RAA 2021 Vol. **21** No. **3**, 63(11pp) doi: 10.1088/1674-4527/21/3/63

http://www.raa-journal.org *http://iopscience.iop.org/raa*

Research in Astronomy and Astrophysics

Adaptive scale model reconstruction for radio synthesis imaging

Li Zhang (张利)[1], Li-Gong Mi[1], Long Xu[2], Ming Zhang[3,4], Dan-Yang Li[1], Xiang Liu[3,4], Feng Wang[5], Yi-Fan Xiao[1] and Zhong-Zu Wu[6]

[1] College of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China *lizhang.science@gmail.com*
[2] Key Laboratory of Solar Activity, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China
[3] Xinjiang Astronomical Observatory, Chinese Academy of Sciences, Urumqi 830011, China
[4] Key Laboratory of Radio Astronomy, Chinese Academy of Sciences, Urumqi 830011, China
[5] Center for Astrophysics, Guangzhou University, Guangzhou 510006, China
[6] College of Physics, Guizhou University, Guiyang 550025, China

Received 2020 June 17; accepted 2020 September 1

Abstract A sky model from CLEAN deconvolution is a particularly effective high dynamic range reconstruction in radio astronomy, which can effectively model the sky and remove the sidelobes of the point spread function (PSF) caused by incomplete sampling in the spatial frequency domain. Compared to scale-free and multi-scale sky models, adaptive-scale sky modeling, which can model both compact and diffuse features, has been proven to have better sky modeling capabilities in narrowband simulated data, especially for large-scale features in high-sensitivity observations which are exactly one of the challenges of data processing for the Square Kilometre Array (SKA). However, adaptive scale CLEAN algorithms have not been verified by real observation data and allow negative components in the model. In this paper, we propose an adaptive scale model algorithm with non-negative constraint and wideband imaging capacities, and it is applied to simulated SKA data and real observation data from the Karl G. Jansky Very Large Array (JVLA), an SKA precursor. Experiments show that the new algorithm can reconstruct more physical models with rich details. This work is a step forward for future SKA image reconstruction and developing SKA imaging pipelines.

Key words: methods: data analysis — techniques: image processing — techniques: interferometric

1 INTRODUCTION

The Square Kilometre Array (SKA) (Wu 2019), a large international scientific project in which China participates, is the world's largest synthesis aperture radio telescope array jointly constructed by more than 10 countries around the world, with a frequency coverage of 50 MHz to 20 GHz and an equivalent receiving area of up to a square kilometer. The SKA will have an extremely high sensitivity, a wide field of view, ultra-high spatial, frequency and time resolutions, and ultra-fast survey speed (Wu 2019; An 2019), which will provide an unprecedented powerful performance and is expected to answer fundamental questions such as the origin of the universe, the nature of gravity, magnetic field of the universe and life, which will bring revolutionary changes in many major fields of natural science and provide humans with a great opportunity to explore and understand the universe. However, it also brings huge technical challenges.

The SKA will increase the sensitivity of the telescope by about 50 times compared to the largest existing radio telescope array, Karl G. Jansky Very Large Array (JVLA) (Wu 2019). The extremely high sensitivity can image a large number of faint compact emission and faint structures of diffuse emission, so that signal features within the imaging region become very complicated. This provides a wealth of materials for exploring the universe, but it also brings a big challenge to sky modeling – effective sky modeling is needed to reconstruct such images.

CLEAN deconvolution is a very widely used and particularly effective image reconstruction method in radio astronomy (Bhatnagar & Cornwell 2004; Cornwell 2008;

Zhang 2018; Zhang et al. 2020), which employs iterative methods to continuously accumulate model components to achieve a model for the sky. CLEAN deconvolution first finds a credible component in the residual images (the dirty image for the first component), and then subtracts the dirty-beam effects of this component, and iterates in this way continuously until there is no obvious signal or close to noise in the residuals, and accumulate all components form the sky model. According to the parameterization of sky modeling, CLEAN deconvolution can be divided into three categories (Zhang et al. 2020): scale-free CLEAN, which constructs a scale-free sky model, multi-scale CLEAN, which constructs a multi-scale sky model, and adaptive-scale CLEAN, which constructs an adaptive-scale sky model.

Adaptive-scale sky models have been proven to have better sky modeling capabilities (Bhatnagar & Cornwell 2004; Zhang et al. 2020). However, there are three problems with the current adaptive-scale sky models:

1. Negative components are allowed in these models.
2. They have not been verified in wideband/multi-frequency cases commonly used in modern telescopes such as the SKA.
3. They have not been tested by real observation data.

To apply adaptive-scale sky models to future SKA imaging, these three problems need to be addressed. In this paper, we propose a new adaptive-sale sky model algorithm, which is applied to simulated SKA data and real observation data of the JVLA.

To better understand the whole problem and our algorithm, the theories of radio synthesis imaging, multi-frequency synthesis and parameterized sky models are introduced in Section 2. The algorithm proposed in this paper is discussed in detail in Section 3. The results of the proposed algorithm applied to simulated and real wideband observations are discussed in Section 4. The final summary is provided in Section 5.

2 THEORY

2.1 Radio Synthesis Imaging

Radio interferometry (Pawsey et al. 1946; Ryle & Vonberg 1948; Thompson et al. 2017) employs multiple telescopes to simultaneously obtain samples in the spatial frequency domain, and uses Earth-rotation synthesis to increase samples to measure more spatial frequencies. For narrowband observations, this measurement can be expressed as follows:

$$V^{\mathrm{obs}} = SV^{\mathrm{sky}} + SN, \qquad (1)$$

where V^{obs} is the observed data, which is called visibilities in radio astronomy, S is a sampling matrix determined by the UV coverage of the observation, V^{sky} is the visibility function which is the Fourier transform of the sky brightness distribution I^{sky} (please refer to the van Cittert-Zernike theorem in Thompson et al. (2017) and Taylor et al. (1999)), and N is the noise in the spatial frequency domain. The dirty image I^{obs} is computed by the inverse Fourier transform,

$$I^{\mathrm{obs}} = F^{-1}V^{\mathrm{obs}} = I^{\mathrm{dbeam}}(I^{\mathrm{sky}} + I^{\mathrm{noise}}), \qquad (2)$$

where F^{-1} is the inverse Fourier transform, $I^{\mathrm{noise}} = F^{-1}N$ is the noise in the image domain, I^{dbeam} is a Toeplitz matrix where each row is a shifted version of the dirty beam which is the inverse Fourier transform of the sampling matrix S. The dirty beam often has different levels of sidelobes due to factors such as the limited number of telescopes and observation times. It, together with the noise term, contaminates astrophysics objects, which often seriously affects their astrophysics analysis. Therefore, an effective sky model is needed to represent the astrophysics objects that are submerged in the dirty beam and noise.

