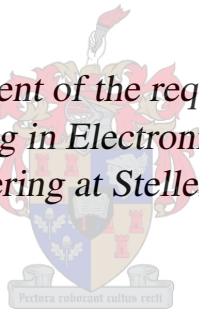
Image Reconstruction in Radio Astronomy with Non-Coplanar Synthesis Arrays

by

Lee Goodrick



Thesis presented in fulfillment of the requirements for the degree of Master of Engineering in Electronic Engineering in the Faculty of Engineering at Stellenbosch University

Supervisor: Prof. David Bruce Davidson
Co-supervisor: Dr. Andre Young

Department Electrical and Electronic Engineering
University of Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.

March 2015

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own original work and that I am the owner of the copyright thereof (unless to the extent explicitly otherwise stated). Furthermore I declare that I have not previously, in its entirety or in part, submitted it for obtaining any qualification.

Lee Goodrick
March 2015

Abstract

Traditional radio astronomy imaging techniques assume that the interferometric array is coplanar, with a small field of view, and that the two-dimensional Fourier relationship between brightness and visibility remains valid, allowing the Fast Fourier Transform to be used. In practice, to acquire more accurate data, the non-coplanar baseline effects need to be incorporated, as small height variations in the array plane introduces the w spatial frequency component. This component adds an additional phase shift to the incoming signals. There are two approaches to account for the non-coplanar baseline effects: either the full three-dimensional brightness and visibility model can be used to reconstruct an image, or the non-coplanar effects can be removed, reducing the three dimensional relationship to that of the two-dimensional one.

This thesis describes and implements the w-projection and w-stacking algorithms. The aim of these algorithms is to account for the phase error introduced by non-coplanar synthesis arrays configurations, making the recovered visibilities more true to the actual brightness distribution model. This is done by reducing the 3D visibilities to a 2D visibility model. The algorithms also have the added benefit of wide-field imaging, although w-stacking supports a wider field of view at the cost of more FFT bin support. For w-projection, the w-term is accounted for in the visibility domain by convolving it out of the problem with a convolution kernel, allowing the use of the two-dimensional Fast Fourier Transform. Similarly, the w-Stacking algorithm applies a phase correction in the image domain to image layers to produce an intensity model that accounts for the non-coplanar baseline effects.

This project considers the KAT7 array for simulation and analysis of the limitations and advantages of both the algorithms. Additionally, a variant of the Högbom CLEAN algorithm was used which employs contour trimming for extended source emission flagging. The CLEAN algorithm is an iterative two-dimensional deconvolution method that can further improve image fidelity by removing the effects of the point spread function which can obscure source data.

Opsomming

Tradisionele beeldvormingstegnieke in radio-astronomie aanvaar dat die interferometriese skikking samevlakkig is. Dit beteken dat die twee-dimensionele Fourier verhouding tussen helderheid en sigbaarheid geldig bly en dat die Vinnige Fourier Transform aangewend kan word. Klein hoogtevariasies in die skikkingsvlak bring die w-ruimtelike frekwensiekomponent mee, wat 'n faseverskuiwing in die inkomende seine tot gevolg het. Dus, in praktyk, moet die bydrae van die nie-samevlakkige basislyneffekte in ag geneem word om sodoende die akkuraatheid van die data te verhoog. Twee benaderings kan gevolg word om die nie-samevlakkige basislyneffekte in ag te neem: Metodes wat die volle drie dimensionele helderheid en sigbaarheidsmodel gebruik kan toegepas word om 'n beeld te herbou, andersins kan die nie-samevlakkige effekte verwyder word om sodoende die drie-dimensionele verhouding te verminder tot 'n twee-dimensionele verhouding.

Hierdie tesis beskryf en implementeer die 'w-projeksie' en 'w-stapel' algoritmes. Die doel van hierdie algoritmes is om die fasefout wat deur nie-samevlakkige sinteseskikkingskonfigurasies veroorsaak word, reg te stel. Hierdie regstelling maak die herwinde sigbaarheid van die beeld meer getrou aan die werklike helderheidsverspreidingsmodel. 'n Bykomende voordeel van die algoritmes is beeldvorming van wye-veld ruimtewaarnemings. In 'w-projection' word die w-term in die sigbaarheidsdomein in ag geneem deur die ruimtelike frekwensiekomponent met behulp van 'n konvolusiekern vanuit die probleem te verwyder. Die twee-dimensionele Vinnige Fourier Transform kan gevolglik toegepas word. Soortgelyk hieraan, wend die 'w-Stacking' algoritme 'n fasekorreksie aan tot 'n reeks beeldlae, om sodoende 'n beeld te verkry wat die nie-samevlakkige basislyneffekte in ag neem.

Die KAT7 teleskoop is gebruik in die simulasie en analiese van die tekortkominge en voordele van beide algoritmes. 'n Hibriede weergawe van die Högbom CLEAN algoritme is bykomend oorweeg. Hierdie algoritme is 'n iteratiewe twee-dimensionele dekonvolusiemetode wat die betroubaarheid van beelde verbeter deur die verskansingseffek van puntverspreidingsfunksies te verwyder. Verder gebruik die Högbom CLEAN algoritme kontoersnoeiing om uitgebreide bron-emisies te identifiseer.

Contents

List of Tables

List of Figures

Nomenclature

Abbreviations and Acronyms

1D	One-Dimensional
2D	Two-Dimensional
3D	Three-Dimensional
ASKAP	Australian Square Kilometer Array Pathfinder
CASA	Common Astronomy Software Applications
CTC	Computational Time Cost
DDE	Direction Dependent Effects
Dec	Declination
FFT	Fast Fourier Transform
FoV	Field of View
IFFT	Inverse Fast Fourier Transform
PSF	Point Spread Function
RA	Right Ascension
SKA	Square Kilometer Array
SNR	Signal-to-Noise Ratio
SUN	Stellenbosch University

Symbols & Greek Letters

θ	Angle
δ_d	Declination Angle
η	System Noise
ψ	Dirty Beam Error
∞	Infinite Value
γ	Loop Gain
ϵ	Absolute Average Energy

ϑ	Threshold Factor
Γ	Gain Ratio
σ	Standard Deviation
μ	Mean
δ	Dirac Delta Function
α	Tukey Tapering Parameter

Lowercase Letters

$\bar{\mathbf{b}}$	Baseline Vector
u,v,w	Spatial Frequency Units
l,m,n	Directional Cosine Units
p	Pixel Count
x	Iteration Count
t	Trim Parameter

Uppercase Letters

$\mathbf{B}_r$	Brightness
D	Effective Diameter
N_a	Number of Antennas
N_b	Number of Baselines
X,Y,Z	Cartesian Units
H	Right Ascension / Hour Angle
$\mathbf{A}$	Antenna Direction Dependent Effects Matrix
$\mathbf{S}$	Sampling/Transfer Function
$\mathbf{V}$	Visibility Function
$\mathbf{I}$	Intensity / Brightness Distribution
$\mathbf{B}$	Dirty Beam / Point Spread Function
$\hat{\mathbf{B}}$	Translated Dirty Beam
$\mathbf{D}$	Dirty Map
$\mathbf{R}$	Residual Map
$\mathbf{T}_r$	Trim Contour Value
$\mathbf{G}$	Image Space Projection Kernel

$\bar{\mathbf{G}}$	Convolution Kernel
$\hat{\mathbf{G}}$	Approximated Convolution Kernel
$\mathbf{G}_{full}$	Alternative Image Space Projection Kernel
$\mathbf{I}_p$	Maximum Intensity Energy Peak
$\mathbf{I}_d$	Identity Matrix
$\mathbf{D}_n$	Dirac Identity
T_k	*uv* Tapering Parameter
Q_k	*uv* Density Weighting Parameter
$\mathbf{W}$	Weighting Function
$\mathbf{T}$	Image Tapering Function
$\mathbf{V}_p$	Recovered *w*-Projection Visibility Function
$\mathbf{V}_s$	Recovered *w*-Stacking Visibility Function

Syntax and Style

$\mathbf{x}$	The vector $\mathbf{x}$
$\mathbf{A}$	The matrix $\mathbf{A}$
$*$	Convolution Operator
$\xrightarrow{\mathcal{F}^{-1}}$	Inverse Fourier Transform
$\xrightarrow{\mathcal{F}}$	Forward Fourier Transform
$\mathcal{F}\{f\}$	The Fourier Transform of the function f
$\mathscr{R}\{\mathrm{f}\}$	The real part of the function f
$\lvert\mathbf{x}\rvert$	The absolute value of the vector $\mathbf{x}$
$\mathbf{x} \bullet \mathbf{y}$	The dot-product of the vectors $\mathbf{x}$ and $\mathbf{y}$
$\mathbf{A} \bullet \mathbf{B}$	Element-wise multiplication of the matrices $\mathbf{A}$ and $\mathbf{B}$
$\mathrm{d}x$	Differential element x
$\frac{\mathrm{d}f}{\mathrm{d}x}$	Derivative of function, f, with respect to x
$\int$	Integral Operator
$\sum$	Summation Operator

Acknowledgments

Throughout this project there have been various people who have offered support. While some assisted more than others, I would like to keep this short but give thanks to a few key people who have aided me in this project completion.

Firstly, to my project coordinator, Professor David Davidson - thank you for taking me on as your student. You allowed me to branch into knowledge of Astronomy that has always been of interest to me. I would probably have never taken the spare time to read into it if you did not accept me for both my MEng and BEng final year project, back in 2011. Your input has been invaluable and I really hope you realise how much I appreciate the opportunities you have afforded me. If, in the near future I choose to further my academic career, I would welcome the idea of being under your tutelage again.

Secondly, to my co-supervisor, Dr. Andre Young. Your insight, too, has directly lead me into producing a project that I am truly proud of. My initial goals were extended far more with your help and willingness to take time into understanding issues that were puzzling me, be it small or big. After feeling totally stumped at times, there was always a sense of relief when returning home after our meetings. Thank you!

To those outside of the University: Dr. Nadeem Oozeer, thank you for assisting me with information and showing interest in my work. The course you hosted really allowed me to fill in the gaps in my knowledge that allowed me to grasp familiar concepts with much more ease. Also, a sincere thanks to Dr. Oleg Smirnov and more specifically to Dr. Andre Offringa for readily meeting with me and assisting me in understanding concepts of w-projection and w-stacking. While possibly trivial to you, your help impacted my project direction to beneficial effect.

To the SKA, along with Stellenbosch University, for investing funding into us students - I would not have been able to accomplish my goal of obtaining a MEng degree without these two institutes.

Finally, to my two support structures and proof readers, my mom and sister, who worked through this project at ungodly hours before deadline - I know it must have been a very tedious exercise for you both.

Chapter 1

Introduction

1.1 A Background into Radio Astronomy and Imaging

Radio astronomy is an important means of providing scientific insight into the mysteries of our universe. It uses the radio electromagnetic spectrum and wave propagation theory to analyse and quantify signals from distance space sources. For the most part, modern day radio astronomy often trumps optical-based astronomy as it is not limited to the visual spectrum, allowing much more detail to be gathered about sources as the input data sets are greater [1]. Furthermore, the world has moved into a digital age where analogue systems are being replaced by digital ones, where technological advances extend the capabilities of data acquisition, processing, storage and analysis.

With large scale projects such as the Square Kilometer Array (SKA), the Australian Square Kilometer Array Pathfinder (ASKAP) and augmenting Very Long Baseline Interferometer (VLBI) systems by extending networks both globally and in space, the radio astronomy industry is growing in both the scientific and engineering fields. The technological and theoretical requirements from both these fields pushes scientists, engineers and astronomers to develop state-of-the-art software, hardware and instruments to meet the demands of these projects.

There are many goals associated with each field of radio astronomy, one of them being to image the sky by analysing the radio signals for accurate modeling of sources and space phenomena. Accurate modeling in imaging entails obtaining precise data of source positions and their size and magnitude attributes. With the advent of the aforementioned technological developments and projects, modern day imaging capabilities have excelled in terms of scale and detail which, in turn, has called for more efficient, refined and precise imaging schemes - giving motivation for this project.

1.2 Motivation

Past imaging techniques worked on the assumption that interferometric networks were coplanar and the fractions of the sky being imaged allowed sources to be modeled on a flat plane [2]. These assumptions neglect the w-component phase shifts in the radio signals which, as imaging becomes more demanding in acquiring precise models, can cause more data inaccuracies that effect source size and attributes. Non-coplanar baselines and large fields of view (FoV) are

very real considerations that need to be factored into the picture and various imaging schemes have emerged to incorporate these - however the imaging schemes each have their advantages and disadvantages in terms of efficiency and accuracy.

Two methods, namely *w*-projection [3] and *w*-stacking [4], both take wide-field imaging and non-coplanar baselines into account to obtain more accurate images. They differ in approaches as the one eliminates the additional phase shift in the Fourier space while the other corrects the additional phase shift in the image domain. After the non-coplanar effects are removed, the resulting images are still affected by incomplete sampling in the *uv*-plane. In practice, the CLEAN algorithm [5] can be used to improve the images, which is an algorithm used to deconvolve the effect of the Point Spread Function (PSF) and remove some noise elements. The CLEAN algorithm is used and the combined result of CLEAN with *w*-projection is shown to yield better models of visibility and higher fidelity images.

1.3 Project Aims

The primary goal of the project is to show how non-coplanar baseline effects and imaging large portions of the sky can warp the image data, resulting in an image that is not true to the actual sky model. Then, to implement imaging techniques that account for these effects which result in higher fidelity images and show how their use in practice is extremely beneficial. The secondary goal is to illustrate a full understanding of the entire imaging process through simulation. This involves knowing how the various stages of data are used in the grand scheme of imaging and where the detrimental effects such as aliasing and noise need to be considered to minimise their effect on the data.

1.4 Project Structure

This thesis is structured in the following manner:

I. Chapter 1 includes a project overview and the motivation behind this thesis.

II. Chapter 2 is a detailed overview which is divided into two parts. Part one outlines the fundamentals of radio astronomy that play a role in understanding this project. Part two highlights the importance of taking the FoV and non-coplanar baselines into account, and discusses the detrimental implications on imaging if these factors are ignored.

III. Chapter 3 discusses the first non-coplanar baseline imaging technique, *w*-projection, which was primarily developed by Cornwell et al [3]. This technique projects non-coplanar baseline effects out of the problem in the visibility domain so that standard two-dimensional imaging methods can be used.

IV. Chapter 4 discusses the second technique, *w*-stacking, which is a more recently developed imaging scheme from various contributors in the Astronomy industry. This method splits the visibility data into various layers and corrects the phase shift in the image domain; all of which are then overlayed to produce one final phase-corrected image.

V. Chapter 5 provides a detailed description of the CLEAN algorithm, a two-dimensional iterative deconvolution technique, and its application in imaging. The original Högbom

model was altered to include contour trimming to allow flagging of more complex source structures, not limiting the application of CLEAN to only point sources.

VI. Chapter 6 provides a comparative analysis of the two methods, *w*-projection and *w*-stacking. The methods are implemented on a simulated observed sky model and the resulting image from *w*-projection is then CLEANed. The sky model incorporates a restricted wide FoV and non-coplanar baseline effects along with typical negative effects that may be experienced in practice, such as aliasing and noise. Additionally, a simulated KAT7 array configuration, located in South Africa, is used for the imaging process to show the steps used in imaging such as gridding, weighting and tapering.

VII. Chapter 7 concludes the findings of this thesis and proposes future work recommendations.

Chapter 2

The Fundamentals of Radio Astronomy and Imaging

This chapter is split into two parts. Section 2.1 includes a breakdown of interferometry and covers the fundamental concepts of radio astronomy in both the image and visibility domains. Section 2.2 extends these fundamentals for the case of non-coplanar arrays, but more specifically in relation to imaging.

