

EVLA Memo 231

Source peeling for JVLA data with 3GC tools

Viral Parekh

National Radio Astronomy Observatory, Socorro, NM

May 2, 2024

Abstract: In this memo, we have applied direction-dependent calibration to the three JVLA data sets using CubiCal software. In all of these data, strong and compact radio sources cause artifacts and hamper the detection of faint radio emissions. We used third-generation calibration (3GC) techniques using DDFacet, WSClean, and CubiCal software and performed “peeling” to these JVLA data sets. After peeling, we found significant improvements in images. These 3GC tools are developed for future radio telescopes such as MeerKAT, LOFAR, etc. which have wide-field and wide-band low-frequency receivers. The peeling with CubiCal is an automated technique and one can select multiple strong sources for direction-dependent calibration. CubiCal can solve direction-dependent gain for all of these sources simultaneously, and generate residual columns after subtracting problematic sources, for further imaging. These tools can be used for JVLA data as complementary tools with CASA (for example `awprojection` in `tclean`).

1 Background

The emergence of newer and more sophisticated radio telescopes has included properties such as non-coplanar arrays, large FOVs, and large fractional bandwidths. These upgrades have consequently brought a host of new calibration and observational errors that require new and elaborate calibration techniques. Furthermore, detecting a faint and weak radio emission is extremely hard; sometimes, nearby strong and bright radio sources (of few 100 mJy) limit the ability to observe weak radio signals of astrophysical sources due to artifacts (Parekh et al., 2022, 2021). If any bright source(s) is not characterized by the calibration process (direction-independent self-calibration) and due to the imperfect time and frequency varying primary beam model of the telescope, it limits the dynamic range of the final image (Heywood et al., 2022). Even with such sophisticated instruments and state-of-the-art data processing pipelines, astronomical data is still populated with a host of observation- and calibration-related errors. These errors are classified as DDE errors and are source-specific. They can be visually recognized by the point spread function (PSF) sidelobe-like artifacts they produce around strong sources, most significantly around sources located far away from the phase center towards the edge of the primary beam. Sometimes, these strong artifacts, in radial-spikes form, corrupt the whole image and hamper the achievable DR at the science target position. Accounting for these DDE errors is currently achieved through the use of more modern, third-generation calibration (3GC) techniques (Smirnov et al., 2024).

To do direction-dependent (DD) calibration, there are two main techniques - “peeling” and facet-based DD self-calibration- which can be carried out with 3GC software (Noordam & Smirnov, 2010). The peeling technique is used to perform direction-specific (towards

Table 1: Sample of JVLA data

Object name	JVLA band	Configuration	Date of Observation	Project ID
MOO_0917-0700	S	B	2021-10-15	21B-235
ZwCl2341.1	S	C	2014-09-19	14B-157
J1023PT	X	A	2016-10-21	SH0218

specific strong source(s)) calibration and modeling of this strong point source(s). It can be done for a single or multiple bright sources simultaneously. The facet-based direction-dependent calibration divides the image plane into several tessellations depending on the SNR of the data. Then, various steps in the calibration can be performed in parallel per direction or tessellation and improve the image. In this memo, we have shown the peeling technique for JVLA data with the employing of 3GC calibration tools. Currently, CASA does not support peeling or direction-dependent calibration. But, it is possible to use these tools in the framework of CASA or use these tools in the CASA environment.

2 Third-Generation Calibrations tools

Generally we refer to these new and advanced tools as 3rd generation calibration developed for next-generation radio interferometers (1GC refers to bandpass and complex gain calibrations, 2GC refers to self-calibration). It incorporates a variety of tools for imaging, calibration, and source modeling. One can download these tools from ‘<https://github.com/ratt-ru>’. We have employed the following listed tools in our analysis. [CubiCal](#) (to perform direction-dependent calibration and peeling, Kenyon et al. (2018)), [WSClean](#) (high-speed imaging and generating clean components list, Offringa et al. (2014)), [DDFacet](#) (facet-based imaging and generating full sky models, Tasse et al. (2018)), [breizorro](#) (image masking tool), [Crystalball](#) (source modeling and generating source-specific data column in measurement set). There is a next version of CubiCal that is under development, called [Quartical](#). Quartical is faster and can combine more than one Jones term (G, DE, K, B, etc.) in the calibration process. Facet-based calibration can be done with the [killMS](#) software (Tasse, 2014). The user has to install these tools externally or the user can download docker images and use the [Stimela](#) environment to execute these tools. Stimela also provides a pipeline framework and one can use different radio astronomy packages in one script. We have developed a generic pipeline for JVLA data using 3GC tools.

