UNIVERSITY OF CAMBRIDGE

Lecture 9
Advanced Imaging with Radio Interferometers

Lecturer: Dr Haoyang Ye (hy297)

UNIVERSITY OF CAMBRIDGE

Overview

- Introduction to Big Data Radio Astronomy and Key Science Projects
- Instrument simulations and design tools
- Science Data Processing
    - o Lecture 7: Calibration of radio observations
    - o Lecture 8: Imaging techniques
    - o Lecture 9: Advanced imaging techniques
    - o Lecture 10: Time-domain radio astronomy
- Computing infrastructure
- Advanced ML and Bayesian methods for data analysis and science extraction

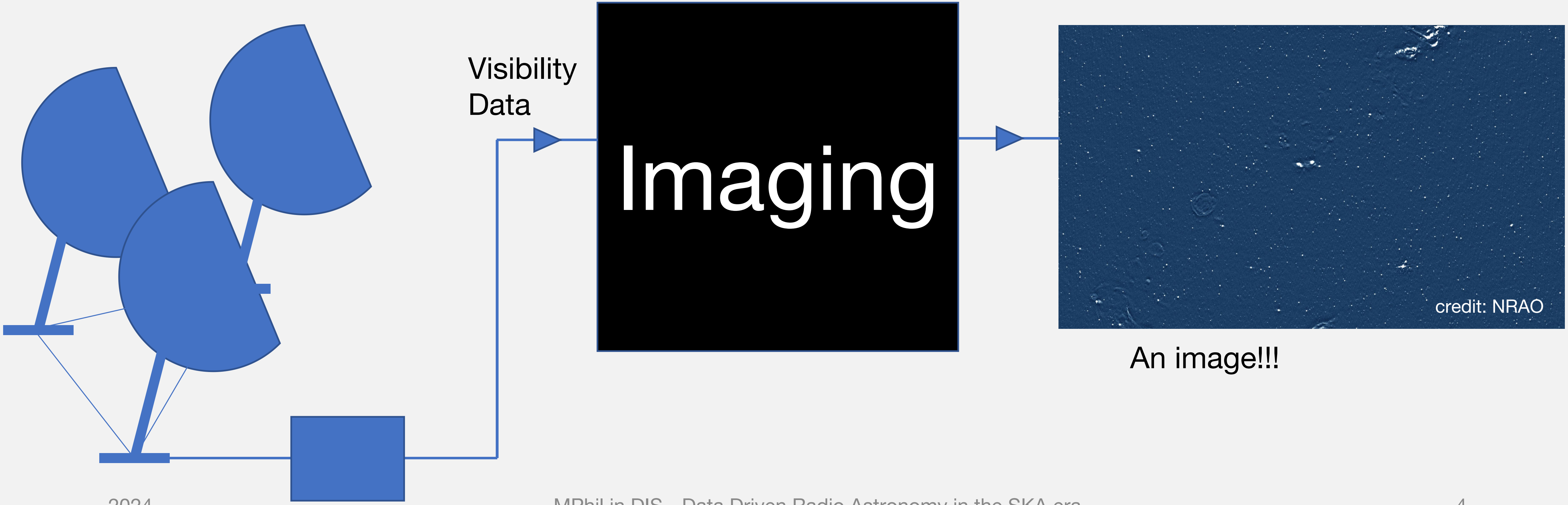
UNIVERSITY OF CAMBRIDGE

In the next one hour

- Join the quiz/discussion
- Ready to code in Python to solve a tiny problem (use QR code to download jupyter notebook)
- Ask questions loud if you need clarification
- You can discuss with each other

UNIVERSITY OF CAMBRIDGE

We wanted to learn:



UNIVERSITY OF CAMBRIDGE

Deconvolution

Observed Visibility data

Inverse Fourier Transform

A DIRTY image!!!

A Deconvolved image!!!

UNIVERSITY OF CAMBRIDGE

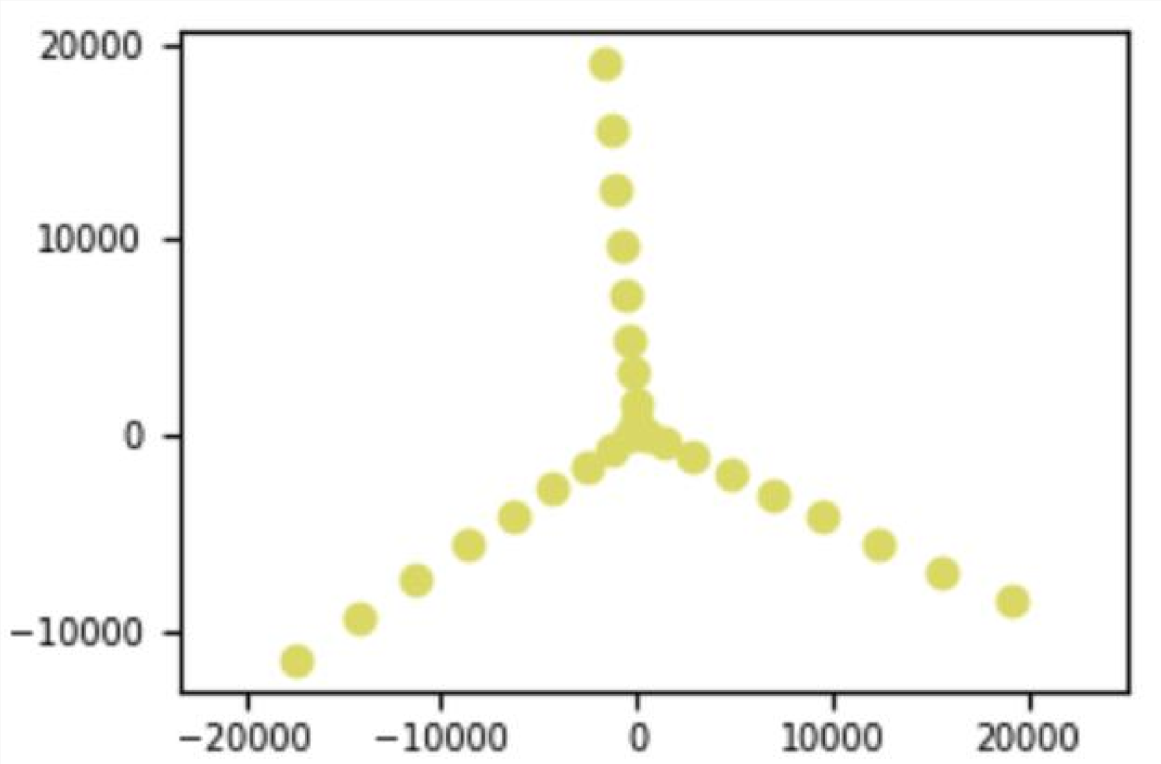
Question Time

Confidence vote on the following concepts:

UNIVERSITY OF CAMBRIDGE

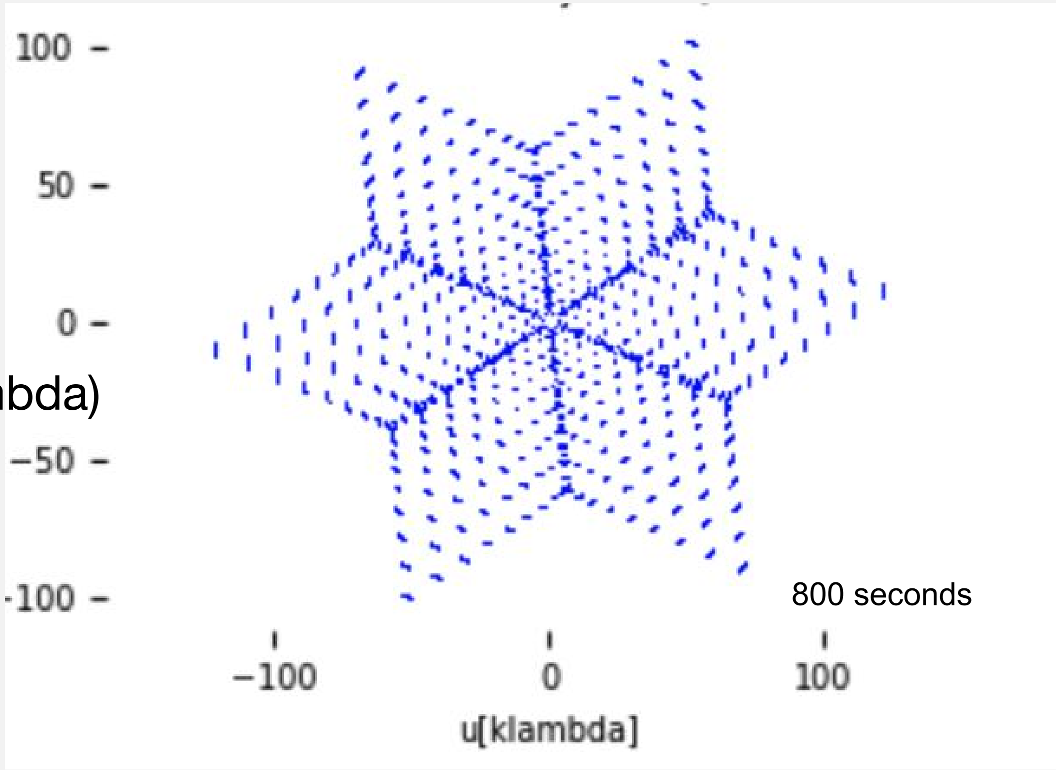
A). uv-coverage

Antenna location



uv-coverage

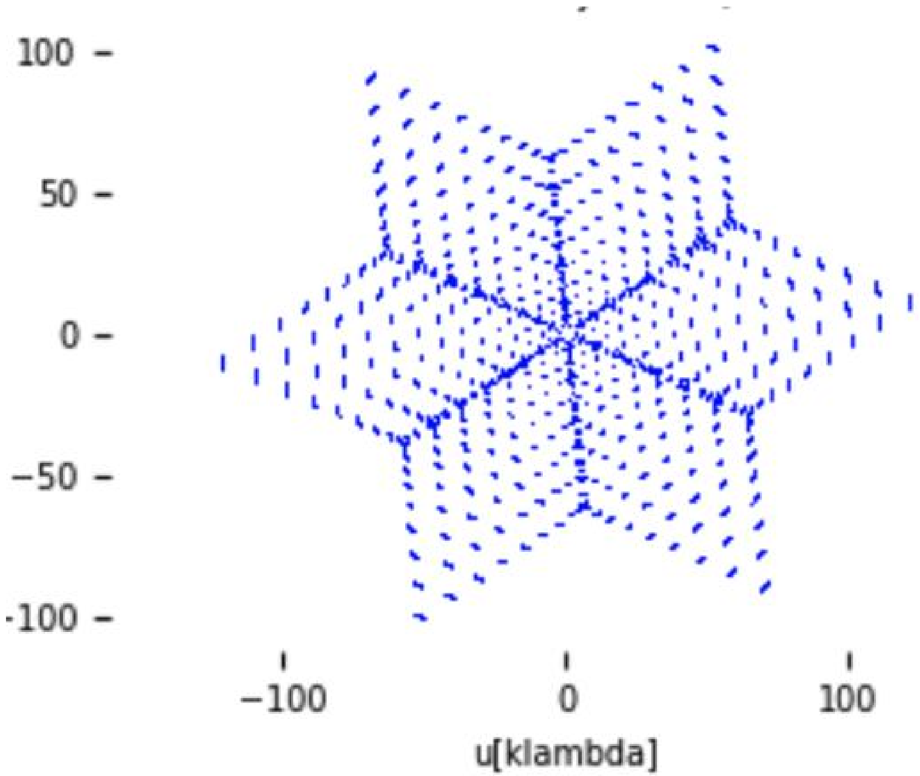
v (klambda)