2.2 Multi-frequency Synthesis

Just like other modern telescope arrays, the SKA will have a wideband imaging capability, which improves the sensitivity of the instruments to measure the spectral structure of astronomical sources in detail, and provides high-dynamic-range imaging performance that is superior to narrowband observations. In wideband measurements, different frequency channels measure different spatial frequency ranges, which increases UV coverage and improves the imaging performance of an interferometric array on the sky brightness distribution. Multi-frequency synthesis (MFS) (Conway et al. 1990) is a technique combining multiple discrete frequency observation, which is applied to multi-frequency image reconstruction. Standard multi-frequency synthesis imaging assumes that the sky brightness distribution does not vary across the total measured frequency bandwidth and grids all observed visibilities from different frequency channels onto the same spatial-frequency grid.

$$V^{\mathrm{obs}} = \sum_{\mathrm{v}} V^{\mathrm{obs}}_{\mathrm{v}}, \qquad (3)$$

where $V^{\mathrm{obs}}_{\mathrm{v}}$ is the visibilities measured at frequency v. MFS can be used to combine multiple narrowband observations to increase UV coverage by switching frequencies during observations, or to eliminate the bandwidth-smearing of wideband observations by splitting the broadband into multiple narrowband frequencies and mapping them to corresponding spatial frequencies. The dirty/observed image is calculated from wideband or multi-frequency visibilities using MFS, and then find the optimal adaptive-scale sky model to represent a sky brightness distribution. Different sky models have significant differences in the ability to represent a sky brightness distribution, which is discussed below.

2.3 Parameterized Sky Models

A scale-free sky model (Högbom 1974; Clark 1980; Schwab 1984; Mei et al. 2018) parameterizes the sky into a series of delta functions,

$$I^{\rm model} = \sum_{i=1}^{N} I_i^{\rm peak} \delta_i (x - x_i, y - y_i), \tag{4}$$

where $I^{\rm model}$ is a scale-free sky model consisting of N components, $\sum$ is a summing operation, $I_i^{\rm peak}$ is the peak amplitude of the ith component and δ_i is a delta function located at $(x - x_i, y - y_i)$. It is very effective for the representation of compact emission, especially the case which there are only a few well-separated point sources in the imaging sky region. However, this model cannot represent the correlation between adjacent pixels. A scale-free sky model represents diffuse emission as a series of isolated delta functions, which is often a physically inaccurate expression. A well-known example is the stripes from the reconstruction of diffuse emission (Clark 1982) and many improvements have been developed in the early days to eliminate the stripes (Palagi 1982; Cornwell 1983; Steer et al. 1984). For example, Cornwell (1983) suppressed stripes by adding a regularization term to the objective function to introduce a smooth prior into the model. However, these did not fundamentally solve the problem. Reconstruction based on a scale-free sky model may not be distinguishable from the real sky due to the unmeasured spatial frequencies that are not limited by this model.

To represent the complex features of astrophysics targets at different scales, a straightforward idea is to introduce some a priori constraints on the scale of sky emission. A multi-scale sky model (Cornwell 2008; Rau & Cornwell 2011) parameterizes the sky in a multi-scale basis, which can express the correlation between pixels and provides a strong constraint to the unmeasured spatial frequencies.

$$I^{\rm model} = \sum_{i=1}^{N} I_i^{\rm comp} (amp, loc, scale), \tag{5}$$

where $I_i^{\rm comp}(amp, loc, scale)$ is the ith component at loc with the amplitude amp and the enumerated scale $scale$. It significantly improves the ability to express diffuse emission and fundamentally eliminates the stripes caused by a scale-free sky model. However, a multi-scale sky model is built on user-specified scales, and the length of the scale list is generally a few due to the limitations of the computer's memory and computational load. This causes sky emission that is not in the specified scale list to be broken into the specified scales, and the sky is forced to be represented as a set of specified scales, which results in an inaccurate representation.

Sky emission obtained by modern ultra-high sensitivity telescopes such as the SKA often has a large number of complex features with uncertain scales, which essentially requires an adaptive-scale sky model. Such cases cannot be well modeled by scale-free and multi-scale sky models. An adaptive-scale sky model (Bhatnagar & Cornwell 2004; Zhang et al. 2016; Zhang 2018; Zhang et al. 2019) uses an adaptive scale basis to parameterize the sky and its scales are changed adaptively with the inherent scales of emission.

$$I^{\rm model} = \sum_{i=1}^{N} I_i^{\rm comp} (p_i), \tag{6}$$

where $I_i^{\rm comp}(p_i)$ is the ith component with an adaptive-scale parameter p_i. The inaccurate representation due to the specified scales in a multi-scale sky model is theoretically eliminated. As mentioned earlier, there are still some problems with adaptive scale models, which will be addressed as follows.

3 THE ALGORITHM OF ADAPTIVE-SCALE SKY MODEL WITH MFS

In this section, we introduce this new algorithm that combines MFS to construct an adaptive scale sky model from wideband or multi-frequency observations. In this adaptive-scale sky model, emission is physically restricted to positive, so we also call it positive-fusedClean or p-fusedClean.

Our algorithm uses the standard reconstruction framework (Venkata 2010; Zhang et al. 2020) to construct the optimal adaptive-scale sky model which is consistent with the measurements and predicts the unmeasured spatial frequencies well by the following iterative method.

1. Compute dirty/residual images;
2. Search model components;
3. Update the model and residuals;
4. Predict the model onto the observed data points when the model accumulates to a certain degree;
5. Compute the visibility residuals;
6. Repeat the above process until the optimal adaptive-scale sky model was found.

This process is performed alternately between the image domain and the visibility domain. Model component search is performed in the image domain and errors such as from gridding are corrected in the visibility domain.

