

## RICA - Radio Imaging Combination Analysis

MILES LUCAS,<sup>1,2</sup> KUMAR GOLAP,<sup>2</sup> AND TAKAHIRO TSUTSUMI<sup>2</sup>

<sup>1</sup>*Iowa State University, Ames, Iowa*

<sup>2</sup>*National Radio Astronomy Observatory, Socorro, New Mexico*

### ABSTRACT

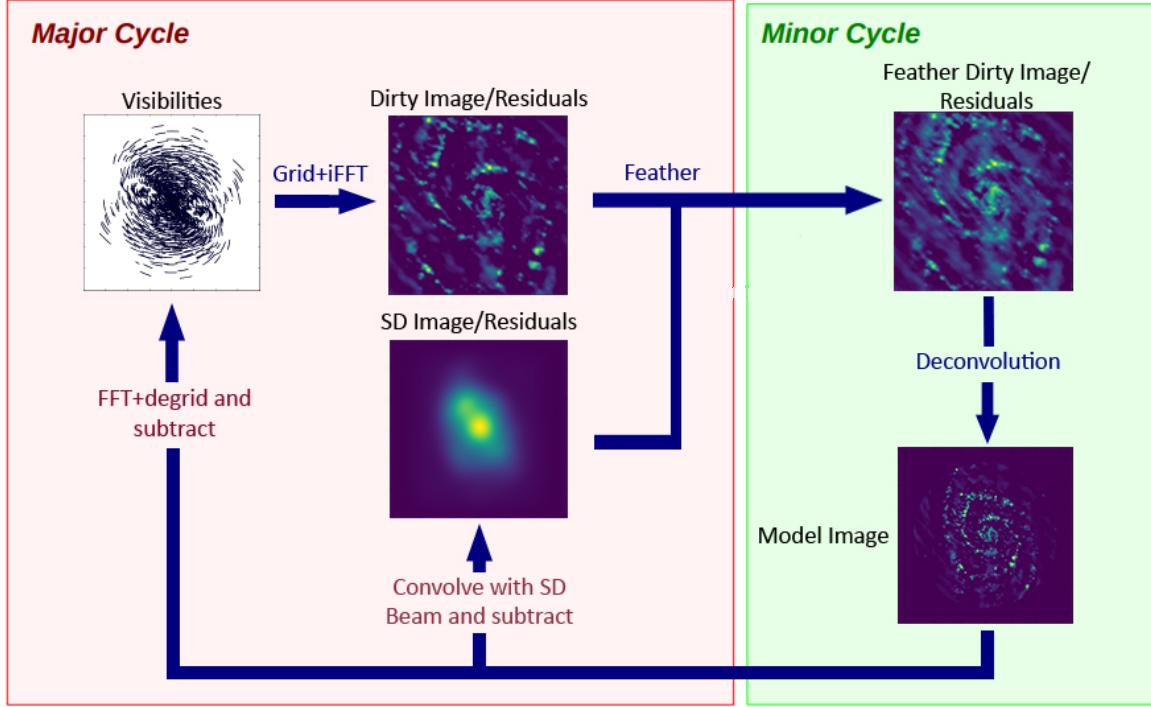
This paper reports on the methodology and some results of testing methods and parameters of combining single dish and interferometer radio images. The metric for analyzing combinations used a ratio of the power spectrum densities of a test image and a true model to determine the accuracy of close UV-spacings. The methods tested were CASA's *feather* task, using a startmodel in CASA's *clean* task, a modified joint-deconvolution CLEAN algorithm, and *tp2vis*. Using simulated images based on the VLA and GBT telescopes of models with various amounts and sizes of structures, many permutations of simulation and imaging parameters were tested along with the combination methods. Some preliminary conclusions are that feathering is weak without adequate single-dish UV coverage overlap, joint deconvolution works well with dirty images, multiscale deconvolution performed well with joint deconvolution with low single-dish UV coverage overlap, and images with few, simple sources do not benefit from combinations in this testing. In addition, this paper presents an open-source codebase to allow researchers to test their own models and parameter permutations to determine their best options for combining data.

### 1. INTRODUCTION

In radio synthesis imaging a common problem arising from interferometry is the lack of zero-spacing data. Without this data, images lack total power information. Single dish radio telescopes retain this total power information but lack the angular resolution capabilities of radio interferometers. To solve this problem, astronomers combine the interferometry data with the total power data. This report seeks to characterize the effectiveness of different combination methods and parameters.

The parameters analyzed were multiscale deconvolution versus normal Hogbom deconvolution, the effect of single-dish size on the combination, masking, and number of total iterations (i.e. depth of clean). One of the previous assumptions is that if the single-dish UV-coverage does not adequately complement the interferometer UV-coverage, feathering will not be effective. Another assumption is that simple point structures should not see much discrepancy in parameters or methods. Multiscale deconvolution is supposed to help pick up fluffy structure better due to its multi-pass method of creating model images in the Cotton-Schwab Cycle, as opposed to Hogbom's  $\delta$ -function method. Therefore, more complex, fluffy structure ought to show more distinction in results.

