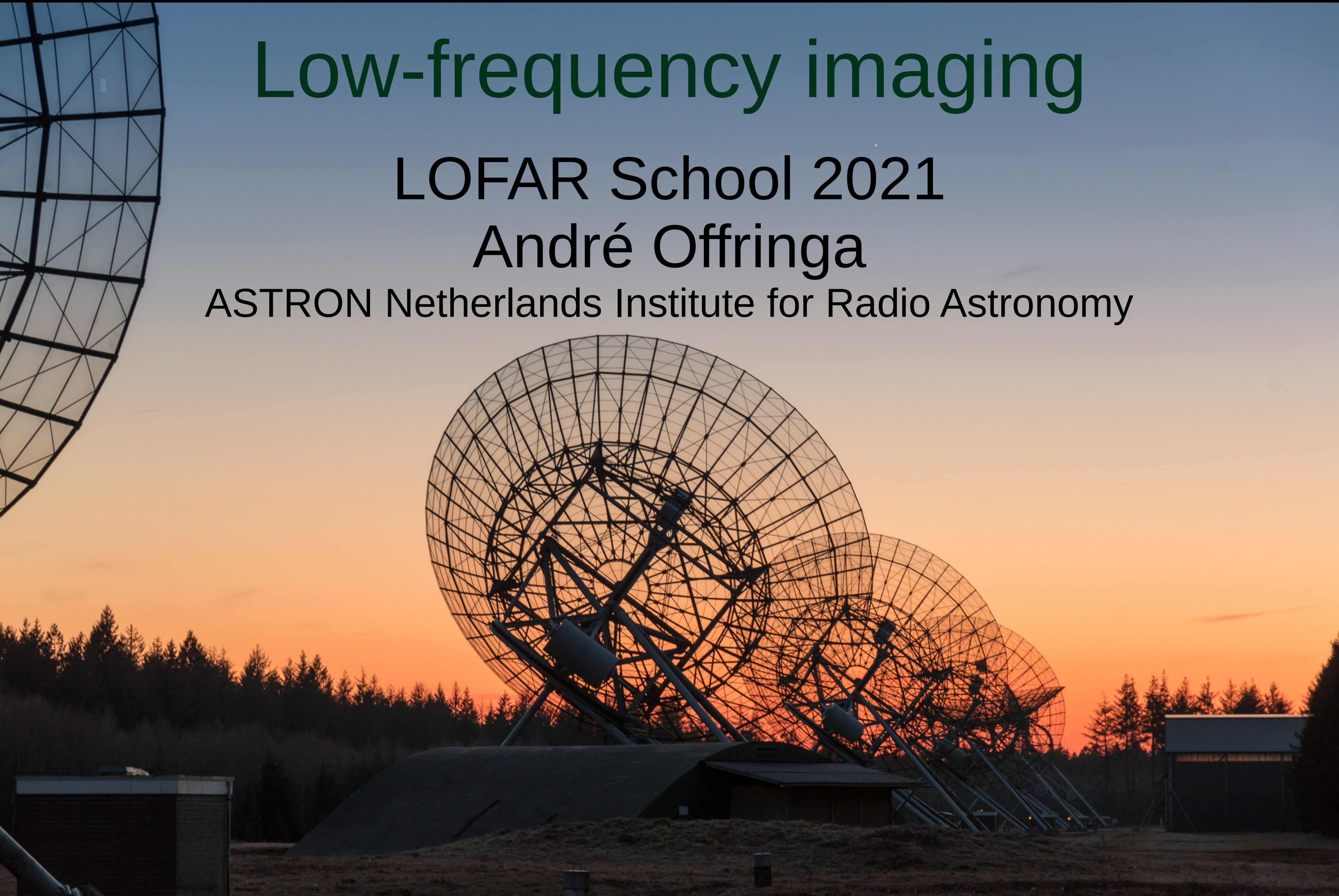
Low-frequency imaging



LOFAR School 2021

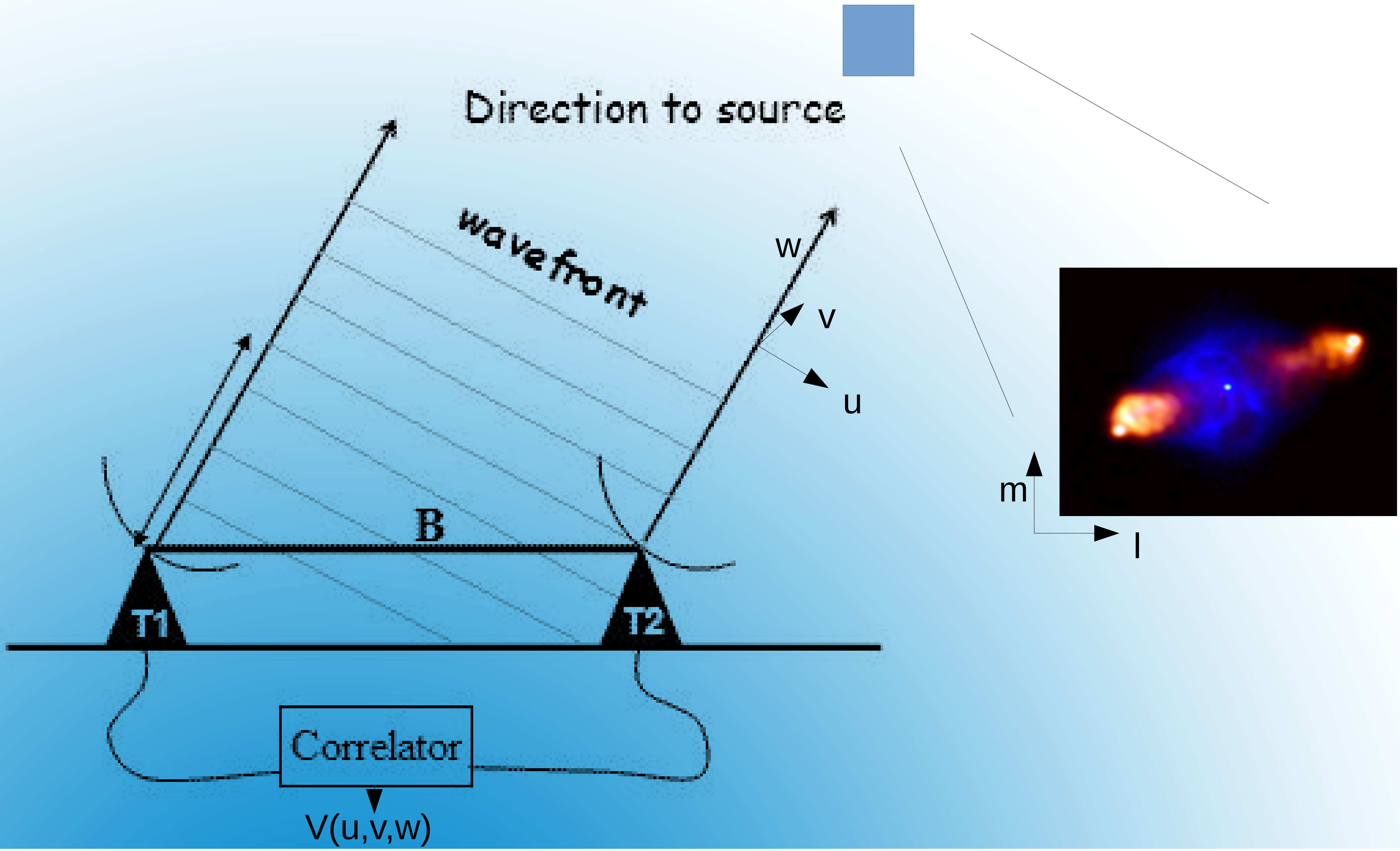
André Offringa
ASTRON Netherlands Institute for Radio Astronomy

Kahoot

- If time permits:

  Kahoot at the end of the lecture

- Nr: 7071233



Coordinates

- Output of an interferometer after calibration:

$$V(u,v,w) = \iint \frac{I(l,m)}{\sqrt{1-l^2-m^2}} e^{-2\pi i\left(ul+vm+w(\sqrt{1-l^2-m^2}-1)\right)} dl dm$$

- (u,v,w) : interferometer's geometrical vector
- (l,m) : position on the sky
- I : sky brightness (“image”)

Imaging : Calculating I(l,m) from V(u,v,w)

Visibility function

- Full visibility function:

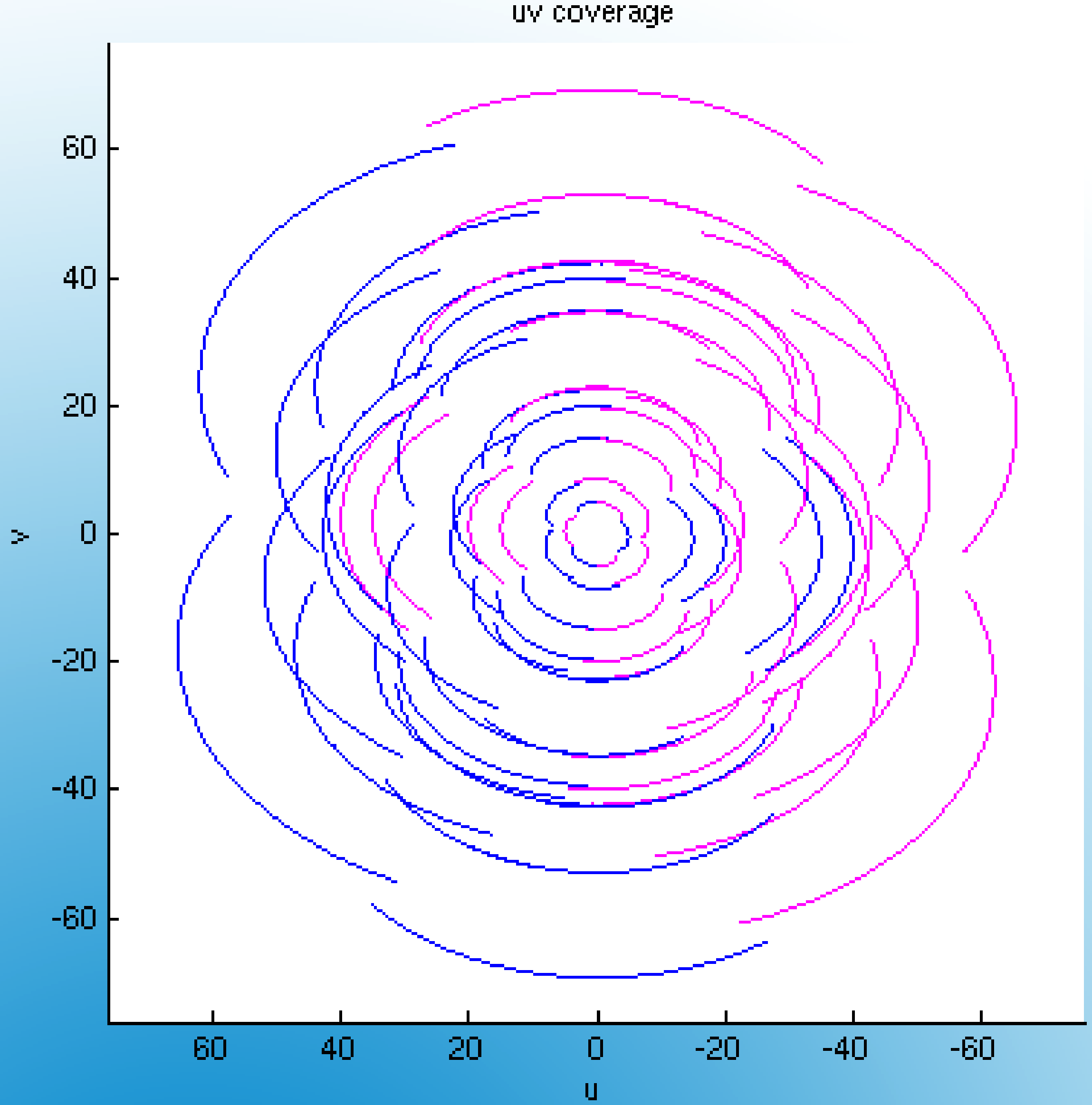
$$V(u,v,w) = \iint \frac{I(l,m)}{\sqrt{1-l^2-m^2}} e^{-2\pi i\left(ul+vm+w(\sqrt{1-l^2-m^2}-1)\right)} dldm$$

- For small field of view (l~0, m~0) or w~0 :

$$V(u,v,\cancel{w}) \approx \iint I(l,m) e^{-2\pi i(ul+vm)} dldm$$

- (u,v,w) : interferometer's geometrical vector
- (l,m) : position on the sky
- $\boldsymbol{I}$: sky brightness (“image”)

Fourier relation



u,v-coverage: values where V is sampled

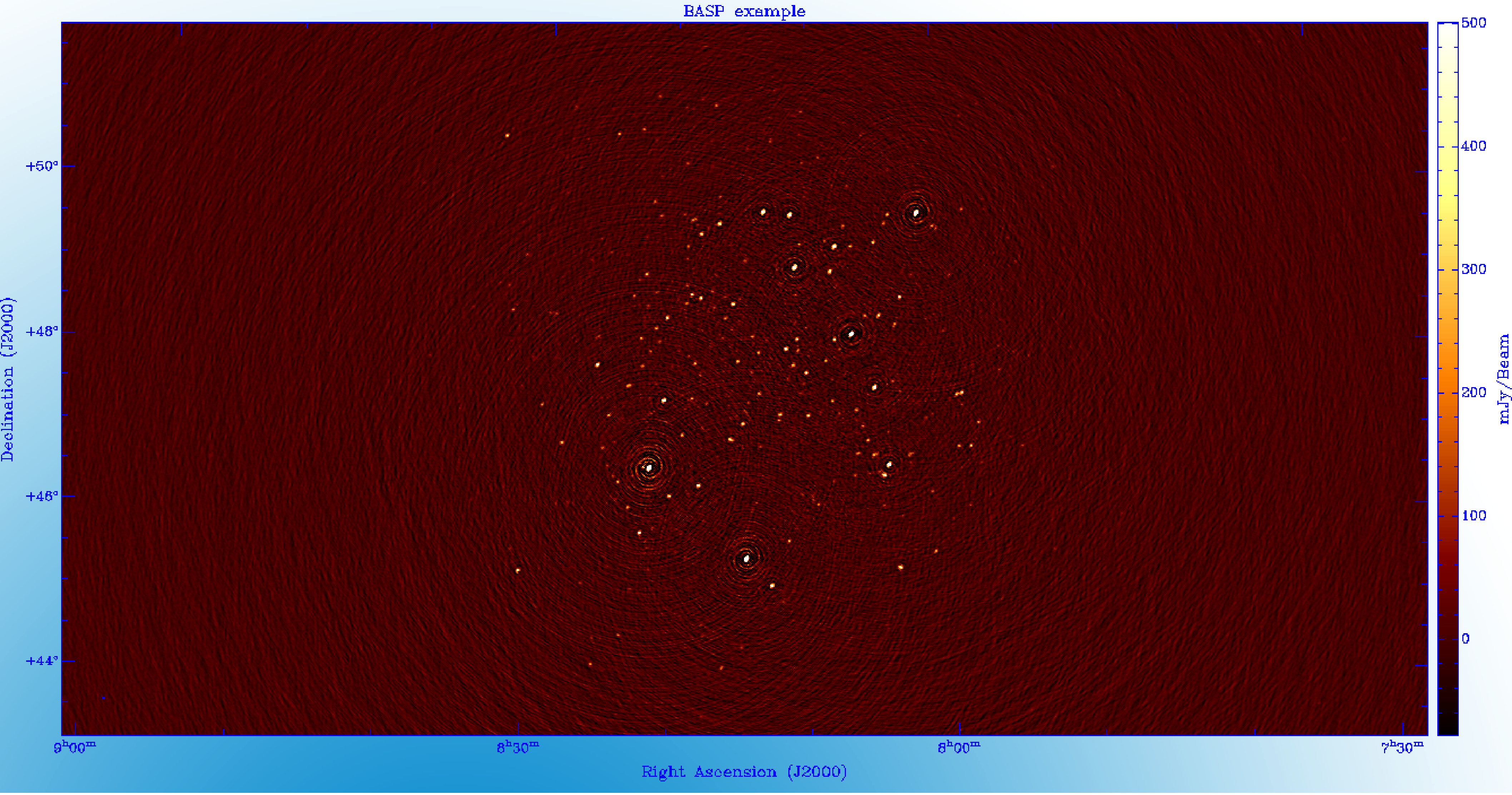
Convolution Theory

Convolution theory says:

$\mathcal{F}$(**A** **B**) = $\mathcal{F}$(**A**) ⊗ $\mathcal{F}$(**B**)
if **A** is the uv coverage (multiplication by ones and zero)
and **B** is the true sky visibility

$\mathcal{F}$(**A**) = PSF
$\mathcal{F}$(**B**) = The sky

$\mathcal{F}$(A) ⊗ $\mathcal{F}$(B) is the "dirty image"
(sky convolved with the PSF)



LOFAR dirty image (3c196)

The dirty image

- Visibility function:

$$V(u,v,w) = \iint \frac{I(l,m)}{\sqrt{1-l^2-m^2}} e^{-2\pi i\left(ul+vm+w(\sqrt{1-l^2-m^2}-1)\right)} dl dm$$

- W-projection: *(Cornwell et al, 2008)*

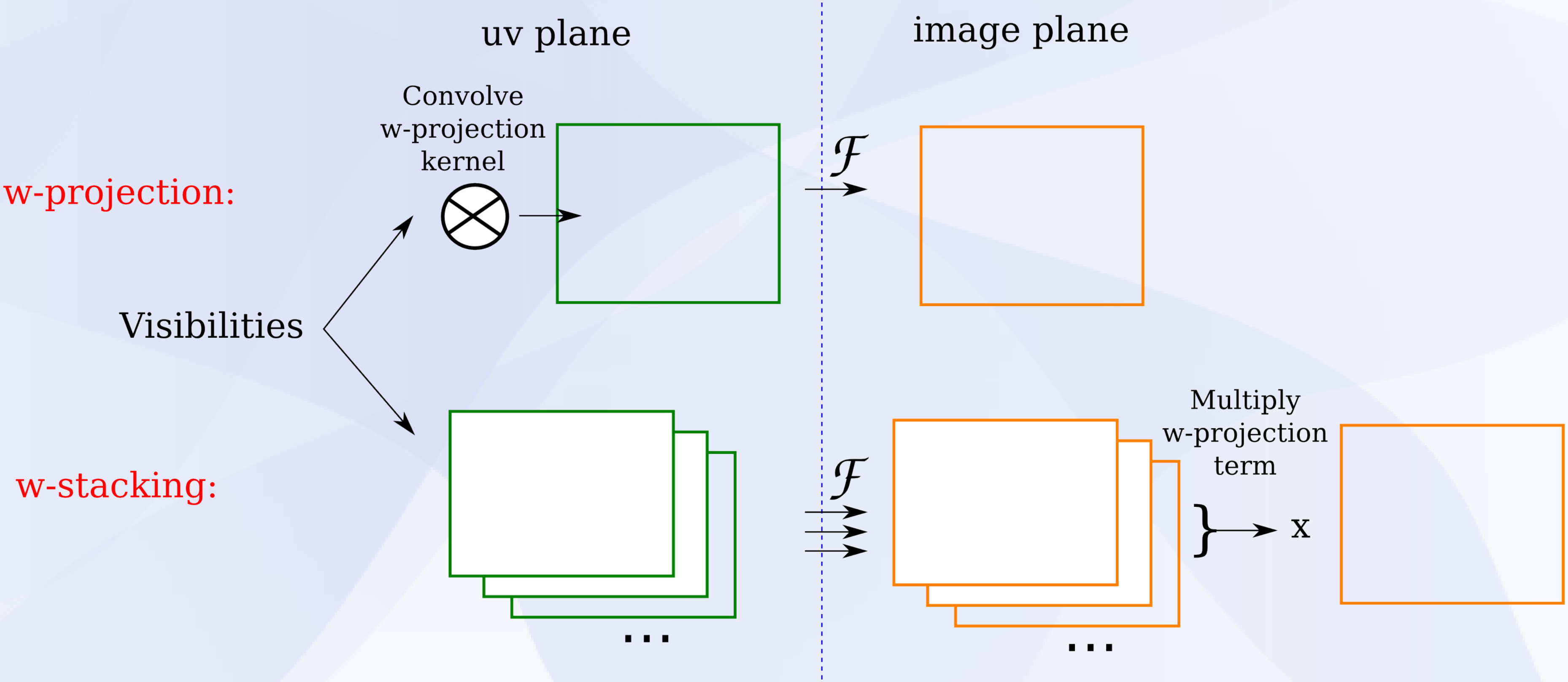
$$V(u,v,w) * \mathcal{F}(e^{-2\pi i w(\sqrt{1-l^2-m^2}-1)}) = \iint \frac{I(l,m)}{\sqrt{1-l^2-m^2}} e^{-2\pi i(ul+vm)} dl dm$$

This convolution turns out
to have a "limited" support

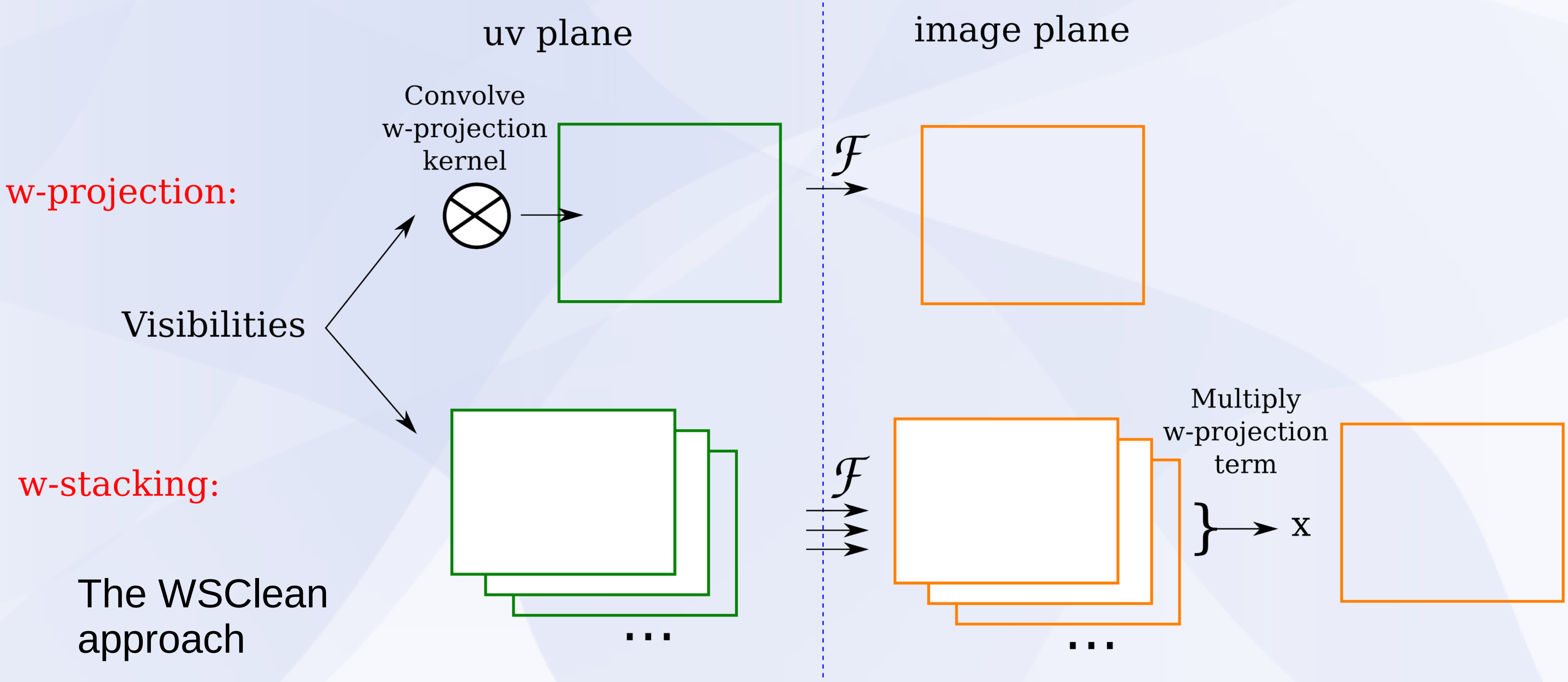
- Performance very dependent on zenith angle, coplanarity of array, field of view and resolution.

w-projection

w-stacking


uv plane
image plane
Convolve
w-projection
kernel
w-projection:
$\mathcal{F}$
Visibilities
Multiply
w-projection
term
w-stacking:
$\mathcal{F}$
x
...
...