2.1 Fundamental Concepts

2.1.1 Coordinate Systems

To define the fundamentals of radio astronomy, the Earth-based and the celestial sphere based coordinate systems must be understood. The celestial sphere is a spherical geometric projection that envelopes Earth, with the celestial poles and equator aligning with those of the Earth. The sources in the universe can be seen as projections onto this sphere whose orientations are defined by the zenith and azimuth angles when using the Horizontal Coordinate System. In the system the zenith points directly upwards from the observer and is normal to the observer's horizontal plane, while azimuth is measured in a clockwise direction from North, increasing towards the East, along the observer's horizontal plane North [2].

The zenith and azimuth angles are based on a local coordinate system and are dependent on the frame of reference of the observer. Another alternative of defining the position of the sources on the celestial sphere is the Equatorial Coordinate System using Declination and Right Ascension, which is more commonly used in astronomy and this thesis. Declination and Right Ascension are based on a global coordinate system; essentially they could be considered measurements of celestial longitude and latitude [2]. Declination is a measure of the angular distance from the celestial equator, corresponding to celestial latitude, whereas Right Ascension is measured from the vernal equinox, corresponding to celestial longitude. Right Ascension is generally defined in hours, which directly relate to degrees; one full Earth rotation would mean 24 hours have passed, or alternatively, the Earth has rotated by 360°. A source position would change in zenith and azimuth degrees as the Earth rotates and the observer changes their frame of reference, but would generally remain fixed when defined in Declination and Right Ascension. Both coordinate systems are depicted in Figure 2.1.

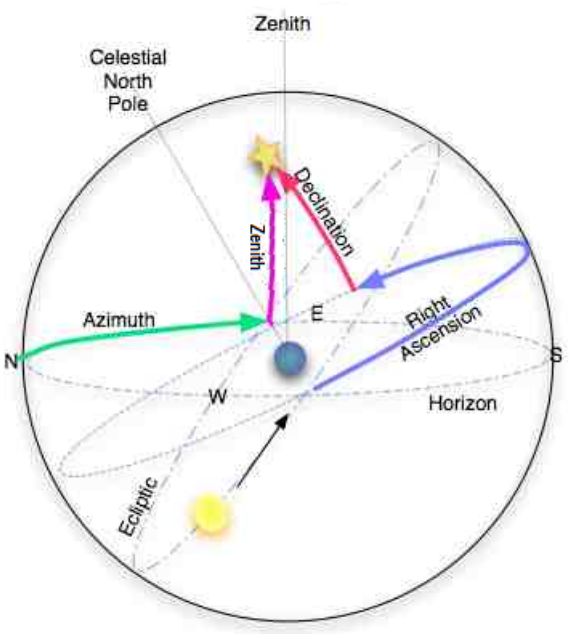


Figure 2.1: The celestial sphere and the corresponding coordinate systems with zenith and azimuth as well as Declination and Right Ascension. Graphics were taken from [6].

2.1.2 Flux and Source Brightness

Radio astronomy uses wave propagation and electromagnetic radiation theory to describe the data from sources in the sky. Sources emit signals in the form of radiation waves which travel radially outward and can be detected by antennas on Earth. The radiation is quantified by flux density which is measured in watts per unit area per hertz, and is given the unit of Jansky [1] where

$$1\,\text{Jy} = 10^{-26}\,\text{Wm}^{-2}\text{Hz}^{-1}. \tag{2.1.1}$$

The Jansky is a measure of the power received over a given aperture size, meaning it is a detector-based quantity. For imaging purposes, it is more convenient to refer to Jy per Beam, having units of $\text{Wm}^{-2}\text{Hz}^{-1}\text{sr}^{-1}$. This gives an indication as to how much power is being received by a receiving antenna beam and is referred to as brightness. Intensity is interchangeable with brightness but describes the flux as either passing from or into the solid angle [2] so that

$$\mathbf{I} = -\mathbf{B}_r, \tag{2.1.2}$$

where $\mathbf{I}$ is the intensity and $\mathbf{B}_r$ is brightness. The apparent brightness depends on distance and can be influenced by other sources of radiation and the interstellar medium; this differs from the intrinsic brightness which is the brightness of the source that is independent of distance and unaffected by interstellar medium. The total power passing through the celestial sphere from the combined effect of various sources is the apparent brightness distribution or intensity map; it is a model of how the flux from various sources is spread across the sky. This apparent brightness distribution or true sky brightness forms the basis for acquiring visibility data which is discussed in the succeeding sections.

2.1.3 Interferometry

For a single dish, the resolution of the image is limited by the size of the dish diameter and lacks the capability to resolve small sources in the sky. The angular resolution, θ_r, is a measure of the minimum angular distance between two objects that will still allow them to be seen as separate entities and is given by

$$\theta_r = \frac{\lambda}{D}, \tag{2.1.3}$$

where λ is the observed wavelength, and D is the effective diameter of the dish.

Interferometric methods were developed in order to increase the effective diameter of the system, allowing smaller sources to be resolved with greater detail and model the true brightness distribution more accurately. For an interferometric system, each antenna is able to measure an incoming signal, whose data is then correlated with other antenna measurements [2]. The combined effect of the separate antennas provide improved signal correlation as the effective diameter becomes the maximum baseline distance, thus improving the angular resolution. Baselines are defined as the vectors between various antennas in the array network.

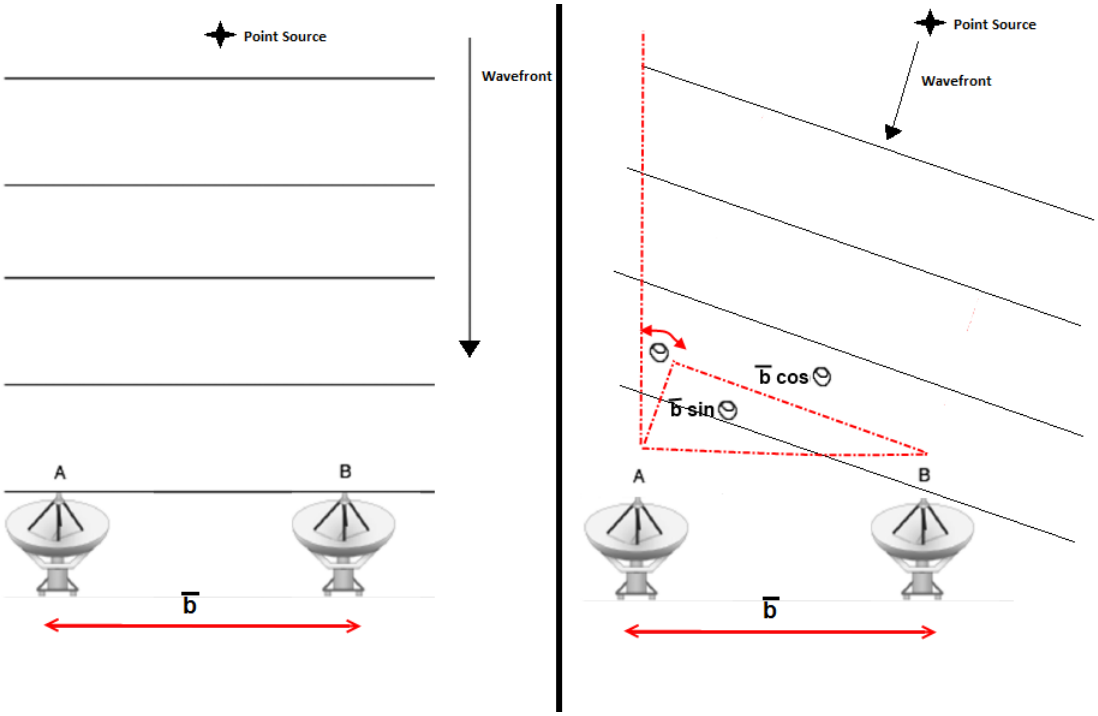


Figure 2.2: Left - A pair of antennas observing in-phase wavefronts from a broadside source. Right - The same configuration with an off-center source whose detected wavefront is now out of phase. Graphics were adapted from [2].

Consider a pair of antennas separated by a baseline distance, $\bar{\mathbf{b}}$, and observing a source whose radiation plane waves are incident to both antennas, as shown on the left in Figure 2.2. These antennas will observe the signal as arriving in phase and the correlation of the two signals would be at a maximum. If the source were slightly off-center at an angle of θ as seen on the

right, the propagated wave arrives at one antenna before the other and the signals are out of phase, decreasing coherency. For increasing θ, where the distance between the path length of the antennas approaches the observed wavelength value, the signals will once again become fully coherent. The increasing and decreasing correlator output produces what is called fringes [1]. These fringes are further enveloped by the individual antenna pattern to produce the fringe pattern, or visibility function [1], illustrated in Figure 2.3.

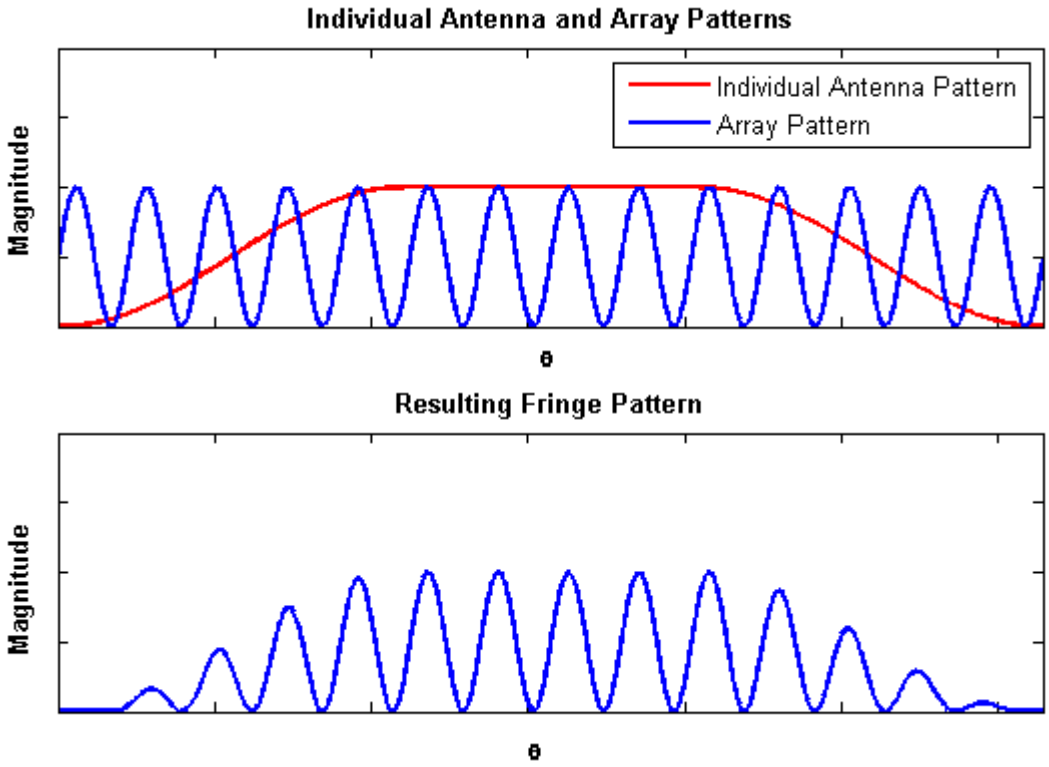


Figure 2.3: Top - The individual antenna pattern, dependent on the antenna design along with the correlation function at the output of a correlator, showing the fringes as the source signal shifts in and out of phase. Bottom - The fringe pattern, or visibility function of a source. The antenna pattern envelopes the fringes. Graphics were adapted from [1].

For extended observations over a longer time period, the rotation of the Earth needs to be taken into account. Furthermore for higher frequency observations, the receiving antenna beam shape becomes increasingly narrower, meaning the source brightness is in the primary beam for a shortened period of time. For these cases, it becomes necessary to track the sources as they pass across the sky. Phase tracking can be accomplished by introducing an instrumental delay into the system which will compensate for the geometric delay of the antenna array [7]. The phase tracking center or phase reference position is a relative reference point for tracking the angles used to specify the source intensity distribution.

Additionally, to improve the quality of the measured brightness distribution, more antennas can be used to increase the effective area so more power can be received. This will increase the number of baselines. The number of baselines increases as

$$N_b = \frac{N_a(N_a - 1)}{2}, \tag{2.1.4}$$

where N_b is the number of projected baselines and N_a is the number of antennas in the array. Increasing the number of baselines effectively improves sampling of the visibility function as each baseline measures a component of the true brightness distribution [1]. The measured brightness will never exactly represent the true brightness distribution as noise and aliasing is introduced when sampling it.

2.1.4 Spatial Frequency and Direction Cosine Components

With the coordinate system and interferometry for the one-dimensional case explained, the theory can be extended to two dimensions to include both East-West and North-South components, allowing sources to be resolved on a plane. It becomes useful to define two sets of planes coordinates, namely the *uv*-plane and the *lm*-plane. The *uv*-plane, also referred to as the Fourier Space [7], contains the spatial frequency components u and v which are in the East-West and North-South directions respectively. The *uv*-points define baseline vectors which describe the distance and orientation between antenna pairs.

Antenna positions are specified in terms of a Cartesian system where X and Y form a plane that is parallel to that of the equator and Z points to the North Pole. To obtain the baseline vector points, or *uv*-data, the transformation matrix below can be used as given by [2]:

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} \sin(H) & \cos(H) & 0 \\ -\sin(\delta_d)\cos(H) & \sin(\delta_d)\sin(H) & \cos(\delta_d) \\ \cos(\delta_d)\cos(H) & -\cos(\delta_d)\sin(H) & \sin(\delta_d) \end{bmatrix} \begin{bmatrix} \mathrm{X}_\lambda \\ \mathrm{Y}_\lambda \\ \mathrm{Z}_\lambda \end{bmatrix}. \tag{2.1.5}$$

Ignore the w-component for the present as this will be discussed at length for non-coplanar arrays. The spatial frequency components are dependent on the observation parameters, such as the frequency, the Right Ascension and the Declination of the source. In (2.1.5), H is the hour angle of the observation, δ_d is the Declination and the frequency has been included in the X_λ,Y_λ and Z_λ parameters as they are expressed in wavelengths.

The *uv*-data points sample the spatial frequencies of the sources in question and produces the sampled visibility function. It is also important to note that the *uv*-plane is perpendicular to the direction of the phase tracking center. The points can be plotted onto a *uv*-plane to give some indication as to where the sampling points are for an instantaneous moment in time as seen in Figure 2.4. With tracking, as the field of interest moves across the sky, the *uv*-data points can be updated throughout the observation to obtain more samples in the *uv*-domain and increase the amount of *uv*-coverage.

