Observatoire de la Côte d’Azur : Introduction to the AADC Consortia

Chiara Ferrari

in collaboration with: A. Ferrari, D. Mary, J. Deguignet, M. Vannier (OCA)
& participants to ANR Project “MAGELLAN”

Observatoire de la CÔTE d'AZUR

LAGRANGE

Université Nice Sophia Antipolis

UNIVERSITÉ CÔTE D’AZUR

Observatoire de la Côte d'Azur (OCA)

- Earth and Space Science research Institute
  Director: Dr. Thierry LANZ (previously at NASA APD)
- 450 employees organised around three poles
  ✓Geophysics and Space geodesy *(Géoazur)*
  ✓Gravitational wave studies *(Artemis)*
  ✓Astrophysics *(Lagrange)*

Observatoire de la Côte d'Azur (OCA)

- Earth and Space Science research Institute
  Director: Dr. Thierry LANZ (previously at NASA APD)

- 450 employees organised around three poles
  ✓Geophysics and Space geodesy *(Géoazur)*
  ✓Gravitational wave studies *(Artemis)*
  ✓Astrophysics *(Lagrange)* →

- Formation and evolution of planetary systems
- Solar and stellar physics
- Physics of galaxies and cosmology
- Signal processing
- Non-linear dynamics

Leading roles in several major international projects, e.g.:

- MATISSE at ESO
- GAIA and Euclid ESA satellites
- Interferometric facility CHARA

Involvement in major radio astronomical projects

LOFAR — Key Project "Surveys" — **Since 2008**

ASKAP — EMU — **Since 2010**

Co-lead of "Galaxy Clusters" WG

MeerKAT — MIGHTEE — **Since 2010**

SKA — Extragalactic Continuum — **Since 2013**

Member of the SWG Core Team

Organisation of the first SKA French Industry Day

Development of algorithms for radio interferometry

Since 2013: CNRS grant to develop a joint research on “Data analysis for Very Large Arrays in Radioastronomy”

Funding used as a leverage to obtain a grant for 2015–2018 from French National Research Agency for the project MAGELLAN

Data processing for radio interferometry

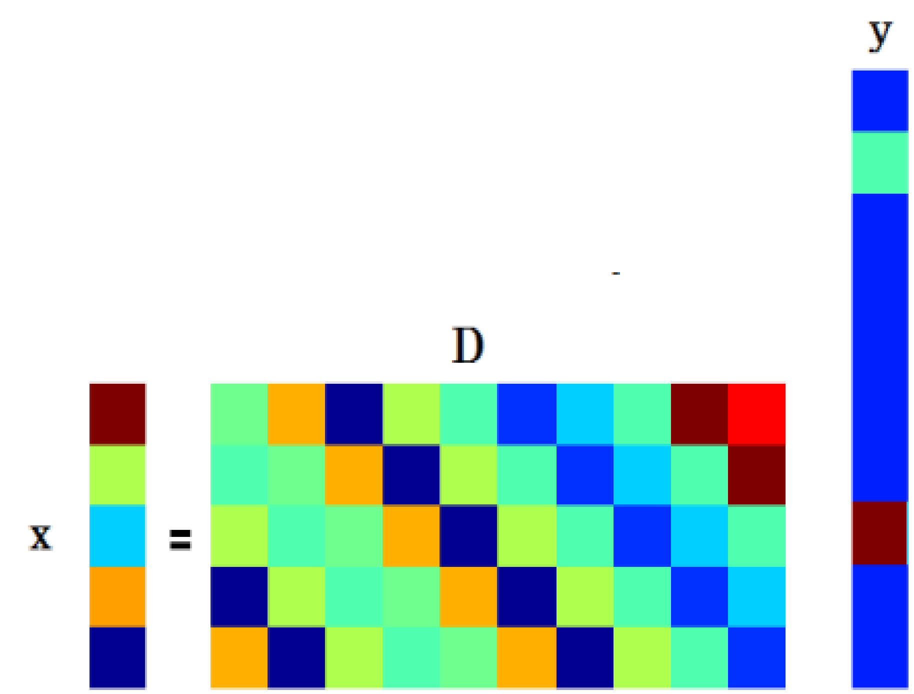
- Image reconstruction
- Calibration

The Project received in July 2015 a grant ($18k) from AstroCompute In The Cloud (SKA/AWS)

Sparse representations: a simplified description

- a **dictionary** is a data representation space where the signal can be *sparsified*
  → the choice of a dictionary depends on the nature of the signal
- an **atom** is a column of a dictionary of the same size as the signal
- a signal is sparse if most of its coefficients are equal to zero

  → natural signals are compressible or weakly sparse, i.e. most of its coefficients have very low amplitudes



Slide courtesy: A. Dabbech

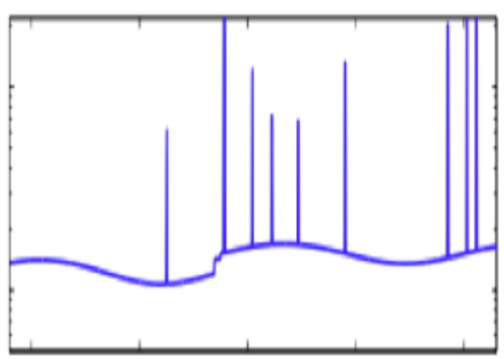
Sparsity promoting approaches

- The **synthesis-based sparsity approach** assumes that the signal to be recovered (of size N) is a linear combination of few atoms of a given over-complete synthesis dictionary ($\boldsymbol{S}$ of size $N \times M$ with $M >> N$).

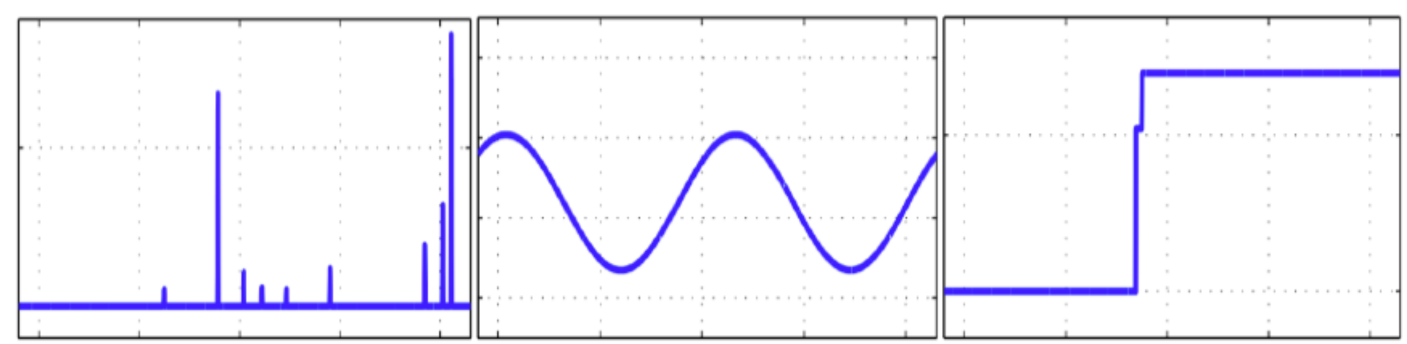
Sparsity promoting approaches

- The **synthesis-based sparsity approach** assumes that the signal to be recovered (of size N) is a linear combination of few atoms of a given over-complete synthesis dictionary (S of size $N \times M$ with $M >> N$).