The adaptive-scale sky model needs to be estimated from the dirty image determined by the following formula,

$$I^{\mathrm{obs}} = F^{-1} \sum_{0}^{N-1} S_{\mathrm{v}} V^{\mathrm{sky}}, \tag{7}$$

where S_{v} is a sampling matrix that changes with frequencies. All the measured S_{v}s after multi-frequency synthesis together determine the entire sampling matrix. In the subsequent correction process, the residual image $V^{\mathrm{vis-resi}}$ needs to be computed, which is estimated from the residual visibilities V^{residual},

$$V^{\mathrm{vis-resi}} = F^{-1} \sum_{0}^{N-1} S_{\mathrm{v}} V^{\mathrm{residual}}. \tag{8}$$

The spatial frequencies increased by multi-frequency synthesis will significantly reduce the sidelobes of the dirty beam, thereby increasing the sensitivity of the measurements. Therefore, there is less overlap between adjacent sources due to the sidelobes of the dirty beam, and the quality of the dirty image is better and easier to reconstruct the original appearance of the sky.

During the model search phase, the adaptive-scale gaussian functions and delta functions are used to model the radio sky. The delta functions are used to construct compact emission, which are effective approximations to the zero-scale gaussian functions. Most of the emission is parameterized by the adaptive-scale gaussian functions, which are obtained by a fast explicit fit. An initial guess of fitting parameters by a matched filtering is used to find these adaptive-scale gaussian components. Each gaussian component is estimated from the current residual image by minimizing the following objective function,

$$\chi^2 = \frac{1}{2}\left(I_{i-1}^{\mathrm{residual}} - I_i^{\mathrm{g-comp}}(p_i)\right)^{\mathrm{T}}\left(I_{i-1}^{\mathrm{residual}} - I_i^{\mathrm{g-comp}}(p_i)\right), \tag{9}$$

where T is the transpose operation, $I_i^{\mathrm{g-comp}}(p_i) = I_i^{\mathrm{gauss}}(a_{\mathrm{ig}}, l_{\mathrm{ig}}, m_{\mathrm{ig}}, \omega_{\mathrm{ig}})$ is a gaussian function with amplitude a_{ig}, position $(l_{\mathrm{ig}}, m_{\mathrm{ig}})$ and scale ω_{ig}, which is estimated by, for example, the Levenberg-Marquardt optimizer (Marquardt 1963) from the current residual image $I_{i-1}^{\mathrm{residual}}$. The optimal component is obtained by continuously updating the direction of the negative gradients,

$$\frac{\partial \chi^2}{\partial p_i} = -\left[I_{i-1}^{\mathrm{residual}}\right]^{\mathrm{T}} \frac{\partial I_i^{\mathrm{g-comp}}}{\partial p_i}, \tag{10}$$

where T is a transpose operation. This optimizer is able to find that the largest scale size in the current residual image, which is consistent with the nature of the image, and the scale size of the component is completely determined by the nature of the image itself. The parameters of a model component can be computed from the optimal gaussian I_i^{gauss} by the following method,

$$l_i = l_{\mathrm{ig}}, \tag{11}$$

$$m_i = m_{\mathrm{ig}}, \tag{12}$$

$$\omega_{\mathrm{ig}}^2 = \omega_i^2 + \omega_{\mathrm{b}}^2, \tag{13}$$

$$a_{\mathrm{ig}}\omega_{\mathrm{ig}}^2 = 2\pi a_{\mathrm{b}} a_i \omega_{\mathrm{b}}^2 \omega_i^2, \tag{14}$$

where l_i and m_i are the location of the ith model component respectively, which are constant during fitting, and ω_i and a_i are the width and amplitude of this model component, respectively. In this method, a gaussian beam with amplitude a_{b} and width ω_{b} needs to be estimated from the dirty beam. The major effect from the main lobe of the dirty beam can be analytically eliminated. Here, the errors caused by the approximation of the dirty beam will be corrected in the next step. The method of analytically eliminating the effects of the main lobe of the dirty beam can significantly improve the computing efficiency by removing a lot of time-consuming convolution calculations in the fitting process.

When compact emission is detected, delta functions are used to represent the components $I_i^{\mathrm{comp}} = I_i^{\mathrm{peak}}(l, m)$. The peak search method is triggered to find compact components when small scale sizes appear frequently in recent model components or the initial scale calculated by the matched filtering method is smaller than a threshold. The number of iterations n_{itr} triggered after the peak search method can be a constant number or proportional to the number t_{trig} of triggers of the peak search method, i.e.

$$n_{\mathrm{itr}} = 100 + 2500\frac{1 - e^{0.05 t_{\mathrm{trig}}}}{1 + e^{0.05 t_{\mathrm{trig}}}}. \tag{15}$$

The specific form of Equation (15) is not important (Zhang 2018), it will have slight differences in performance with different data, but the important thing is that delta functions are used to represent compact emission. The use of delta functions to represent compact emission has been proven to be effective (Högbom 1974; Schwab 1984), and the peak search method is significantly faster than explicit fitting in a model component search. However, the expression of diffuse emission is more effective in constructing adaptive scale models using explicit fitting. The combination of the two has been proven to have a large improvement in model representation and computing performance (Zhang 2018).

There is only one zero-scale in the scale-free sky model and a few scales specified by the user in the multi-scale sky model. These specified scales do not adequately represent radio sky with uncertain scales. The scales inconsistent with the enumerated scales will be broken into several specified scales, which means unmeasured spatial frequencies cannot be accurately predicted and the resulting model is not optimal. The scales of the sky model constructed by our method are adaptive to the nature of the observed image, not limited to the preset scale, which just solves the problem of scale-free and multi-scale key model representation.

Once a model component is found, the model and residual images need to be updated. Before updating the model image, the new model component will be multiplied by a factor term called loop gain,

$$I_i^{\text{model}} = I_{i-1}^{\text{model}} + kI_i^{\text{comp}} \tag{16}$$

where k is the loop gain factor term and its value is between 0 and 1. This defines the step size of the gradient update to prevent overshoot and divergence. In order to reduce the effect of the negative component, the loop gain of the negative component is half of that of the positive component. This can effectively reduce the negative components and increase the extraction of the positive components, so that the model tends to be as positive as possible, which helps to make strict non-negative restrictions in the major cycle. In order to eliminate the influence of the dirty beam of the current model and find new model components, the residual I_i^{residual} is updated,

$$I_i^{\text{residual}} = I^{\text{dirty}} - B^{\text{dirty}} I_i^{\text{model}}. \tag{17}$$

When we are unable to extract a reliable signal from the current residuals (e.g. the current residual peak is a ratio (such as 10%) of the residual peak calculated from the visibility domain), and then perform corrections in the visibility domain. This requires model prediction and then calculation of visibility residuals.