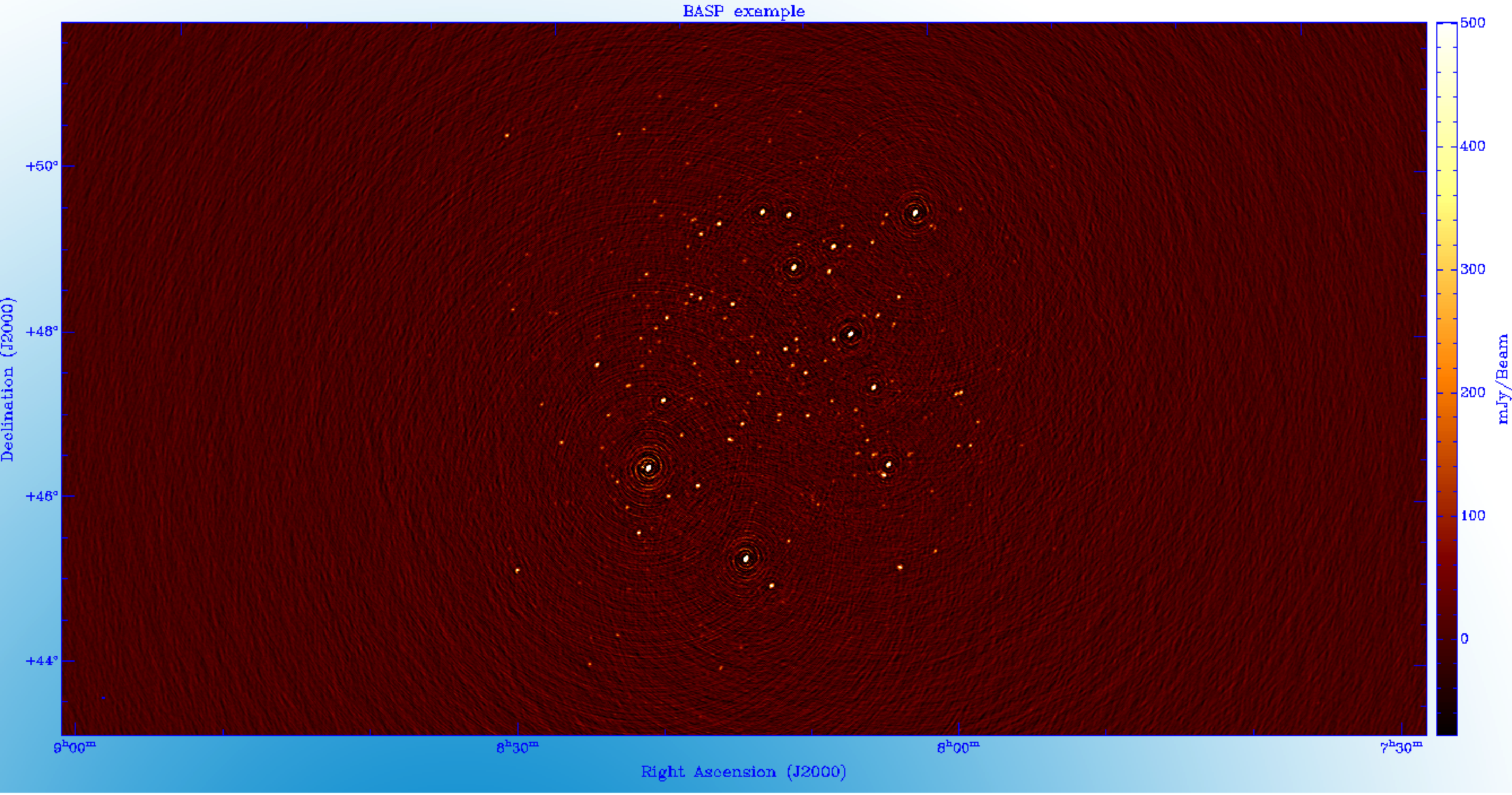
w-stacking



The WSClean approach

WSClean

- WSClean is a generic imager
- Full support for LOFAR
- Integrated in pipelines (Rapthor, Factor)
- When do you run WSClean manually? Eg:
  - For redoing a pipeline’s imaging
  - To self-cal on a VLBI source
  - For inspection
- Docs: https://wsclean.readthedocs.io

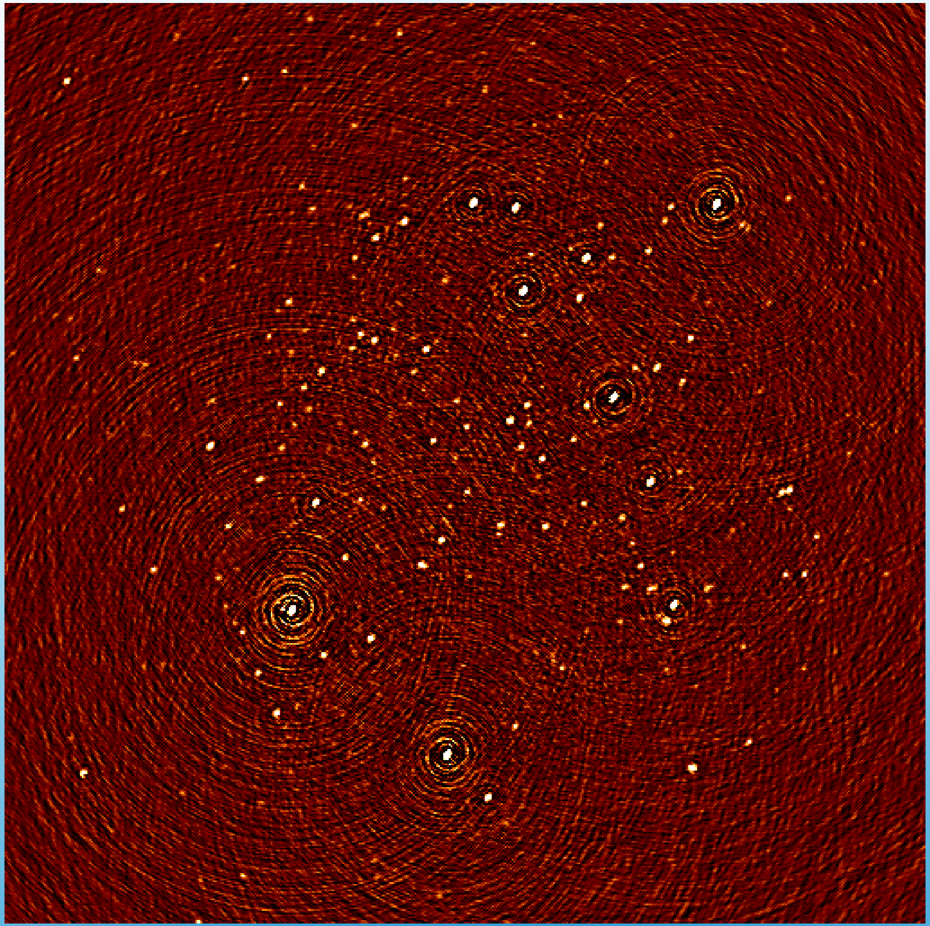


LOFAR dirty image (3c196)

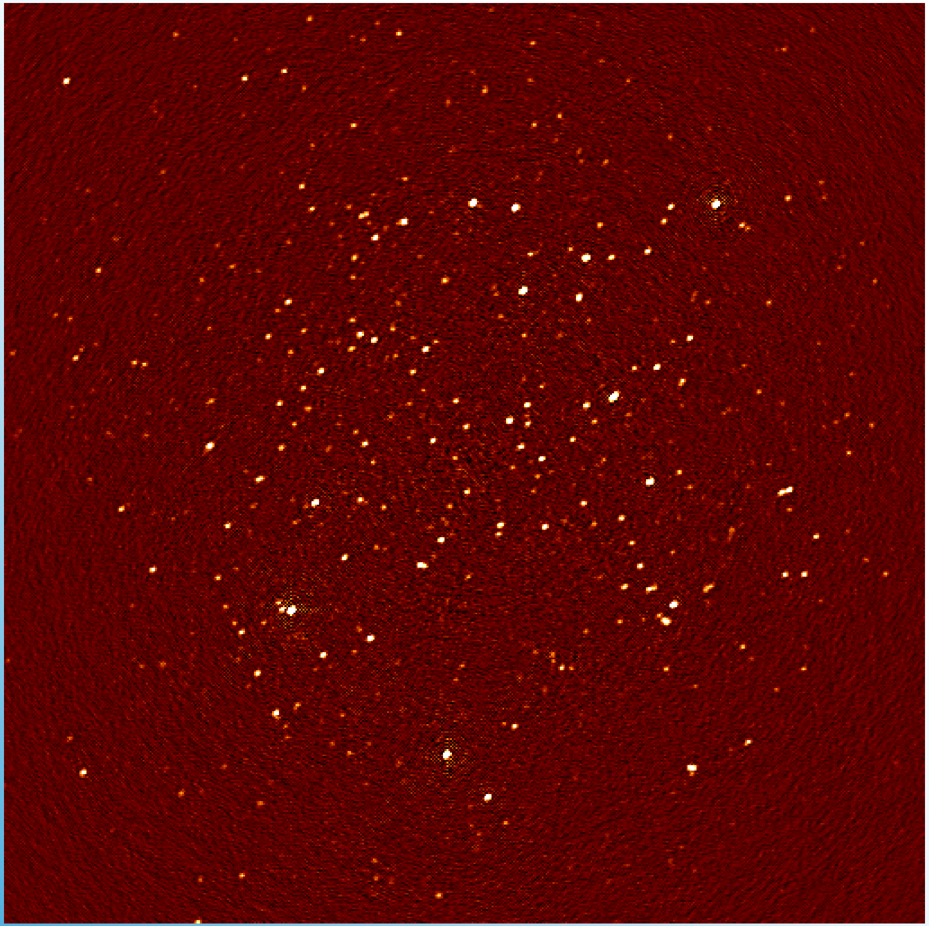
The dirty image

Högbom CLEAN

- Högbom CLEAN algorithm (1974):
  - Find largest peak in image
  - Scale PSF to fraction of peak and subtract
  - Repeat until peak < threshold or nIter > limit
  - Finally: restore subtracted components



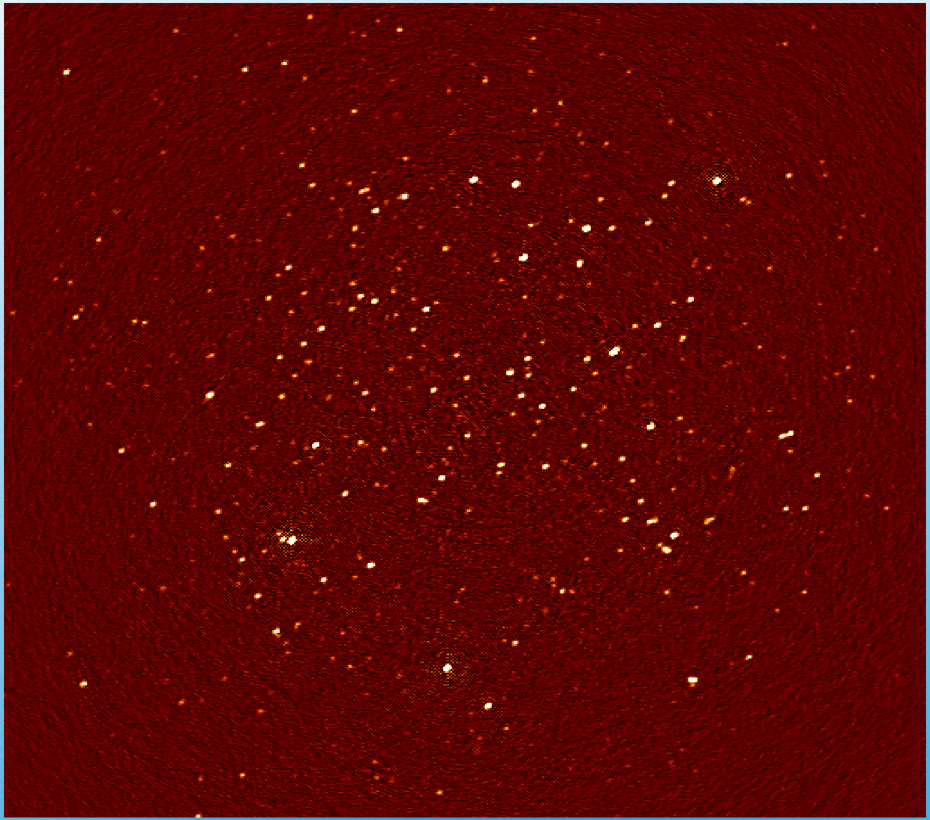
LOFAR undeconvolved (“dirty”) image



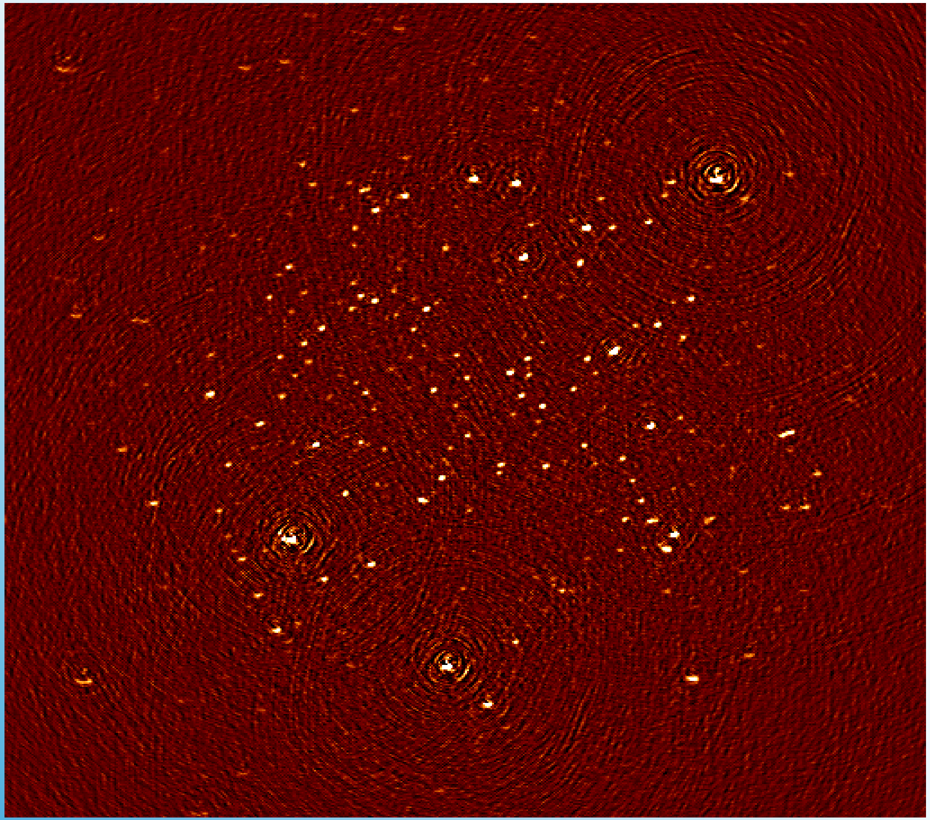
Deconvolved with Högbom CLEAN (“restored” image)

Högbom CLEAN

- 2D FT does not hold for LOFAR: l,m,w >> 0



Correcting w-terms



Without correcting w-terms

The w-term

- Another problem; convolution theorem no longer works when *w*-terms present in

$$V(u,v,w) = \iint \frac{I(l,m)}{\sqrt{1-l^2-m^2}} e^{-2\pi i\left(ul+vm+w(\sqrt{1-l^2-m^2}-1)\right)} dl dm$$

- Högbom CLEAN assumes spatially constant PSF
- But PSF changes (slightly) over the image
- Solved with Cotton-Schwab algorithm (schwab 1984)
- Most imagers will automatically use CS
  - WSClean will do so when `mgain<1`