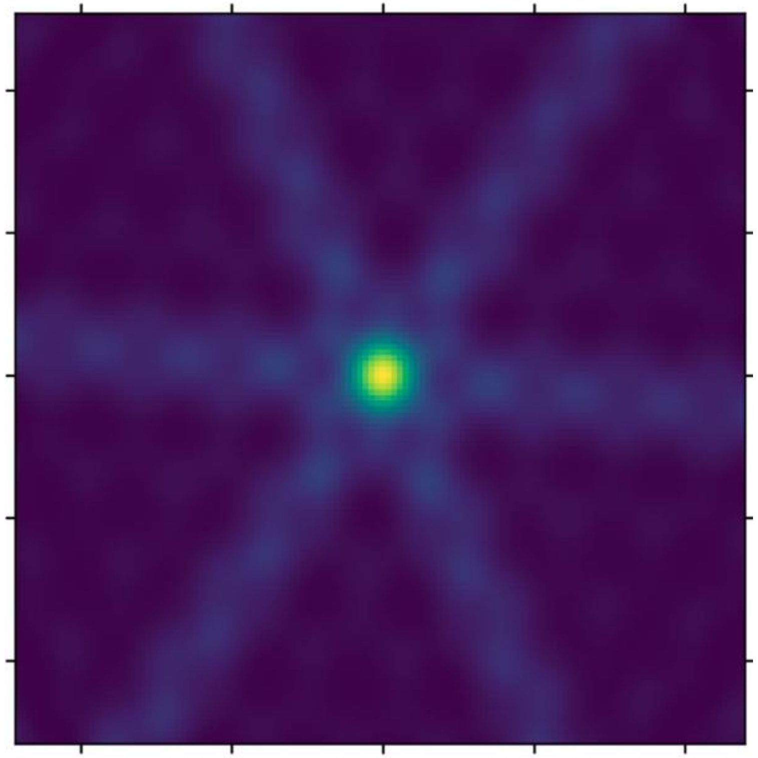
Generated using CASA

UNIVERSITY OF CAMBRIDGE

B). Dirty Beam



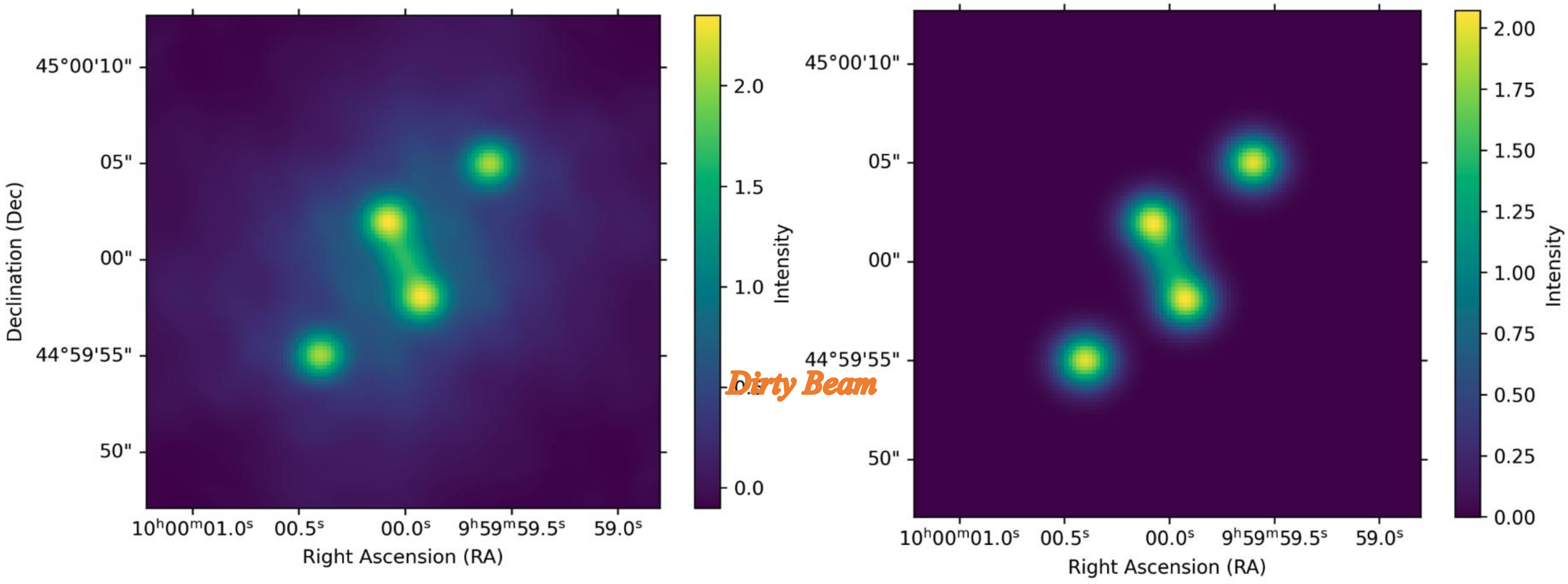
$S(u,v)$



$B(l,m)$

UNIVERSITY OF CAMBRIDGE

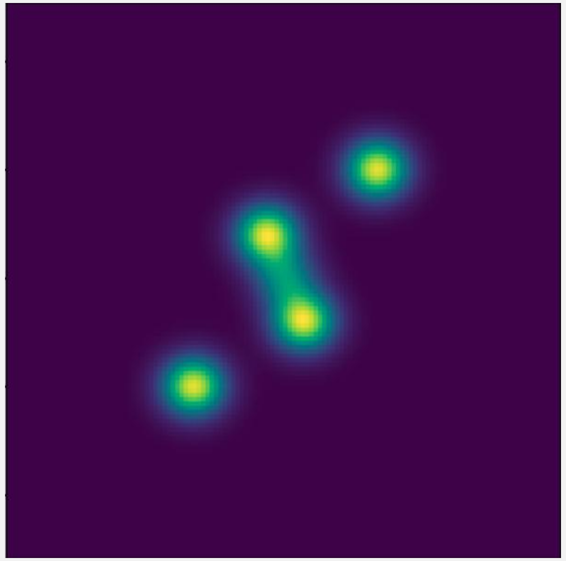
C). Dirty image



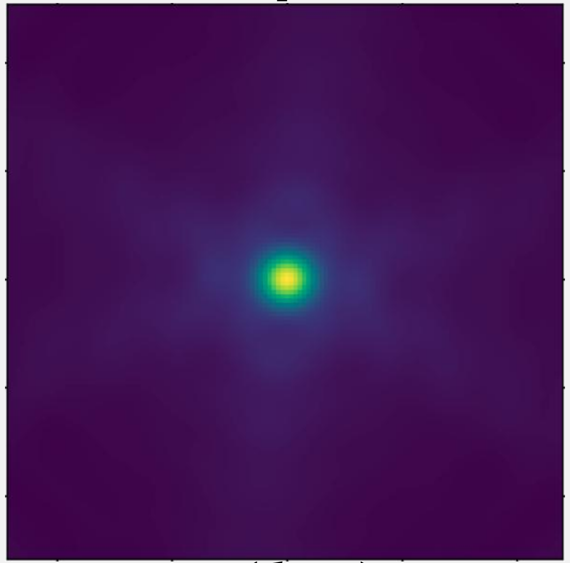
$I(l,m)*B(l,m)$ $I(l,m)$

UNIVERSITY OF CAMBRIDGE

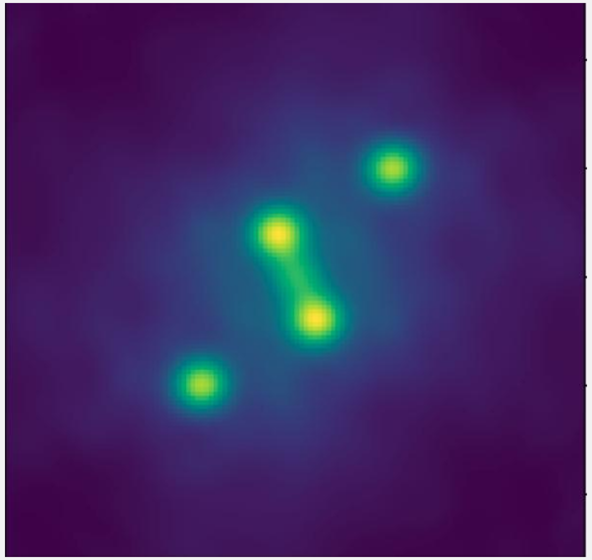
D). Their relationships



$I(l,m)$

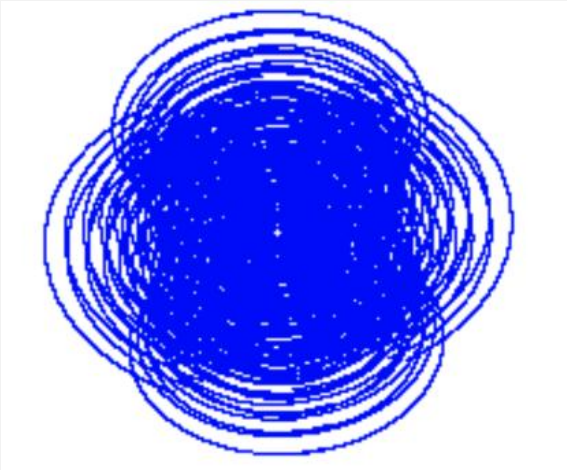


$B(l,m)$

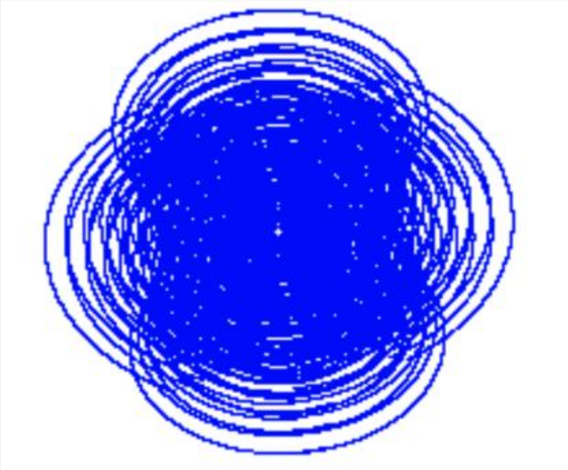


$B(l,m) * I(l,m)$

$V(u,v)$



$S(u,v)$



$S(u,v)V(u,v)$

UNIVERSITY OF CAMBRIDGE

E). Högbom CLEAN - simplest!!

Jan Arvid Högbom (born 3 October 1929) is a Swedish radio astronomer and astrophysicist. Högbom obtained his PhD in 1959 from the University of Cambridge with Martin Ryle.

1. Initialisation: Let $D(x, y)$ represent the dirty image, where (x, y) are the spatial coordinates, and set the residual image $R(x, y)$ equal to the dirty image: $R(x, y) = D(x, y)$.
2. Identify Peaks: Search for the location (x_p, y_p) of the brightest pixel (peak) in the residual image
3. Extract Component: Subtract a (scaled) Gaussian component $G(x, y)$ from the residual image at the location of the identified peak: $R(x, y) \leftarrow R(x, y) - aG(x_p, y_p)$, a is the scaled peak intensity at (x_p, y_p)
4. Update Clean Component List: Add the extracted component to the clean component list
5. Iterate: Repeat steps 2-4 iteratively until one or more termination criteria are met.
6. Final Image Reconstruction: $I(x, y)$ is reconstructed by combining the clean component list with the residual image

UNIVERSITY OF CAMBRIDGE

Question Time

Can I invite someone to explain the Hogbom CLEAN? With or without codes :D

UNIVERSITY OF CAMBRIDGE

About me

- Bachelor: Electronic Engieering
- PhD: Astrophyscis - developing imaging and data anlaysis algorithms for radio interferometers
- Postdoc: worked with FAST*, LOFAR*, now working for SKA*

*FAST, LOFAR, SKA are radio telescope names

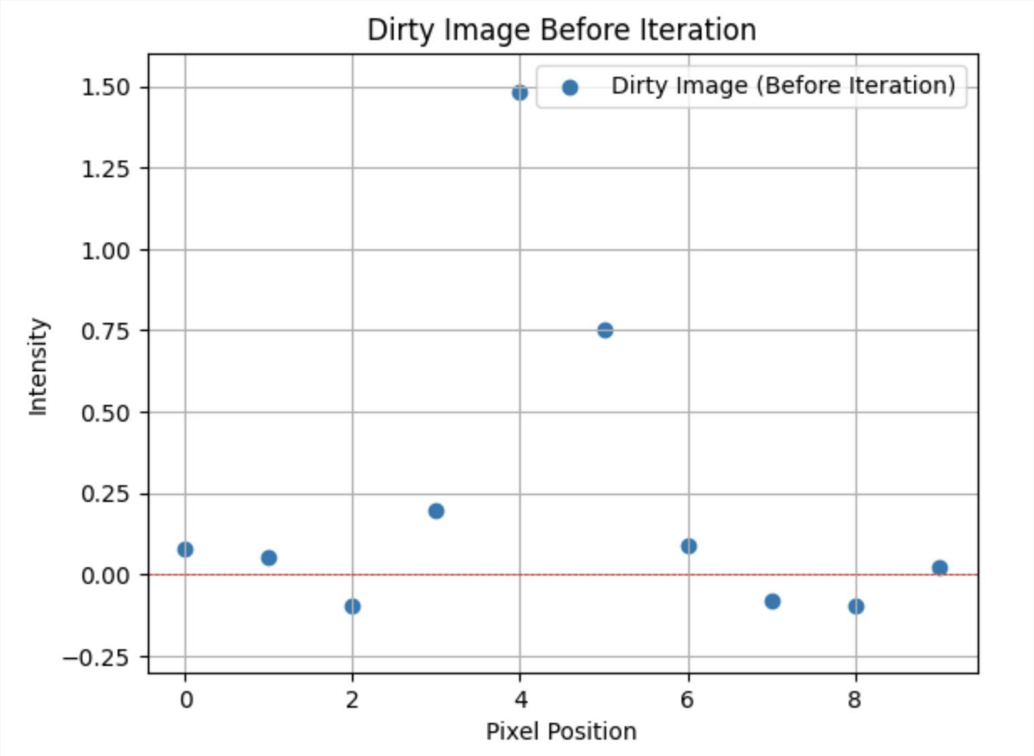
UNIVERSITY OF CAMBRIDGE

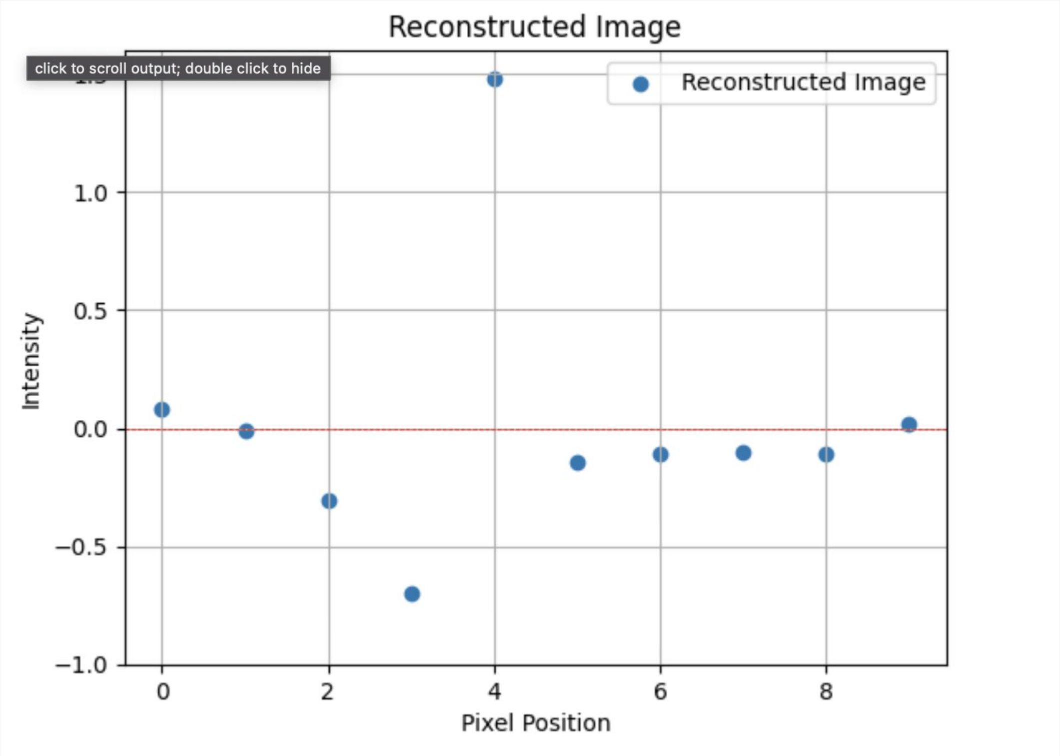
Högbom CLEAN - cons

1. Artifacts: Högbom CLEAN can introduce artifacts, especially in low signal-to-noise scenarios or with strong sources.
2. Parameter Sensitivity: Performance depends on parameter choices, requiring careful tuning.
3. Resolution Limitations: It struggles with sources smaller than resolution limits defined by the CLEAN beam.
4. More suitable for point-like sources: may encounter challenges with extended sources or complex emission structures

UNIVERSITY OF CAMBRIDGE

Högbom CLEAN - cons





Two sources are so close to each other, so that the weaker source didn’t have the chance to be added to the CLEAN component list

UNIVERSITY OF CAMBRIDGE

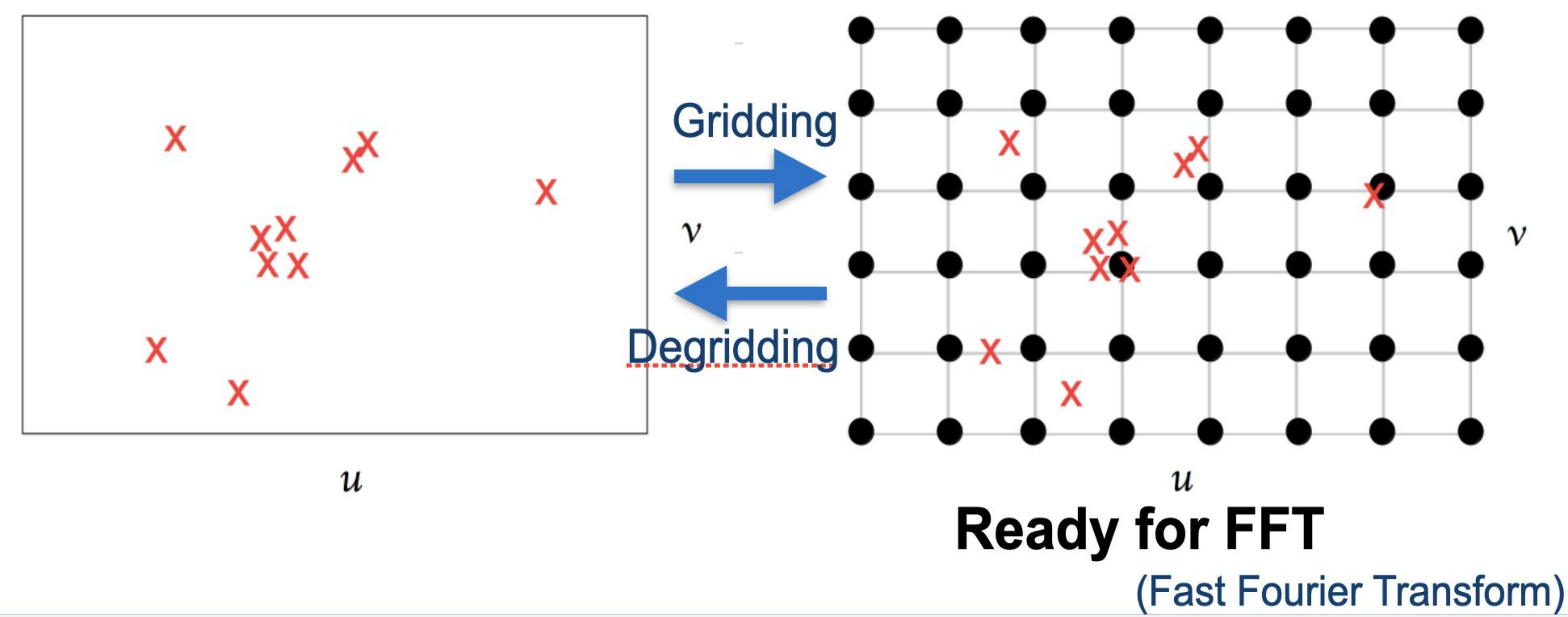
Högbom CLEAN

There are several variants and improvements upon the classic CLEAN algorithm used in radio interferometry imaging.

To understand some CLEAN variants, we need to understand Gridding

UNIVERSITY OF CAMBRIDGE

Gridding and iFFT



UNIVERSITY OF CAMBRIDGE

Gridding and iFFT

Direct Fourier Transform (DFT) is too slow...

↓

use Fast Fourier Transform (FFT)

↓

standard FFT algorithms works with evenly sampled data

↓

resampling the visibility data onto a regular grid in the uv-plane.

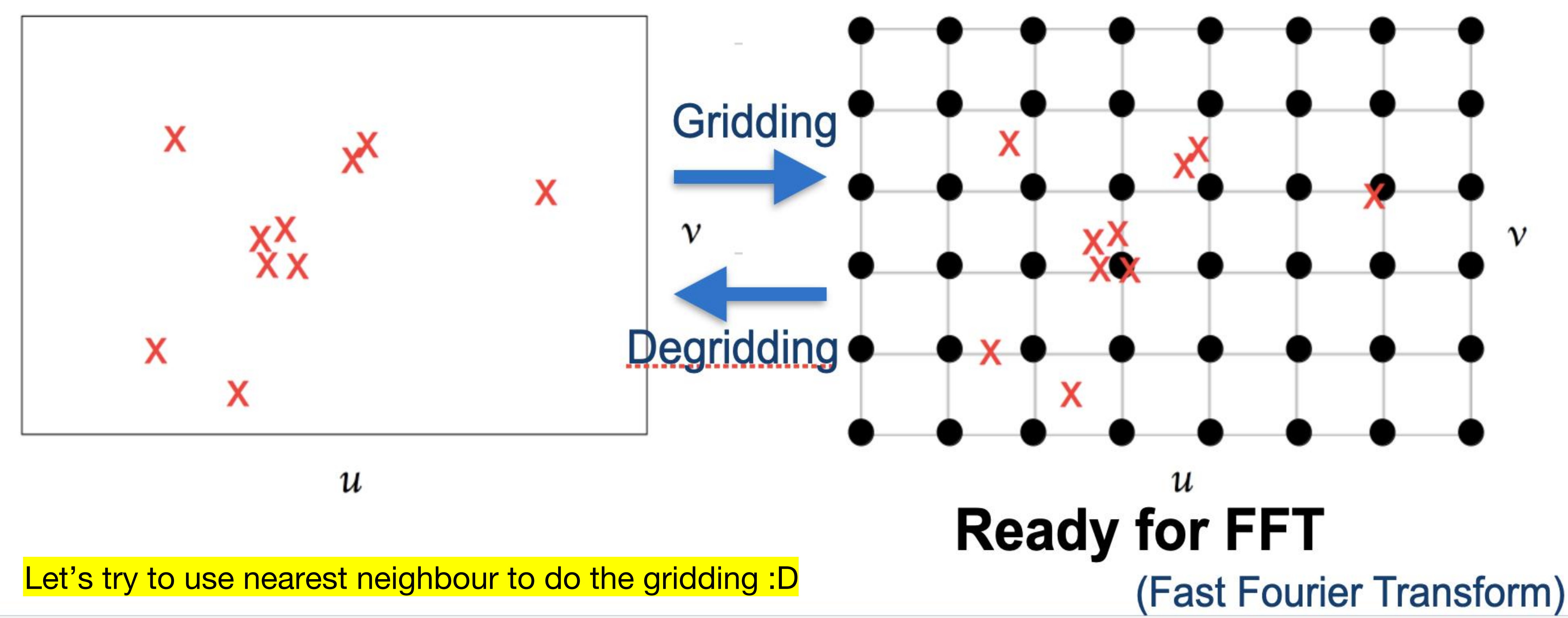
↓

FFT can be applied

Standard FFT algorithms require the data to be regularly sampled so as to exploit the periodicity of the data, meaning that the samples are evenly spaced.

