

Revisiting a flux recovery systematic error arising from common deconvolution methods used in aperture-synthesis imaging

Jack F. Radcliffe,^{1,2,3}★ R. J. Beswick,² A. P. Thomson,² A. Njeri,^{2,4} and T. W. B. Muxlow²

¹*Department of Physics, University of Pretoria, Lynnwood Road, Hatfield, Pretoria, 0083, South Africa*

²*Jodrell Bank Centre for Astrophysics, University of Manchester, Oxford Road, Manchester M13 9PL, UK*

³*National Institute for Theoretical and Computational Sciences (NITheCS) South Africa*

⁴*School of Mathematics, Statistics and Physics, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK*

Accepted 2023 August 31. Received 2023 August 31; in original form 2022 November 28

ABSTRACT

The point-spread function (PSF) is a fundamental property of any astronomical instrument. In interferometers, differing array configurations combined with their uv coverage, and various weighting schemes can produce an irregular but deterministic PSF. As a result, the PSF is often deconvolved using CLEAN-style algorithms to improve image fidelity. In this paper, we revisit a significant effect that causes the flux densities measured with any interferometer to be systematically offset from the true values. Using a suite of carefully controlled simulations, we show that the systematic offset originates from a mismatch in the units of the image produced by these CLEAN-style algorithms. We illustrate that this systematic error can be significant, ranging from a few to tens of per cent. Accounting for this effect is important for current and future interferometric arrays, such as MeerKAT, LOFAR and the SKA, whose core-dominated configuration naturally causes an irregular PSF. We show that this offset is independent of other systematics, and can worsen due to some factors such as the goodness of the fit to the PSF, the deconvolution depth, and the signal-to-noise of the source. Finally, we present several methods that can reduce this effect to just a few per cent.

Key words: techniques: interferometric, photometric – methods: data analysis

1 INTRODUCTION

The point spread function (PSF) is a fundamental property of any astronomical instrument and describes the response of an instrument to a normalised point source. It characterises how much the true sky brightness distribution is distorted by the instrument and is often used to define the angular resolution of the instrument. Understanding and removing the PSF contribution is of critical importance in extracting accurate photometry from your image. For single aperture instruments, the PSF shape is dependent on many factors such as obscuration, aberrations, and pointing errors. This is exacerbated for ground-based instruments where thermal and gravitational effects come into play. As a result, the PSF is often time-varying and thus sophisticated modelling is required to characterise it properly (e.g., [Krist et al. 2011](#)). However, single-aperture instruments have one key advantage and that is a filled aperture. This means that the resultant PSF is often well-behaved, with most of its amplitude located in a main directional lobe and, with minimal side lobes that only affect high dynamic range imaging.

In contrast, the PSF of an interferometer (often known as the ‘dirty beam’ or ‘synthesised beam’) does not have this luxury, and the incomplete Fourier sampling of the aperture means that the PSF often has significant sidelobe structure that can affect the extraction of accurate photometry. However, the PSF of an interferometer is highly deterministic and can be calculated exactly from the visibilities measured by the interferometer. This means that interferometric images

can often be improved by deconvolving the PSF from the image. The most commonly used method of deconvolution is the CLEAN family of algorithms. These algorithms deconvolve through a process of iterative peak finding and PSF subtraction, down to some predetermined threshold (also known as ‘matching pursuit’). However, the estimated sky brightness model obtained from these algorithms is typically comprised of delta functions and/or Gaussians which can often look distinctly unphysical. As a result, a ‘restoration’ step is usually employed. In this step, the model is convolved with a 2D Gaussian fitted to the PSF that estimates the effective resolution of the interferometer as if it had a filled aperture. This is then added to the subtracted image (which can contain low-level unresolved flux) and this restored image is then used for science.

However, this restoration step can result in a systematic error that occurs when extracting the flux densities of sources. This was noted early when the Fourier sampling was sparse (e.g., [Högbom 1974](#)). As a result, early array designs were often optimised to generate an approximately Gaussian PSF (e.g., the Very Large Array - VLA). The error was first quantified by [Jörsäter & van Moorsel \(1995\)](#) who found that there was a mismatch between the flux density of an input model source and the resultant recovered flux densities when CLEAN deconvolution was used. This was found to be from contamination by faint flux that had not been deconvolved. While this effect has been routinely corrected for in spectral line studies (e.g., [Kennicutt et al. 2007](#); [Cannon et al. 2009](#); [Hunter et al. 2012](#); [de Blok et al. 2018](#); [Novak et al. 2020](#)), it had sparingly been taken into account during continuum imaging studies (e.g., [Benisty et al. 2021](#); [Heywood et al. 2022](#); [Booth et al. 2023](#)).

★ E-mail: jack.radcliffe@up.ac.za

This is understandable, as the effect of this systematic error had been lessened through the improved uv coverage and array designs that resulted in PSFs that more closely resemble a filled aperture. However, the designs of modern interferometric arrays, such as MeerKAT, ASKAP, LOFAR, ALMA, the upcoming Square Kilometre Array (SKA) and next-generation VLA (ngVLA), are driven by multiple, diverse scientific objectives that require a high sensitivity to diffuse structures, and high angular resolutions, and non-imaging applications, such as pulsar timing, all at the same time. These competing scientific requirements have resulted in the configuration of these arrays being a compromise, resulting in the majority being core-dominated with a sparser distribution of antennas on longer baselines. As a result, the large number of shorter baselines produces a naturally irregular PSF.

It is in this context that we are re-examining this systematic effect as will be important to account for when using these modern interferometric arrays (as long as the use of CLEAN-style deconvolution methods persists). In this paper, we shall use a controlled simulation of a core-dominated radio interferometer array to quantify this effect on recovered flux densities, and present some methods that can help correct it so that interferometric surveys can take this systematic error into account.