- Synthesis sparse : $\mathbf{x} = \mathbf{S}\theta$, with θ sparse



$\mathbf{x}$



$\mathbf{x} = \mathbf{S}_{\text{peaks}}\theta_{\text{peaks}} \quad + \quad \mathbf{S}_{\text{sinus}}\theta_{\text{sinus}} \quad + \quad \mathbf{S}_{\text{steps}}\theta_{\text{steps}}$

$\mathbf{x} = \mathbf{S}\theta$ with $\mathbf{S} = [\mathbf{S}_{\text{peaks}}\ \mathbf{S}_{\text{sinus}}\ \mathbf{S}_{\text{steps}}]$

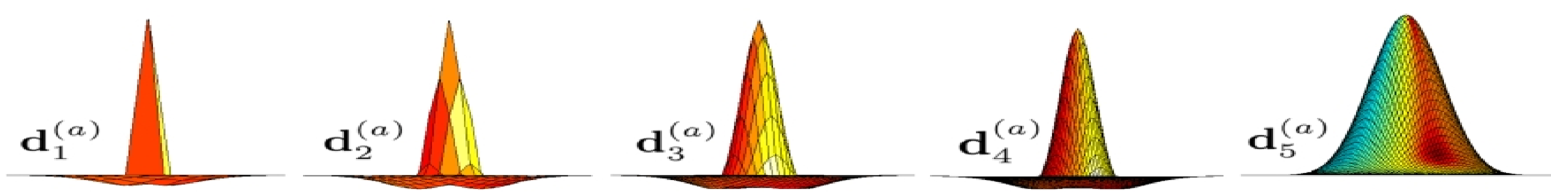
Sparsity promoting approaches

- The **analysis-based sparsity approach** assumes that the projection of the signal $\boldsymbol{x}$ to be recovered onto a given over-complete analysis dictionary ($\mathbf{A}$ of size $N \times M$ with $M >> N$) is sparse.

Sparsity promoting approaches

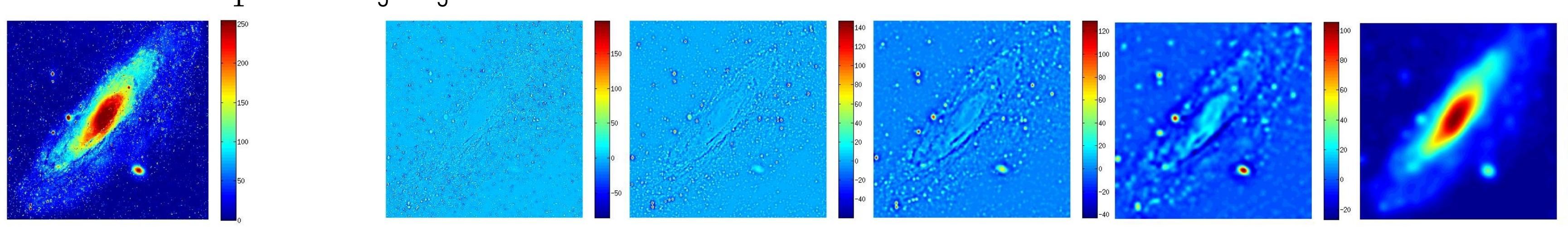
- The **analysis-based sparsity approach** assumes that the projection of the signal x to be recovered onto a given over-complete analysis dictionary (A of size $N \times M$ with $M >> N$) is sparse.

✯ The IUWT up to a scale J→ analysis dictionary $\mathbf{A} = [\mathbf{A}_{(1)}, .. , \mathbf{A}_{(J+1)}]$,
concatenation of J+1 sub-dictionaries of size N×N



✯ The IUWT analyzes an image $\mathbf{y}$ of size N into a set of coefficients:

$\mathbf{a} = \mathbf{A}^\top \mathbf{y} = [\mathbf{w}_1^\top, \ldots, \mathbf{w}_J^\top, \mathbf{c}_J^\top]^\top$ of size N×(J+1)



$\mathbf{y}$ $\quad \mathbf{a} =$ $\mathrm{A}_{(1)}{}^\top \mathbf{y}$ $\mathrm{A}_{(2)}{}^\top \mathbf{y}$ $\mathrm{A}_{(3)}{}^\top \mathbf{y}$ $\mathrm{A}_{(4)}{}^\top \mathbf{y}$ $\mathrm{A}_{(5)}{}^\top \mathbf{y}$

Application I -
Radio interferometric deconvolution

Well posed problem:

- A solution exists
- It is unique
- Its behaviour changes continuously with initial conditions

- Fourier space : linear filtering by sampling function

$$\underbrace{V(u,v) = \mathcal{F}\{O(\alpha,\beta)\} \cdot T(u,v)}_{\textit{Linear filtering}}$$

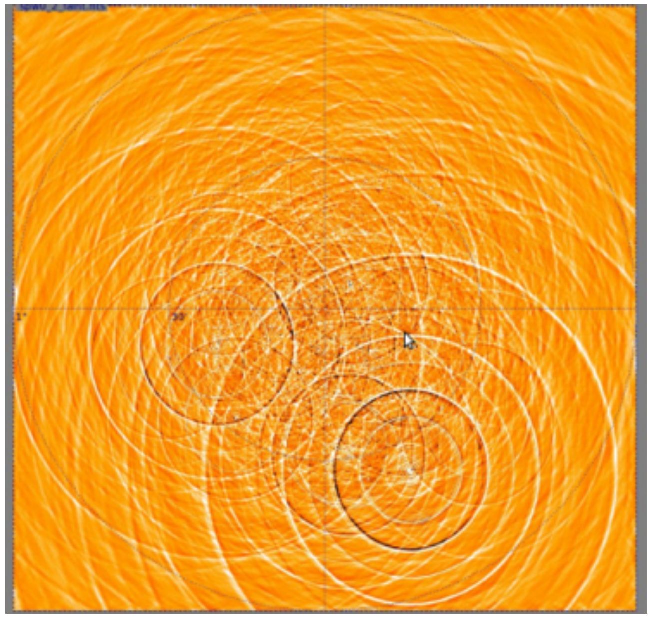
- Image space : convolution by PSF (dirty beam)

$$\underbrace{I(\alpha,\beta) = O(\alpha,\beta) \star PSF(\alpha,\beta)}_{\textit{Convolution}}$$

- Discrete model in matrix form :

Fourier space : $\mathbf{v} = \mathbf{TFx} + \mathbf{n}$

Image space : $\mathbf{y} = \mathbf{F}^{\dagger}\mathbf{TFx} + \epsilon = \mathbf{Hx} + \epsilon$



Deconvolution is an **ill posed problem** due to missing information in the uv coverage → **Infinity number of “skies” that can fit the dirty map**

Application I - Radio interferometric deconvolution

- Maximum likelihood approaches aiming to estimate the signal that best fits the data are not sufficient to solve ill posed problems.
- Additional information on the solution is required and must be added to the inverse problem as **regularization terms**.

The inverse problem takes the form

$$min_{\mathrm{x}} \quad \Phi_{fidelity}(y, x) \quad s.t. \quad \Phi_{prior}(x)$$

- $\Phi_{fidelity}(y, x)$ = fidelity to data term, i.e. distance between the available noisy data **y** and the possible measured data given the signal **x**
  $\rightarrow \Phi_{fidelity}(y, x) = \frac{1}{2}\|Y - HX\|^2$ in this case
- $\Phi_{prior}(x)$ = regularization term corresponding to the additional information on the signal to be estimated
  $\rightarrow$???

- Besides the positivity prior, the **sparsity priors** have been very successful in image reconstruction.

Regularisation through Synthesis & Analysis

$$min_x \quad \Phi_{fidelity}(y, x) \quad s.t. \quad \Phi_{prior}(x)$$

- **Synthesis:** the regularisation term takes the form $\Phi_{prior}(x) = \mu||\gamma||_n$, with, in general, $n = 1$.

The synthesis model $\mathbf{y} = \mathbf{HS}\boldsymbol{\gamma} + \mathbf{n}$ where $\boldsymbol{\gamma}$ is **sparse**

ex. $\hat{\mathbf{x}}_s = \mathbf{S}.\{\arg\min_{\gamma} \parallel \mathbf{y} - \mathbf{HS}\gamma \parallel_2^2 + \mu \parallel \gamma \parallel_p^p\}.$

Regularisation through Synthesis & Analysis

$$min_x \quad \Phi_{fidelity}(y,x) \quad s.t. \quad \Phi_{prior}(x)$$

- **Synthesis:** the regularisation term takes the form $\Phi_{prior}(x) = \mu||\gamma||_n$, with, in general, $n = 1$.