When plotted on the *uv*-plane, the uv-points create tracks which are ellipsoidal in shape. The degree of ellipticity is dependent on the Declination of the observation. This is because for extended observations, the Earth rotates which changes the antenna positions. If these positions were projected and traced onto a plane, unless looking at a Declination of either 90° or 0°, the rotation would form traces of ellipses. A Declination of 90° would result in perfect concentric circles, while a declination of 0° would produce a horizontal line. There are also limitations as

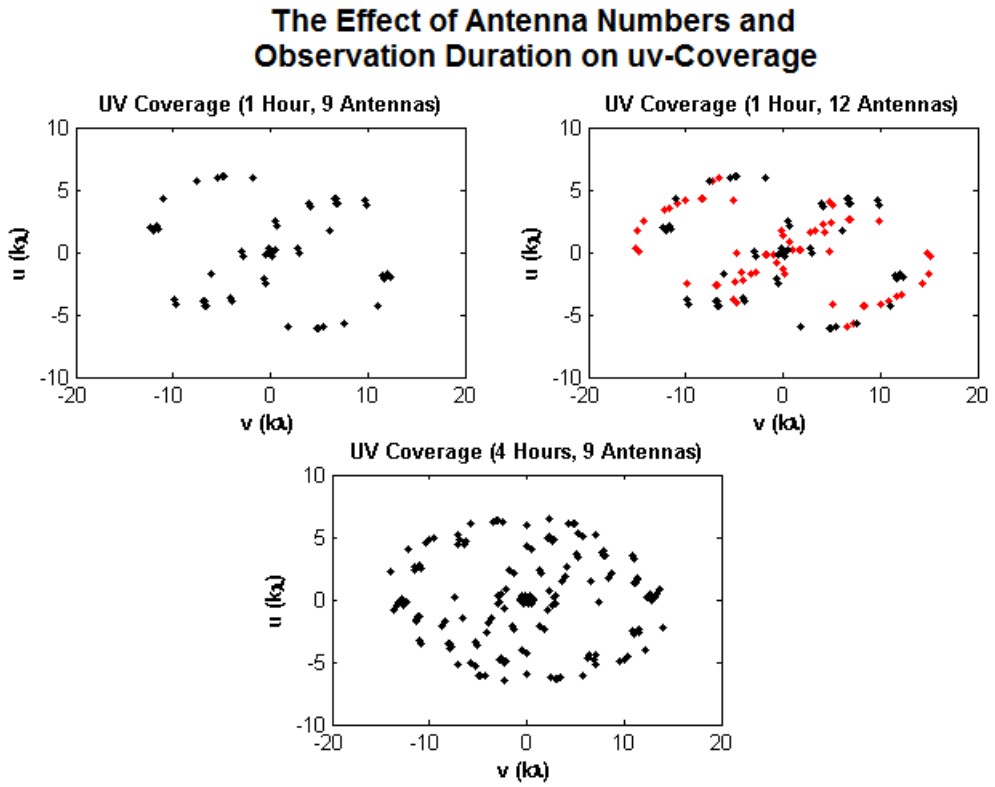


Figure 2.4: Various uv-coverage plots for a Declination of -25°, a frequency of 1.4GHz and a fixed sample interval of 30mins for a randomly generated array configuration. With increasing antenna numbers the coverage improves as there are more baselines to sample. The additional baselines for an increase of 3 antennas are indicated in red. The bottom image shows how increasing the observation length can further improve coverage for the same array with no added antennas.

to what Declination an array can view due to geographical locations, as an array situated in the Southern hemisphere would not be able to view a fraction of sky in the Northern Hemisphere due to the positioning. This is called shadowing, where sources are considered to be out of the observable FoV, shadowed or blocked out by the horizon [2].

In current practices such as for the MeerKAT telescope [8], the general trend for *uv*-coverage is that the core is more dense with data points as the arrays are designed for shorter baselines. As the distance from the origin increases, there are fewer points as these represent longer baselines which are less common. This is dependent on the antenna positions in the array configuration. Various *uv*-coverages are shown in Figure 2.4, illustrating that with more antennas or incorporating Earth rotation, the coverage can increase and the better the visibility function will become in representing the exact model. Figure 2.4 also shows a typical example of an array where the core is dense, with fewer points as the distance increases from the array site center.

When imaging the sky, there may be many sources contained in an image and these need to be individually defined. The phase tracking center becomes the reference point for describing the source locations. The source positions are specified in terms of directional cosines, l and m, with origin at the phase center. Figure 2.5 illustrates both the uv and lm coordinates. These directional cosines now correspond to points on a plane in the image domain on which the

brightness distribution lies that is both tangential to the celestial sphere at the phase tracking center as well as being parallel to the *uv*-plane. Note that all images, graphs and FoV data are described using the more general l and m directional cosines as they are relative to the phase reference center and are more general, rather than specifying exact Right Ascension and Declination Coordinates.

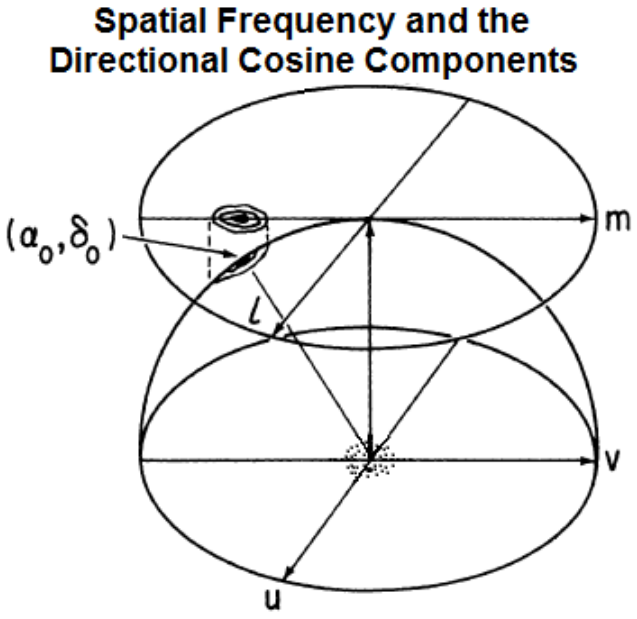


Figure 2.5: The source can be described by using the directional cosine components l and m, which are relative to the spatial frequency components. The u and v coordinates correspond to axes measured along the East-West and North-South directions respectively. Graphics were adapted from [7].

2.1.5 The van Cittert-Zernike Theorem

The full relationship between visibility and brightness is a three-dimensional model, confined to a sphere. This is shown in (2.1.6),

$$\mathbf{V}(u,v,w) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\mathbf{A}(l,m)\mathbf{I}(l,m)}{\sqrt{1-l^2-m^2}} e^{[-2\pi i(ul+vm+w(\sqrt{1-l^2-m^2})-1)]}\,\mathrm{d}l\,\mathrm{d}m, \tag{2.1.6}$$

where **A**(*l*,*m*) is the DDE which include the effects of polarisation and atmospheric effects on the signals, **V**(*u*,*v*,*w*) is the visibility, **I**(*l*,*m*) is the brightness distribution, and *u*, *v*, *w* and *l*, *m*, *n* are the spatial frequency and directional cosines used to describe the Fourier and image spaces, respectively.

For coplanar arrays and using the assumption that the FoV is narrow, the model can be simplified. This simplified model assumes the small angle approximation and that all the source data is on a tangential plane to the phase reference center. This implies

$$w \approx 0, \qquad l \approx 0, \qquad m \approx 0, \tag{2.1.7}$$

resulting in the third exponential term in (2.1.6) becoming

$$w(\sqrt{1-l^2-m^2}-1) \approx 0, \tag{2.1.8}$$

and a scaling factor of

$$\sqrt{1-l^2-m^2}) \approx 1. \tag{2.1.9}$$

In practice these assumptions are often made, as most arrays are coplanar or have sufficiently small variation in coplanarity that the effects can be ignored. In addition, the FoV is typically small meaning deviation from the tangential plane and the actual points on the celestial sphere is minimal. Furthermore, the DDE are not considered such as the antenna radiation patterns, polarisation and ionospheric effects, so $\mathbf{A}(l,m)$ can be incorporated into $\mathbf{I}(l,m)$. With these assumptions, (2.1.6) becomes the van Cittert-Zernike Theorem. The full model is discussed in Section 2.2 when non-coplanar arrays are used.

With a mathematical relationship existing between the spatial frequency components and the directional cosines, it implies that there is a mathematical relationship between the true Brightness distribution, $\mathbf{I}(l,m)$ and the observed visibility, $\mathbf{V}(u,v)$. The van Cittert-Zernike Theorem states that there exists a two-dimensional Fourier relationship between the brightness distribution in the image domain and visibility in the Fourier domain. Using the defined *uv* and *lm* coordinates and substituting (2.1.8) and (2.1.9) into (2.1.6), the relationship is defined as follows:

$$\mathbf{V}(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbf{I}(l,m) e^{-2\pi i(ul+vm)}\,\mathrm{d}l\,\mathrm{d}m, \tag{2.1.10}$$

or inversely,

$$\mathbf{I}(l,m) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbf{V}(u,v) e^{2\pi i(ul+vm)}\,\mathrm{d}u\,\mathrm{d}v. \tag{2.1.11}$$

This is extremely beneficial as it means if any one of these functions, $\mathbf{V}(u,v)$ or $\mathbf{I}(l,m)$ are known, a two-dimensional Fourier Transform can be used to obtain the other,

$$\mathbf{I}(l,m) \xrightarrow{\mathscr{F}^{-1}} \mathbf{V}(u,v) \qquad \mathbf{V}(u,v) \xrightarrow{\mathscr{F}} \mathbf{I}(l,m). \tag{2.1.12}$$

This is more complicated in practice as only discrete *uv* and *lm*-points are known, since for a practical system, $\mathbf{V}(u,v)$ is sampled at finite discrete points.

2.1.6 The Point Spread Function and Dirty Maps

The *uv*-coverage can be considered as a two-dimensional sampling function that samples the visibility in the Fourier domain. The sampled visibility can be seen as the product of the full continuous visibility function, $\mathbf{V}(u,v)$ and the sampling function, $\mathbf{S}(u,v)$. Applying the convolution theorem, if the sampled visibility were transformed into the image domain, it implies that the brightness distribution is convolved with the Fourier Transform of the *uv*-coverage, as shown in (2.1.13), to give

$$\mathscr{F}\{\mathbf{V}(u,v)\cdot\mathbf{S}(u,v)\} = \mathscr{F}\{\mathbf{V}(u,v)\} * \mathscr{F}\{\mathbf{S}(u,v)\} = \mathbf{I}(l,m) * \mathbf{B}(l,m), \tag{2.1.13}$$

where $\mathbf{B}(l,m)$ is the Fourier Transform of the sampling function or *uv*-coverage, not to be confused with brightness, $\mathbf{B}_r$. This is often referred to as the dirty beam or Point Spread Function (PSF) as it is the synthesised beam generated by the array configuration [7].

The PSF is a measure of how much a point in the image domain will be spread when viewed by the array configuration — essentially it is the impulse response of a point source. It is an important concept as the PSF is convolved with the sky intensity function, which can produce severely warped images if the intensity function contains multitudes of sources, or if the PSF is widely spread. For the ideal case, the *uv*-coverage will have sample points that cover the entire plane, resulting in $\mathbf{S}(u,v) = 1$. If this were transformed to the image domain, it would result in a Dirac delta function of unity weight, and the PSF would be an exact approximation of a point source, excluding the effect of noise. However, this is unrealistic and source data in the sky is often warped by the dirty beam, resulting in a dirty map. Dirty maps are the distorted models of the sky, as seen by an array, which in the Fourier space can be directly related to missing visibility data due to insufficient sampling.

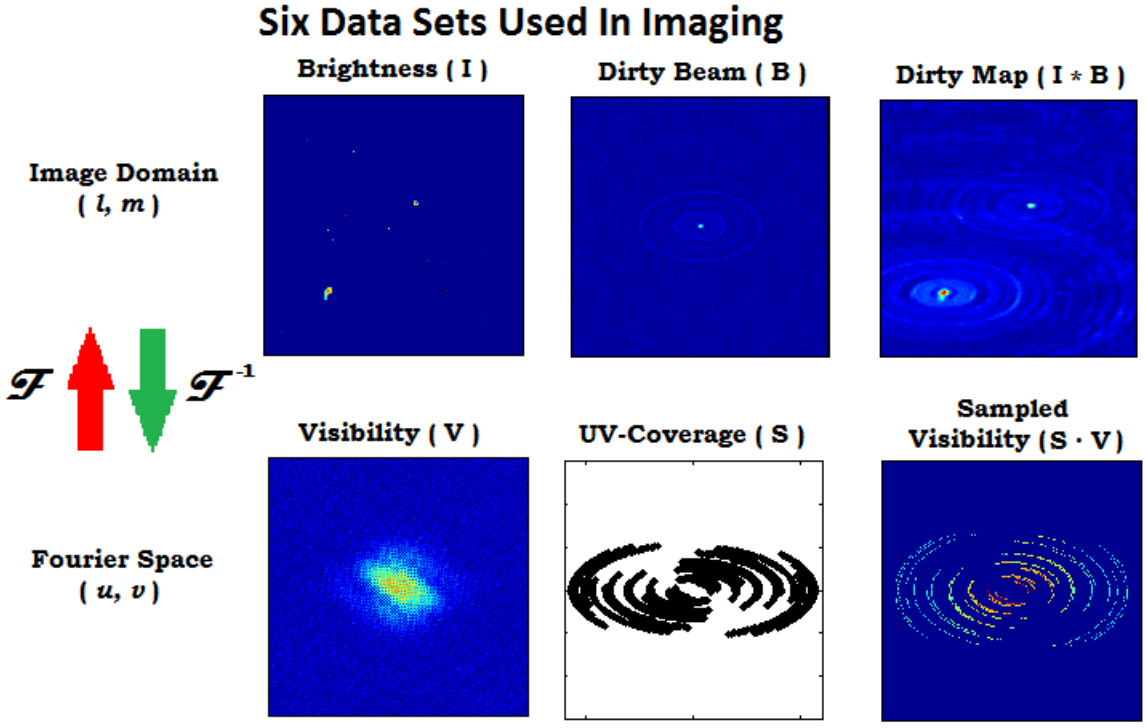


Figure 2.6: A graphical representation of the various data sets and how each set is related to another. The top row relates to data in the image space, while the bottom is in the Fourier space.

Figure 2.6 shows the six sets of data that are essential to imaging: The image domain consists of the sky intensity distribution, the PSF or dirty beam and the resulting dirty map. In the Fourier domain there is the visibility function, the *uv*-coverage or sampling function and the resulting sampled visibilities. In practice, the goal of imaging techniques is often to remove

the distortion, noise, aliasing and other effects that influence the quality of the dirty map, to more accurately model the true brightness distribution. With the fundamentals covered, the theory can now be expanded to include the effects of non-coplanar arrays which has become ever-increasingly more widespread as the demand for better quality images and more accurate visibility data grows.

2.2 Imaging for Non-Coplanar Arrays

2.2.1 The *w* Spatial Frequency Component

It is common practice to make various assumptions about the visibility, $\mathbf{V}(u,v)$, and the brightness distribution, $\mathbf{I}(l,m)$, to ensure that the two-dimensional Fourier Transform relationship still holds. However, as imaging techniques become more advanced, the full three dimensional relationship needs to be considered to obtain more accurate data. This is further motivated by the development of advanced array instrumentation and further, more computationally effective software which can support larger and more complex data structures.

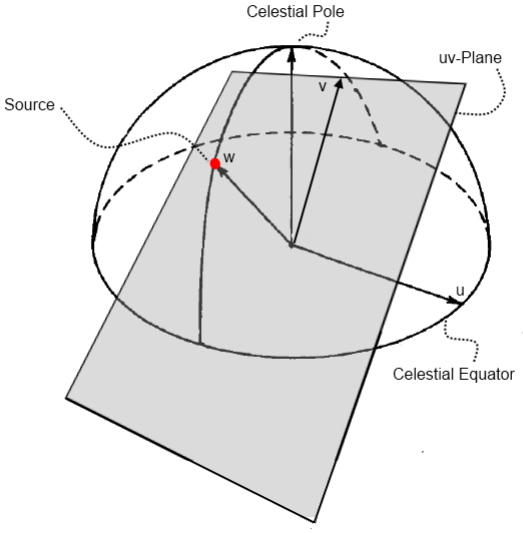


Figure 2.7: The uv-Plane with w pointing to the source, located at the phase reference center. Baselines are projected onto this plane. Graphics were adapted from [2].