2.1 Source peeling with CubiCal

CubiCal employs complex optimization routines to carry out fast radio interferometric calibration. It can be used for both direction-independent (phase only and amplitude and phase self-calibration) and direction-dependent calibration. CubiCal performs direction-dependent calibration using a simultaneous form of the peeling approach called differential gains (Smirnov, 2011a). It can apply corrections to many sources simultaneously from the self-calibrated sky model, as compared to the one-by-one iterative algorithm used by the typical peeling method (Williams et al., 2019). These sources are manually marked and contain information on the specific direction in which to perform peeling. The radio interferometer measurement equation (RIME, Smirnov (2011a,b)) used for the differential gains takes the form:

$$D_{l,m} = G_l \left(\sum_{l,m} \sum_s \Delta E_l P_l S_{l,m,s} P_m^H \Delta E_m^H \right) G_m^H \quad (1)$$

G are direction-independent errors affecting the entire FOV, constrained to phase-only solutions while ΔE are direction-dependent errors that are notably present around strong sources far away from the phase center. These errors are constrained to fully complex 2×2 Jones matrices. P terms are implemented for antenna-specific primary beam models to correct for time and frequency-varying gains, and S terms denote sky models or clean component models of the given observation. The sum is over all sources S and their respective directions l, m . CubiCal solves for the G terms (DI) on small-time/frequency scales by the field as a whole while simultaneously solving for ΔE terms (DDE) on larger time/frequency scales for a subset of sources. The advantage of using the above RIME is that most of the clean components or all-sky flux are taken into account while solving direction-dependent gains. This, to a large extent, avoids the flux suppression typically observed in the self-calibration (a&p) process. This RIME equation also ensured that direction-independent gains were not affected by flux variability because of the effects of the antenna’s far-field response on bright off-axis sources. Below, we describe two such peeling methods that are tested and could be implemented in the automated data analysis pipeline.

3 “peeling” method for JVLA data

We have applied the above mentioned 3GC tools to the JVLA data (Table 1). This data has substantial direction-dependent errors visible even after self-calibration. Below we outline our procedure.

First we downloaded the raw data from the JVLA archive and then we ran an automated SRDP calibration pipeline (flagging, Bandpass, and complex gain calibration, split the target and initial imaging) and self-calibration pipeline on each data. Then we used two different methods - (1) DDFacet and CubiCal and (2) WSClean and CubiCal to experiment with peeling techniques on JVLA data. Both of these techniques are similar and it depends on the user which one to use. For example, if the user has used WSClean in the self-calibration rounds, then the user can directly use the final visibility for the CubiCal step and avoid the initial imaging steps to generate the all-sky model.

DDFacet+CubiCal: The following steps have been performed to do the peeling.

1. We first run the DDFacet on the self-calibrated target’s measurement set. In the first run, we use only a few thousand iterations and do shallow cleaning.
2. We generate the 5σ mask from (1) using breizorro masking tool.
3. We again run the DDFacet with the tight mask and a few ten thousand iterations. This will generate the full sky model of the target visibility. This model is stored as "DicoModel" in the form of a Python dictionary for each of the spws present in the data.
4. Then we manually select the bright source(s)’ location (which needs to be peeled) using the ds9 viewer and save it in the region file.
5. Next we supply the ds9 region file and all sky clean component models (from 3) to the CubiCal software. CubiCal will perform direction-dependent calibration towards the selected bright and faulty source against the full sky model and generate a residual column in the measurement set after subtracting the strong source.

```
--model-list MODEL_DATA+-image_DI.DicoModel@bright_src.reg:image_DI.DicoModel@bright_src.reg
```

"`model-list`" is one of the CubiCal parameters. It can read the multiple model components. Here, it takes the direction-independent full sky model from `MODEL_DATA` column subtracting all fluxes from the selected region, and uses this product as a direction-independent model. The specification after the semi-column in "`model-list`" parameter specifies a number of direction-dependent models from the component model generated by the DDFacet and ds9 region file.