The synthesis model $\mathbf{y} = \mathbf{HS}\gamma + \mathbf{n}$ where γ is **sparse**

ex. $\hat{\mathbf{x}}_s = \mathbf{S}.\{\arg\min_{\gamma} \| \mathbf{y} - \mathbf{HS}\gamma \|_2^2 + \mu \| \gamma \|_p^p\}.$

- **Analysis**: the regularisation term takes the form $\Phi_{prior}(x) = \mu||A^\top x||_n$, with, in general, $n = 1$.

The analysis model $\mathbf{y} = \mathbf{Hx} + \mathbf{n}$ where $\mathbf{A}^\top\mathbf{x}$ is **sparse**

ex. $\hat{\mathbf{x}}_a = \arg\min_{\mathbf{x}} \| \mathbf{y} - \mathbf{Hx} \|_2^2 + \mu \| \mathbf{A}^\top \mathbf{x} \|_p^p$

MORESANE:

Model Reconstruction by Synthesis-Analysis Estimators

- **Signal model**

$$\mathbf{x} = \sum_{i}^{P} \mathbf{o}^{(i)} = \mathbf{X}\boldsymbol{\theta}, \quad \text{where } \mathbf{o}^{(i)} = \boldsymbol{\theta}_i \mathbf{X}_i \qquad \text{P<<N}$$

$\mathbf{X}$ synthesis dictionary, whose atoms are $\mathbf{X}_i = \frac{\mathbf{o}^{(i)}}{||\mathbf{o}^{(i)}||_2}$

- **Data model**

$$\mathbf{y} = \mathbf{H}\mathbf{X}\boldsymbol{\theta} + \mathbf{n}$$

sparse synthesis problem with **unknown** synthesis dictionary

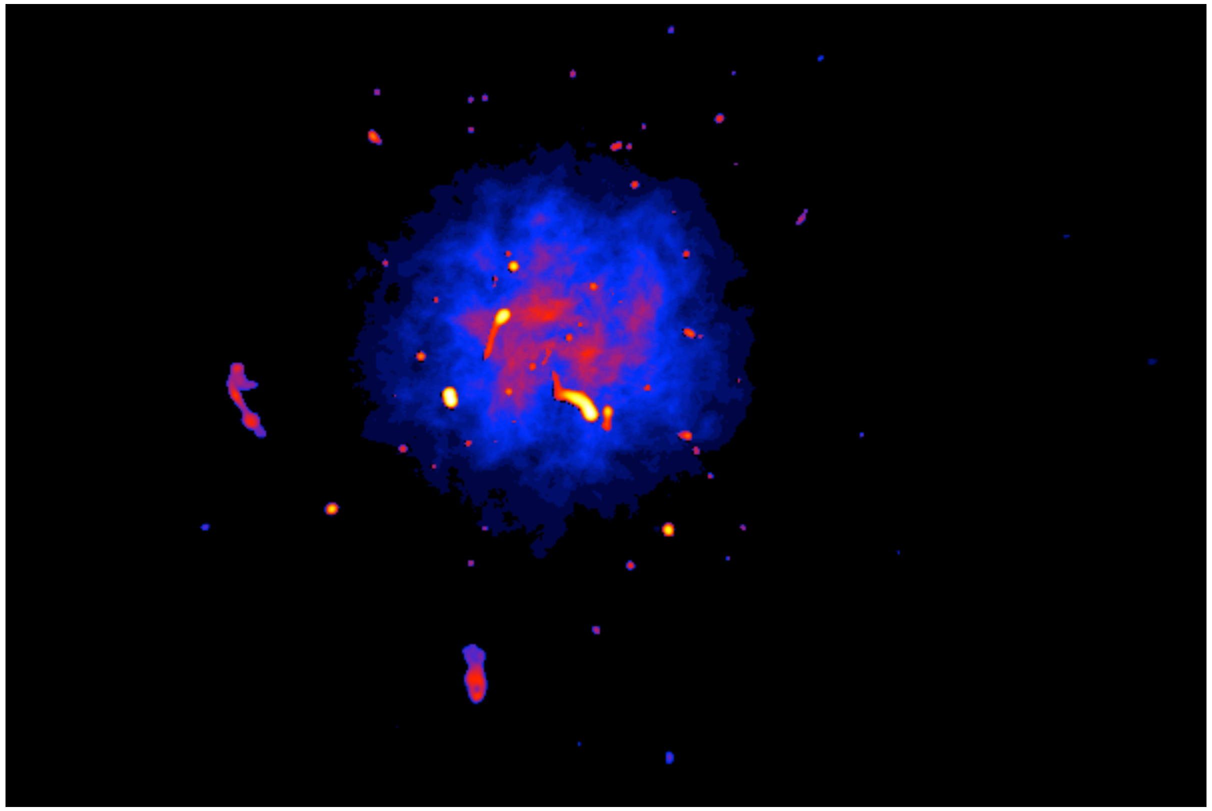
- Greedy, iterative
- The atoms $\mathbf{X}_i$ are estimated using **analysis** priors

Dabbech+ 12, 15
See also PhD Dabbech (on ADS)

MORESANE: application to simulated data

Radio galaxies +

Radio halo ($P_{1.4\ \mathrm{GHz}} \sim 1 \times 10^{24}$ W/Hz)



Relativistic electron population

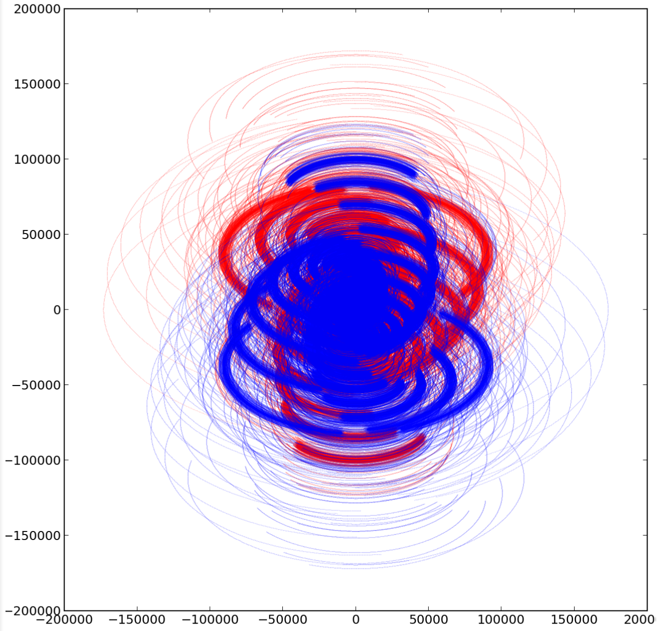
+ Magnetic field model

Faraday tool (Murgia+ 04)

