VLASS memo 16: Clean/Dirty Beam Bias and Correction for Taylor term VLASS images

Jürgen Ott

28 October 2021

Abstract: Cleaned images in radio aperture synthesis are typically composed of the clean components convolved with a clean beam plus the residual image. Although the units of the former is Jy/(clean beam) and the latter Jy/(dirty beam), the composite image is treated as having Jy/(clean beam) units. This misnomer results in a clean/dirty beam bias where image flux densities deviate from the true flux densities as a function of the ratio of the clean to dirty beam area. The clean/dirty beam bias therefore depends on the uv-coverage, gridding, and weighting of the visibilities, all of which defines the differences between the clean and dirty beam. In general, the clean/dirty beam bias may therefore vary in magnitude and sign, underestimating or overestimating image values. The strength of the effect depends also on the S/N of the images as well as on the extent of the sources.
In this memo, we describe the effect and introduce a method for calculating a correction for Taylor term expansion images in frequency space. For two 2-term VLA Sky Survey (VLASS) images in the GOODSN field, cleaned to a threshold of 3σ, we quantify the bias for image flux densities and spectral indices, and demonstrate our correction method. The application reduces the flux density error from up to $\sim 8\%$ to $\sim 2\%$ or better within the $3 < S/N < 5$ interval, when compared to deep cleaned (unbiased) images. Spectral indices improve from an initial maximum deviation from a deep cleaned image of ~ 0.3 to ~ 0.1 after the correction. Higher S/N components result in smaller correction factors, and at $S/N > 10$, the correction is typically $< 1\%$. For a constant dirty-to-clean beam ratio, one would expect constant ϵ correction factors across the image and S/N regimes. We use the median of all individual, source-based ϵ values to estimate common values. We find that the correction with median ϵ factors is of similar magnitude than a source-by-source based correction, but can lead to instabilities. Given that VLASS images are large mosaics with a spatial variation of the dirty/clean beam ratio, we recommend a source-based correction. We do not see a reduction in the clean/dirty beam bias based on different selections of multiscale scales.

1 Introduction

The CLEAN algorithm is an iterative deconvolution process that finds peaks in the dirty image (D), from which clean components in the form of delta functions or large paraboloids (multiscale clean; Cornwell 2008, Rau & Cornwell 2011), are subtracted. Together, the clean components are called a clean model. After a subtraction of the clean components from the data, the remaining image content is called the residual image (R). Depending on the choice and number of clean components, the residual image still contains relevant flux. The 'dirty' beam is replaced by a 'clean' beam that is determined by a Gaussian fit to the central portion of the dirty beam. The final image (I) is then constructed by adding the clean model, convolved by the clean beam (C), to the residual (R). I is expressed in units of Jy/(clean beam). A depiction of the image composition is shown in Fig. 1.

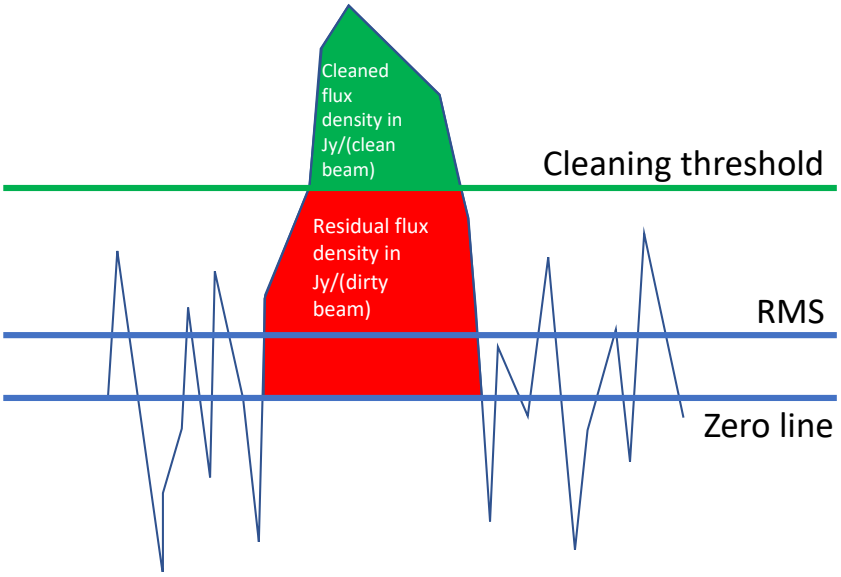


Figure 1: Clean/dirty beam bias: Down to the clean threshold, an image is expressed as cleaned flux density to which the remaining residual image with residual flux density is added. Typically, the cleaning threshold is chosen to be a few times the rms of the image to avoid cleaning into noise. Whereas the entire image is expressed in Jy/(clean beam) this is only correct for the cleaned portion of the image since the residual component has in fact units of Jy/(dirty beam). The resulting error of the mislabelling is the clean/dirty beam bias. The effect will be reduced by cleaning very deeply inside masks and therefore minimizing or eliminating any residual flux density. When this is not practical, a correction for the flux density can be computed.

Problem: The residual image R has units of Jy/(dirty beam), whereas the clean model convolved with the clean beam C has units of Jy/(clean beam). The image I=C+R is expressed in Jy/(clean beam) units, which is technically incorrect given the different units of the residual image and the convolved clean model. This error is important when significant flux density is left in the residual image that has not been deconvolved.