6. Lastly we image the residual column to get the final image. The final image will be free from artifacts and have a better image quality. Users can use this improved map for further science analysis.

WSClean+CubiCal: The following steps have been performed to do the peeling.

1. We first run the WSClean on the self-calibrated target's measurement set. In the first run, we use only a few thousand iterations and do shallow cleaning.
2. We generate the 5σ tightest mask from (1) using breizorro masking tool.
3. We again run the WSClean with a tight mask and a few ten thousand iterations. This will generate the full sky model component list (with `--save-source-list` parameter) for the target visibility.
4. Then the user can select the bright source(s) (which need to be peeled) location using the ds9 viewer and save the region file.
5. Using crystalball software, we can predict the bright source model column in the measurement set. This column has only a model for the problematic source. One can predict as many model columns into the measurement set for a set of bright sources.
6. CubiCal can read both columns and will perform direction-dependent calibration towards the bright and faulty source against the full sky model and generate residual column in the measurement set after subtracting strong source.

```
--model-list MODEL_DATA+-bright_source_column:bright_source_column
```

Here `MODEL_DATA` column is generated by WSClean and has all sky model components. `bright_source_column` is predicted by crystalball and it has only model of bright point source. One can include several directions here (e.g. `bright_source_column1:bright_source_column2 ...`)

7. Finally the user can image the residual column to get the final image. The final image will be free from artifacts and have a better image quality. Users can use this improved map for further science analysis.

4 Results and Discussion

We applied the “peeling” method to three JVLA data sets (Table 1). We show our results in Figs 1, 2, and 3. In these experiments, for objects 1 and 2, we used DDFacet+CubiCal and for object 3, we used WSClean+CubiCal. In all images, the left image is without peeling (or only a self-calibrated DI map) and the right image is after the peeling of the bright source. Generally, DDFacet takes 3-4 hours to generate the map of ~ 6 hrs of data. WSClean also takes 3-4 hours. CubiCal also takes 4-5 hours to generate the calibrated residual data. The whole process takes 12-14 hours to get the final image after direction-dependent calibration.

In object 1, there are two compact and bright point sources located at 8.7' from the phase center in the South. The combined flux density is ~ 35 mJy. We derived complex DD gain solutions toward these sources and then subtracted them from the data. One can see that after peeling, the image has been improved and artifacts are removed from the data.

In object 2, there are two diffuse radio relics located in the center. A bright source (of flux density ~ 26 mJy) is located at $\sim 9'$ in the northwest direction from the phase center, causing artifacts that affect the science target (double relics) at the center. After peeling, there is a significant improvement in the central region of the map.

In object 3, a complex radio source is located at the edge of the primary beam ($\sim 2'$ from the phase center) in the West. This radio source is causing artifacts. We modeled this radio source with `crystalball` and predicted the model into the data and calibrated this model against the all-sky model followed by subtraction from the data. Finally, we imaged the residual visibility and improved the image quality by removing artifacts.

These maps are dynamic range limited to only a few thousand. There are not many bright sources present in the images (except strong sources causing artifacts). After peeling, we found improvement in the image quality and minimize the artifacts around strong sources and hence local rms improvements. We also compared fluxes and positions of radio sources between images generated with the CASA pipeline (`tclean`) and `DDFacet` for the same object. Our comparison is plotted in Fig 4. As both imagers have different imaging parameters and cleaning algorithms, both images are not identical. In both images, we used `mosaic` gridding and `mtmfs` ($n=2$). There is flux dispersion between `tclean` and `DDFacet` is $< 10\%$.

In this memo, our objective is to explore the 3GC tools for JVLA data and apply the “peeling” technique to improve the image quality (for those images that have strong sources and artifacts). Generally, DD errors are more prone to low-frequency observations where FOV is large and the ionosphere is also playing a role. In our current work, we also noticed DD errors at X-band data, hence this error could arise at higher frequency also.

5 Future plan

It has been noted that DDE errors are visible in the VLASS data where the bright unresolved source(s) cause errors and generate artifacts. These newer tools can be applied to the VLASS data. `DDFacet` and `WSClean` software are mainly used for a single field and do not support the OTF mosaics scanning mode data, hence we need to use `tclean` and `CubiCal` to apply direction-dependent calibration to the VLASS data. `CubiCal` can also apply primary beam corrections to the visibility while applying gain corrections. It can take eight beam models, i.e. a real and imaginary part for each of the four Jones terms (RR, RL, LR, and LL). There are beam models available for the JVLA S-band ([EVLA Memo 195](#)) which can also be used for this work.