8 hours observations

60 sec integration time

50 MHz BW starting @ 1415 MHz

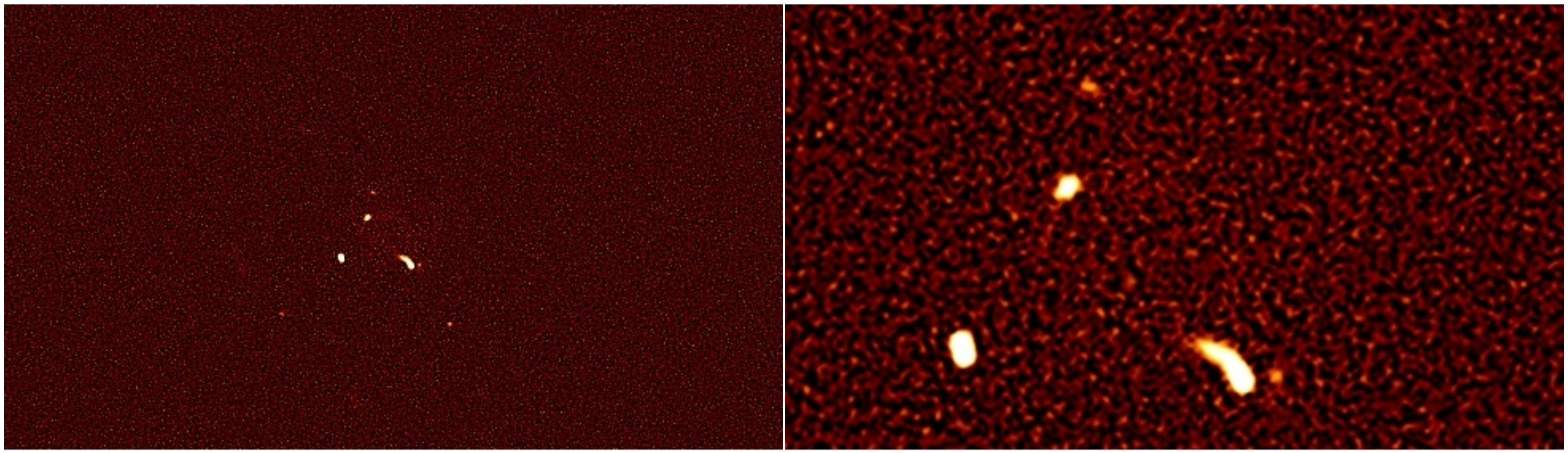


Simulations of SKA1 MID

observations

MeqTrees tool (Noordam & Smirnov 10)

