# PolyCLEAN

A Polyatomic CLEAN-like algorithm for sparse Bayesian imaging

Adrian Jarret
PhD Student @EPFL

Swiss SKA Days 2023

EPFL

## 01
## Background
Radio Interferometry
and the CLEAN realm

## 02
## MAP estimation
Optimization problems and
numerical challenges

## 03
## PolyCLEAN
Convex optimization solved in
an atomic manner

## 04
## Demonstration
Performances and
experimental reconstructions

# 01. Background

Radio Interferometric Imaging

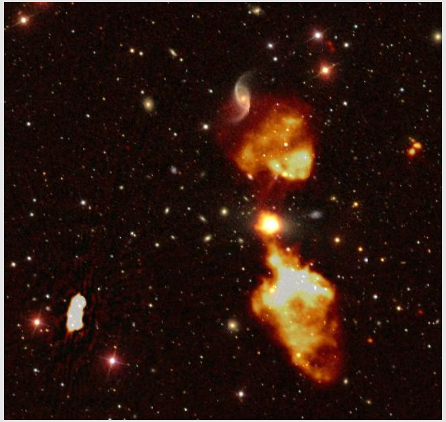
# Linear Inverse Problem

$$\boxed{\mathbf{V} = \mathbf{\Phi} \mathbf{I}}$$

# Linear Inverse Problem

$$\boxed{\mathbf{V} = \mathbf{\Phi} \mathbf{I}}$$

Observed area of the sky



[Credits: Cyril Tasse and the LOFAR surveys team.]

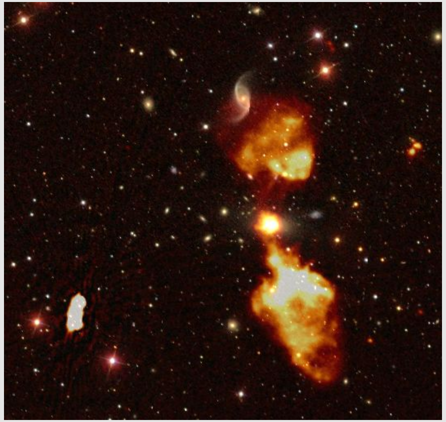
# Linear Inverse Problem

*Visibility* measurements

Spatial frequency information,
*Fourier-like* measurements,
complex valued

$$\boxed{\mathbf{V} = \mathbf{\Phi} \mathbf{I}}$$

Observed area of the sky



[Credits: Cyril Tasse and the LOFAR surveys team.]

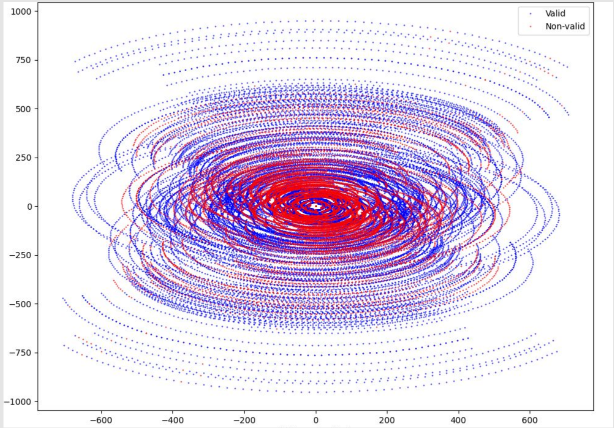
# Linear Inverse Problem

Observed area of the sky

*Visibility* measurements

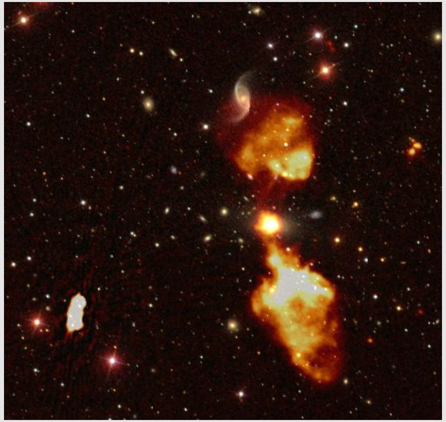
Spatial frequency information,
*Fourier-like* measurements,
complex valued

$$\mathbf{V} = \mathbf{\Phi I}$$




Interferometry operator

Depends on the location of the
antennas (*baselines)*
*And the observed wavelength*



[Credits: Cyril Tasse and the LOFAR surveys team.]

# Measurement Operator

Image

Sky coordinates

$$\Phi(\mathbf{I})_k = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\mathbf{I}[i,j]}{n(l_i, m_j)} e^{-\mathrm{j}2\pi(u_k l_i + v_k m_j)} \mathcal{W}(l_i, m_j; w_k)$$

Baseline coordinates

Linear IP

$$\boxed{\mathbf{V} = \mathbf{\Phi I}}$$

# Challenges of RI

- Noisy measurements

$$\mathbf{V} = \mathbf{\Phi I} + \varepsilon$$

Linear IP

$$\boxed{\mathbf{V} = \mathbf{\Phi I}}$$

# Challenges of RI

- Noisy measurements

$$\mathbf{V} = \mathbf{\Phi I} + \varepsilon$$

- Ill-posed problem

$$\mathrm{Null}(\mathbf{\Phi}) \neq \{0\}$$

Linear IP

$$\boxed{\mathbf{V} = \mathbf{\Phi}\mathbf{I}}$$

# Challenges of RI

- Noisy measurements

$$\mathbf{V} = \mathbf{\Phi}\mathbf{I} + \varepsilon$$

- Ill-posed problem

$$\mathrm{Null}(\mathbf{\Phi}) \neq \{0\}$$

Use of priors for reconstruction!

Linear IP

$$\boxed{\mathbf{V} = \mathbf{\Phi I}}$$

# Challenges of RI

- Noisy measurements

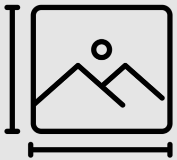
$$\mathbf{V} = \mathbf{\Phi I} + \varepsilon$$

- Ill-posed problem

$$\mathrm{Null}(\mathbf{\Phi}) \neq \{0\}$$

Use of priors for reconstruction!

- Huge volumes of data



# The CLEAN Algorithm

Linear IP

$$\boxed{\mathbf{V} = \mathbf{\Phi}\mathbf{I}}$$

## Parametric Shape of the Solutions

$$\mathbf{I}^*[n] = \sum_k \alpha_k g(n - n_k)$$

## Matching Pursuit Algorithm

Iterative atomic updates

$$\mathbf{I}^{(k+1)} \leftarrow \mathbf{I}^{(k)} + \alpha \mathbf{g}(\cdot - n_k)$$

**Empirical sparsity along iteration**

# The CLEAN Algorithm

Linear IP

$$\boxed{\mathbf{V} = \boldsymbol{\Phi}\mathbf{I}}$$

## Parametric Shape of the Solutions

$$\mathbf{I}^*[n] = \sum_k \alpha_k \boxed{\delta(}n - n_k)$$

Point sources

## Matching Pursuit Algorithm

Iterative atomic updates

$$\mathbf{I}^{(k+1)} \leftarrow \mathbf{I}^{(k)} + \alpha \mathbf{g}(\cdot - n_k)$$

**Empirical sparsity along iteration**

# The Algorithm

**Algorithm 1** Högbom CLEAN Algorithm (Major cycles only)

**Parameters :** $k_{\max}$ (iterations), $\alpha > 0$ (gain)

**Initialisation :** $\mathbf{I}^{(0)} = \mathbf{0}$, $\mathbf{I}_D = \mathbf{\Phi}^* \mathbf{V}$

**for** $k = 1, 2, \cdots, k_{\max}$ **do**

1. Compute the dirty residual: $\mathbf{I}_R^{(k)} = \mathbf{I}_D - \mathbf{\Phi}^* \mathbf{\Phi} \mathbf{I}^{(k-1)}$
2. Find the location of the next reconstructed source: $s^{(k)} = \arg\max_{(i,j)} \left| \mathbf{I}_R^{(k)}[i,j] \right|$
3. Update the iterate: $\mathbf{I}^{(k)} = \mathbf{I}^{(k-1)} + \alpha (\max \mathbf{I}_R^{(k)}) \boldsymbol{\delta}_{s^{(k)}}$

**end for**

**Output:**

Postprocess $\mathbf{I}^{(k)}$ (convolution with synthetic beam, add residual image)

# CLEAN-Like methods

- ✓ Atomic method (scalable)
- ✓ A lot of hacks and tips to make them very fast
- ✓ Developed and maintained by the astronomers
- ✓ Long date expertise
- ✓ Calibration-compliant

# CLEAN-Like methods (continued)

✓ Atomic method (scalable)

✓ A lot of hacks and tips to make them very fast

✓ Developed and maintained by the astronomers

✓ Long date expertise

✓ Calibration-compliant

✗ Only denoising = enforcing the prior model

✗ Very sensitive to stop

✗ Objective function unclear

# 02. Bayesian MAP Estimation

A principled way to introduce prior information

# LASSO as a MAP estimator

$$\arg\min \underbrace{\frac{1}{2}\|\mathbf{V} - \mathbf{\Phi I}\|_2^2}_{\text{Log likelihood}} + \underbrace{\lambda\|\mathbf{I}\|_1}_{\text{Prior distribution}}$$

- Convex optimization methods
- Sparse solutions => Well suited for **Point Sources**

# Sparse Dictionary reconstruction

$$\arg\min_{\boldsymbol{\theta}} \frac{1}{2}||\mathbf{V} - \boldsymbol{\Phi}\boxed{\boldsymbol{\Psi}}\boldsymbol{\theta}||_2^2 + \lambda||\boldsymbol{\theta}||_1$$

$\boldsymbol{\Psi} \in \mathbb{R}^{N\times M}$ Dictionary synthesis operator

$\boldsymbol{\theta} \in \mathbb{R}^{M}$ Dictionary coefficients

# Optimization methods

- ✓ Denoising (with only one parameter!)
- ✓ Excellent reconstruction quality demonstrated
- ✓ Can handle very complex priors
- ✓ Fast principled algorithms

# Optimization methods (continued)

- ✓ Denoising (with only one parameter!)
- ✓ Excellent reconstruction quality demonstrated
- ✓ Can handle very complex priors
- ✓ Fast principled algorithms

- ✗ Completely different implementation paradigm (proximal method)
- ✗ Memory scalability
- ✗ Non calibration-compliant
- ✗ Shrinkage of the reconstructed intensity

# 03. PolyCLEAN

Convex optimization for RA
enabled with an atomic method

## 1. Optimization method

Penalty-based prior

## 2. Atomic behavior

CLEAN-like algorithmic structure
and minor cycles

## 3. Focus on Scalability

Sparsity-informed computations with
Pycsou and NUFFT

## 1. Optimization method

Penalty-based prior

$$\boxed{\lambda \|\mathbf{I}\|_1 \, , \, \mathbf{I} \geqslant 0}$$

## 2. Atomic behavior

CLEAN-like algorithmic structure
and minor cycles

## 3. Focus on Scalability

Sparsity-informed computations with
Pycsou and NUFFT

## 1. Optimization method

Penalty-based prior

$$\boxed{\lambda \|\mathbf{I}\|_1 \,, \; \mathbf{I} \geqslant 0}$$

## 2. Atomic behavior

CLEAN-like algorithmic structure and minor cycles

$$\boxed{\mathbf{I} = \sum \alpha_k \boldsymbol{\delta}_{i_k}}$$

## 3. Focus on Scalability

Sparsity-informed computations with Pycsou and NUFFT

## 1. Optimization method

Penalty-based prior

$$\boxed{\lambda \|\mathbf{I}\|_1 \,,\; \mathbf{I} \geqslant 0}$$

## 2. Atomic behavior

CLEAN-like algorithmic structure
and minor cycles

$$\boxed{\mathbf{I} = \sum \alpha_k \boldsymbol{\delta}_{i_k}}$$

## 3. Focus on Scalability

$$\boxed{\boldsymbol{\Phi}\left(\mathbf{I}^{(k)}\right)}$$

Sparsity-informed computations with
Pycsou and NUFFT

# The Algorithm

**Algorithm 2** PolyCLEAN

**Initialisation :** $\mathbf{I}^{(0)} = \mathbf{0},\ \mathcal{S}^{(0)} = \mathrm{Supp}(\mathbf{I}^{(0)}) = \emptyset,\ \ \mathbf{I}_D = \mathbf{\Phi}^*\mathbf{V}$

**while** `stopping_criterion`$(\mathbf{I}^{(k)})$ not reached **do**

1. Compute the dirty residual: $\mathbf{I}_R^{(k)} = \mathbf{I}_D - \mathbf{\Phi}^*\mathbf{\Phi}\mathbf{I}^{(k-1)}$

2. Place many candidate sources: $s_1^{(k)}, s_2^{(k)}, \cdots =$ `highest_level_set` $(\mathbf{I}_R^{(k)})$

   Update active set : $\mathcal{S}^{(k)} \leftarrow \mathcal{S}^{(k-1)} \cup \{s_1^{(k)}, s_2^{(k)}, \ldots\}$

3. Update the iterate:

$$\mathbf{I}^{(k)} = \underset{\substack{\mathrm{Supp}(\mathbf{I}) \subset \mathcal{S}^{(k)} \\ \mathbf{I} \geqslant 0}}{\arg\min}\ \frac{1}{2}\left\|\mathbf{V} - \mathbf{\Phi}\mathbf{I}\right\|_2^2 + \lambda \left\|\mathbf{I}\right\|_1 \tag{R}$$

**end while**

# Focus on the sparse iterates

- Beneficial only if handled correctly
  - → Low memory requirement
  - → Simple model
  - → Fast computation

- NU Fourier Transform: Type II -> Type III

**Symbiosis with HVOX**

# Optimization methods (continued)

- ✓ Denoising (with only one parameter!)
- ✓ Excellent reconstruction quality demonstrated
- ✓ Can handle very complex priors
- ✓ Fast principled algorithms

- ✗ ~~Completely different implementation paradigm (proximal method)~~
- ✗ ~~Memory scalability~~
- ✗ ~~Non calibration-compliant~~
- ✗ Shrinkage of the reconstructed intensity

# The Landscape of Methods

CLEAN

Convex Optimization

- L1 penalty and positivity
- Objective function
- Interpretability

- Atomic algorithm
- Scalability

PolyCLEAN

# 04. Numerical Results

It works, and it is fast

# Performance benchmark

1. Pick an interferometer radius and image size

# Performance benchmark

1. Pick an interferometer radius and image size
2. Simulate a source sky image

# Performance benchmark

1. Pick an interferometer radius and image size
2. Simulate a source sky image
3. Simulate a measurement set

# Performance benchmark

1. Pick an interferometer radius and image size
2. Simulate a source sky image
3. Simulate a measurement set
4. Solve with LASSO solvers:
   a. PolyCLEAN
   b. APGD
   c. MonoFW

# Performance benchmark

1. Pick an interferometer radius and image size
2. Simulate a source sky image
3. Simulate a measurement set
4. Solve with LASSO solvers:
    a. PolyCLEAN
    b. APGD
    c. MonoFW
5. Solve with WS-CLEAN