w-projection

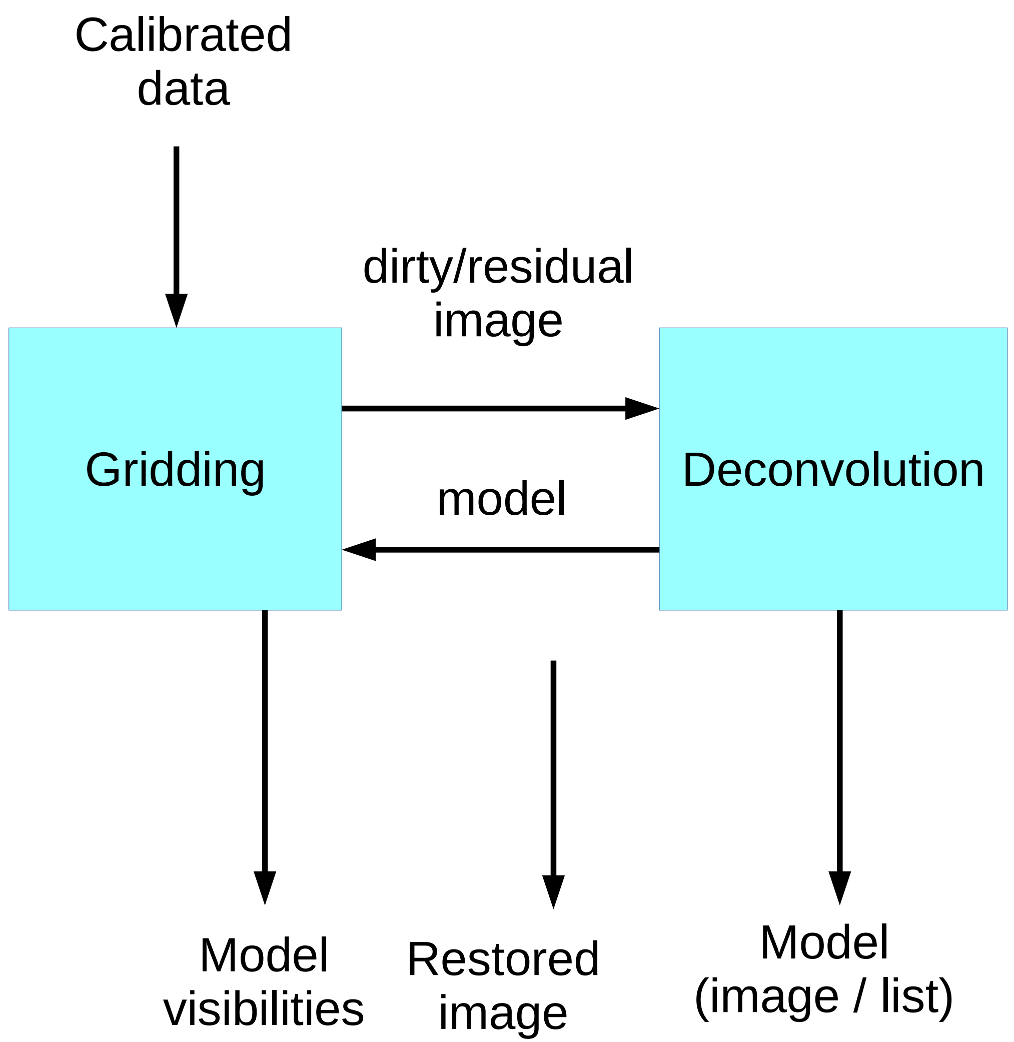
Imaging

Direct FT

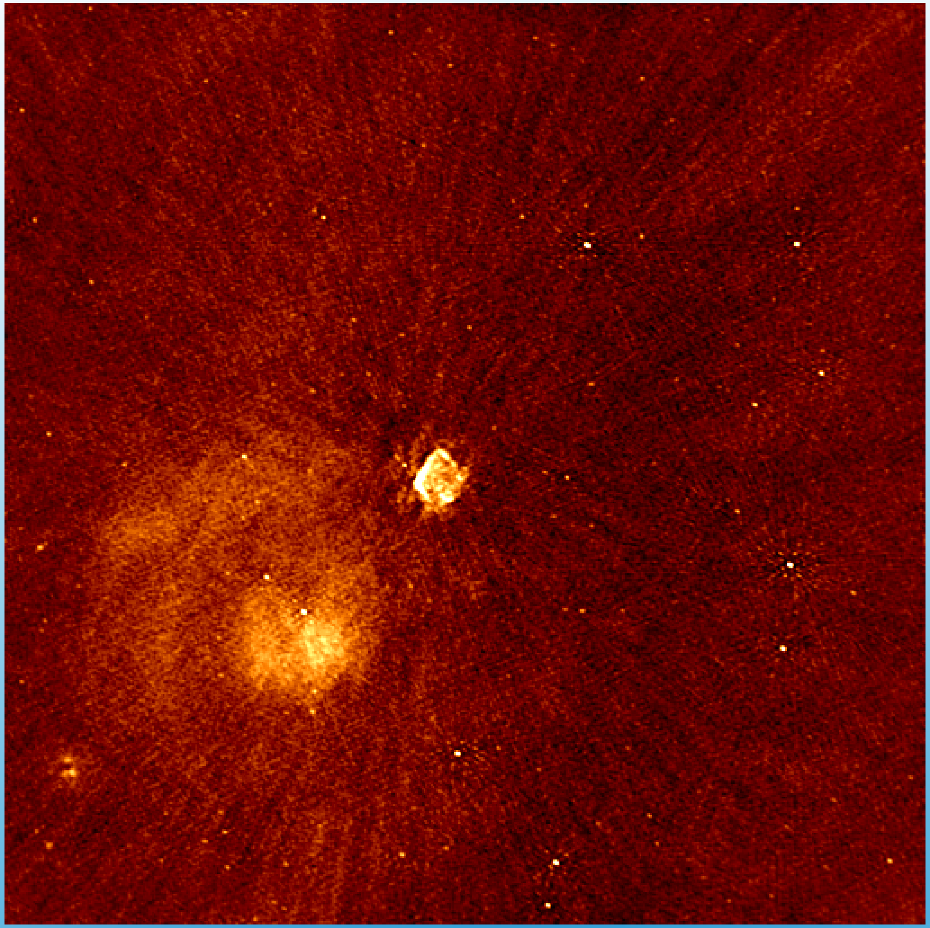
W-stacking
Standard, simple gridding

W-gridder
Written by Martin Reinecke
More accurate
Faster for some cases

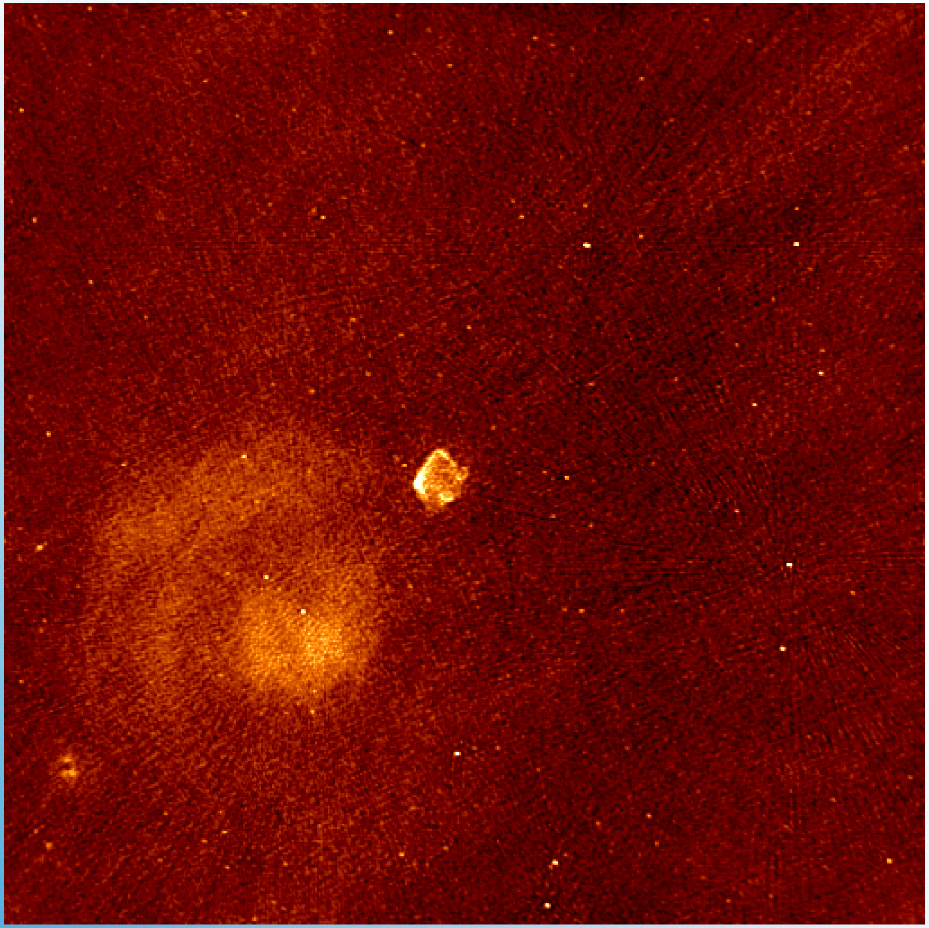
Image Domain Gridding (IDG)
Use GPUs
Grid with beam
  (Also allows mosaicking)
Grid with ionospheric
  corrections (aterms)
Superior accuracy



WSClean’s algos:
Multi-scale
Multi-frequency
Auto-masking
Component list
Local RMS cleaning
Parallel
  deconvolution

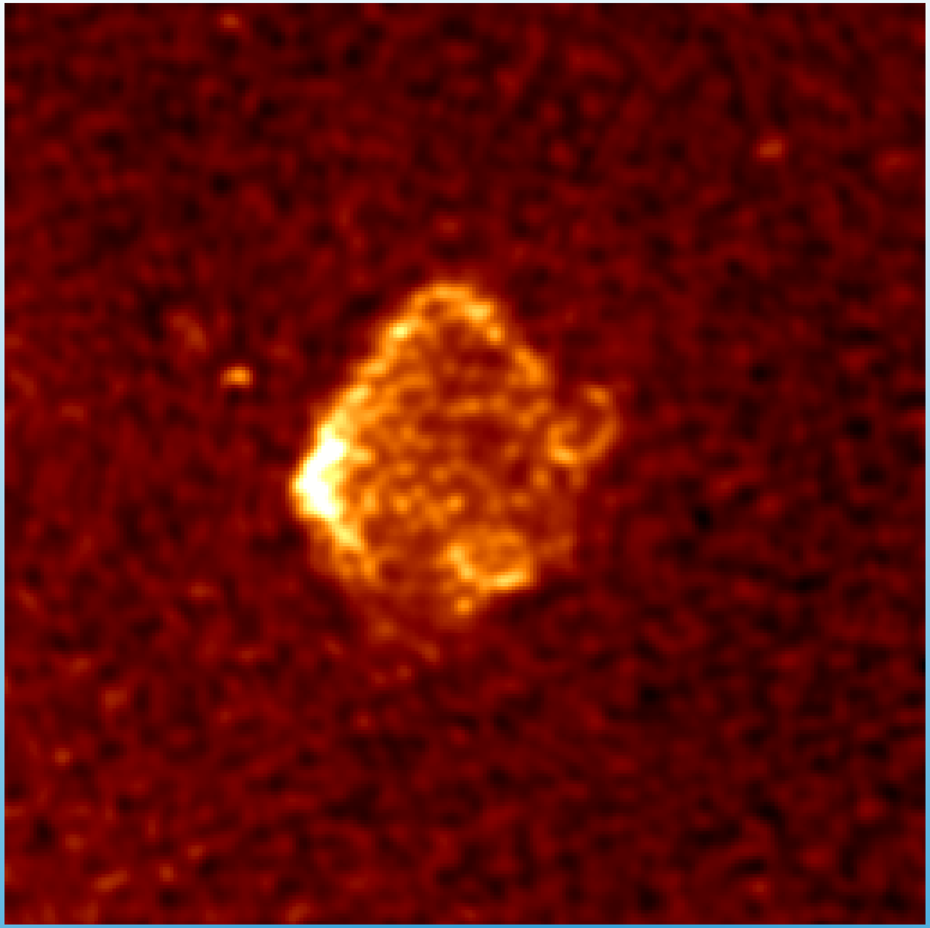


Undeconvolved “dirty” image

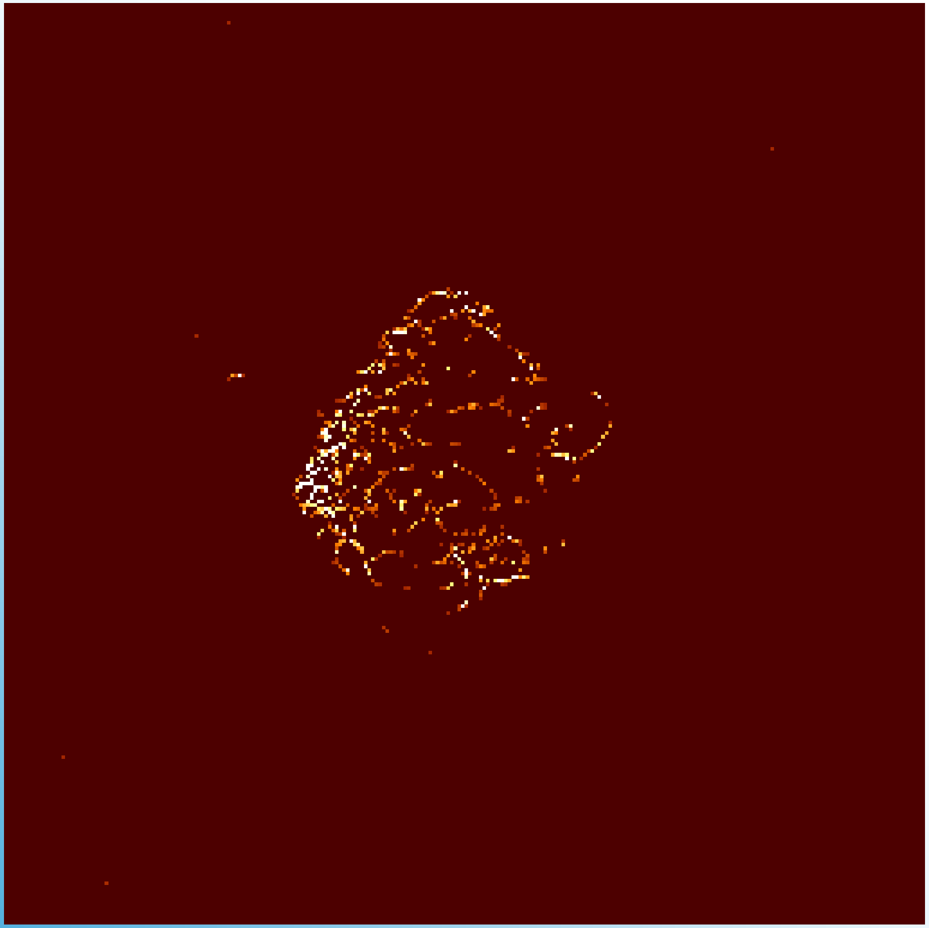


Deconvolved image with Högbom CLEAN

Deconvolving diffuse structures

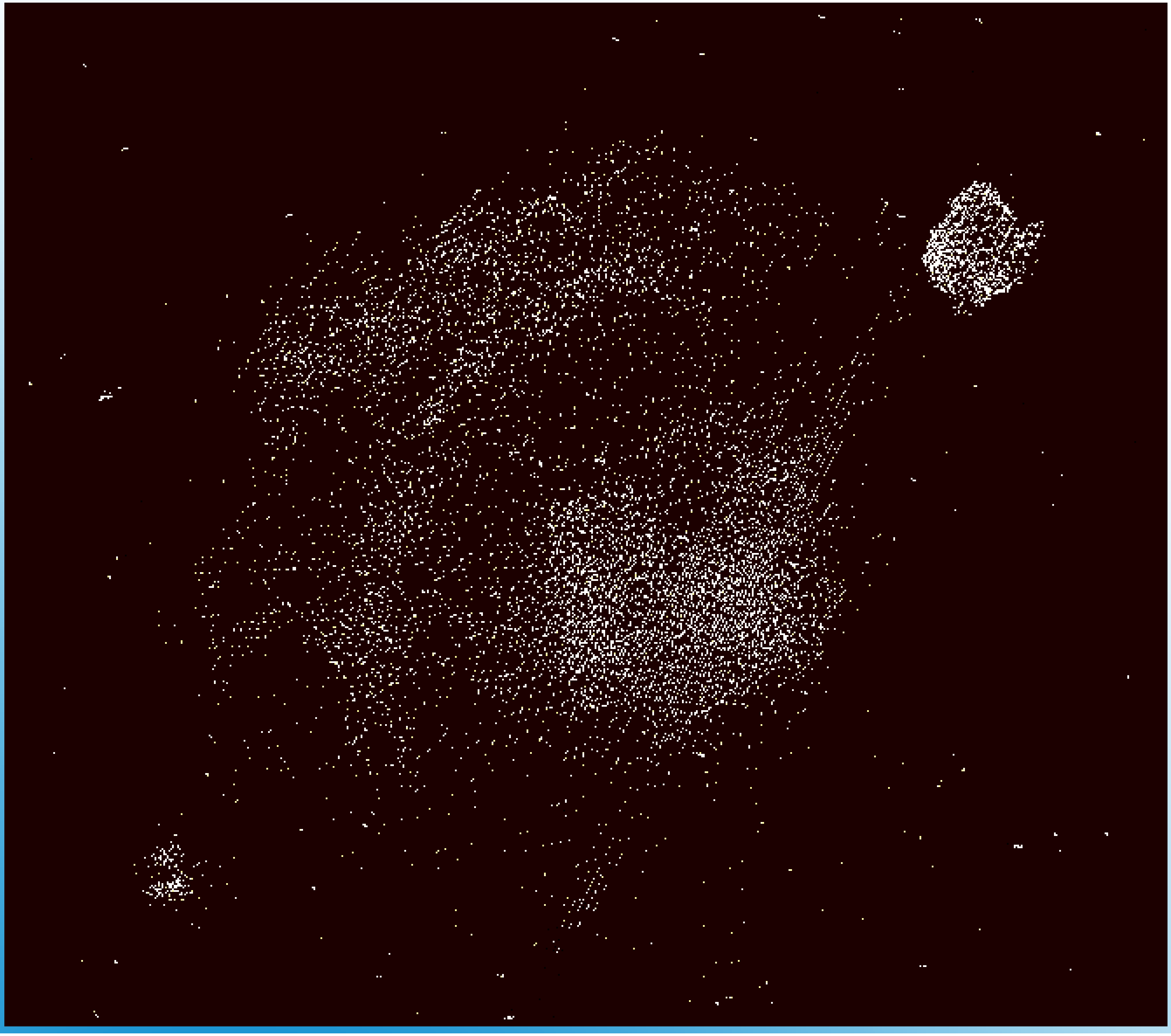


Deconvolved image (Högbom CLEAN)



Högbom CLEAN constructed model

Deconvolving diffuse structures



Model created by Högbom clean

Improved algorithm by Cornwell (2008) :

- “Multi-scale clean”

- Fits small smooth kernels (and delta functions) during a Högbom CLEAN iteration

Multi-scale CLEAN

Multi-scale kernel

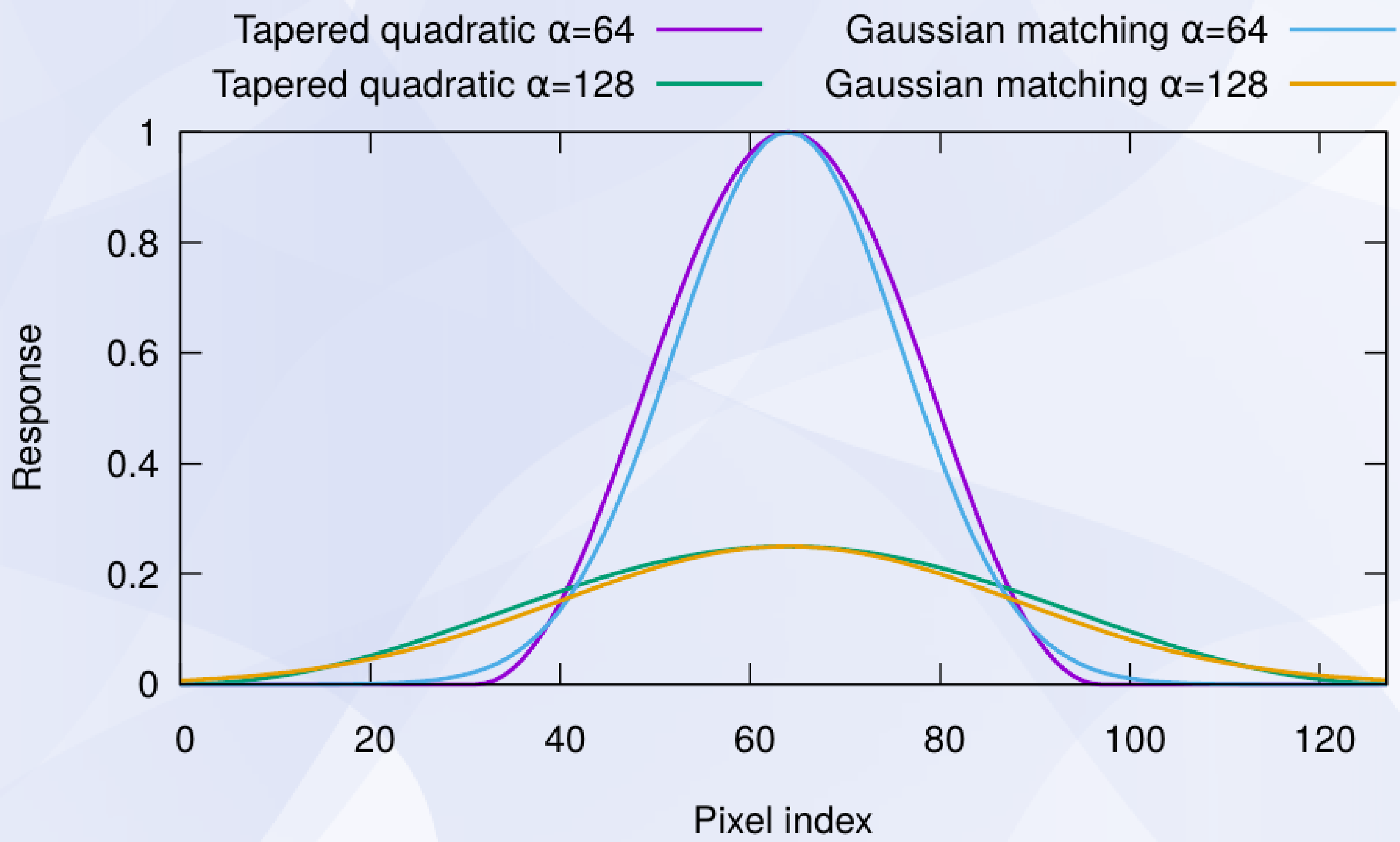


Figure 1. Shape functions for scales $\alpha = 64$ pixels and $\alpha = 128$ pixels.