Solution: The best solution to this problem is to remove all flux density from the residual image. This can be done by cleaning very deeply, or by using better clean component functions like multi-scale clean paraboloids of different widths. If this is not possible or practical (e.g. when deep images diverge or are computationally expensive), one can estimate a correction factor ϵ by the ratio of the clean to dirty beam, correct R and consequently I to I^{corr}:

$$I^{corr} = C + \epsilon R \tag{1}$$

Since the dirty beam is not an analytical function, the ϵ factor needs to be determined over some beam area and can only be approximated. It also assumes that the dirty beam is constant across the image, which is not always the case, in particular it is not valid for mosaic images.

A different method to determine ϵ and therefore the corrected flux density is described in Jörsäter & van Moorsel (1995; JvM). This method can be described as follows.

Clean one image to a certain depth, and introduce the correction factor ϵ as a variable:

$$I^{corr} = C_1 + \epsilon R_1 \tag{2}$$

Clean the same image to a different depth, leading to a different clean model and a different residual. Per definition, the corrected image is the same as in the first cleaning procedure:

$$I^{corr} = C_2 + \epsilon R_2 \tag{3}$$

If the image is not cleaned at all, then $C = 0$ and R is the same as the dirty image D:

$$I^{corr} = \epsilon D \tag{4}$$

Equations 2 and 4 can now be combined to solve for both, ϵ and I^{corr}; we can now drop the index 1 for the first cleaning stage. First we equate I^{corr}:

$$C + \epsilon R = \epsilon D \rightarrow \epsilon = \frac{C}{D - R} \tag{5}$$

Inserting ϵ into Eq. 4 provides the corrected flux density as

$$I^{corr} = \epsilon D = \frac{DC}{D - R} \tag{6}$$

This equation works for all values above the cleaning threshold. For sources below, no cleaning occurs, i.e. $R = D$ and $C = 0$, leaving the equation undefined.

It has been shown (e.g., Ott et al. 2012; Walter et al. 2008; Del Rio et al. 2004; Ott et al. 2001) that this method leads to very reliable results and delivers the same I^{corr}, almost independent of cleaning depth or parameters that shape the dirty beam, like visibility weighting schemes.

Although, in principle, the JvM method should lead to a single ϵ per image that is representative for the clean to dirty beam ratio, it is usually better to calculate ϵ per aperture/source and image plane separately. That way, changes in beam shape across the image cube, e.g. due to channel flagging, or mosaicking, will be calculated and corrected on a source by source basis. This will also omit any scaling of the noise. For sources with S/N below the cleaning threshold, ϵ cannot be determined per source (see above). For those cases a median ϵ derived from sources above the cleaning threshold will at least provide an approximate correction.

2 Clean/dirty beam bias for images with Taylor term expansion

The JvM method has been used for many years on regular cleaned continuum images or plane by plane in image cubes, assuming they are monochromatic. The JvM corrected images and cubes can be used to compute corrected spectral indices by comparing images or image planes at different frequencies. For wideband continuum images, a new method has been proposed by Rau & Cornwell (2011) that performs a Taylor term expansion in frequency during imaging, thus deriving flux density, spectral index, spectral curvature, and higher order spectral images in the process. These images can be computed from the individual Taylor term tt images that are defined as:

$$I = \sum_{tt} I_{tt}^{sky} \left(\frac{\nu - \nu_0}{\nu_0} \right)^{tt} \tag{7}$$

where I_{tt}^{sky} are Taylor term images, ν the frequency, and ν_0 the reference frequency. From these, for example, the flux density at the reference frequency is derived from I_{tt0}^{sky}, and the spectral index from

$$I_{tt1}^{sky} / I_{tt0}^{sky}. \tag{8}$$

The Taylor term images themselves are composed of the respective Taylor term residuals and clean components convolved with the clean beam via the inverse of a Hessian matrix $\mathbf{H}$:

$$I(tt0, tt1, ...) = \mathbf{H^{-1}} C(tt0, tt1, ..) + \mathbf{H^{-1}} R(tt0, tt1, ...) \tag{9}$$

In CASA's TCLEAN task[1], the Hessian matrix $\mathbf{H}$ is provided in the logger output. For TCLEAN note that a) for multi-scale images the delta function scale is relevant here, b) that the Hessian matrix needs to be normalized before inversion, and c) that $\mathbf{H^{-1}}$ is already applied in the model output of C. Therefore, the equation reduces to

$$I(tt0, tt1, ...) = C(tt0, tt1, ...) + \mathbf{H^{-1}} R(tt0, tt1, ...) \tag{10}$$

using a Hessian matrix notation like

$$\mathbf{H^{-1}} = \begin{pmatrix} h_{00}^{-1} & h_{01}^{-1} & h_{02}^{-1} & ... \\ h_{10}^{-1} & h_{11}^{-1} & h_{12}^{-1} & ... \\ h_{20}^{-1} & h_{21}^{-1} & h_{22}^{-1} & ... \\ ... & ... & ... & ... \end{pmatrix} \tag{11}$$

this turns into

$$\begin{aligned} I(tt0) &= C(tt0) + h_{00}^{-1} R(tt0) + h_{01}^{-1} R(tt1) + h_{02}^{-1} R(tt2) + ... \\ I(tt1) &= C(tt1) + h_{10}^{-1} R(tt0) + h_{11}^{-1} R(tt1) + h_{12}^{-1} R(tt2) + ... \\ I(tt2) &= C(tt1) + h_{20}^{-1} R(tt0) + h_{21}^{-1} R(tt1) + h_{22}^{-1} R(tt2) + ... \\ ... \end{aligned} \tag{12}$$