References

- Heywood, I. et al. 2022, MNRAS, 509, 2150, 2110.00347
- Kenyon, J. S., Smirnov, O. M., Grobler, T. L., & Perkins, S. J. 2018, MNRAS, 478, 2399, 1805.03410
- Noordam, J. E., & Smirnov, O. M. 2010, A&A, 524, A61, 1101.1745
- Offringa, A. R., McKinley, B., & Hurley-Walker, N. 2014, MNRAS, 444, 606, 1407.1943

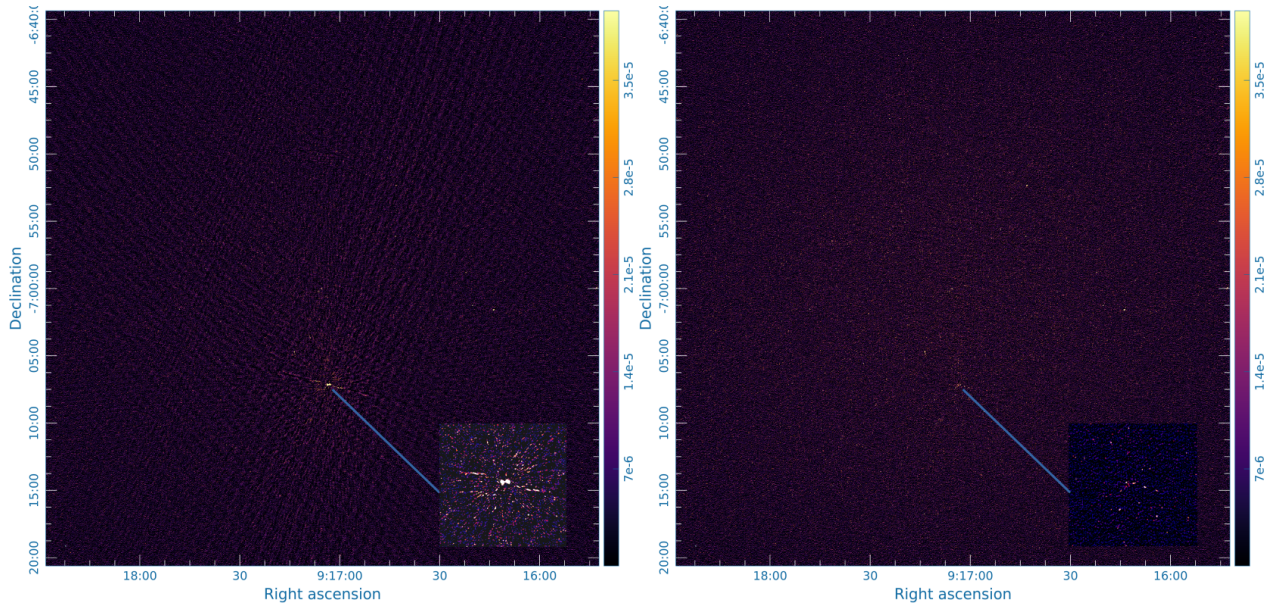


Figure 1: MOO_0917-0700. Left is a direction-independent calibrated image. The right is a direction-dependent calibrated image. Both images have global rms of $\sim 10 \mu\text{Jy beam}^{-1}$. We measured local rms around strong source is $\sim 17 \mu\text{Jy beam}^{-1}$ (DI) and $12 \mu\text{Jy beam}^{-1}$ (DD).