For the standard two-dimensional visibility model, the approximation is frequently made that the synthesis array is coplanar with $w \approx 0$, or that the imaging data is near the phase reference center. For non-coplanar baselines and wide-field imaging, this may not be the case, requiring the full three-dimensional visibility model to be used [9], as expressed in (2.1.6).

The *w*-component is determined by the array configuration along with observation parameters such as the source position and the observed frequency, as with its counterpart components, *u* and *v*. For an array configuration pointing, the *u*-component is the baseline projected along the East-West direction, the *v*-component is along the North-South component and *w*-points

to the phase reference center — this is seen in Figure 2.7. Any deviation in the perpendicular height with respect to the *uv*-plane will increase the *w*-component because of the extra delay. Additionally, a source located away from the phase tracking center will result the direction cosines increasing. The combination of both of these cases results in the additional third exponential phase term in (2.1.6).

Due to the spatial component *w* in (2.1.6), the relationship is not a two-dimensional Fourier Transform; the visibility is a function of three variables, while the brightness is only a function of two independent variables — so the estimated sky brightness cannot be calculated from a Fourier inversion. The implications on imaging and visibility prediction are further discussed in the chapters to follow. Note that $\mathbf{A}(l,m)$, the direction dependent effects in this visibility model are incorporated into $\mathbf{I}(l,m)$ and this assumption will be used throughout the thesis.

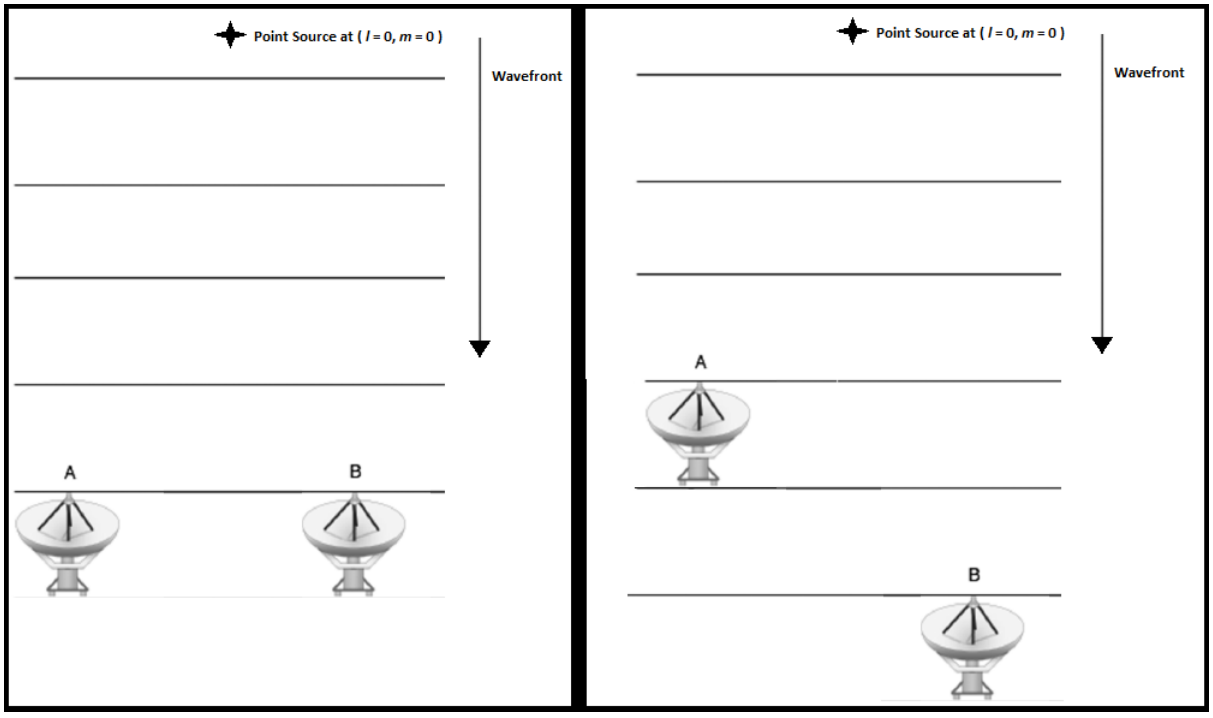


Figure 2.8: Left) Coplanar Antenna Configuration - Antenna A will receive the wavefront at the same instant as Antenna B, resulting in a perfectly coherent signal. Right) Non-coplanar Antenna Configuration - Antenna A will receive the wavefront before Antenna B, resulting in an incoherent signal. Graphics were adapted from [10].

Consider an East-West, two-element interferometer with no North-South component and with a single point source directly above at the phase reference center, as seen in Figure 2.8. For the coplanar case, the two antennas will receive the radiation wavefront from the source exactly in phase and the measured correlation of this data will be the standard Fourier relationship. Now consider the non-coplanar case. When one of the antennas is shifted to an offset above or below the *uv*-plane, the radiation wavefront from the source will not simultaneously reach each receiver, introducing a phase shift. This phase shift must be taken into consideration, which leads to the breakdown of the Van Cittert-Zernike Theorem [10]. The phase shift can be corrected by inserting a delay but this is not possible for multiple directions simultaneously.

2.2.2 Tangent Plane Approximation and *n*

The full visibility model is not solely dependent on the spatial frequency components, *u*, *v* and *w*, but also the directional cosines, *l*, *m* and *n*. In (2.1.6), the third exponential term containing *w*, has a dependence on the variable *n*. These directional cosines denote a position on the celestial sphere, parameterised by

$$l^2 + m^2 + n^2 = 1, \tag{2.2.1}$$

where *n* is restricted to one hemisphere and has a dependence on *l* and *m* given by

$$n = \pm\sqrt{1 - l^2 - m^2}. \tag{2.2.2}$$

Substitution of *n* into equation (2.1.6) results in

$$\mathbf{V}(u, v, w) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\mathbf{I}(l, m)}{n} e^{[-2\pi i(ul+vm+w(n-1))]}\,\mathrm{d}l\,\mathrm{d}m. \tag{2.2.3}$$

where the third exponential term contains *w* and *n*. The variable *n* can be seen as a measure of error for the tangent plane approximation; the more it deviates from unity, the greater the phase shift error becomes. For both (2.1.6) and (2.2.3), this extra term has been shifted by one unit radius for phase tracking purposes [9].

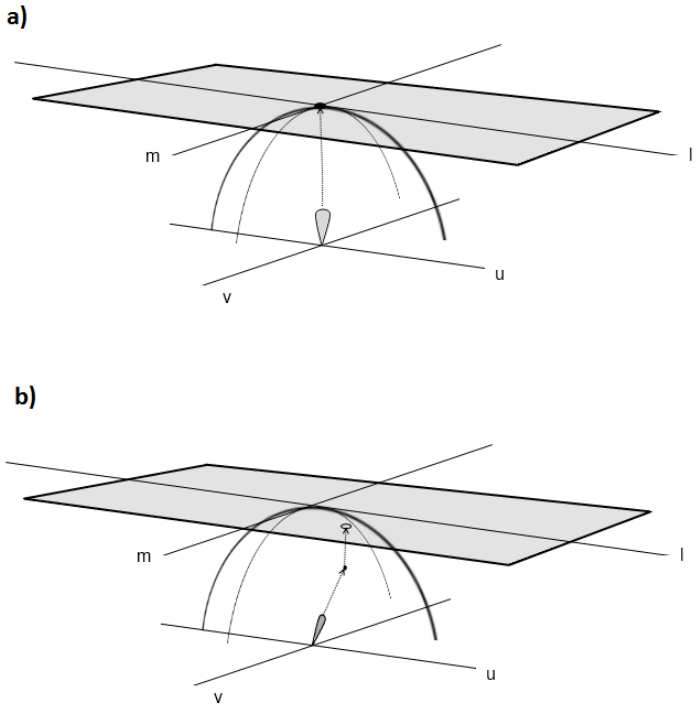


Figure 2.9: a) Projection of a point at the phase reference center, indicating no error when projecting onto the Tangent Plane (n = 1). b) Projection of a point away from the phase reference center, indicating the need to approximate the value on the Tangent Plane. Graphics were adapted from [10].

This additional directional cosine can be seen as changing the image plane to that of an image volume or cube [9]; however n restricts the relevant $\mathbf{I}(l,m)$ values to lie on the celestial sphere of unit radius. Any other values that do not lie on this sphere are meaningless. The visibilities predicted are obtained from the image plane, however the actual brightness values lie on the unit sphere. This leads to distortion when reconstructing the image. An easier way to view the relationship would be to view it as the 3D brightness values being projected onto a 2D image tangent plane, which is then used to acquire the 3D visibilities.

When the brightness is projected from the unit sphere onto the tangent plane, it introduces an error which increases as the point under consideration shifts from the array tracking center, as the baseline projections are dependent on the direction of observation. While this is only a geometric representation of the projection, Figure 2.9 illustrates this error more clearly.

2.2.3 Implications On Imaging

With both variables w and n explained, it is evident that they have a detrimental influence on phase shifting and the combination of these variables in the exponential term of (2.2.3) requires careful consideration when trying to accurately model the mutual coherence data or reconstructing a high fidelity image. For non-zero w and n values that digress from unity, the exact additional phase shift is given by

$$\phi = 2\pi w(\sqrt{1 - l^2 - m^2} - 1) = 2\pi w(n - 1), \tag{2.2.4}$$

and becomes increasingly important for:

I. Wide Field of View imaging - the directional cosine, n, as defined in (2.2.2), depends on the exact sky position and large l and m affect phase shifting and limits imaging capability — the greater the FoV, the larger these variables become and the larger the phase shifts.

II. Non-coplanar baselines - height deviations in the array that cause visibility data to lie outside the Fourier uv-plane which is perpendicular to the phase tracking center, result in a frequency dependent, non-zero w-component [11].

Only these two are considered in the scope of this analysis; however one must take note that DDE's, such as primary antenna beam patterns, tropospheric and ionospheric effects, Faraday Rotation and polarisations, all impact $\mathbf{V}(u,v,w)$ for both non-coplanar arrays and wide FoV imaging [12].

The phase shift in (2.2.4) is non-linear due to the dependence on the directional cosines, while w acts as a scaling factor to the error, illustrated by Figure 2.10. It must also be noted that w is baseline dependent — each baseline in the uv-plane will have an associated w-component — generating a unique phase shift for each $\mathbf{V}(u,v,w)$ value. There is also a restriction on the FoV, where l and m are bound to

$$l^2 + m^2 \leq 1, \tag{2.2.5}$$

or else n becomes a complex quantity. Also, n is always positive as it is restricted to one hemisphere.

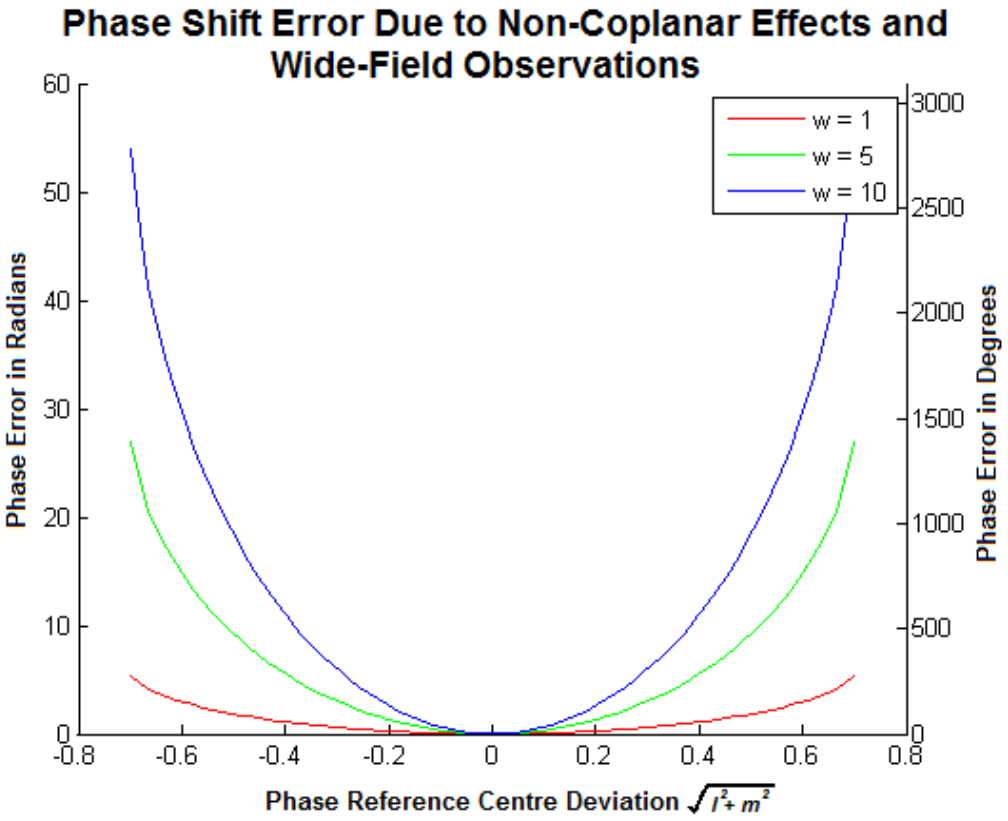


Figure 2.10: The phase shift due to various w-values and increasing FoV size which affects n.

The effect of the phase shift, if it is not accounted for, will result in distorted images when trying to reconstruct the intensity models from the visibilities or vice versa. The greater the w offset, the more warped an image becomes as illustrated in Figure 2.11. For a single point source located in the FoV, other than at the phase center as this would result in $n = 1$, the visibility data was calculated using the full 3D model. A uniform grid of uv-points were simulated, where each uv-coordinate was separated by 1λ and the maximum baseline was 60λ. This was chosen purely for the sake of simulation, as using a uniform grid meant that no gridding function was needed. Values of w were selected to be 0, 2, 10 and 20 to generate a phase shift. The visibilities were then Inverse Fourier Transformed using the van Cittert-Zernike Theorem, so that the phase shift was not accounted for when trying to image the sky model. By ignoring the phase shift caused by the w offset, the transformation from the Fourier space to the image space becomes warped, making the source appear as if it is smeared, thereby affecting the image fidelity.

2.2.4 Correcting for the *w*-Term

In the following two chapters, w-projection and w-stacking are discussed but it is important to note that these are not the only methods to compute non-coplanar visibilities. Various algorithms suggest both practical and purely theoretical methods to compute the full three-dimensional visibility model using non-coplanar baseline effects, with the most common being Warped Snapshots and uvw-Facets [13]. The most accurate approach would be to reformulate (2.1.6) to take the form of a 3D Fourier Transform [3]. With the advances in technology, it has been argued that the computational requirements to perform the three-dimensional Fourier Transform are not so demanding nowadays in terms of the number of operations required; however, the 3D Fourier Transform can still become expensive.