We retain the possibility of negative components during the model search phase, which helps to increase the robustness of our algorithm. However, the negative components are not physical, so the negative components in the model need to be eliminated before the prediction,

$$I_j^{\text{p}-\text{model}} = PI_i^{\text{model}}, \tag{18}$$

where $I_j^{\text{p}-\text{model}}$ is the model modified before the jth prediction, and P is the modification operation which makes the model nonnegative. In the modified model, emission is restricted to positive, which is more consistent with astrophysics. In order to correct the errors of operations from visibilities to images, the current model needs to be predicted onto the measurement points. The predicted visibilities

$$V_j^{\text{pre}} = AI_j^{\text{p}-\text{model}}, \tag{19}$$

where A represents the set of operations from the model image to the predicted visibilities. This is the inverse of computing the dirty image from visibilities (Venkata 2010; Zhang et al. 2020). Then we use the predicted visibilities to calculate the visibility residuals,

$$V_j^{\text{residual}} = V^{\text{obs}} - V_j^{\text{pre}}. \tag{20}$$

Then we return to the first step to calculate the residual image to correct the errors.

Our algorithm is capable of processing wideband or multi-frequency observations and building a physically adaptive-scale sky model. The adaptive-scale sky model constructed by our algorithm in this paper has no negative components allowed in the current other adaptive-scale sky model algorithms (Bhatnagar & Cornwell 2004; Zhang et al. 2016; Zhang 2018). In Bhatnagar & Cornwell (2004), the active-set is used to construct orthogonal basis functions to reconstruct an adaptive-scale radio sky, however, building an orthogonal adaptive scale sky is very time-consuming. As in Zhang (2018), our algorithm relaxes the orthogonal assumption and significantly accelerates the reconstruction process while maintaining the reconstruction quality. Instead of calculating a model component after obtaining a gaussian component from the dirty image, a model component in our method is directly optimized during the fitting process. This eliminates the manual critical judgment rules for distinguishing between compact and extended features, making reconstruction more accurate and automated. Another notable difference is that our algorithm can handle wideband or multi-frequency observations while other adaptive-scale sky models are only studied in narrowband mode. The next section will demonstrate the performance of our algorithm with simulated SKA data and real wideband observations of the JVLA.

4 RESULTS AND DISCUSSION

4.1 Performance on Simulated SKA Data

In order to show the ability of the sky model construction of this algorithm proposed in this paper, here we apply simulated SKA data where the reference/true image is known. Using the Radio Astronomy Simulation, Calibration and Imaging Library (RASCIL) [1], we make an SKA observation simulation with the configuration ‘LOWBD2’ on this widely researched medium-complex radio source ‘M31’ (Fig. 2 *left*). This simulation was

[1] *https://developer.skatelescope.org/projects/sim-tools/en/latest/* or *https://github.com/SKA-ScienceDataProcessor/rascil*; this library is officially developed by the SKA for radio interferometry calibration and imaging algorithms using Python and numpy.

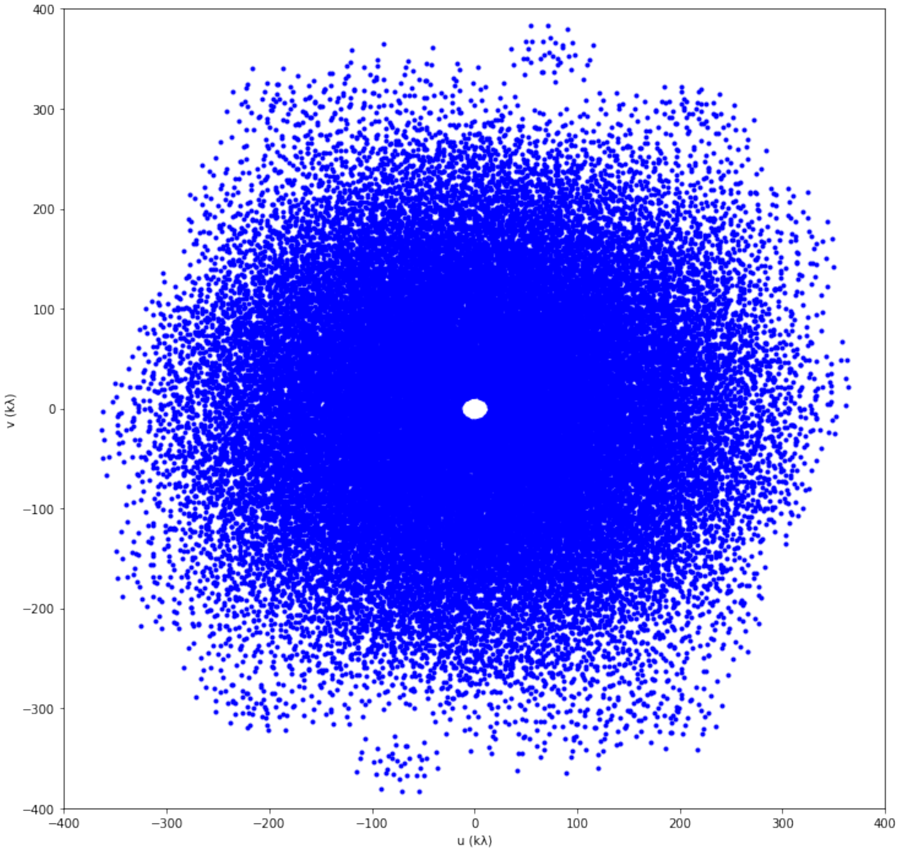

Fig. 1: This is the UV coverage from the simulation of SKA observation at 100 megahertz with the 'LOWBD2' configuration.