The paper is organised as follows. In Section 2, we introduce the simulations used. Our results are shown in Section 3, and we explain the origin of the offsets in Section 4. We present the methods to solve for the effects of the PSF in Section 5, and discuss the implications and limitations of these results in Section 6.

Throughout this paper, we define the spectral index, α , as $S_\nu \propto \nu^\alpha$ where S_ν is the specific integrated flux density (per unit frequency, ν). We use the convention used in the Common Astronomy Software Applications (CASA) package (The CASA Team et al. 2022) for the robustness factor (R) of the Briggs’s weighting scheme (Briggs et al. 1999) which is defined to be between -2 and 2 . These values approach uniform and natural weighting, respectively.

2 SIMULATIONS

For the simulations, we opted to generate a combined VLA (in A-configuration) and e -MERLIN array. This combination has a well-documented irregular PSF due to the short spacings of the VLA compared to the e -MERLIN array (e.g., see Muxlow et al. 2020). We also have generated a representative MeerKAT array to prove that this effect is evident in modern core-dominated interferometers (see Section 6).

To build the simulated data, `simms`¹ was used to create empty CASA measurement sets with a delay-tracking centre located at a Right Ascension of 12^{h} and a Declination of $+60^\circ$ (or -30° for MeerKAT), and a total on-source integration time of 12 hours. These data sets have four correlations present with circular feeds for VLA and e -MERLIN, and linear feeds for MeerKAT. We used a bandwidth representative of a standard L-band (~ 1.5 GHz) observation with each array. The long integration times (10 s) and a small number of frequency channels were used to keep the data sizes as small as possible without adversely changing the PSF profile. The exact specifications of each simulated data set are summarised in Tables 1 and 2.

Thermal noise was inserted into the visibilities for each baseline, V_{pq} , that comprises antennas p and q . We assumed that the system

Table 1. The properties of the simulated data sets. The sensitivities were calculated using a natural weighting scheme when imaging.

Array	Frequency (GHz)	SEFD (Jy)	# channels	Sensitivity ($\mu\text{Jy beam}^{-1}$)
VLA	1.024–2.048	420	32	1.91
e -MERLIN	1.254–1.766	see Table 2	64	6.04
MeerKAT	0.900–1.677	450	32	1.15

Table 2. System equivalent flux densities used for the e -MERLIN antennas.

Antenna	SEFD (Jy)
Lovell	40
Knockin	400
Pickmere	450
Darnhall	450
Defford	350
Cambridge	175

thermal noise was broadband and so can be described using a circularly complex Gaussian probability distribution \mathcal{N} that has a zero mean and a variance of σ_{pq}^2 per visibility. The standard deviation of the visibility scatter, σ_{pq} , is given by the radiometer equation that, per visibility, is given by,

$$\sigma_{pq} = \frac{1}{\eta_c} \left(\frac{\text{SEFD}_p \times \text{SEFD}_q}{2t_{\text{int}}\delta\nu} \right)^{\frac{1}{2}}, \quad (1)$$

where SEFD is the antenna’s System Equivalent Flux Density, η_c is the efficiency factor (taken to be 0.88 assuming 4-level quantisation), t_{int} is the integration time, and $\delta\nu$ is the channel bandwidth. This was then added into the visibilities of the measurement set as an additive term per polarisation,

$$V'_{pq} = V_{pq} + \mathcal{N}(0, \sigma_{pq}^2). \quad (2)$$

To ensure that the natural weighting scheme generates an image with the optimal sensitivity, the weights of each visibility, W_{pq} , were set such that,

$$W_{pq} = \frac{1}{\sigma_{pq}^2}. \quad (3)$$

This is important for the heterogeneous e -MERLIN array where the majority of the sensitivity is provided by the baselines to the more sensitive Lovell and Cambridge telescopes. These visibilities should now be representative of the thermal noise of the array given their nominal antenna sensitivities. Next, elevation-dependent flagging was conducted by removing data from each antenna when the pointing centre is below 5 degrees elevation for that antenna.

With the noise inserted into the data, we then generated the sky model. To keep the number of variables as small as possible, our input sky model was comprised of a total of 13,500 point sources. This was split into three individual realisations each with 4,500 sources. Point sources were selected as our model to eliminate any complications arising from the ‘resolving out’ of flux due to the incomplete Fourier sampling of an interferometric array. These sources had randomly assigned integrated flux densities which were designed to cover a signal-to-noise (S/N) range of 5–70. The sources were distributed randomly with the caveat that each source was at least $10''$ from its nearest neighbour. This was chosen to eliminate any complications

¹ <https://github.com/ratt-ru/simms>

arising from source confusion when attempting to measure the individual flux densities. In addition, the point sources were defined to be located at an integer number of pixels in the model image to prevent any pixelation errors when it was transformed into model visibilities.

This sky model was inserted into the measurement sets by converting the model sky (via de-gridding) into model visibilities using the `WSCLEAN` (Offringa et al. 2014) package. This method ensures that wide-field effects, such as baseline non-coplanarity (known as the w -term), are not present in the modelled sky. This model was then added into the visibilities using the CASA task `UVSUB`. For the subsequent imaging runs, `WSCLEAN` was used due to its versatility and speed. We cross-checked the imaging outputs of `WSCLEAN` and CASA's `TCLEAN` task when using identical parameters and we found an average fractional difference of just 0.3%. We are therefore confident that our results are minimally impacted by the choice of imaging software. The exact imaging runs are presented in Section 3.