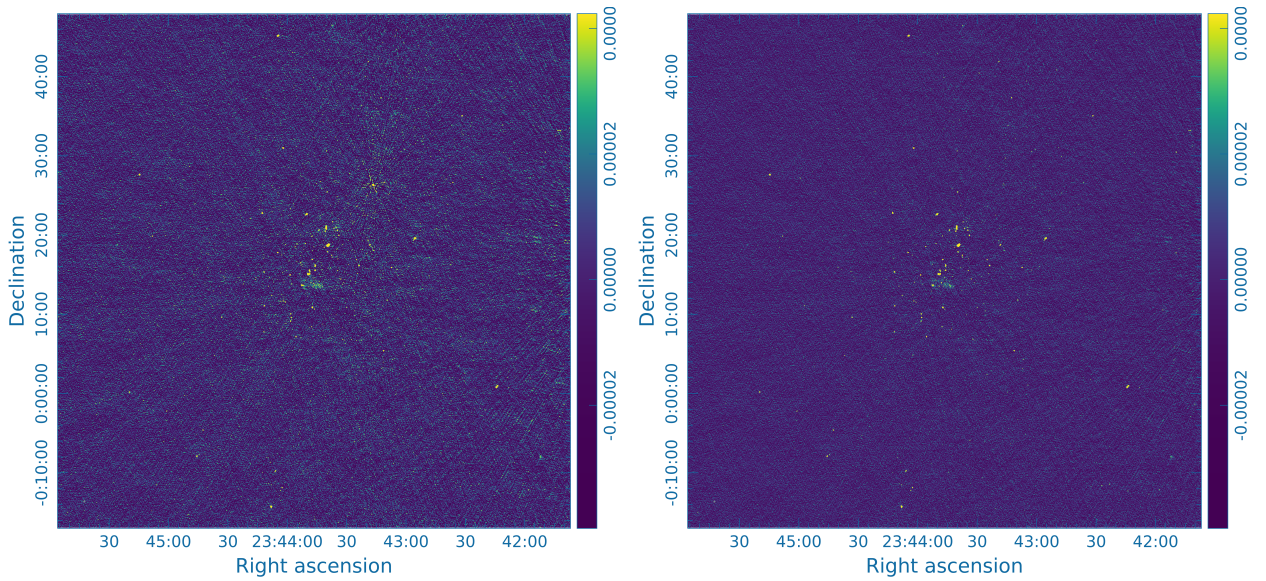


Figure 2: ZwCl2341.1. Left is a direction-independent calibrated image. The right is a direction-dependent calibrated image. Both images have global rms of $\sim 9 \mu\text{Jy beam}^{-1}$. We measured local rms around strong source is $\sim 15 \mu\text{Jy beam}^{-1}$ (DI) and $10 \mu\text{Jy beam}^{-1}$ (DD).

Parekh, V., Kincaid, R., Hugo, B., Ramaila, A., & Oozeer, N. 2021, *Galaxies*, 9, 90, 2111.01858

Parekh, V. et al. 2022, *MNRAS*, 509, 3086, 2110.07713

Smirnov, O. M. 2011a, *A&A*, 527, A106, 1101.1764

———. 2011b, *A&A*, 527, A107, 1101.1765

Smirnov, O. M. et al. 2024, *Monthly Notices of the Royal Astronomical Society*, 528, 6517, <https://academic.oup.com/mnras/article-pdf/528/4/6517/56720461/stae303.pdf>