# Performance benchmark

1. Pick an interferometer radius and image size
2. Simulate a source sky image
3. Simulate a measurement set
4. Solve with LASSO solvers:
    a. PolyCLEAN
    b. APGD
    c. MonoFW
5. Solve with WS-CLEAN
6. Compare reconstruction time

# Performance benchmark

1. Pick an interferometer radius and image size
2. Simulate a source sky image
3. Simulate a measurement set
4. Solve with LASSO solvers:
   a. PolyCLEAN
   b. APGD
   c. MonoFW
5. Solve with WS-CLEAN
6. Compare reconstruction time

Largest measured frequency ↗
Image size (in px) ↗
Pixel size ↘

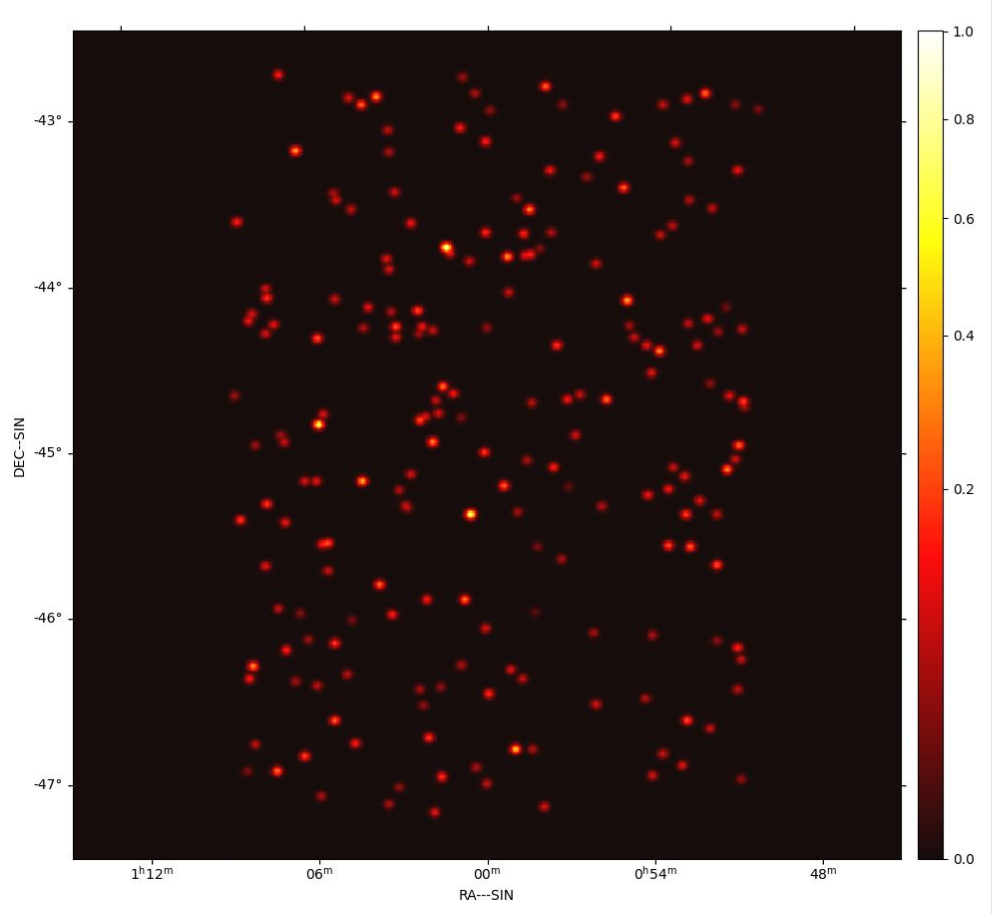
→ Observe scalability

# Simulated Source image

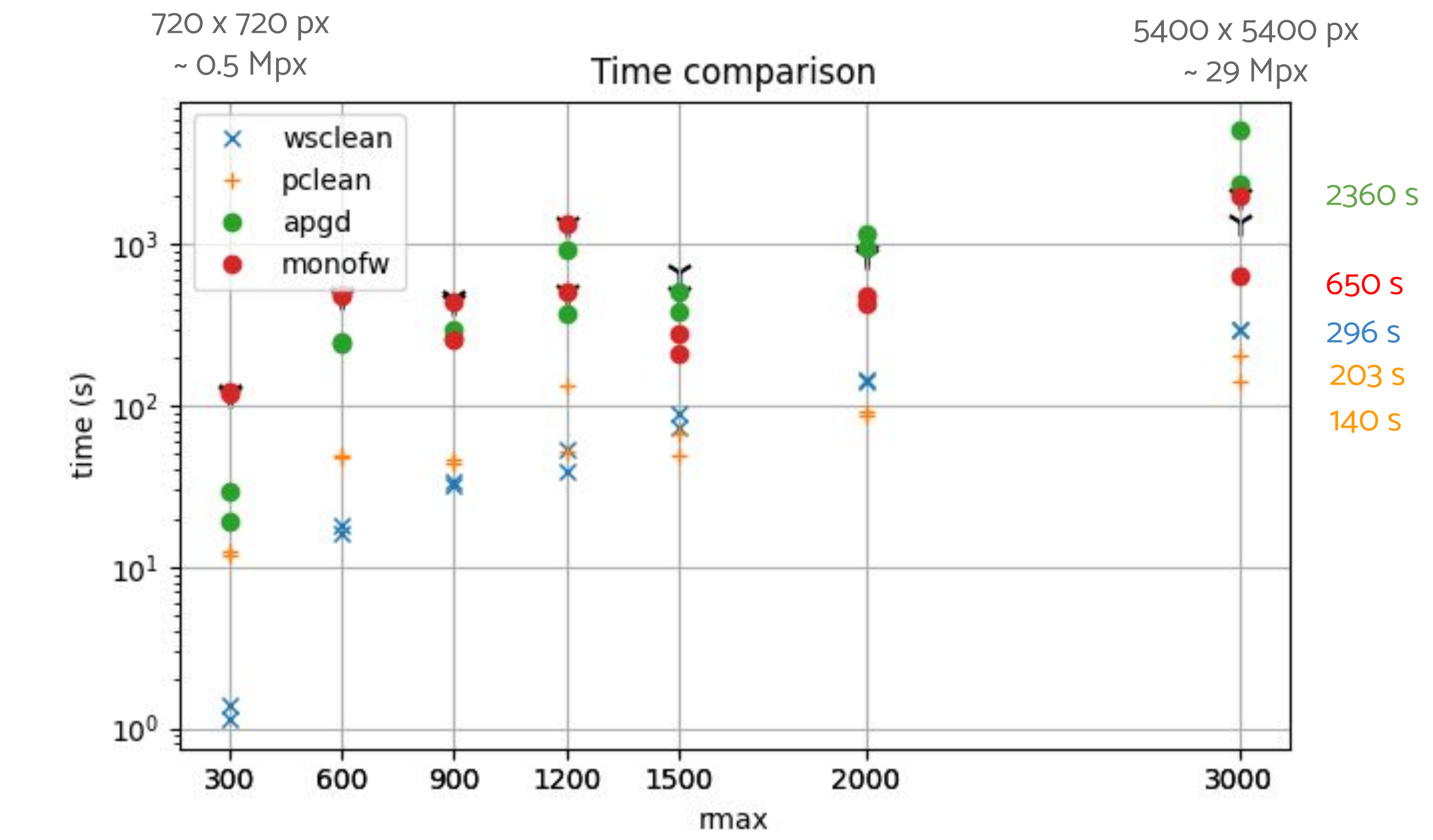
200 point sources

5° x 5° FOV

Image size: 720 -> 5400 pixels



# Reconstruction benchmark



# Real world measurement

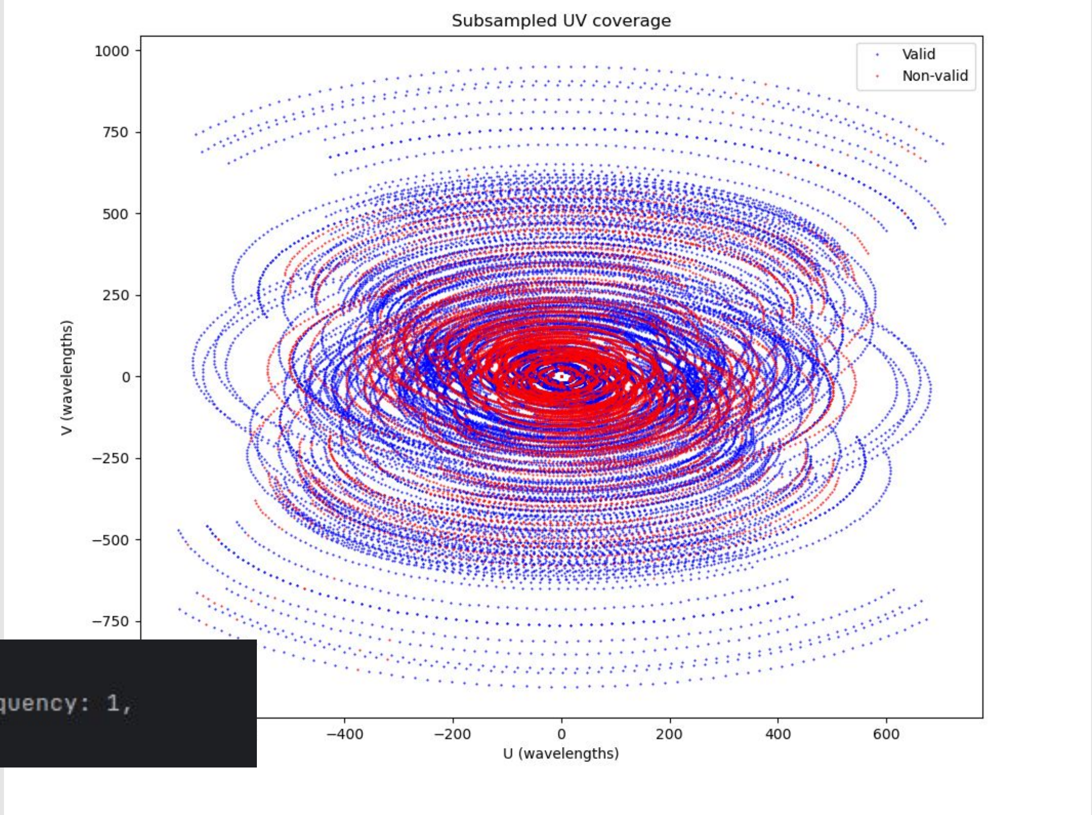
~ 400 MB

```
<xarray.Visibility>
Dimensions:             (time: 3595, baselines: 1953, frequency: 1,
                         polarisation: 1, spatial: 3)
Coordinates:
  * time                (time) float64 4.914e+09 4.914e+09 ... 4.914e+09
  * baselines           (baselines) object MultiIndex
  * antenna1            (baselines) int64 0 0 0 0 0 0 0 ... 58 59 59 59 60 60 61
  * antenna2            (baselines) int64 0 1 2 3 4 5 6 ... 61 59 60 61 60 61 61
  * frequency           (frequency) float64 1.458e+08
  * polarisation        (polarisation) <U1 'I'
  * spatial             (spatial) <U1 'u' 'v' 'w'
```

# Selection of antennas

1/50 measurement times

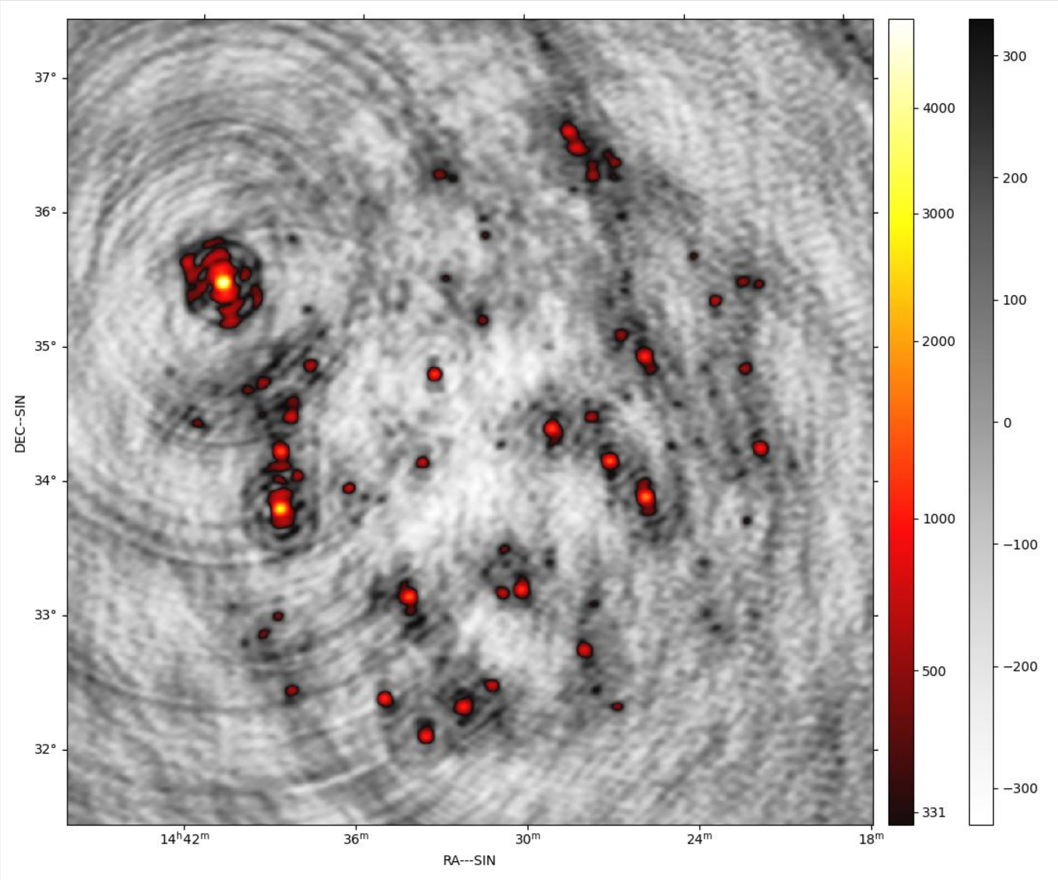
24 antennas




```
<xarray.Visibility>
Dimensions:            (time: 72, baselines: 406, frequency: 1,
                        polarisation: 1, spatial: 3)
```