3 RESULTS

To show the effects of an irregular PSF upon the resultant measured flux densities, the combined VLA and e -MERLIN array were imaged using a robust value of zero. This produces an irregular PSF with a secondary component on larger angular scales that is caused by the short spacings provided by the VLA. This is shown in the right-hand panel of Figure 1 which shows a 1D slice through the PSF. Each image was $25k \times 25k$ pixels in size with each pixel corresponding to an angular size of $0''.05$. These choices enable all 4,500 sources of each realisation to be imaged at once. The PSF was deconvolved from the source brightness using the Clark algorithm (Clark 1980). We used multiple different deconvolution thresholds (i.e., the flux density level where deconvolution stops) ranging from 1–4 times the local root-mean-square (rms) noise levels (σ) so we can illustrate how the deconvolution process can drastically affect the recovered fluxes. Once deconvolution had ceased, the source brightnesses were restored by convolving the model with two different 2D Gaussian fits to the PSF; a $\sim 0''.3$ full-width half-maximum (FWHM) fit that was produced CASA's `TCLEAN` task and a $\sim 0''.6$ FWHM fit from `WSCLEAN`. The fact that two different imaging packages produce completely different fits is indicative of the difficulty of assuming a Gaussian fit to a highly non-Gaussian PSF. The convolved model is then added back to the residual image (where the source flux densities have been removed and deconvolved). To summarise, we produced a total of eight images, with four different deconvolution thresholds, and two different fits to the PSF.

Flux density extraction was conducted using aperture photometry (implemented using `PHOTUTILS`; Bradley et al. 2020). We used an aperture that is 2.5 times larger than the major axis of the fitted PSF. This enables us to recover $> 99\%$ of the source's flux density while compensating for the widening of the source structure due to the convolution of the model sky with the restoring beam. By using aperture photometry and the knowledge of where our sources are, we can mitigate the well-known effect of 'flux-boosting' (e.g., Jauncey 1968). This occurs because typical source-detection routines use a sigma detection threshold to determine what is a true source above the noise. In the low S/N regime, local noise variations can cause some sources to be co-located with a noise peak which results in the source exceeding the detection threshold, while those sources on a negative noise peak would not be detected. This will bias the distribution of sources upward in flux density in the low S/N regime. In addition, the fitting routines often employed perform inadequately in the low S/N regime (e.g., Hale et al. 2019). Aperture photometry will not

suffer from this issue as the noise contribution should (statistically) average to zero, while all the real flux located within the aperture will be detected. In addition, no sources will be missed (e.g., due to being co-located on a negative noise peak) as we know the source positions beforehand.

Figure 1 shows the results of the flux extraction for the eight images described earlier. In the left and centre panels, we present the ratio of measured flux densities (S_{measured}) to the model flux densities (S_{model}) against the S/N ratio for a range of deconvolution thresholds. The flux densities of the sources are median binned into 7 bins with each containing approximately an equal number of sources. The error bars and filled-in areas correspond to the 1σ standard deviation of the flux densities within each bin. In the right panel, we show a 1D slice of the PSF with the corresponding Gaussian fit which is used to produce the output images.

We would expect that the ratio of $S_{\text{measured}}/S_{\text{model}}$ would be approximately equal to one if the flux densities were recovered perfectly. However, as we can see in the left and centre panels of Figure 1, the flux densities recovered differ significantly from those inputted. There are four main results from this that we need to draw attention to.

Firstly, regardless of the fit, the recovered flux densities are different to the model flux densities. These offsets can be significant and can range from just 5% to 250%. Such a systematic error can surpass the typical amplitude errors of interferometers which are often taken to be of the order of 10%. Secondly, the extent of the flux density offsets worsens when we deconvolve to higher thresholds (or a 'shallower' deconvolution). This indicates that the flux densities recovered in any interferometric array are highly dependent on the deconvolution depth. For interferometric arrays with a large number of elements (e.g., LOFAR, MeerKAT, SKA), the additional computational requirements to deconvolve to much lower flux densities could be prohibitive, subsequently adversely affecting the measured flux densities of these arrays.

Thirdly, the flux density offset becomes more severe towards the low S/N regime showing that flux uncertainties at the noise limit are increasingly uncertain irrespective of the increasing fitting errors. The PSF must be taken into account within this regime. Finally, the flux density offsets are more extreme when the fit between the restoring beam is worse. In the next section, we shall explain how the flux density systematics arise and their relationship to both the S/N and deconvolution.

4 THE ORIGIN OF THE MEASURED FLUX DENSITY ERRORS

As inferred previously, these trends are all related to the deconvolution method used. To illustrate this, we are going to use a toy model of the deconvolution process. For simplicity, we are going to assume that there is no thermal noise present and that we are observing a simple 1 Jy point source. This source is observed with an instrument that has the same irregular PSF profile as shown in the right panel of Figure 1. As expected, the measured sky brightness is just equal to the PSF response. The key here is that the interferometer measures the source brightness in units of Jansky per PSF.

The PSF contribution is then deconvolved from the true source brightness using the Högbom (1974) CLEAN algorithm. In this method and many other deconvolution methods used in radio interferometry, the pixel of peak brightness is found and then a percentage of this brightness is convolved with the PSF and then subtracted from the image. The value and position of the brightness that is re-