Fast multi-scale deconvolution

- In Cornwell’s (2008) multi-scale method, the appropriate scale is determined every minor iteration
- This algorithm can be made a lot faster by keeping the scale fixed for a while
- This is the algorithm implemented in WSClean
  - Parameter: `−multiscale`

Fast multi-scale deconvolution

A multi-scale deconvolution iteration in WSClean:

- Convolve the images with several scales
- Determine the scale with the most significant peak
- Prepare PSF and enter a “subminor” loop:
  - Find peak in scale-convolved image
  - Subtract (double) scale-convolved PSF
  - Iterate until start peak < (1-gain) x new peak

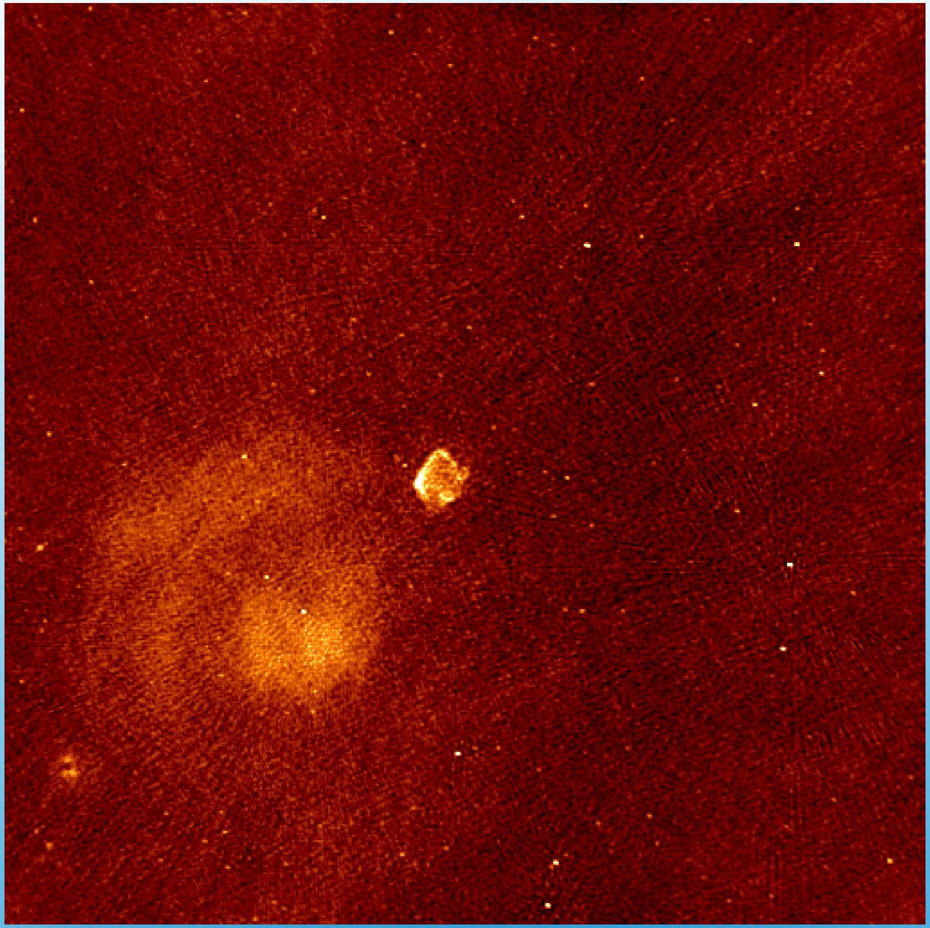
Fast multi-scale deconvolution

A multi-scale deconvolution iteration in WSClean:

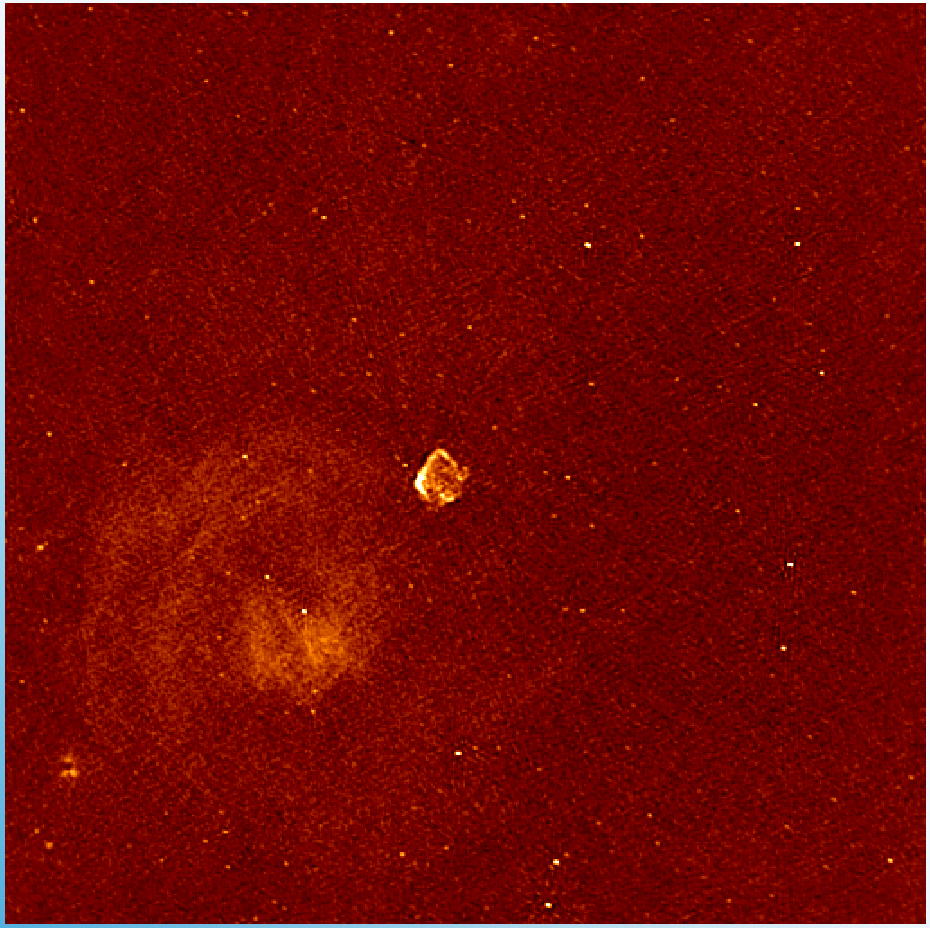
- Convolve the images with several scales

These iterations are as fast as a regular clean minor iteration

- Prepare PSF and enter a "subminor" loop:
  - Find peak in scale-convolved image
  - Subtract (double) scale-convolved PSF
  - Iterate until start peak < (1-gain) x new peak

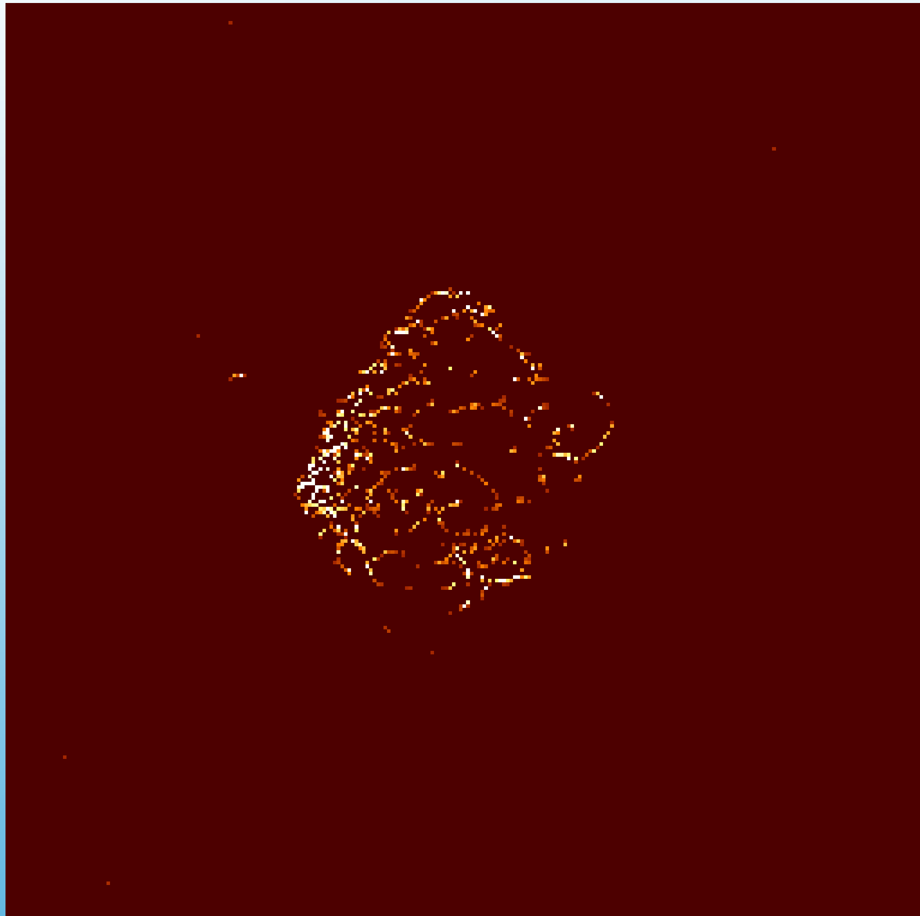


Normal Högbom CLEAN

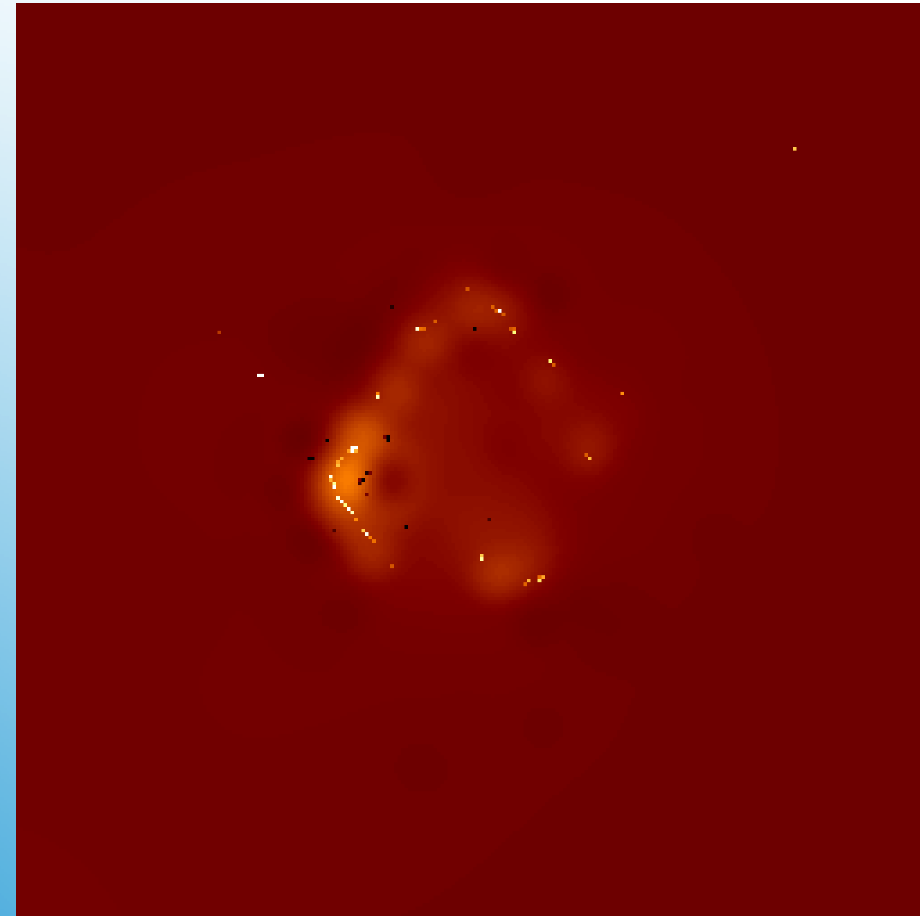


Multi-scale CLEAN
(as implemented in WSClean)

Multi-scale CLEAN



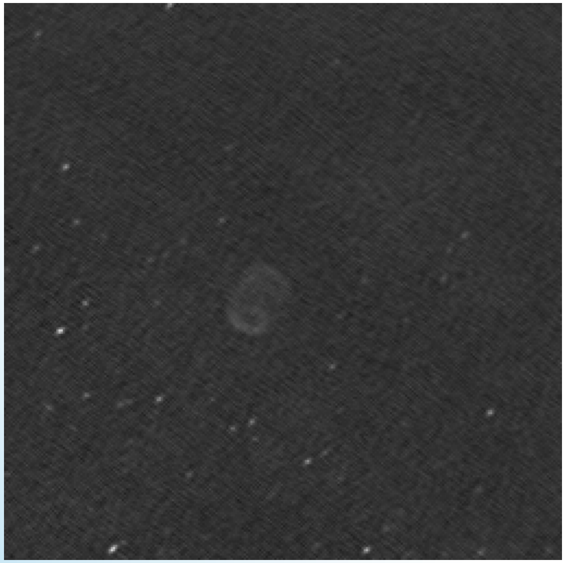
Normal Högbom CLEAN
Output model



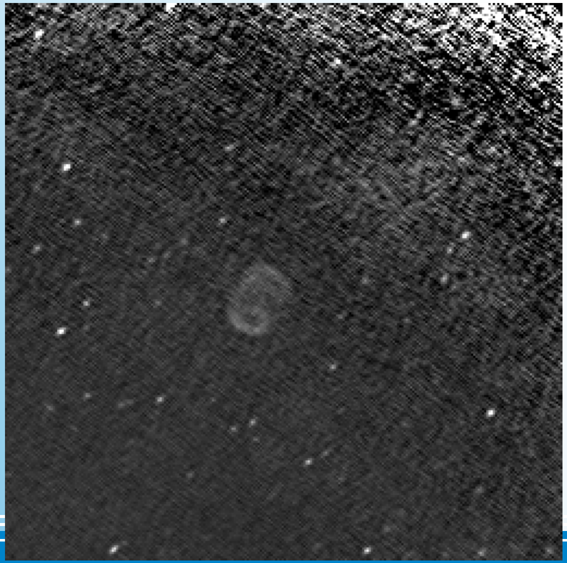
Multi-scale CLEAN
(as implementation in WSClean)

Multi-scale CLEAN

- To get proper flux density values, images need to be corrected for the LOFAR beam.
- WSClean can do this automatically
  - Uses “EveryBeam” library for correction
  - Most simple correction enabled with `-apply-primary-beam`.

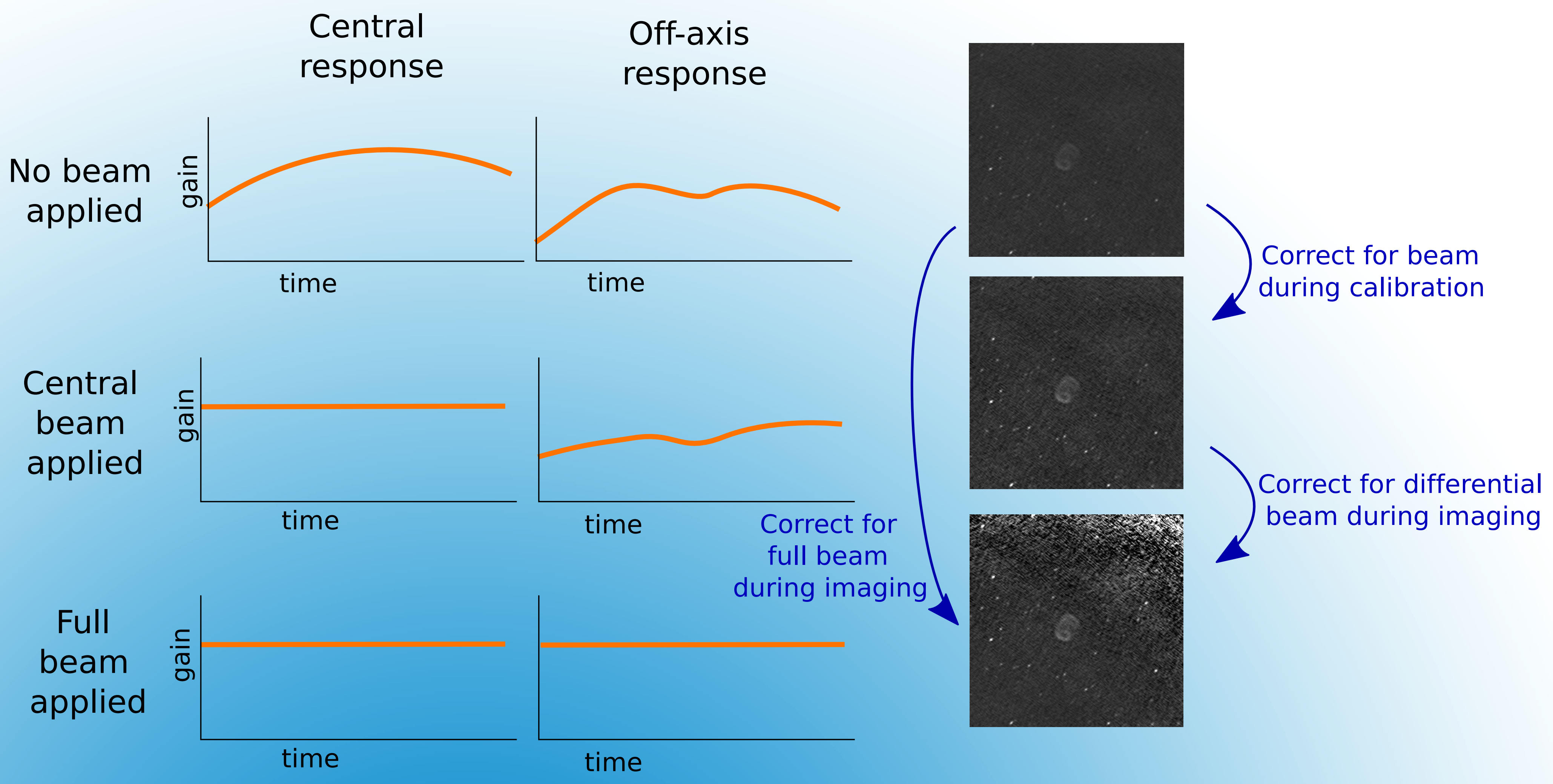


Beam not applied



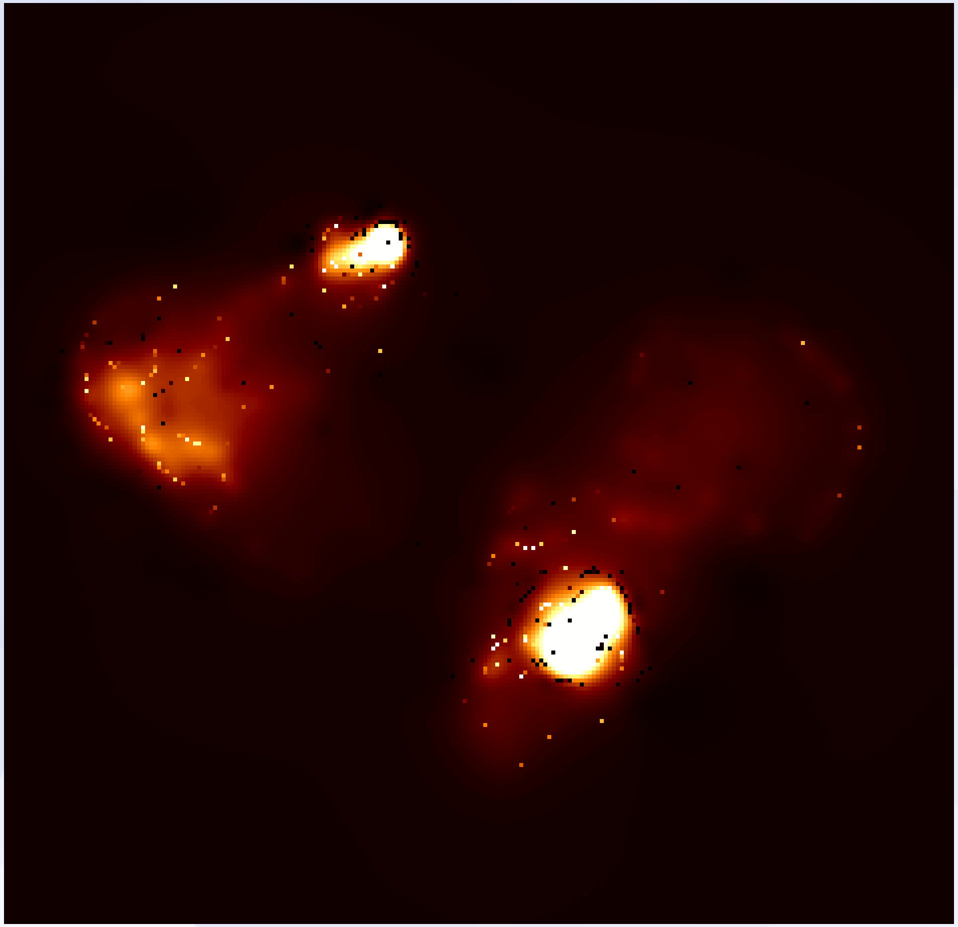
Beam applied

Applying the LOFAR beam



LOFAR beam correction

Multi-scale model



Units

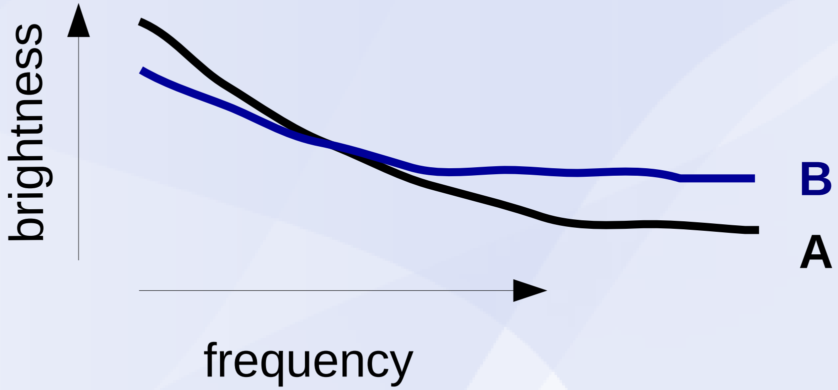
- Visibilities are in units of Jansky (Jy)
- Restored images are in units of Jy/Beam:
  - Flux for a point source is given by its peak flux
  - Flux of a resolved source = spatial integration & dividing out the beam
  - Conversion to surface brightness:

$$T(l,m,\nu) = S(l,m,\nu)\frac{10^{-26}c^2}{2k_B\nu^2\Omega_{\rm psf}}$$

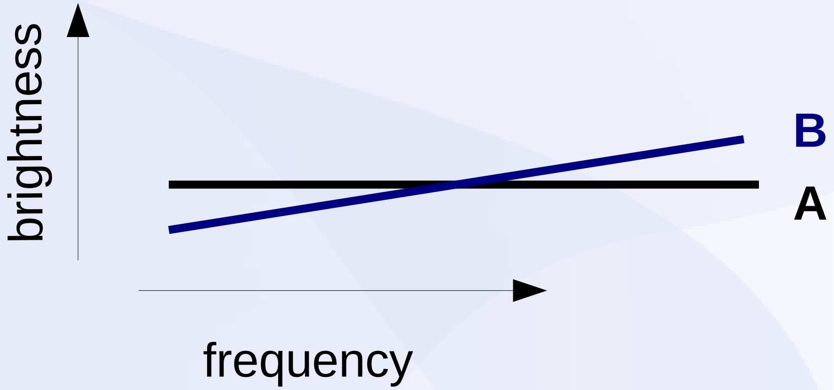
- Model images are in units of Jy/pixel
- Dirty & residual images are in Jy/Beam, but very hard to interpret!