UNIVERSITY OF CAMBRIDGE

Gridding and iFFT



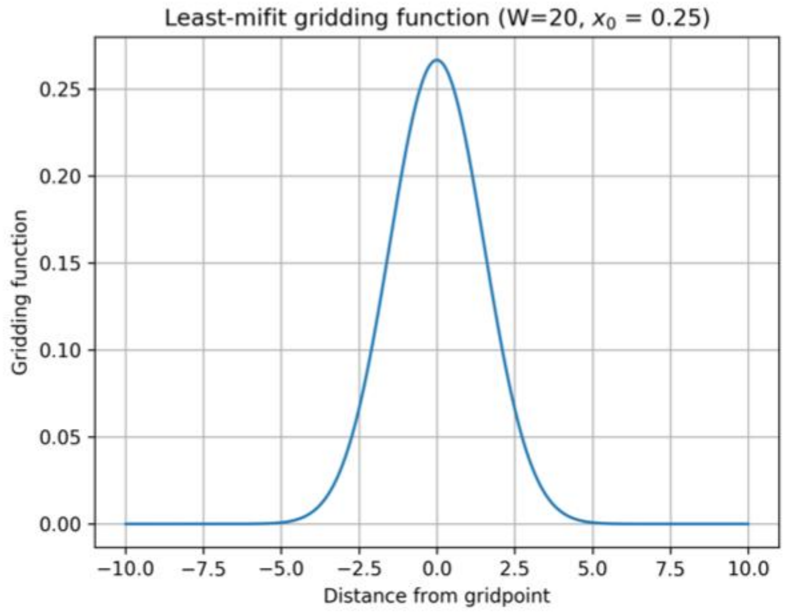
Let's try to use nearest neighbour to do the gridding :D

UNIVERSITY OF CAMBRIDGE

Gridding and iFFT

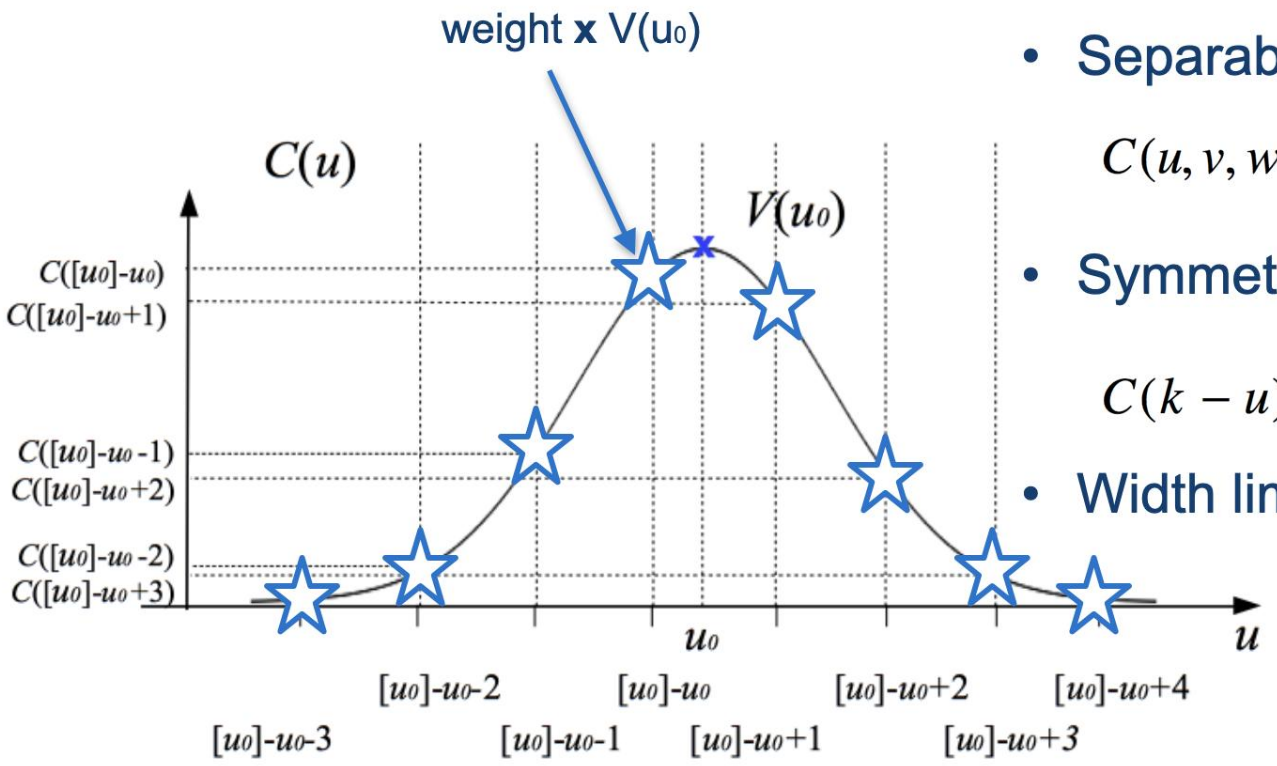
Gridding function C(u,v)

- Exponential function (including Gaussian)
- Sinc function
- Exponential times Sinc function
- Kaiser-Bessel function
- Least-misfit function (my PhD work)



UNIVERSITY OF CAMBRIDGE

Gridding function



- Separable

$$C(u, v, w) = C_u(u)C_v(v)C_w(w)$$

- Symmetric

$$C(k-u) = C(u-k)$$

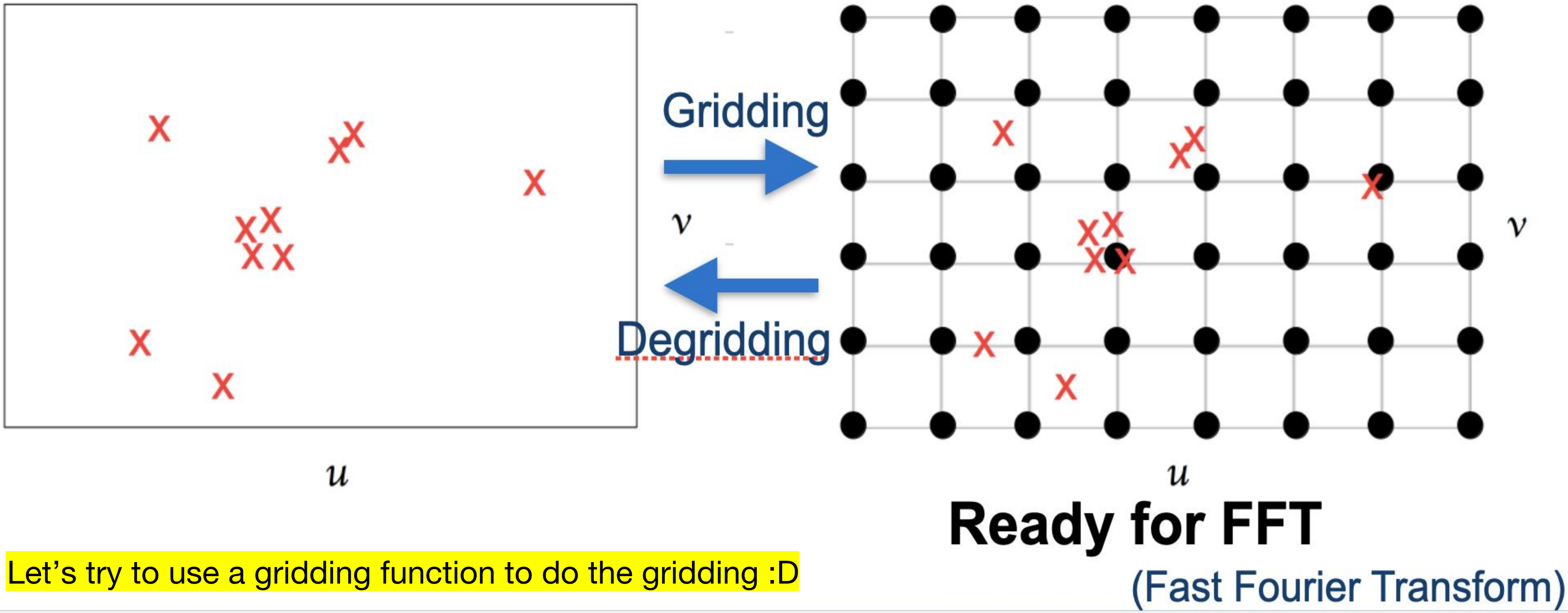
- Width limit *W*

$$\mathrm{III}(u,v)\Big(\big(V(u,v)S(u,v)\big) * C(u,v)\Big)$$

$[u]$ is the largest integer smaller than u, e.g. $[3.5] = 3$

UNIVERSITY OF CAMBRIDGE

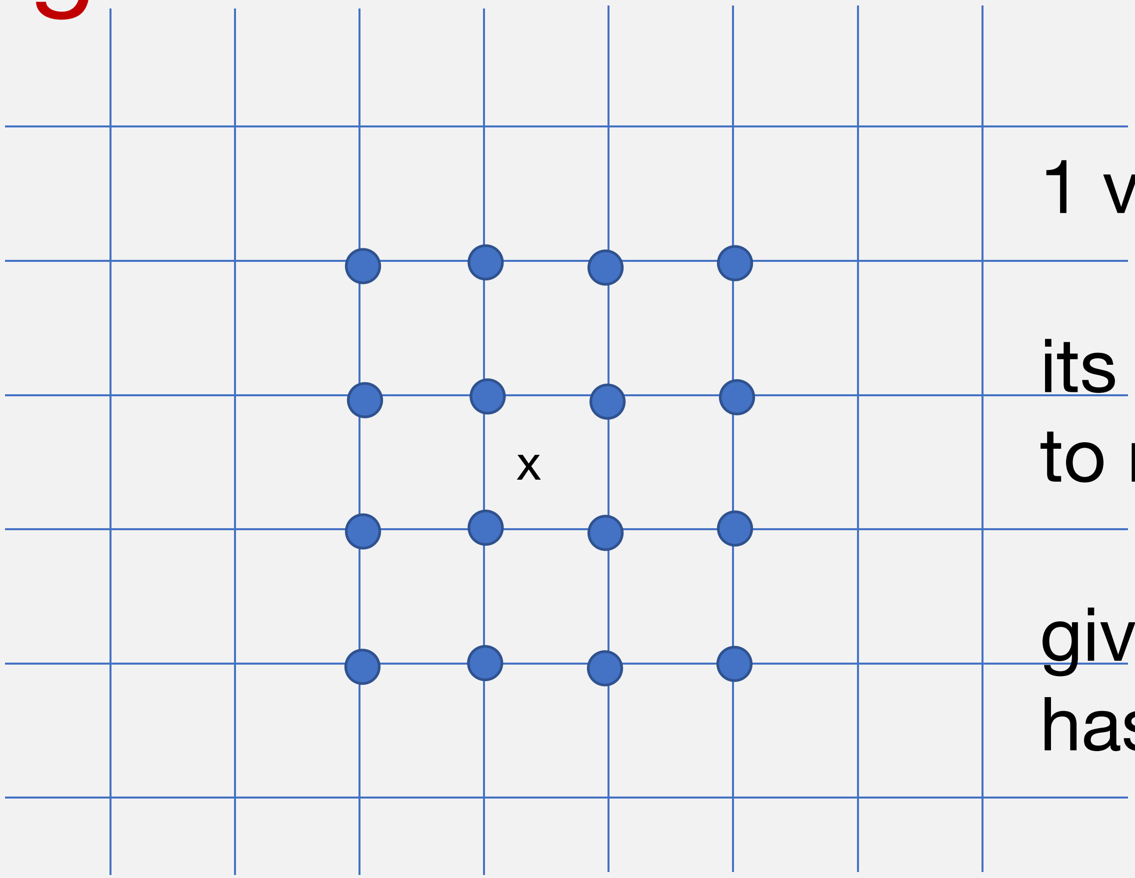
Gridding process



Let’s try to use a gridding function to do the gridding :D

UNIVERSITY OF CAMBRIDGE

Gridding and FFT



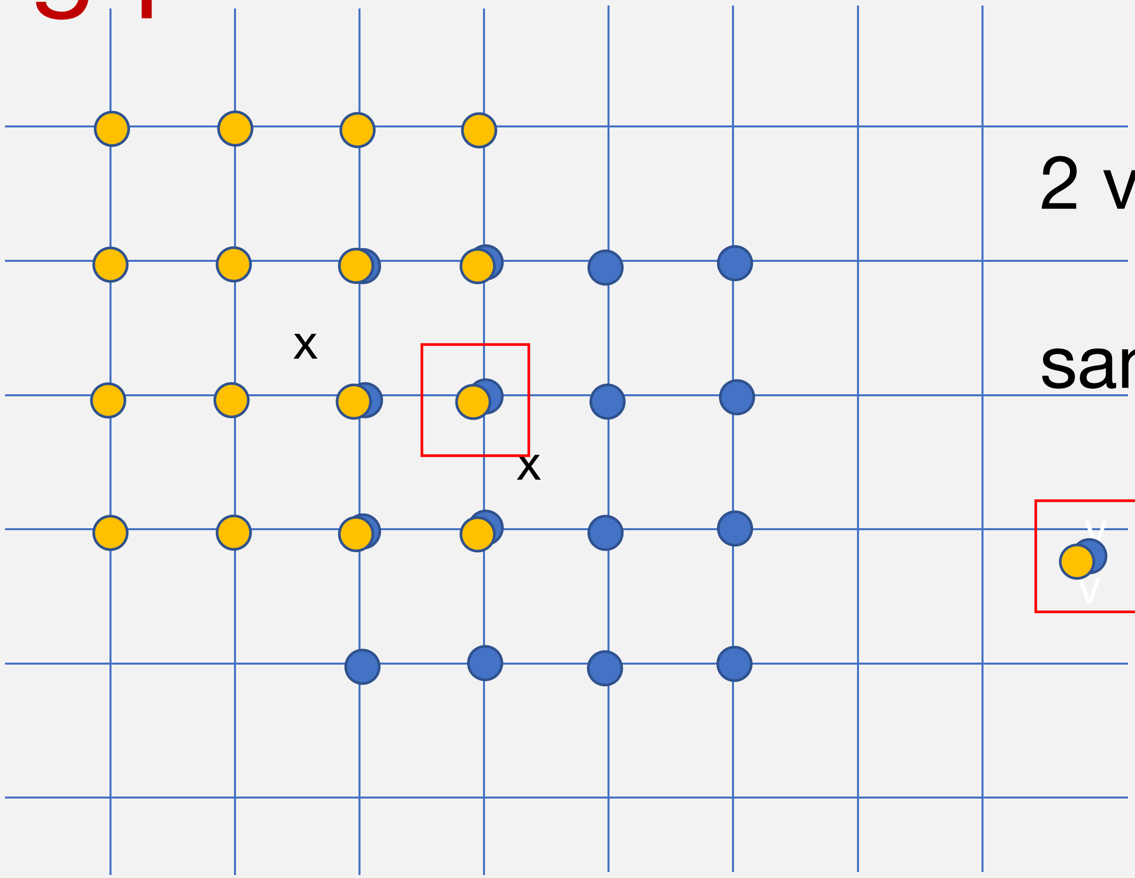
1 visibility

its value is to be assigned to nearby 4 by 4 grids

given the gridding function has a width of 4

UNIVERSITY OF CAMBRIDGE

Gridding process



2 visibility

same gridding function

sum the values

UNIVERSITY OF CAMBRIDGE

Gridding function

Apply inverse Fourier transform to gridded visibilities:

$$\mathcal{F}^{-1}(V(u,v)\ (u,v) *\ C(u,v)) = \mathcal{F}^{-1}(V(u,v)S(u,v))\ \boxed{\mathcal{F}^{-1}(C(u,v))}$$

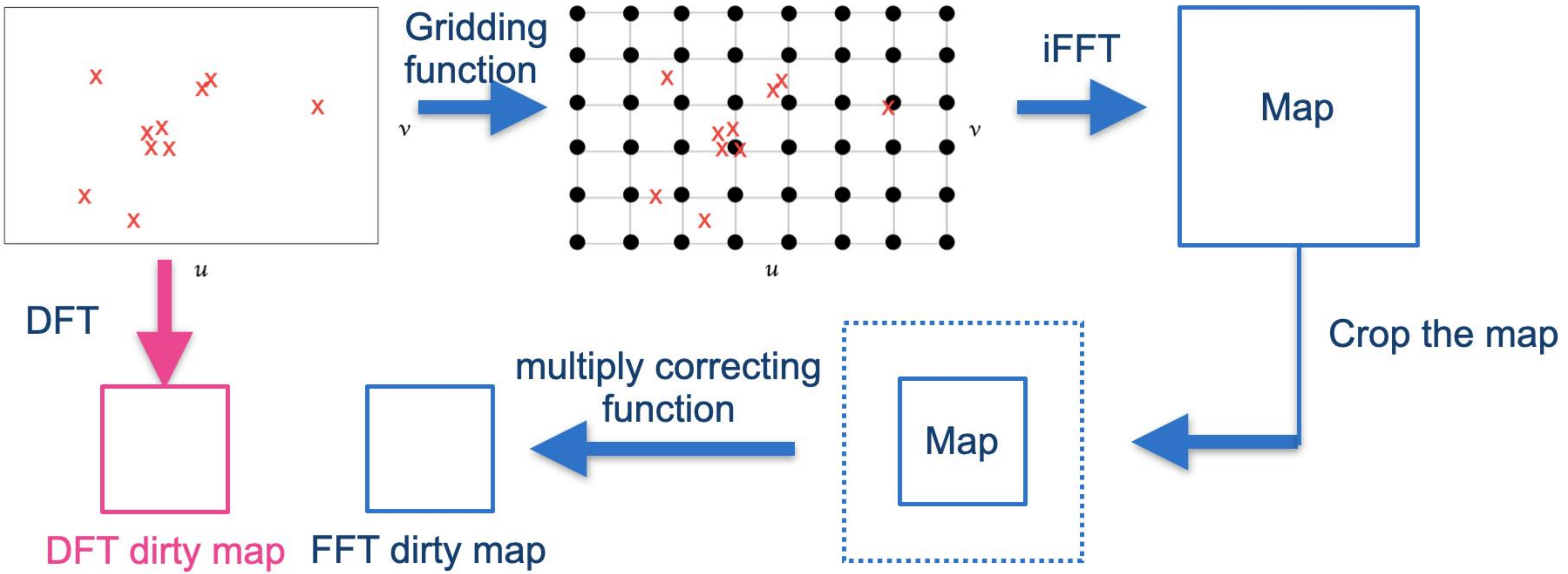
However, the FFT of the gridded data is not the dirty image. Instead, we multiply it with a "correcting function" to obtain the dirty image from the result image.

This correcting function corrects for the effects of the gridding operation on the Fourier transform.

UNIVERSITY OF CAMBRIDGE

Gridding and iFFT

1) Gridding visibility with a chosen gridding function on the *u*v plane.
2) Apply the inverse FFT to the gridded visibility data to obtain an image.
3) Crop the image and multiply the correcting function with the cropped image.



UNIVERSITY OF CAMBRIDGE

Gridding and iFFT

Why do we throw away the edge of the image?