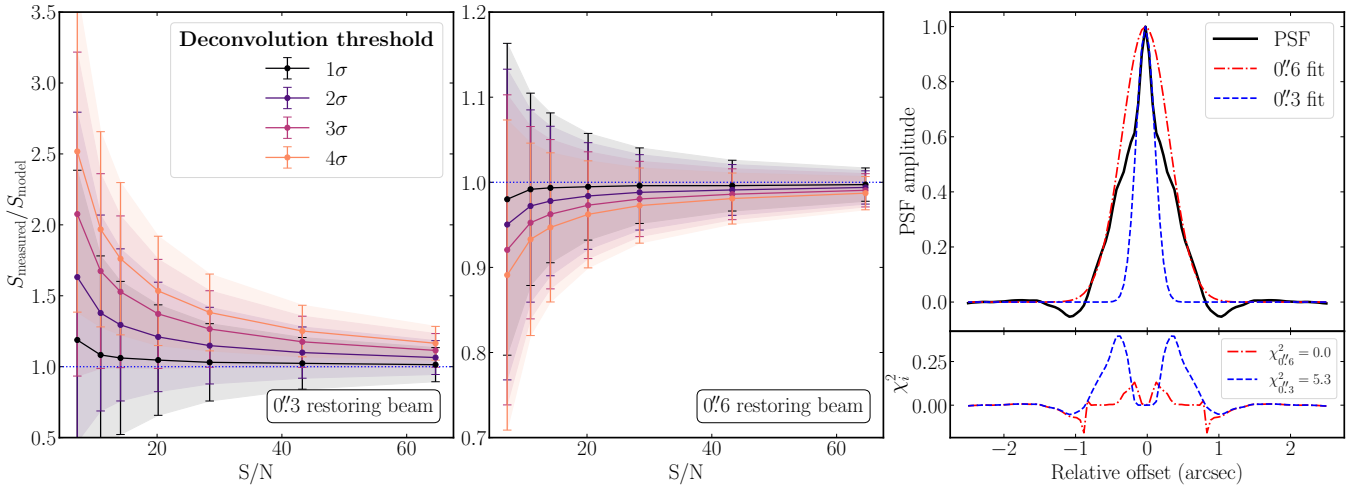


Figure 1. *Left and centre panels:* The flux density recovery performance of the simulated combined VLA and *e*-MERLIN arrays using two different Gaussian fits to the PSF. The binned S/N of the 13,500 sources are plotted against the ratio of the measured to model flux densities for a range of deconvolution thresholds. The error bars and shaded regions correspond to 1σ deviations within the bins. The fits are ~ 0.3 (left panel) and ~ 0.6 (centre panel) which correspond to the resultant fits by CASA’s `TCLEAN` task and the stand-alone `WSCLEAN` package. *Right panel:* A 1D slice through the PSF with the resultant fits overlaid. The bottom panel shows the residuals are plotted per pixel ($\chi_i^2 = (\text{PSF} - \text{fit})^2 / \text{PSF}$) and the $\chi^2 = \sum \chi_i^2$ values are presented.

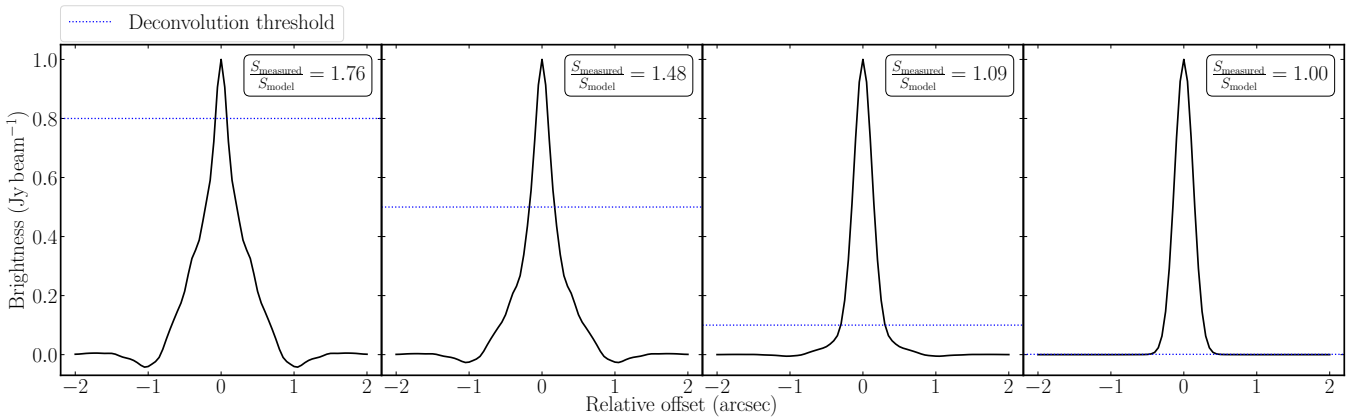


Figure 2. A 1D toy model of the deconvolution process using the PSF profile presented in Figure 1. The deconvolution threshold is shown by the blue dotted line. As deconvolution proceeds to lower thresholds (from the left to the right panel), the number of beam widths which the flux density is measured across tends to be one. Theoretically, we would get a perfect recovery of the flux density if we were able to deconvolve to a 0 Jy threshold.

moved are then stored as a sky brightness model (as delta functions in the case of the Högbom (1974) algorithm). This process is then re-iterated until a stopping criterion is met. Often this criterion is a threshold where deconvolution stops when the brightest pixel in the subtracted image falls below a pre-defined value. This threshold is often a few σ above the noise for multiple reasons such as reducing computational expenditure, preventing the divergence of the deconvolution process, ensuring that a reliable sky model is generated (e.g., for self-calibration), or preventing unnecessary over-deconvolution which can lead to other systematics such as CLEAN bias (e.g., Becker et al. 1995).

The model sky brightness is then convolved with a restoring beam which comes from a Gaussian fit to the PSF and is added back to the residuals (the image after the PSF multiplied by the brightness is subtracted). A Gaussian is often chosen as it is a good approximation for the idealised response of the interferometer as if it has complete *uv*

coverage. The restored image is assumed to have units of Jansky per beam where the beam is defined by the angular size of the Gaussian. A flux density is then measured by summing the brightnesses over the number of beams that the source brightness extends over.

For example, in our toy model, the PSF is fitted with a ~ 0.3 Gaussian that is used as the restoring beam. This gives us a fixed solid angle which we can integrate over to obtain a flux density. In Figure 2, we show the deconvolution process at four progressively lower thresholds. At the lowest threshold (0.001 Jy; first panel from the right in Fig. 2), the brightness in the output image covers almost a single beam area therefore we recover the input flux density exactly. However, when the PSF is still present in the output image (as in the other three panels), the brightness of the source is summed over more than one beam area hence a higher flux density is measured when compared to the input model. This explains why the 0.6 fit, which is larger than the true PSF, causes a reverse effect where the measured

flux density is smaller than the input model flux density (as shown in the central panel of Fig. 1). In this case, there are fewer beam elements that the brightness is observed over, hence the recovered flux density is measured to be lower than the model.