There are four methods of combination that were tested. The first is CASA's *feather* task. The process of feathering starts with regridding the lower resolution to the higher resolution image. Each is then Fourier transformed and gridded, while the lower resolution data is scaled by the ratio of the clean beams (high resolution/low resolution). The high-resolution grid is then added to this, scaled



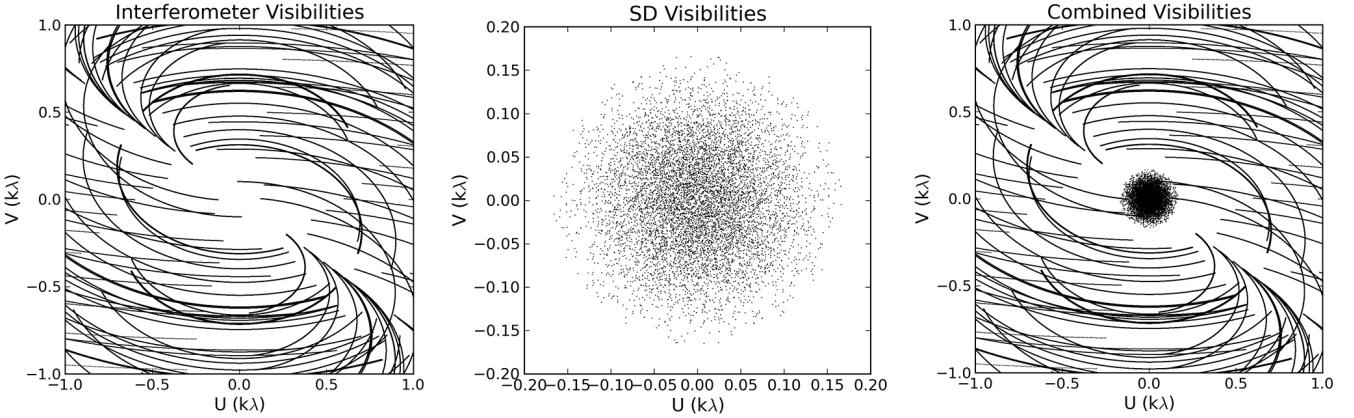
**Figure 1.** A flow chart describing the joint-deconvolution approach to combining interferometer and single dish data. Adapted from Rau, U. & Naik, N. 2018 (in prep).

by a weighting that starts at 0 and increases to 1 as the wavelength increases. This final combination is then inverse Fourier transformed back into the image plane.

The next method is using the total power image as the starting model in CASA's *tclean* task. In the Cotton-Schwab CLEAN algorithm (CSCLEAN), each major-cycle begins with a blank model image and constructs an image on it by using minor-cycle iterations. By using a starting model, this model is no longer blank, but is passed as a parameter. By using the total power image as the starting model, we can hope to retain the total power information through the CLEAN algorithm.

Another modification to the Cotton-Schwab cycle in order to combine images is by doing a modified joint deconvolution. In each major Cotton-Schwab cycle, the visibilities are gridded and inverse Fourier transformed to create an image residual, which is then feathered with the total power residual before deconvolving. The deconvolved model is then convolved with the single-dish beam and subtracted from the previous single-dish residual. The deconvolved model is also Fourier transformed, degridded, and subtracted from the visibilities to create a new residual image. These residuals are then degridded and transformed to finish the major cycle. See Figure 1 for a visual flowchart of the process.

The last method of combination we use is *tp2vis*, which takes the total power image and spoofs a measurement set. The way *tp2vis* accomplishes this is by taking Fourier data from the total power image and sparsely sampling the Fourier data at close spacings (see Koda et al. 2011). This data can be concatenated with the interferometer data and deconvolved in *tclean*. Figure 2 shows how the single-dish is simulated and concatenated with the interferometer data. This should help fill the close-spacing gap created by the interferometer data and subsequently help increase total power information.



**Figure 2.** A look at how the visibilities are added through using `tp2vis` for a model simulated on the VLA A configuration. In the center of the interferometer image in the first plot shows sparsity that is filled in on the third plot. The second plot shows the approximation `tp2vis` makes of the single dish.

## 2. METHODS

In order to evaluate the different combination methods, a metric needed to be created and tested on a suite of models. The metrics used were CLEAN residuals, fidelity images, and a ratio of the power spectrum densities (PSD). The fidelities are given [Equation 1](#)

$$\text{Fidelity}(i, j) = \frac{|\text{Model}(i, j)|}{\max(|\text{Difference}(i, j)|, 0.7 \times \text{rms}(\text{Difference}))} \quad (1)$$