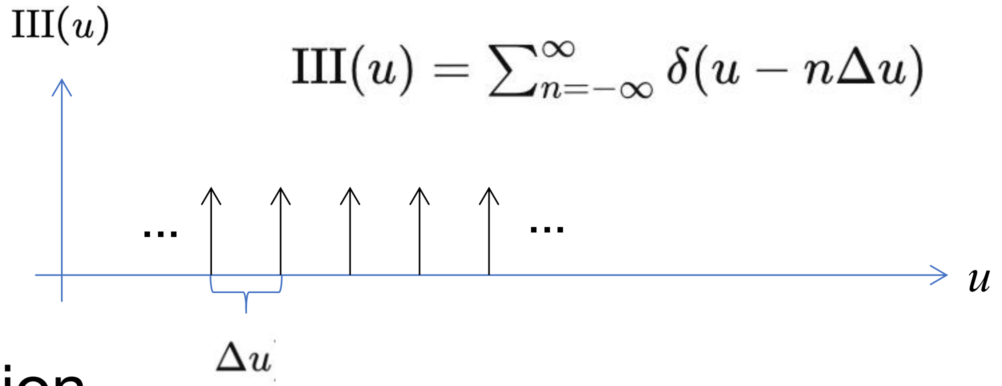
UNIVERSITY OF CAMBRIDGE

Gridding and iFFT

The Shah function is used to resample visibilities on grids:

$$\mathrm{III}(u,v)\Big(\big(V(u,v)S(u,v)\big) * C(u,v)\Big)$$

$$\mathrm{III}(u) = \sum_{n=-\infty}^{\infty} \delta(u - n\Delta u)$$



Its Fourier transform is also a Shah function.
But this Shah function is truncated because we have finite number of u-grids

To understand more about Shah function, please refer to
https://www.cs.unm.edu/~williams/cs530/shannon5.pdf

UNIVERSITY OF CAMBRIDGE

Gridding and iFFT

$$I_D = \mathcal{F}^{-1}\left\{ \text{III}_{\text{truncated}}(u,v)\left(\left(V(u,v)S(u,v)\right) * C(u,v)\right)\right\}$$

$$= \boxed{\mathcal{F}^{-1}\left\{\text{III}_{\text{truncated}}(u,v)\right\}} * \mathcal{F}^{-1}\left\{\left(V(u,v)S(u,v)\right) * C(u,v)\right\}$$

Periodic!!!

To understand more about Shah function, please refer to
https://www.cs.unm.edu/~williams/cs530/shannon5.pdf

UNIVERSITY OF CAMBRIDGE

Gridding and iFFT

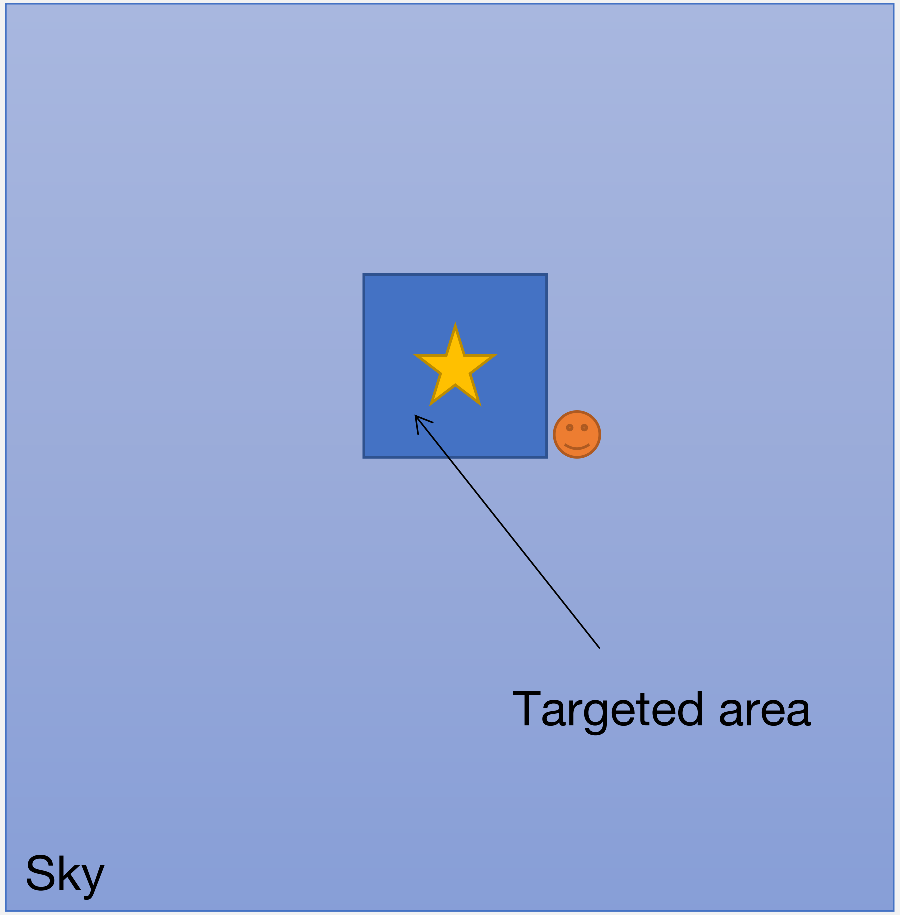
But there are two effects of using Gridding + iFFT that we need to consider:

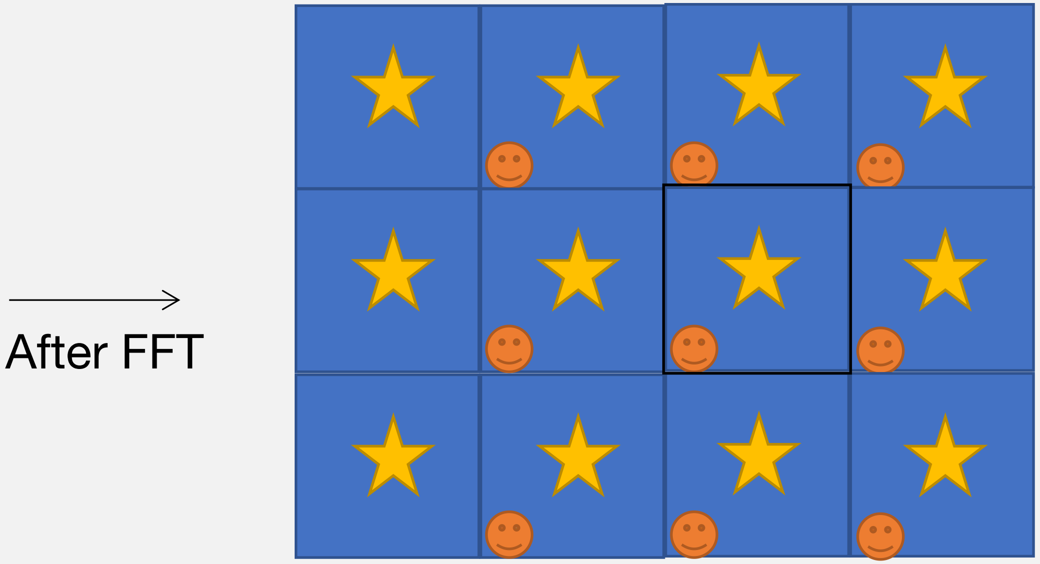
1) Ghost source: aliasing

2) Undersampling

UNIVERSITY OF CAMBRIDGE

Ghost source: alising





UNIVERSITY OF CAMBRIDGE

Ghost source: alising

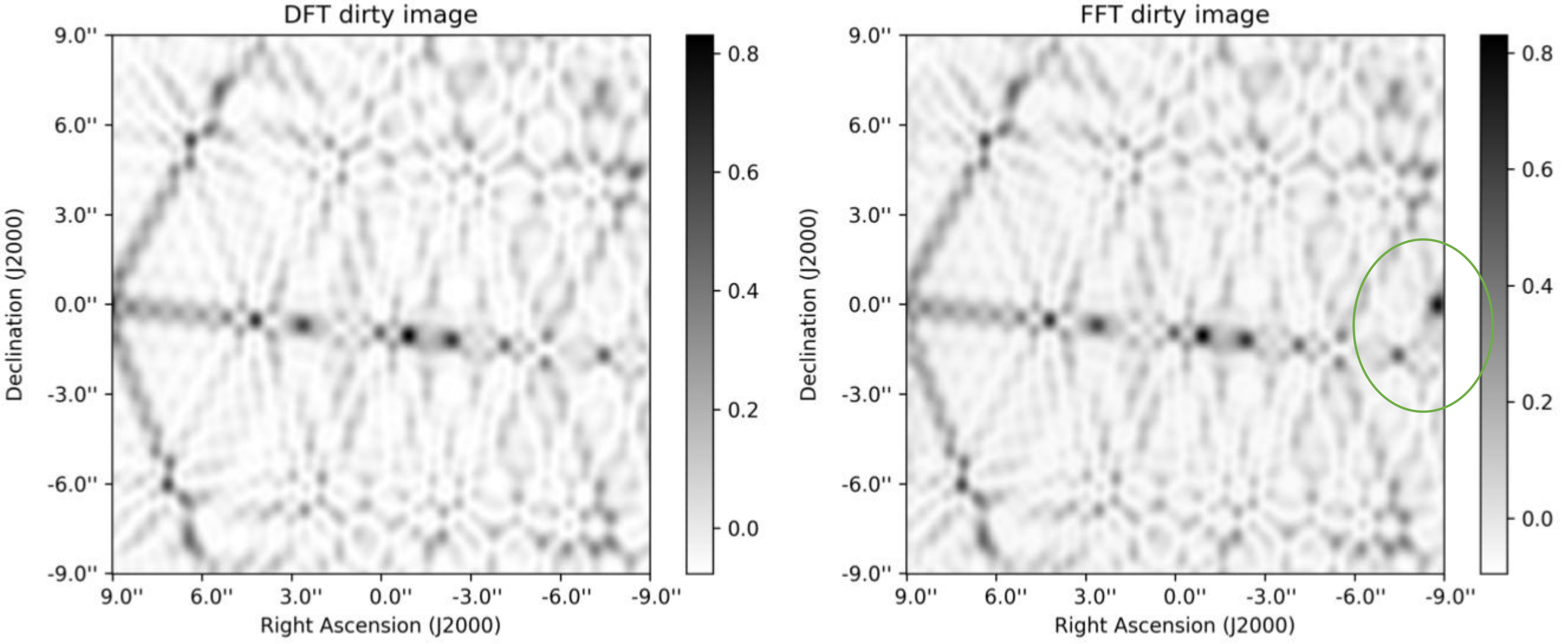


Figure 3.1: Demonstration of the aliasing effect due to the periodicity of FFT. DFT dirty image (top left) and FFT dirty image (top right) on the simulated point–source visibility data. The alias is obvious in the FFT dirty image, with a brightness of 0.82 Jy.

UNIVERSITY OF CAMBRIDGE

Gridding and iFFT

But there are two effects of using Gridding + iFFT that we need to consider:

1) Ghost source: aliasing

2) Undersampling

UNIVERSITY OF CAMBRIDGE

Undersampling

To ensure that the visibility data are properly sampled according to the Nyquist sampling theorem, you need to determine an appropriate grid size in the spatial frequency domain (uv-space)

$$\Delta u \leq \frac{1}{2u_{max}}$$

You can have smaller grid size than this, but no bigger than it....

UNIVERSITY OF CAMBRIDGE

Gridding and iFFT

1) Gridding + iFFT
2) Direct Fourier Transform (DFT)

Which operation is the most computationally expensive?

Gridding and iFFT - computaional cost

- DFT operation has a complexity of $O(N^2)$, N is the number of visibility data
- FFT operation has a complexity of $O(M\log M)$, M is the number of grid points in the uv-plane, usually in the same order of magnitude as the image size
- Gridding operation has a complexity of $O(W^2N)$, where W is the width of the gridding function, usually is less than 7, and much less than N, therefore can be simpfied to O(N)

Gridding and Degridding

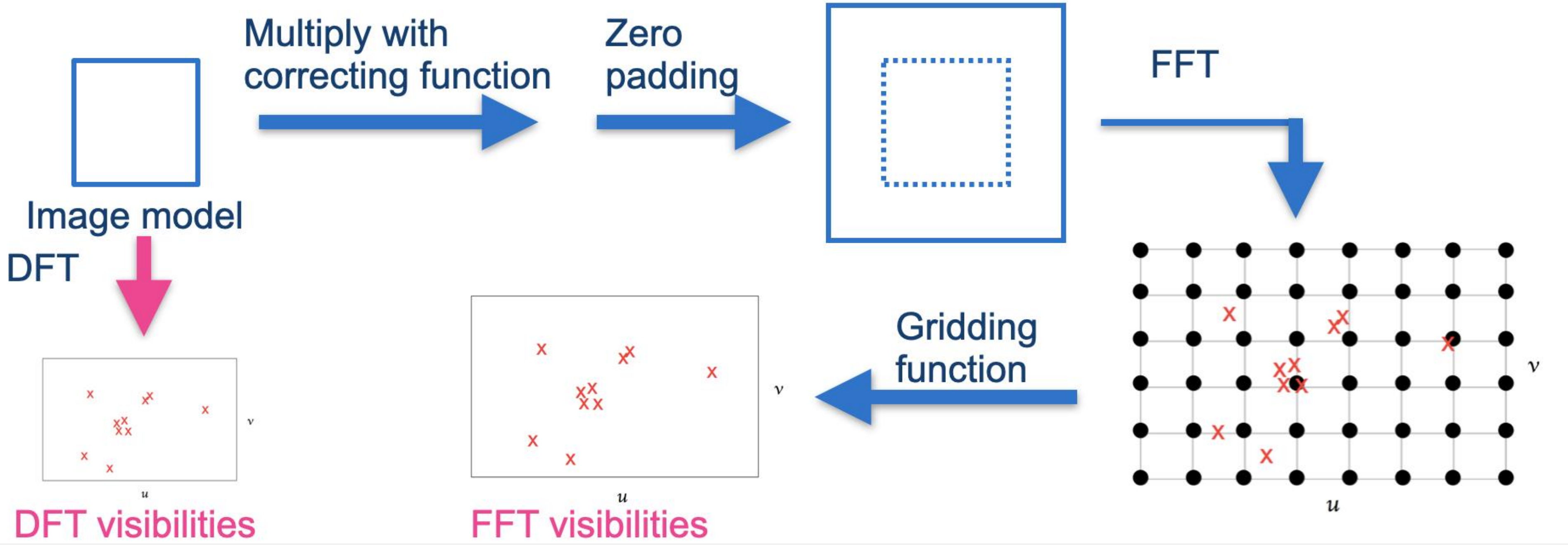
Gridding involves mapping unevenly distributed visibility measurements onto a regular grid in the uv-plane. This process prepares the data for the inverse Fast Fourier Transform (iFFT), which is then applied to generate an image from the gridded visibilities.

Degridding is the reverse process. It entails taking an image, performing FFT to transform it back to the visibility plane, and restoring all visibilities to their original grid positions.

UNIVERSITY OF CAMBRIDGE

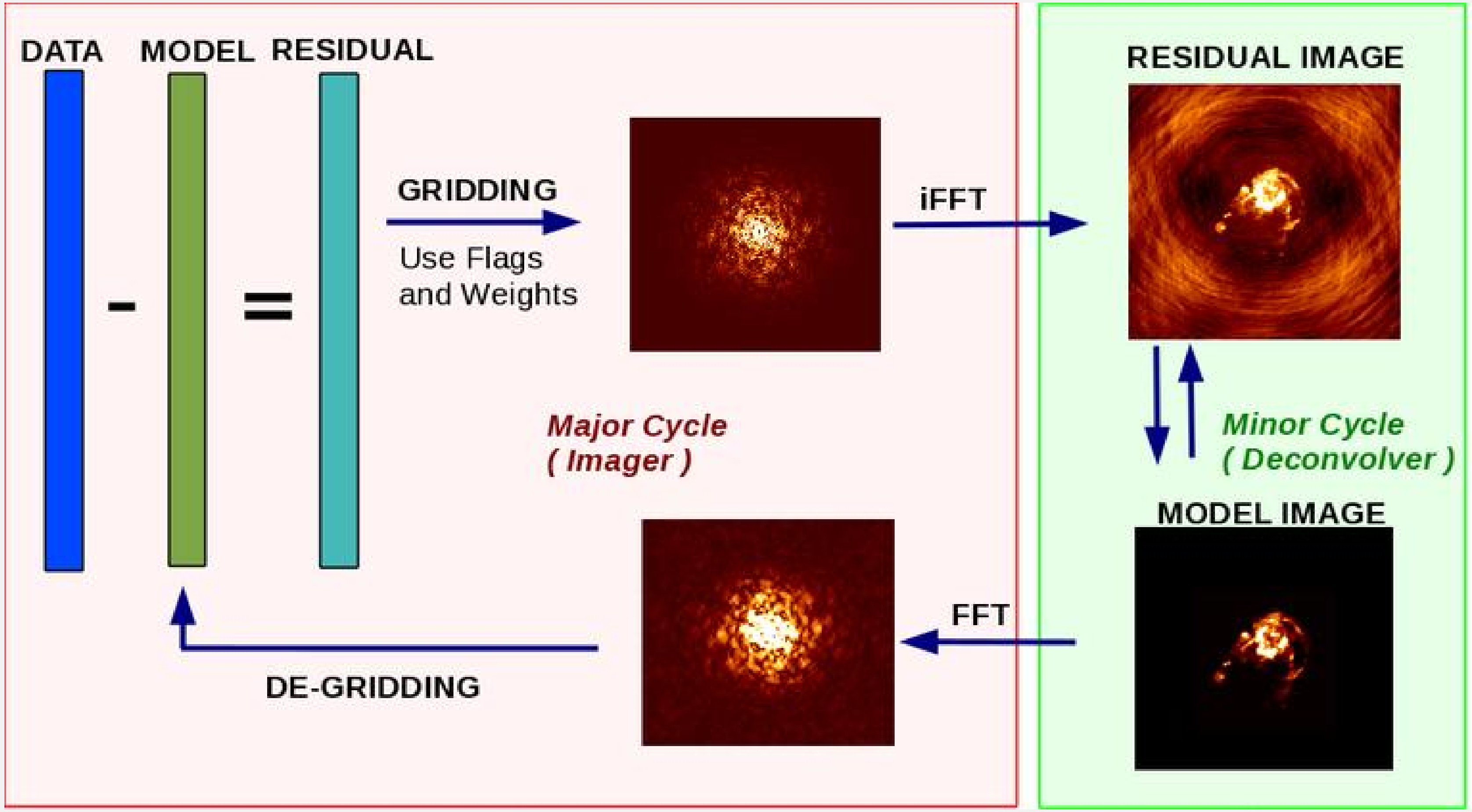
Degridding

1) Multiply the model image by the correcting function.
2) Pad zeros to the outer edge of the corrected model image, which is the opposite operation of image cropping during gridding.
3) Apply FFT to the corrected image.
4) Correlate the gridding function with the FFT results to obtain the visibility model.



UNIVERSITY OF CAMBRIDGE

CLEAN variant - Major and Minor Cycles



Resource: https://casadocs.readthedocs.io/en/stable/notebooks/synthesis_imaging.html#Introduction

CLEAN variant - Major and Minor Cycles

Each of you will explain one step :D

UNIVERSITY OF CAMBRIDGE

Cotton-Schwab CLEAN

Schwab, F. R. (1984b), ``Relaxing the isoplanatism assumption in self-calibration; applications to low-frequency radio interferometry'', Astron. J., 89, 1076-1081.

```
D = initialise_dirty_image()  # Initialise with observed visibility data
R = D.copy()  # Residual image starts as the dirty image
clean_components = [] # Initialise clean component list
clean_components = []

# Major cycle
while not convergence_criteria_met():
    # Minor cycle
    while not minor_cycle_convergence_criteria_met():
        # Extract CLEAN components
        peak = find_peak(R)
        component = estimate_clean_component(peak)
        clean_components.append(component)
    # Degridding and subtraction
    model_visibility_data = degrid(clean_components)
    model_visibility_data = correct_gridding_errors(model_visibility_data)
    residual_visibility_data = subtract_clean_from_visibility(D, model_visibility_data)

    # Update residual image at the end of major cycle
    R = residual_visibility_data

# Final residual image is now the model visibility data
final_residual_image = R

# Final image reconstruction
final_image = reconstruct_image(D, clean_components)
```

UNIVERSITY OF CAMBRIDGE

CLEAN variant - Major and Minor Cycles

Why do we want to work on the visibility plane?