In other words, above the deconvolution threshold, the recovered sky brightness has units of Jy beam^{-1} , but below the deconvolution threshold, the flux density is still in units of Jy PSF^{-1} . This means that the source of this error occurs when the residuals (which has units of Jy PSF^{-1}) are added back to the convolved model (which has units of Jy beam^{-1}), but we extract the flux densities from the image assuming units of Jy beam^{-1} throughout. The subsequent mismatch between PSF and beam area below the deconvolution threshold, therefore, causes the flux density to be miscalculated. These flux density offsets get progressively worse as the deconvolution threshold increases, as is shown in Figure 2. This is because the percentage of flux restored with the fitted beam width/solid angle (i.e., converted from units of Jy PSF^{-1} to Jy beam^{-1}) is increased when using a lower threshold. This argument also explains the trends with S/N seen in Figure 1. When a fixed deconvolution threshold level is used, sources with a lower S/N have less of their total flux density deconvolved and restored with the fitted beam which has a well-defined solid angle. This results in a larger proportion of the flux density being incorrectly estimated once measured.

5 POSSIBLE SOLUTIONS

Now that the issue has been outlined and explained, we shall introduce some possible methods that can reduce these flux density offsets. These solutions are not only motivated by the simulations presented but also through experiences with real data.

5.1 Making the PSF closer to a Gaussian through re-weighting

The motivation for this solution comes from the *e*-MERGE survey (Muxlow et al. 2020). This survey uses a combined VLA and *e*-MERLIN array which results in a highly non-Gaussian PSF which caused flux density offsets in these data.² As part of this survey, solutions to this problem were formed and implemented.

There are three main ways in which the flux density offsets can be reduced. The first method is to deconvolve to deeper thresholds. This is shown in the left and centre panels of Figure 1. However, it is worth noting that this may not be possible for real data. Firstly, deconvolution cannot continue indefinitely as the computational expense can become prohibitive but also, for many modern sensitive arrays, such as MeerKAT, it is often residual calibration errors (such as pointing errors) and direction-dependent effects that limit to which deconvolution can be performed.

It is therefore practically impossible to deconvolve all flux (especially that below the noise level) and fully remove this effect. The extent of this systematic could be removed in statistical studies (e.g., source counts) using simulations as the exact extent of the flux density offsets will be sensitive to any process that changes the PSF shape. This would include factors such as the exact *uv* coverage,³ and data weighting. However, on a source-by-source basis, extra care would still need to be taken to account for this systematic effect.

² In particular, see Section 2.5.1 of Muxlow et al. (2020) where this is discussed.

³ This is in itself dependent upon a multitude of factors including the number of antennas and their location, source location, flagging of data and more.

The second method would be to re-weight the data so that the resultant PSF more closely represents a Gaussian. This means that the systematic effect is reduced and prevents complications caused by deconvolving too deeply to be mitigated. We achieved this with the simulated data sets using a Gaussian taper of 1 arcsecond. The results are shown in Figure 3. As shown, the flux density offsets are still present (due to the imperfect Gaussian PSF), but to a lesser extent due to the much better fit. The offsets are only a few per cent compared to Figure 1 which could be more than 5–10%. In addition, the spread between the various deconvolution thresholds has been reduced with differences of around $\sim 1\%$ on the lowest S/N sources. However, as with any re-weighting scheme, there are always sacrifices that have to be made whether it be resolution or point-source sensitivity. In this case, the sensitivity of the observation worsened by 90% which would be detrimental to many surveys. In such cases, a compromise should be required and so would need to subsume this systematic effect into their error budget for individual sources.

This was the approach that was used for the *e*-MERGE survey (Muxlow et al. 2020). We found for core-dominated arrays, as that previously encountered in *e*-MERGE, applying a Tukey taper (Tukey 1962) allowed us to make the PSF more closely represent a Gaussian, and reduce the systematic flux offsets to just a few per cent. The taper allowed us to smoothly reduce the re-weight of the shorter baselines which in turn, reduces the amplitude of the shoulders in the PSF. While the sensitivity was degraded by around 30% the resultant flux densities were now reliable when compared to other surveys.

Unfortunately, the sheer number of different weighting schemes for interferometric data means that it should be viable to obtain these compromises, but the exact parameters are often difficult to determine analytically as the resultant PSF shape is dependent upon many interconnected factors. This means that a Monte-Carlo approach to the exact compromise would often be required for each observation.

5.2 Rescaling the residuals

An alternative method was proposed by Jörsäter & van Moorsel (1995) (see their Appendix A.2) and is the method commonly used in many spectral line studies. Here, the mismatch between the restoring beam and PSF areas, which causes the flux density systematic error, can be estimated and corrected by rescaling the measured flux densities to be in units of per beam rather than per PSF. Their formulation is as follows. If we take a residual flux density, R , and a restored/deconvolved flux density, C , the unknown true source flux density, G , is given by residual map, R , rescaled by a factor ϵ by,

$$G = C + \epsilon R. \quad (4)$$

The factor ϵ is unknown and is required for us to be able to rescale the residual flux from units of Jy PSF^{-1} to Jy beam^{-1} . This factor is related to the ratio of the areas through,

$$\epsilon = \frac{\text{Restoring beam area}}{\text{PSF area}}. \quad (5)$$

To estimate ϵ , Jörsäter & van Moorsel (1995) suggested that you can conduct two separate deconvolution runs to different thresholds. This will result in two recovered fluxes, C_1 and C_2 , and two residual fluxes, R_1 and R_2 . These can be related to the true flux density, G , by the following two equations,

$$G = C_1 + \epsilon R_1 \quad \text{and} \quad G = C_2 + \epsilon R_2. \quad (6)$$