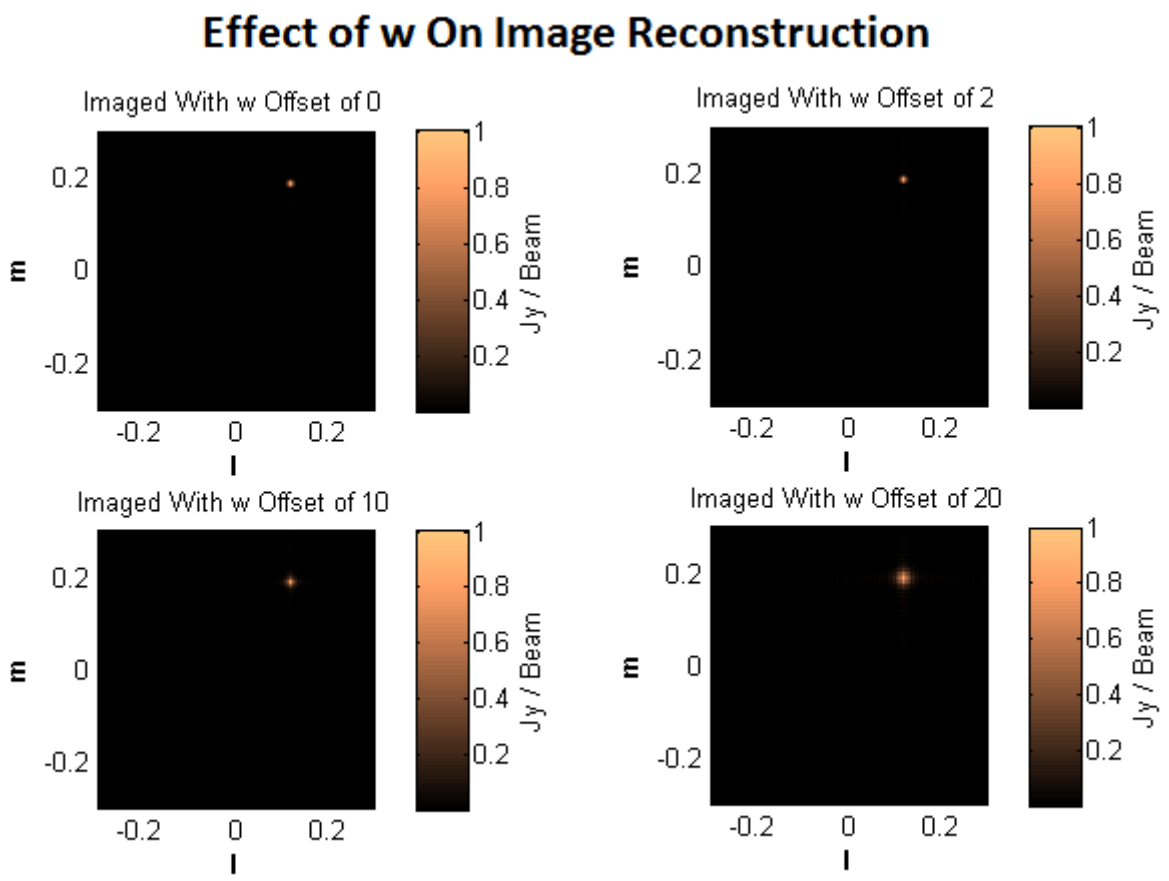


Figure 2.11: An off-center point source located in the FoV. The images are reconstructed using the direct two-dimensional Fourier Transform in (2.1.11)*, discarding the effect of w. For greater w-values, the point source becomes more distorted.*

Even with advances in computing, the problem remains that much of the data in the image volume is non-physical and transforming every point within this cube is inefficient, as all data points are computed but those that lie outside the unit sphere are discarded. This is especially true for the additional w-term, as this dimension is sparse but the three dimensional FFT still requires a full uniform grid, so this sparse space is still used. In some instances, using the sparsity of this dimension can work to the three dimensional FFT's benefit such as Spread Spectrum [14] or Compressed Sensing techniques [15]. Furthermore, gridding also becomes more expensive as one must grid the extra dimension data on a uniform grid for the 3D FFT and to do this the curved sphere surface must be interpolated.

Another workaround would be to change the coordinate system to only account for the true emission points that lie on the unit sphere, but this requires each baseline projection to be converted to spherical coordinates. In turn, the Fourier Transform must then utilise first and second order Bessel Functions [16]. This is computationally demanding for long observations but the benefit is that the only error in the visibility data will be due to aliasing, numerical precision capabilities and the model being discrete, as no approximations are used. However, the complexity of this operation much outweighs that of the approximation methods. Section 6.2.3 contains a computational analysis for the 3D FFT, w-projection and w-stacking methods.

2.3 Other Considerations

The non-coplanar effect and large field observations are not the only factors to consider when imaging. Aside from DDE's, two other aspects that detrimentally affect image quality and fidelity are discretisation and aliasing when using the Fast Fourier Transform (FFT). In (2.2.3), both the sky brightness and the mutual coherence functions are continuous; however, they are subjected to sampling where they become finite and discrete. The discretised form of the three variable van Cittert-Zernike integral for an **I**(*l*,*m*) that only contain point sources which are represented by Dirac deltas, results in (2.1.6) becoming a summation of data, rather than integration. **I**(*l*,*m*) becomes a weighted sum of Dirac delta functions and the distribution can be related to the image pixels by

$$\mathbf{V}(u,v,w) = \sum_{p=1}^{P} \frac{M_p \delta(l_p, m_p)}{\sqrt{1-l_p^2-m_p^2}} \; e^{[-2\pi i(ul_p+vm_p+w(\sqrt{1-l_p^2-m_p^2})-1)]}, \tag{2.3.1}$$

where *P* is the total number of pixels, *p* is the pixel index and M_p is the weighting or magnitude of the source in the portion of the FoV at pixel (l_p,m_p).

With discretisation comes possible aliasing issues. For wide-field imaging the sampling of the *uv*-points needs to be frequent enough to avoid aliasing issues. It is analogous to the time-frequency FFT relationship more familiar to electronic engineers in the context of signal processing theory: the image space is comparable to the time domain, and the visibilities to that of temporal frequency. All variable pairs, namely *u* and *l*, *v* and *m*, or *w* and *n*, for one-, two- and three dimensions, respectively, must satisfy Nyquist sampling conditions.

For the one-dimensional case consider the signal processing analogy as seen below. For a given frequency, ω, the signal needs to be sampled at a rate above the Nyquist rate to avoid aliasing. For the inverse case, the correct frequency components must be used for the time signal to be perfectly reconstructed.

$$\begin{aligned} u \quad &\text{(Baseline)} \implies \omega \quad \text{(Frequency)}, \\ l \quad &\text{(Directional Cosine)} \implies t \quad \text{(Time)}, \\ \mathbf{V}(u) &= \int_{-\infty}^{\infty} \mathbf{I}(u) \; e^{-2\pi i u l} \, \mathrm{d}l, \\ \mathbf{I}(l) &= \int_{-\infty}^{\infty} \mathbf{V}(u) \; e^{2\pi i u l} \, \mathrm{d}u, \end{aligned}$$

In radio astronomy, if the maximum spacing between various baseline points Δu becomes too large, the FoV will have to be confined to a smaller fraction of the sky, to limit the directional cosine *l*, else aliasing will occur. Vice versa with reproducing the image and using the inverse transform, for a given *l*, the baseline spacings must be small enough to resolve the sources in the FoV, or else aliasing and ghost images of the frequency components will be generated. This model can be extended to encompass the second and third dimensions, where all boundaries must hold. The exact boundaries for sampling are defined by [11]

$$\Delta l \leq \frac{1}{2u_{max}}, \qquad \Delta m \leq \frac{1}{2v_{max}}, \qquad \Delta n \leq \frac{1}{2w_{max}}, \tag{2.3.2}$$

or in the reverse case of going to the image domain,

$$\Delta u \leq \frac{1}{2l_{max}}, \qquad \Delta v \leq \frac{1}{2m_{max}}, \qquad \Delta w \leq \frac{1}{2n_{max}}. \tag{2.3.3}$$

For the case where the 2D and 3D models are used, the sampling must satisfy all the restrictions in (2.3.2) and (2.3.3); if one of the combinations does not satisfy the Nyquist criteria while the other two are met, then the overall image or visibility data will be incorrect, resulting in duplication of sources within the same image or large sidelobes which distort magnitude information [11].

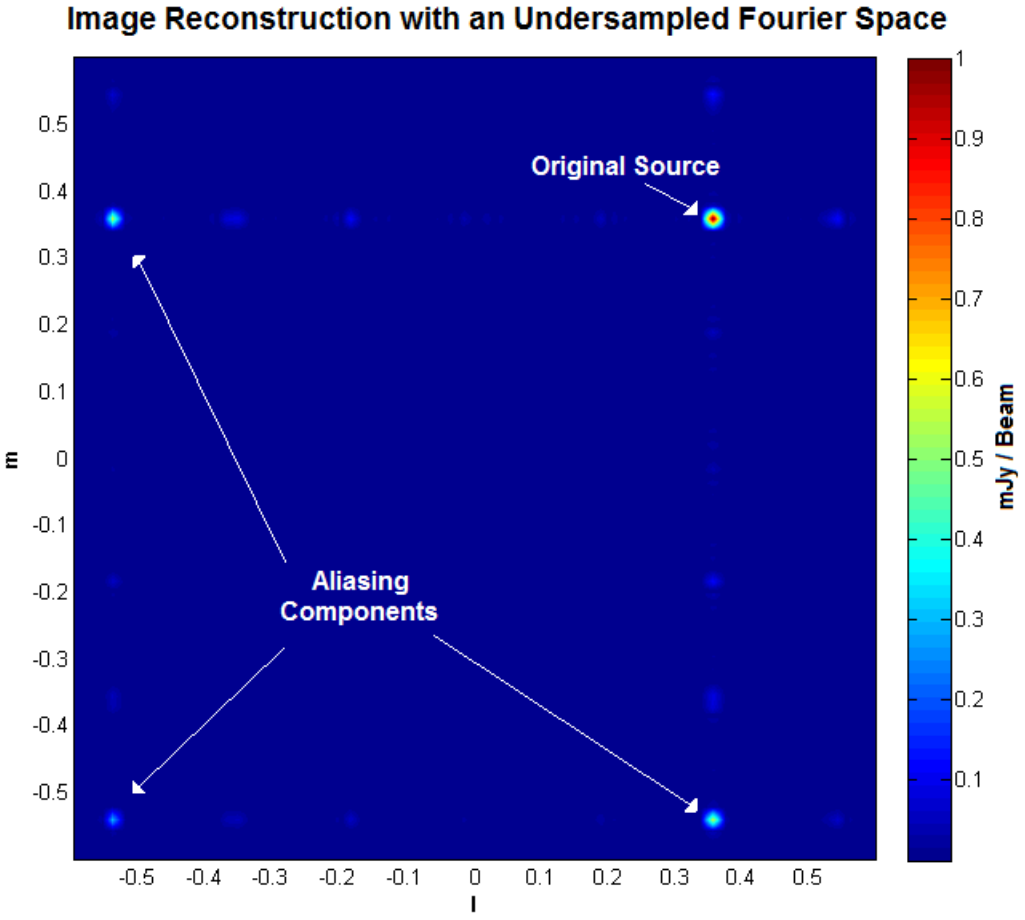


Figure 2.12: The original map contained a single, off-center source, but due to undersampling in the uv-plane, the mirrored frequency components appear in the image in both the vertical and horizontal directions because of the two-dimensional IFFT. This aliasing effect can be avoided by proper sampling to meet Nyquist conditions.

Figure 2.12 shows the aliasing effect. For this case, both the Δu and Δv conditions in (2.3.3) were not satisfied. Again, a uniform grid of *uv*-points was generated and the 2D Fourier Transform was used to generate $\mathbf{V}(u,v)$ from a sky model containing one off-center source. The image was processed with Δu being too large in the *uv*-plane. The sampled points did not satisfy the Nyquist criteria, which lead to aliasing. These frequency components that mirror back into the image could be mistaken for ghost images. The frequency components that start to appear in the image are along both *l* and *m* axes, as both the Δu and Δv criteria were unsatisfied.

It is clear from the aforementioned issues with non-coplanar baselines that the w-component detrimentally affects the image data if it is ignored. The extra phase shift can distort the intensity and visibility functions, affecting the fidelity of the images. Furthermore, it can introduce aliasing components into the images. The next chapter proposes an algorithm that removes the effect of the additional phase shift, essentially projecting the non-coplanar visibilities to coplanar points. This algorithm is called w-projection.

Chapter 3

The W-Projection Algorithm

3.1 The *w*-Projection Imaging Technique

A popular technique used in non-coplanar imaging was suggested by Cornwell et al. [3], where they proposed one could shift points located in the *uvw*-space, to other points located in the same space but at a different *w*-value. This method of shifting a point to another location is called projection, forming the basis of this chapter and *w*-projection. It has been fully integrated into modern day imaging software such as CASA and Meqtrees, and algorithms such as CLEAN, which can utilise *w*-projection as an added parameter.

The idea is to project a three-dimensional visibility point located in the *uvw*-space, where *w* is at zero, to another point at different specified *w*. This was initially proved by Frater and Docherty [17] for a single point, where *w*-projection adapted this theory to encompass many *w*-points. This process shifts the visibility values, essentially the 2D visibility values at $\mathbf{V}(u,v,w{=}0)$, to the 3D location at $\mathbf{V}(u,v,w)$.

The operation can also be reversed by applying the inverse of the projection algorithm, projecting a visibility value of non-zero *w*, to the *w*=0 plane. Note that projection can be used to describe both the forward and reverse cases but re-projection specifically refers to the projection of a point to *w*=0. The implication of projecting the visibilities to or from the *w*=0 plane, is that the *w*-term in the full 3D mutual coherence and brightness relationship, is completely eliminated from the problem so that the van Cittert-Zernike Theorem still holds true and only a single 2D Fourier transform is needed to produce an image model.

To reiterate,

$$\mathbf{V}(u,v,w) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\mathbf{I}(l,m)}{\sqrt{1-l^2-m^2}} e^{-2\pi i[ul+vm+w(\sqrt{1-l^2-m^2}-1)]}\,\mathrm{d}l\,\mathrm{d}m, \tag{3.1.1}$$

which would reduce to

$$\mathbf{V}(u,v,w=0) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\mathbf{I}(l,m)}{\sqrt{1-l^2-m^2}} e^{[-2\pi i(ul+vm)]}\,\mathrm{d}l\,\mathrm{d}m, \tag{3.1.2}$$

once the value is projected to the *w*=0 plane. While imaging a large FoV will still cause some magnitude distortion due to the directional cosines *l* and *m*, which acts as a scaling factor as

seen in the denominator in (3.1.2), the third phase term in the exponent argument has been removed. The following section elaborates on the details of the algorithm and how to project the points.