Aliasing and Gridding Errors: Visibility data can suffer from aliasing and gridding errors, especially when transforming between the image plane and the visibility plane.

Efficiency: Some deconvolution algorithms, like CLEAN, are more computationally efficient when applied in the visibility plane.

Handling of Noise: Noise properties can be better understood and managed in the visibility domain, which can improve the accuracy of deconvolution algorithms. Additionally, noise statistics in the visibility plane are typically simpler than in the image plane.

UNIVERSITY OF CAMBRIDGE

Maxmum Entropy Method (MEM)

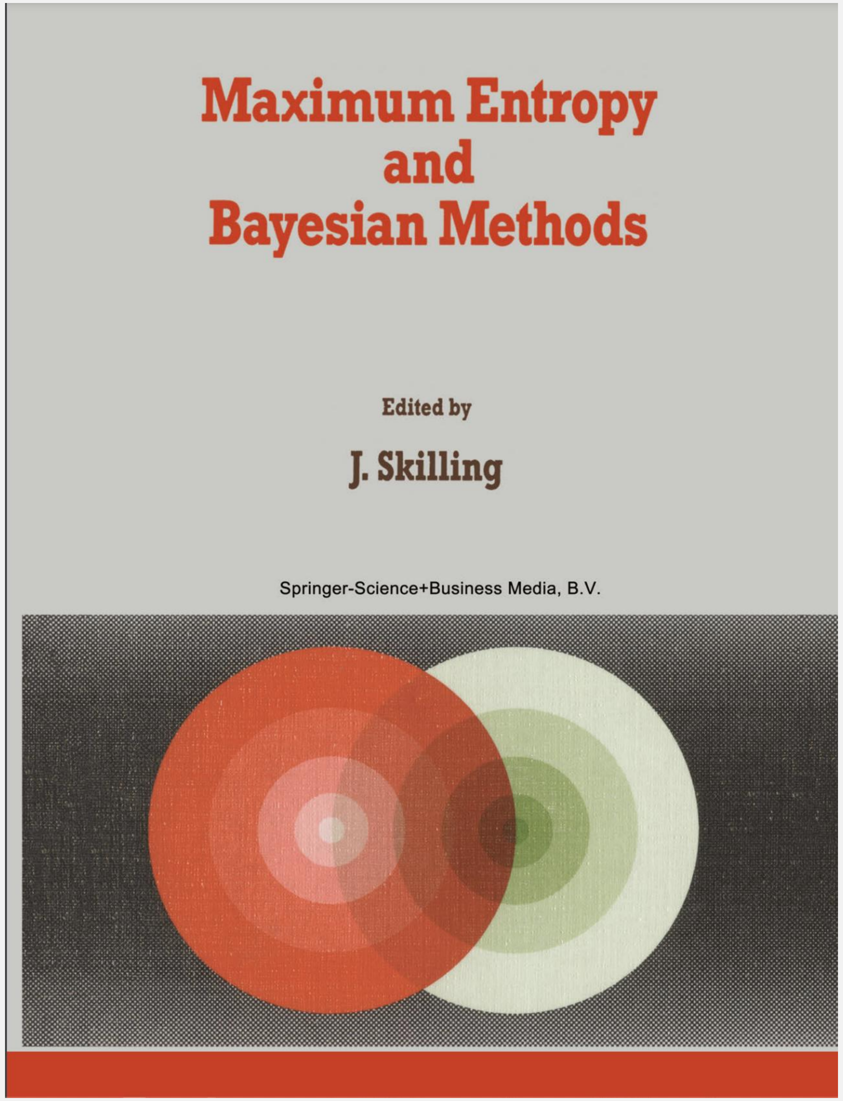
Pros

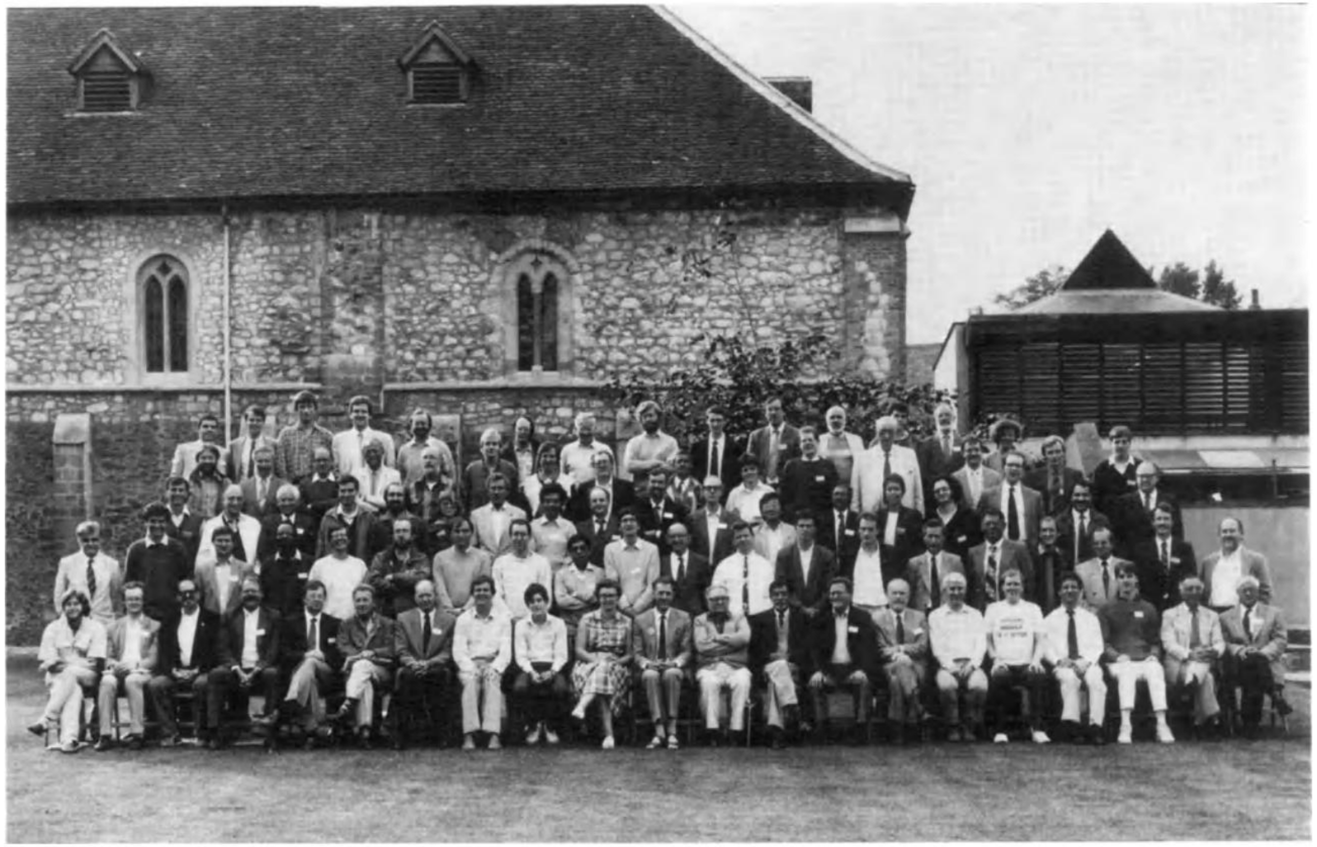
- Good with extended sources, complex source structures
- Incorporation of prior knowledge or constraints about the image
- Noise Robustness

Cons

- Difficult to understand
- Computationally intensive
- Risk of overfitting
- Sensitivity to Model Assumptions
- Parameter Tuning

UNIVERSITY OF CAMBRIDGE

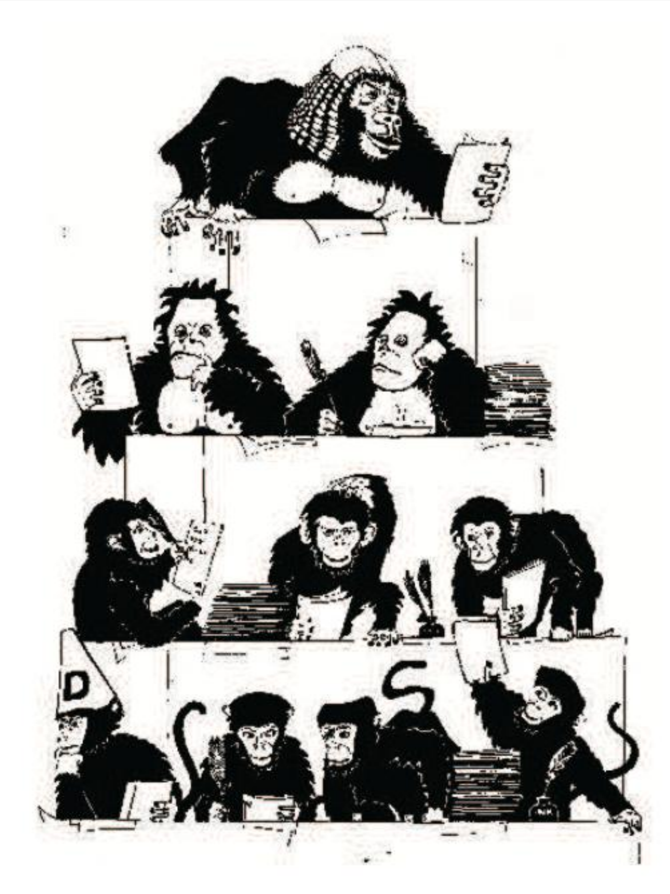




Photograph courtesy of W. Eaden Lilley, Cambridge

Cambridge, England, 1988

Maxmum Entropy Method (MEM)



Credit: Prof. Stephen Gull

Infinite Monkey Theorem

Given enough time, a monkey randomly hitting keys on a typewriter would eventually produce the complete works of William Shakespeare

Therefore, the monkeys can also throw a large number of grains of luminance ('atoms') into the pixels of the object.

UNIVERSITY OF CAMBRIDGE

Maxmum Entropy Method (MEM)

Steve Gull and John Skilling
Maximum Entropy Data Consultants Ltd.

The programs (old and not so old)

- MemSys5 (1990) Quantified MaxEnt for generalised imaging problems and spectroscopy.
- MemSys4/RaMem (1991) Quantified MaxEnt for non-linear problems.
- MitSyS (1998) Atomic priors for spectral line and image reconstruction.
- BayeSys1 (1999) Parameter fitting using Markov Chain Monte Carlo.
- NeuroSys (2000) MemSys4/RaMem used for neural net training.
- BayeSys2 (2001) Atomic priors with many parameters — replaces MitSys and BayeSys1.

Credit: Prof. Stephen Gull

https://www.mpe.mpg.de/~aws/integral/issw/memsys5.pdf

Infinite Monkey Theorem

From all the possible images, MEM selects a positive and smooth image that

1) fits the data and any imposed constraints, to within the noise level, and

2) also has maximum entropy

MEM reconstruction typically involves an iterative optimisation process. Starting with an initial guess for the image, the algorithm iteratively updates the image to increase its entropy while minimising the difference between the observed and predicted data.

UNIVERSITY OF CAMBRIDGE

Maxmum Entropy Method (MEM)

1) fits the data and any imposed constraints, to within the noise level, and →

Chi-square minimisation:

$$\chi^2 = \sum_{u,v} \frac{|V_{\text{observe}}(u,v) - V_{\text{predict}}(u,v)|^2}{\sigma^2(u,v)}$$

The goal is to minimize the value to effectively find the best-fitting model that explains the observed data within the noise levels.

---

2) also has maximum entropy. →

In information theory, entropy is a measure of uncertainty or disorder in a system. In the context of image reconstruction, higher entropy implies greater uncertainty about the pixel values in the image.

MEM seeks to avoid overfitting the data. By favoring solutions with higher entropy, MEM tends to produce smoother and more generalised images that are less susceptible to noise.

UNIVERSITY OF CAMBRIDGE

Bayesian MEM - explained

Bayes' Theroem

$$\text{Inference} = \frac{\text{Likelihood} \times \text{Prior}}{\text{Evidence}}$$

$$P(I|D) = \frac{P(D|I)P(I)}{P(D)}$$

D: Visibility Observation (can be denoted as V)
I: Predicted Image
$S(I)$ is the entropy of the image I
α controls the strength of the entropy regularisation

In the Bayesian MEM framework, the prior distribution on the image is chosen to be proportional to the exponential of the entropy of the image, which encodes the available information efficiently while making minimal assumptions beyond what is known.

$$P(I) \propto exp\,[\alpha S(I)]$$

Objective function can be something like:

$$I_{\mathrm{MAP}} = \mathrm{argmax}_I \ \mathrm{logP(D|I)}$$

$$= \mathrm{argmax}_I \Big(\mathrm{logP(D|I)} + \mathrm{logP(I)} - \mathrm{logP(D)}\Big)$$

$$\propto \mathrm{argmax}_I \Big(\mathrm{logP(D|I)} + \alpha S(I)\Big)$$

UNIVERSITY OF CAMBRIDGE

Bayesian MEM - explained

$$\mathrm{argmax}_I\Big(\log \mathrm{P}(\mathrm{D}|\mathrm{I}) + \boxed{\alpha S(I)}\Big)$$

One possible Entropy Definition

$$\mathrm{S}(\mathrm{I}) = \sum_{x,y}(I(x,y) - M(x,y) - I(x,y)\log(I(x,y)/M(x,y)))$$

Where $I(x,y)$ is the value of our predicted image at (x,y), and $M(x,y)$ is the image value of the “default image”

The “default image” is usually chosen at the beginning of the reconstruction process and remains fixed throughout the iterations. It provides a baseline against which the variation or change in the reconstructed images is measured. By keeping the default image constant, the regularisation encourages the reconstructed images to be similar to the default image in terms of entropy, while also fitting the observed data.

For example, the default image could be set to a uniform image representing the absence of any sources, or it could be initialiaed with a low-resolution image reconstructed from the initial visibility data.

This default image serves as a starting point for the iterative reconstruction process and helps guide the regulariaation to produce solutions that are consistent with the underlying structures in the observed data.

UNIVERSITY OF CAMBRIDGE

Bayesian MEM - explained

$$\mathrm{argmax}_I\left(\boxed{\log \mathrm{P}(\mathrm{D}|\mathrm{I})} + \alpha S(I)\right)$$

One possible likelihood

$$\mathrm{P}(\mathrm{D}|\mathrm{I}) = \frac{1}{\sqrt{(2\pi)^N \det(\mathrm{C})}} \exp\left(-\frac{1}{2}\,(D - V_{\mathrm{predicted}})^{\mathrm{T}}\,\mathrm{C}^{-1}\,(D - V_{\mathrm{predicted}})\right)$$

N is the total number of visibility measurements.
D is the vector of observed visibility data.
$V_{\mathrm{predicted}}$ is the vector of predicted visibility data based on the current reconstructed image
C is the covariance matrix representing the noise characteristics of the visibility data.

The likelihood function quantifies the probability of observing the data D given the current reconstructed image I, taking into account the noise characteristics of the visibility data. By maximising the likelihood function, we aim to find the reconstructed image that best explains the observed data within the framework of the model.

UNIVERSITY OF CAMBRIDGE

Bayesian MEM - explained

$$\mathrm{argmax}_I \left(\log P(D|I) + \alpha \sum_{x,y} \left(I(x,y) - M(x,y) - I(x,y) \log\left(I(x,y)/M(x,y) \right) \right) \right)$$

Step1: Initialisation
Initialise the reconstructed image I with a default image or an initial guess.

Step2: Iterative Update
Repeat the following steps until convergence or a maximum number of iterations is reached:

1) Use the current reconstructed image I to predict the visibility data $V_{predict}$
2) Calculate likelihood $P(D|I)$ based on the visibility data $V_{observe}$ and $V_{predict}$
3) Use the current reconstructed image I to calculate Entropy $S(I)$
4) Minimise the objective function and obtain the updated reconstructed image I

Stopping criteria
- Convergence of the reconstructed image
- Convergence of the objective function
- Maximum number of iterations reached
- Stability of other metrics

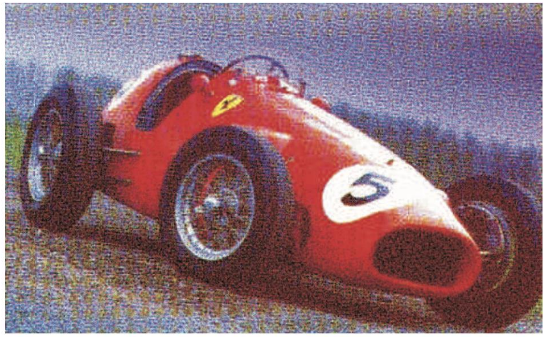
Bayesian MEM - explained

- Positivity Constraint: Ensure all pixel intensities are non-negative.
- Smoothness Constraint: Penalise abrupt intensity changes across neighboring pixels.
- Sparsity Constraint: Encourage solutions with a small number of non-zero pixel intensities.
- Total Variation (TV) Constraint: Promote smooth intensity gradients and sharp edges in the image.
- Flux Density Constraint: Ensure the total flux density of the reconstructed image matches the observed value.
- Prior Information Constraint: Incorporate prior knowledge or information about the sources being imaged.