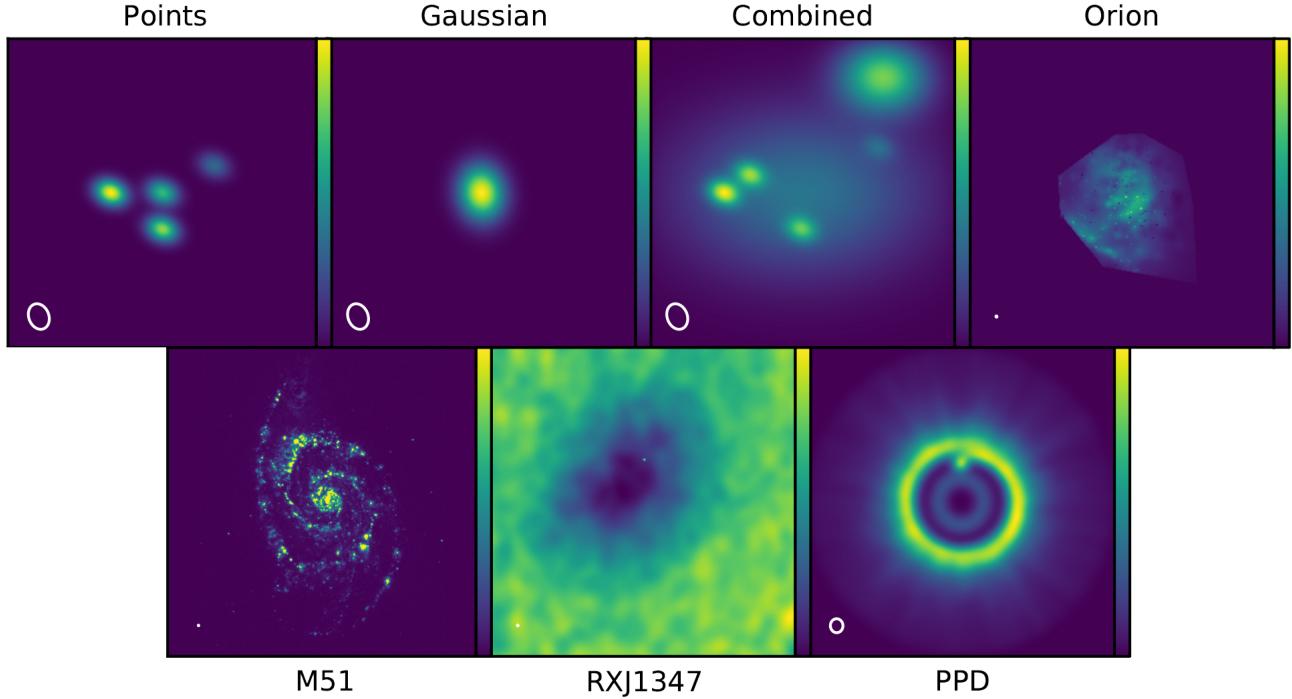
where *model* is the reference image and *diff* is the difference between the test and reference images ([Pety et al. 2011](#), p.19). The exigence for using PSDs is that the zero-spacing power is readily visible for every image. By using these ratios, the closer the ratio is to 1.0 at short spacings (when compared to the true model), the more accurate the combination. In addition, when compared with the total power image, it shows how much effective weight is given to the total power image in the combination. The ratio of the PSDs is given by [Equation 2](#)

$$\text{Ratio}(UV) = \frac{\text{Power}_{test}(UV)}{\text{Power}_{ref}(UV)} \cdot \frac{\text{BA}_{ref}}{\text{BA}_{test}} \quad (2)$$

$$\text{err}(UV) = \left( \frac{\text{Power}_{ref}(UV) + \text{Power}_{test}(UV)}{2} \right)^{-1} \quad (3)$$

where *Power* is the power from the PSD, *BA* is the beam area, *ref* refers to the reference image (model or SD), and *test* refers to whichever image is being tested. The uncertainty given is [Equation 3](#) such that the uncertainty is the reciprocal of the average of the PSDs.

Each combination was compared to both the true model and the single dish, total power image. Many models were tested with various extra parameters. Three models were generated from component lists. One has 4 point sources, one has a single Gaussian source, and one has a mixture of 4 point sources, one very broad Gaussian, and one off-center, stronger Gaussian. There were also various models based off real structure, including M51 (based off an H- $\alpha$  image), Orion, RXJ1347, which is a source displaying the Sunyaev-Zel'dovich effect ([Kitayama et al. 2016](#)), and a protoplanetary disk



**Figure 3.** The models used for testing combination methods convolved with some restoring beam to show structure better.

(PPD) simulation. Figure 3 shows all of the models. It is important to note that these models have been regridded onto a common coordinate system that is not representative of the true astronomical targets. This was in effort to simplify the simulation process for creating measurement sets.

The code base for testing the effectiveness is hosted publicly for anyone to use<sup>1</sup>. In the source code there are many scripts and methods to facilitate simulating, combining, and comparing. Every model is described in `src/_models.py` with a dictionary defining the simulation and `tclean` parameters. The pipeline for testing these models was as follows:

1. Simulate measurement set (MS) based on VLA configurations
2. Simulate single dish image by convolving model with Gaussian
3. Image the MS
4. Do the combinations (feather, startmodel, joint deconvolution, and `tp2vis`)
5. For each combination, do a comparison with the true model (convolved with the restoring beam) and the single dish image

and is implemented in `src/pipeline.py`.

### 2.1. *Simulations and Deconvolution*

All simulations were done using the VLA for the telescope model except for the PPD simulation, which was done using `simalma`. Each test model could define which of the VLA configurations to

<sup>1</sup> <https://gitlab.com/mileslucas/rica>

use, from A, B, BnA, C, CnB, D, and DnC. The spectral window for the simulations mimic the VLA C-band at 4.5 GHz to 5.5 GHz (central 5 GHz) with 101 channels at 10 MHz intervals. The field observed was J2000 21h28m31.0 45°0'0.0''. The observation date and time was 2018/06/01 at 12:00:00.0 UTC. The integration time was 10 s each with a total of 30 000 s. Finally, the data was corrupted with 1 mJy of simple noise.

The PPD model was created by following the [CASA guide for simulation](#). The base model is from [Wolf & D'Angelo \(2005\)](#), which is a simulation of ALMA data at 672 GHz. This involved one pointing at J2000 18h00m00.031s -22d59m59.6s observed for 20 min in the ALMA configuration 20. The total power image was created using *simalma* to recreate the Atacama Compact Array (ACA) with its 12 m single dish telescopes.