We are now introducing the correction factor ϵ at this stage. The form of ϵ is not entirely defined. The best match would be a $n \times n$ ϵ matrix. To determine the matrix elements, however, n^2 elements need to be computed, which requires additional images, cleaned to different depths. Our approach is to use scalars instead, one for each Taylor term and introduce them in the following way:

$$\begin{aligned} I^{corr}(tt0) &= C(tt0) + \epsilon_0 [h_{00}^{-1} R(tt0) + h_{01}^{-1} R(tt1) + h_{02}^{-1} R(tt2) + ...] \\ I^{corr}(tt1) &= C(tt1) + \epsilon_1 [h_{10}^{-1} R(tt0) + h_{11}^{-1} R(tt1) + h_{12}^{-1} R(tt2) + ...] \\ I^{corr}(tt2) &= C(tt2) + \epsilon_2 [h_{20}^{-1} R(tt0) + h_{21}^{-1} R(tt1) + h_{22}^{-1} R(tt2) + ...] \\ ... \end{aligned} \tag{13}$$

The second set of equations is given by the dirty image which implies $C = 0$ and $R = D$:

$$\begin{aligned} I^{corr}(tt0) &= \epsilon_0 [h_{00}^{-1} D(tt0) + h_{01}^{-1} D(tt1) + h_{02}^{-1} D(tt2) + ...] \\ I^{corr}(tt1) &= \epsilon_1 [h_{10}^{-1} D(tt0) + h_{11}^{-1} D(tt1) + h_{12}^{-1} D(tt2) + ...] \\ I^{corr}(tt2) &= \epsilon_2 [h_{20}^{-1} D(tt0) + h_{21}^{-1} D(tt1) + h_{22}^{-1} D(tt2) + ...] \\ ... \end{aligned} \tag{14}$$

We can now solve this set of equations for ϵ_i.

[1] using CASA 6.1 for this memo

2.1 Solving for two Taylor terms

Since VLASS is imaged with 2 Taylor terms (nterms=2 in TCLEAN), we will now explicitly derive solutions for this case. Solving for ϵ_0 and ϵ_1 we equate equations 13 and 14 to obtain for all:

$$\begin{aligned} \epsilon_0[h_{00}^{-1}D(tt0) + h_{01}^{-1}D(tt1)] &= C(tt0) + \epsilon_0[h_{00}^{-1}R(tt0) + h_{01}^{-1}R(tt1)] \\ \epsilon_1[h_{10}^{-1}D(tt0) + h_{11}^{-1}D(tt1)] &= C(tt1) + \epsilon_1[h_{10}^{-1}R(tt0) + h_{11}^{-1}R(tt1)] \end{aligned} \quad (15)$$

and find

$$\begin{aligned} \epsilon_0 &= \frac{C(tt0)}{[h_{00}^{-1}D(tt0) + h_{01}^{-1}D(tt1)] - [h_{00}^{-1}R(tt0) + h_{01}^{-1}R(tt1)]} \\ \epsilon_1 &= \frac{C(tt1)}{[h_{10}^{-1}D(tt0) + h_{11}^{-1}D(tt1)] - [h_{10}^{-1}R(tt0) + h_{11}^{-1}R(tt1)]} \end{aligned} \quad (16)$$

Inserting equations 16 into 14 then provides the corrected $I^{corr}(tt0)$ and $I^{corr}(tt1)$:

$$\begin{aligned} I^{corr}(tt0) &= \epsilon_0[h_{00}^{-1}D(tt0) + h_{01}^{-1}D(tt1)] = & (17) \\ &= \frac{C(tt0)[h_{00}^{-1}D(tt0) + h_{01}^{-1}D(tt1)]}{[h_{00}^{-1}D(tt0) + h_{01}^{-1}D(tt1)] - [h_{00}^{-1}R(tt0) + h_{01}^{-1}R(tt1)]} & \\ I^{corr}(tt1) &= \epsilon_1[h_{10}^{-1}D(tt0) + h_{11}^{-1}D(tt1)] = & (18) \\ &= \frac{C(tt1)[h_{10}^{-1}D(tt0) + h_{11}^{-1}D(tt1)]}{[h_{10}^{-1}D(tt0) + h_{11}^{-1}D(tt1)] - [h_{10}^{-1}R(tt0) + h_{11}^{-1}R(tt1)]} & (19) \end{aligned}$$

As in for non-Taylor term images, the solution is undefined below the cleaning threshold, where $C = 0$ and $D = R$.