# Dirty image - Point Sources

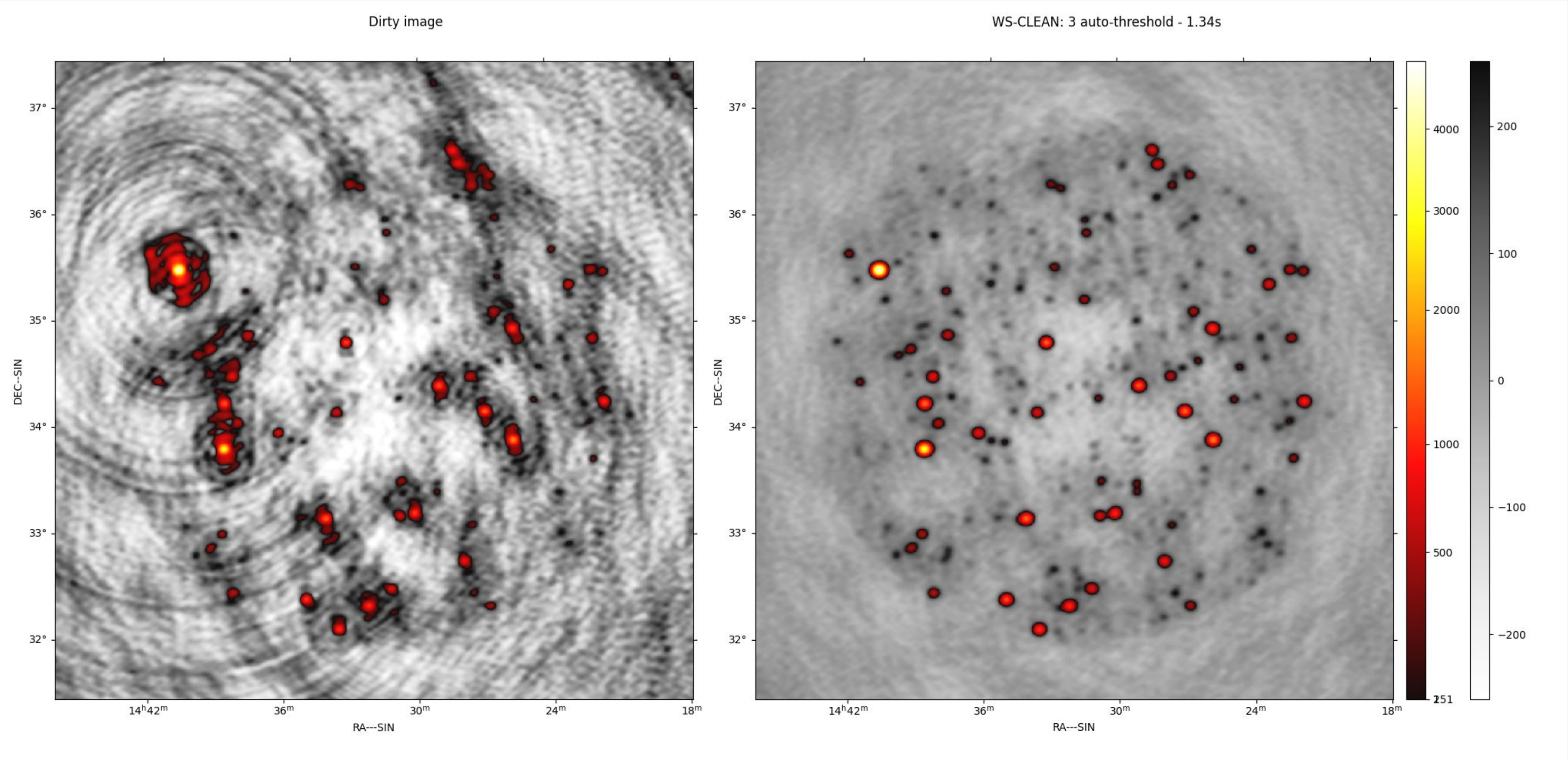
1024 pixels
=
6° FOV

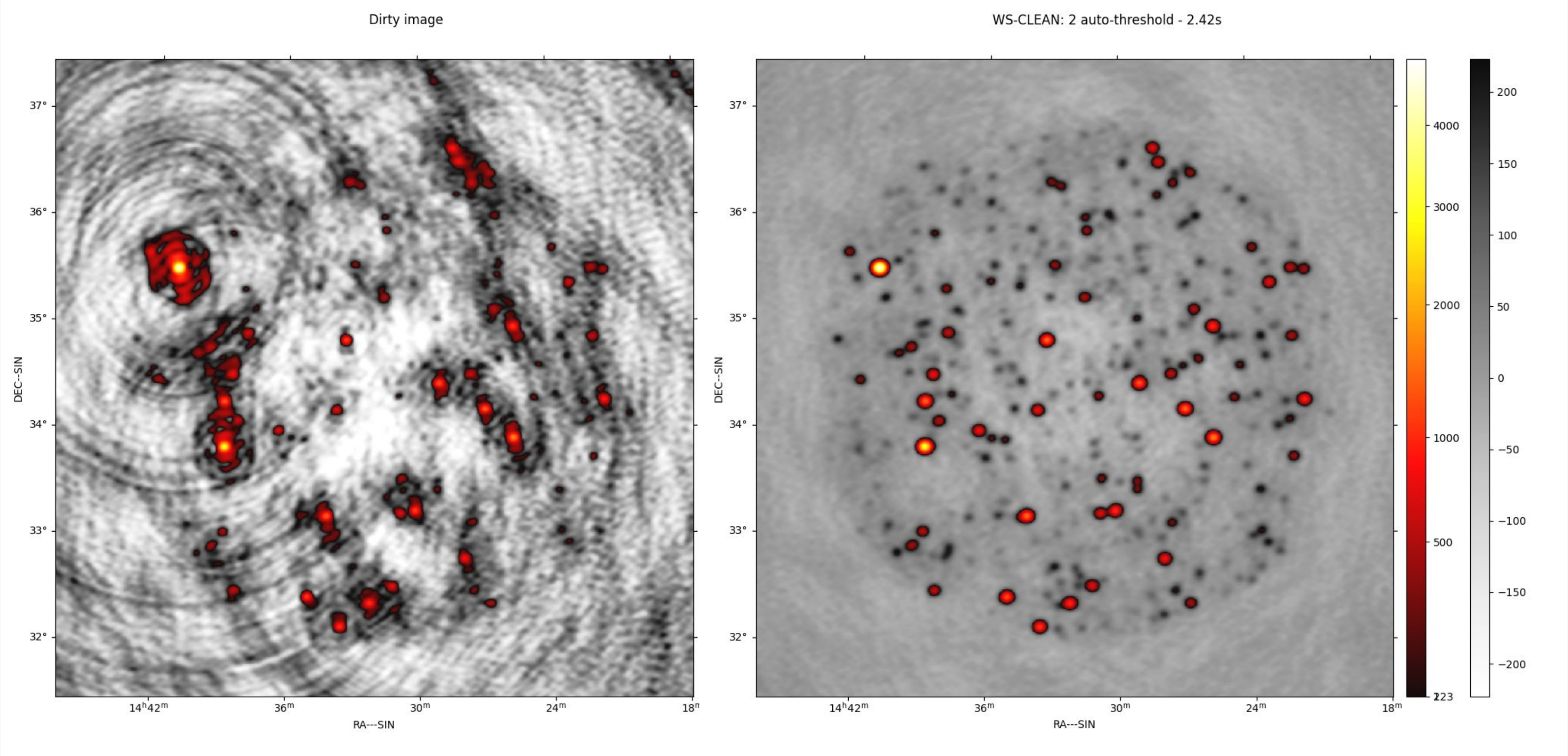


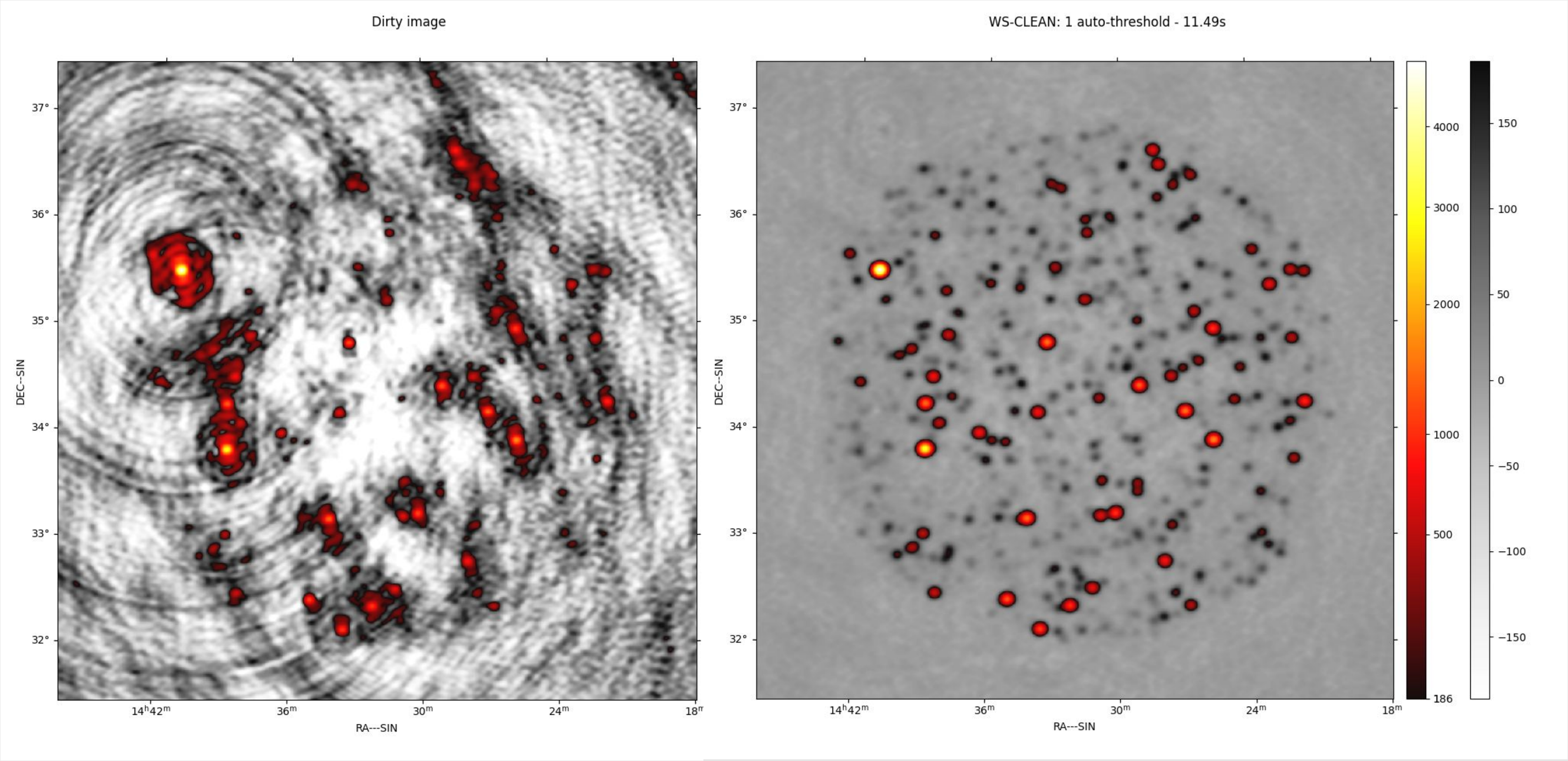

$$\mathbf{I}_D = \mathbf{\Phi}^* \mathbf{V}$$

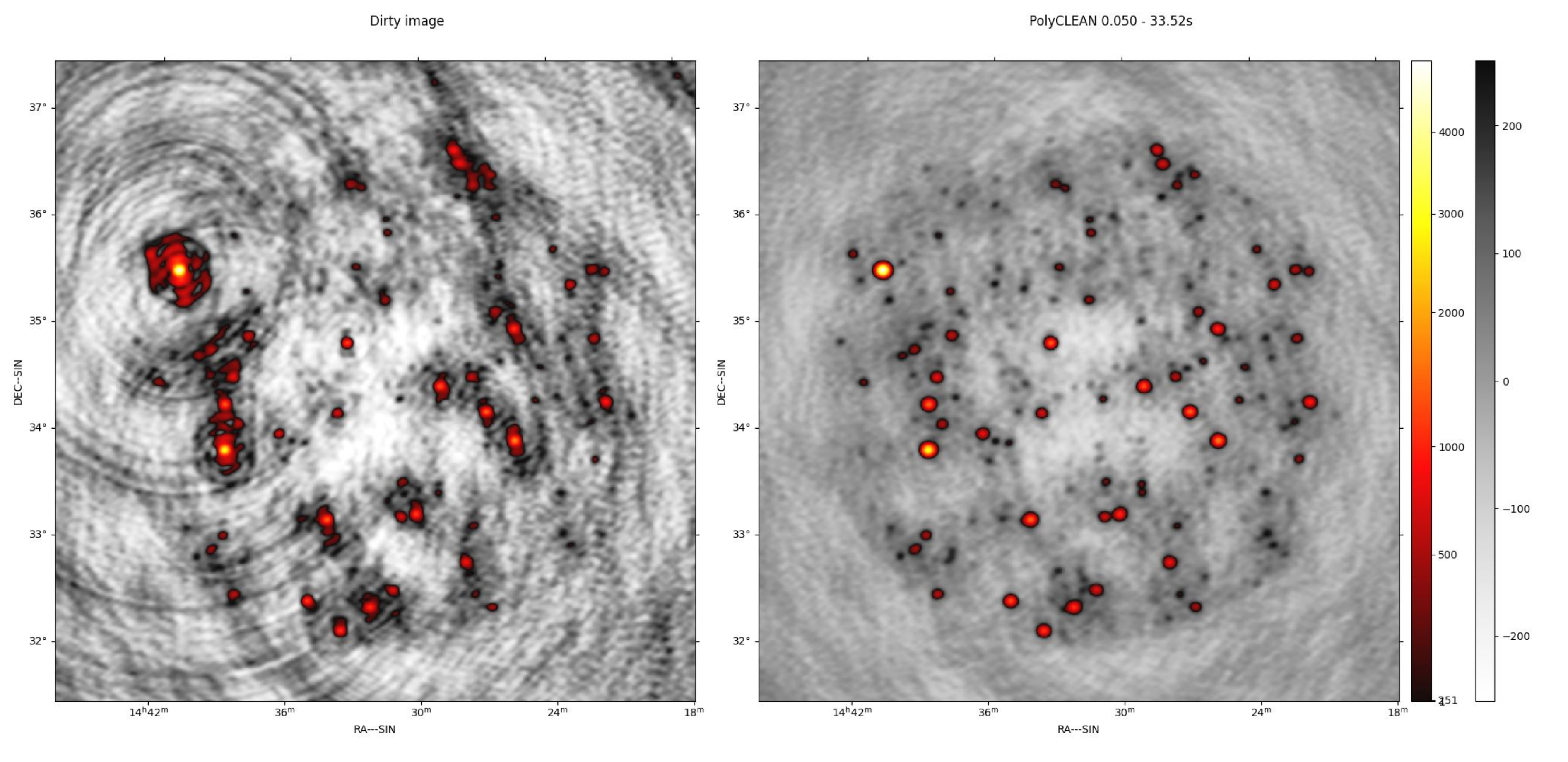
# Reconstruction parameters

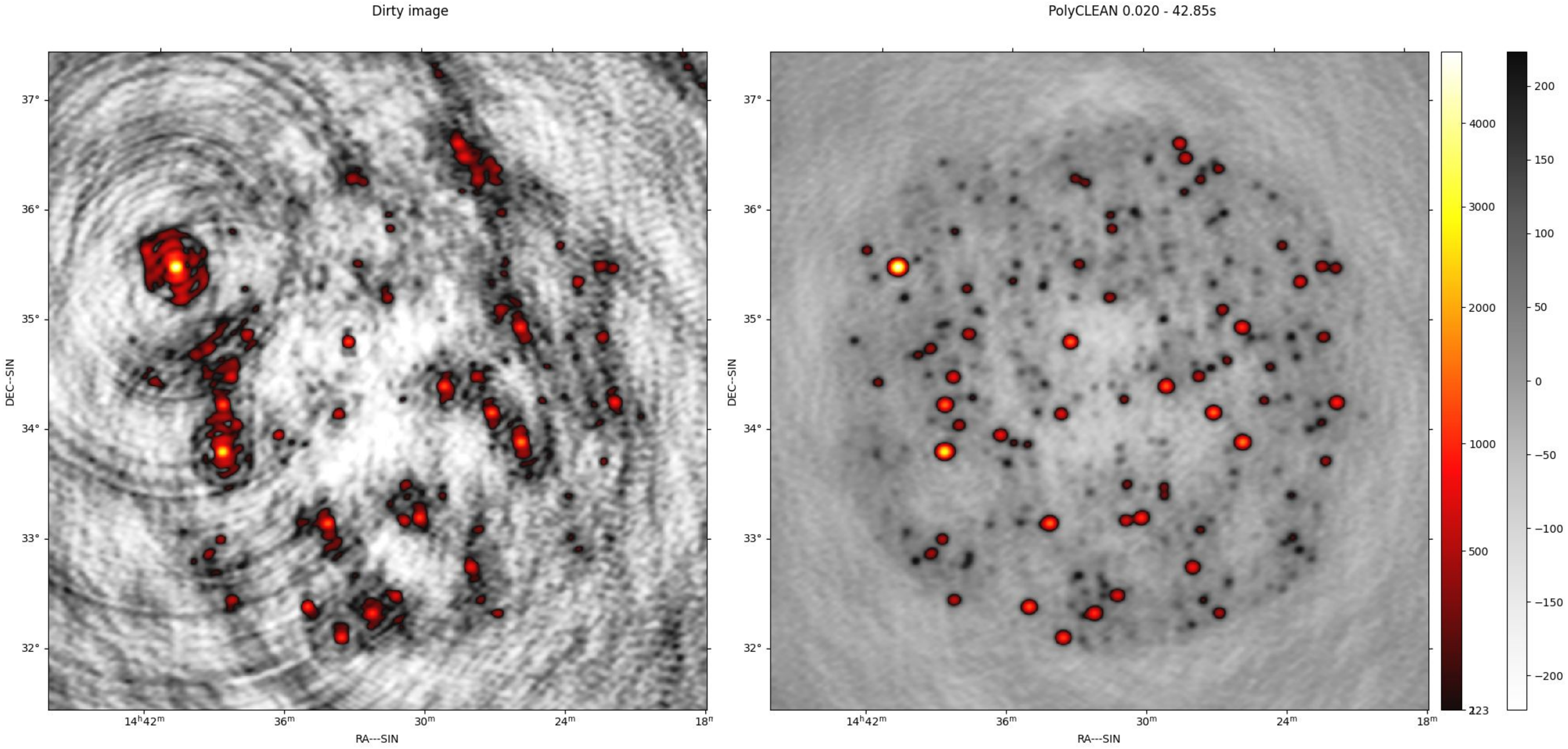
| WS-CLEAN<br>Auto-threshold parameter | PolyCLEAN<br>Penalty parameter |
|---|---|
| 3 σ | 5% |
| 2 σ | 2% |
| 1 σ | 0.5% |


Dirty image
WS-CLEAN: 3 auto-threshold - 1.34s
DEC--SIN
RA---SIN
37°
36°
35°
34°
33°
32°
14h42m
36m
30m
24m
18m
4000
3000
2000
1000
500
251
200
100
0
−100
−200

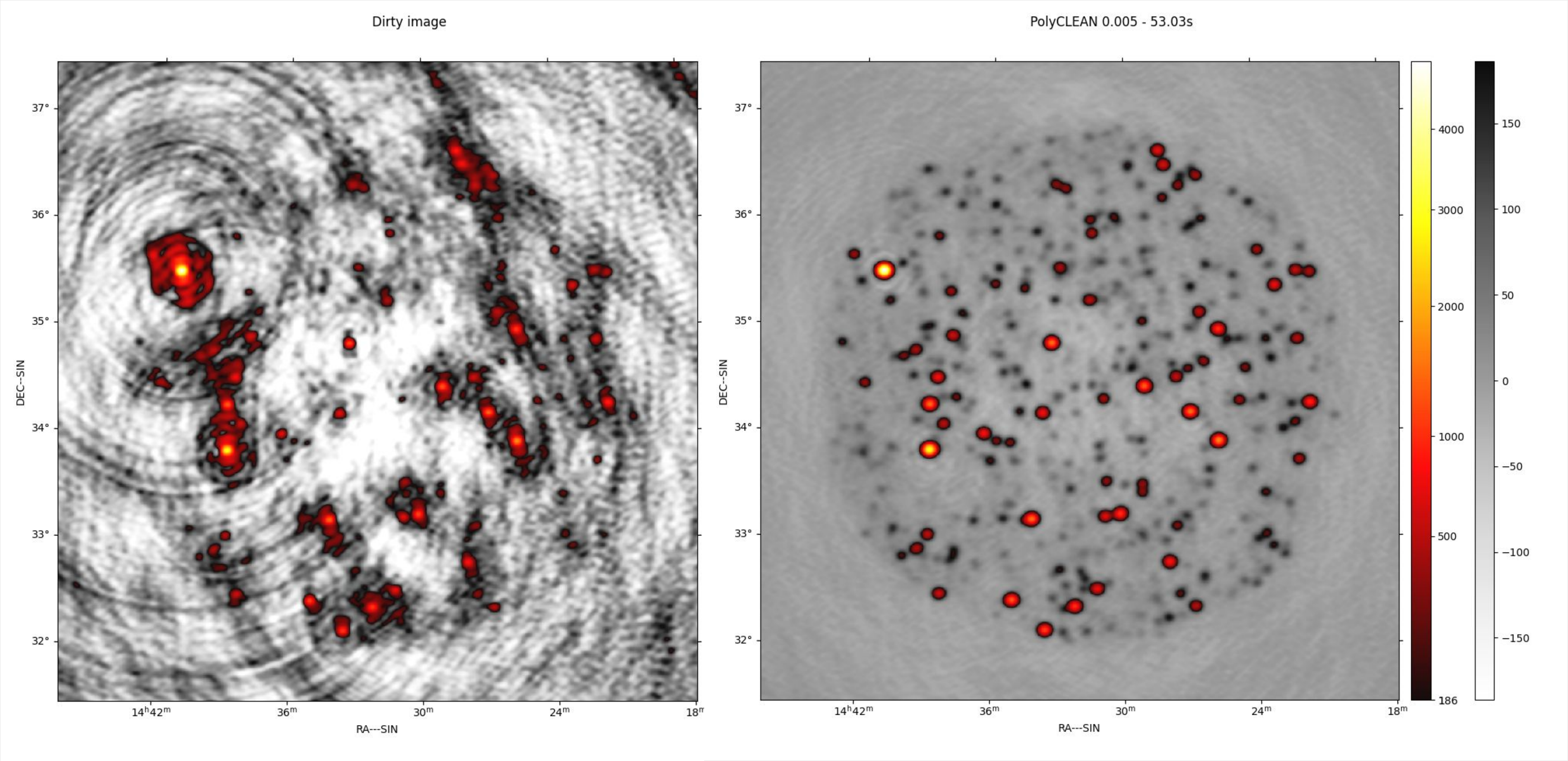

Dirty image
WS-CLEAN: 2 auto-threshold - 2.42s
37°
36°
35°
34°
33°
32°
DEC--SIN
14h42m
36m
30m
24m
18m
RA---SIN
4000
3000
2000
1000
500
223
200
150
100
50
0
−50
−100
−150
−200


Dirty image