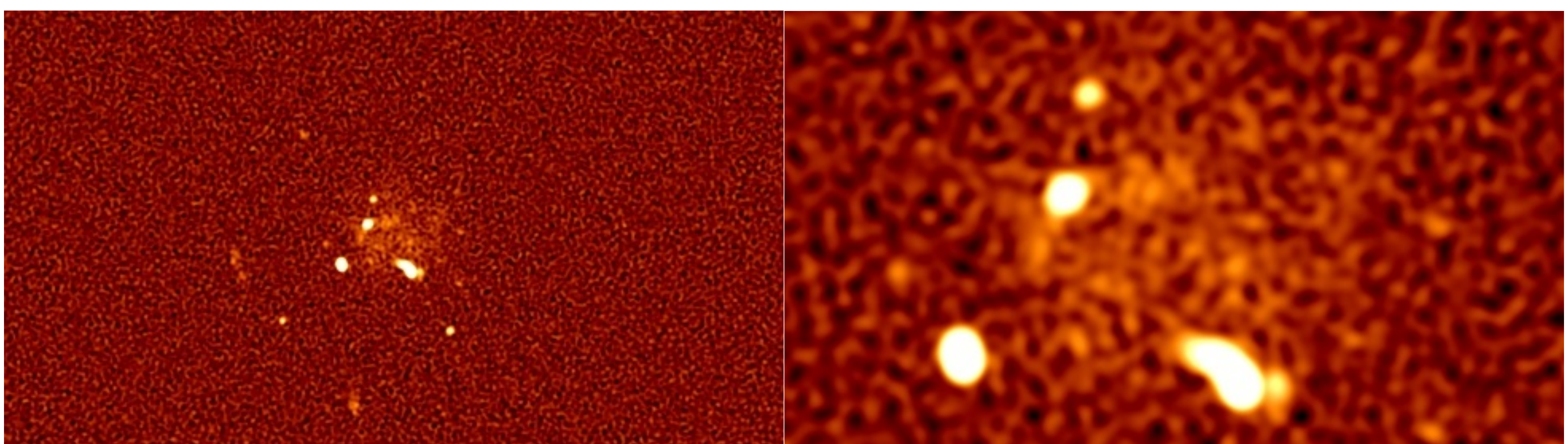
Simulated SKA1-MID observations



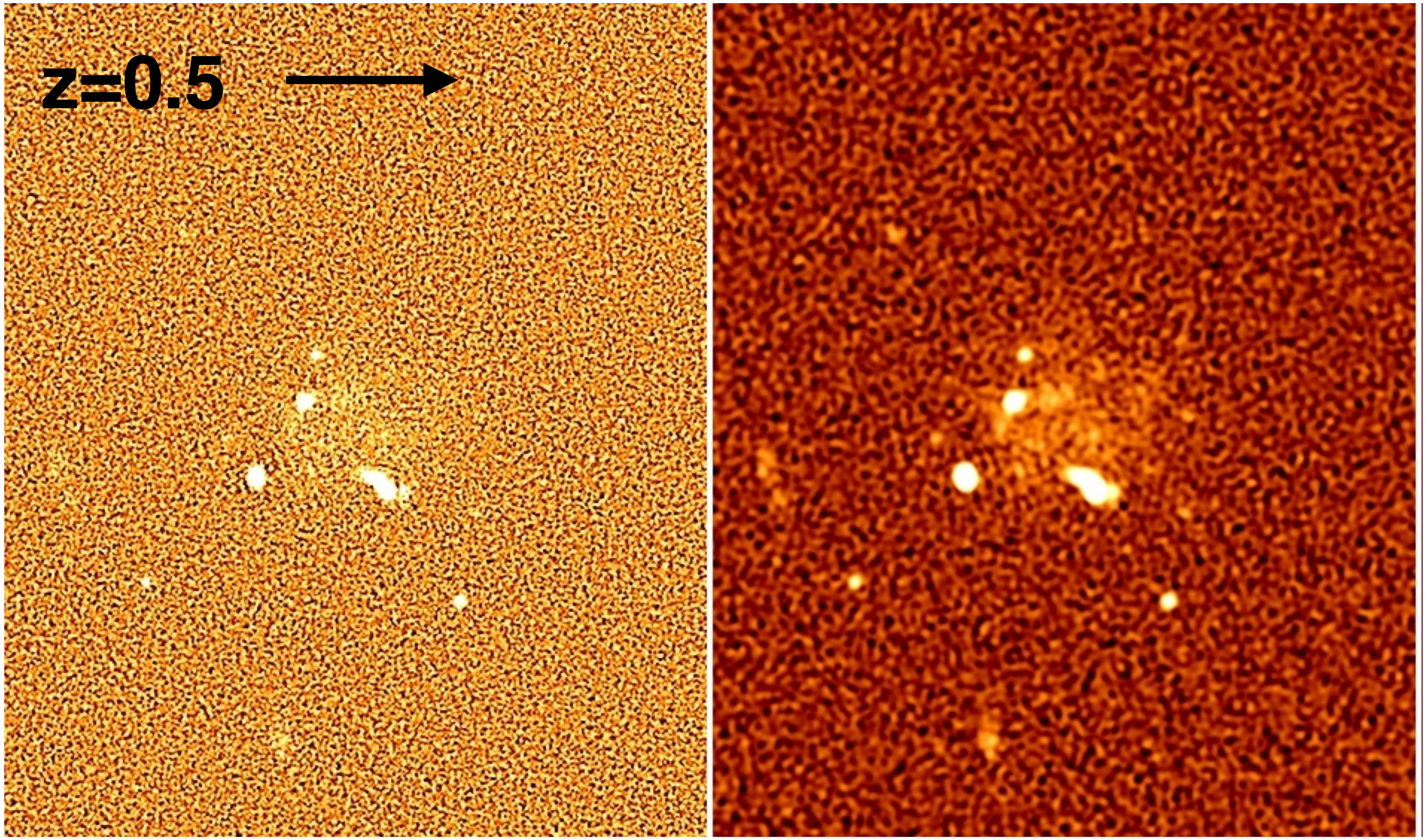
Dirty maps

@ 1 & 5 arcsec resolution

Ferrari+ 15 - PoS(AASKA14)075
See also Dabbech+ 15



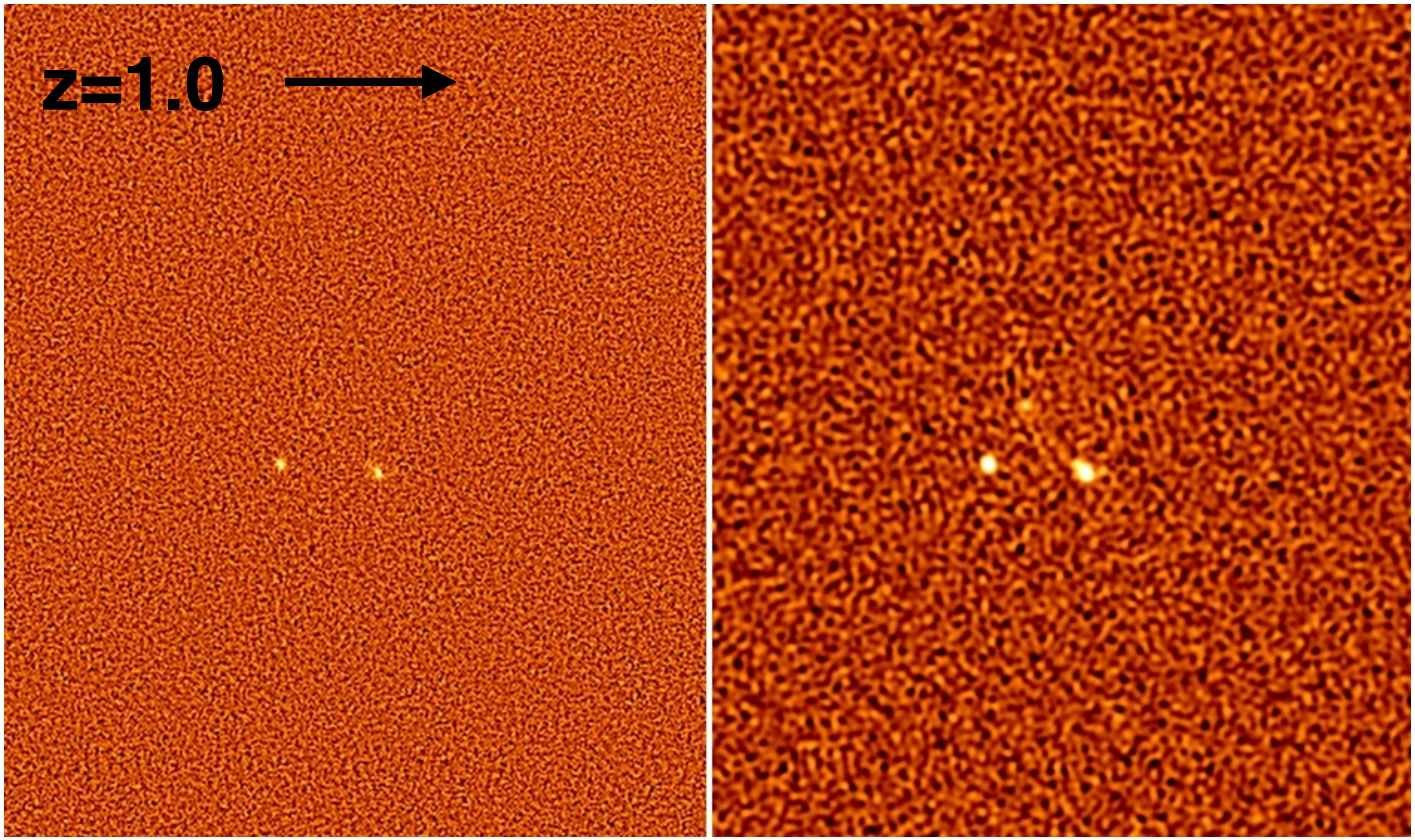
Up to which z can we detect galaxy clusters? Results with CLEAN



Uniform weighting
res ~ 1.8 arcsec
rms ~ 2.4 μJy/b

Uniform weighting
res ~ 4.5 arcsec
rms ~1.7 μJy/b

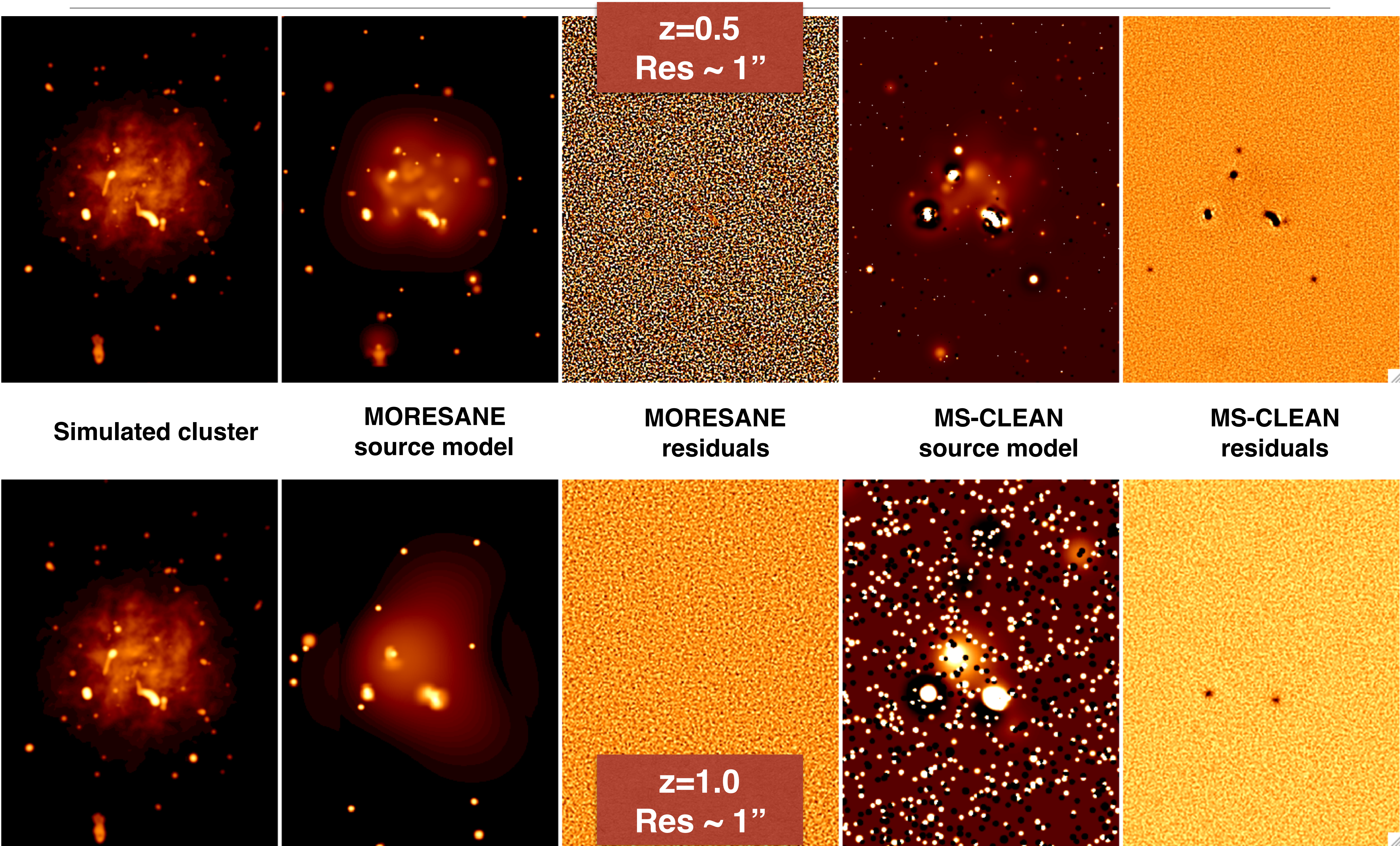
Up to which z can we detect galaxy clusters?
Results with CLEAN