Note: TCLEAN does not output C directly, so it has to be reconstructed from I and R first, via:

$$C(tt0) = I(tt0) - h_{00}^{-1}R(tt0) - h_{01}^{-1}R(tt1) \quad (20)$$

$$C(tt1) = I(tt1) - h_{10}^{-1}R(tt0) - h_{11}^{-1}R(tt1) \quad (21)$$

3 Application to VLASS Data

For the following, we use two VLASS data sets, both toward the GOODS-N area, tile T26t10, measurement set TSKY0001.sb36463619.eb36473386.58555.315533263885:

- J125918+643000, phase center J2000 12:59:18.674 +64.23.00.913
- J124103+643000, phase center J2000 12:41:3.666 +64.30.0.0000

The data were calibrated with the quick look pipeline and imaged with a single epoch imaging script that was the prototype for the production version in summer 2021. The S-band data are large mosaics, observed interferometrically on the fly, and imaged with the aw-projection gridder using 32 w-projection planes. The Briggs weighting robust parameter was set to 1 and masks were produced during the processing to only clean areas with emission. The data were deconvolved with multi-scale cleaning, choosing scales of 0, 5, 12 pixels (pixel cell size 0.6"). The fluxes of the sources were extracted from within the masks that are produced in the imaging script. Every source within a mask is corrected for the clean/dirty beam bias separately.

3.1 Hessian Matrix

The Hessian matrix is provided in the TCLEAN log. For J125918 the output was:

```
task_tclean::MultiTermMatrixCleaner::computeHessianPeak
The Matrix [H] for 0 pixel scale is :
Axis Lengths: [2, 2] (NB: Matrix in Row/Column order)
task_tclean::MultiTermMatrixCleaner::computeHessianPeak + [1, 0.0444914
task_tclean::MultiTermMatrixCleaner::computeHessianPeak + 0.0444914, 0.0322221]
task_tclean::MultiTermMatrixCleaner::computeHessianPeak
The Matrix [H] for 5 pixel scale is :
Axis Lengths: [2, 2] (NB: Matrix in Row/Column order)
task_tclean::MultiTermMatrixCleaner::computeHessianPeak + [0.426282, 0.00507535
task_tclean::MultiTermMatrixCleaner::computeHessianPeak + 0.00507535, 0.0134031]
task_tclean::MultiTermMatrixCleaner::computeHessianPeak
The Matrix [H] for 12 pixel scale is :
Axis Lengths: [2, 2] (NB: Matrix in Row/Column order)
task_tclean::MultiTermMatrixCleaner::computeHessianPeak + [0.136754, -0.00026639
task_tclean::MultiTermMatrixCleaner::computeHessianPeak + -0.00026639, 0.00424926]
```

As mentioned in Sect. 2, the relevant numbers are in the '0 pixel scale' of the multi-scale kernels, i.e. $\mathbf{H} = \begin{pmatrix} 1 & 0.0444914 \\ 0.0444914 & 0.0322221 \end{pmatrix}$. This matrix now needs to be normalized for the h_{00} component (in this case this has been done internally) and inverted to $\mathbf{H^{-1}}$ for the clean/dirty beam bias correction.

3.2 Assessing the Clean/Dirty Beam Bias

The data were cleaned to different depths at and beyond the single epoch imaging script of 3σ rms within the clean masks, a process that took about 60 h using 4 nodes, 8 cores each. We cleaned these images to lower 2.0σ, 1.5σ, 1.0σ, 0.8σ, and 0.6σ thresholds. In an additional step, the J125918 image was further cleaned to a very deep 0.3σ level. To go from the regular 3σ-cleaned image to the 0.6σ image, additional processing of roughly 10 days were needed. The processing to the 0.3σ threshold took an additional 26 days over the regular 3σ cleaning. At that level some image artifacts can occur, e.g. due to oscillations

of the algorithm. Therefore, we consider the 0.6σ-level cleaned images the 'deep images' that have a minimum of clean/dirty beam bias and we consider them as the true sky values (also see Sec. 3.3).

In general, the clean/dirty beam bias may over or underestimate the flux density values, depending on the shape differences between the clean and dirty beams. Those depend on the uv-coverage, gridding, and weighting of the data. In addition, lower S/N sources and sources with large extent exhibit a lower cleaned to residual flux density ratio, thus increasing the clean/dirty beam bias. Deeper cleaned images reduce the clean/dirty beam bias.

For our VLASS images, we find that flux densities are overestimated (Fig. 2 (a) and (b)) and that the clean/dirty beam bias depends on the S/N of the sources as expected. The effect for VLASS data, which contains only slightly extended sources, is significant below S/N ratios of ~ 10. The median flux density ratios, and therefore the overestimates of the 3σ cleaned images to the unbiased 0.6σ images are given in Table 1 (left) for different S/N intervals.

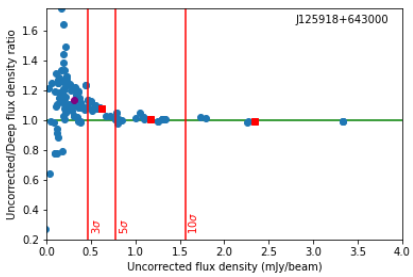
In Fig. 3 (a) and (b) we show the effect on the spectral index α of the images (computed as in eq. 8). The clean/dirty beam bias tends to produce smaller values for α and the median differences are provided in Table 1 (left) as a function of S/N interval.