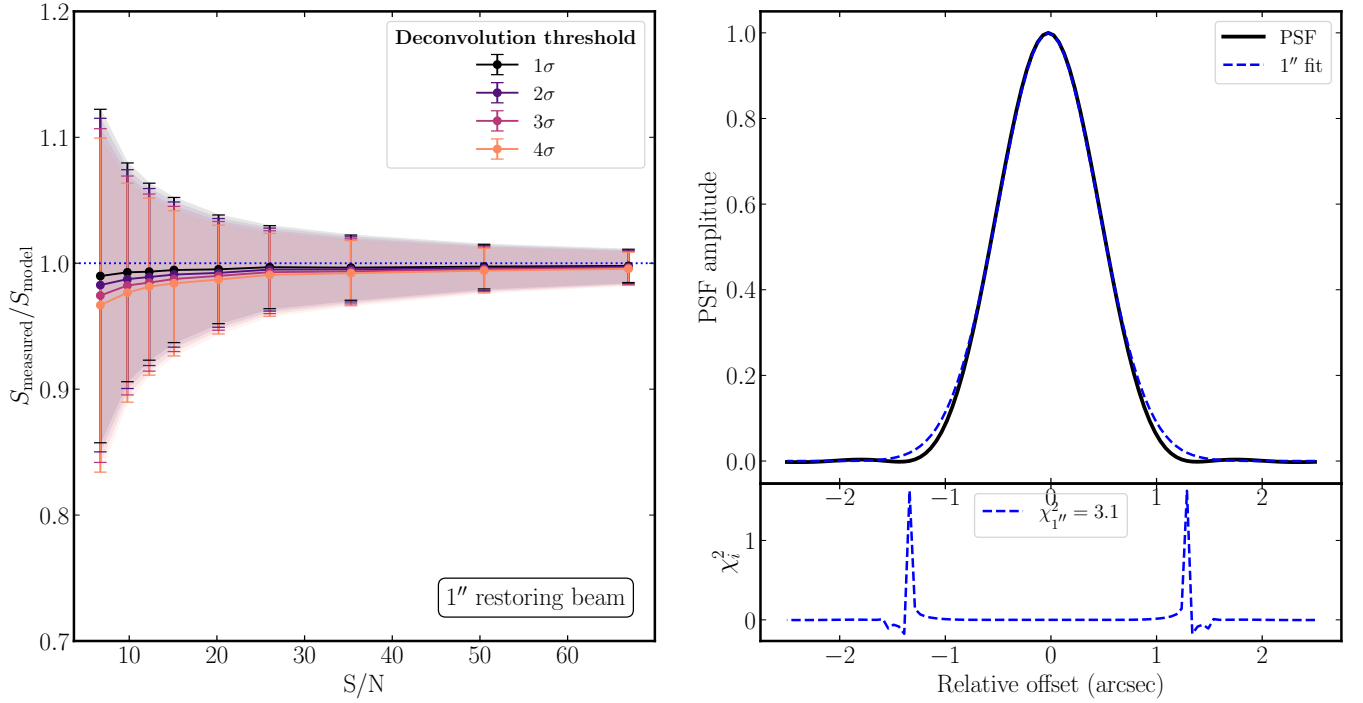


Figure 3. *Left panel:* The flux density recovery performance of the simulated combined VLA and *e*-MERLIN arrays using a reweighing scheme that forces the resultant PSF to closely represent a Gaussian with FWHM of $1''$. The S/N of the 4500 sources is plotted against the binned ratio of the measured to model flux densities for a range of deconvolution thresholds. The vastly improved performance compared to Figure 1 is evident and the flux density offsets are reduced to just a few per cent for the faintest sources. *Right panel:* A 1D slice through the PSF with the resultant fit overlaid. The bottom panel shows the residuals with the same definition as Figure 1. The slight overestimation of the PSF size causes a slight underestimate of the measured flux densities.

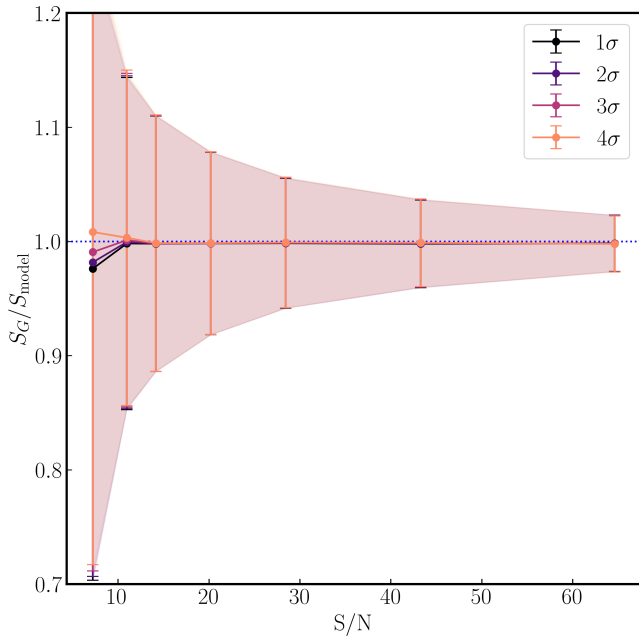


Figure 4. The estimated true flux density S_G achieved by rescaling the dirty map using Equation 8. The restoring beam was chosen to be of $0''.6$ as in the centre panel of Figure 2.

These can be solved for the two unknowns, G and ϵ to give,

$$\epsilon = \frac{C_2 - C_1}{R_1 - R_2}, \quad (7)$$

$$G = \frac{C_1 R_2 - R_1 C_2}{R_2 - R_1}.$$

This can be simplified by using a single deconvolution run and the un-deconvolved or dirty map so $C_2 = 0$. This reduces Eqn. 7 to,

$$\epsilon = \frac{C_1}{R_2 - R_1} \quad \text{and} \quad G = \epsilon R_2. \quad (8)$$

We applied Eqn. 8 to the imaging runs that used a $0''.6$ restoring beam (see the centre panel of Figure 1). We found that the true flux densities are almost perfectly recovered for all deconvolution thresholds at S/N larger than 15, far exceeding the performance of the re-weighting method presented in Section 5.2 (see Figure 4). However, at low S/N , the recovery deviates with a few per cent errors on all deconvolution depths. In addition, the calculation of ϵ for all sources increases the scatter in an individual measurement whereas the binned statistical population is correct.