WS-CLEAN: 1 auto-threshold - 11.49s
DEC--SIN
RA---SIN
37°
36°
35°
34°
33°
32°
14h42m
36m
30m
24m
18m
4000
3000
2000
1000
500
186
150
100
50
0
−50
−100
−150

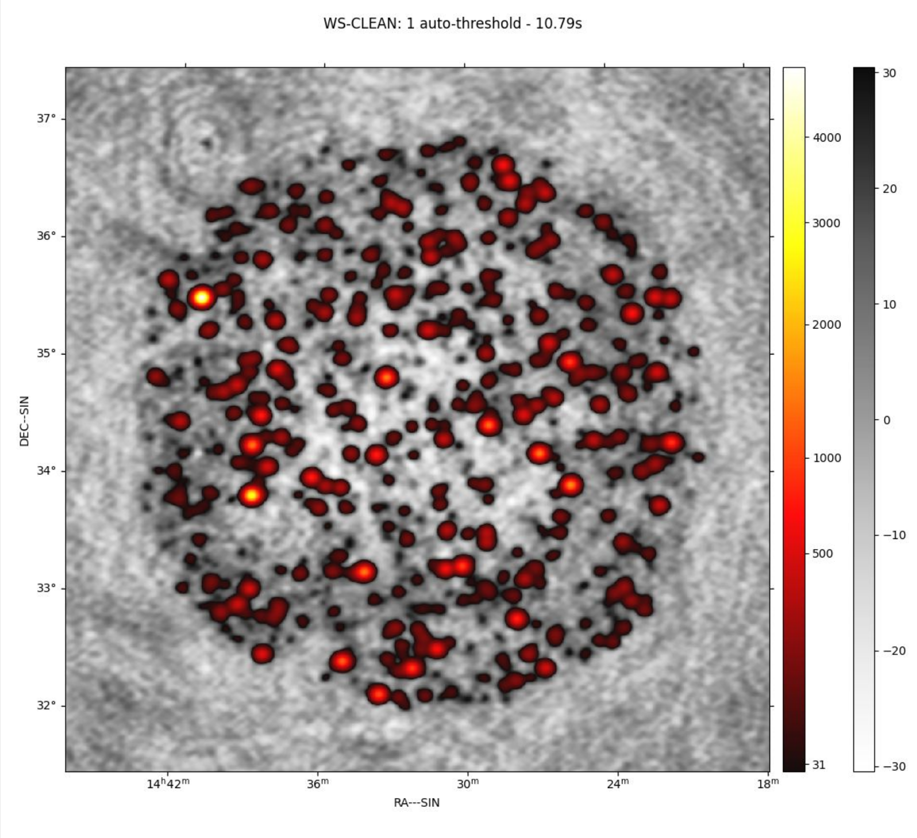

Dirty image
PolyCLEAN 0.050 - 33.52s
DEC--SIN
RA---SIN
37°
36°
35°
34°
33°
32°
14h42m
36m
30m
24m
18m
4000
3000
2000
1000
500
251
200
100
0
−100
−200

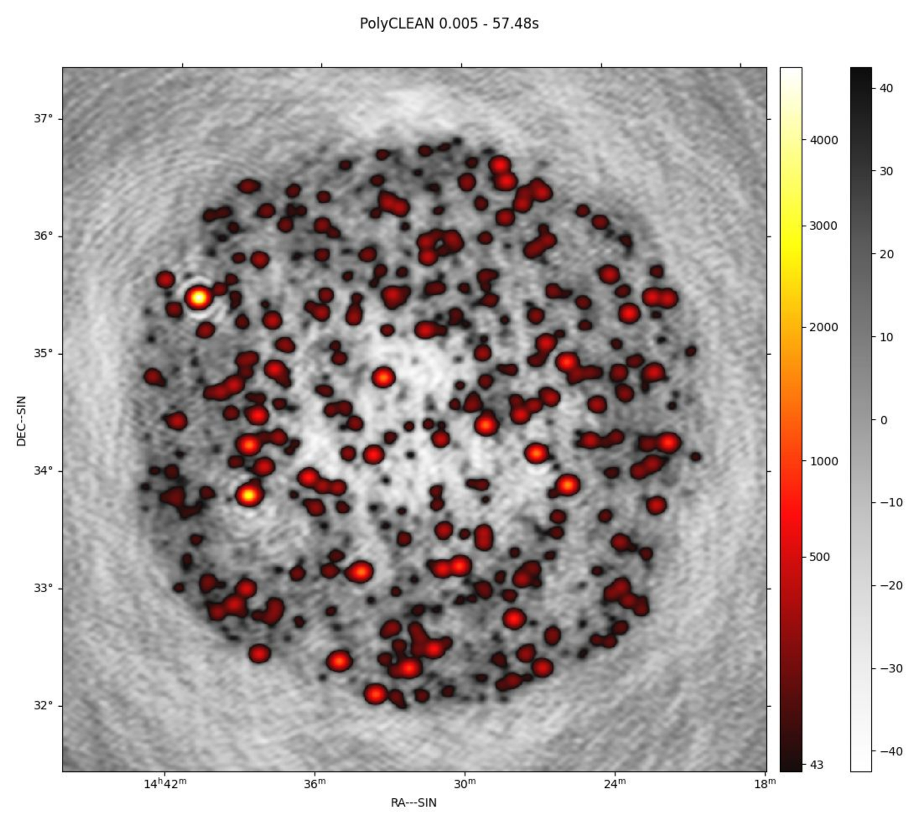


Dirty image
PolyCLEAN 0.020 - 42.85s
DEC--SIN
RA---SIN
37°
36°
35°
34°
33°
32°
14h42m
36m
30m
24m
18m
4000
3000
2000
1000
500
200
150
100
50
0
−50
−100
−150
−200


Dirty image
PolyCLEAN 0.005 - 53.03s
37°
36°
35°
34°
33°
32°
DEC--SIN
14h42m
36m
30m
24m
18m
RA---SIN
4000
3000
2000
1000
500
186
150
100
50
0
−50
−100
−150

# Longest deconvolution example



# Dual certificate

Definition:

$$\mu_\lambda = \frac{1}{\lambda}\mathbf{\Phi}^*(\mathbf{V} - \mathbf{\Phi}\mathbf{I}^*)$$