Taking the exponent in the integrand of (3.1.1) and using the exponential laws, one can separate the exponential into two parts

$$e^{-2\pi i[ul+vm+w(\sqrt{1-l^2-m^2}-1)]} = e^{-2\pi i(ul+vm)}\ \mathbf{G}(l,m,w), \tag{3.1.3}$$

$$\text{where} \quad \mathbf{G}(l,m,w) = e^{-2\pi iw(\sqrt{1-l^2-m^2}-1)}. \tag{3.1.4}$$

The separated *w*-term, **G**(*l*,*m*,*w*), is considered to be a multiplicative distortion in the image space, and is vital to the projection. **G**(*l*,*m*,*w*) has variables in the image domain, but also has a dependency on *w*, which is located in the Fourier space. According to the convolution theorem, two functions being multiplied in a point-wise manner, and then Fourier transformed, will produce the same result as separately transforming each function and then convolving the Fourier transformed factors, as seen in (3.1.5),

$$\mathscr{F}\{a \cdot b\} = \mathscr{F}\{a\} * \mathscr{F}\{b\}, \tag{3.1.5}$$

where * is the convolution operator. The desired objective is to find a convolution kernel, $\bar{\mathbf{G}}$(*u*,*v*,*w*), that can satisfy (3.1.6),

$$\mathbf{V}(u,v,w) = \bar{\mathbf{G}}(u,v,w)\ *\ \mathbf{V}(u,v,w=0), \tag{3.1.6}$$

where $\bar{\mathbf{G}}$(*u*,*v*,*w*) is a convolution kernel that will shift the values of **V**(*u*,*v*,*w*=0) to the *w*-points determined by the array and the observation parameters. $\bar{\mathbf{G}}$(*u*,*v*,*w*) is used to indicate the Fourier space function and to distinguish it from **G**(*l*,*m*,*w*) in the image space. To compute $\bar{\mathbf{G}}$(*u*,*v*,*w*), the theorem in (3.1.5) must be applied to produce a convolution kernel that will be the key to *w*-projection. With **G**(*l*,*m*,*w*) having variable dependencies in both the image domain and the Fourier space, if the 2D Fourier transform is employed with respect to the image domain variables, *l* and *m*, then *w* will remain unaffected by the transform, and will produce a Fourier space kernel, $\bar{\mathbf{G}}$(*u*,*v*,*w*), as

$$\bar{\mathbf{G}}(u,v,w) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbf{G}(l,m,w)e^{-2\pi i(ul+vm)}\,\mathrm{d}l\,\mathrm{d}m, \tag{3.1.7}$$

where **G**(*l*,*m*,*w*) is defined in (3.1.4). A closed form solution of this integral is not known and the suggested approach by Cornwell et al. [11] is to use numerical analysis and simplify the model using the small angle approximation. The small angle approximation uses Taylor expansion to approximate the exponential phase shift term, resulting in

$$e^{-2\pi iw(\sqrt{1-l^2-m^2}-1)} \approx e^{\pi i[w(l^2+m^2)]} \qquad \text{for } l,m \ll 1. \tag{3.1.8}$$

The substitution of (3.1.8) into (3.1.7) gives

$$\hat{\mathbf{G}}(u,v,w) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{\pi i[w(l^2+m^2)]} e^{-2\pi i(ul+vm)}\,\mathrm{d}l\,\mathrm{d}m = \frac{i}{w}\, e^{-\pi i[\frac{u^2+v^2}{w}]}. \tag{3.1.9}$$

$\hat{\mathbf{G}}(u,v,w)$ is used to denote the convolution kernel that is determined by the approximated version of $\mathbf{G}(l,m,w)$, which differs from that of $\bar{\mathbf{G}}(u,v,w)$, which uses the non-approximated result of (3.1.7). The expression in (3.1.9) is the desired kernel we need to project the visibility values to the given w-points using (3.1.6). This form of $\hat{\mathbf{G}}(u,v,w)$ for a constant w has a constant magnitude in the uv-plane but the phase shift is a 2D Gaussian, where the shift is scaled by each w-value. While seemingly simple the application of this kernel is extremely useful: one is able to determine the 3D visibility data, $\mathbf{V}(u,v,w)$, by using only the 2D data $\mathbf{V}(u,v)$. This process can also be reversed to re-project a visibility value at a specific w, to the $w = 0$ plane, so that the visibilities only lie within the uv-plane, where normal imaging techniques can be used that do not need to account for the additional phase shift as it has been removed.

Alternatively, looking at the real part of $\mathbf{G}(u,v,w)$,

$$\mathscr{R}\{\mathbf{G}(u,v,w)\} = \frac{\sin(\frac{\pi(u^2+v^2)}{w})}{w}, \tag{3.1.10}$$

it can be seen that for increasing w, the oscillations of the sine term increase and become smaller in magnitude. Additionally, for w=0, the kernel diverges [18]. This is because the convolution kernel was based on the small angle approximation and the $\sqrt{1-l^2-m^2}$ scaling factor was not included in the definition of $\mathbf{G}(l,m,w)$ as suggested by Humphreys [18], to avoid issues of divergence. This factor can be included to get

$$\mathbf{G}_{full}(l,m,w) = \frac{e^{-2\pi iw(\sqrt{1-l^2-m^2}-1)}}{\sqrt{1-l^2-m^2}}. \tag{3.1.11}$$

Note that the right hand side of (3.1.9) is only used to show the form of the kernel. To coincide with Cornwell [3], $\hat{\mathbf{G}}(u,v,w)$ is determined by Fourier transforming the approximated $\mathbf{G}(l,m,w)$ in (3.1.8) which is determined numerically using the l, m and w-points. It has been noticed that for different literature sources such as [18], the convolution kernels and their Fourier counterparts have a sign difference in their definitions. This is due to the definitions of the forward or reverse Fourier transforms, where the polarity of the exponential term is either positive or negative. The derivations below are based on the forward transform having a negative exponent, coinciding with the majority of the literature regarding w-projection.

It was shown in Chapter 2 that non-coplanar antennas lead to decorrelation of sources. This can be better understood by using this projection kernel. The kernel is applied in the uv-plane to all the visibility samples, which are essentially points. By applying the convolution along this plane, with each convolution being dependent on the w-value associated with the visibility sample, one is convolving a Dirac Delta with the complex Gaussian, which results in another complex Gaussian that is weighted by the visibility point. Note that each kernel is unique to each w-value.

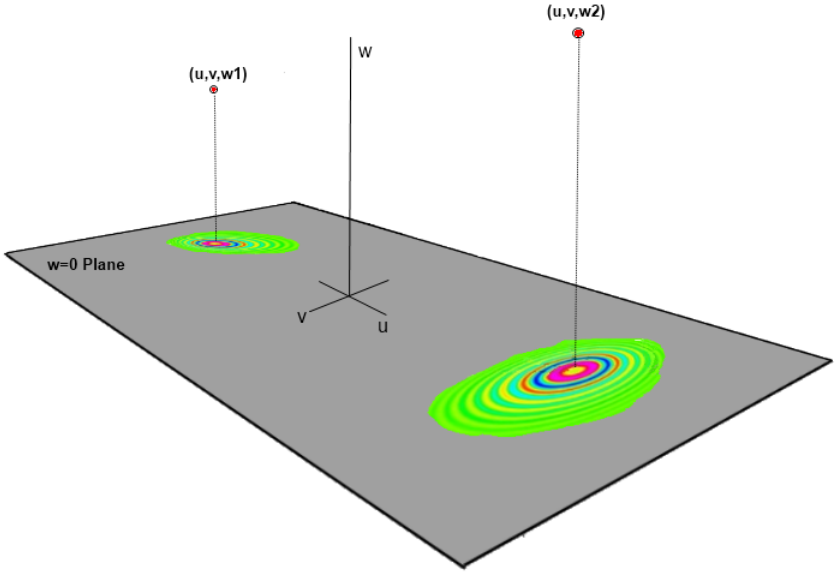


Figure 3.1: A graphical representation of the point spreading when projecting a point to another location. Graphics were adapted from [13].

Figure 3.1 is a graphical representation of how sampled visibility points for non-zero w-values, appear to be smeared when they are assumed to be on the w=0 plane. This will cause distortion in the image domain if these visibilities are transformed using the van Cittert-Zernike theorem.

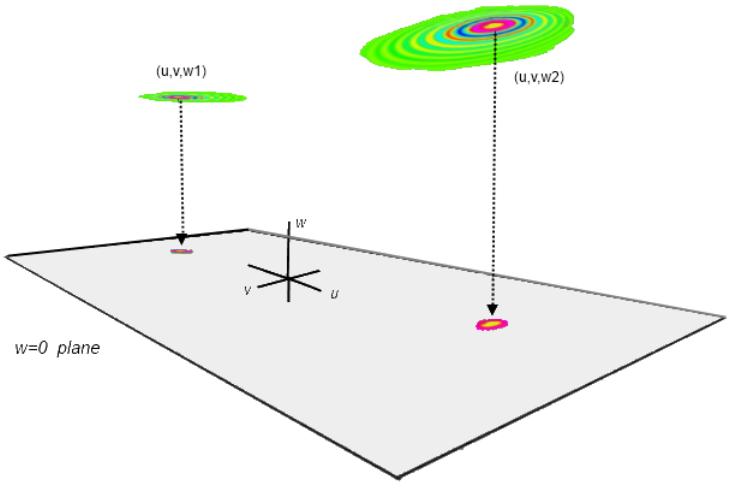


Figure 3.2: The effect of smearing can be minimised when the sample is projected onto the w=0 plane and then imaged using the Van Cittert-Zernike theorem. Graphics were adapted from [13].

In imaging, the reverse case is of interest — one wants to shift the visibility points to the same plane as shown in Figure 3.2, where two visibility points, that are of different w-values are shown. By convolving with their corresponding $\hat{\mathbf{G}}(u,v,w)$, this can be achieved. A point at $\mathbf{V}(u,v,w)$ can be projected to $\mathbf{V}(u,v,w{=}0)$ so that the seemingly smeared visibility data for non-zero w becomes more condensed when it is transformed into the image domain. This is because the visibilities transformed by the van Cittert-Zernike theorem, are no longer phase

shifted by the non-coplanarity; the phase is removed by making $w \approx 0$. It will now resemble a point source more accurately — the source will be less distorted and true to its actual form. There will still be a spreading of the point in the image domain as the projection is an approximation method and there are other factors such as sampling, aliasing and gridding that further distort the data; however, this is a clear graphical representation of what is happening with *w*-projection.

3.2 The Algorithm Structure

The algorithm steps for both projection and re-projection are highlighted below. Using projection to predict visibilities from the image domain:

I. Given a sky model, perform the two dimensional FFT to obtain $\mathbf{V}(u,v,w{=}0)$.

II. Compute the convolution kernel, $\hat{\mathbf{G}}(u,v,w)$, for the corresponding *w*-values.

III. Apply the convolution to project $\mathbf{V}(u,v,w{=}0)$ to the correct $\mathbf{V}(u,v,w)$ plane.

Theoretically, this is useful as the visibilities can be predicted by taking the *w* into account and also has imaging benefits for CLEANing algorithm models that use the predicted visibilities such as Cotton-Schwab [4]. If the $\mathbf{V}(u,v,w)$ were inverse Fourier transformed, it would result in a dirty map where all the sources experience some form of data smearing depending on their position in the FoV. However, in practice one would want to go in the opposite direction — taking the sampled $\mathbf{V}(u,v,w)$ function and acquire the 2D $\mathbf{V}(u,v)$, which has eliminated the non-coplanar baseline effects and produces an accurate sky model on the tangent plane [10].

Going from the Fourier space for non-zero *w* to $w = 0$, and then imaging the 2D function proceeds as follows:

I. Numerically evaluate the convolution kernel $\hat{\mathbf{G}}(u,v,w)$, for the reciprocal $\mathbf{G}^{-1}(l,m,w)$ at the given *w*-values.

II. Given the sampled 3D visibilities, $\mathbf{V}(u,v,w)$, use $\hat{\mathbf{G}}(u,v,w)$ to project the visibilities to the $w{=}0$ plane.

III. Inverse FFT the gridded $\mathbf{V}(u,v,w{=}0)$ to obtain the phase corrected dirty map.

3.3 Understanding the functions $\mathbf{G}(l,m,w)$ and $\hat{\mathbf{G}}(u,v,w)$

Projection is based on the two functions $\mathbf{G}(l,m,w)$ and $\hat{\mathbf{G}}(u,v,w)$, where understanding the form and use of each of these gives insight into why *w*-projection works and what is happening to the data in both the Fourier and image domains.

Regarding the function in the image domain, $\mathbf{G}(l,m,w)$, the approximated version is in the form of a 2D chirp function. The function has a constant magnitude of unity across the plane, but a varying phase distribution, where the combination of $|l|$ and $|m|$ define the distance to the origin of the phase plane, and *w* defines the rate of the phase variation [19].

For a given field of view bounded by $\sqrt{l^2 + m^2} \leq 0.5$, $\mathbf{G}(l,m,w)$ was numerically calculated for various w-values. In Figure 3.3, the phase shift distribution shows that the rate at which the phase shifts increases and is proportional to the w-value [14]. While it is the Fourier transform of this that is used for projection, one can see that the phase shift that is applied to the brightness distribution is severely affected. Furthermore, Figure 3.3 only highlights the effect of $\mathbf{G}(l,m,w)$ for a constant w. It is the combination of all the image domain distortions for all baselines, individually determined by the corresponding w-values, that results in the final image. Understanding what happens in the image domain for projection is complex, which is why it is done in the Fourier domain, where it is more practical and easily analysed.

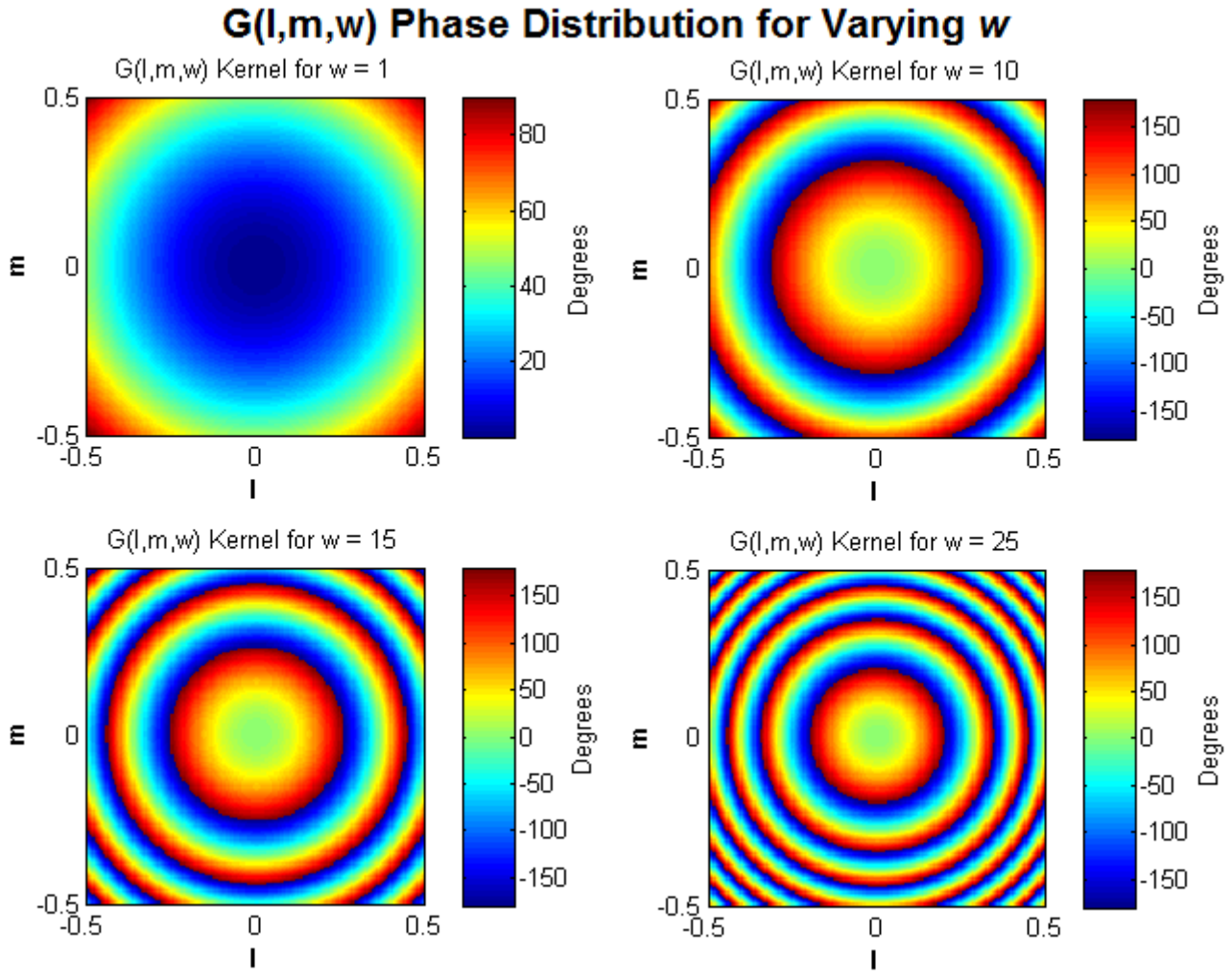


Figure 3.3: G(l,m,w) phase distribution highlighting the rapid increase of the spatial frequency of the phase shifts as w increases.