An alternative method for correcting for this effect comes from the MeerKAT MeerKAT International Gigahertz Tiered Extragalactic Explorations (MIGHTEE) survey (Heywood et al. 2022). In this paper, they deal with the PSF mismatch by using a PSF homogenisation kernel which is convolved with the residual image (post-deconvolution model removal) to make the residuals match the restoring beam sizes. The model is then convolved with the restoring beam and added to the rescaled residuals. This ensures that the whole image is now in the same units. This method has the distinct advantage in removing the need to reweigh the data to get Gaussian

beams but the open question still arises as to whether this artificially affects the Fourier scales that are expressed in the image. We leave this investigation to a future paper that will look at resolved sources.

6 DISCUSSION AND CONCLUSIONS

To conclude, the point spread function of an interferometer is integral to every observation but can have far-reaching effects when we are attempting to extract accurate flux density measurements. Whilst the effect of the PSF and CLEAN-based deconvolution methods have been noted for many decades (e.g., Jörsäter & van Moorsel 1995), there have been only a few considerations in the literature (e.g., Benisty et al. 2021; Heywood et al. 2022).

The flux density systematic error shown in this paper affects *all* interferometric images as long as CLEAN-style deconvolution methods are still employed. While many previous interferometric arrays (e.g., the VLA) were originally designed to reduce the irregularities / non-Gaussianity of the PSF response, modern and upcoming interferometers (e.g., MeerKAT, LOFAR, SKA and the next-generation VLA), with their core dominated array configuration will naturally produce irregular PSFs. For example, in Figure 5 where we have simulated a 12-hour, 64-antenna MeerKAT observation at $\delta = -30^\circ$. The left-hand panel shows the resultant PSF which shows a widening due to the large number of short baselines present in the MeerKAT array. This is more clearly shown in the 1D slice of the PSF in the right-hand panel.

It is likely that CLEAN-based deconvolution algorithms will still be used for many years to come due to their speed and versatility, therefore this effect should be taken into account when analysing any interferometric images. This means that users will need to adopt some of the solutions that were presented in Section 5 and include these in their data analysis/reduction pipelines. However, the methods presented in this paper are not a comprehensive encyclopedia of all possible methods that could alleviate this effect but instead, we wish to open a discussion on the matter and ensure that this systematic does not need to be revisited in the future.

To mitigate this issue in its entirety, we will likely need to deviate from CLEAN-based deconvolution methods and adopt the more statistically robust algorithms that can produce more physically meaningful sky models (e.g., Carrillo et al. 2012; Arras et al. 2021); however, it is still to be proven how these algorithms could be implemented on the datasets of the size that the next generation of interferometers will produce.

ACKNOWLEDGEMENTS

We would like to extend our gratitude to the referee Prof. O. Smirnov and the scientific editor Prof. T. Pearson who helped improve this paper. We acknowledge the use of the ilifu cloud computing facility - www.ilifu.ac.za, a partnership between the University of Cape Town, the University of the Western Cape, the University of Stellenbosch, Sol Plaatje University, the Cape Peninsula University of Technology and the South African Radio Astronomy Observatory. The ilifu facility is supported by contributions from the Inter-University Institute for Data Intensive Astronomy (IDIA - a partnership between the University of Cape Town, the University of Pretoria and the University of the Western Cape), the Computational Biology division at UCT and the Data Intensive Research Initiative of South Africa (DIRISA). We also acknowledge the Jodrell Bank Centre for Astrophysics, which

is funded by the STFC. R.J.B. acknowledges funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004719 (Option-RadioNet PILOT). The RADIOBLOCKS project will receive funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101093934. A.N. has been supported by the Development in Africa with Radio Astronomy (DARA) project funded by STFC. *e*-MERLIN and formerly, MERLIN, is a National Facility operated by the University of Manchester at Jodrell Bank Observatory on behalf of STFC. This research makes use of PYTHON and the following packages: MATPLOTLIB (Hunter 2007; Caswell et al. 2019), ASTROPY (Astropy Collaboration et al. 2013, 2018), NUMPY (Oliphant 2006; Van Der Walt et al. 2011), and SCIPY (Virtanen et al. 2020).

DATA AVAILABILITY

The data underlying this article will be shared upon reasonable request to the corresponding author.