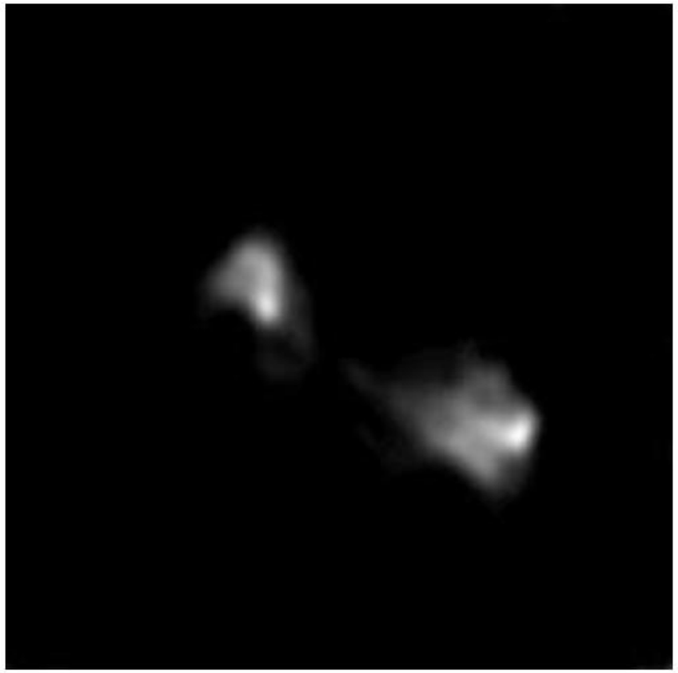
Bayesian MEM - explained

MAXENT IMAGE RECONSTRUCTION

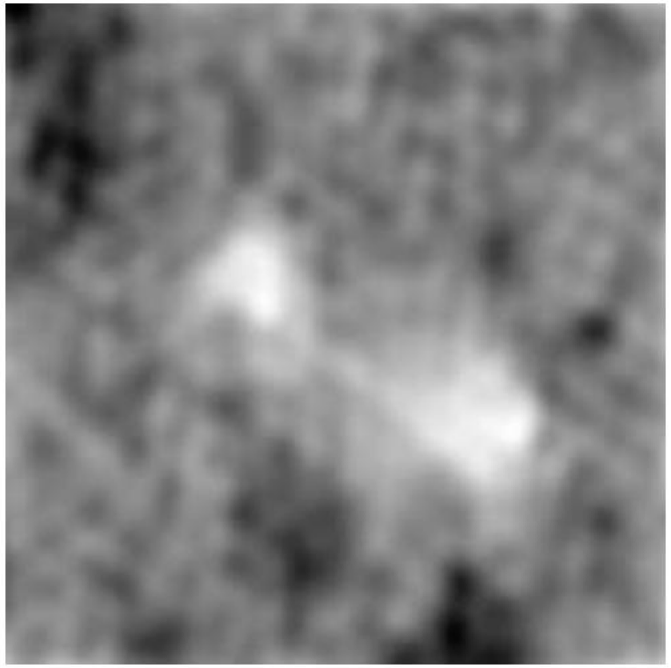
Out of focus Ferrari image



MaxEnt reconstruction

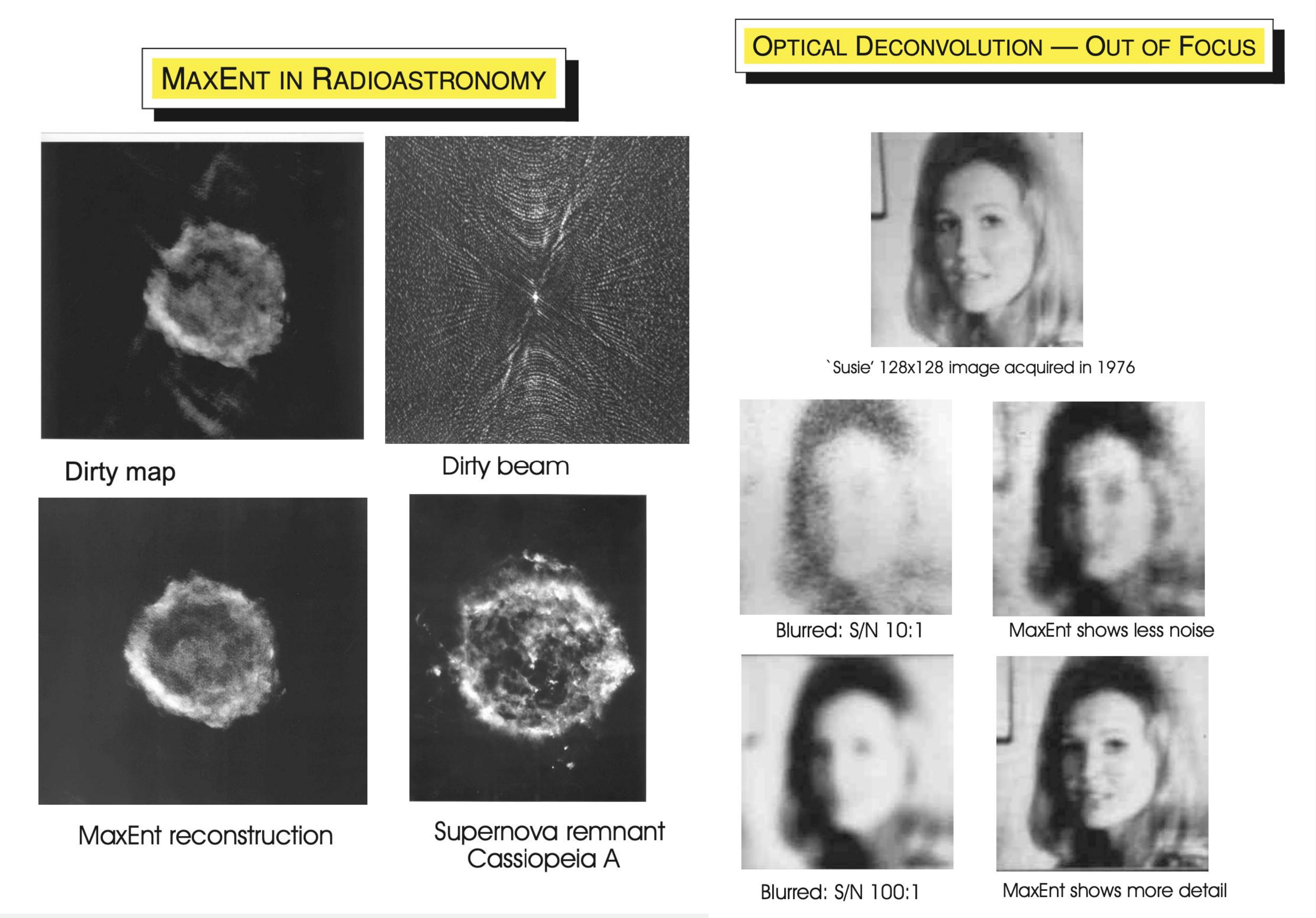


MaxEnt Reconstruction



Dirty Residuals (0.2%)

Credit: Prof. Stephen Gull



1980s

Credit: Prof. Stephen Gull

UNIVERSITY OF CAMBRIDGE

Deconvolution

Inverse
Fourier Transform

Observed
Visibility data

A DIRTY
image!!!

A Deconvolved
image!!!

UNIVERSITY OF CAMBRIDGE

Deconvolution

Observed Visibility data

Inverse Fourier Transform

A DI
imag

A Dec
image!!!

UNIVERSITY OF CAMBRIDGE

Wide Field Imaging

$$\boxed{\left(\sqrt{1-l^2-m^2}-1\right)w \approx 0}$$

$$V(u,v) = \int\int I(l,m)e^{-2\pi i\,(ul+vm)}\,dl\,dm$$

$$\Updownarrow$$

$$I(l,m) = \int\int V(u,v)\,e^{2\pi i\,(ul+vm)}\,du\,dv$$

UNIVERSITY OF CAMBRIDGE

Wide Field Imaging

When the filed of view is wide, the w-term cannot be ignored.

$$\boxed{V(u,v,w)} = \int\int \frac{\mathrm{d}l\,\mathrm{d}m}{\sqrt{1-l^2-m^2}} \boxed{I(l,m)} \exp[-i2\pi(ul+vm+\boxed{w(\sqrt{1-l^2-m^2}-1)}]$$

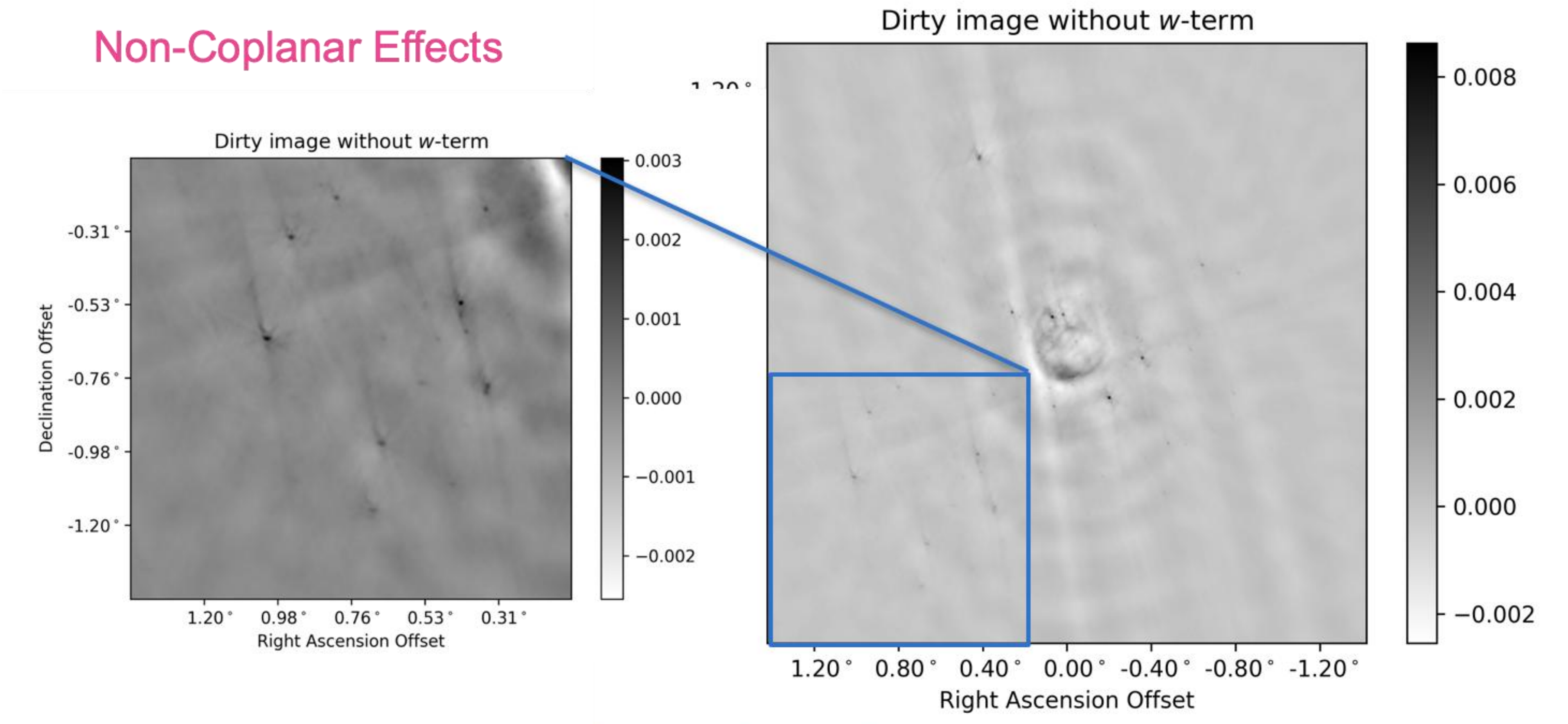
Visibilities

Sky brightness

w-term

UNIVERSITY OF CAMBRIDGE

Wide Field Imaging



VLA D-array observation
of the supernova remnant
G55.7+3.4

UNIVERSITY OF CAMBRIDGE

Wide Field Imaging

$$\boxed{V(u,v,w)} = \int\int \frac{dl\,dm}{\sqrt{1-l^2-m^2}} \boxed{I(l,m)} \exp[-i2\pi(ul + vm + \boxed{w(\sqrt{1-l^2-m^2}-1)})]$$

Visibilities

Sky brightness

w-term

There are several existing methods to take care of the w-term,

Today, we introduce the W-Stacking method.

UNIVERSITY OF CAMBRIDGE

W-Stacking method

W-Stacking method is implemented in WSCLEAN (Offringa et al. 2014).

$$V(u, v, w) = \int\int \frac{\mathrm{d}l\mathrm{d}m}{\sqrt{1-l^2-m^2}} I(l, m) \exp[-i2\pi(ul + vm + w(\sqrt{1-l^2-m^2} - 1)]$$

$$\frac{I(l,m)(w_{max}-w_{min})}{\sqrt{1-l^2-m^2}} \propto \int_{w_{min}}^{w_{max}} \int\int \mathrm{V(u,v,w)exp}\left(i2\pi\left(\mathrm{ul}+\mathrm{vm}+\mathrm{w}\left(\sqrt{1-l^2-m^2}-1\right)\right)\right)dudvdw$$

$$= \int_{w_{min}}^{w_{max}} \exp\left(i2\pi\mathrm{w}\left(\sqrt{1-l^2-m^2}-1\right)\right)\mathrm{d}w \int\int \mathrm{V(u,v,w)exp}(i2\pi(\mathrm{ul}+\mathrm{vm}))dudv$$

$$= \sum_{n=0}^{N_w-1} \exp\left(i2\pi w_n\left(\sqrt{1-l^2-m^2}-1\right)\right)\mathrm{d}w \int\int \mathrm{V(u,v,w)exp}(i2\pi(\mathrm{ul}+\mathrm{vm}))dudv$$

UNIVERSITY OF CAMBRIDGE

W-Stacking method

$$\frac{I(l,m)\left(w_{max}-w_{min}\right)}{\sqrt{1-l^2-m^2}} \propto \sum_{n=0}^{N_w-1} \exp\left(i2\pi w_n\left(\sqrt{1-l^2-m^2}-1\right)\right)\mathrm{d}w \int\int \mathrm{V}(\mathrm{u},\mathrm{v},\mathrm{w})\exp\left(i2\pi(\mathrm{ul}+\mathrm{vm})\right)dudv$$

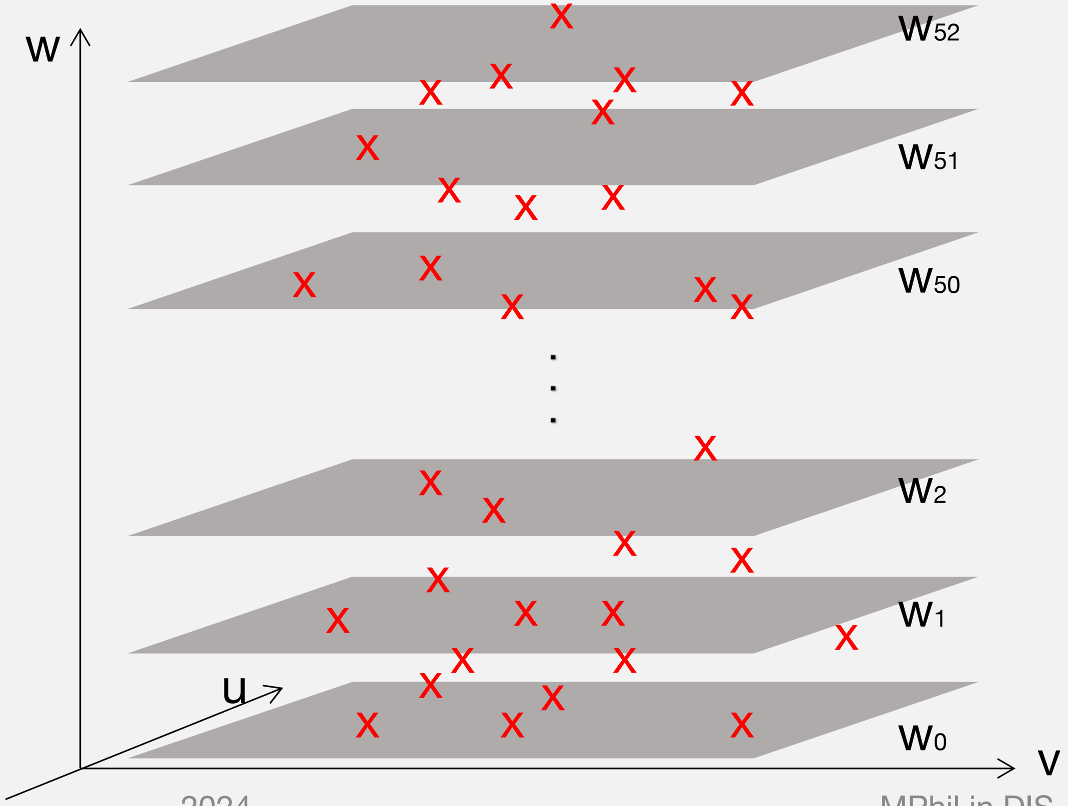
In theory, we can have one w-plane for each w value...

In practice, we don't...