Multi-frequency deconvolution

- Standard clean assumes all sources are flat spectrum sources
-



Without self-cal

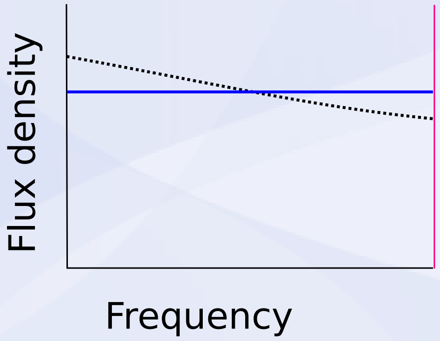


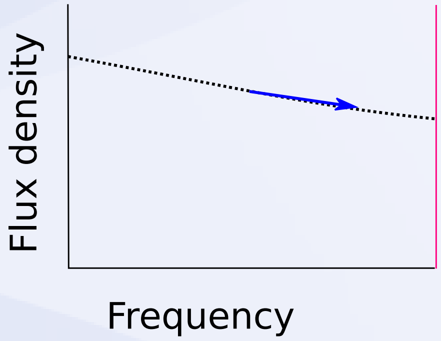
After self-cal on Source A

Multi-frequency deconvolution

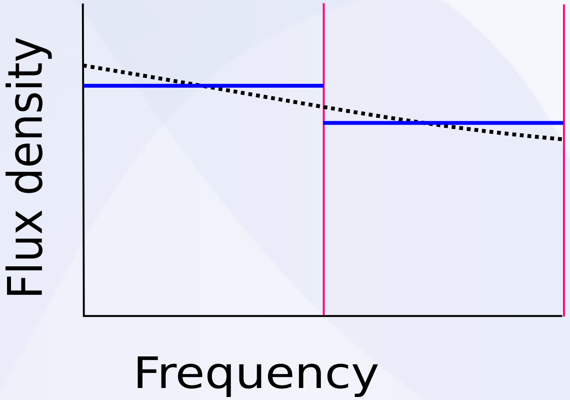
- Common approach in MF deconvolution is imaging / predicting “frequency derivative” images (“nterms>1”, the Sault & Wieringa (1994) method).

That results in:



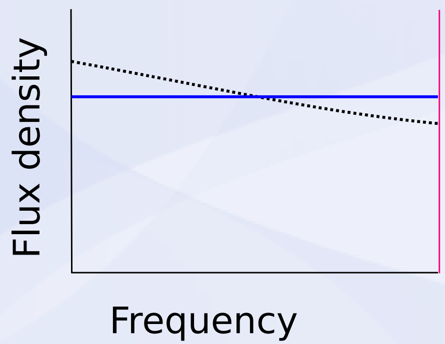


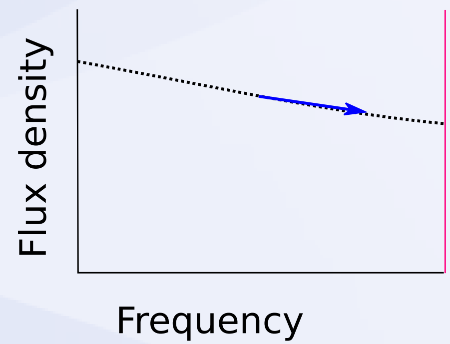
Instead, WSClean splits the bandwidth and creates separate images for each part:



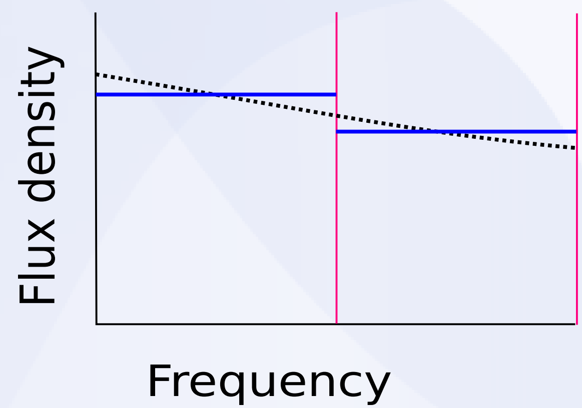
Multi-frequency deconvolution

- Common approach in MF deconvolution is imaging /
derivative" images ("nterms>1", the Sault &





an splits the
and creates
r each part:



- Of course, these contain the same information

  (they can even be converted from one to the other)

- But the second option is easier/more intuitive to clean...

Multi-frequency deconvolution

- Common approach in MF deconvolution is imaging / ... derivative” images (“nterms>1”, the Sault & ...

- Of course, these contain the same information

  (they can even be converted from one to the other)

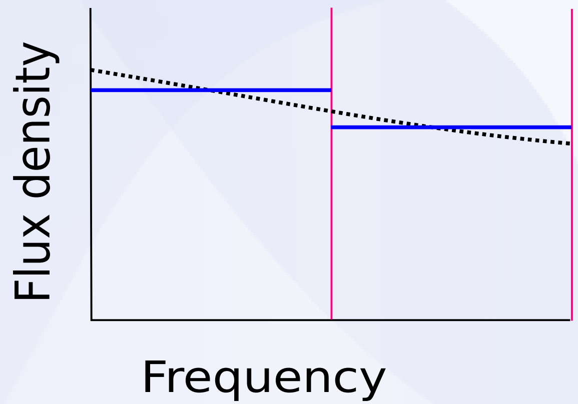
- But the second option is easier/more intuitive to clean...

Flux density

$$S_i = \frac{1}{\xi_i - \nu_i} \int_{\nu=\nu_i}^{\xi_i} c_1 + c_2\nu + c_3\nu^2 + \ldots$$

$$= \sum_{t=1}^{N_{\text{terms}}} c_t \frac{\xi_i^t - \nu_i^t}{t\left(\xi_i - \nu_i\right)}$$

...an splits the ...and creates ...r each part:



Multi-frequency deconvolution

WSClean’s Multi-frequency clean algorithm: (1 maj iter)

- Make residual images at different frequencies
- Start cleaning:
  - Find a peak in the **integrated** image
  - Measure the flux at this position in the subband images
  - Subtracted the correct PSF from each subband image.
- ...Until major iteration threshold is reached
- (Optionally) convert to Taylor-term images and predict

Multi-frequency deconvolution

WSClean's Multi-frequency clean algorithm: (1 maj iter)

- Make residual images at different frequencies
- Start cleani
  - Find a   ge
  - Meas   the subband ima
  - Subtracted the correct PSF from each subband image.
- ...Until major iteration threshold is reached
- (Optionally) convert to Taylor-term images and predict

This is called
"joined channel cleaning"
in WSClean

Multi-frequency deconvolution

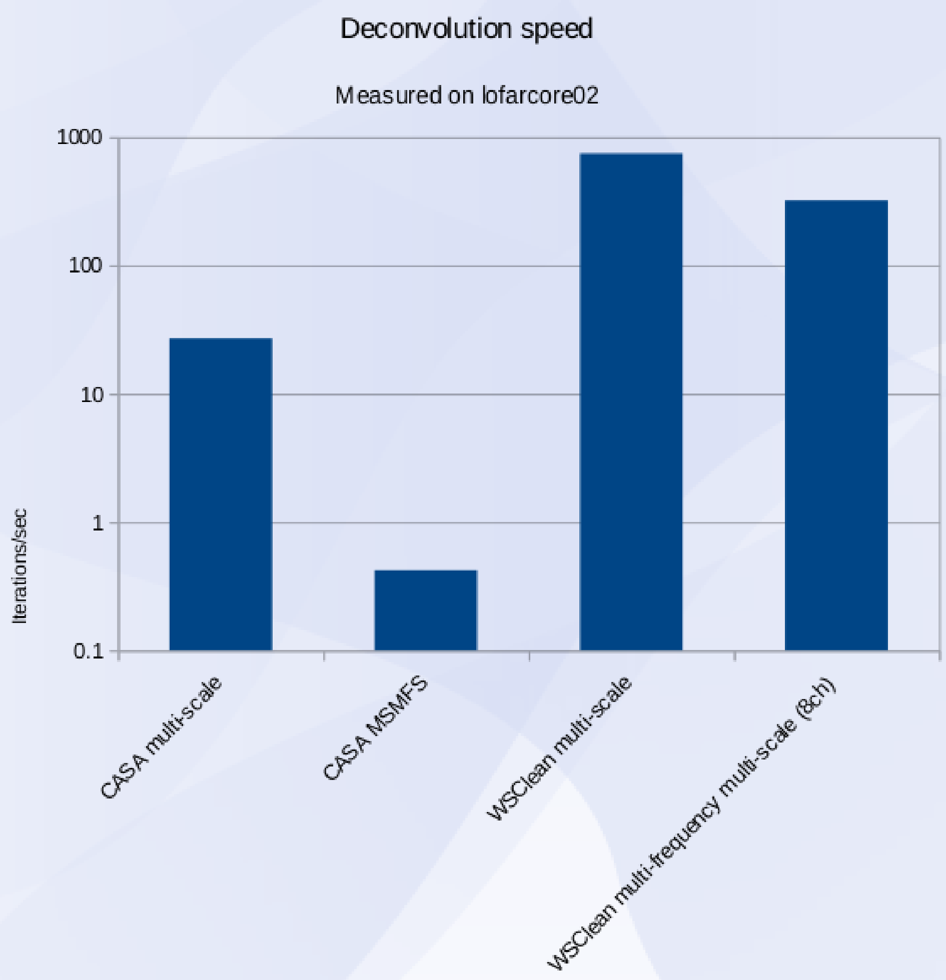
To improve the signal to noise:

- When imaging/predicting many channels, force the components to lie on some spectral function
  (e.g. log-polynomial, polynomial, Zernike-polynomials, etc.)

  (WSClean parameter: `` `-fit-spectral-pol <terms>` ``)

- When imaging Q & U, it is possible to fit a rotation measure during deconvolution.
  - Peak finding on the sums of $Q^2 + U^2$ of all channels

Deconvolution performance



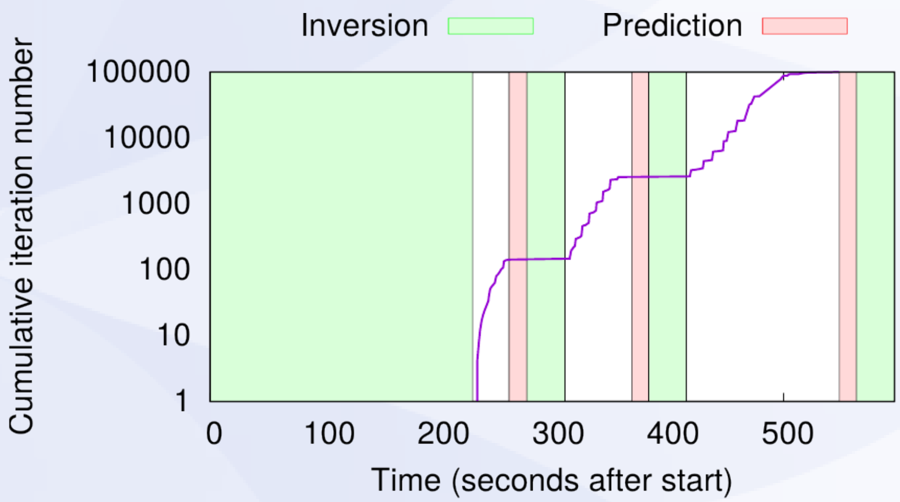
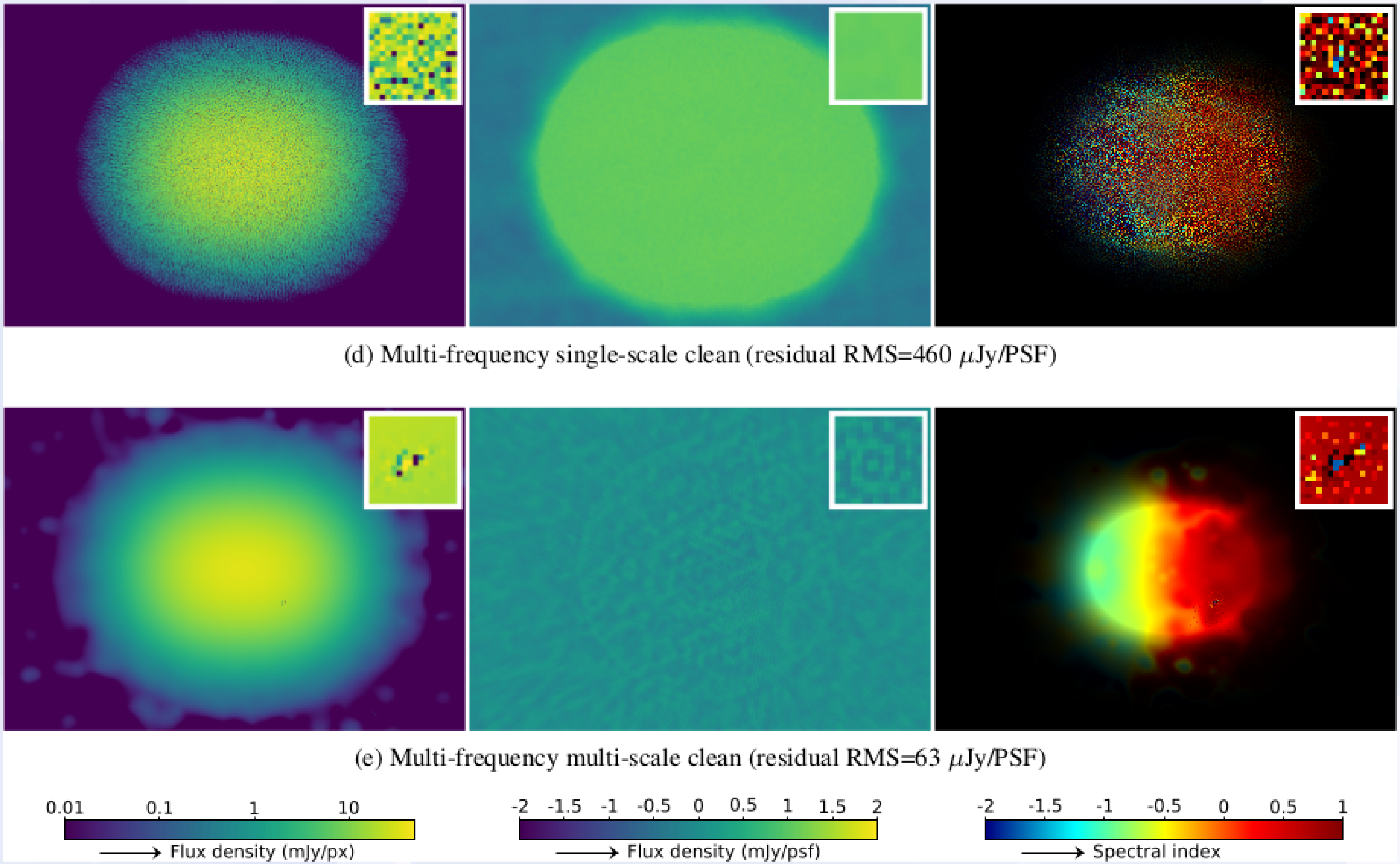


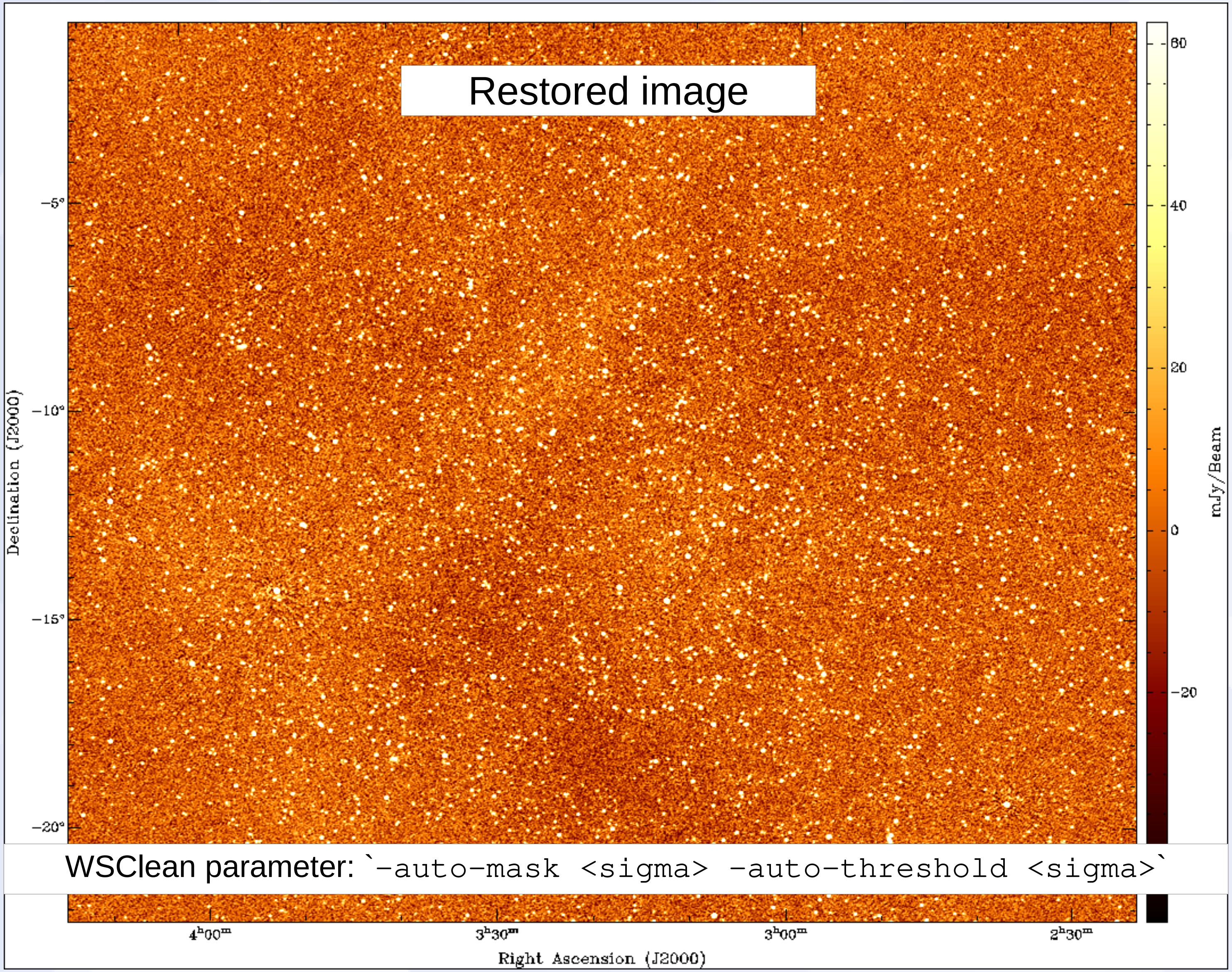
Figure 12. Example of the progression over time when using the new multi-scale clean algorithm on a 2048×2048 image.



- Comparison of WSClean MF single scale and multi-scale cleaning
- Simulated bandwidth of 30 MHz at 150 MHz.
- MWA layout, 2 min snapshot

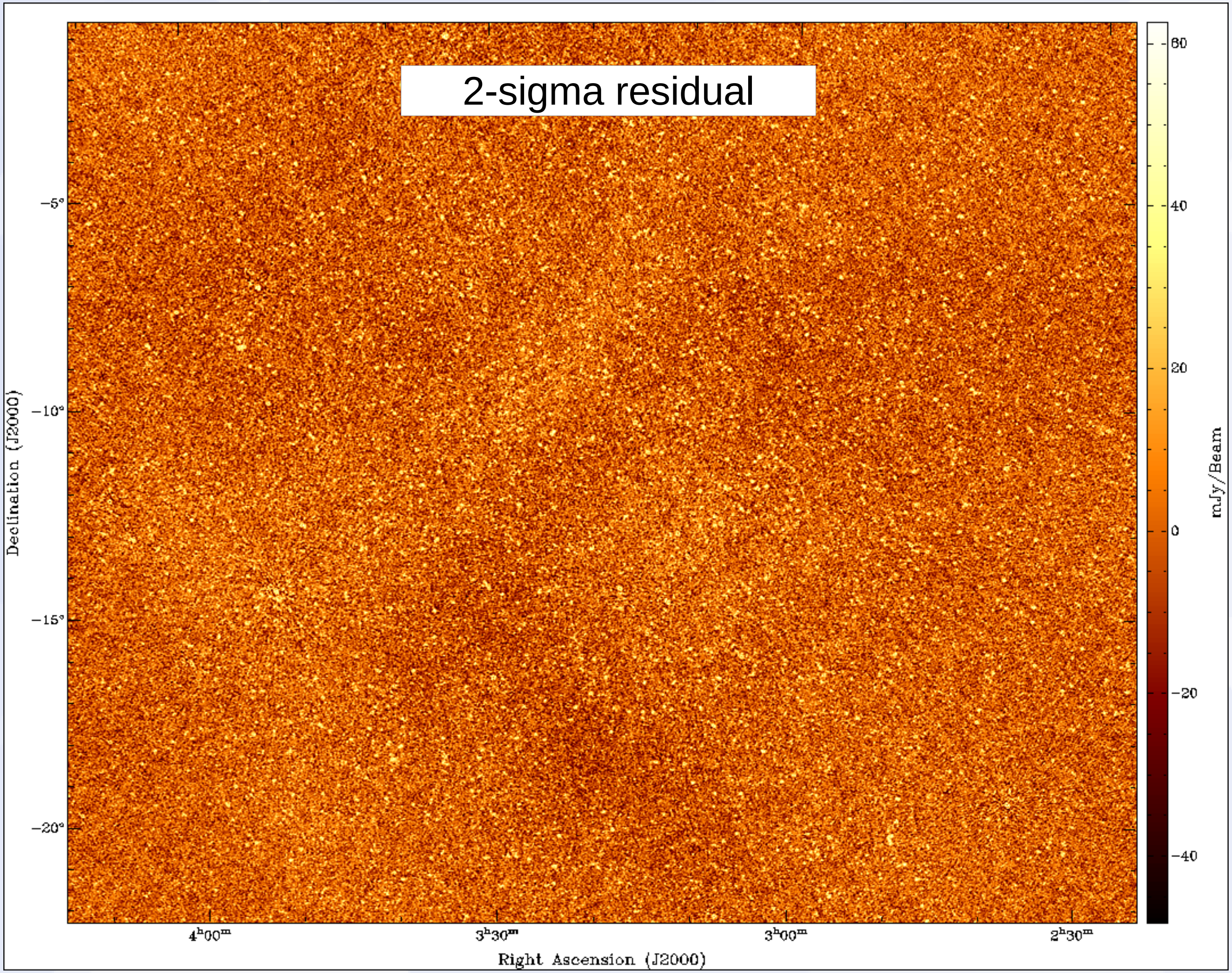
From Offringa & Smirnov (2017)