The single dish images were created by taking the true model and convolving with a Gaussian equivalent to the single dish beam at the given frequency.

$$\text{FWHM}(\nu, D) = \frac{3.66 \times 10^8 \text{ m s}^{-1}}{\nu \cdot D} \quad (4)$$

Equation 4 gives the Gaussian kernel full-width half-maximum (FWHM) in radians for a frequency in Hz and dish diameter in meters. For instance, to simulate data from the Green Bank Telescope (GBT) with a dish diameter of 100 m at 5 GHz gives a FWHM of 2'31'', which is in agreeance with values from the GBT proposer's guide<sup>2</sup>. Note that this is not the most accurate equation for other telescopes, because there is a constant involved in this equation based off the taper-length of the telescope's feedhorns. In addition, these single dish images have  $5 \times 10^{-6}$  uniform noise added.

Each model defines its own parameters for cleaning. [Appendix A](#) lists every single parameter for each model. As a default, a model simulated in VLA D configuration has an image size of 64 with a cell size of 7.0''. The restoring beam for the D configuration is approximately 20.41'' by 15.60'' so the cell accounts for half to a third of the beam width. In general, cleaning was done with very high numbers of iterations, relying on a threshold to stop the algorithm for consistency. This differs for the models tested at explicitly different cleaned levels.

## 2.2. Comparisons

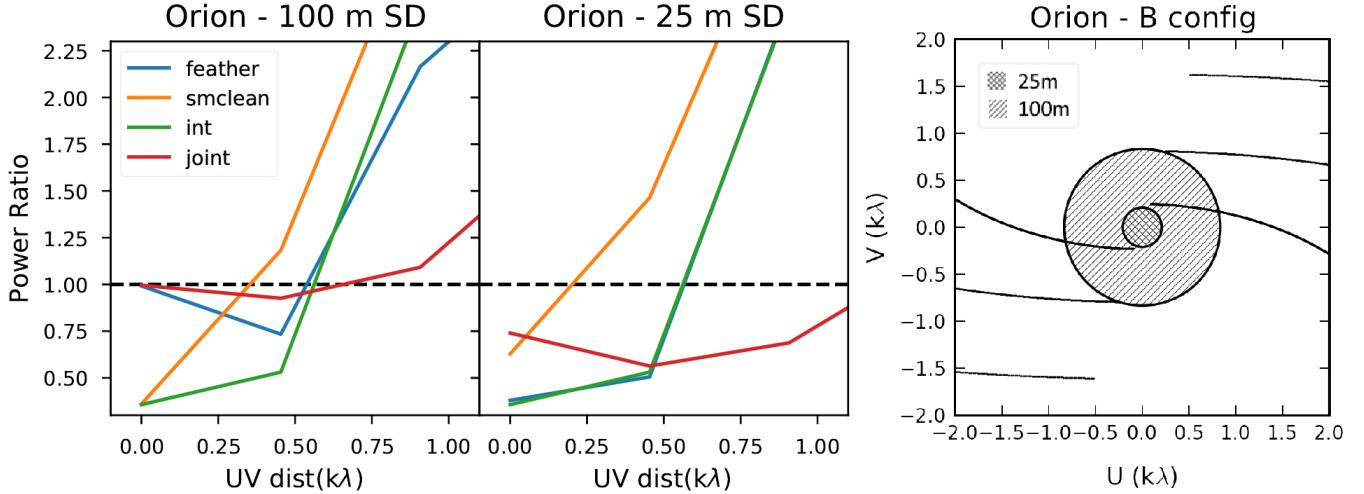
Each comparison was ran with the same clean parameters for consistency. Because *feather* does not actively clean, it has its own set of parameters, but all of the models tested used the default parameters. After all the combination permutations were imaged, they were compared against the true model convolved with the restoring beam and the single dish image alongside the uncombined interferometer image. All of the ratio were saved into individual CSV files for further analysis. In all, each model has 10 comparisons (5 against the models and 5 against the single-dish images) and 4 separate cleans to perform.

In order to run a new model to test comparison methods, there are two general methods. If there exists a true model, simply add an entry to the models dictionary defining the simulation and clean parameters. It is also important to edit *src/simulate.py* to accommodate the new model. Follow the existing code base for guidance. It is also important to create a copy of the model and regrid it to the common coordinate system (use an existing model to retrieve the coordinate system). The model can then be ran through the pipeline along with any other models using *src/pipeline.py*.

<sup>2</sup> p.9, <https://science.nrao.edu/facilities/gbt/proposing/GBTpg.pdf>

If there is no true model, it is still useful to test and compare against the total power image, alone. In this case, using a cleaned image, a measurement set, and a total power image with the `src/combine.py` script will produce all of the combinations. These can then be compared using `src/compare.py`. It is possible to edit `src/pipeline.py` to accommodate special models, see how the PPD model is handled in the pipeline.