Numerically determining $\mathbf{G}(l,m,w)$ and transforming it to the visibility domain, as stated in (3.1.9), gives the projection kernel, $\hat{\mathbf{G}}(u,v,w)$. Figure 3.4 illustrates what the kernel looks like in the visibility domain for various w-planes. With increasing w comes an increase in the kernel size, and as explained in Section 3.1, a single visibility point will be convolved with this function and result in the point data changing, essentially projecting the point. For densely sampled uv-planes, the convolution kernels may cause the point data to affect neighbouring

visibility samples, so the more accurate the kernel, the better. Higher fidelity images can be further be achieved by anti-aliasing functions, zero padding and restricting the FoV so the small angle approximation does not deviate too much from the actual model - this is discussed in 3.5.1.

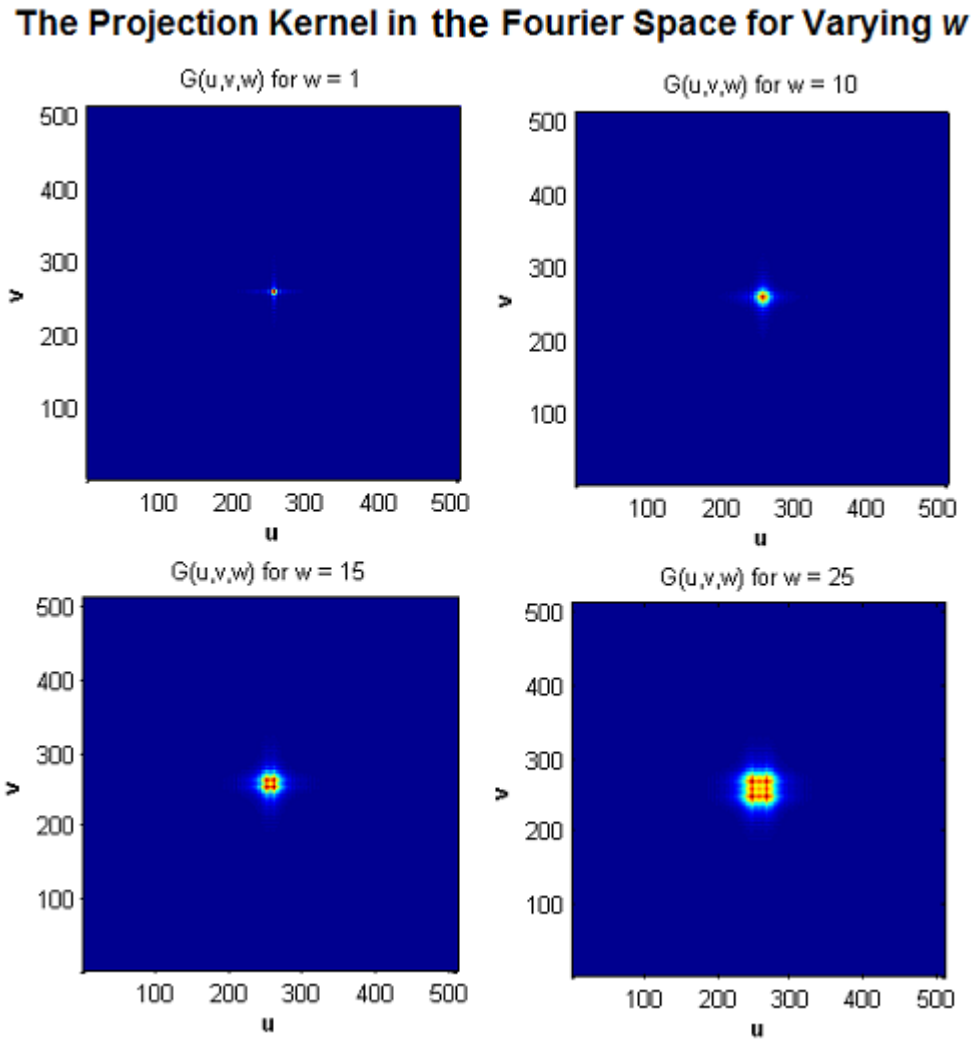


Figure 3.4: The absolute value of the projection kernel, $\hat{G}(u,v,w)$ as seen in the visibility domain by transforming the approximated G(l,m,w) with various w-planes. With increasing w, the required support for the kernel increases.

Fourier analysis of a slice in the u-column of the projection kernel, $\hat{\mathbf{G}}(u,v,w)$, shows that for increased w-planes, the kernel itself becomes more intricate as higher frequencies become more prominent in the data. For smaller w, the required support for the kernel is minimised, resulting in faster computation as opposed to a much larger w. This is because the kernel frequency components become spread out throughout the spectrum for increased w values, as elucidated in Figure 3.5. The same $\hat{\mathbf{G}}(u,v,w)$ data from Figure 3.4 were used to show the required number of bins to represent the projection kernel.

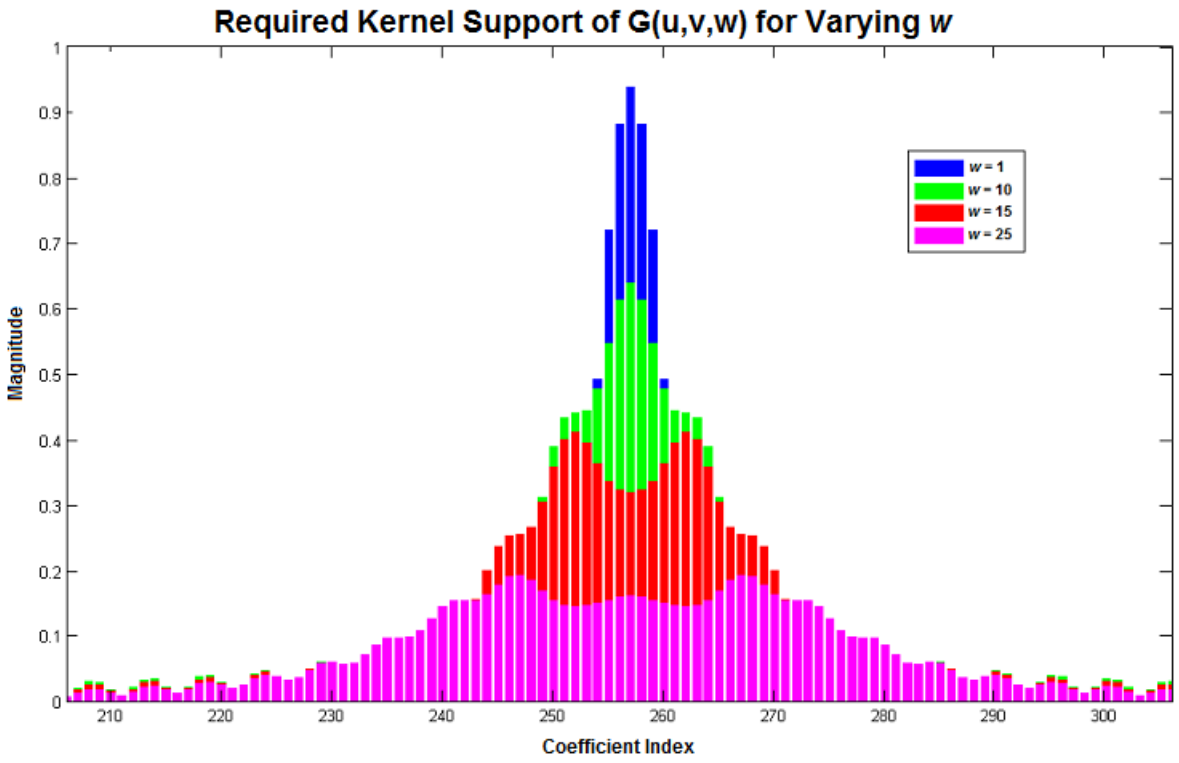


Figure 3.5: The Fourier coefficients required to model a column of $\hat{G}(u,v,w)$. With increasing w, the required kernel contains higher frequency content and the required support for the kernel increases.

With insight into the development and structure of the *w*-projection algorithm, astronomers can use it to remove non-coplanar baseline effects and obtain more precise images and visibility information. Image reconstruction, along with a comparative analysis of a different method, *w*-stacking, are the subject matter of Chapters 4 and 5.

3.4 Using a *w*-Key for Faster Computation

With each baseline having a *w* value associated with it, it can be quite computationally expensive to calculate this kernel for every point and then perform the two dimensional convolution, especially for long observations with synthesis arrays that have many antennas which will translate to thousands of baseline points. One method of reducing the expense of the projection is to generate a *w*-Key or *w*-Planes [3]. The process works as follows:

I. Determine the minimum and maximum values of *w* for the given baselines; w_{min} and w_{max} respectively.

II. Create a *w*-key by discretising the range into a user specified number of *w*-planes.

III. For each *w*-plane value in the key, calculate and store the convolution kernel $\hat{\mathbf{G}}(u,v,w)$.

IV. Compare each original baseline *w* value with the values contained within the key, and shift them to the closest value.

V. Once the baselines have been shifted, apply the correct convolution kernel associated with the key value.

The *w*-key is a user-specified discretised range of *w* values that are within the w_{min} and w_{max} boundaries. For more accurate data, the key must contain more *w*-values to describe the space; however, this will take longer to compute with the most accurate case being when each plane is associated with each unique *w*-value, but this may be impractical.

The original baseline *w*-values are then compared to these values in the *w*-key and are shifted to the closest value. Once this shift is performed, the corresponding convolution kernel is applied to the shifted baseline so the visibility samples can be grouped, making use of one convolution for multiple points, which have different *w*-values. This improves computational costs when compared to performing one convolution per visibility with a unique *w*. This process repeats for all *w*-key values until the entire projection is complete. It should be noted that the shifted data will differ from the data when using each exact *w*-value as the baselines are slightly altered when they are shifted. If the alteration is too great and affects the image quality too severely, then the key must be increased to contain more *w*-points.

3.5 Limitations and Considerations

In practice, numerically computing $\hat{\mathbf{G}}(u,v,w)$ relies upon the small angle approximation and improving the algorithm efficiency with a *w*-key, alters the original visibility data. Aliasing is also an issue as the sampled brightness and visibility functions are discrete. This can be prevented with the use of tapering functions and sufficient sampling, among other techniques. These limitations are discussed below, along with other considerations when imaging using *w*-projection.

3.5.1 The Small Angle Approximation

In the derivation of $\hat{\mathbf{G}}(u,v,w)$, the small angle approximation is used to simplify $\mathbf{G}(l,m,w)$, which becomes problematic for wide-field imaging as larger angle scopes for *l* and *m* are required. This affects the fidelity of the data as the greater the deviation from 'small' both *l* and *m* become, the greater the error. While there is no specified limit as to when the small angle approximation breaks down, the approximation states [3] that for small angles

$$-2\pi w(\sqrt{1-l^2-m^2}-1) \approx \pi w(l^2+m^2). \tag{3.5.1}$$

This result was simulated to gather some insight into when the approximation becomes invalid as seen in Figure 3.6. The two results are shown to start noticeable RMS deviation from the true angle, regardless of the *w*-value, when

$$\sqrt{l^2+m^2} \geq 0.4, \tag{3.5.2}$$

which, for a *w* of 10, results in an error of $\phi \approx 70°$. As stated by Humphreys [18], although the projection kernel is used for wide-field imaging, it must be calculated numerically and is only valid for a restricted FoV, else the small angle approximation conditions will break and the visibility points will not be accurately projected to the zero plane. This concurs with the trend in Figure 3.6, which suggests that *w*-projection be used for a limited wide-field. This deviation will cause the convolution kernel to project the *w*-values to a value that is close to, but not exactly zero. For wide-fields of view where *n* is large, the deviations in *w* from zero will result in a phase shift, when one is expecting it to be completely removed.

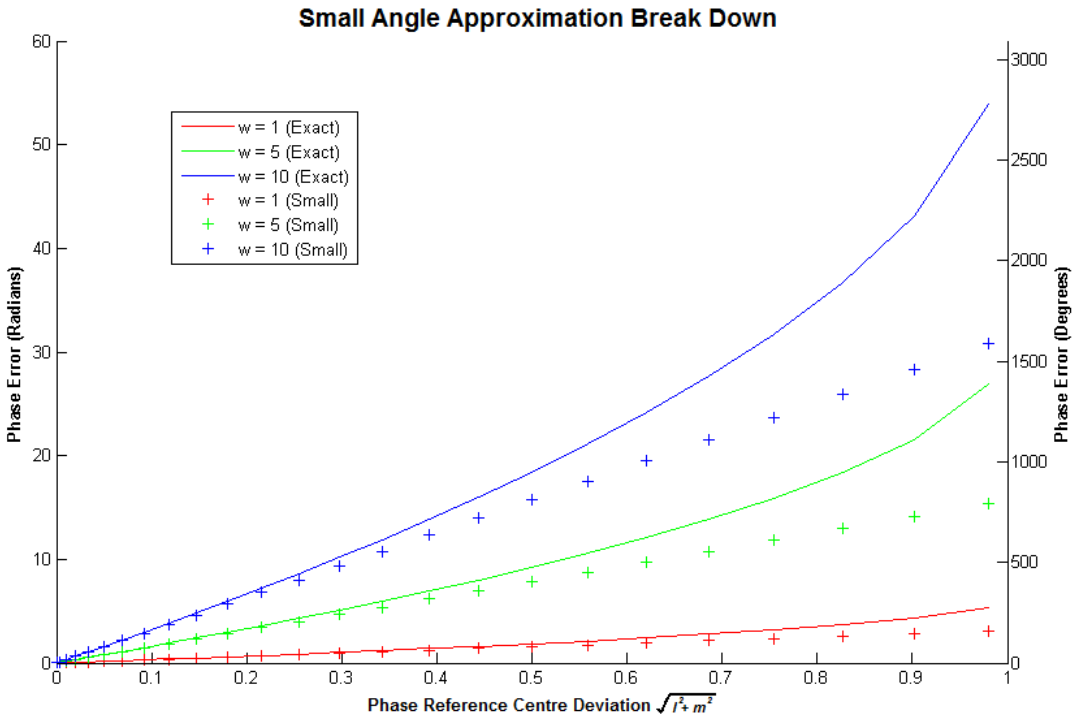


Figure 3.6: The phase deviation of the small angle approximated phase and the exact phase for various w-values. The data deviates from the start but there is noticeable difference by 0.4 for all simulated cases.