REFERENCES

- Arras P., Bester H. L., Perley R. A., Leike R., Smirnov O., Westermann R., Enßlin T. A., 2021, *A&A*, **646**, A84
- Astropy Collaboration et al., 2013, *A&A*, **558**, A33
- Astropy Collaboration et al., 2018, *AJ*, **156**, 123
- Becker R. H., White R. L., Helfand D. J., 1995, *ApJ*, **450**, 559
- Benisty M., et al., 2021, *ApJ*, **916**, L2
- Booth M., et al., 2023, *MNRAS*, **521**, 6180
- Bradley L., et al., 2020, *astropy/photutils*: 1.0.0, [doi:10.5281/zenodo.4044744](https://doi.org/10.5281/zenodo.4044744), <https://doi.org/10.5281/zenodo.4044744>
- Briggs D. S., Schwab F. R., Sramek R. A., 1999, in Taylor G. B., Carilli C. L., Perley R. A., eds, *Astronomical Society of the Pacific Conference Series* Vol. 180, *Synthesis Imaging in Radio Astronomy II*. p. 127
- Cannon J. M., Salzer J. J., Rosenberg J. L., 2009, *ApJ*, **696**, 2104
- Carrillo R. E., McEwen J. D., Wiaux Y., 2012, *MNRAS*, **426**, 1223
- Caswell T. A., et al., 2019, *matplotlib/matplotlib*: REL: v3.1.1, [doi:10.5281/zenodo.3264781](https://doi.org/10.5281/zenodo.3264781), <https://doi.org/10.5281/zenodo.3264781>
- Clark B. G., 1980, *A&A*, **89**, 377
- Hale C. L., Robotham A. S. G., Davies L. J. M., Jarvis M. J., Driver S. P., Heywood I., 2019, *MNRAS*, **487**, 3971
- Heywood I., et al., 2022, *MNRAS*, **509**, 2150
- Högbom J. A., 1974, *A&AS*, **15**, 417
- Hunter J. D., 2007, *Computing in Science & Engineering*, **9**, 90
- Hunter D. A., et al., 2012, *AJ*, **144**, 134
- Jauncey D. L., 1968, *ApJ*, **152**, 647
- Jörsäter S., van Moorsel G. A., 1995, *AJ*, **110**, 2037
- Kennicutt Robert C. J., et al., 2007, *ApJ*, **671**, 333
- Krist J. E., Hook R. N., Stoehr F., 2011, in Kahan M. A., ed., *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* Vol. 8127, *Optical Modeling and Performance Predictions V*. p. 81270J, [doi:10.1117/12.892762](https://doi.org/10.1117/12.892762)
- Muxlow T. W. B., et al., 2020, *MNRAS*, **495**, 1188
- Novak M., et al., 2020, *ApJ*, **904**, 131
- Offringa A. R., et al., 2014, *MNRAS*, **444**, 606
- Oliphant T. E., 2006, *Guide to NumPy*, 2nd edn. Trelgol Publishing, USA
- The CASA Team et al., 2022, arXiv e-prints, [p. arXiv:2210.02276](https://arxiv.org/abs/2210.02276)
- Tukey J. W., 1962, *Ann. Math. Statist.*, **33**, 1
- Van Der Walt S., Colbert S. C., Varoquaux G., 2011, *Computing in Science & Engineering*, **13**, 22
- Virtanen P., et al., 2020, *Nature Methods*, **17**, 261
- de Blok W. J. G., et al., 2018, *ApJ*, **865**, 26

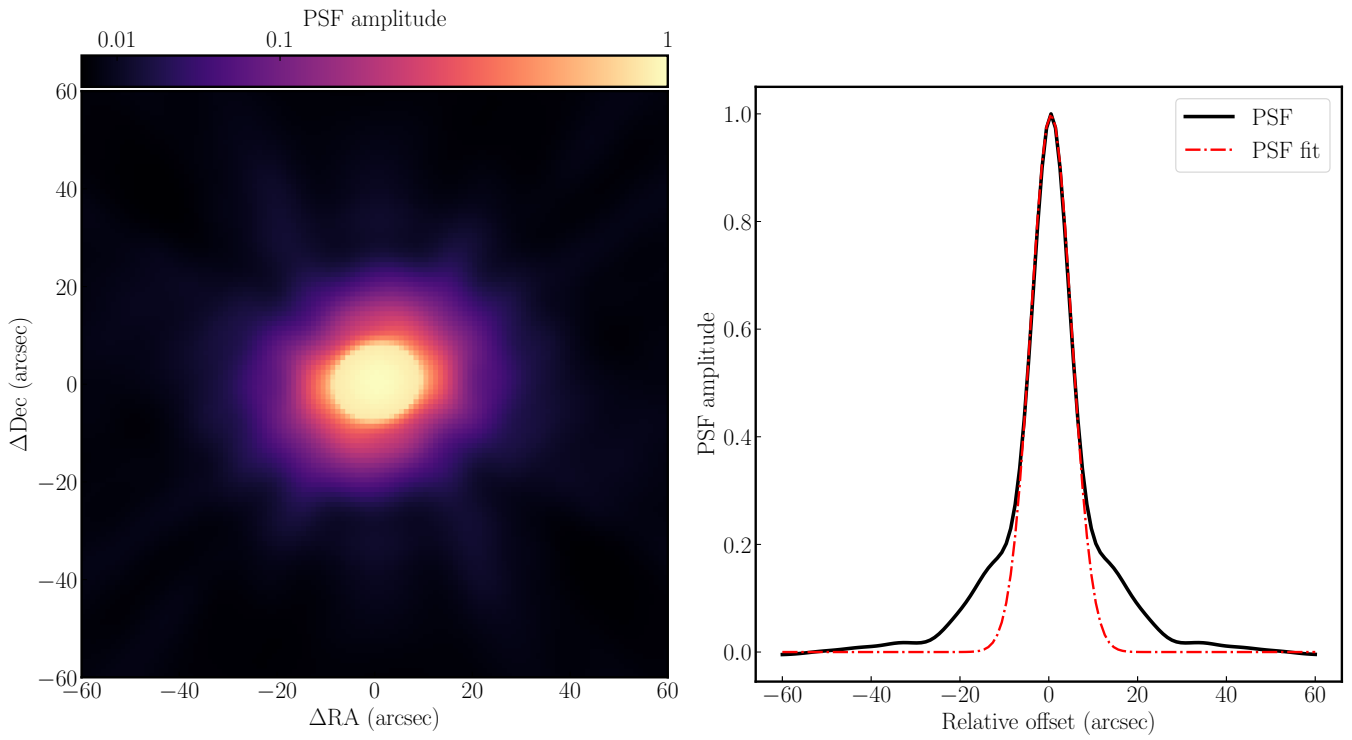


Figure 5. The PSF of a simulated 12 hr MeerKAT observation of a source at $\delta = -30^\circ$ using robust weighting of 0.5. The left panel is a $1' \times 1'$ cutout of the PSF while the right panel is a 1D slice at $\Delta Dec = 0$ with the Gaussian fit to the PSF over-plotted. The shoulder of the PSF due to the large number of short baselines present in both panels which results in a Gaussian unable to fit the PSF.

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.