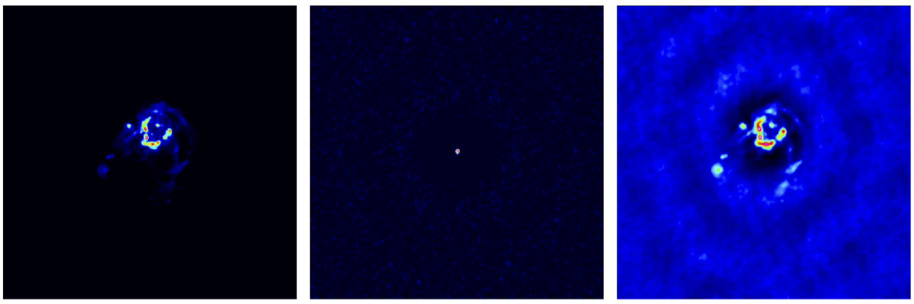

Fig. 2: The simulation of SKA observation. *Left*: the reference /true image 'M31' used in this simulation, *middle*: the spread point function with the logarithmic scaling (CASA scaling power cycles $= -0.9$) to show more details of these sidelobes, *right*: the dirty image corrupted by the point spread function (PSF).

performed at a center frequency of 100 megahertz and a bandwidth of 1 megahertz. This source is considered to be at right ascension 15 degrees and declination -45 degrees. This simulated UV coverage is shown in Figure 1 and the simulated observation image (i.e. dirty image) is shown in Figure 2 *right*. We then reconstruct a sky model from the dirty image that is close to the reference image, which is also the work of deconvolution.

In Figure 3, we show the reconstruction results from the multiscale CLEAN and our algorithms. From these model images, our algorithm can reconstruct the extended features very well. We can see that our algorithm can construct a model close to the reference image. Obviously, the residual image of our algorithm contains fewer signals, which shows that our algorithm extracts signals more fully. In reconstruction, we use a CLEAN window of $1/4$ image size. The results show that the residual image obtained by our algorithm has less signal residue. The final restored image is shown in Figure 4. Our restored image (Fig. 4 *right*) shows that our algorithm has well eliminated the

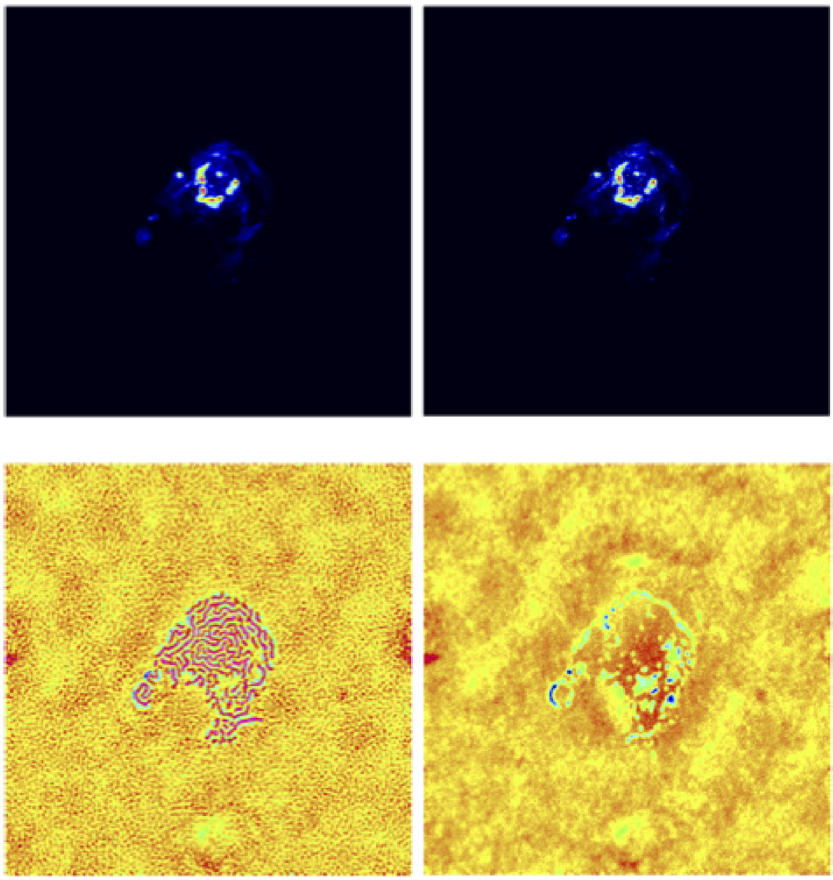

Fig. 3: Model and residual images from multiscale CLEAN and our algorithms. In *the first column* from the multiscale CLEAN algorithm: *top*: the model image, *bottom*: the residual image. In *the second column* from our algorithm: *top*: the model image, *bottom*: the residual image.

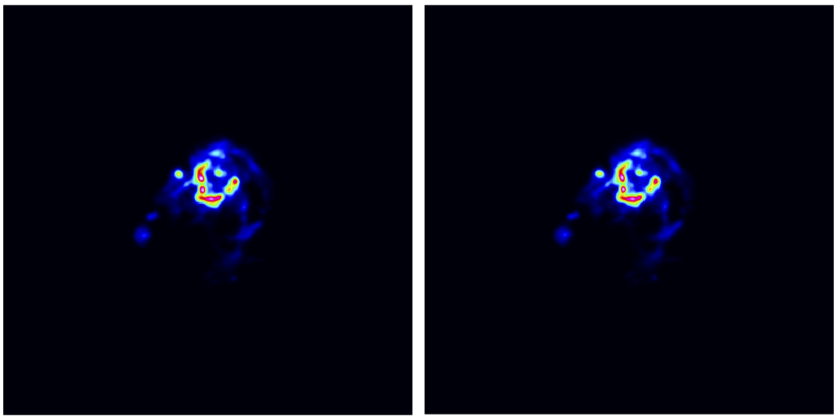

Fig. 4: Restored Images. *Left*: from the multiscale CLEAN algorithm, *right*: from our algorithm.

sidelobe effects in the dirty image displayed in Figure 2, so that there are no longer significant features outside the source.

