imaging. *Astronomy & Astrophysics*, 646: A44.

10. Van Der Walt, S., Schönberger, J., Nunez-Iglesias, J., et al. 2014. Scikit-image: image processing in Python. *PeerJ*, 2: e453.
11. Wayth, R. B., Lenc, E., Bell, M. E., et al. 2015. GLEAM: The galactic and extragalactic all-sky MWA survey. *Proceedings of the Astronomical Society of Australia*, 32: e025.
12. Hurley-Walker, N., Callingham, J. R., Hancock, P. J., et al. 2017. Galactic and Extragalactic All-sky Murchison Widefield Array (GLEAM) survey-I. A low-frequency extragalactic catalogue. *Monthly Notices of the Royal Astronomical Society*, 464(1): 1146–1167.
13. Zhang, M., Jackson, N., Porcas, R. W., et al. 2007. A search for the third lensed image in JVAS B1030+074. *Monthly Notices of the Royal Astronomical Society*, 377(4): 1623–1634.
14. Bonaldi, A., Bonato, M., Galluzzi, V., et al. 2019. T-RECS: Tiered Radio Extragalactic Continuum Simulation, Astrophysics Source Code Library, <https://ascl.net/1906>.
15. Bonaldi, A., Bonato, M., Galluzzi, V., et al. 2019. The Tiered Radio Extragalactic Continuum Simulation (T-RECS). *Monthly Notices of the Royal Astronomical Society*, 482(1): 2–19.
16. CASA Team, Bean, B., Bhatnagar, S., et al. 2022. CASA, the common astronomy software applications for radio astronomy. *Proceedings of the Astronomical Society of Australia*, 134(1041): 114501.

Note: The figure and table markers referenced in the text ([Figure 1: see original paper], [Figure 2: see original paper], etc., ,) correspond to the original publication’s figures and tables, which are not reproduced here.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv — Machine translation. Verify with original.