UNIVERSITY OF CAMBRIDGE

W-Stacking method

$$\frac{I(l,m)\left(w_{max}-w_{min}\right)}{\sqrt{1-l^2-m^2}} \propto \sum_{n=0}^{N_w-1} \exp\left(i2\pi w_n\left(\sqrt{1-l^2-m^2}-1\right)\right)\mathrm{d}w \int\int \mathrm{V}(\mathrm{u},\mathrm{v},\mathrm{w})\exp\left(i2\pi(\mathrm{ul}+\mathrm{vm})\right)dudv$$



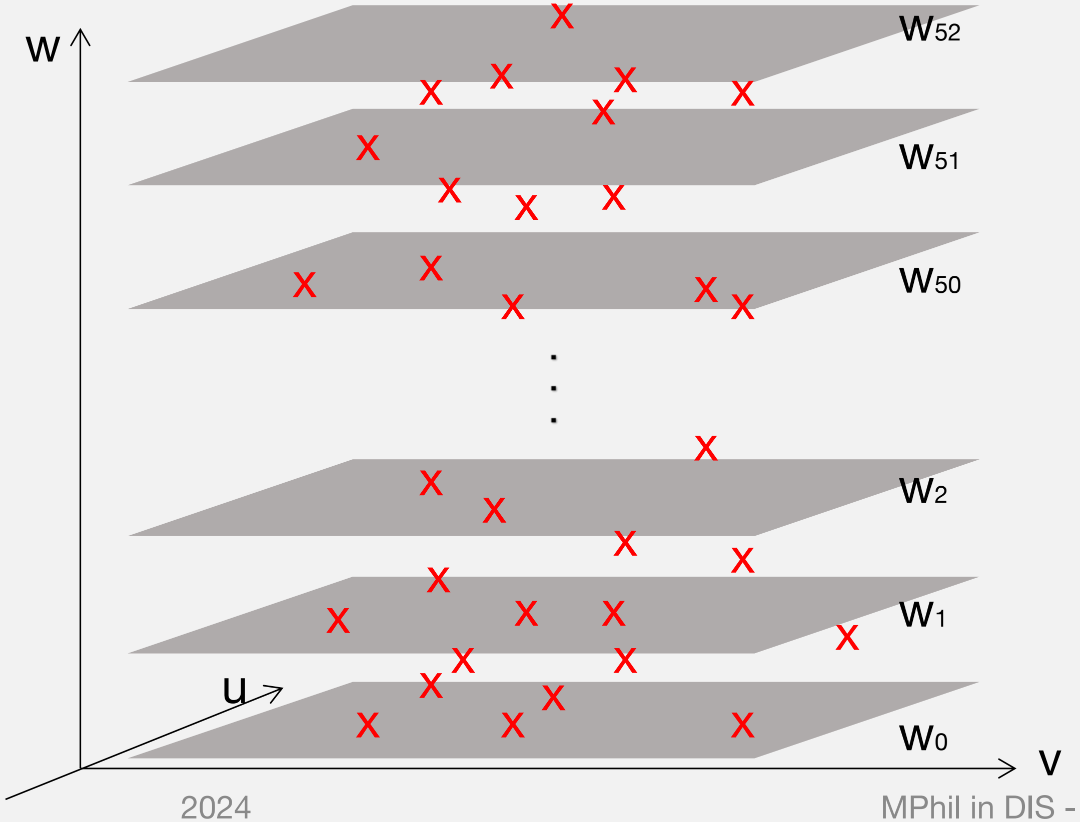
We use limited number of w-planes,

and assign visibilities to its nearest w-plane

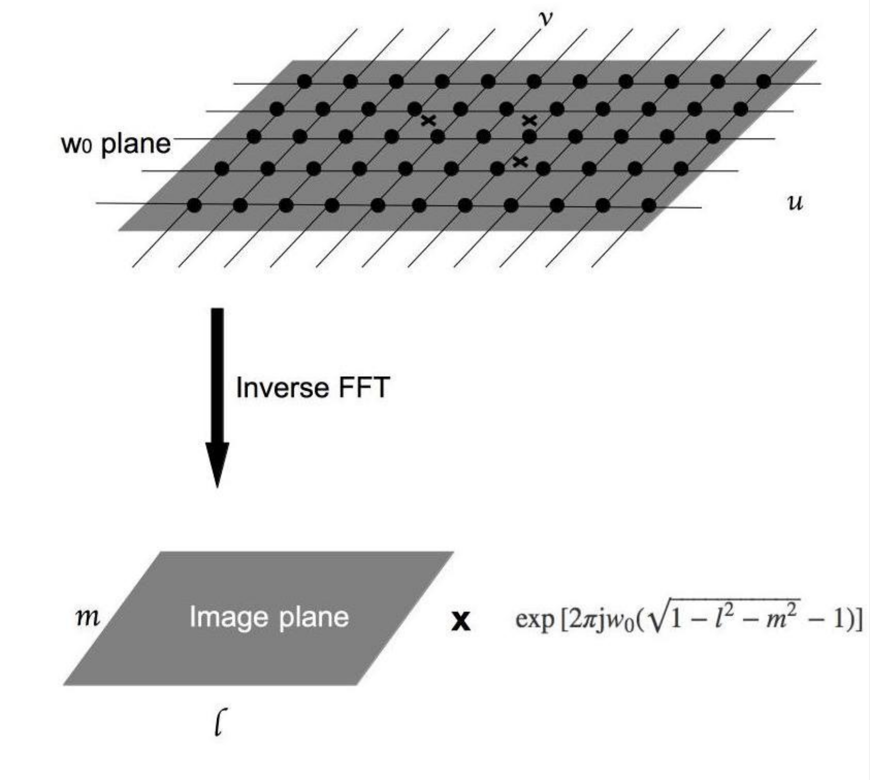
UNIVERSITY OF CAMBRIDGE

W-Stacking method

$$\frac{I(l,m)\left(w_{max}-w_{min}\right)}{\sqrt{1-l^2-m^2}} \propto \sum_{n=0}^{N_w-1} \exp\left(i2\pi w_n\left(\sqrt{1-l^2-m^2}-1\right)\right)\mathrm{d}w \int\int \mathrm{V(u,v,w)}\exp\left(i2\pi(\mathrm{ul}+\mathrm{vm})\right)dudv$$



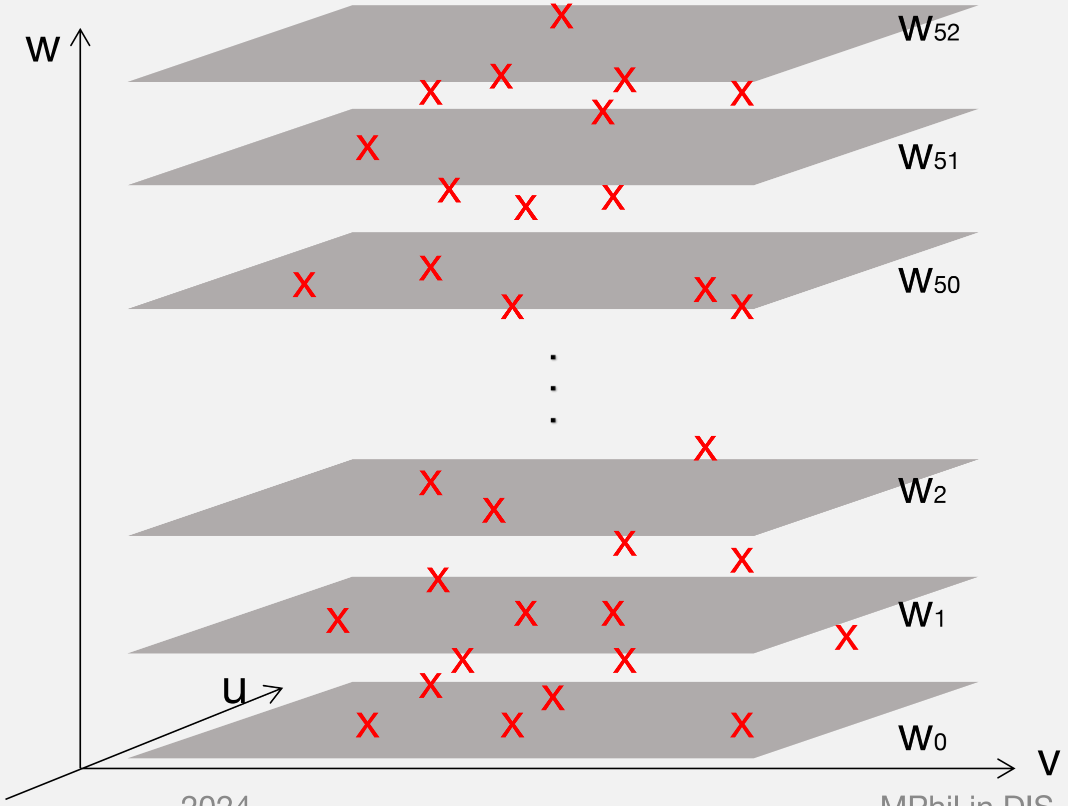
For a certain w-plane



UNIVERSITY OF CAMBRIDGE

W-Stacking method

$$\frac{I(l,m)\left(w_{max}-w_{min}\right)}{\sqrt{1-l^2-m^2}} \propto \sum_{n=0}^{N_w-1} \exp\left(i2\pi w_n\left(\sqrt{1-l^2-m^2}-1\right)\right)\mathrm{d}w \int\int \mathrm{V}(\mathrm{u},\mathrm{v},\mathrm{w})\exp(i2\pi(\mathrm{ul}+\mathrm{vm}))dudv$$



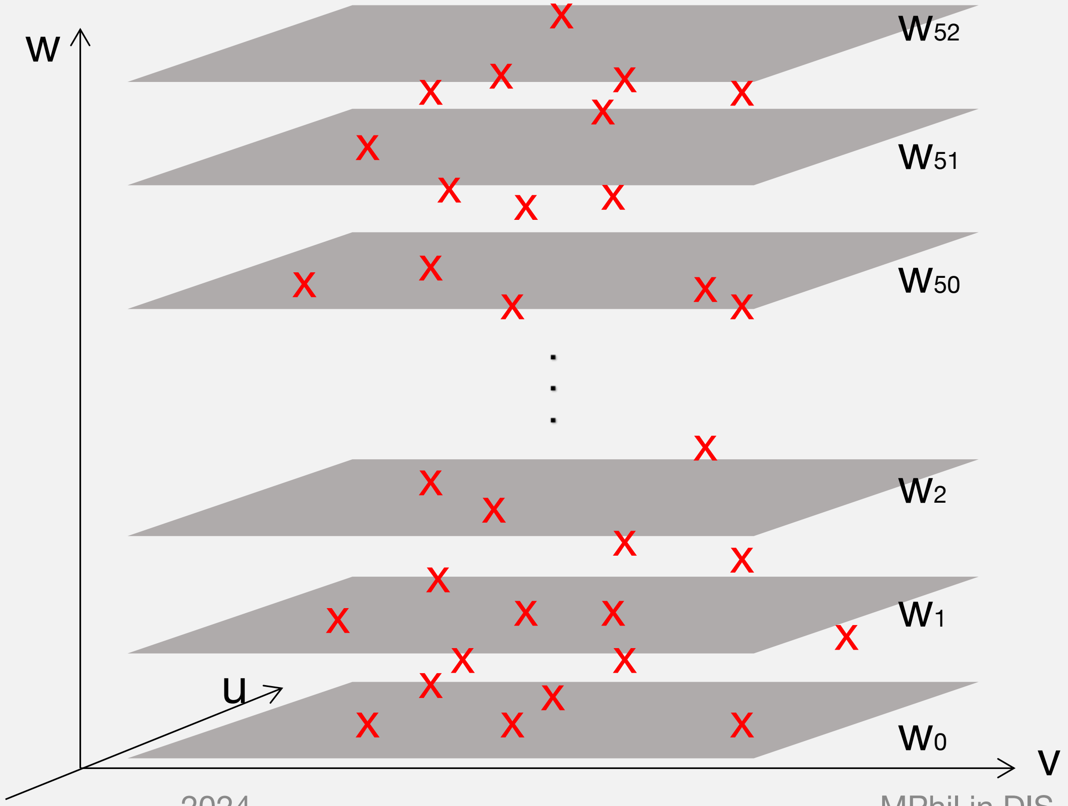
Repeat for all w-planes

$$\sum_{n=0}^{N_w-1} I_{w_n}(l,m)\ \exp\left(i2\pi w_n\left(\sqrt{1-l^2-m^2}-1\right)\right)$$

UNIVERSITY OF CAMBRIDGE

W-Stacking method - improved

$$\frac{I(l,m)\left(w_{max}-w_{min}\right)}{\sqrt{1-l^2-m^2}} \propto \sum_{n=0}^{N_w-1} \exp\left(i2\pi w_n\left(\sqrt{1-l^2-m^2}-1\right)\right)\mathrm{d}w \int\int \mathrm{V(u,v,w)}\exp\left(i2\pi(\mathrm{ul}+\mathrm{vm})\right)dudv$$



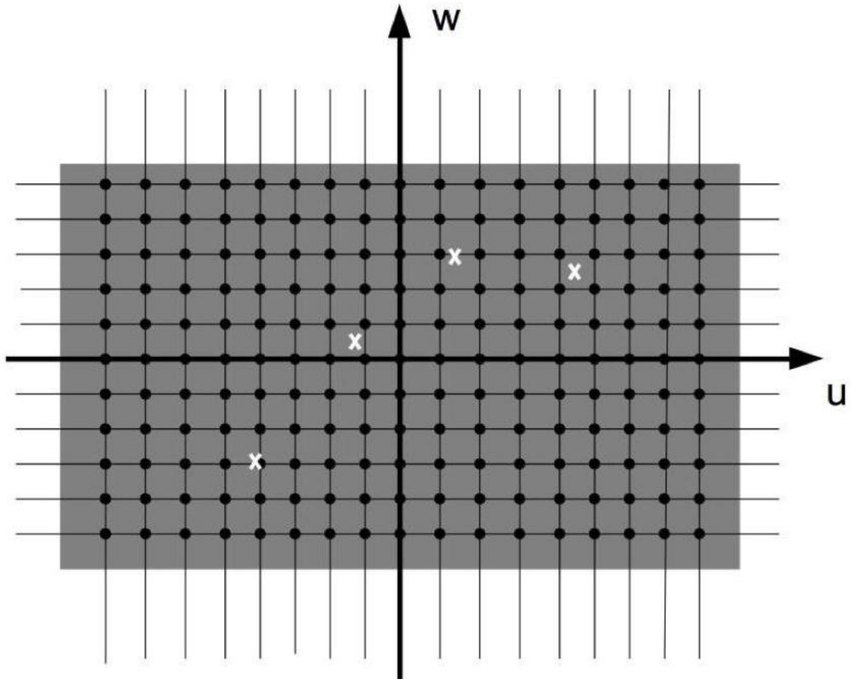
Instead of assiging the visibility value to its nearest w-plane,

We use w-gridding (H. Ye et al. 2022)

UNIVERSITY OF CAMBRIDGE

W-Stacking method - improved

$$\frac{I(l,m)\left(w_{max}-w_{min}\right)}{\sqrt{1-l^2-m^2}} \propto \sum_{n=0}^{N_w-1} \exp\left(i2\pi w_n\left(\sqrt{1-l^2-m^2}-1\right)\right)\mathrm{d}w \int\int \mathrm{V}(\mathrm{u},\mathrm{v},\mathrm{w})\exp(i2\pi(\mathrm{ul}+\mathrm{vm}))dudv$$

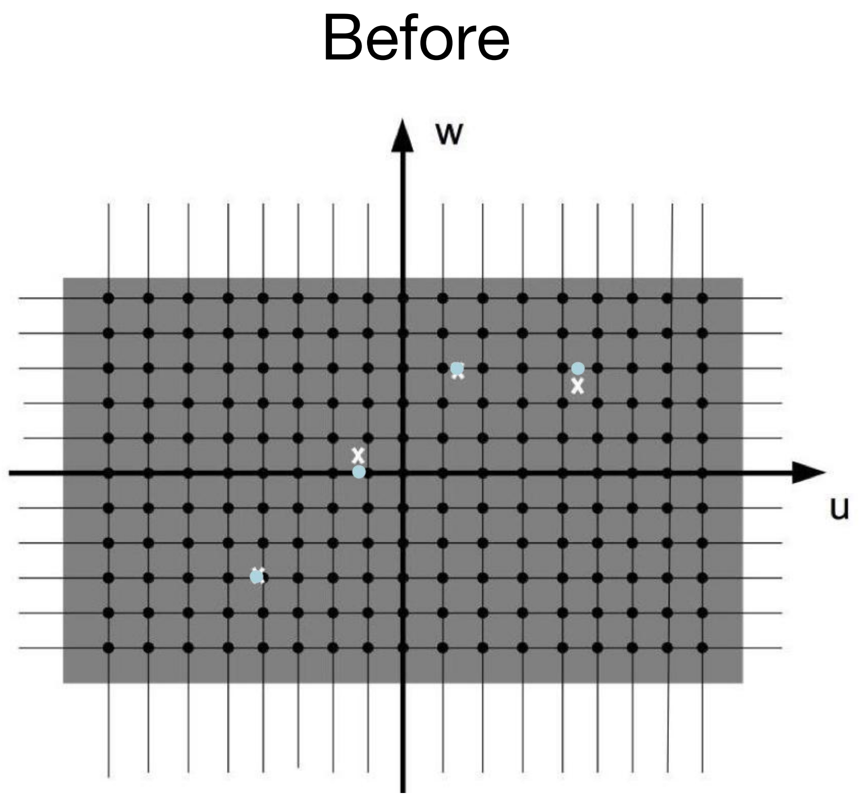


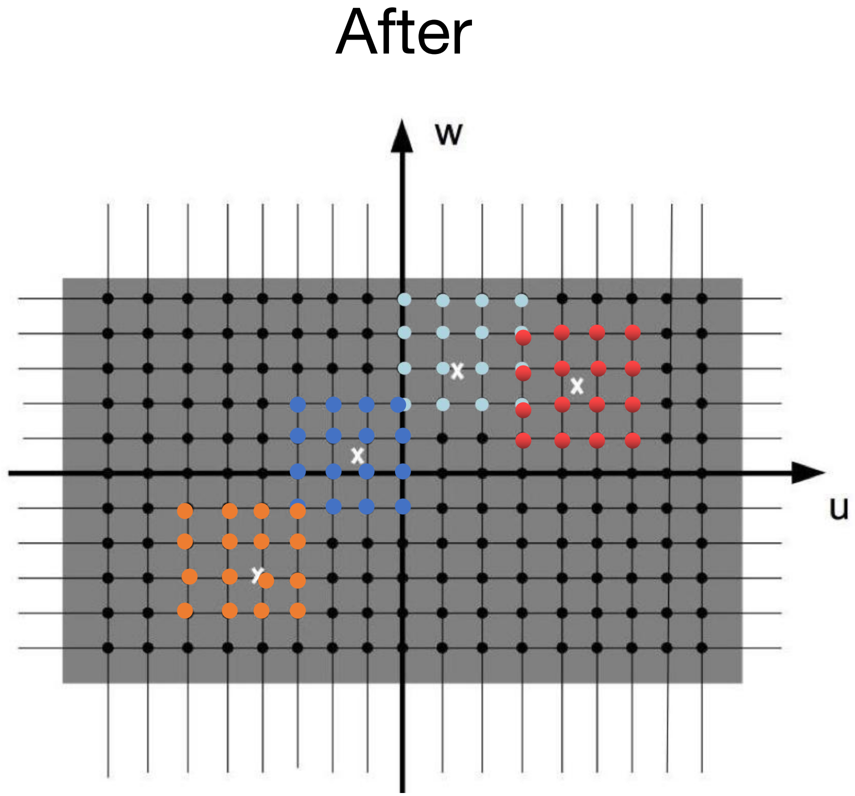
Instead of assiging the visibility value to its nearest w-plane,

We use w-gridding ([H. Ye et al. 2022]())