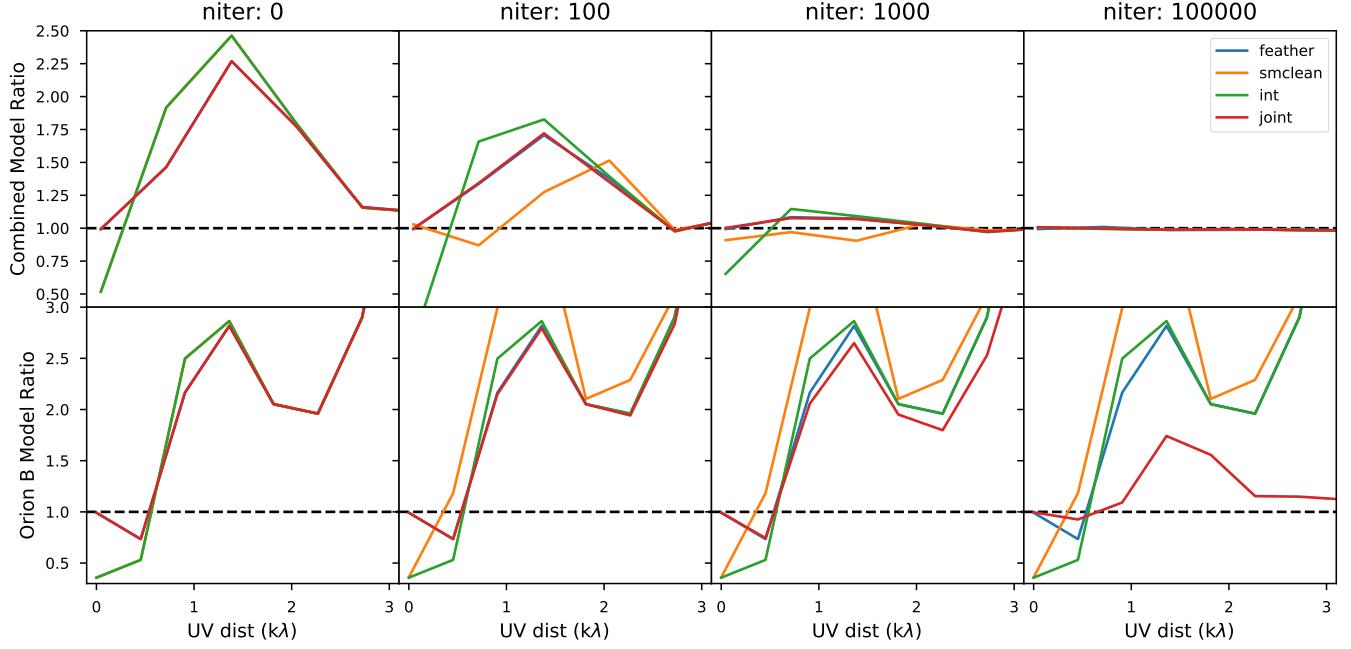
### 3. RESULTS



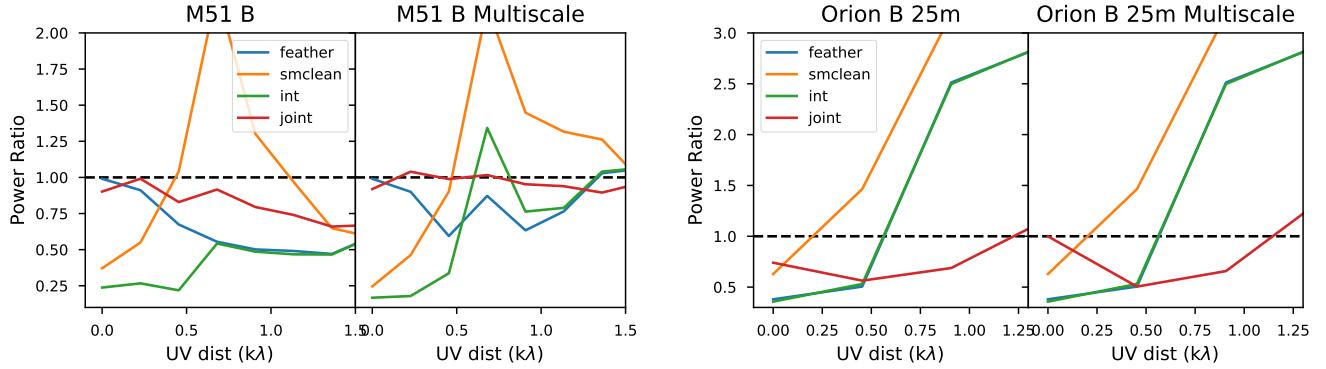
**Figure 4.** The left two plots are ratios for each combination method between two different single dish sizes. The horizontal line marks the ideal ratio for close UV spacing. The right plot shows the UV coverage of the interferometer with the coverage of the 100 m and 25 m single dishes overlaid.

For fidelity images and ratios of every comparison, see Appendix B or the [gitlab page](#). This section will only comment on a few models and comparison highlighting trends. The effect of single dish size was apparent on the accuracy of combinations. Figure 4 shows how the ratios compare between the 100 m single dish size for the Orion model in B configuration and the 25 m single dish size of the same model. Notably, *feather* does not work as well for small single dish sizes. Both start model and joint deconvolution combination methods are able to do better than the plain interferometer with the small dish, but still do not reach that ideal ratio that is possible for the larger single dish. Looking at the coverage of the single dishes, the 25 m dish does not overlap with the interferometer data at all.

Next, Figure 5 shows the effect that depth of cleaning has on the power ratios. Each row corresponds to each model and each column is a depth of cleaning based on the number of iterations of deconvolution. The ‘combined’ model is a much simpler model than the ‘orion-b’ model. First, the joint deconvolution worked well regardless of the depth of cleaning for both models. Even completely dirty images had great total power information. After some amount of cleaning, *feather* quickly rose to the same level as joint deconvolution and in the ‘combined’ model, using a start model started working after some cleaning. In the simpler, ‘combined’ model, all combination methods converged at high enough iterations- even the interferometer image. However, for the more complex, ‘orion-b’ model, start model never converged and the values of the ratios remained erratic even at full cleaning depth.