	Uncorrected		Corrected		Constant ϵ corrected	
	J125918	J124103	J125918	J124103	J125918	J124103
	Flux density ratio					
$S/N < 3$	$\sim 13\%$	$\sim 10\%$	...	...	4%	12%
$3 < S/N < 5$	$\sim 8\%$	$\sim 3\%$	$\lesssim 2\%$	$\lesssim 2\%$	12%	3%
$5 < S/N < 10$	$\sim 1\%$	$< 1\%$	$\lesssim 2\%$	$\lesssim 2\%$	$\lesssim 1\%$	9%
$S/N > 10$	$\lesssim 1\%$	$\lesssim 1\%$	$\lesssim 2\%$	$\lesssim 2\%$	$\lesssim 1\%$	$\lesssim 1\%$
	Spectral Index difference					
$S/N < 3$	-0.61	-0.37	...	...	-0.20	0.08
$3 < S/N < 5$	-0.30	-0.25	-0.1	-0.45	-0.1	-0.31
$5 < S/N < 10$	-0.16	-0.05	-0.01	-0.08	-0.03	-0.05
$S/N > 10$	-0.08	-0.03	-0.01	-0.03	-0.07	0.02

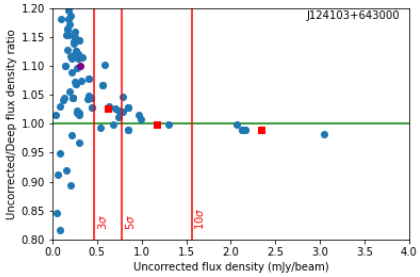
Table 1: Flux density ratios and spectral index differences comparing the 3.0σ to the deep images. The columns are the uncorrected clean/dirty beam bias, the source-by-source corrected values, and the correction, taking a single, median ϵ_0 and ϵ_1 per field. The values are median values for all the sources in each listed S/N bin. Data is presented for both fields, J125918 and J124103.

3.3 Correction

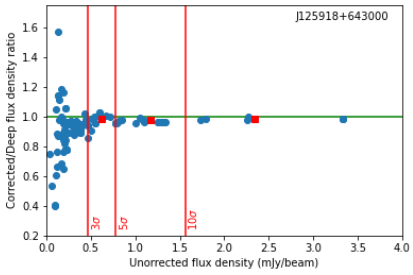
We calculate the ϵ correction factors for the two images and in Fig. 2 (c) and (d) show the ratio of the corrected flux densities and the deep cleaned (0.6σ) images. For both, J125918 and J124103 the clean/dirty beam bias is is reduced