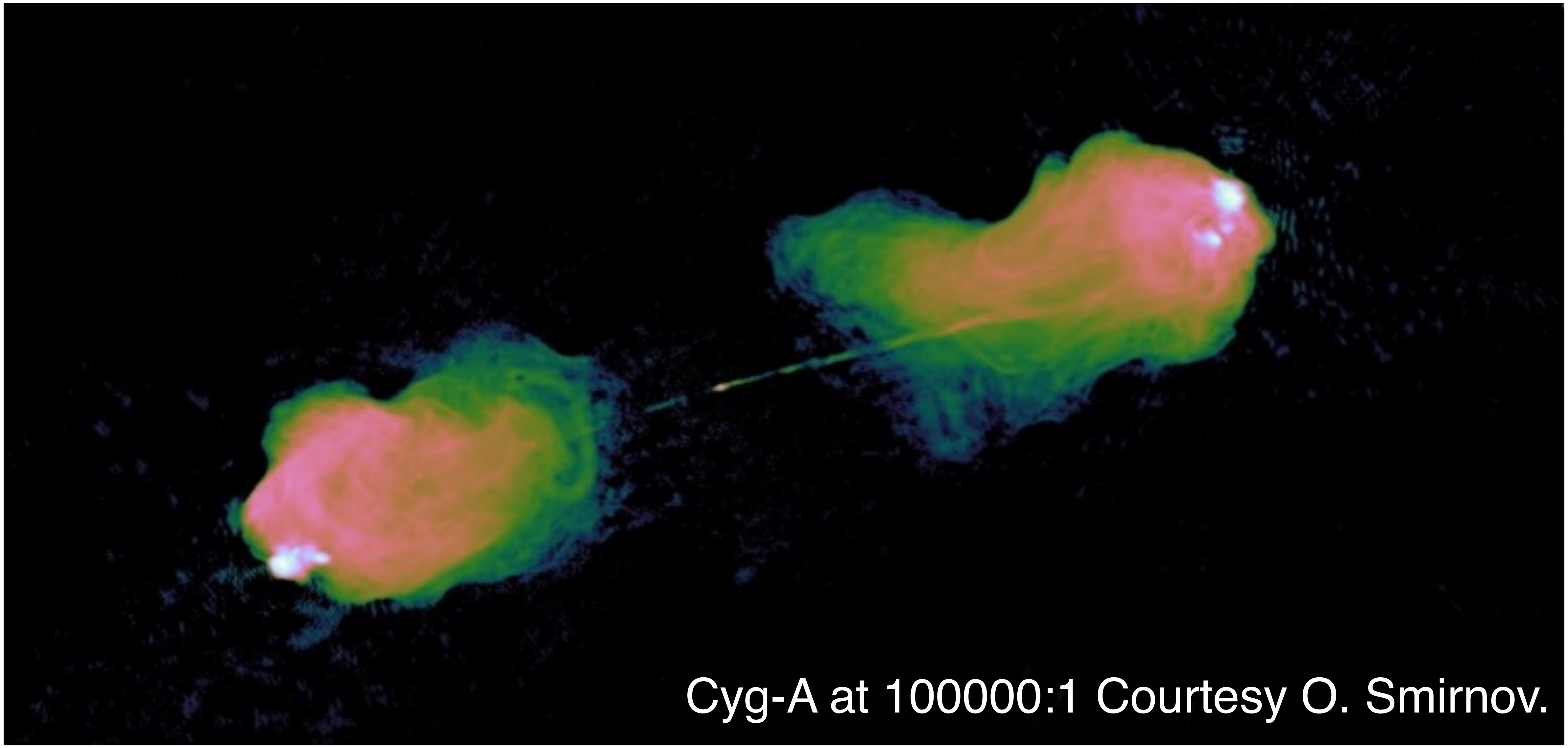
Uniform weighting	**Uniform weighting**
res ~ 1.8 arcsec	res ~ 4.5 arcsec
rms ~ 2.5 μJy/b	rms ~1.9 μJy/b

Up to which z can we detect galaxy clusters?

Results with CLEAN & MORESANE



MORESANE: application to real observations



https://github.com/ratt-ru/PyMORESANE

(A python and pyCUDA-accelerated implementation of MORESANE by J. Kenyon, Rhodes University, available on GitHub)

Multi-frequency image reconstruction

Rely on a model of the frequency-dependent brightness distribution

Taylor expansion of a power law around a reference frequency.

"Double deconvolution" of a "wide-band" dirty image

- "double deconvolution", (Conway et al. 1990).
- "double CLEAN", (Sault and Wieringa 1994).

Taylor expansion of a power law (with a varying index).

Multi-frequencies deconvolution:

- Multi-scale multi-frequency CLEAN, (Rau and Cornwell 2011)
- Maximum entropy algorithm, (Bajkova and Pushkarev 2011).

Power law.

Multi-frequency deconvolution using a Bayesian estimation framework, (Junklewitz et al. 2014).

MUFFIN:

Multi-Frequency Image Reconstruction for Radio Interferometry

Ferrari, A. +15

The regularization game

Objective function = data fidelity + regularization

$$\min_X \frac{1}{2}\|Y - HX\|^2 + f_{\text{reg}}(X)$$

Sparse analysis prior

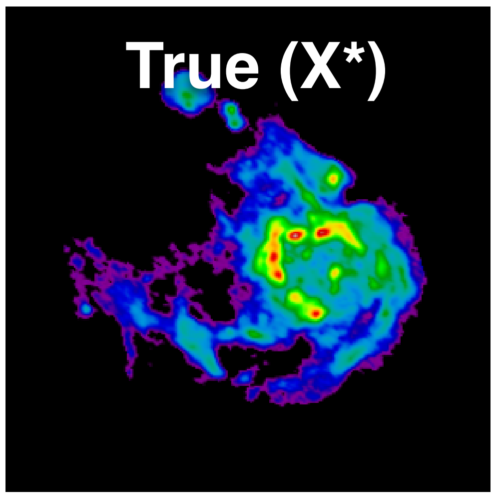
$$f^{\text{reg}}(X) = \mathbf{1}_{\mathbb{R}+}(X) + \frac{\mu_\varepsilon}{2}\|X\|_F^2 + \mu_s\|W_s X\|_1 + \mu_\nu\|XW_\nu\|_1$$

- $\mathbf{1}_{\mathbb{R}+}(X)$: positivity constraint, $\|X\|_F^2$: Tikhonov (if needed).
- $\|W_s X\|_1 = \|W_s x_1\|_1 + \ldots + \|W_s x_L\|_1 \to W_s$ analyses the columns of X, i.e. the single frequency images.
  The ℓ_1 norm promotes sparsity in the analysis coefficients $W_s x_\ell$.
- $\|XW_\nu\|_1 = \|W_\nu^\top X^\top\|_1 \to W_\nu^\top$ analyses the lines of X, i.e. the spectra.