Table 1 records the residual RMS and dynamic range of the restored images of different algorithms. In this example, our algorithm can achieve over 66% improvement in dynamic range and over 30% reduction

Table 1: Numerical Comparison of Different Algorithms for the SKA Simulation

	Off-source RMS (10^{-3})	Full RMS (10^{-3})	Dynamic Range (10^{3})
Multiscale	4.435	5.982	4.224
Our	2.664	3.962	7.028

Off-source RMS is calculated from the non-source area of the residual image while full RMS is calculated from the entire residual image.

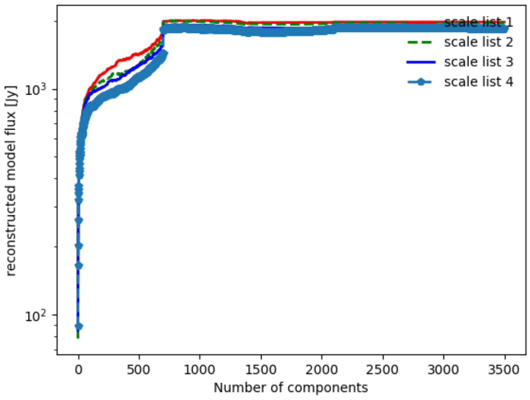

Fig. 5: The behavior of the reconstructed model flux with different initial parameters as the components decrease. Notes: 'scale list 1' =[0,1,2,4] pixels, 'scale list 2' =[0,2,4,6] pixels, 'scale list 3' =[0,2,5,8] pixels, 'scale list 4' =[0,3,6,10] pixels. These scales are the full width at half maximum of the Gaussian functions, which are used to find the best initial fitting parameters in the matched filtering process. Figs. 6 and 7 are the same.

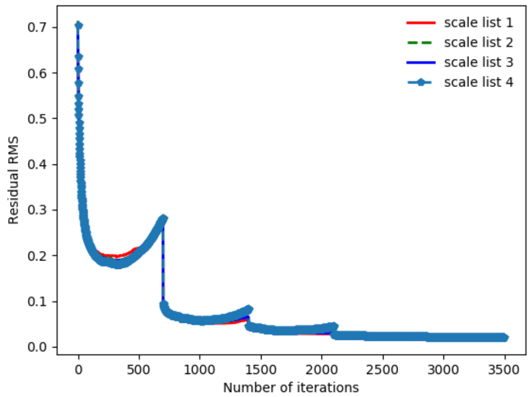

Fig. 6: The RMS behaviors of the entire residual images with different initial parameters.

in residual RMS. This further confirms that our algorithm can obtain better reconstruction quality. This improvement is likely to come from the better representation of adaptive scale model for sky emission.

In our algorithm, the acquisition of optimization components requires initial parameters. The effect of different initial parameters on the accumulation process of the reconstructed model is shown in Figure 5. It can be seen that although different initial parameters have a slight effect on the accumulation process of the reconstructed model, this does not affect the convergence of the algorithm to the same level. If we look at the effect of different parameters on the residual RMS (Fig. 6), it

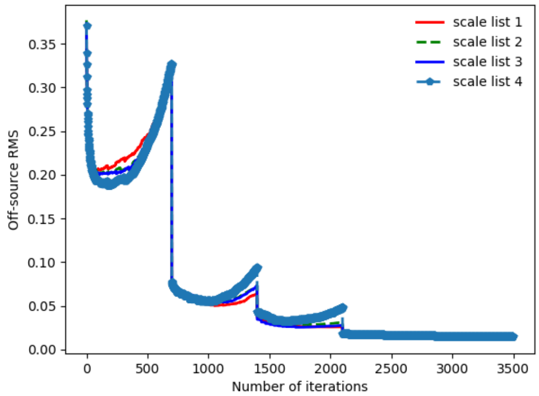

Fig. 7: The off-source RMS behaviors of the residual images with different initial parameters.

can quickly drop to very low levels. At the same time, we can also see that there is some recovery in RMS during the descent, which is caused by the accumulation of errors during model construction, but our algorithm can quickly correct it and continue to move forward until convergence. In addition, RMS is insensitive to these initial parameters. We can get similar conclusions from the residual off-source RMS (Fig. 7). These show that our algorithm is not significantly affected by the parameters, that is, robust. This enables our algorithm to obtain a stable result under different parameters.

4.2 Performance on Real Data

Now that we know the performance of our algorithm in SKA simulated data, we now test its performance in real observation data. The demonstration example for real data is the supernova remnants G55.7+3.4, observed with the D configuration of the JVLA. This 8-hour observation uses the entire 1 GHz bandwidth at the L band from 1 GHz to 2 GHz (Bhatnagar et al. 2011). The frequency resolution is 2 MHz and the time resolution is 1 s. During the imaging, the 1s resolution data is averaged to the 10 s time resolution to reduce the amount of the data. J1925–2106 is used for a bandpass and phase calibrator and 3C147 for standard flux calibration. The robust weighting is used during the gridding. The cell size is 8 arcsec, and the imaging region is approximately 34 arcmin. All observations of different frequency windows are combined by the MFS in our algorithm, which significantly increases the observed UV coverage (Fig. 8). The wideband dirty image shown in Figure 9 has an extended source with different angular scales and many compact sources in the background. The representation of such a sky region essentially requires an

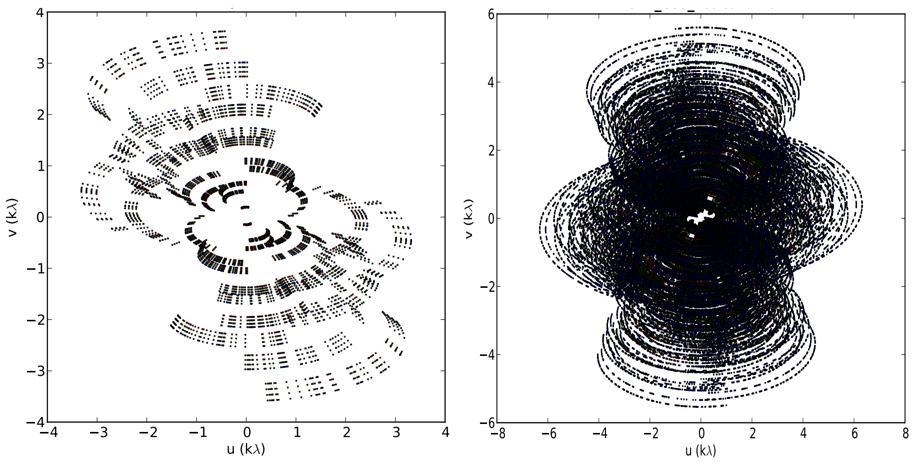

Fig. 8: *Left*: the UV coverage from one frequency measurements, *right*: the full UV coverage, which combines all observations from 1 to 2 GHz by the MFS.

adaptive-scale sky model, which models diffuse emission at different scales and compact emission well.