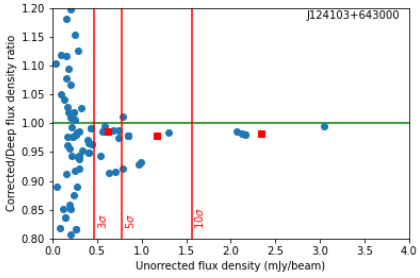
(a) Uncorrected Field J125918+643000



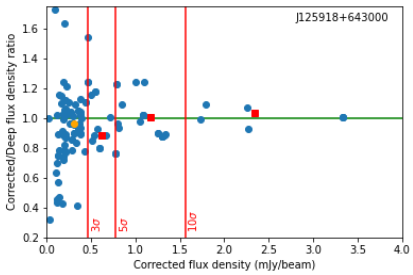
(b) Uncorrected field J124103



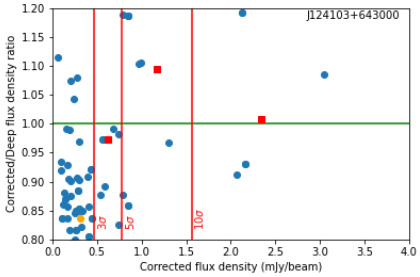
(c) Corrected field J125918



(d) Corrected field J124103

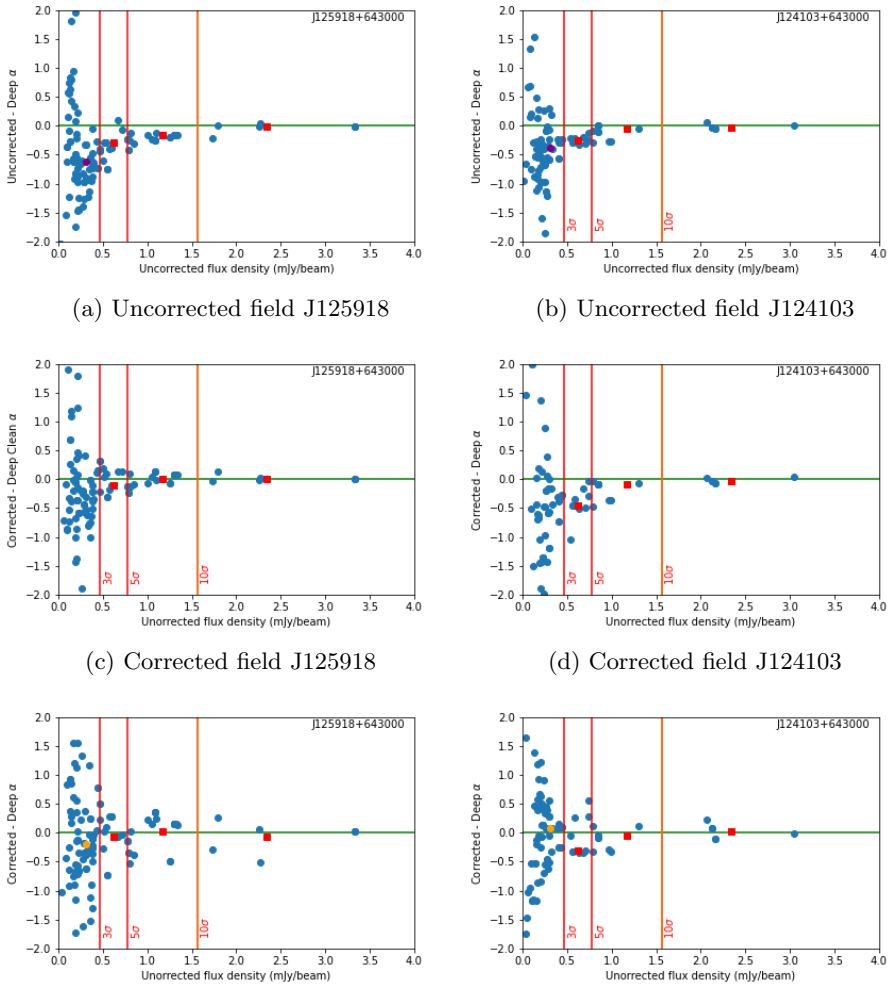


(e) Single-ϵ corrected field J125918



(f) Single-ϵ corrected field J124103

Figure 2: Flux density ratio of sources between a 3.0σ and a deep, unbiased (0.6σ) cleaning depth as a function of (uncorrected) source flux density for the two fields. Panels *(a)* and *(b)* show the distribution before and *(c)* and *(d)* after correction for the clean/dirty beam bias. Panels *(e)* and *(f)* show corrected flux density ratios using single, median ϵ values. The vertical lines depict the 3, 5, and 10σ rms noise levels of the images. The red squares are the median values in the $3 < S/N < 5$, $5 < S/N < 10$ and $S/N > 10$ intervals. The purple dots are the median values at $S/N < 3$, and the orange dots are median corrected values when a median ϵ is applied for the $S/N < 3$ interval.



(a) Uncorrected field J125918

(b) Uncorrected field J124103

(c) Corrected field J125918

(d) Corrected field J124103

(e) Single-ϵ corrected field J125918

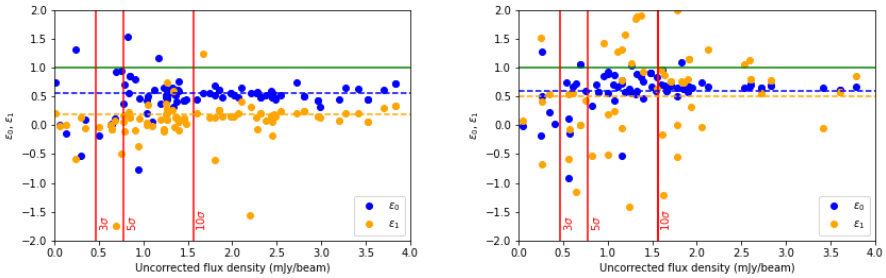
(f) Single-ϵ corrected field J124103

Figure 3: Spectral index differences of sources between a 2.0σ and deep 0.6σ cleaning depth as a function of (uncorrected) source flux density. Panels *(a)* and *(b)* show the distribution before and *(c)* and *(d)* after correction for the clean/dirty beam bias on a per source basis. Panels *(e)* and *(f)* show corrected flux ratios using single, median ϵ values. The vertical lines depict the 3, 5, and 10σ rms noise levels of the images. The red squares are the median vlues in the $3 < S/N < 5$, $5 < S/N < 10$ and $S/N > 10$ intervals. The purple dots are the median values at $S/N < 3$, and the orange dots are median corrected values when a median ϵ is applied for the $S/N < 3$ interval.

to $\sim 2\%$ for all $S/N > 3$ intervals, where the correction is defined (Table 1 [center]).

For the spectral index, the correction is demonstrated in Fig. 3 (c) and (d) and tabulated in Table 1 (center). We see clear improvements for J125918. For J124103, however, the difference after correction is -0.45 in the $3 < S/N < 5$ bin and -0.08 at $5 < S/N < 10$, which is actually slightly worse than the uncorrected values.

In Fig. 4 we show the distribution of the ϵ factors. For J125918, ϵ_0 hovers around a median value of 0.56 and ϵ_1 around a median of 0.18 using all data points with $S/N > 3$ for the calculation. For J124103, ϵ_0 has a median of 0.60 and ϵ_1 of 0.51.



(a) Field J125918+643000 (b) Field J124103+643000

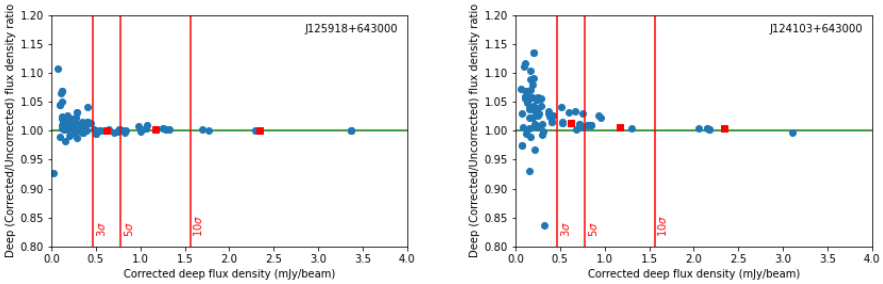
Figure 4: The distribution of ϵ correction factors as a function of source flux density. The dashed lines are the medians through the respective ϵ_0 and ϵ_1 distributions and the solid line corresponds to $\epsilon = 1$, i.e. no correction.