Auto-masking on point sources

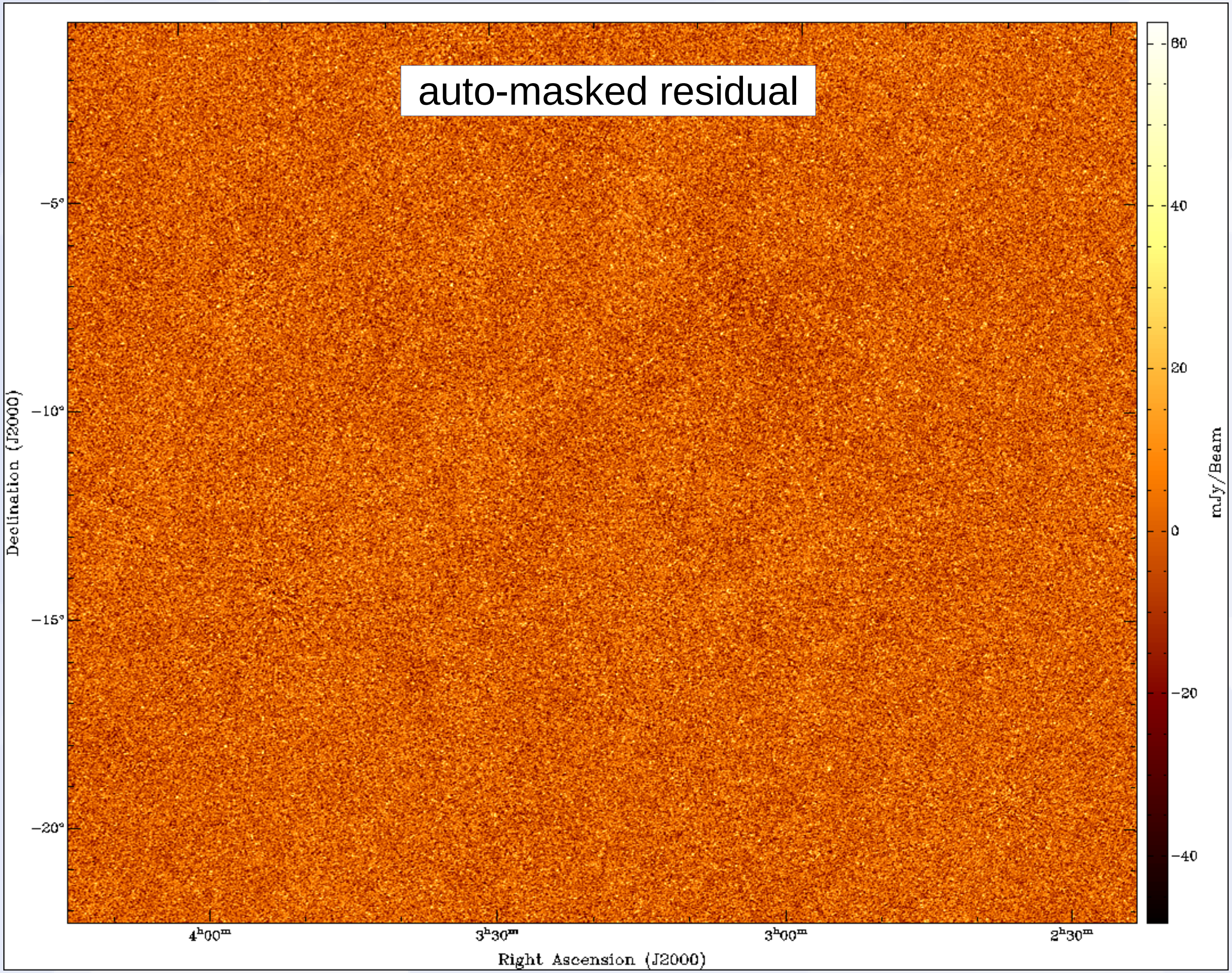


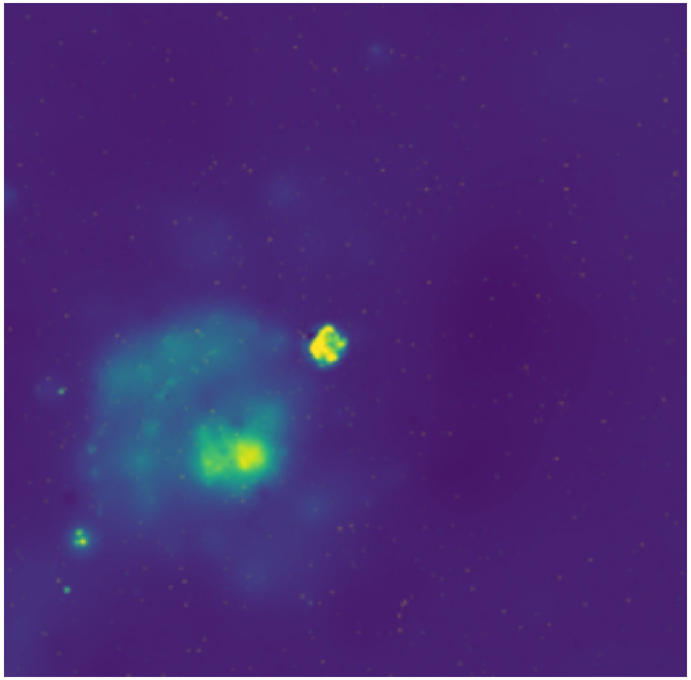
WSClean parameter: `-auto-mask <sigma> -auto-threshold <sigma>`

Auto-masking on point sources

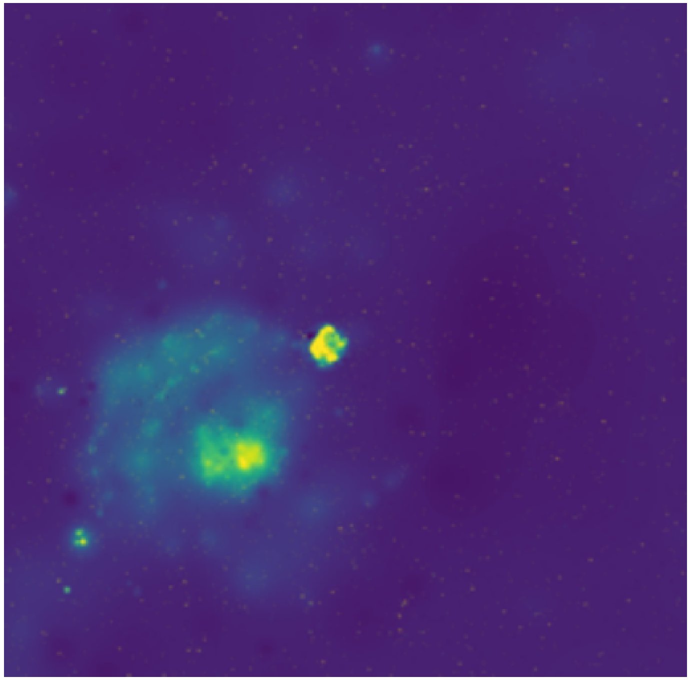


Auto-masking on point sources

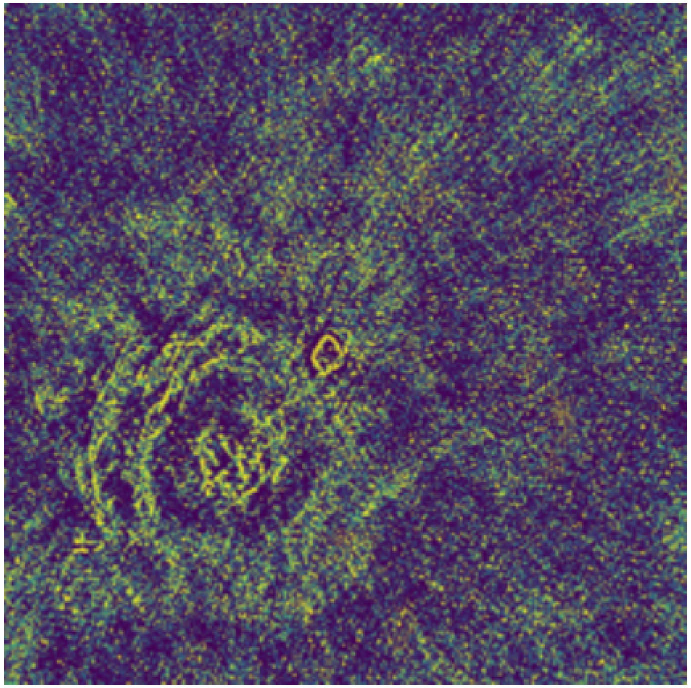




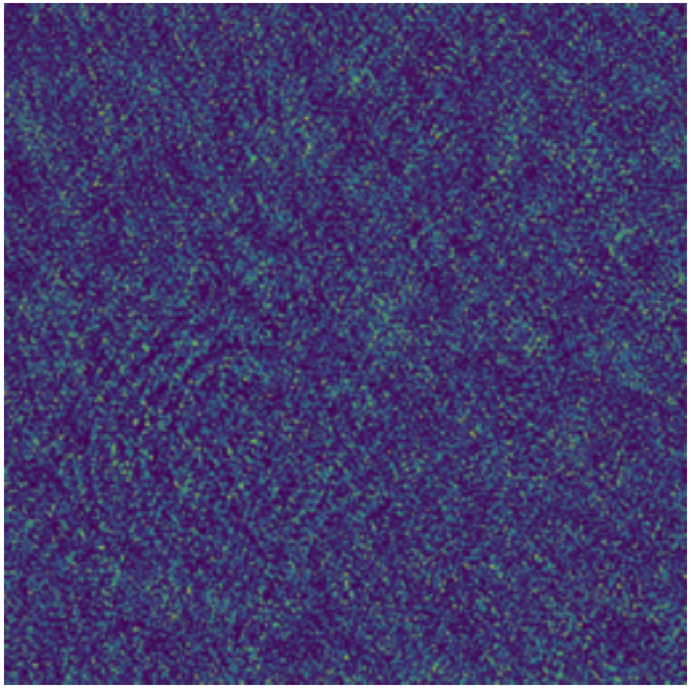
(a) Multi-scale model image without masking



(b) Multi-scale model image with automatic masking

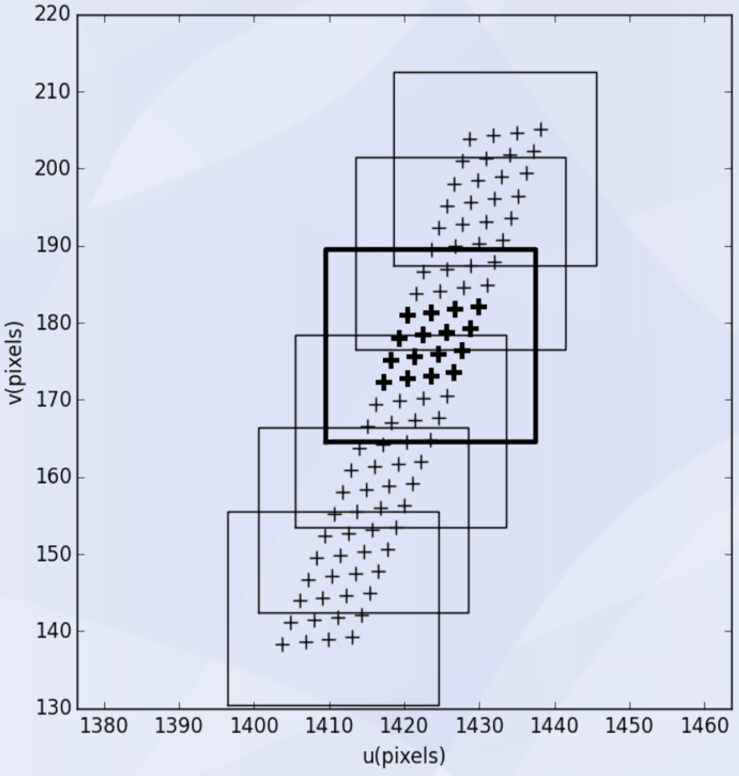


(c) Multi-scale residual without masking (rms=50 mJy/B)

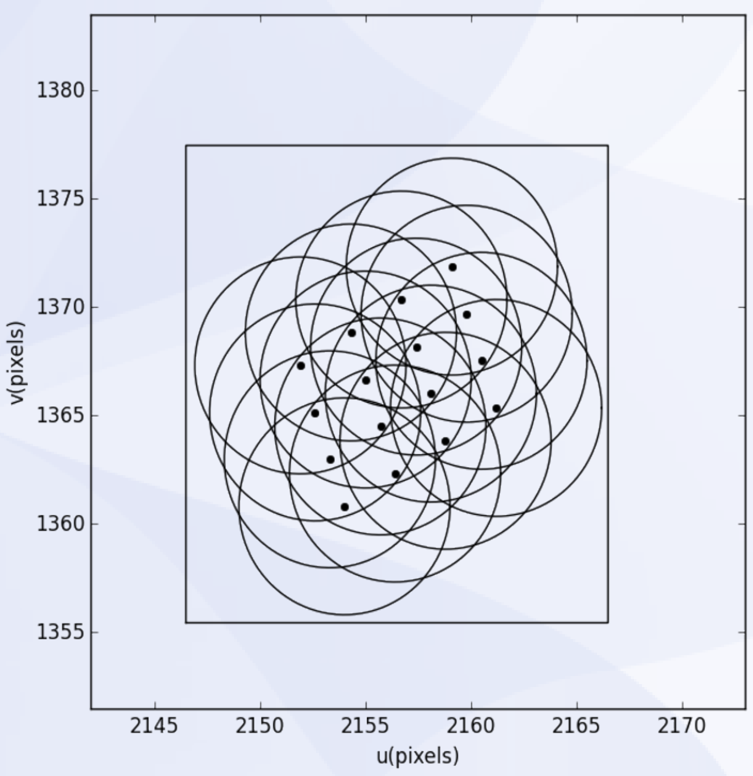


(d) Multi-scale residual with automatic masking (rms=38 mJy/B)

Image Domain Gridding (IDG)



(a)



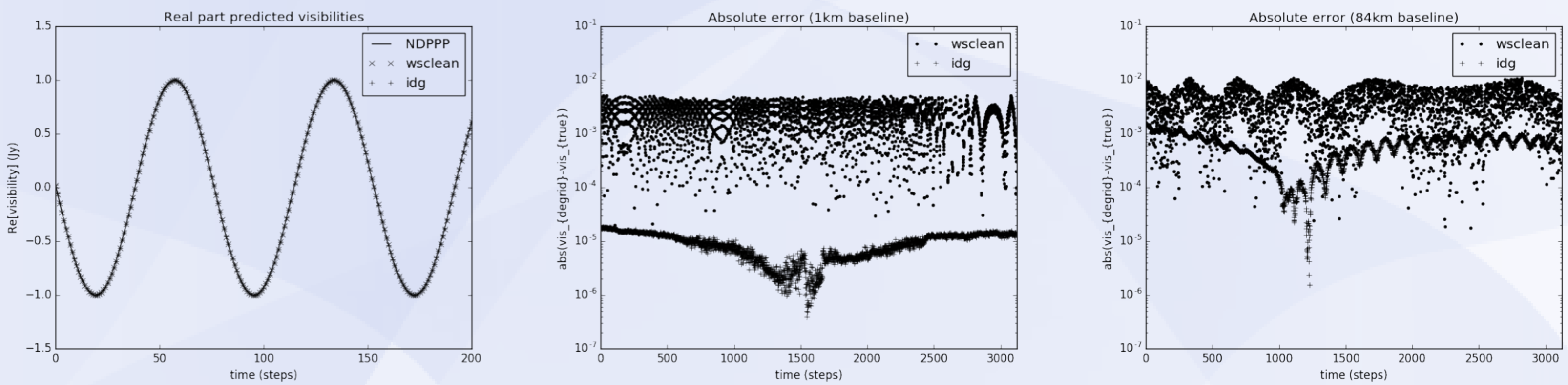
(b)

(a) uv track for a single baseline and multiple channels. The boxes indicate the position of the subgrids. The bold box correspond to the bold samples. (b) A single subgrid (box) encompassesing all affected pixels in the uv grid. The support of the convolution function is indicated by the circles around the samples.

Van der Tol, Veenboer & Offringa (A&A, 2018)

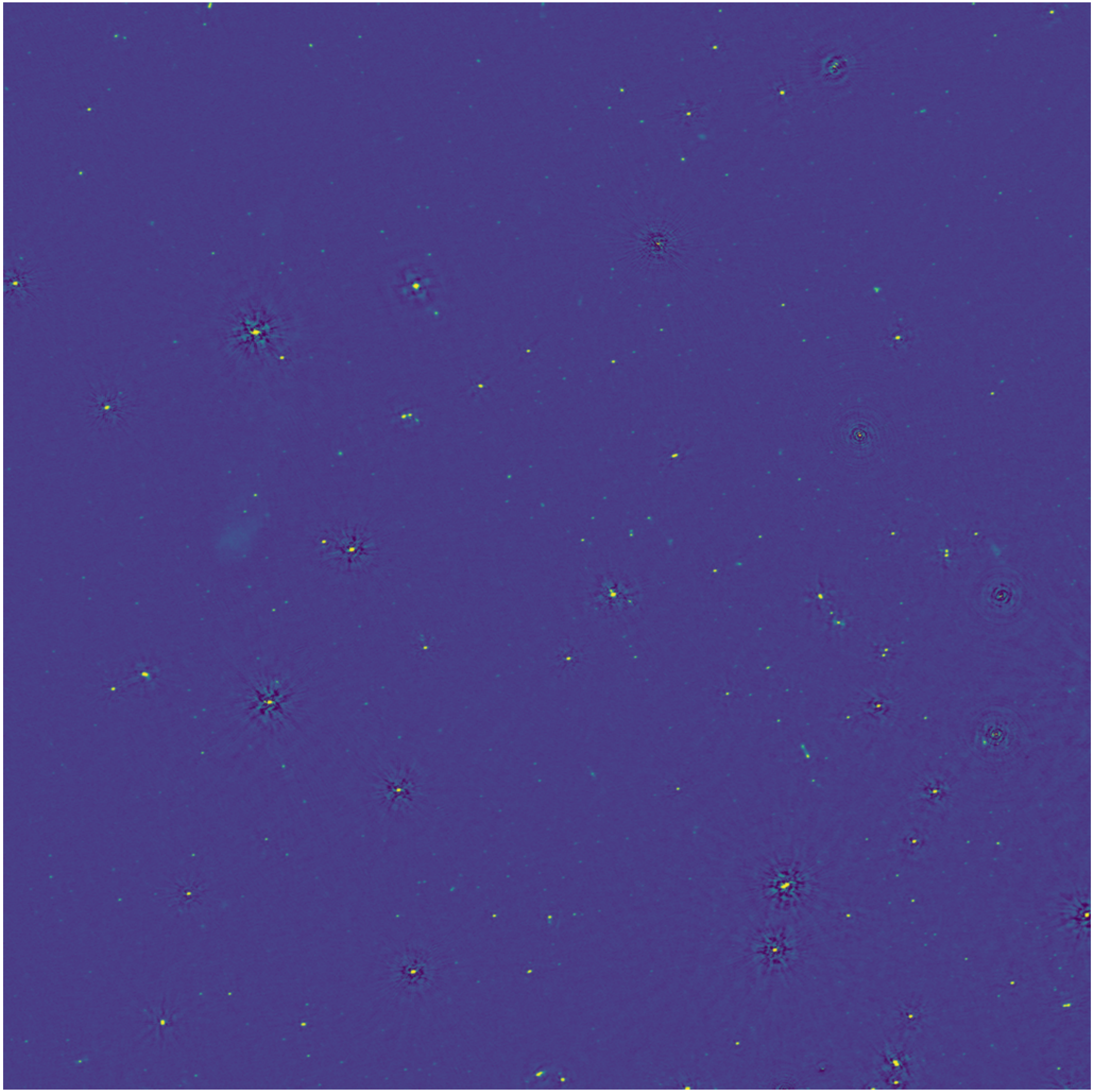
Image Domain Gridding (IDG)

- Compared to normal gridding, IDG does (on first order) not change the amount of operations to be performed
- However, parallelizes extremely well, even more so on GPUs
- W & A-term (beam/ionosphere) correction “for free”
- Results in very high gridding accuracy:



Left: visibilities for a point source as predicted by direct evaluation of the ME, and degridding by the classical gridder and image domain gridder. The visibilities are too close together to distinguish in this graph. The plot and the middle and on the right show the absolute value of the difference between direct evaluation and degridding for a short (1km) and a long (84km) baseline. On the short baseline the image domain gridder rms error of 1.03×10^{-5} Jy is about 242 times lower than the classical gridder rms error of 2.51×10^{-3} Jy. On the long baseline the image domain gridder rms error of 7.10×10^{-4} Jy is about 7 times lower than the classical gridder error of 4.78×10^{-3} Jy.