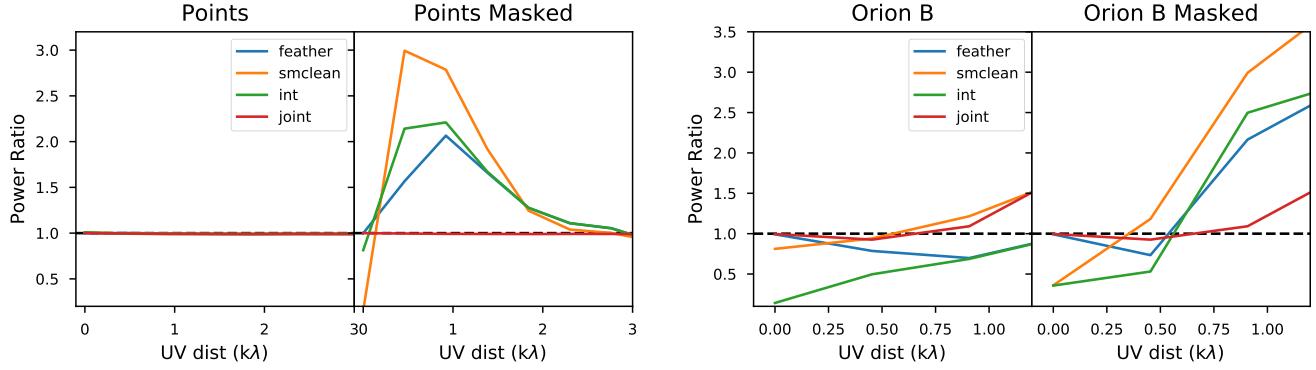
**Figure 5.** A comparison showing the effect on the combinations due to the depth of cleaning for two models, ‘combined’ and ‘orion-b’. The dashed black line represents the ideal ratio for the true model comparison.



**Figure 6.** A comparison showing the effect on the combinations due to multiscale deconvolution for two models, ‘m51-b’ and ‘orion-b’. The dashed black line represents the ideal ratio for the true model comparison.

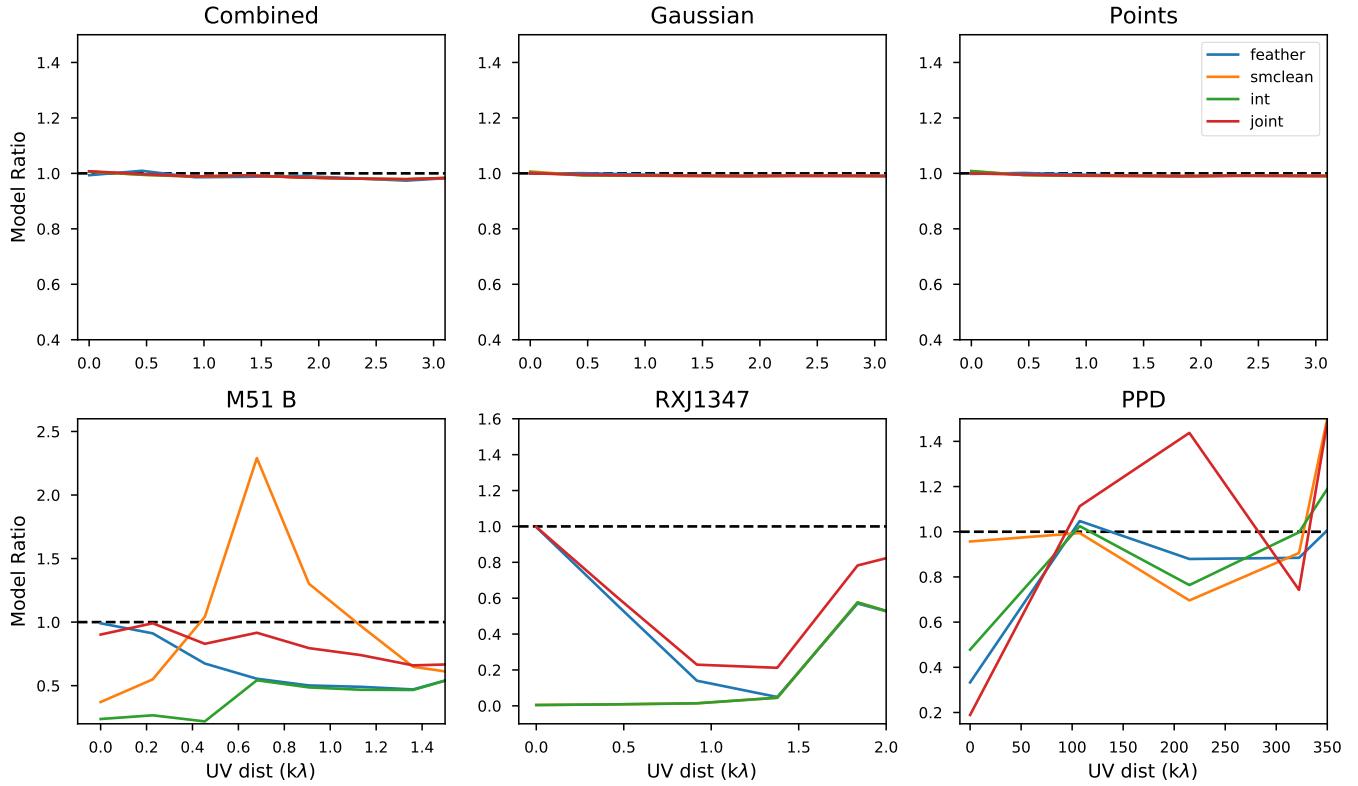
Figure 6 shows how multiscale deconvolution affects the combinations. For each model, there are some slight variations, but in general the ratios stay consistent. The ‘orion-b-25’ multiscale model shows how multiscale can improve efficiency for joint deconvolution when the single dish overlap is low. The ‘orion-b’ multiscale model (not shown) does not show much change between deconvolvers except for an unexplained overshoot in the multiscale joint deconvolution.