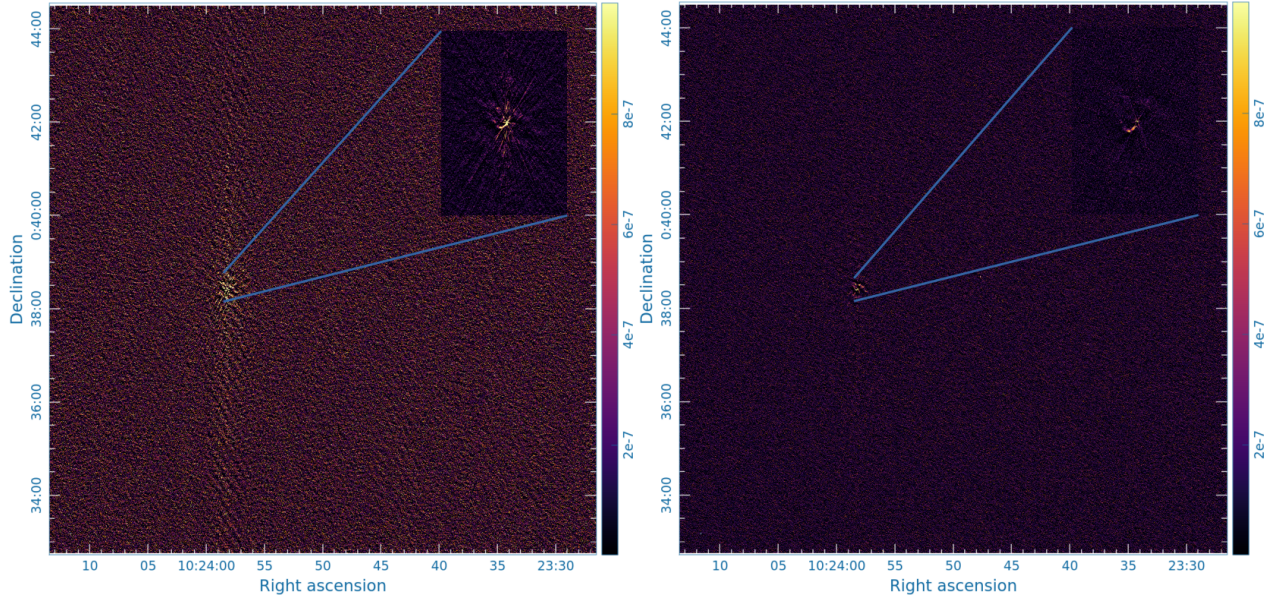


Figure 3: J1023PT. Left is a direction-independent calibrated image. The right is a direction-dependent calibrated image. Both images have global rms of $\sim 19 \mu\text{Jy beam}^{-1}$. We measured local rms around strong source is $\sim 26 \mu\text{Jy beam}^{-1}$ (DI) and $18 \mu\text{Jy beam}^{-1}$ (DD).

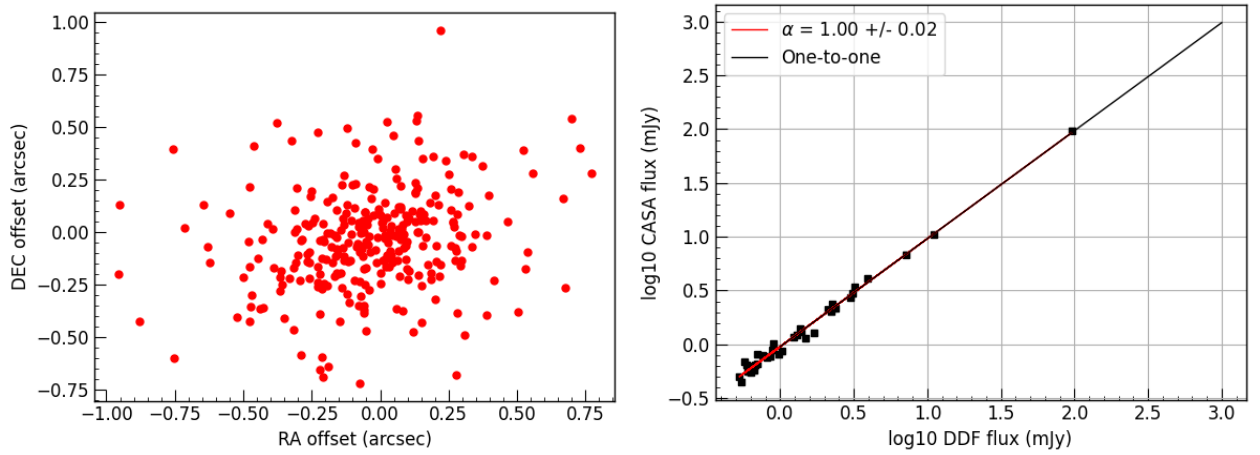


Figure 4: (Left) Astrometry between `tclean` and `DDFacet`. (Right) Flux values comparison between `tclean` and `DDFacet`.

Tasse, C. 2014, arXiv e-prints, arXiv:1410.8706, 1410.8706

Tasse, C. et al. 2018, *A&A*, 611, A87, 1712.02078

Williams, P. K. G., Allers, K. N., Biller, B. A., & Vos, J. 2019, *Research Notes of the American Astronomical Society*, 3, 110