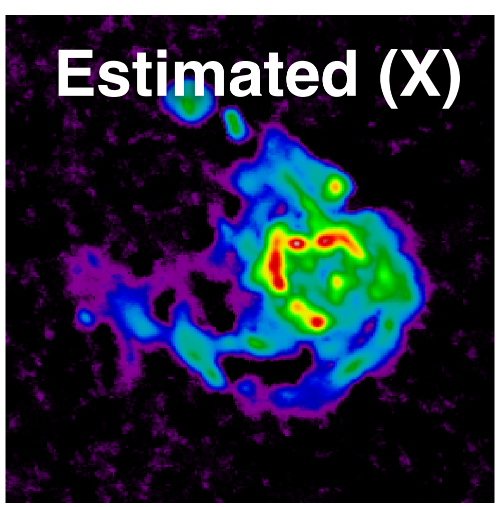
https://github.com/andferrari/Muffin.jl

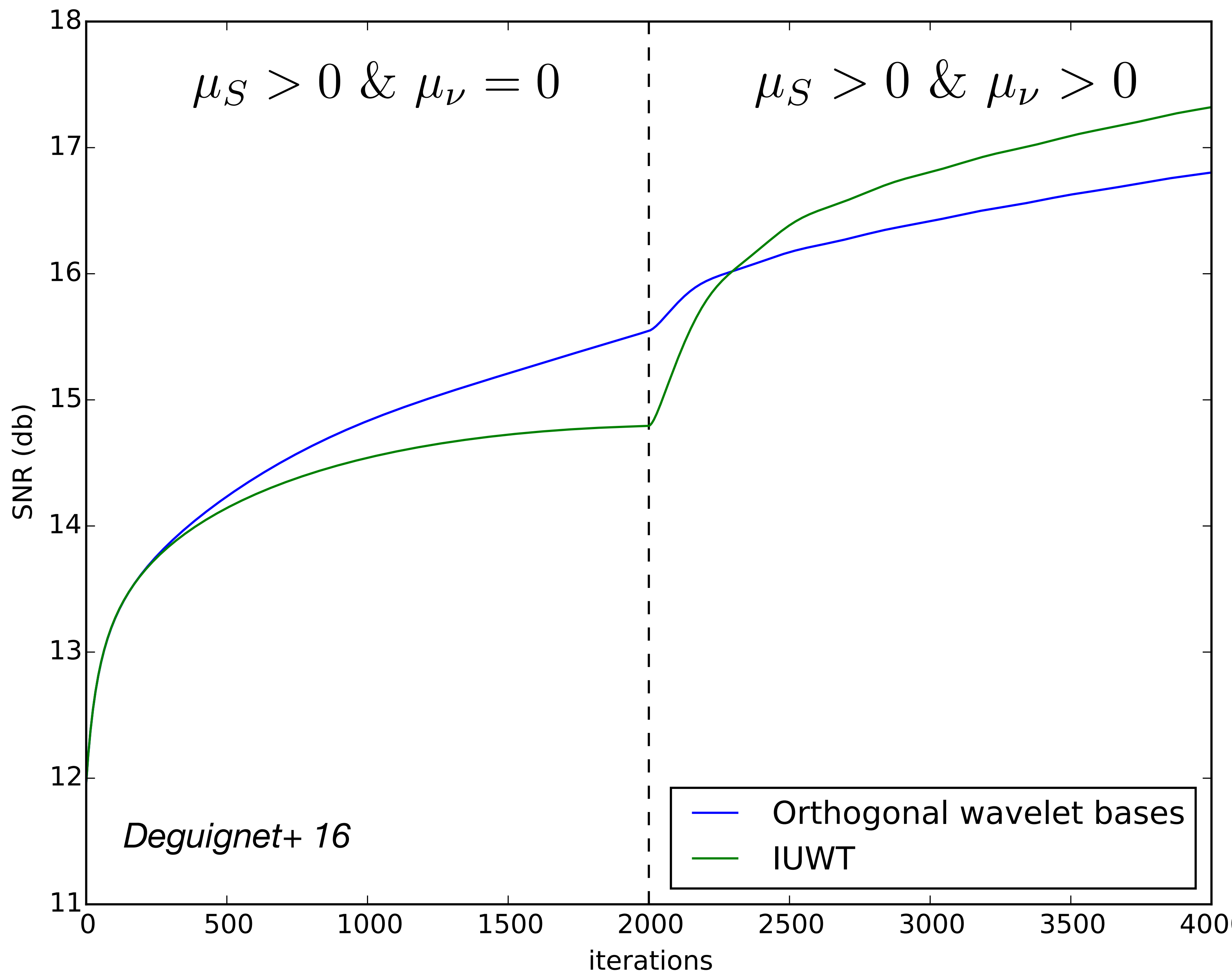
(A Julia parallel version of MUFFIN available on GitHub)

MUFFIN: RESULTS ON SIMULATIONS

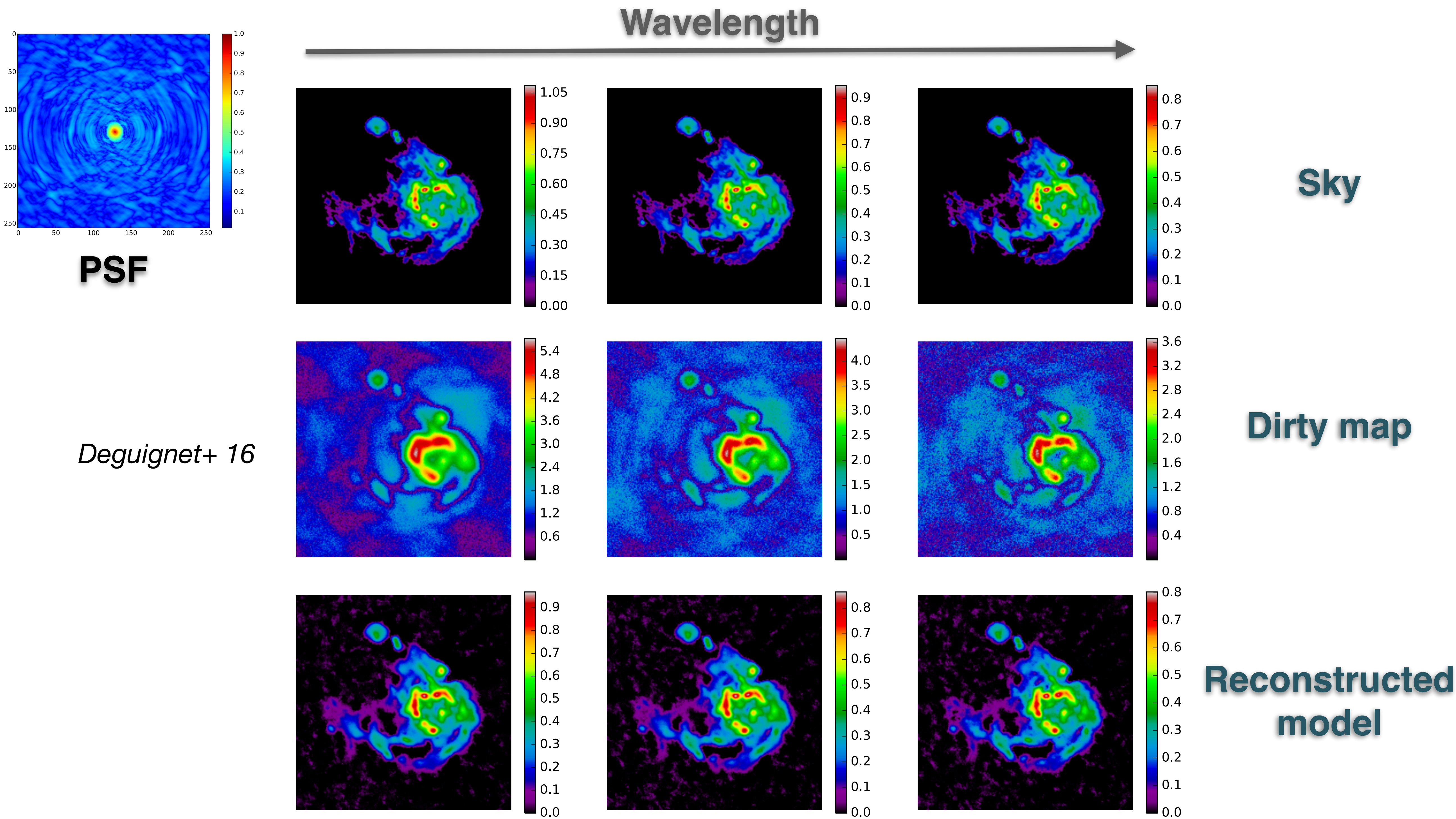


$$\mathrm{SNR}(\boldsymbol{X}, \boldsymbol{X}^\star) := 10 \log_{10} \left(\frac{\|\boldsymbol{X}^\star\|_2^2}{\|\boldsymbol{X} - \boldsymbol{X}^\star\|_2^2} \right)$$