To check if the deep cleaned images are representative for having no clean/dirty beam bias, we also corrected the 0.6σ images themselves using the equations in Sect. 2. The ratio of the 0.6σ corrected and uncorrected flux densities are shown in Fig. 5. The correction in these images amounts to a maximum median value of $\sim 1\%$ in the $3 < S/N < 5$ interval of J124103 and less than 1% in all other intervals above the cleaning threshold. This indicates that the 0.6σ images are indeed cleaned deep enough to be good representations of the true, unbiased flux densities.

3.4 Constant ϵ

Different observations or weighting of the data will lead to different values of ϵ. Within an image, however, the clean to dirty beam ratio should be constant, and one would expect constant values of ϵ_0 and ϵ_1 (cf. Eqns. 1 and 13). For large-field VLASS mosaic images, however, we have a situation where the dirty beam is not constant across the entire image, and therefore ϵ may show a directional dependence.

When we use the median values of the correction factors ϵ_0 and ϵ_1, as shown



(a) Field J125918 (b) Field J124103

Figure 5: Ratio of fluxes of the deep (0.6σ) corrected and deep uncorrected flux density ratios as a function of deep corrected flux density. The vertical lines show the 3, 5, and 10σ rms noise of the images. The red dots are the median values in the $3 < S/N < 5$, $5 < S/N < 10$ and $S/N > 10$ intervals.

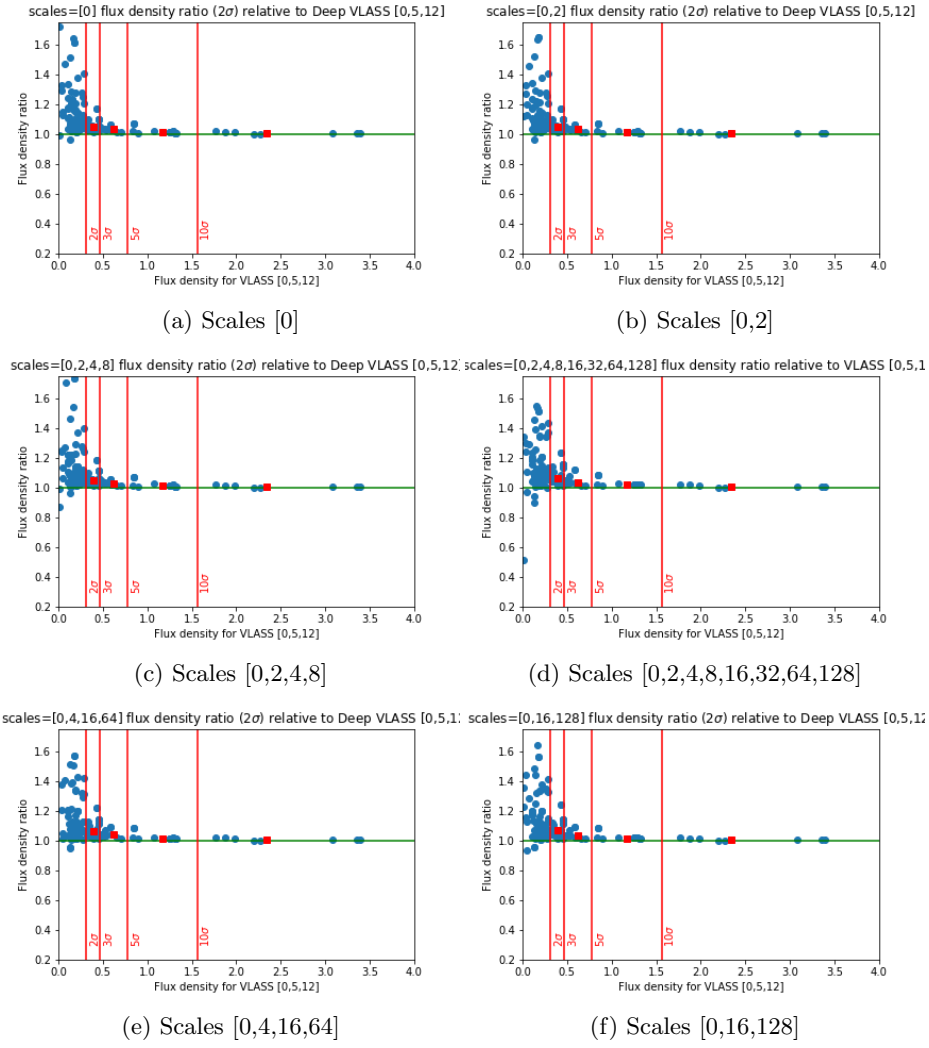
in Fig. 4, however, we clearly see some deviations (Fig. 2(e) and (f))). They are likely due to the spatial variation of the dirty beam in the mosaic. The flux density and spectral index corrections for the two fields are listed in Table 1 (right). The scatter in the plots is rather large. This may be a reflection of Fig. 4(b), which also shows a large scatter of ϵ values.