Van der Tol, Veenboer & Offringa (A&A, accepted)



Zoomed 25k x 25k image

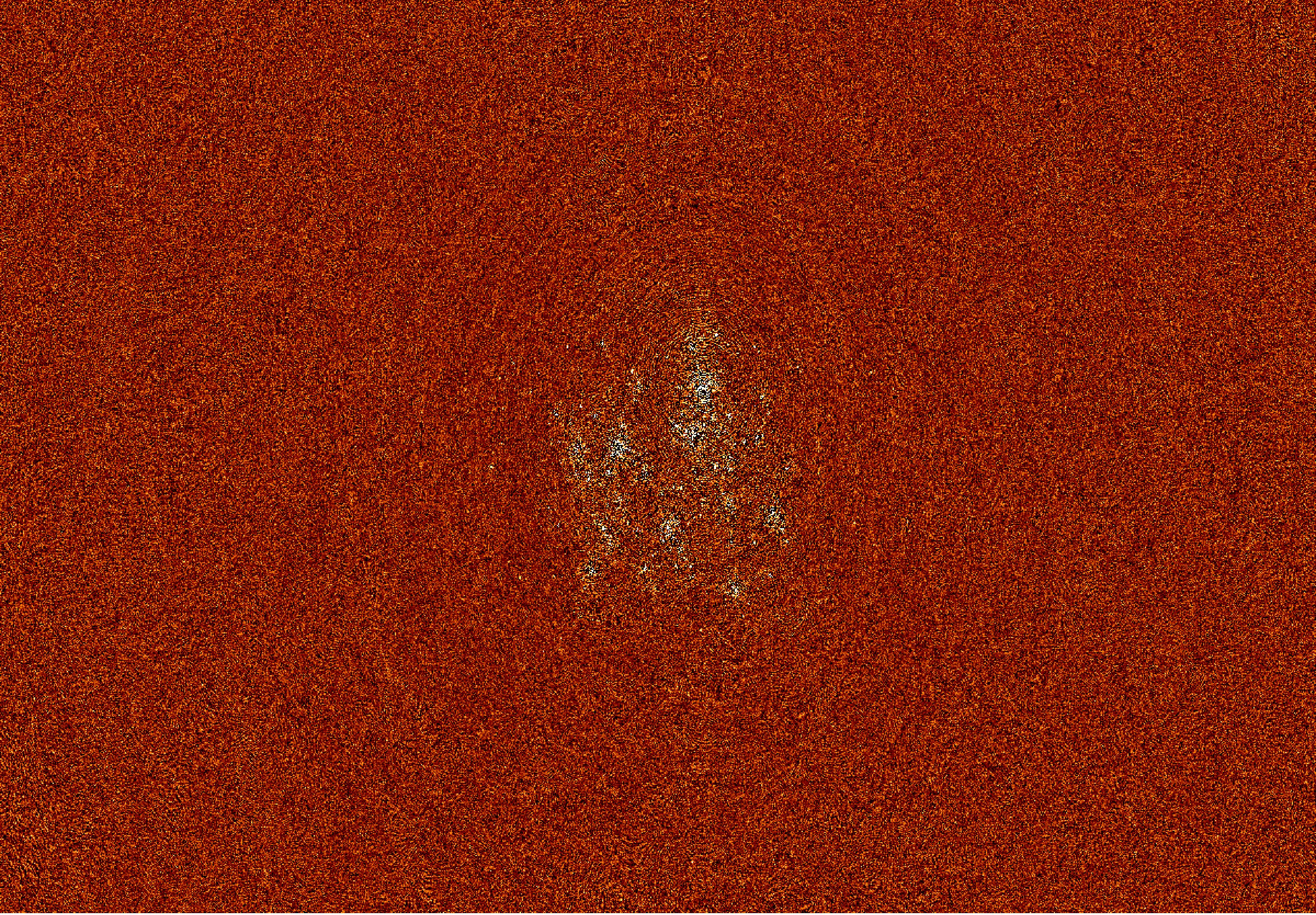
LOFAR, 48 MHz 6 h

Gridding with IDG

IDG + WSClean

(Both are publicly available)

Fully multi-scale multi-frequency cleaned IDG 25k x 25k result

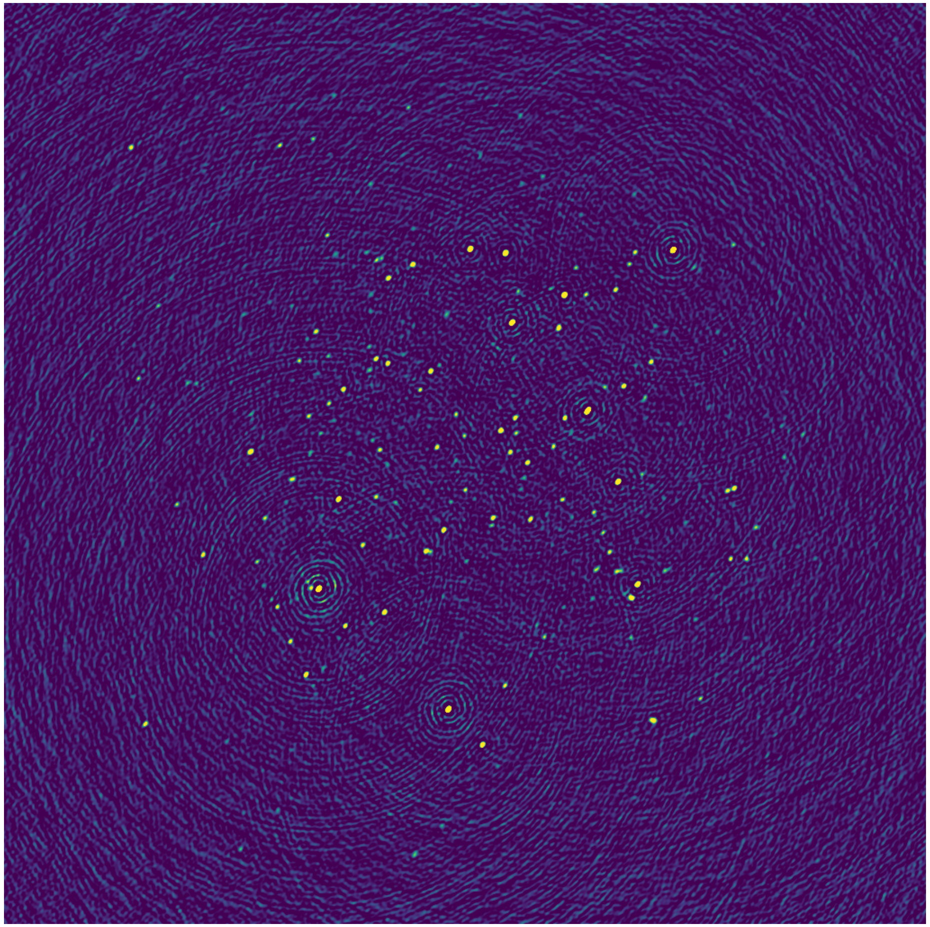


100K x 100K image using facetting

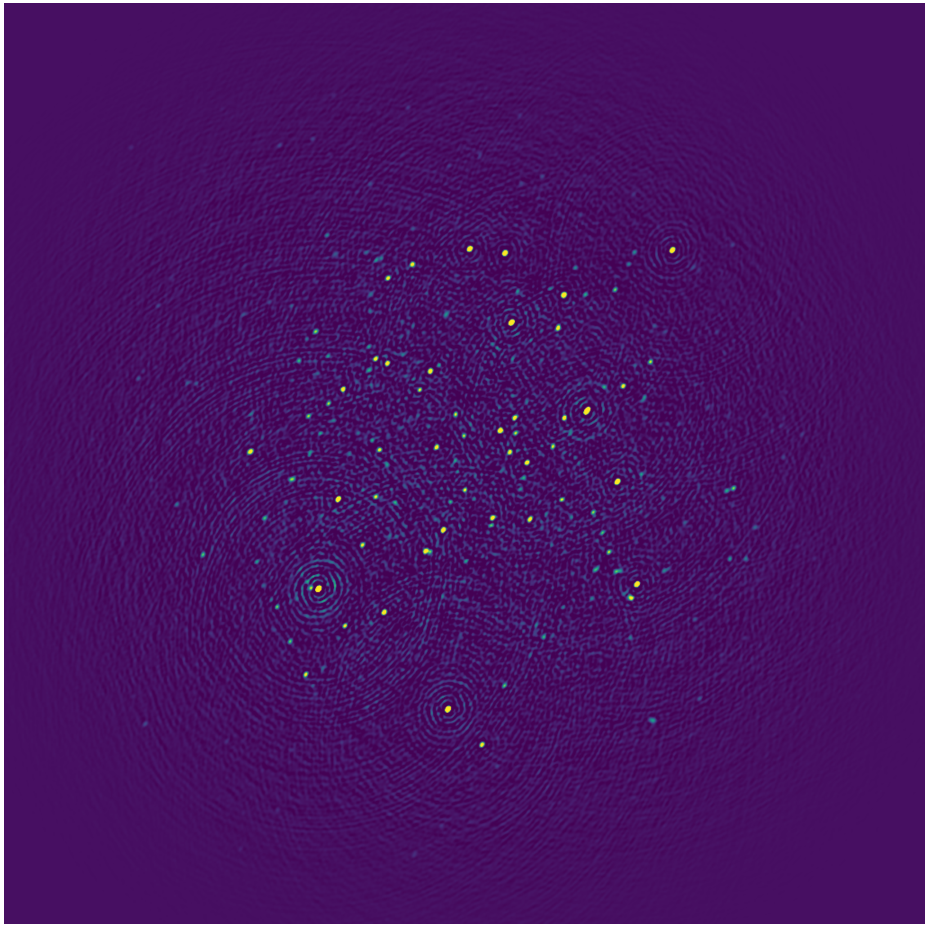
Applying a-term with IDG

Next step: apply LOFAR beam

- Applies full-Jones antenna beam in forward and backward imaging step
- No extra computational cost. WSClean parameter: `-grid-with-beam`



Normal imaging with w-stacking gridder
(no beam)

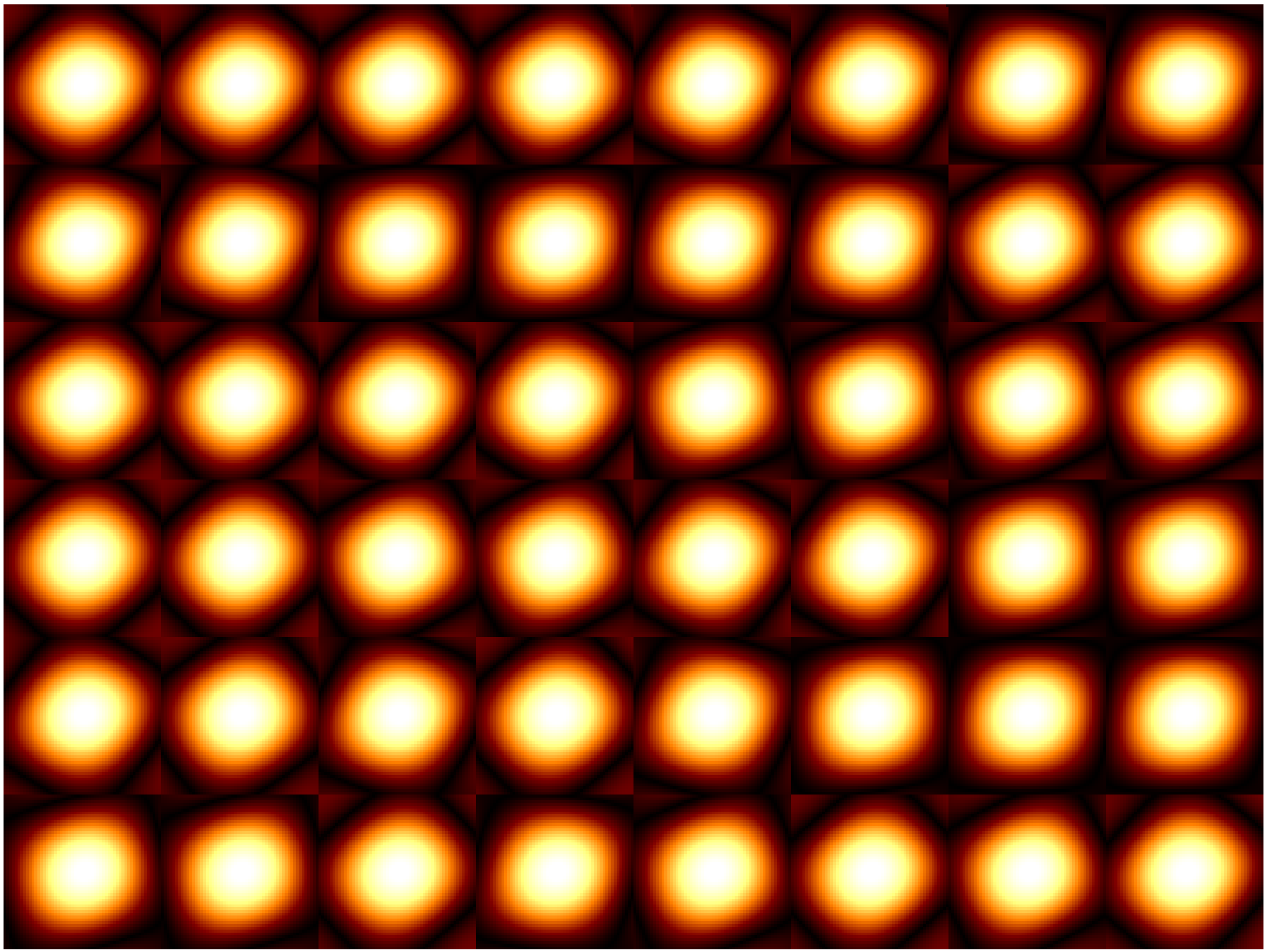


LOFAR beam applied during imaging stage
Producing "optimally weighted" image
(this is the "raw" uncorrected output)

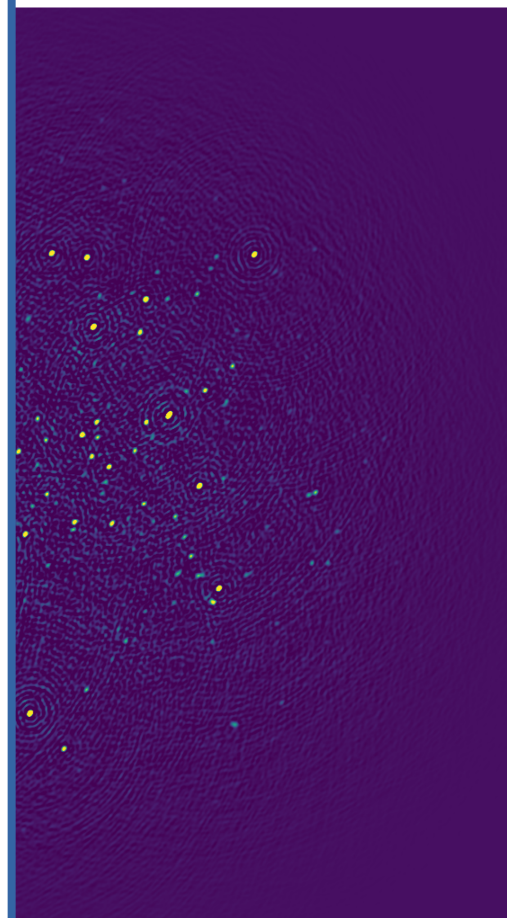
Applying a-term with IDG

Next step: apply LOFAR beam

Snapshot of the LOFAR beam for the 48 stations:



ward imaging step



ed during imaging stage

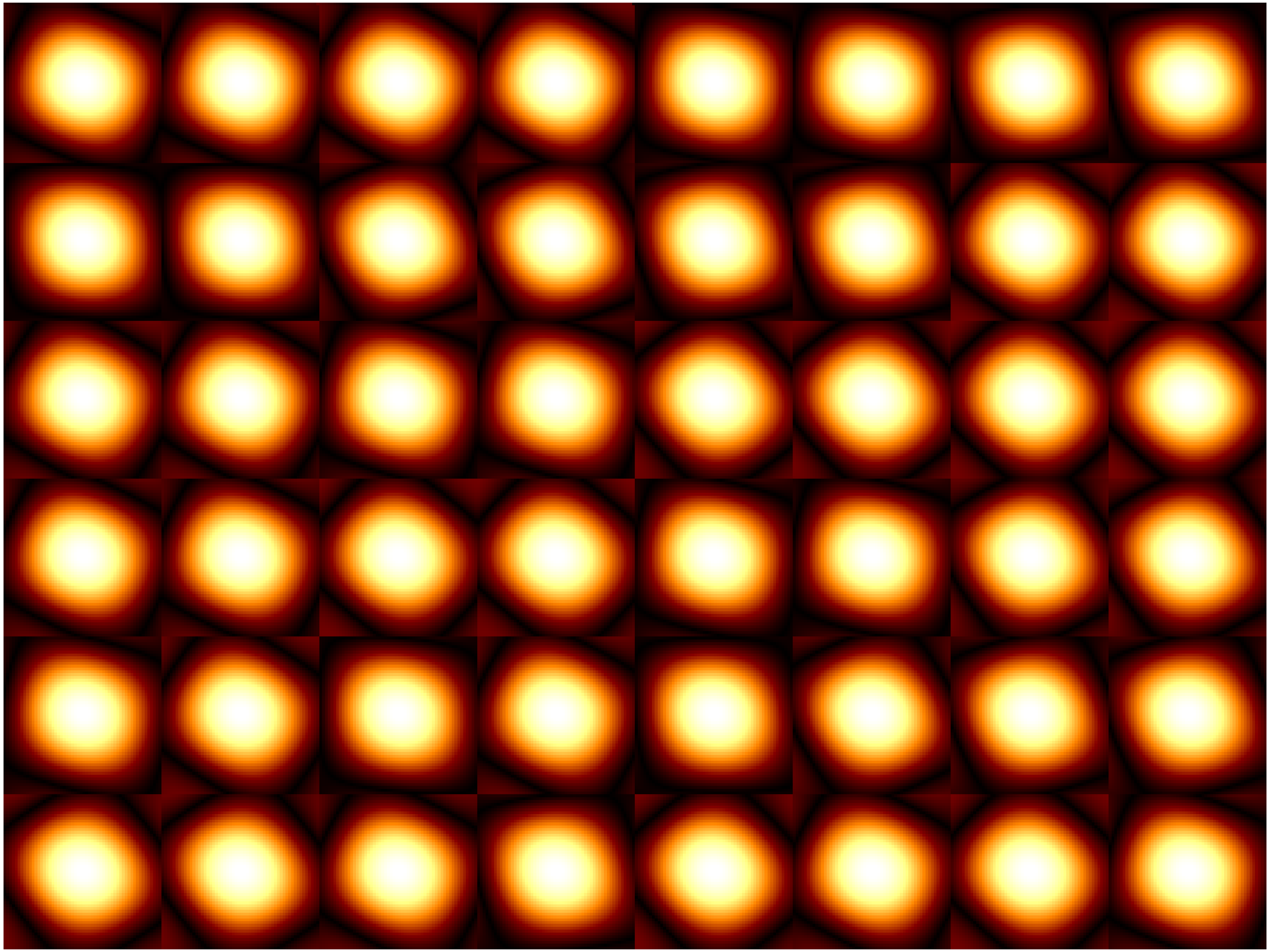
(no beam)

Producing “optimally weighted” image

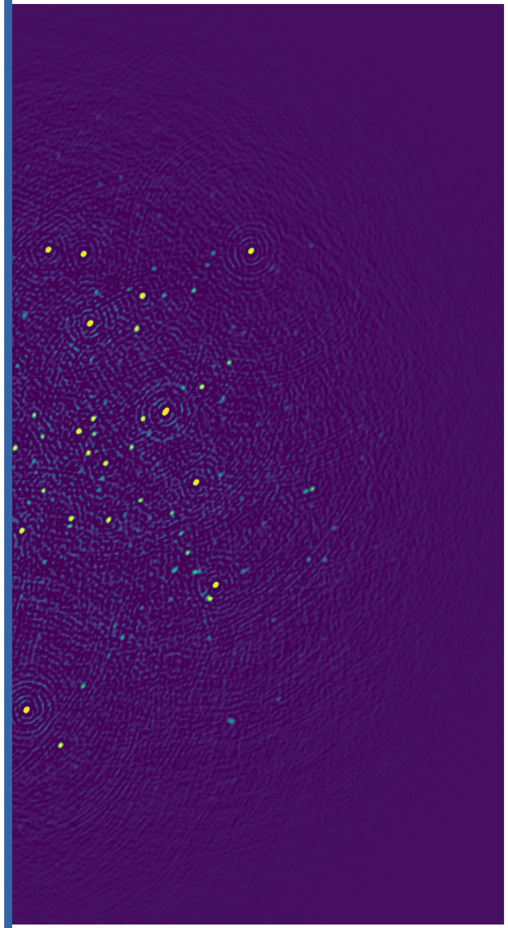
Applying a-term with IDG

Next step: apply LOFAR beam

Snapshot of the LOFAR beam for the 48 stations:



ward imaging step

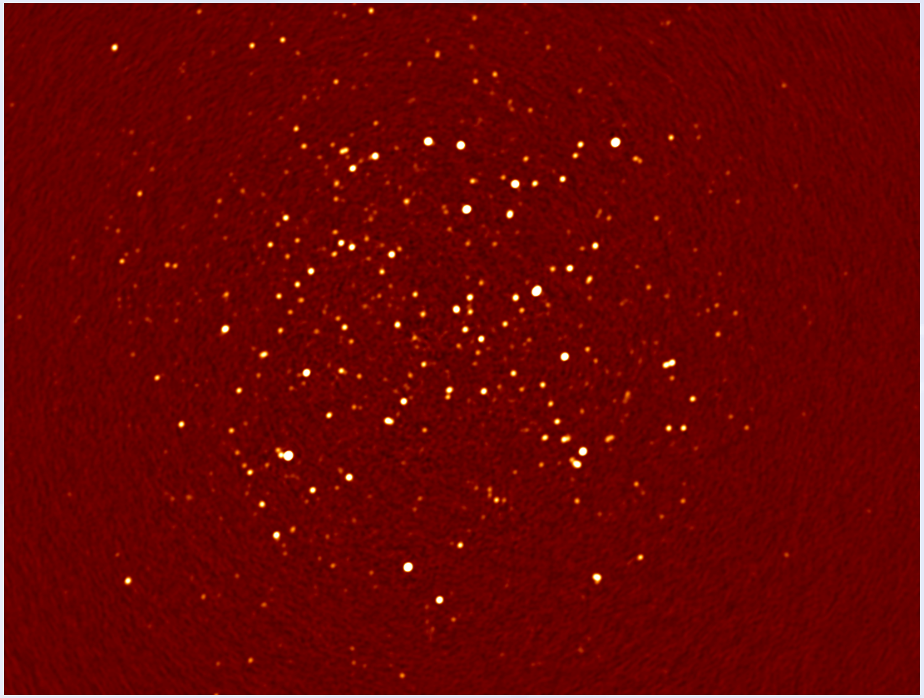


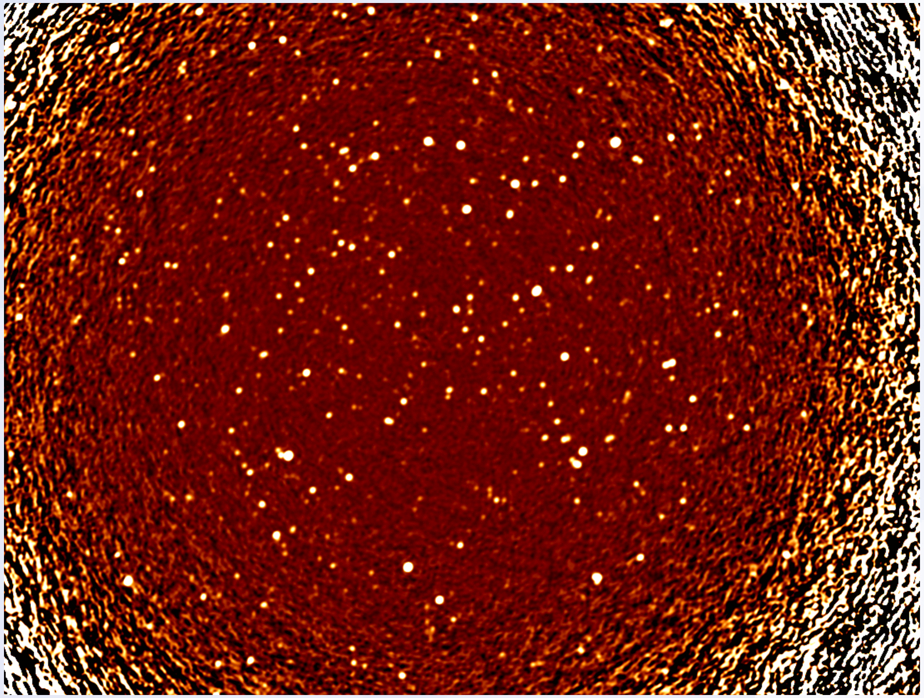
(no beam)

ed during imaging stage

Producing “optimally weighted” image

A-term correction & cleaning





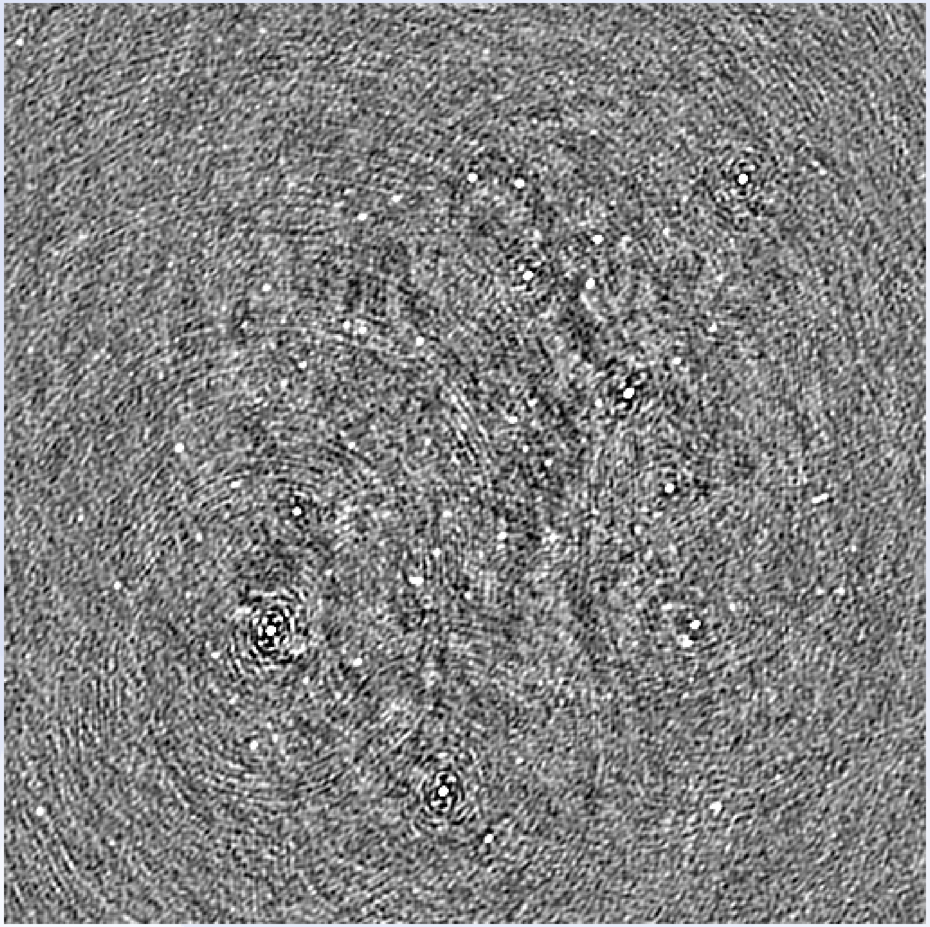
"Flat noise" vs "flat gain"

Used for cleaning

Final WSClean output

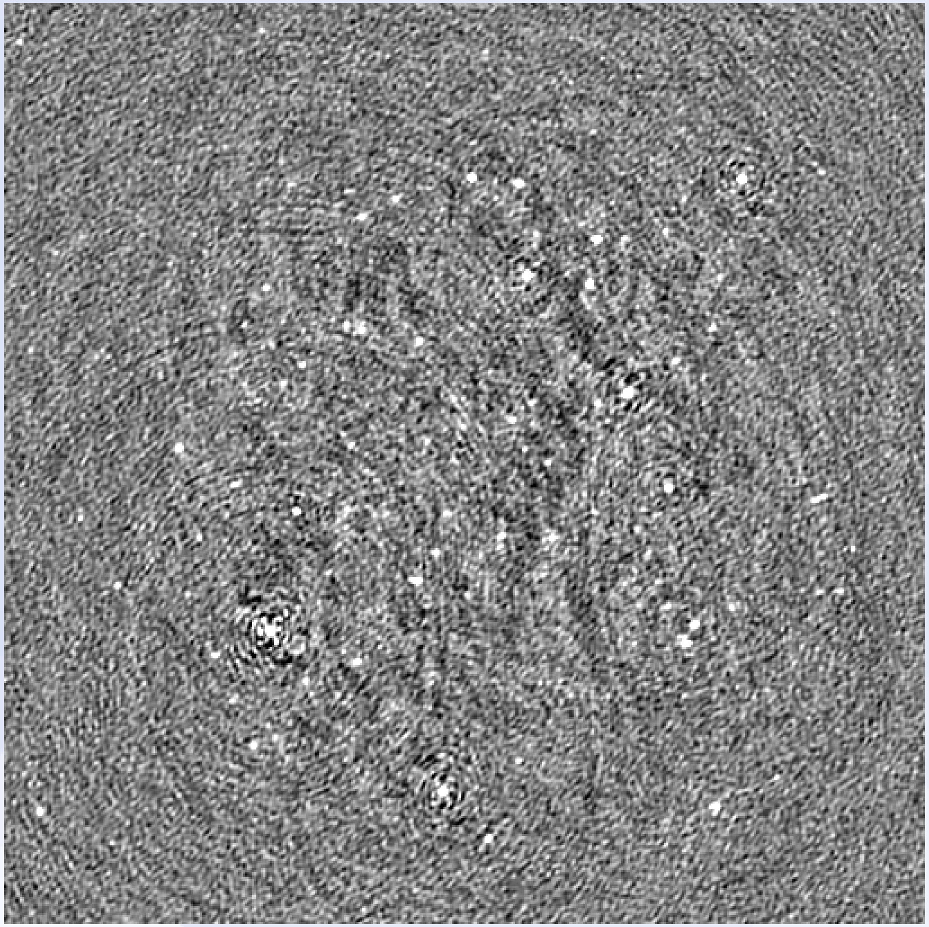
Imaging without beam

dirty image – Stokes Q (single subband)



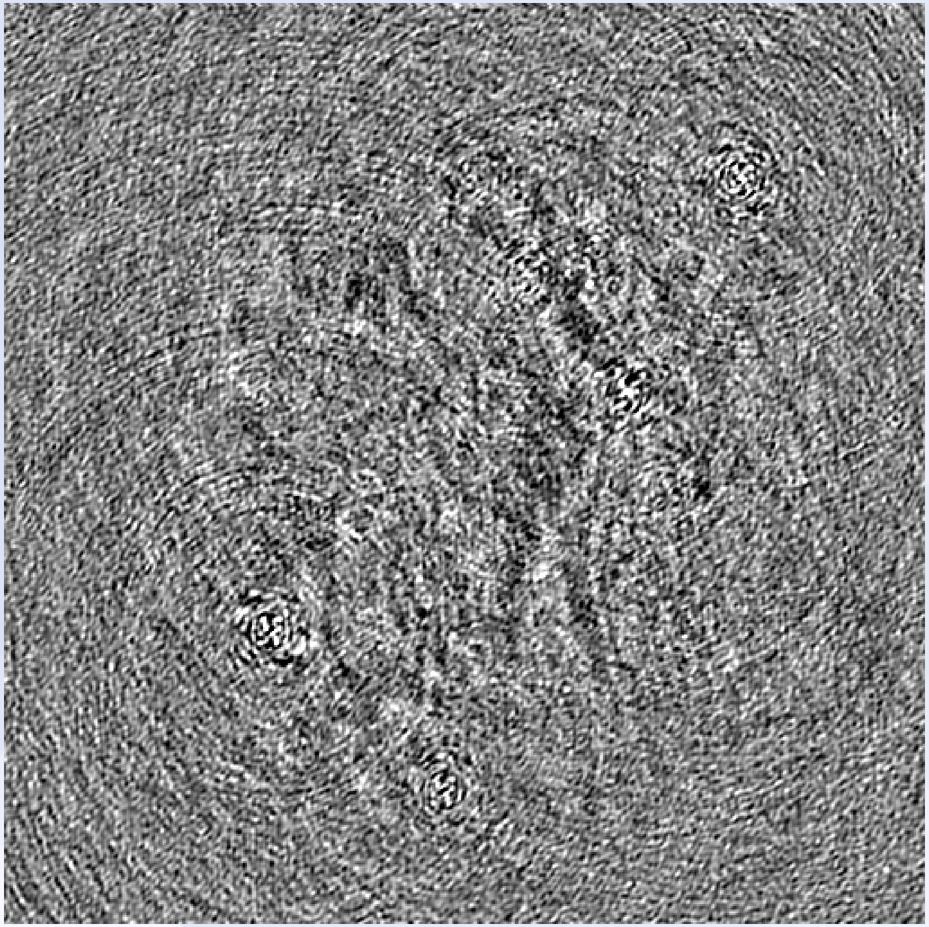
Imaging without beam

cleaned image – Stokes Q (single subband)



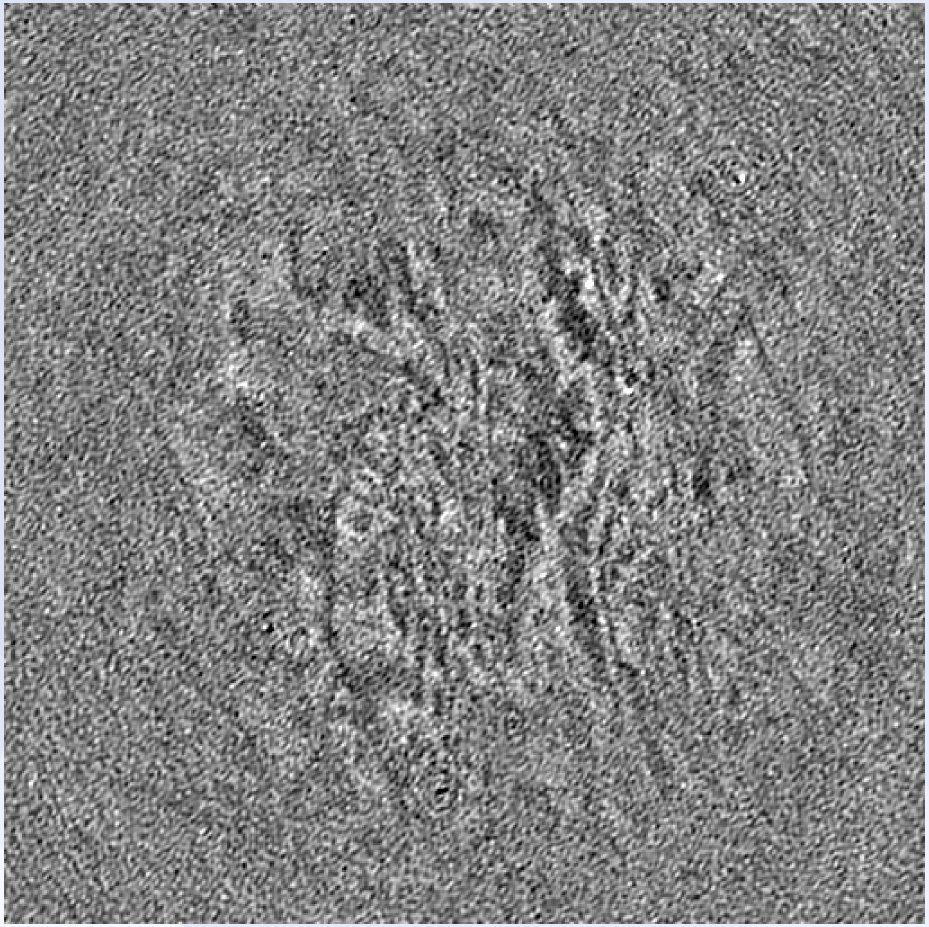
Imaging with beam

dirty image – Stokes Q (single subband)



Imaging with beam

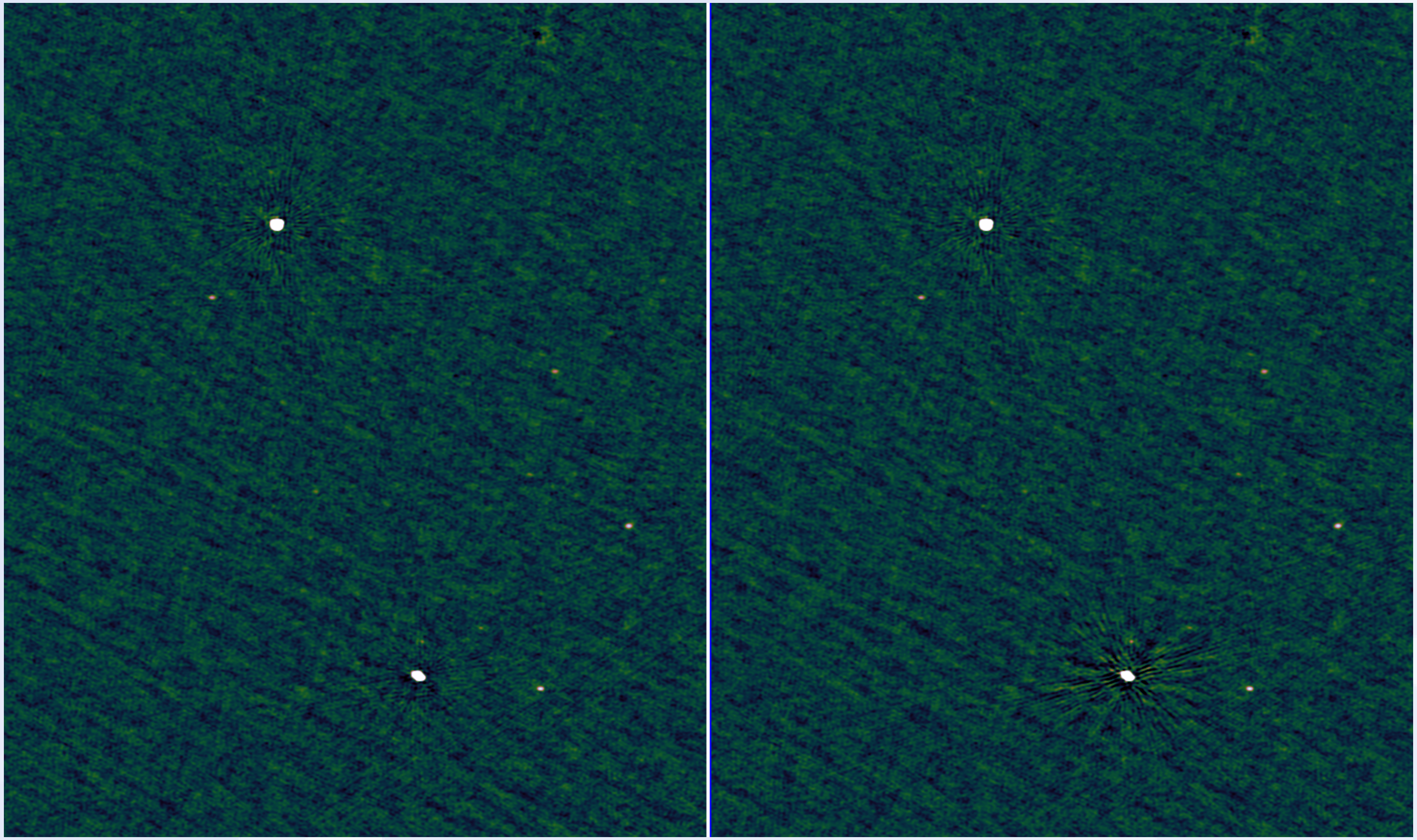
cleaned image – Stokes Q (single subband)



Apply TEC screen

- IDG can apply “screens”
- Benefit:
  - No facets (/ edges)
  - Good for diffuse emission
- But more complex:
  - Have to convert solutions to screen
- See Rapthor talk by David!

Screen vs facets

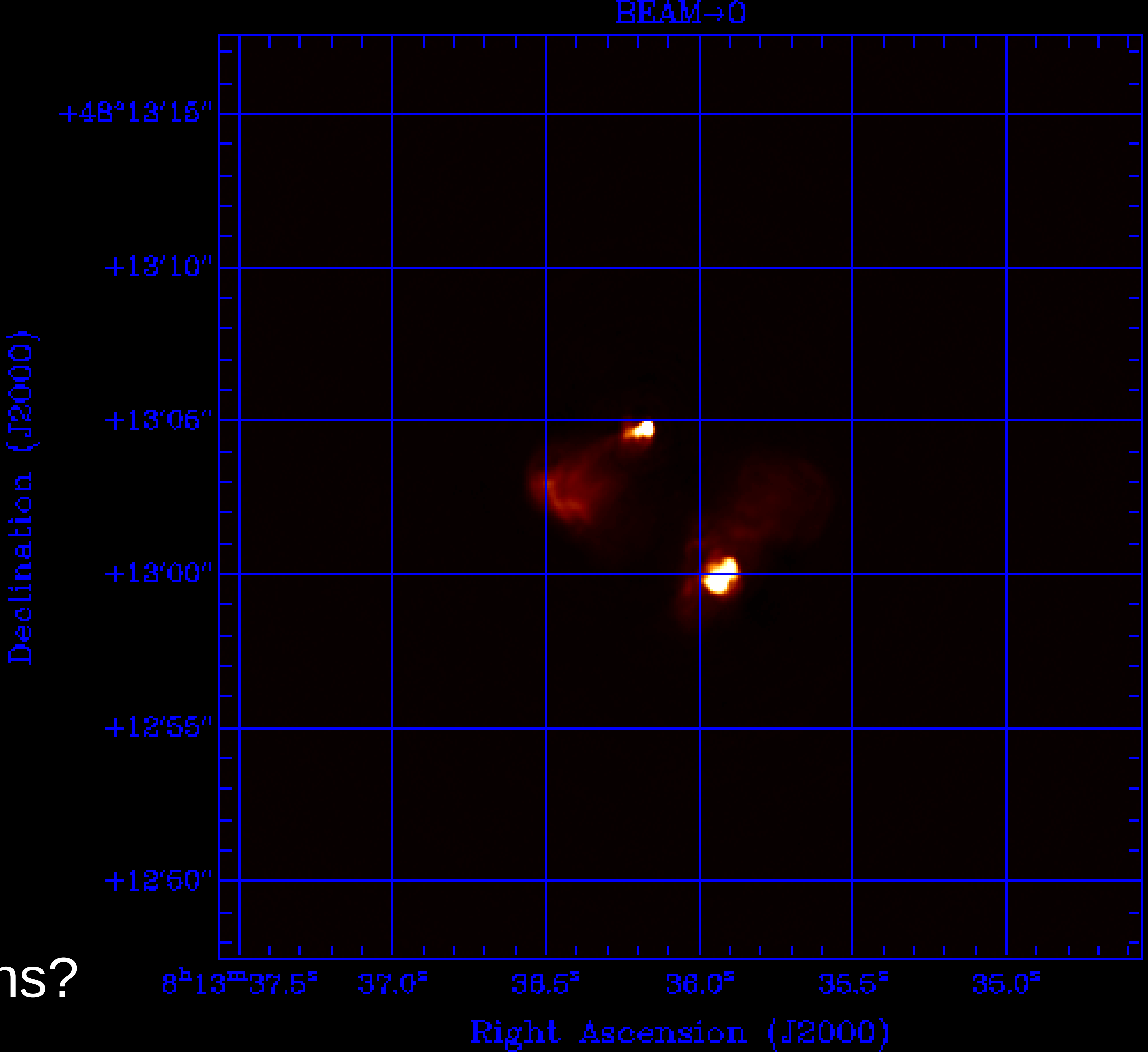


Screen

Faint source gets interpolate solution from screen

Facet

Faint source gets same solution as bright source


BEAM→0
+48°13'15"
+13'10"
+13'05"
+13'00"
+12'55"
+12'50"
Declination (J2000)
8h13m37.5s
37.0s
36.5s
36.0s
35.5s
35.0s
Right Ascension (J2000)

Questions?