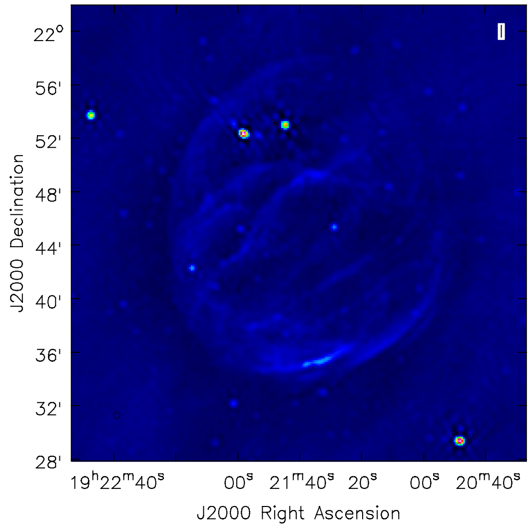

Fig. 9: This is the wideband dirty image, which combines all observations from 1 to 2 GHz by the MFS.

Model images reconstructed from three different sky models are displayed in Figure 10. Obviously, the scale-free model image (Fig. 10 *left*) represents the sky as a collection of points with zero scales (delta functions), which cannot represent the correlation between adjacent pixels in such an extended emission as G55.7+3.4. In addition, the scale-free sky model has difficulty reconstructing faint extended features. Compared to the scale-free sky model, the model image reconstructed by the multi-scale sky model (Fig. 10 *middle*) has improved significantly on the representation of the extended emission. However, the enumeration scale of the multi-scale sky model often keeps the reconstructed model image away from the intrinsic scale characteristics of astrophysical targets. It is obvious that the adaptive-scale model image (Fig. 10 *right*) is the best reconstructed model image, which is basically consistent with the nature of astrophysical targets and has rich details. This shows that our proposed adaptive scale sky model can reconstruct the intrinsic scale characteristics of astrophysical targets well.

In addition, scale-free and multi-scale sky models allow negative components in model images. These negative structures can be seen in the beamed model images (shown in Fig. 11), which are obtained by convolving model images and the restored beam. A physically credible model should not contain any negative structures, so we removed the negative components from the beamed model images shown in Figure 11. Therefore, these non-negative model images in Figure 12 are astrophysical. As can be seen from Figure 12, the scale-free model image recovers the fewest physically credible features, while the adaptive-scale model image recovers the richest detail and looks like more consistent with the features contained in the observed/dirty image. This happens because the adaptive-scale sky model presented in this paper can represent the intrinsic scale characteristics of astrophysical targets well and does not allow the presence of non-physical negative components, which makes the reconstructed model images closer to astrophysical targets. Accurate reconstruction of the adaptive scale sky model will provide more powerful material for astrophysical research.

5 SUMMARY

In this paper, we propose a new adaptive-scale sky model and apply it to data processing for simulated SKA observation and real wideband measurements. Compared to scale-free and multi-scale sky models, the adaptive-scale sky model can better represent the sky. In the

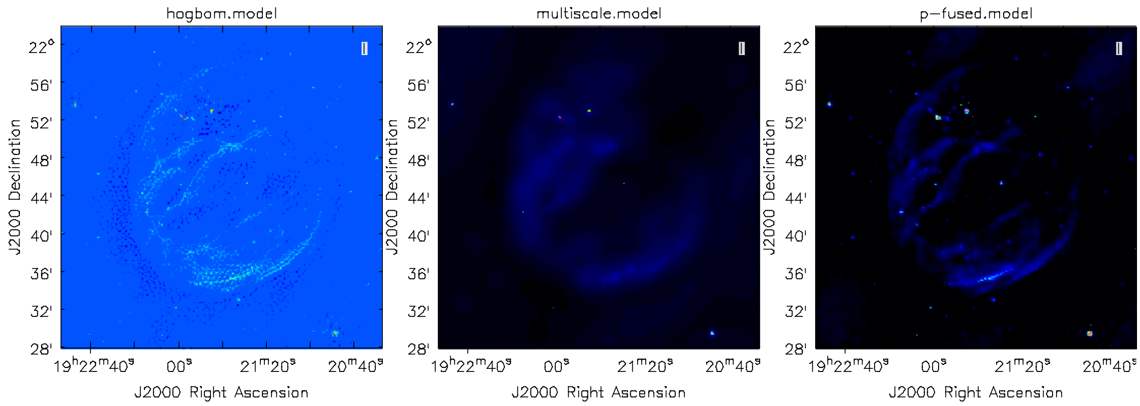

Fig. 10: These model images from different sky models. *Left:* the scale-free model image, *middle:* the multi-scale model image, *right:* the adaptive-scale model image reconstructed by our algorithm.

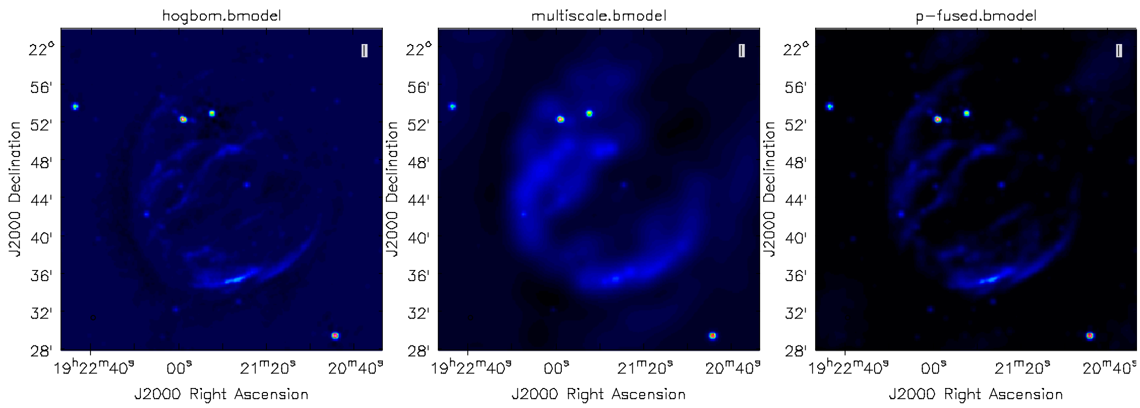

Fig. 11: These beamed model images from different sky models. *Left:* the scale-free beamed model image, *middle:* the multi-scale beamed model image, *right:* the adaptive-scale beamed model image reconstructed by our algorithm.