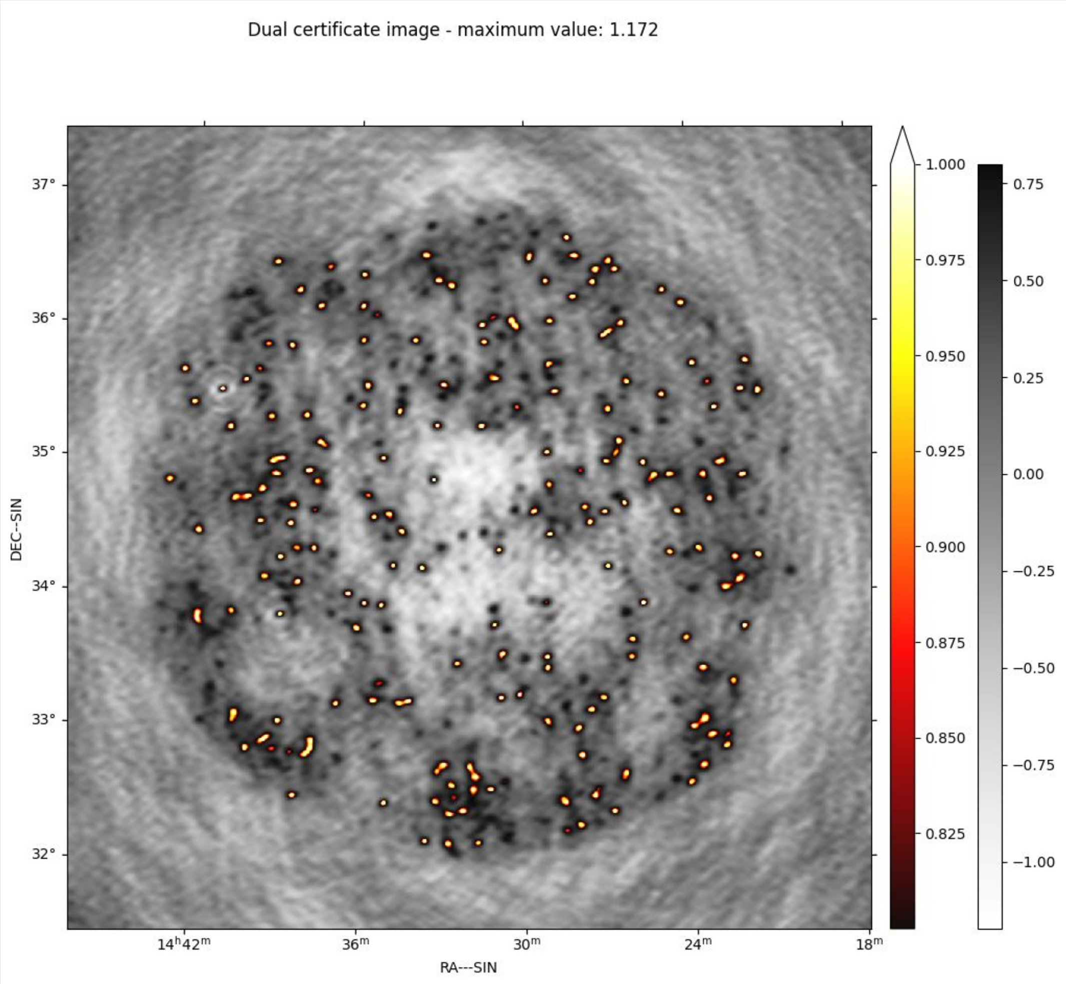
Properties:

$$\|\mu_\lambda\|_\infty \leq 1$$

$$\langle \mu_\lambda, \mathbf{I}^* \rangle = \|\mathbf{I}^*\|_1$$

Usages:

- Convergence
- Saturation set

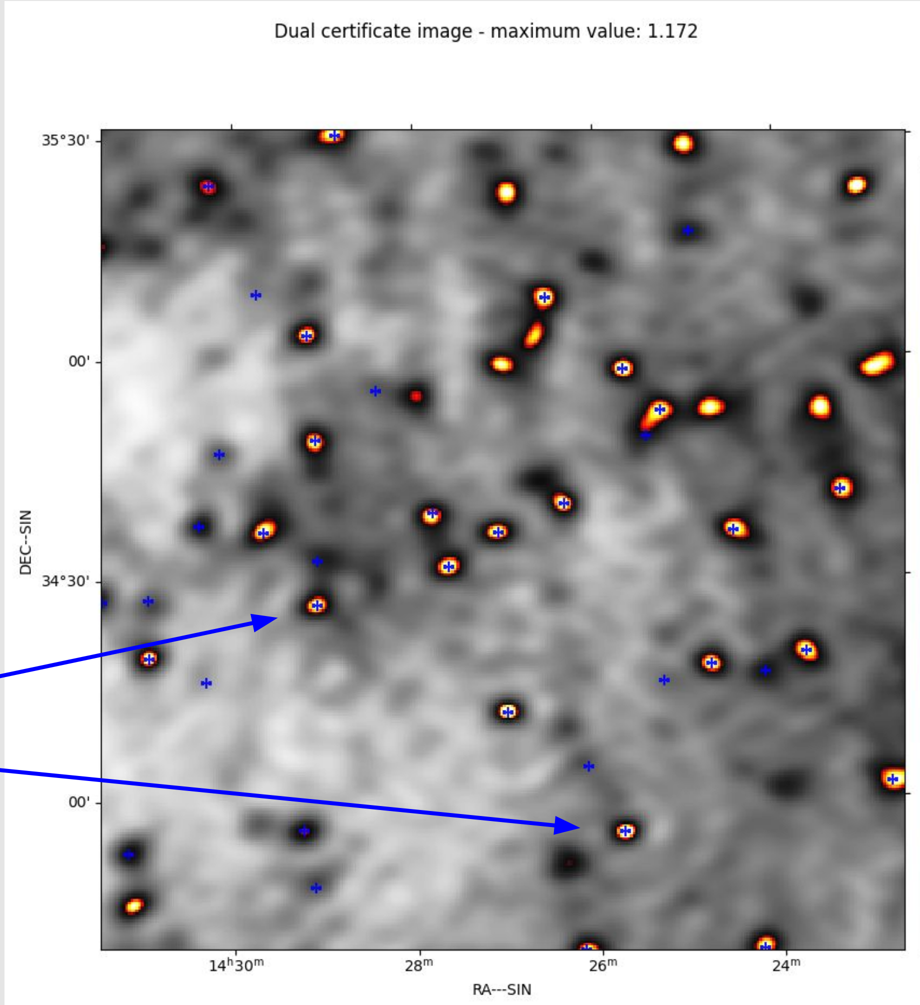


# Dual certificate

Rate of information explained

$$\|\mu_\lambda\|_\infty = \frac{\|\mathbf{\Phi}^*(\mathbf{V}-\mathbf{\Phi}\mathbf{I}^*)\|_\infty}{0.02\ \ \|\mathbf{\Phi}^*\mathbf{V}\|_\infty} \leq 1$$

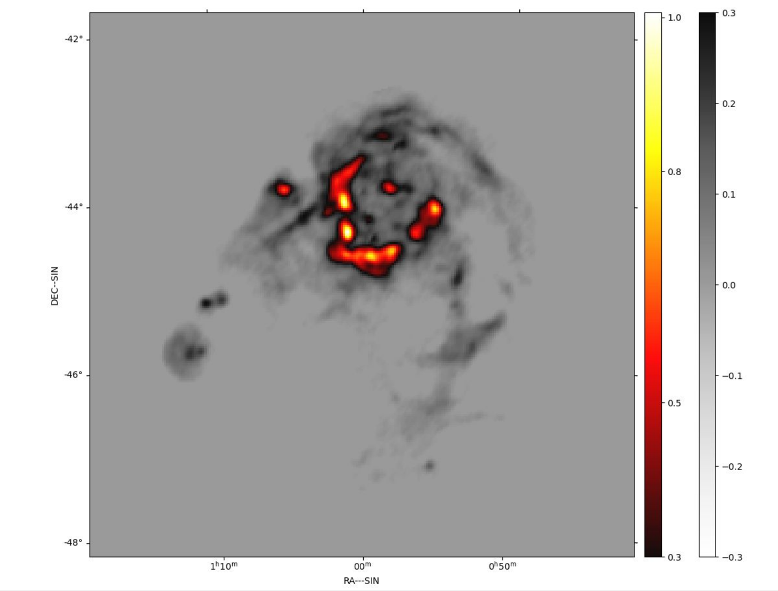
Interpretation of the saturation set



# Extended Sources: Simulations

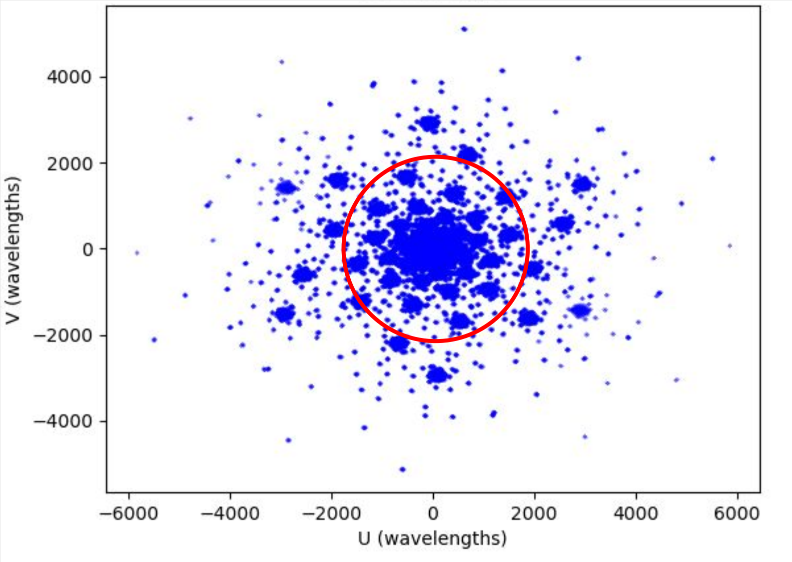
M31 image:

6.5 degrees -> 256 pixels / side




Baselines:

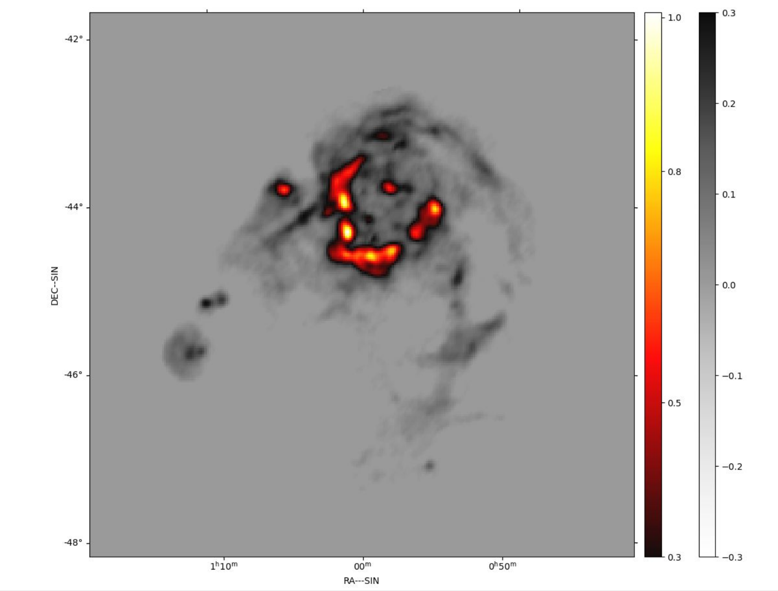
rmax = 1000m -> 31500 baselines



# Extended Sources: Simulations

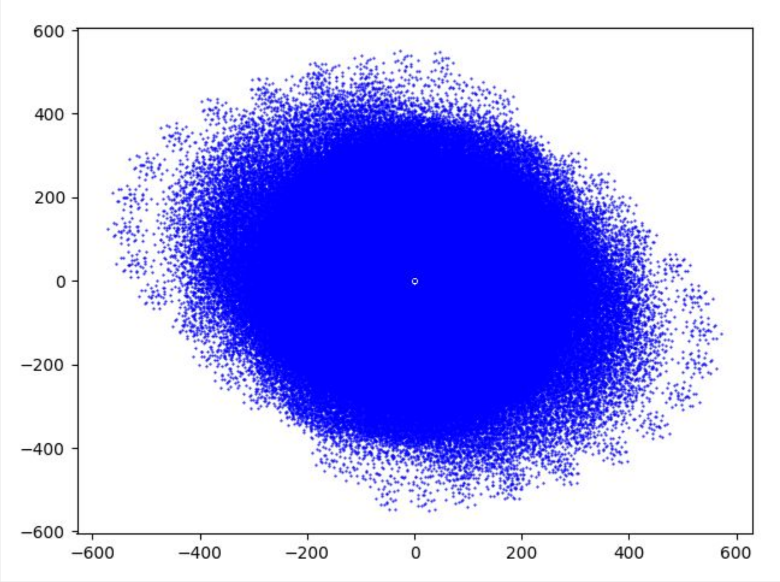
M31 image:

6.5 degrees -> 256 pixels / side




Baselines:

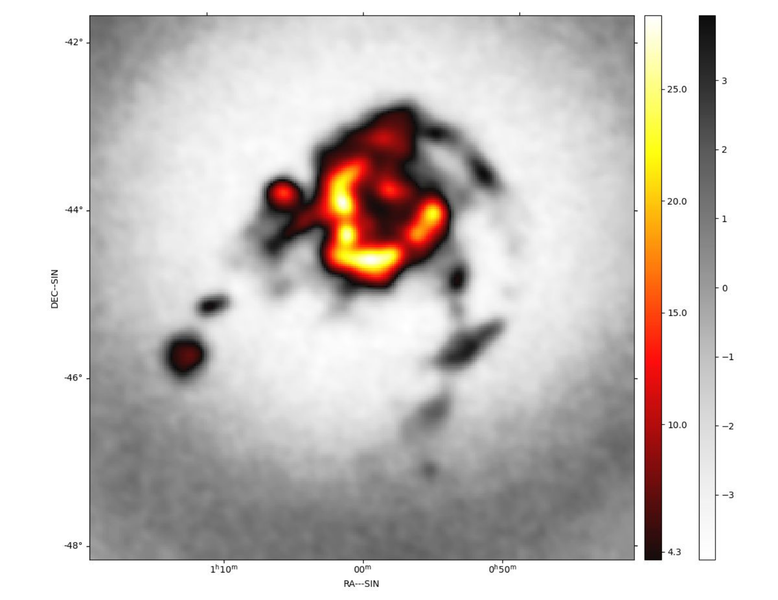
rmax = 1000m -> 31500 baselines,
**11 measurement times**



# Extended Sources: Simulations

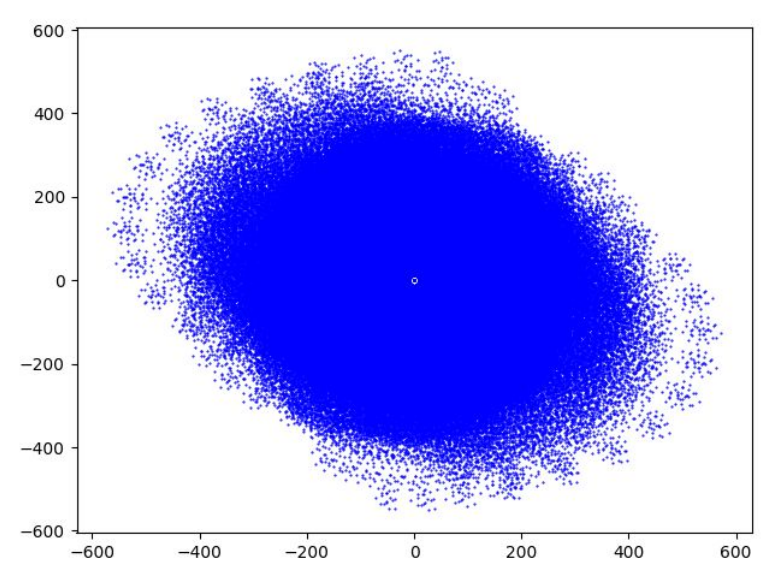
M31 image:

6.5 degrees -> 256 pixels / side

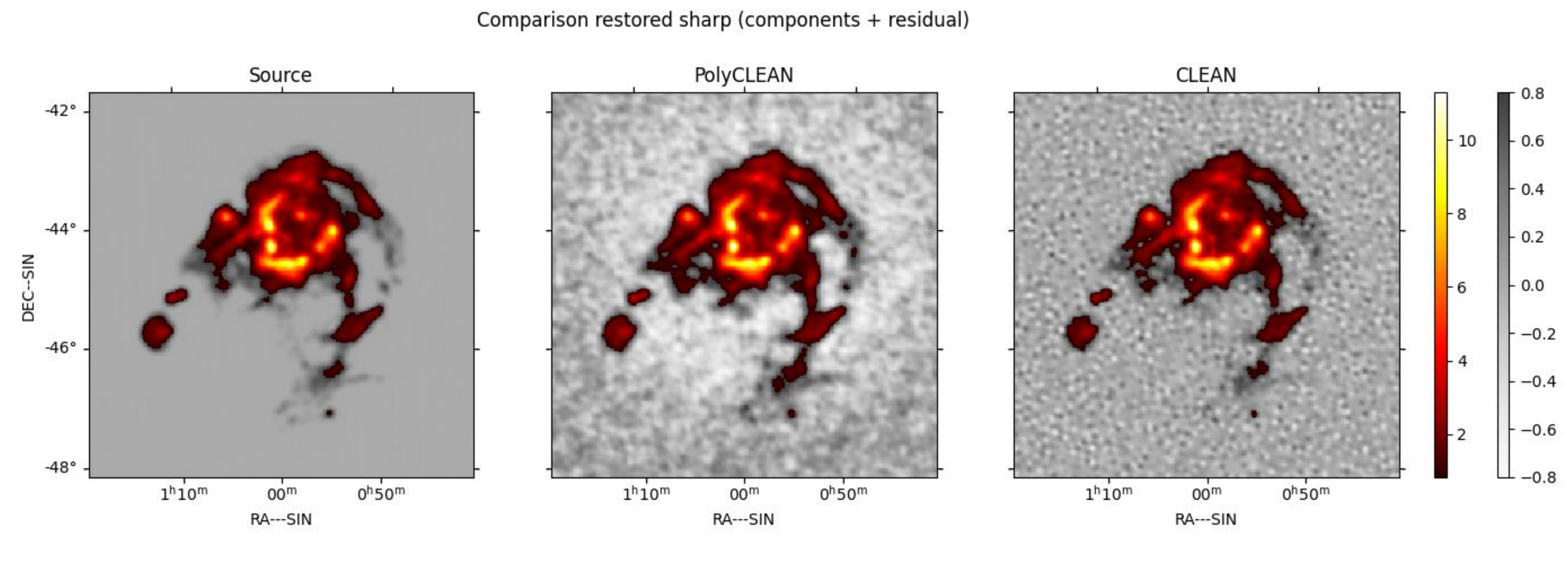



Baselines:

rmax = 1000m -> 31500 baselines,
**11 measurement times**

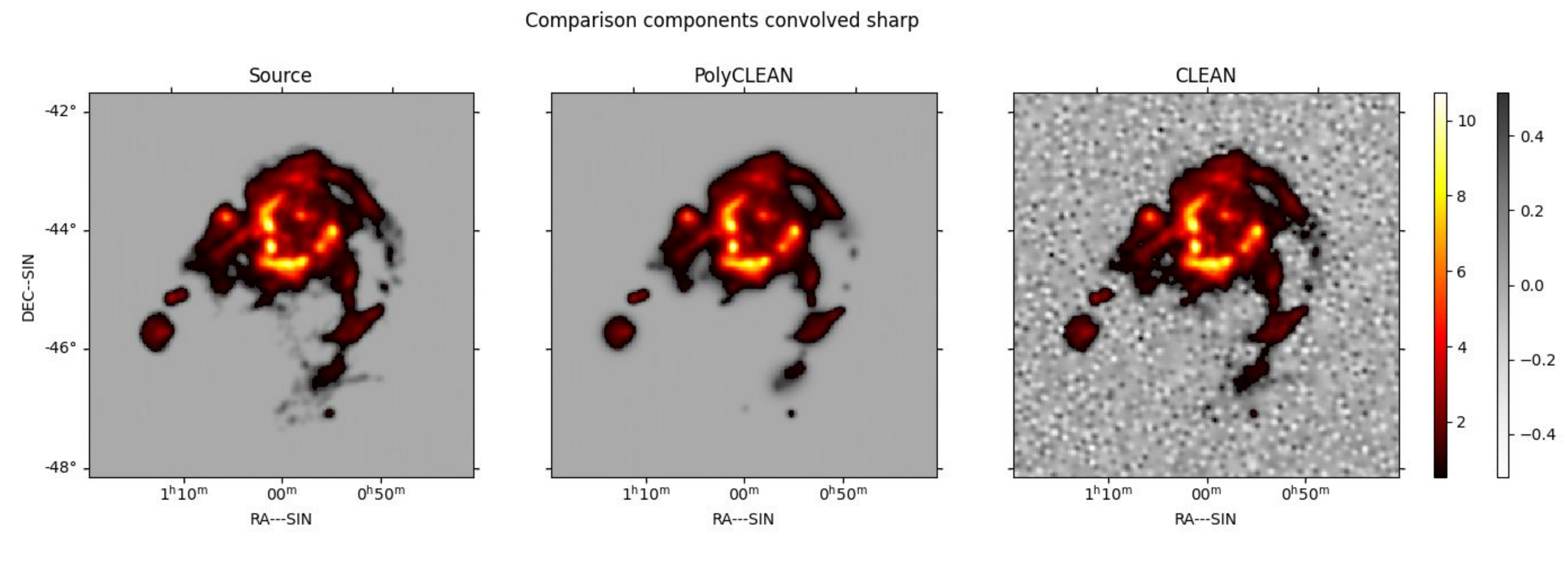


# Reconstructions



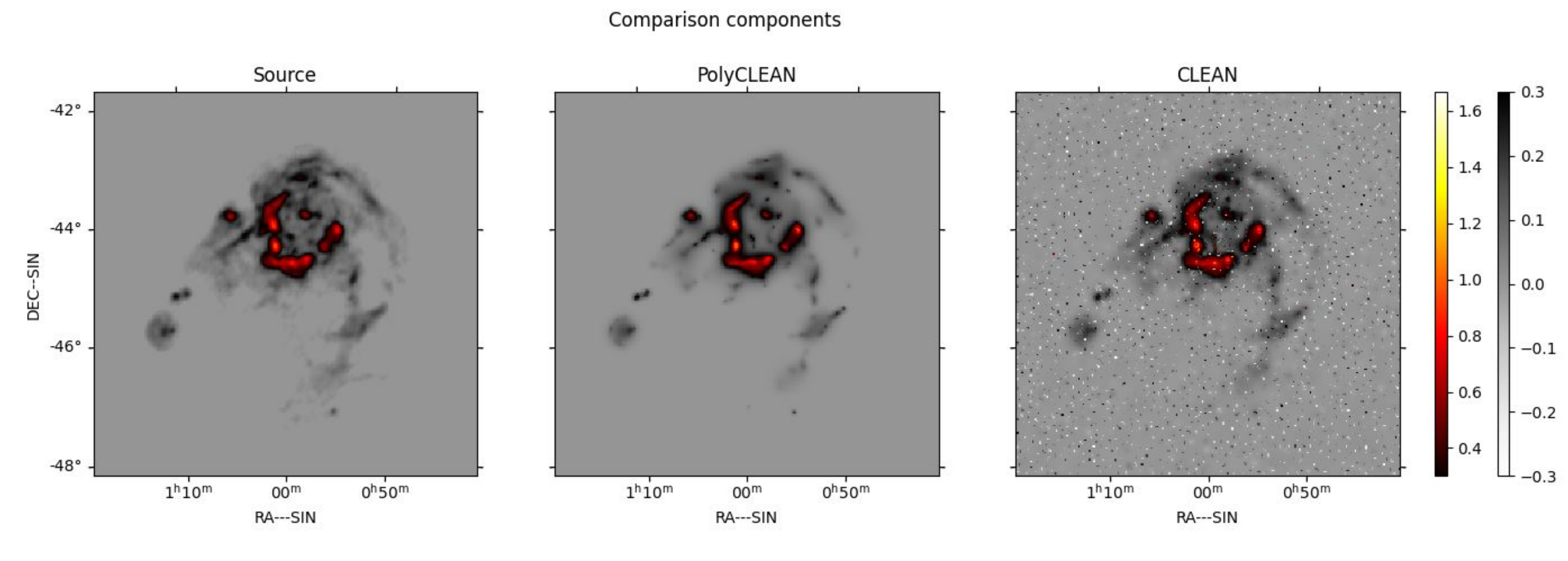

| | PolyCLEAN | CLEAN |
|---|---|---|
| MSE: | 5.5 x 10^-2 | 1.4 x 10^-2 |
| MAD: | 1.9 x 10^-1 | 8.9 x 10^-2 |

# Reconstructions (without res.)




| | PolyCLEAN | CLEAN |
|---|---|---|
| MSE: | 1.2 x 10^-2 | 1.0 x 10^-2 |
| MAD: | 4.6 x 10^-2 | 7.5 x 10^-2 |

# Reconstructions (model)




| | PolyCLEAN | CLEAN |
|---|---|---|
| MSE: | 1.5 x 10^-4 | 1.2 x 10^-3 |
| MAD: | 4.3 x 10^-3 | 1.2 x 10^-2 |

# Summary

## 1. Numerical performance

- Scalability achievement
- Optimization method up to speed with atomic method
- Sparsity-based method

## 2. Versatility

- Adaptable:
  - Tune the parameters
  - Fine control
- Point and extended sources

## 3. Ongoing research work

- **Dual certificate**: promising new scientific tool for RI image reconstruction
- Room for improvement in the code as well as in the algorithm
- Bayes estimation of the parameters

# Thanks!

| | CLEAN | MAP Estimation | PolyCLEAN |
|---|---|---|---|
| Sparse iterates | ✅ | ❌ | ✅ |
| Flexible priors | ~ | ✅ | ~ |
| Fast solvers | ✅ | ~ | ✅ |
| Calibration compliant | ✅ | ❌ | ✅ |
| Interpretable obj. function | ❌ | ✅ | ✅ |

# Multi-Scales dictionaries

Redundant dictionary:

$$\boxed{\boldsymbol{\Psi} = [\boldsymbol{\Psi}_1 \dots \boldsymbol{\Psi}_D]} \quad \in \mathbb{R}^{N \times DN}$$

# Multi-Scales dictionaries

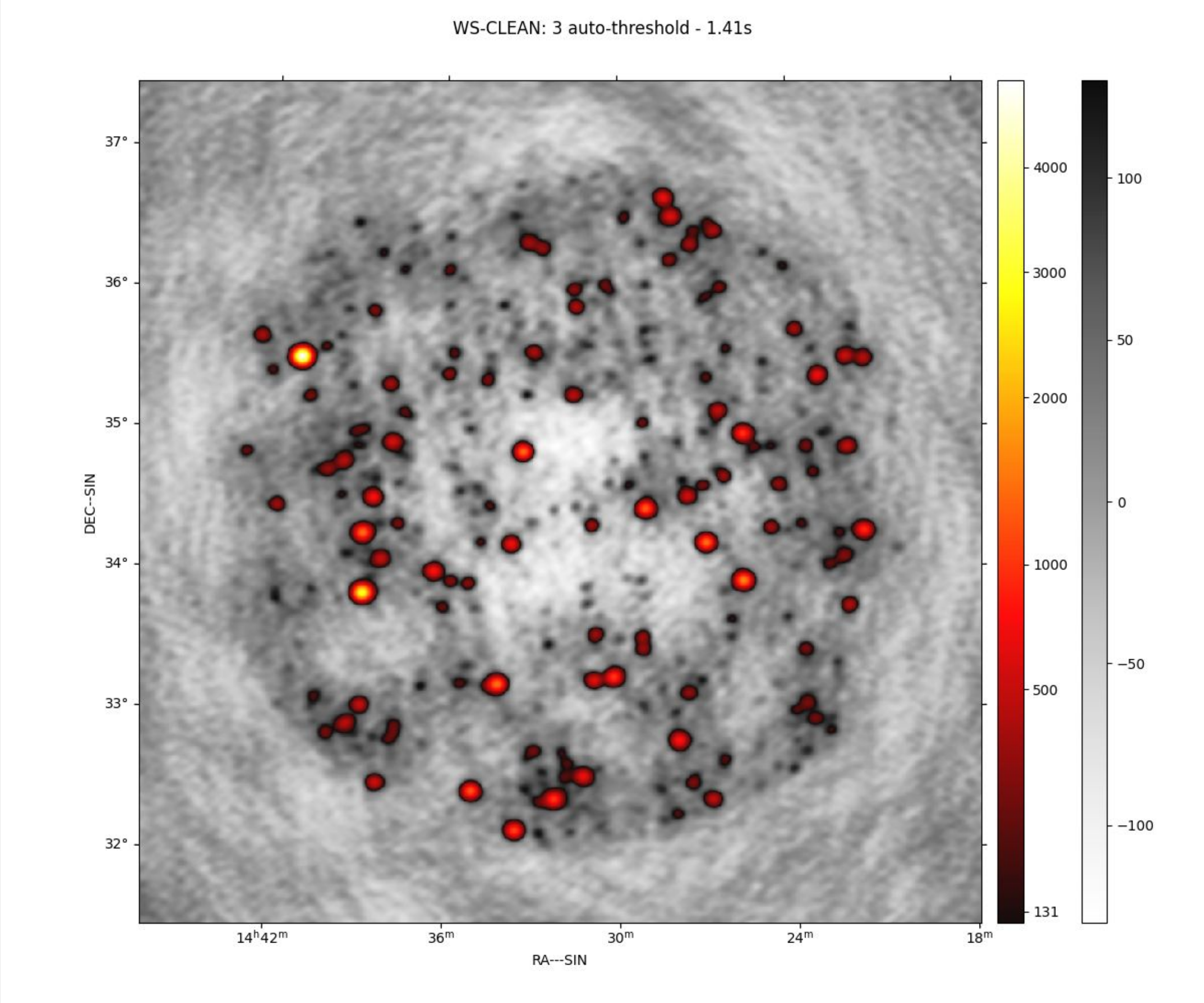
Redundant dictionary:

$$\boxed{\boldsymbol{\Psi} = [\boldsymbol{\Psi}_1 \ldots \boldsymbol{\Psi}_D]} \quad \in \mathbb{R}^{N \times DN}$$