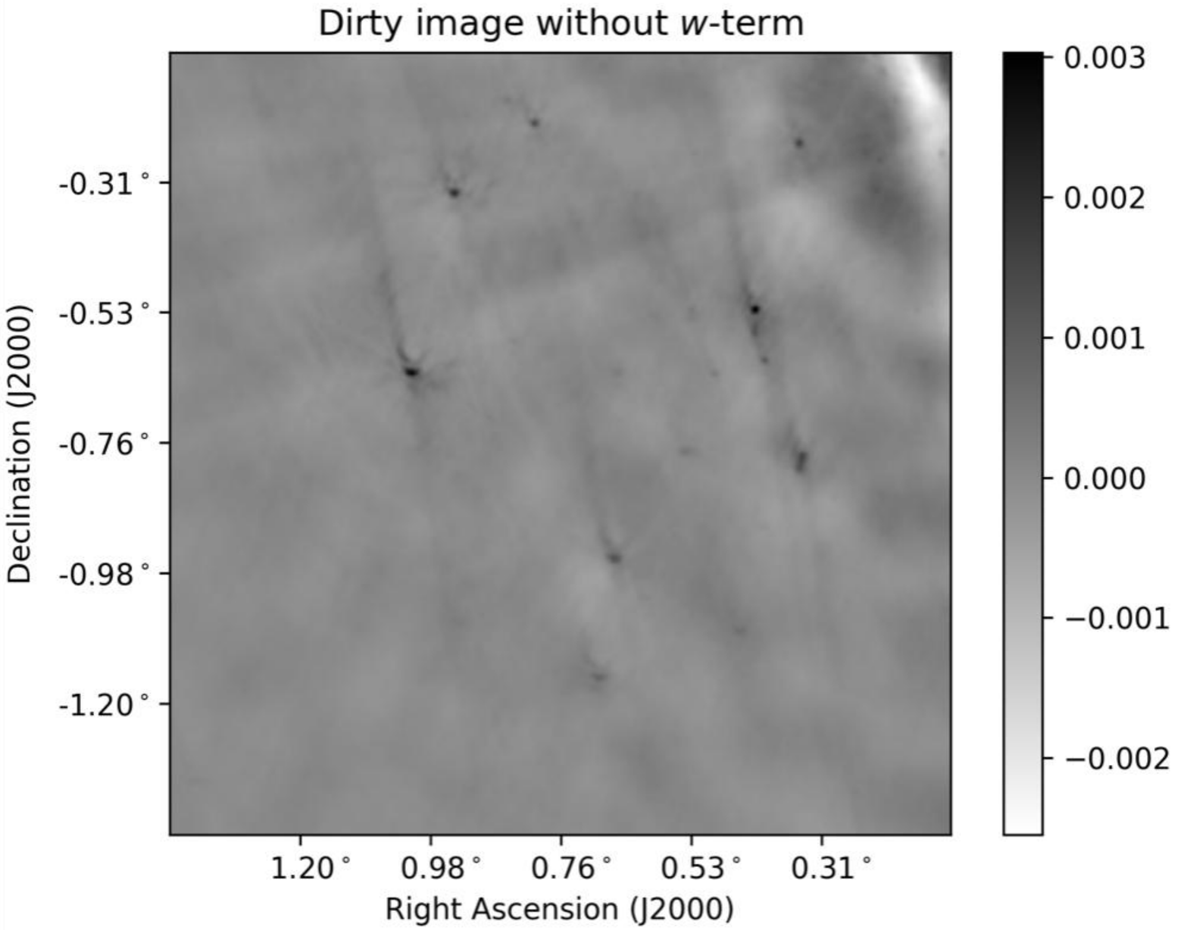
UNIVERSITY OF CAMBRIDGE

W-Stacking method - improved

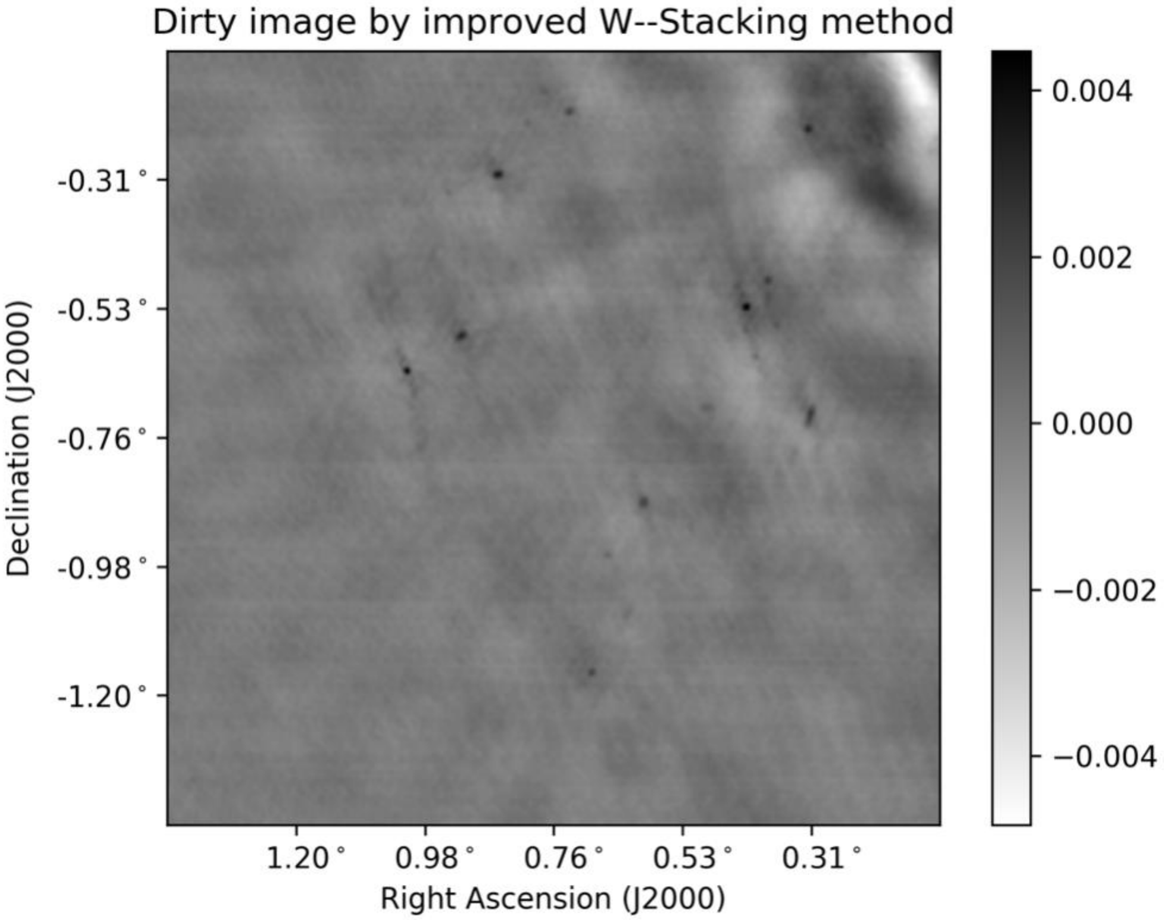




Improved W-Stacking method

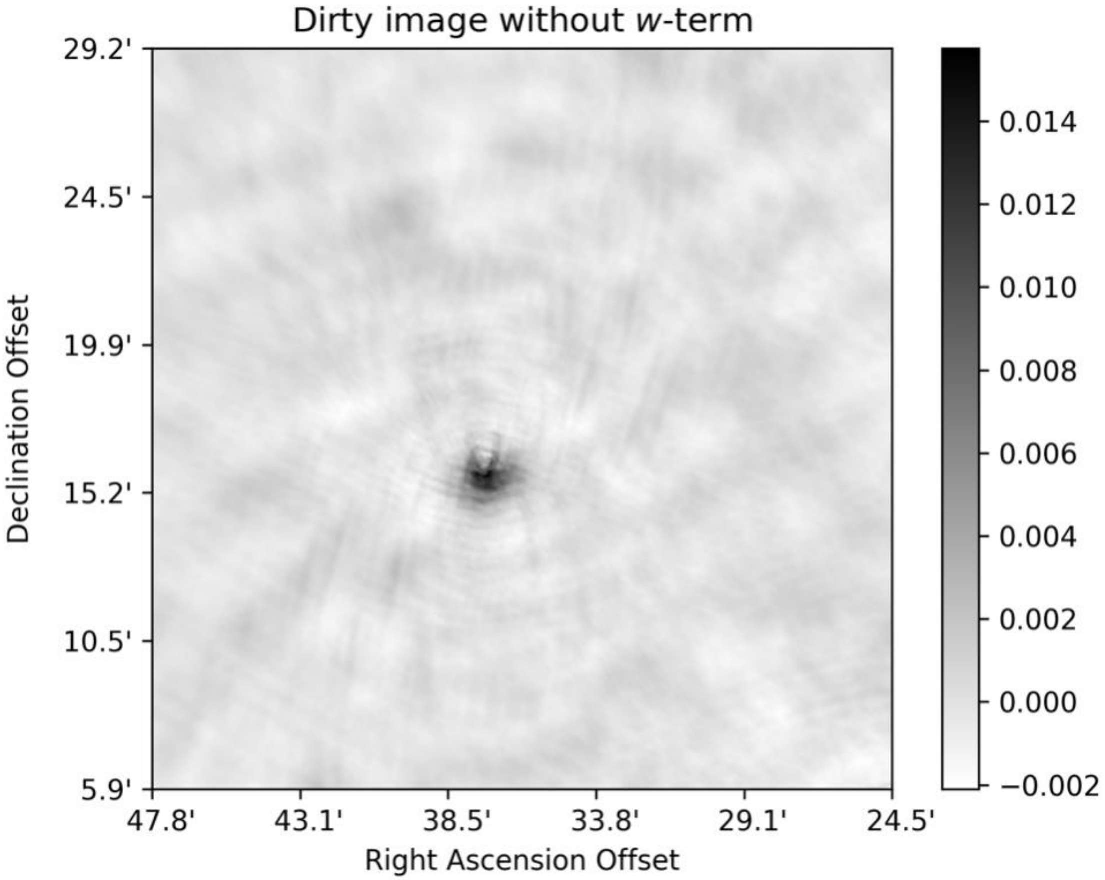


Dirty image without considering w-term

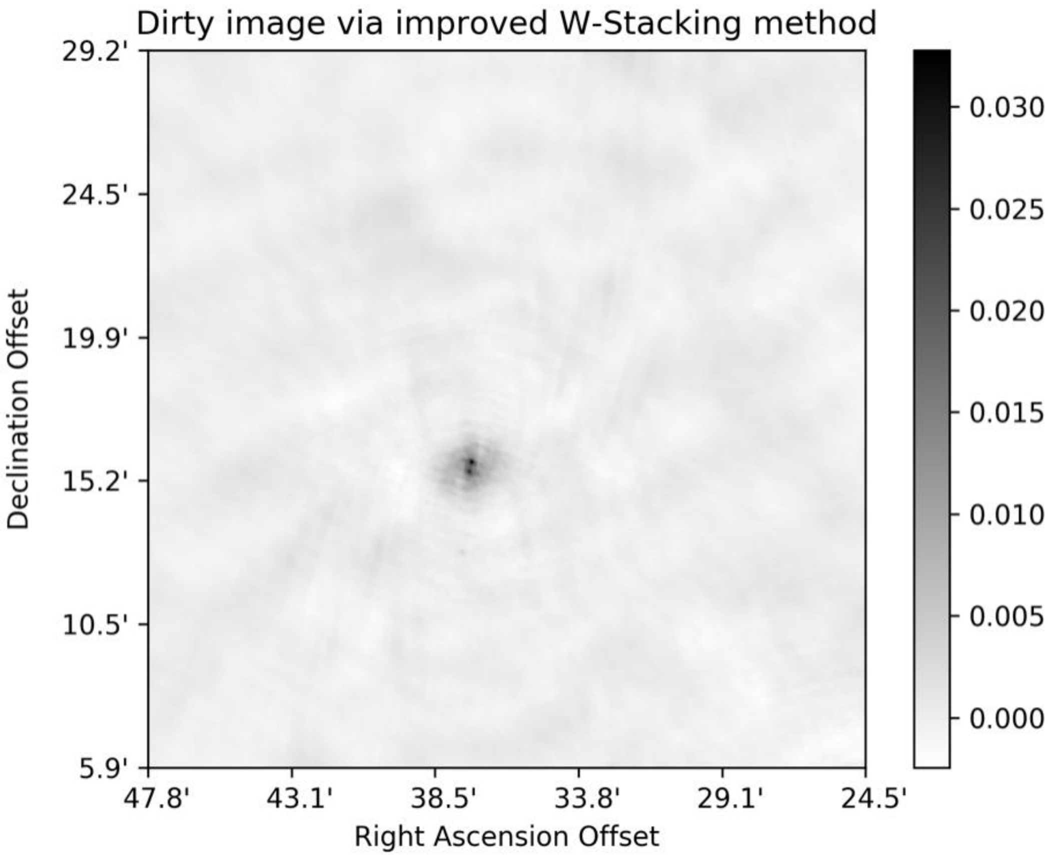


Dirty image using improved W-Stacking method

Improved W-Stacking method



Dirty image without considering w-term



Dirty image using improved W-Stacking method

UNIVERSITY OF CAMBRIDGE

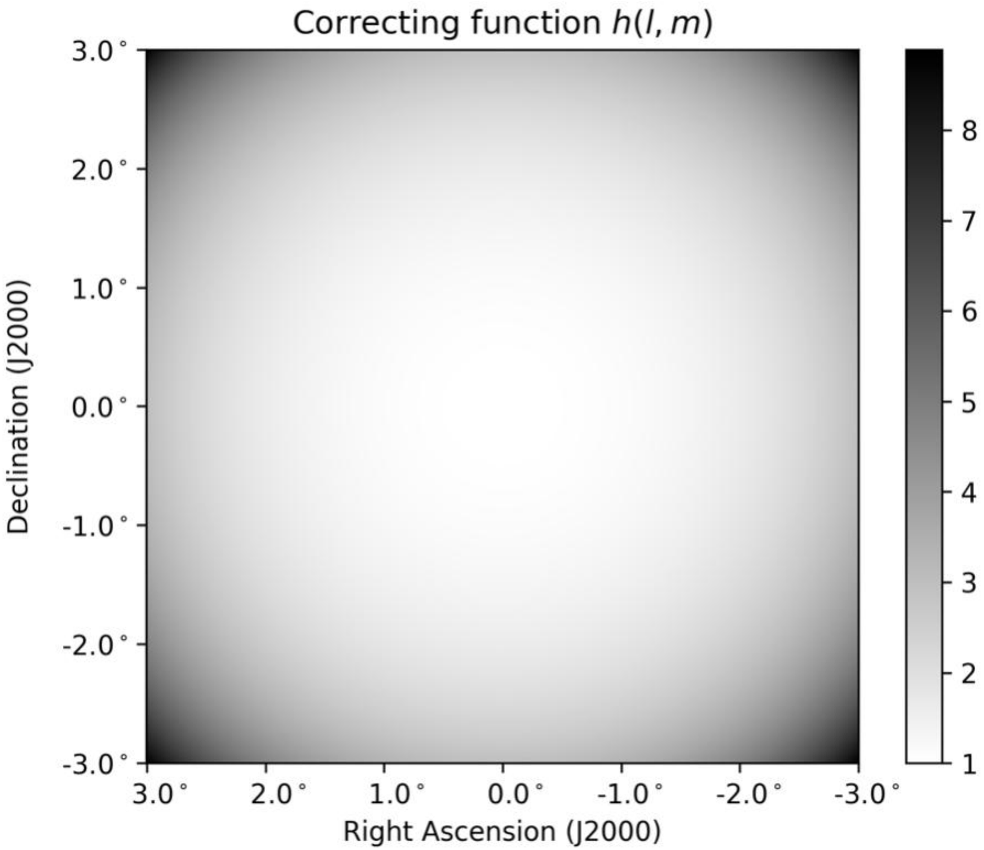
Any questions?

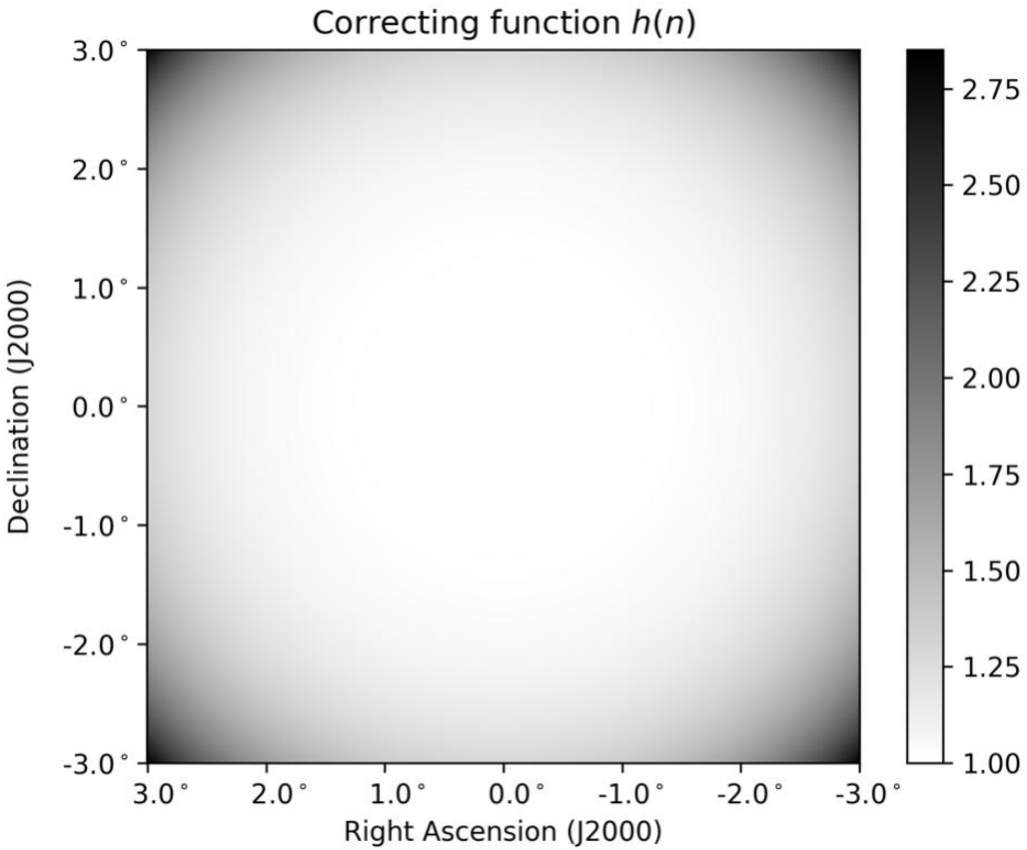
UNIVERSITY OF CAMBRIDGE

Reading

- Accurate image reconstruction in radio interferometry by Haoyang Ye: https://www.repository.cam.ac.uk/items/b4a86e62-6ca4-4f47-9156-ca92de84df3d
- Maximum Entropy and Bayesian Methods Edited by John Skilling, Springer0Science+Business Media, B.V. https://gwern.net/doc/statistics/bayes/1988-jaynes-maximumentropyandbayesianmethods.pdf
- Imaging and Deconvolution by David J. Wilner: https://science.nrao.edu/science/meetings/2018/16th-synthesis-imaging-workshop/talks/Wilner_Imaging.pdf
- Wide Band and Wide Field Imaging by Urvashi Rau: https://science.nrao.edu/science/meetings/2018/16th-synthesis-imaging-workshop/talks/Rao_Wide_1.pdf/view

Improved W-Stacking method





Correcting function h(l,m)

Correcting function h(n)

UNIVERSITY OF CAMBRIDGE

$$N_{w'} \geq \frac{\max_{l,m}(1-\sqrt{1-l^2-m^2})(w_{\max}-w_{\min})}{x_0} + W$$

x0 is usually taken as 0.25
+W: the visibility near the top or bottom w-plane can be gridded

$$N_w \geq 2\pi(w_{\max}-w_{\min})\max_{l,m}(1-\sqrt{1-l^2-m^2})$$