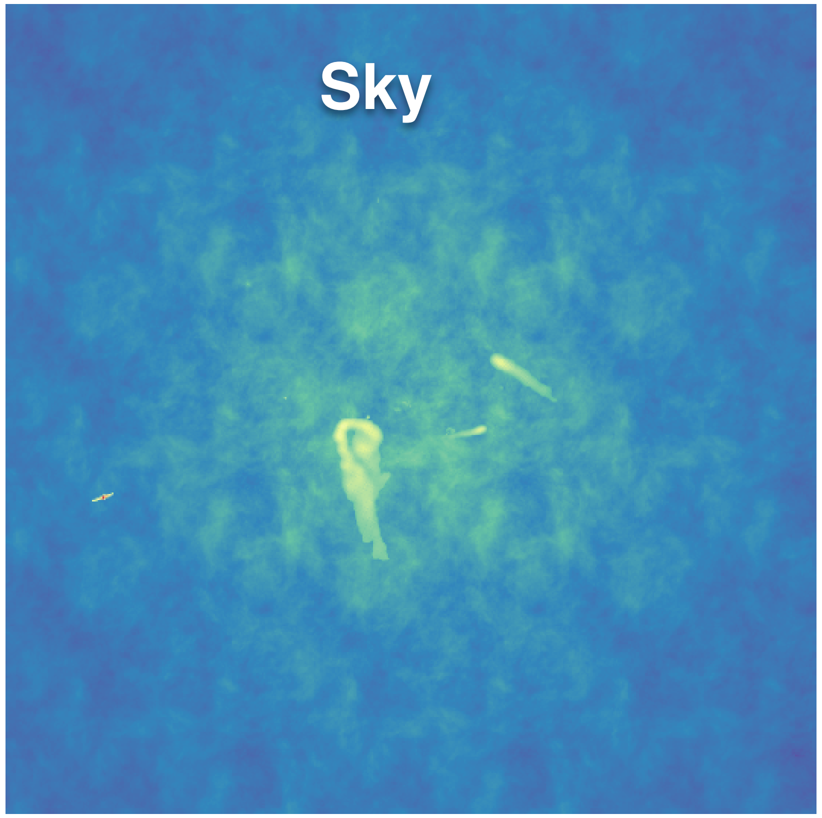


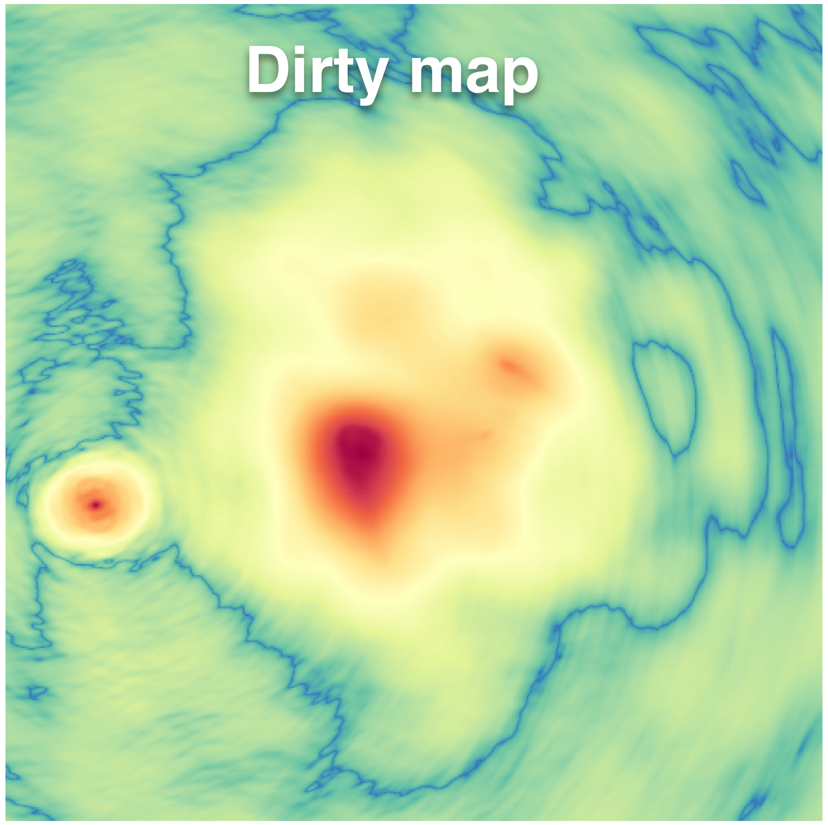
MUFFIN: RESULTS ON SIMULATIONS

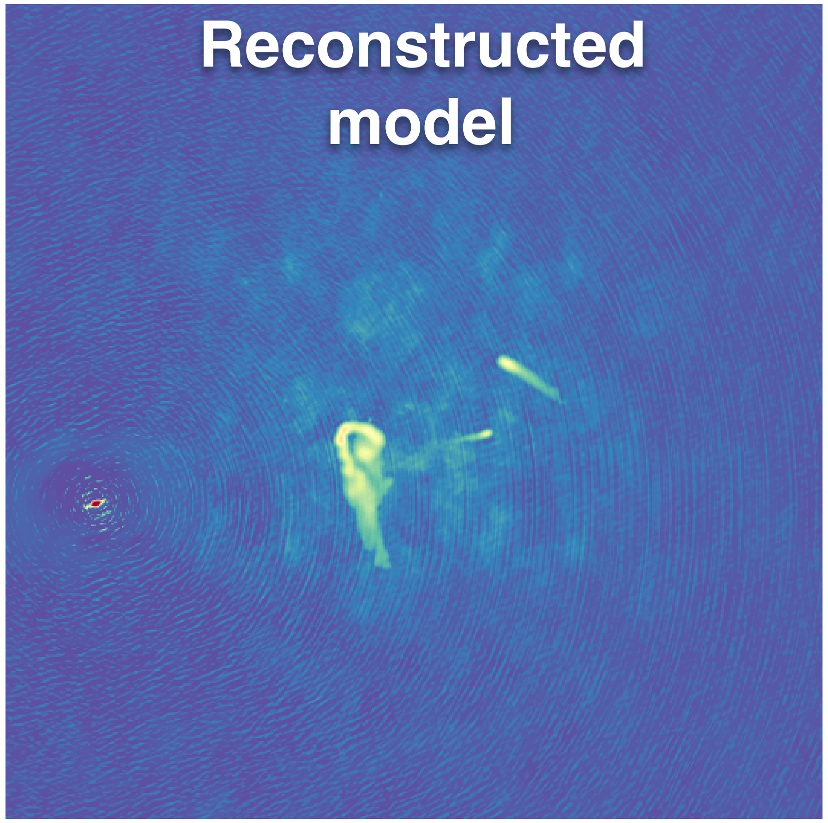


Deguignet+ 16

MUFFIN: RESULTS ON SIMULATIONS







Based on galaxy cluster simulated images by

F. Loi, M. Murgia, F. Govoni (INAF, Cagliari)

2048 x 2048 x 64 cubes

Comparison with MFMS CLEAN (CASA implementation)

Deguignet+ in prep.

Application II -
Radio interferometric calibration

Designing super-resolution calibration algorithms for modern radio interferometers

- **Robustness** : wider distribution class than the Gaussian one, to model the noise
- **Computational efficiency** : based on an iterative maximum-likelihood estimator for the case of non-uniform white and coloured noise and sparse representation framework

M.N. El Korso, R. Boyer and P. Larzabal :

organisers of **special session "Advanced methods in calibration for interferometry phased array in radio-astronomy"**

EUSIPCO 2016

See first results in:

[1] M. Brossard, M. N. El Korso, M. Pesavento, R. Boyer and P. Larzabal, "Calibration of Radio Interferometers Using a Sparse DOA Estimation Framework", http://arxiv.org/abs/1603.00263

[2] V. Ollier, M. N. El Korso, R. Boyer, P. Larzabal and M. Pesavento, "Relaxed concentrated MLE for robust calibration of radio interferometers", http://arxiv.org/abs/1603.01070

AADC @ OCA: algorithmic developments

Participation to the AADC consortia in Nice

(1.7 PY in S2 for a total cost of ~150 k€)

- Chiara FERRARI, Astronomer at OCA
- André FERRARI, Professor at UNS
- David MARY, Professor at UNS
- Jérémy DEGUIGNET, PhD student at OCA
- Martin VANNIER, Research engineer at OCA

Extra-galactic team @ OCA

Signal Processing team @ OCA

External collaborators

- Group in École Normale Supérieure de Cachan
- Group in Télécom ParisTech