DDFacet parameters

```
DDF.py --Data-MS data.ms
--Data-ChunkHours 1
--Data-Sort False
--Data-ColName DATA
--Output-Images all
--Output-Name=image_DI
--Output-Also DdPMRIikz
--Parallel-NCPU=10
--Deconv-CycleFactor=0
--Deconv-MaxMajorIter=20
--Deconv-MaxMinorIter 500000
--Deconv-Mode Hogbom
--Deconv-Gain 0.1
--Deconv-FluxThreshold 1.0e-6
--Deconv-AllowNegative True
--Deconv-PeakFactor 0.25
--Deconv-RMSFactor=0.000000
--Weight-Mode Briggs
--Weight-Robust 0.0
--Image-NPix=7250
--Image-Cell 0.9
--CF-wmax 10000.0
--CF-Nw 100
--Facets-NFacets=20
--Freq-NBand 10
--Freq-NDegradBand 15
--Mask-External=None
--Mask-Auto False
--Mask-SigTh 10
--Mask-FluxImageType ModelConv
--Cache-Dir=.
--Cache-Reset True
--Hogbom-PolyFitOrder 2
--Facets-DiamMax 180
--Facets-DiamMin 0
--Facets-PSFOversize 1.0
--Facets-PSFFacets=0
--Facets-Padding 1.7
--Predict-colName MODEL_DATA
--Misc-ConserveMemory=1
--Log-Boring True
--Log-Memory True
--Predict-MaskSquare=None
```

CubiCal parameters

```
gocubical --data-ms data.ms
--data-column DATA
--out-column SUBDD_DATA
```

```

--out-model-column MODEL_OUT
--out-name ddcals_0
--out-overwrite 1
--out-mode sr #solve and generate corrected residuals
--out-casa-gaintables False
--out-subtract-dirs 1: #subtract all direction models except all-sky model (1)
--out-plots 0
--bbc-plot 0
--log-verbose solver=2
--weight-column WEIGHT
--flags-auto-init legacy
--flags-apply FLAG
--flags-reinit-bitflags False
--sol-jones G,DD #Jones terms to be solved (complex gain and direction-dependent)
--sol-min-bl 00.0
--sol-term-iters [50,90,50,90] #Number of iterations solved per Jones term
--g-solvable True
--g-type complex-2x2
--g-max-prior-error 0.33
--g-max-post-error 0.33
--g-update-type phase-diag
--g-save-to ampphase-0.paramdb
--g-clip-high 1.7
--g-clip-low 0.25
--g-time-int 60
--g-freq-int 0
--model-ddes auto
--model-list MODEL_DATA+-image_DI.DicoModel@bright_src.reg:image_DI.DicoModel@bright_src.reg
--data-freq-chunk 100
--data-time-chunk 100
--madmax-enable 1
--madmax-plot 0
--madmax-threshold [0,10]
--madmax-global-threshold 0,12
--madmax-estimate corr
--dd-dd-term 1
--dd-time-int 100
--dd-freq-int 100
--dd-clip-low 0.0
--dd-clip-high 0.0
--dd-solvable 1
--dd-type complex-2x2
--dd-save-to ddjone_cal_1.paramdb
--dd-fix-dirs 0
--dd-max-prior-error 0.35
--dd-max-post-error 0.35
--dd-prop-flags always
--degridding-FITsFeedSwap 1
--degridding-OverS 11
--degridding-Support 7

```

```
--degridding-Nw 100
--degridding-wmax 0
--degridding-Padding 1.7
--degridding-NDegradBand 15
--degridding-MaxFacetSize 0.15
--degridding-MinNFacetPerAxis 1
--dist-max-chunks 25
--dist-ncpu 5
--dist-nworker 25
--dist-nthread 15
--dist-safe 0
```

WSClean parameters

```
wsclean -name image_DI
-j 5 -mem 100
-weight briggs 0.0
-super-weight 1.0
-weighting-rank-filter-size 16
-taper-gaussian 0
-size 7250 7250
-scale 0.9asec
-channels-out 10
-pol I
-data-column DATA
-niter 50000
-auto-threshold 0.05
-gain 0.1
-mgain 0.9
-join-channels # activate wideband imaging
-fit-spectral-pol 3
-multiscale-scales True
-multiscale-scale-bias 0.6
-fit-beam
-elliptical-beam
-padding 1.3
-save-source-list # save the clean components list file
-parallel-deconvolution 1024 data.ms
```

Crystalball parameters

```
crystalball data.ms
-sm target-sources.txt # clean components list
-w ds9.reg # source region file
-o bright_source_column # output source model column in ms file
-f 2
```

breizorro parameters

```
breizorro --restored-image image_DI.fits
--threshold 5
--outfile image_DI.fits.mask.fits
```