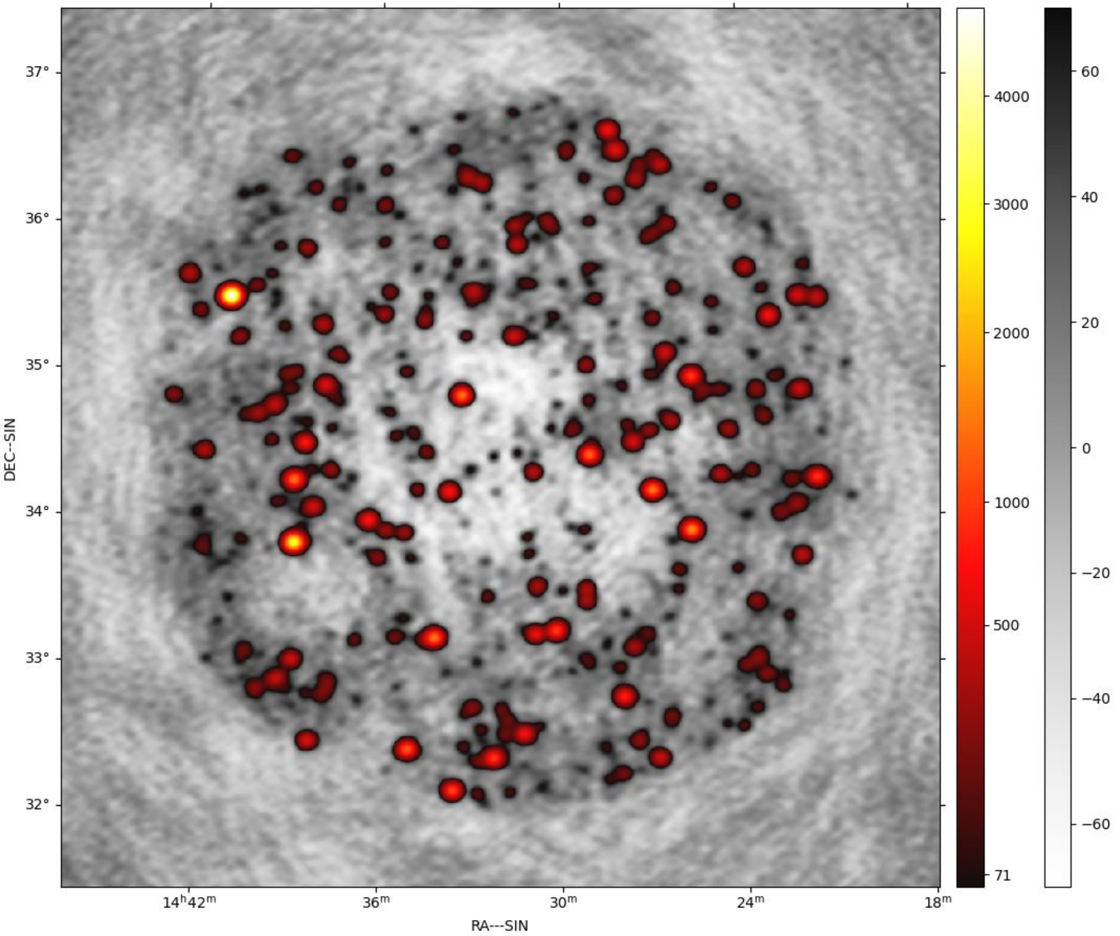
- Wavelets: $$\boldsymbol{\Psi}_d = \mathbf{W}_d \qquad \in \mathbb{R}^{N \times N}$$
- Gaussian kernels: $$\boldsymbol{\Psi}_d \boldsymbol{\theta}_d = \boldsymbol{\theta}_d * \mathbf{g}_{\sigma_d} \qquad \in \mathbb{R}^{N \times N}$$

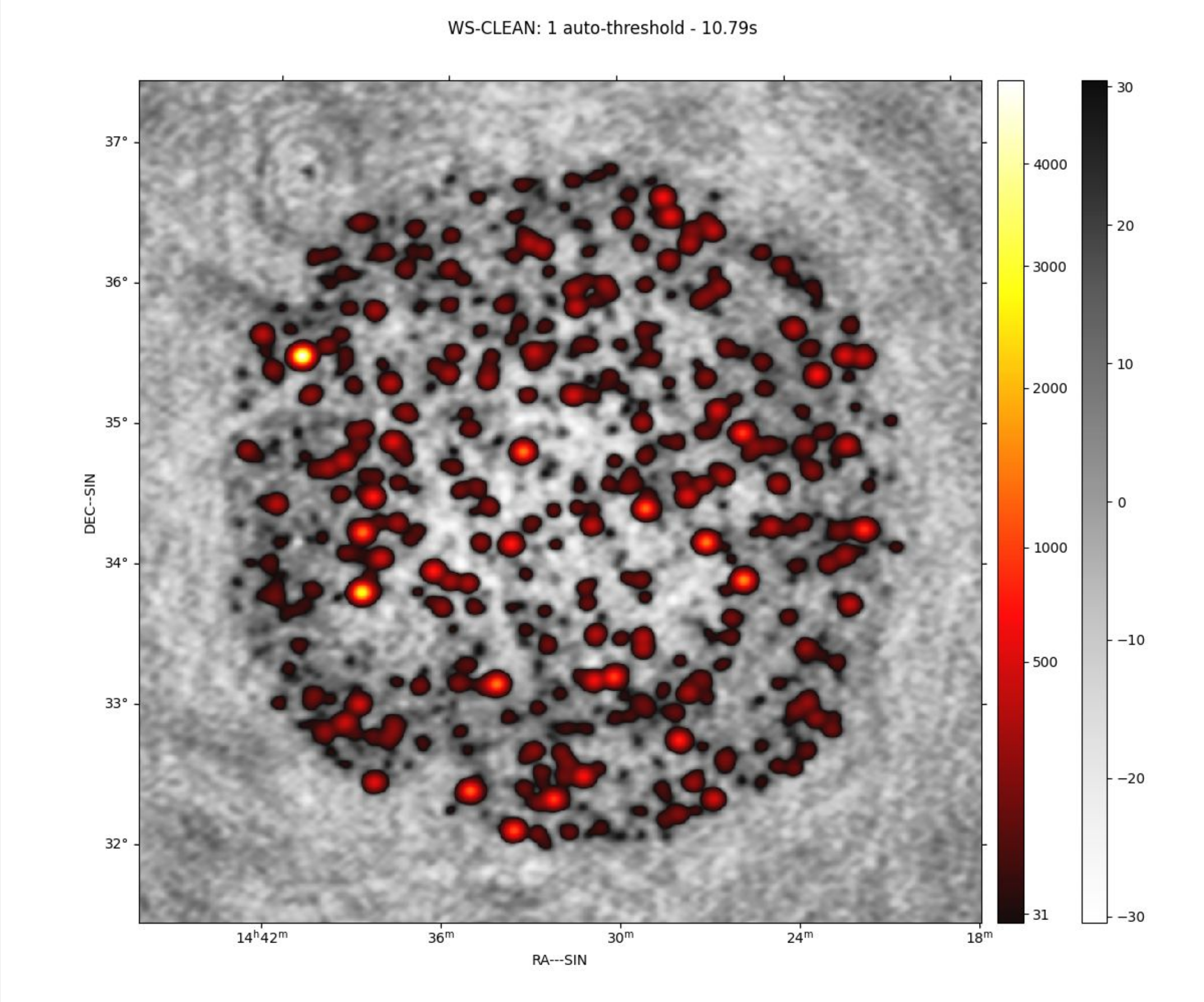
# A Frank-Wolfe Solver

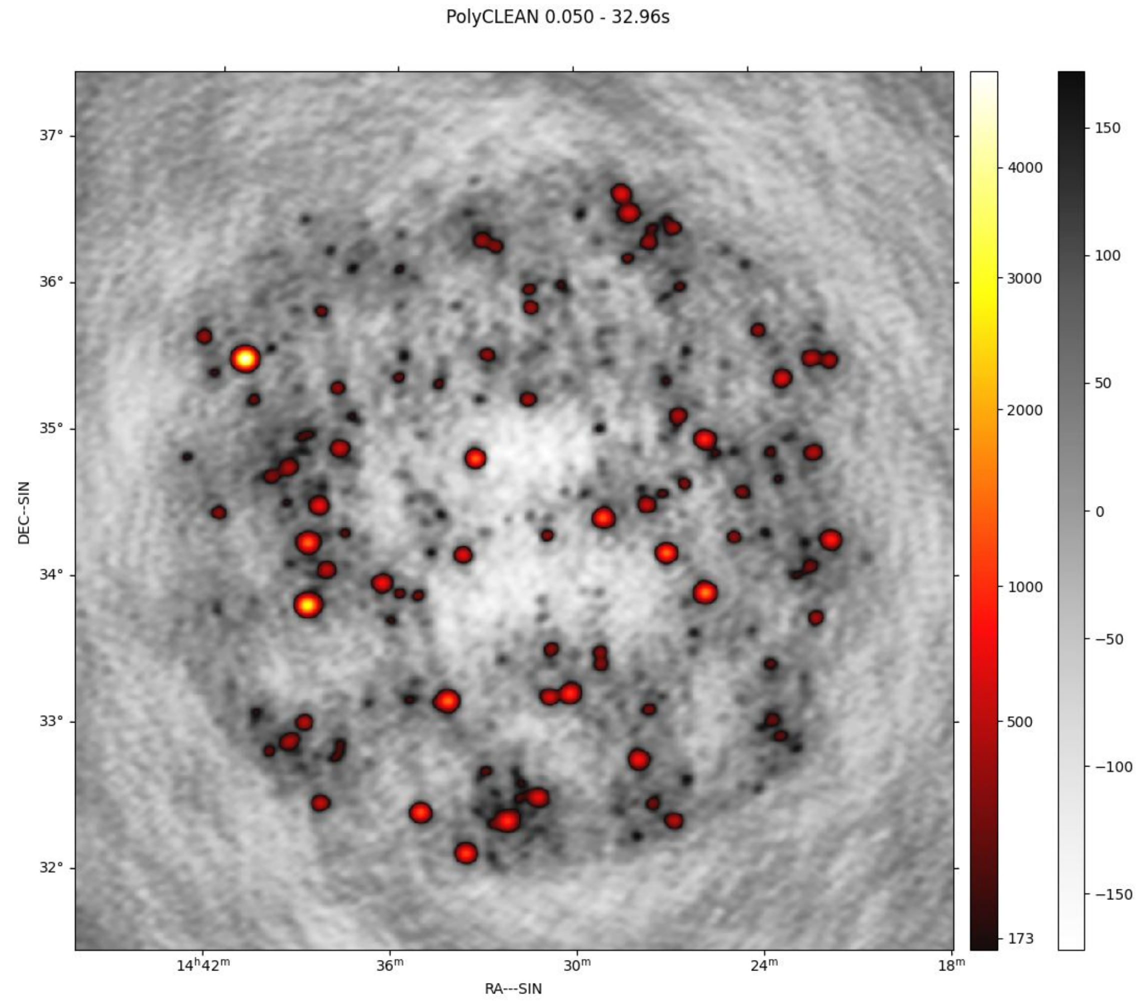
- Convex optimization method:
  - → Convergence guarantee
- Frank-Wolfe for atomic norm:
  - → Atomic behavior
  - → Sparse iterates
- Polyatomic variation:
  - → Fast solver

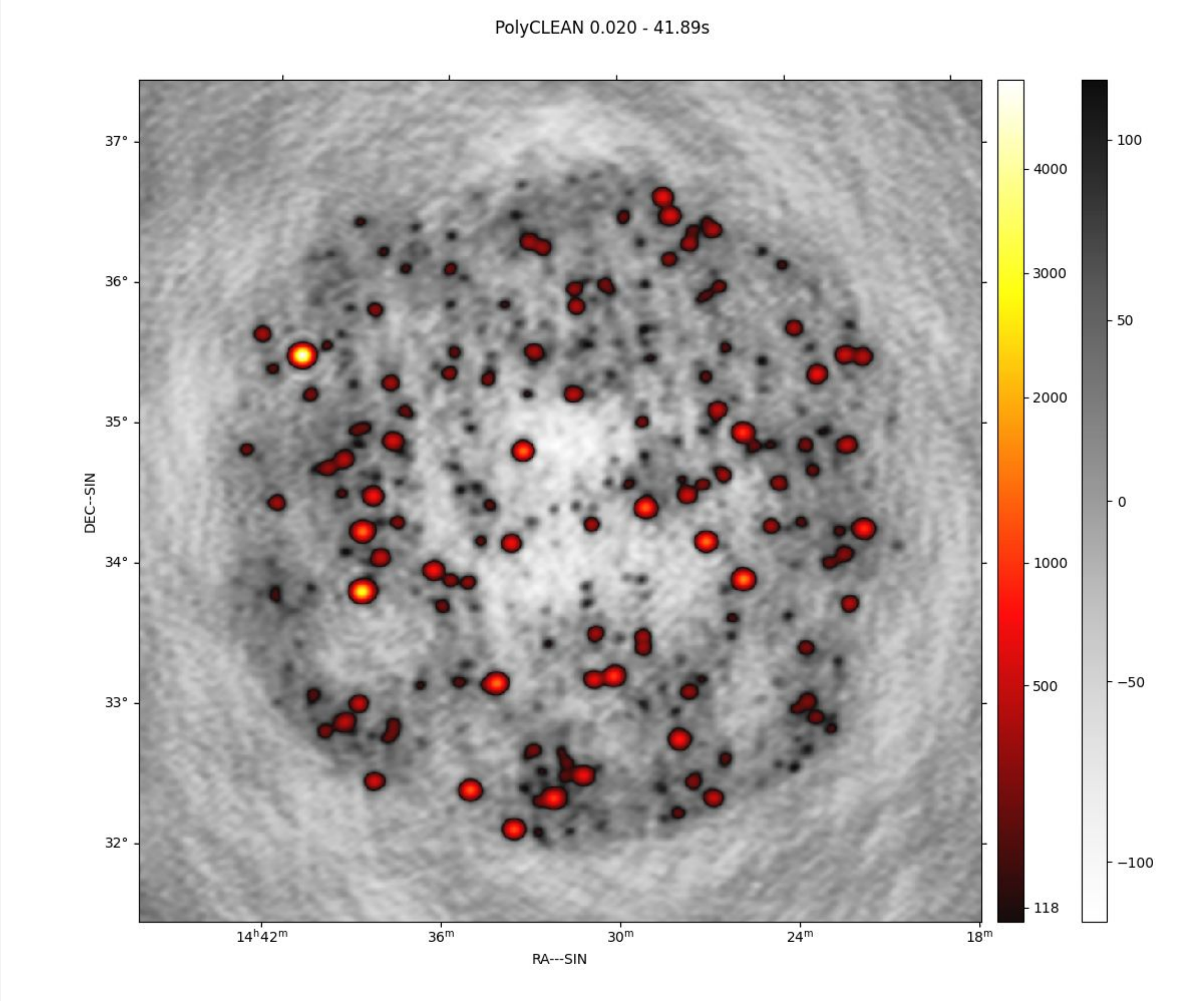

WS-CLEAN: 3 auto-threshold - 1.41s
37°
36°
35°
34°
33°
32°
DEC--SIN
14h42m
36m
30m
24m
18m
RA---SIN
4000
3000
2000
1000
500
131
100
50
0
−50
−100