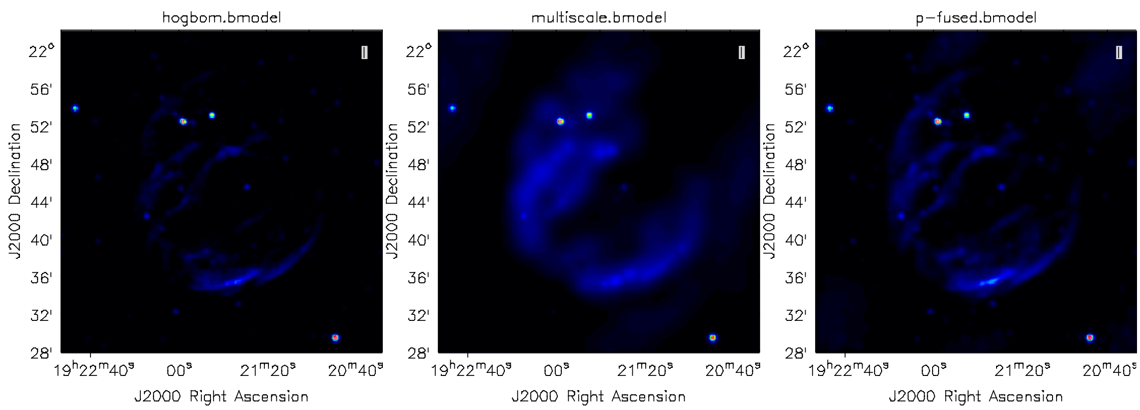

Fig. 12: These non-negative beamed model images from different sky models, which are physical representations. *Left:* the physical scale-free beamed model image, *middle:* the physical multi-scale beamed model image, *right:* the physical adaptive-scale beamed model image reconstructed by our algorithm.

other adaptive-scale sky models, negative components are allowed in the final model image, but the negative components do not conform to the nature of astrophysics. So we introduce positive restrictions on emission, which makes the adaptive-scale sky model physically represent the sky, which will make the sky model constructed by our algorithm more consistent with the nature of astrophysics. These are very important for high dynamic range imaging. In addition, the current other adaptive-scale sky models are designed for narrowband simulation observations, which cannot meet the requirements of SKA's wideband observations. However, our algorithm can have wideband imaging capabilities, which effectively combines observations of different frequencies, which is conducive to increasing measurement sensitivity and coping with bandwidth smearing issues. Effective sky models and processing capabilities for wideband observations are very important for SKA imaging. This is also quite useful for the development of the SKA imaging pipeline.

Acknowledgements This work was partially supported by the National Key R&D Program of China (2018YFA0404602 and 2018YFA0404603), the National SKA Program of China (2020SKA0110300), the National Natural Science Foundation of China (NSFC, 11963003, 11763002, 61572461, 11790305, U1831204, U1931141, 11961141001 and 11903009), the Guizhou Science & Technology Plan Project (Platform Talent No.[2017]5788, [2017]5781), the Youth Science & Technology Talents Development Project of Guizhou Education Department (No. KY[2018]119 and [2018]433), and the Guizhou University Talent Research Fund (No.(2018)60).

References

An, T. 2019, Science China Physics, Mechanics, and Astronomy, 62, 989531

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

Bhatnagar, S., Rau, U., Green, D. A., & Rupen, M. P. 2011, ApJL, 739, L20

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

Clark, B. 1982, in Synthesis Mapping, In: Synthesis Mapping, NRAO Workshop No. 5, eds Thompson, A. R., & DAddario, L. R., NRAO, Greenbank, West Virginia, 10

Conway, J. E., Cornwell, T. J., & Wilkinson, P. N. 1990, MNRAS, 246, 490

Cornwell, T. J. 1983, A&A, 121, 281

Cornwell, T. J. 2008, IEEE Journal of Selected Topics in Signal Processing, 2, 793

Högbom, J. A. 1974, A&AS, 15, 417

Marquardt, D. W. 1963, Journal of the Society for Industrial and Applied Mathematics, 11, 431

Mei, Y., Wang, F., Wang, W., et al. 2018, PASP, 130, 014503

Palagi, F. 1982, A&AS, 49, 101

Pawsey, J. L., Payne-Scott, R., & McCready, L. L. 1946, Nature, 157, 158

Rau, U., & Cornwell, T. J. 2011, A&A, 532, A71

Ryle, M., & Vonberg, D. D. 1948, Proceedings of the Royal Society of London Series A, 193, 98

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

Steer, D. G., Dewdney, P. E., & Ito, M. R. 1984, A&A, 137, 159

Taylor, G. B., Carilli, C. L., & Perley, R. A. 1999, in Astronomical Society of the Pacific Conference Series, 180, Synthesis Imaging in Radio Astronomy II

Thompson, A. R., Moran, J. M., & Swenson, George W., J. 2017, Interferometry and Synthesis in Radio Astronomy (3rd ed. Springer)

Venkata, U. R. 2010, Parameterized Deconvolution for Wide-Band Radio Synthesis Imaging, PhD thesis, New Mexico Institute of Mining and Technology

Wu, X. P. 2019, China SKA Science Report (Science press, ISBN: 9787030629791)

Zhang, L. 2018, A&A, 618, A117

Zhang, L., Bhatnagar, S., Rau, U., & Zhang, M. 2016, A&A, 592, A128

Zhang, L., Xu, L., & Zhang, M. 2020, PASP, 132, 041001

Zhang, L., Xu, L., Zhang, M., & Wu, Z.-Z. 2019, RAA (Research in Astronomy and Astrophysics), 19, 079