Figure 7 shows the effect masking has on the combinations. For the simple ‘points’ model masking causes issues, especially with startmodel. Also, in the ‘orion-b’ model, the startmodel value falls at 0 UV spacing, as well as the plain, cleaned image. In general, the masking decreases the convergence of the combination methods at the same clean level and may have further affects by changing the



**Figure 7.** A comparison showing the effect on the combinations due to masking for two models, ‘points’ and ‘orion-b’. The dashed black line represents the ideal ratio for the true model comparison.

performance of the uncombined image. This is also apparent in the ‘RXJ1347’ masked model, where the 0 UV spacing value does not change but the convergence is definitely affected.



**Figure 8.** A comparison of many models. The top row are ‘simple’ models that do not contain much structure or extended emission, whereas the second row contains more complex models with different structural scales and interesting phenomenon (Sunyaev–Zel’dovich effect in RXJ1347).

Figure 8 shows an overview of the performance of different models. Simpler models with point sources and Gaussians converge at a ratio of 1 and show very little difference until scales at hundredths of a ratio. The more complex models have no easy, clear trend. Each one of these combinations is

highly dependent on the structure and parameters. It is interesting that the ALMA ‘PPD’ model only performs well with startmodel and that *feather* and joint deconvolution actually perform worse than the uncombined image.

#### 4. CONCLUSION

From the results of the single dish sizes, the main factor in algorithm success is the amount of overlap of the single dish and interferometer data. This could be the reason why feather fails so much and why joint deconvolution and start model are not quite as good, either. Future studies might find the point at which there is enough overlay for each method to “succeed”.

In regards to the depth of cleaning, in general, more is better. Joint deconvolution seemed to work for any dirty image with enough single dish coverage, but for future studies, it would be nice to try and see if joint deconvolution works for dirty images of many more models. In conjunction with the study on single dish size, there were no combinations that retrieved the total power information well from the dirty image due to insufficient single dish coverage.

The multiscale analysis shows that using the multiscale deconvolver can make up for lack of single dish UV coverage in some instances. When there is already decent single dish UV coverage, though, there does not seem to be a significant effect. Also, simple models of few Gaussian or point sources converge when cleaned appropriately, regardless of the methods or parameters of combination, even skipping any combination method retained the close-spacing information well enough in this study.

There are many more permutations of simulation and imaging parameters to test, though. Especially simulating using telescopes other than the VLA and GBT. In addition, results from *tp2vis* would be a nice supplement to the current results. The codebase of RICA lends itself to open-source development and use, though. So it is encouraged to attempt different models and parameters instead of relying on this sole paper for seeing the differences given your own datasets. Hopefully this code will allow a better understanding of the differences in single dish combination methods.

#### ACKNOWLEDGMENTS

Thank you to Dr. Kumar Golap and Dr. Takahiro Tsutsumi for their guidance and assistance. This work was funded by the National Science Foundation in partnership with the National Radio Astronomy Observatory and Associated Universities Incorporated.

*Software:* CASA, tp2vis (Koda et al. 2011)

#### REFERENCES

- |   |   |
|---|---|
| Kitayama, T., Ueda, S., Takakuwa, S., et al. 2016,<br>PASJ, 68, 88    | Pety, J., Gueth, F., & Guilloteau, S. 2011, ALMA<br>Memo #398 Impact of ACA on the Wide-Field<br>Imaging Capabilities of ALMA, Tech. rep. |
| Koda, J., Sawada, T., Wright, M. C. H., et al.<br>2011, ApJS, 193, 19 | Wolf, S., & D’Angelo, G. 2005, ApJ, 619, 1114   |

## APPENDIX

### A. MODEL PARAMETERS

The following table describes the parameters used for cleaning the model images. The only exception is the model ‘combined-sp’, which additionally uses `specmode='cube'` and `nchan=-1` to facilitate cube imaging. For more details, look at the `src/_models.py` file, which this table is generated from.