As discussed in Sect. 2, the correction breaks down below the cleaning threshold. What we can do, however, is to apply the ϵ_0 and ϵ_1 values that are obtained from sources above the cleaning threshold to those below. This is indicated by the orange points in Figs. 2 (e) and (f) and 3 (e) and (f) and values are given in Table 1 (right). Flux densities and spectral indices after correction are similar or somewhat worse than in a source-by-source correction, and the corrections are more volatile.

4 The effect of using different multi-scale clean scales on the clean/dirty beam bias

We also tested the effect of the clean/dirty beam bias for different different selections of scales in multi-scale clean for source J125918. Flux density differences to the nominal, deep VLASS image that was cleaned with [0, 5, 12] pixel sized paraboloids, are shown in Fig. 6. The pixel scales were selected to have different numbers of scales and different granularity. For these tests, we compare images cleaned to 2σ with the 0.6σ reference.

There are differences, with medians up to $\sim 7\%$ in the $2 < S/N < 3$, $\sim 9\%$ in the $1 < S/N < 3$, $\sim 3\%$ in the $3 < S/N < 5$ intervals, and less than 2% at higher S/N. Overall, however, we do not see any preferred choice of scale that would reduce the clean/dirty beam bias at the 3σ cleaning level compared to the regular VLASS image as represented in Fig. 2(a) and (b).



(a) Scales [0]

(b) Scales [0,2]

(c) Scales [0,2,4,8]

(d) Scales [0,2,4,8,16,32,64,128]

(e) Scales [0,4,16,64]

(f) Scales [0,16,128]

Figure 6: Flux densities of images run with different multiscale parameters. The flux density ratios are relative to the 2σ cleaned VLASS reference image with [0,5,12] scales. The vertical lines show the 3, 5, and 10σ rms noise of the images. The red dots are the medians in the $S/N < 1$, $1 < S/N < 3$, $3 < S/N < 5$, $5 < S/N < 10$ and $S/N > 10$ intervals.

5 Summary

The clean/dirty beam bias is introduced by residual flux density with units of Jy/(dirty beam) that is, however, assumed as Jy/(clean beam) in the composite image when added to the clean beam-convolved clean components. When the residual flux density is not negligible compared to the cleaned flux density, this difference can be very significant with tens of percent of flux density difference to the true flux density. The effect is thus stronger for sources that are more extended and have lower S/N. VLASS at high Galactic latitude detects relatively compact sources, so the clean/dirty beam bias is not as strong as it would be for more extended emission. We tested two VLASS GOODSN images, cleaned to 3σ, and find differences of median values up to $\sim 13\%$ in flux density as compared to deeply cleaned, unbiased images (cleaned to 0.6σ, images that have virtually all flux cleaned out). The spectral index difference is up to 0.4. These differences are found at flux densities in the $1 < S/N < 3$ interval and are less prominent at higher S/N values with almost no difference above $S/N > 5$. The bias appears not to be a strong function of the selection of scales in multiscale deconvolution mode. A method for clean/dirty beam bias correction has been introduced by JvM. In this memo, we expand upon this method for Taylor term images. To avoid multiple cleaning depths, we use scalar-based correction factors ϵ_i rather than tensors. This simplified calculation is capable of reducing the clean/dirty beam bias from $\sim 8\%$ to $\sim 3\%$ in the $3 < S/N < 5$ interval and effectively removes the clean/dirty beam bias to a scatter of $< 2\%$ at higher S/N values. The spectral index improves from a difference of 0.3 to 0.1 at the maximum for J125918, but it degrades the the accuracy for J124103 by up to 0.2. For sources below the cleaning threshold, it is not possible to calculate the correction directly. One can, however, apply the median ϵ_0 and ϵ_1 values obtained for sources above the cleaning threshold to sources below. That improves the fluxes from 13% to 4% from the unbiased fluxes and -0.61 to -0.20 in the best case for the $1 < S/N < 3$ range. VLASS images, however, are large-scale mosaics, and a median value for ϵ across all sources (implying a constant clean to dirty beam ratio) is not always suitable for an appropriate correction. Therefore, in general, the corrections should be calculated on a source by source basis. For sources below the cleaning threshold the median value correction is the only option, but may lead to instable results. An application of a common ϵ below the cleaning threshold needs to be investigated on a larger number of datasets.

We recommend to apply the demonstrated flux density and spectral index corrections to all $S/N < 5$ sources in any catalog with VLASS data as additional columns. The correction should be calculated on a source-by-source basis. It may be acceptable to calculate median correction values from sources with flux densities above the cleaning threshold and apply them to sources with flux densities below the cleaning threshold. For easier calculation, we recommend that future version of CASA's TCLEAN *performs this calculation directly, or at least delivers a machine-readable version of the Hessian matrix, its inverse, and the clean component images convolved with the clean beam as additional output.*

6 Bibliography

Cornwell, T. J., IEEE Journal of Selected Topics in Signal Processing, 2, 5, 793-801
Del Rio, et al., 2004, , AJ, 128, 89
Jörsäter & van Moorsel 1995, AJ, 110, 2037 (JvM)
Ott et al., 2001, AJ, 122, 3070
Ott, J. et al., 2012, AJ, 144, 123
Rau & Cornwell, 2011, A&A, 532, 71
Walter, F. et al., 2008, AJ, 136, 2563