WS-CLEAN: 2 auto-threshold - 2.21s

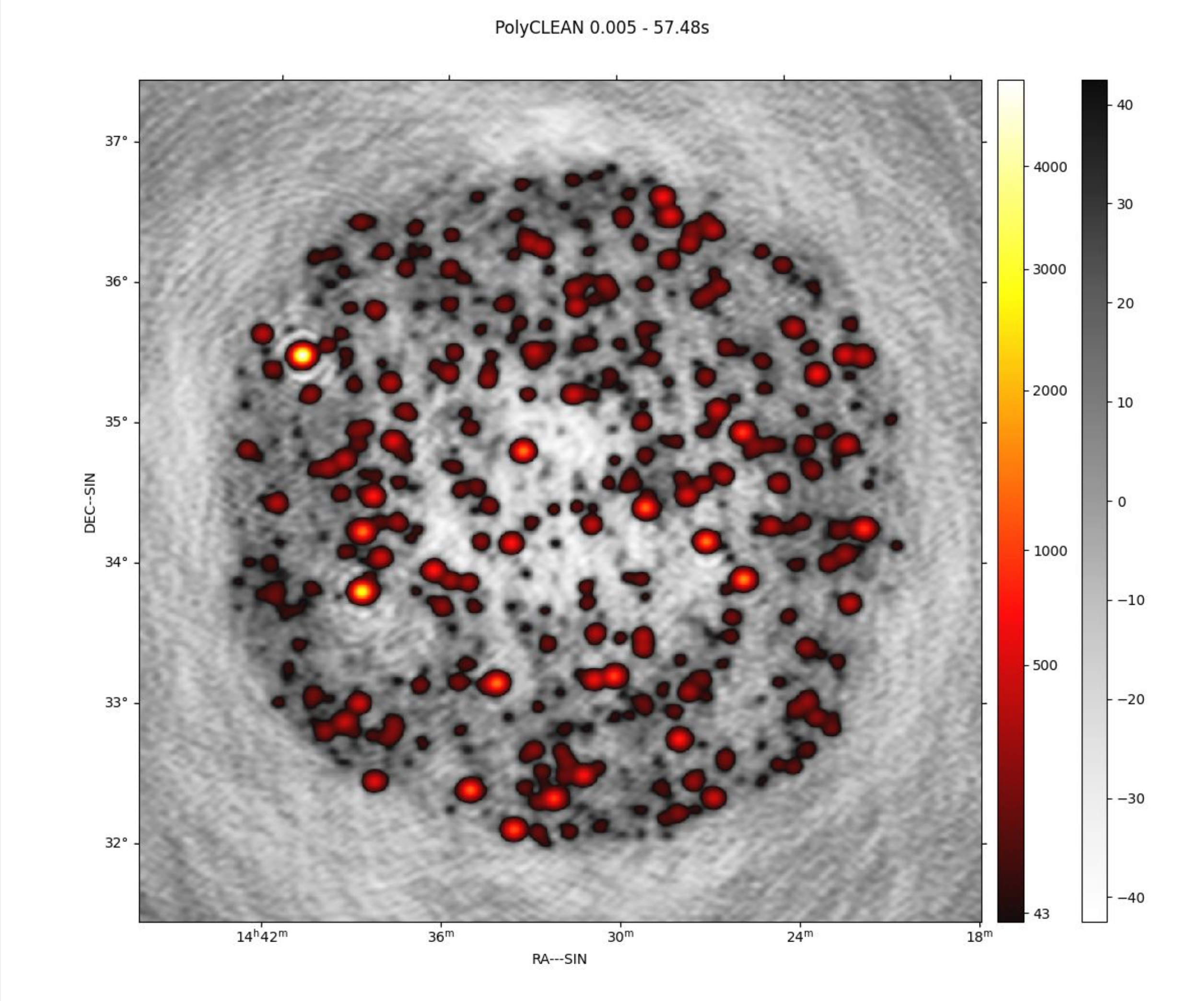


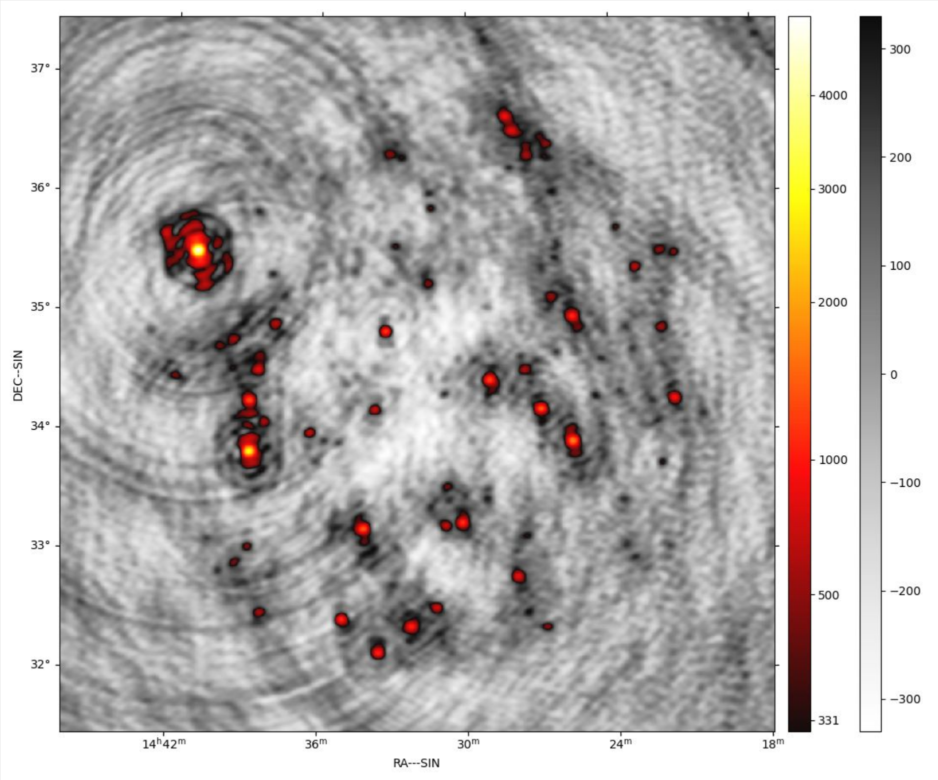

WS-CLEAN: 1 auto-threshold - 10.79s
DEC--SIN
RA---SIN
37°
36°
35°
34°
33°
32°
14h42m
36m
30m
24m
18m
4000
3000
2000
1000
500
31
30
20
10
0
−10
−20
−30


PolyCLEAN 0.050 - 32.96s
37°
36°
35°
34°
33°
32°
DEC--SIN
14h42m
36m
30m
24m
18m
RA---SIN
4000
3000
2000
1000
500
173
150
100
50
0
−50
−100
−150

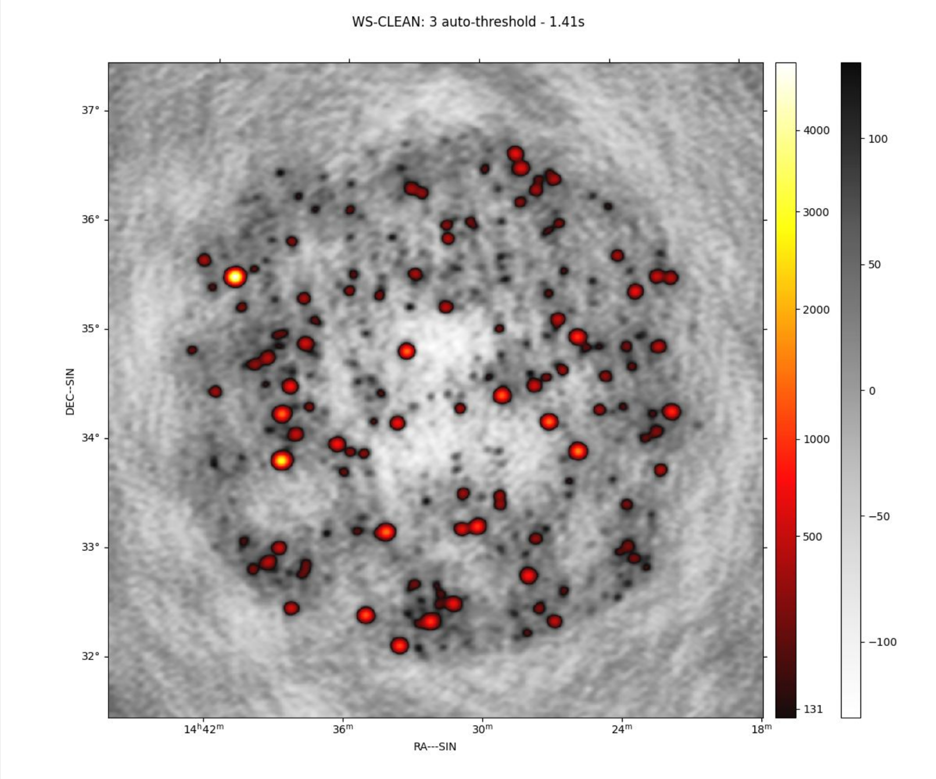


PolyCLEAN 0.020 - 41.89s
37°
36°
35°
34°
33°
32°
DEC--SIN
14h42m
36m
30m
24m
18m
RA---SIN
4000
3000
2000
1000
500
118
100
50
0
−50
−100


PolyCLEAN 0.005 - 57.48s
37°
36°
35°
34°
33°
32°
DEC--SIN
14h42m
36m
30m
24m
18m
RA---SIN
4000
3000
2000
1000
500
43
40
30
20
10
0
−10
−20
−30
−40

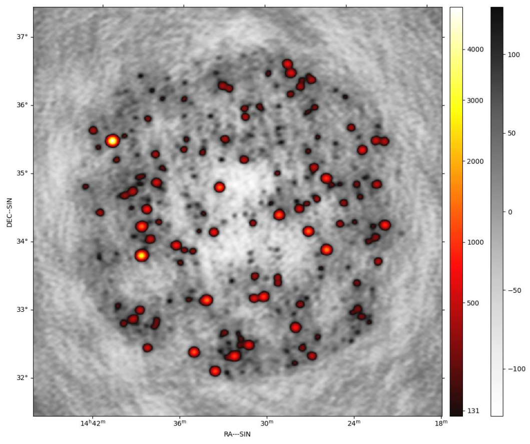


DEC--SIN
RA---SIN
37°
36°
35°
34°
33°
32°
14h42m
36m
30m
24m
18m
4000
3000
2000
1000
500
331
300
200
100
0
−100
−200
−300

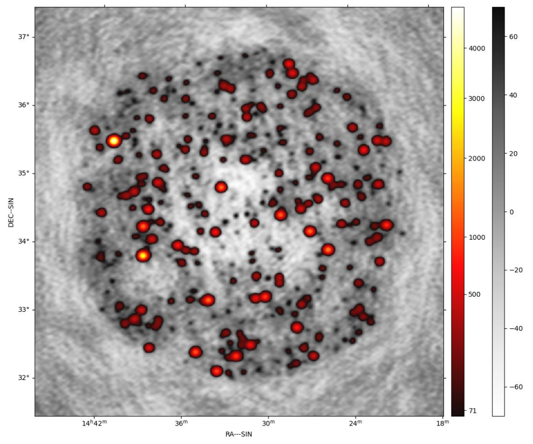



WS-CLEAN: 3 auto-threshold - 1.41s
DEC--SIN
RA---SIN
37°
36°
35°
34°
33°
32°
14h42m
36m
30m
24m
18m
4000
3000
2000
1000
500
131
100
50
0
−50
−100

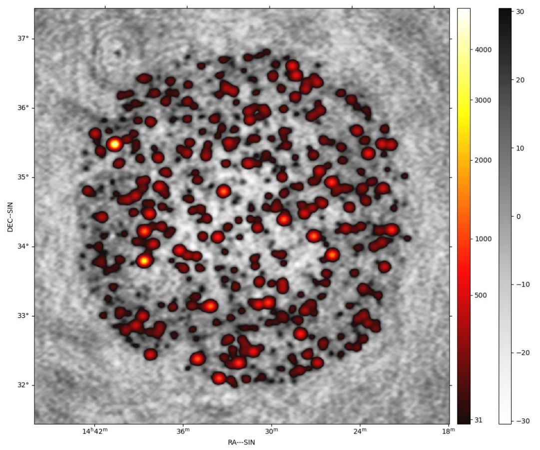
WS-CLEAN: 3 auto-threshold - 1.41s



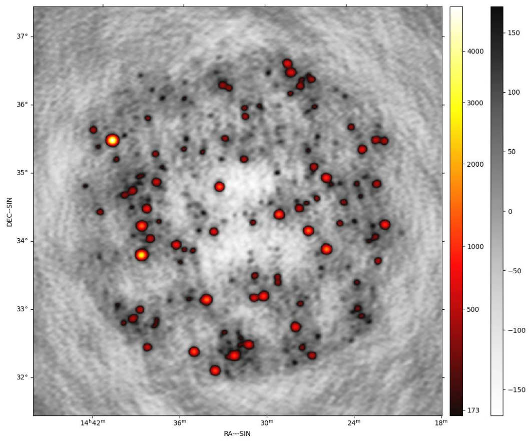

WS-CLEAN: 2 auto-threshold - 2.21s



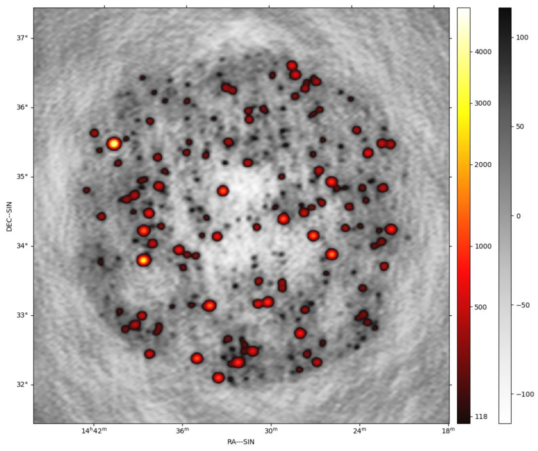

WS-CLEAN: 1 auto-threshold - 10.79s




PolyCLEAN 0.050 - 32.96s




PolyCLEAN 0.020 - 41.89s




PolyCLEAN 0.005 - 57.48s