**Table 1.** The parameters for every model ran through the pipeline

Name	Model	Config	SD size	imsize	cell	niter	cycleniter	threshold	deconvolver	scales	mask
points	points	VLA d	100 m	[64, 64]	7.00arcsec	100000	100	1e-5Jy	hogbom	No	Yes
points-masked	points	VLA d	100 m	[64, 64]	7.00arcsec	100000	100	1e-5Jy	hogbom	No	No
gauss	gauss	VLA d	100 m	[64, 64]	7.00arcsec	100000	100	1e-5Jy	hogbom	No	No
combined	combined	VLA d	100 m	[64, 64]	7.00arcsec	100000	100	1e-5Jy	hogbom	No	No
combined-sp	combined	VLA d	100 m	[64, 64]	7.00arcsec	100000	100	1e-5Jy	hogbom	No	No
combined-n0	combined	VLA d	100 m	[64, 64]	7.00arcsec	0	100	1e-5Jy	hogbom	No	No
combined-n100	combined	VLA d	100 m	[64, 64]	7.00arcsec	100	100	1e-5Jy	hogbom	No	No
combined-n1000	combined	VLA d	100 m	[64, 64]	7.00arcsec	1000	100	1e-5Jy	hogbom	No	No
combined-ms	combined	VLA d	100 m	[64, 64]	7.00arcsec	100000	100	1e-5Jy	multiscale	[0, 3, 10, 30]	No
m51	m51	VLA d	100 m	[128, 128]	7.00arcsec	100000	100	1e-12Jy	hogbom	No	No
m51-ms	m51	VLA d	100 m	[128, 128]	7.00arcsec	100000	100	1e-12Jy	multiscale	[0, 3, 10, 30]	No
m51-b	m51	VLA b	100 m	[1400, 1400]	0.65arcsec	100000	100	1e-12Jy	hogbom	No	No
m51-b-ms	m51	VLA b	100 m	[1400, 1400]	0.65arcsec	100000	100	1e-12Jy	multiscale	[0, 3, 10, 30]	No
orion-b-unmasked	orion	VLA b	100 m	[700, 700]	0.65arcsec	100000	400	1e-9Jy	hogbom	No	Yes
orion-b	orion	VLA b	100 m	[700, 700]	0.65arcsec	100000	400	1e-9Jy	hogbom	Yes	Yes
orion-b-n0	orion	VLA b	100 m	[700, 700]	0.65arcsec	0	100	1e-9Jy	hogbom	Yes	Yes
orion-b-n100	orion	VLA b	100 m	[700, 700]	0.65arcsec	100	100	1e-9Jy	hogbom	Yes	Yes
orion-b-n1000	orion	VLA b	100 m	[700, 700]	0.65arcsec	1000	100	1e-9Jy	hogbom	Yes	Yes
orion-b-ms	orion	VLA b	100 m	[700, 700]	0.65arcsec	100000	400	1e-9Jy	multiscale	[0, 3, 10, 50]	Yes
orion-b-25	orion	VLA b	25 m	[700, 700]	0.65arcsec	100000	400	1e-9Jy	hogbom	Yes	Yes
orion-b-25-n0	orion	VLA b	25 m	[700, 700]	0.65arcsec	0	100	1e-9Jy	hogbom	Yes	Yes
orion-b-25-n100	orion	VLA b	25 m	[700, 700]	0.65arcsec	100	100	1e-9Jy	hogbom	Yes	Yes

**Table 1** continued on next page

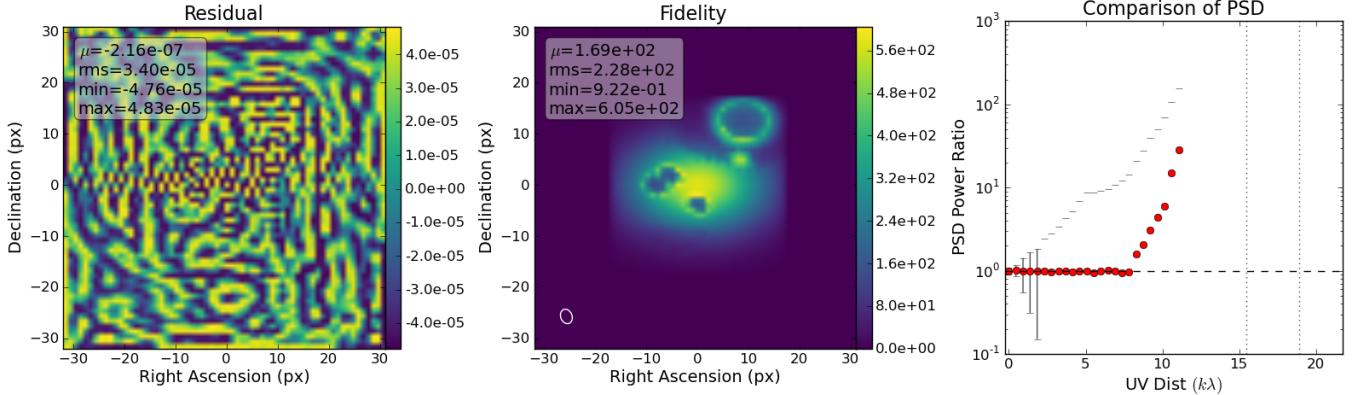
**Table 1** (*continued*)

Name	Model	Config	SD size	imsize	cell	niter	cycleniter	threshold	deconvolver	scales	mask
orion-b-25-n1000	orion	VLA b	25 m	[700, 700]	0.65arcsec	1000	100	1e-9Jy	hogbom		Yes
orion-b-25-ms	orion	VLA b	25 m	[700, 700]	0.65arcsec	100000	400	1e-9Jy	multiscale	[0, 3, 10, 30]	Yes
orion-c	orion	VLA c	100 m	[216, 216]	2.12arcsec	100000	400	1e-9Jy	hogbom		No
RXJ1347	RXJ1347	VLA a	100 m	[2268, 2268]	0.20arcsec	100000	100	1e-6Jy	hogbom		No
RXJ1347-masked	RXJ1347	VLA a	100 m	[2268, 2268]	0.20arcsec	100000	100	1e-6Jy	hogbom		Yes
ppd	ppd	ALMA 20	12 m	[192, 192]	0.01arcsec	100000	100	1e-7Jy	hogbom		No

## B. MODEL RESULTS

For data tables of every single comparison, visit the [RICA repository](#).

**Fig. Set 1. Model Results**



**Figure 9.** Comparison of feather combination to the true model for combined simulation. The ratio is Equation 2. The black ticks are error bars on a scaled uncertainty based on Equation 3. The vertical lines are the maximum UV distance of the interferometer at the longest and shortest wavelengths, respectively. The complete figure set (125 figures) is available at the [RICA repository](#).