3.5.2 Restricting the Convolution Process and Zero Padding

A small consideration to improve the resolution of the convolution kernel, **G**(*l*,*m*,*w*), is to make use of zero padding. For the standard discrete 1D convolution of two vectors of length *A* and *B*, the length of the output vector after convolution becomes

$$Output\ Vector\ Length = A + (B - 1). \tag{3.5.3}$$

Similarly, this occurs in the 2D case where the convolution of two matrices, of size *A*x*C* and *B*x*D* will result in a larger output matrix of size

$$Output\ Matrix\ Size = [A + (B - 1)] \quad \times \quad [C + (D - 1)]. \tag{3.5.4}$$

When applying the convolution kernel to the visibility data, the output matrix would have increased in size, affecting the range of visibility points in the *uv*-plane. Only the central part of the convolution must be kept, by discarding all data points that lie outside of the initial plane or image dimensions, to ensure that the FoV and *uv*-plane remain true to their initial proportions and where the data generated is usable.

One can use zero padding to increase the resolution and ensure the visibility projections are accurate. The projection kernel in the image space is zero padded along both the *l* and *m* dimensions prior to the Fourier transformation. When the support of the kernel is small, it is still necessary to pad, as more precise data is always desirable. However, it becomes increasingly important when the projection influences a large portion of the *uv*-plane, as samples close to

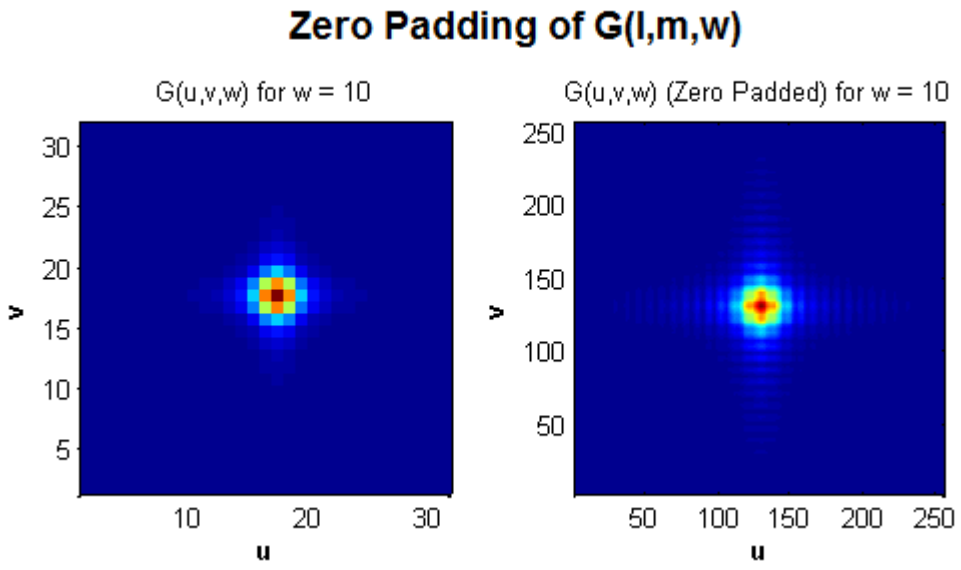


Figure 3.7: Left) $\hat{G}$(u,v,w) in the visibility domain, without G(l,m,w) being zero padded. Right) $\hat{G}$(u,v,w) in the visibility domain, with G(l,m,w) zero padded by to 256 showing a more resolved projection kernel.

each other will affect each other. The more accurate the projected point, the less error will develop when two sample points are spread in the *uv*-plane and interfere. Figure 3.7 shows the zero padding where **G**(*l*,*m*,*w*) is padded from 32 to 256 points along *l* and *m*.

3.5.3 Anti-Aliasing with Image Tapering

To limit the effect of aliasing, an anti-aliasing function, or tapering function, can be applied to the image plane. This is especially useful for those sources that are positioned close to the border of the defined FoV as it will prevent unwanted frequency components being generated in the image that are not actually there. The tapering function, **T**(*l*,*m*), is implemented in the image domain and applies to both the forward and reverse cases.

For the forward case, one applies it to the entire image before it is Fourier transformed. After transformation to Fourier space, the convolution is applied. For the reverse case, one must project to the 2D visibility plane using **G**(*l*,*m*,*w*), then perform the 2D inverse Fourier transform. Finally, **T**(*l*,*m*) must be divided out of the image to remove the tapering and obtain the dirty map [3].

Applying the tapering function, **T**(*l*,*m*), to (3.1.1), gives

$$\mathbf{V}(u,v,w) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\mathbf{T}(l,m)\,\mathbf{I}(l,m)}{\sqrt{1-l^2-m^2}} e^{-2\pi i[ul+vm+w(\sqrt{1-l^2-m^2}-1)]}\,\mathrm{d}l\,\mathrm{d}m, \tag{3.5.5}$$

which can be rearranged so that the convolution kernel becomes [3]

$$\mathbf{G}(l,m,w) = \mathbf{T}(l,m)\, e^{-2\pi i w(\sqrt{1-l^2-m^2}-1)} \approx \mathbf{T}(l,m)\, e^{\pi i[w(l^2+m^2]}. \tag{3.5.6}$$

The two considerations for tapering models were a 2D Tukey function or a 2D Welch function [20] as they are both functions that attenuate data values as one moves outwards from the origin. Note that Cornwell et al. [3] suggests using a tapering function that will also perform gridding of the *uv*-data as well as anti-aliasing; however, these functions were selected only to avoid aliasing components being folded into the image and they are not used as an gridding method. The Tukey function is a variation on a raised cosine model in combination with a rectangular window where α determines what portion of the window is the phase-shifted cosine. The Welch window is a parabolic function.

The objective is to reduce sidelobes, with a large primary lobe and to retain as much of the original data as possible when tapering the image, and to reduce the power received by unwanted sidelobes. For a 1D analysis, as shown in Figure 3.8, three windows were generated for a given FoV. The windows were then Fourier transformed to show the spectral profile, an indication of what would be happening in the visibility domain once the tapered image is transformed. Note that the spectral profile values for the Welch window were zoomed in on as plotting the data on the same decibel scale as the Tukey windows made the data unreadable as the oscillations are too rapid and fast decreasing. The window functions for one dimension are defined as

$$x(n_i) = 1 - \left(\frac{n_i - (\frac{N-1}{2})}{\frac{N-1}{n_i}} \right)^2 \tag{3.5.7}$$

for the Welch window and

$$x(n_i) = \begin{cases} 0.5\left[1 + cos\left(\pi(\frac{2n_i}{\alpha(N-1)} - 1)\right)\right] & 0 \leq n_i \leq \frac{\alpha(N-1)}{2} \\ 1 & \frac{\alpha(N-1)}{2} \leq n_i \leq (N-1)(1-\frac{\alpha}{2}) \\ 0.5\left[1 + cos\left(\pi(\frac{2n_i}{\alpha(N-1)} - \frac{2}{\alpha} + 1)\right)\right] & (N-1)(1-\frac{\alpha}{2}) \leq n_i \leq (N-1). \end{cases} \tag{3.5.8}$$

for the Tukey Window. N is the total window length, and n_i is the sample number. The windows were determined along each dimension of the image and then multiplied to give the 2D window functions.

The two Tukey windows that were analysed have a different α value, where α=0.1 relates to a near-rectangular window, and α=0.9 corresponds to a Hann window [21]. There is a trade-off between conserving more of the original image by having a flatter top for the majority of the window, but having greater sidelobes, the first being of magnitude -15dB, and a smaller main lobe with the smaller values of α. Greater sidelobes mean that more noise from other radiation sources may potentially be included in the data, obscuring weaker sources.

In contrast, using a more gradual decline with a greater α value, results in smaller sidelobes, with a magnitude of -24dB and a larger primary lobe, but more of the FoV is suppressed. The third is the Welch function which suppressed the sidelobes comparable to that of the higher α values in the Tukey function, roughly -23dB, however the primary lobe is much narrower and the image data will experience magnitude distortion at every point other than the phase reference center.

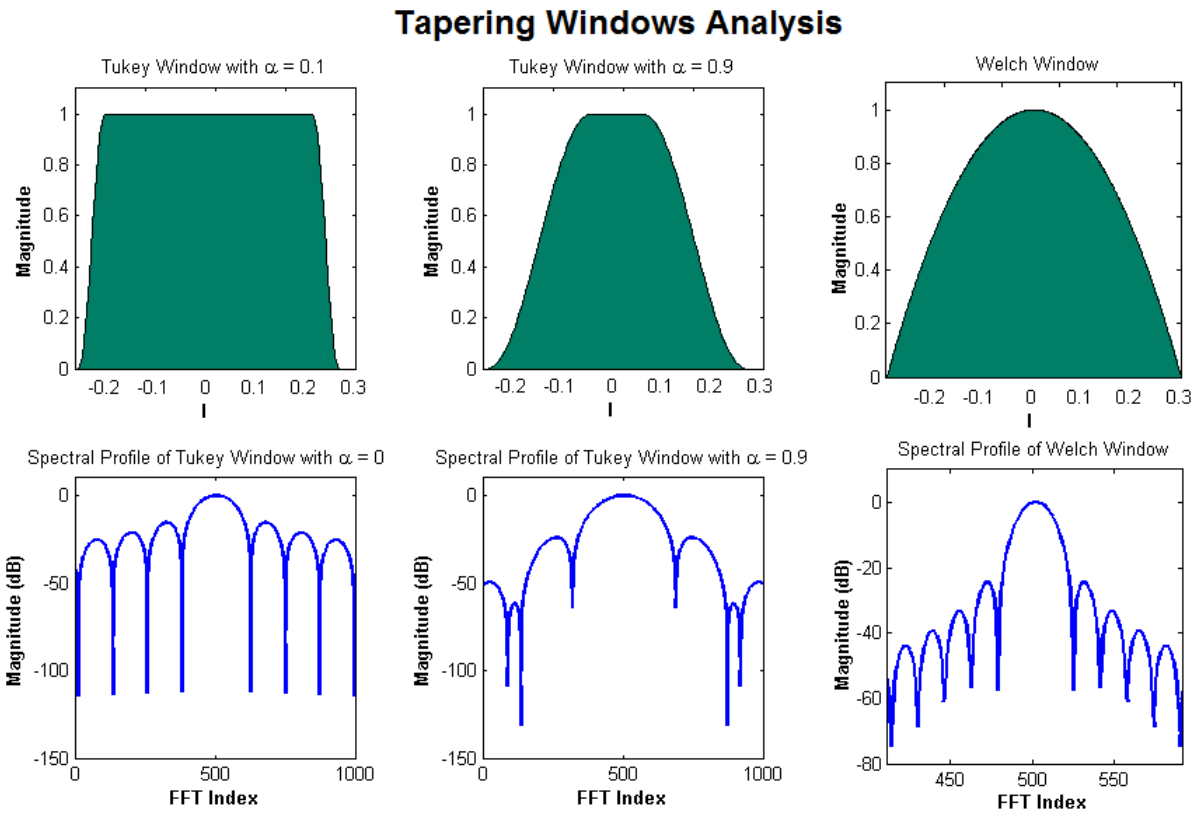


Figure 3.8: The top column shows the 1D window functions and how they will taper the values of the sky brightness distribution along a given axis; this is extended to 2D when applying the taper to an image. The bottom are their corresponding spectral profiles in decibels. Note that for clarity the scale of the vertical axis of the Welch window spectral profile has been enlarged.

With the above windows considered, a 2D Tukey function with α=0.25 was chosen as an appropriate tapering function, which preserved most of the image domain data but would have a beneficial effect on anti-aliasing by removing spectrally folded components. This was selected by analysis of the two α=0.1 and α=0.9 windows, where a window closer to that of the rectangular shape was preferred, such as α=0.25. In Figure 3.9, the 2D Tukey taper was generated and then employed in the image to act as the desired 2D low pass filter. The sky model was randomly generated to show the effect of the tapering for a FoV confined to $\sqrt{l^2 + m^2} \leq 0.3$; as this is slightly below 0.4, the limit where the small angle approximation was determined to break down.

As one moves closer to the boundary of the image, there is a inner boundary at approximately$\sqrt{l^2 + m^2} \geq 0.25$, where the sources become dampened, with those near the image boundary almost completely diminished by the tapering function. This creates a type of buffer for the aliasing artifacts that would otherwise appear in the image if the effects of aliasing became evident in the image. It is important to note if tapering has been implemented, it needs to be removed again when trying to get the most accurate models, be it visibility or sky brightness, as it alters the magnitudes of the sources in the FoV. A Tukey window of α=0.25 was implemented in Chapter 6.

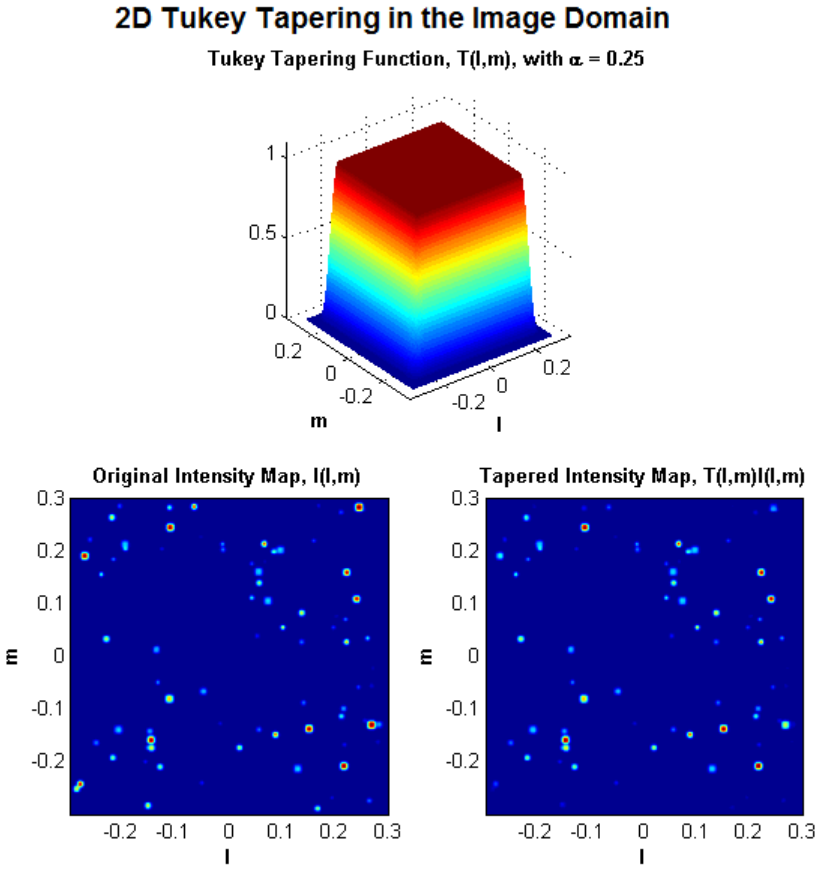


Figure 3.9: Top) A 2D Tukey Tapering function with α=0.25. Bottom Left) The original sky brightness model without any tapering. Bottom Right) The tapered image showing source magnitude distortion for $\sqrt{l^2 + m^2} \geq 0.25$.

3.6 Concluding *w*-Projection

To conclude this chapter, the *w*-projection algorithm is evidently useful in improving the quality of the brightness and visibility data, making modeling of the true brightness distribution more accurate. Obtaining the convolution kernel is non-trivial and the convolution itself may be computationally expensive and introduce aliasing into the data, if not implemented correctly. However, by projecting the *w*-component out of the equation in the Fourier space, the objective of removing the additional phase shift can be achieved. The small angle approximation based kernel discussed in this chapter will be used in Chapter 6 to obtain the reconstructed images.

An alternative means of correcting for the non-coplanar baseline effects is to apply a corrective phase shift in the image domain. This is the approach taken by *w*-stacking and forms the basis of the next chapter.

Chapter 4

The *w*-Stacking Algorithm

4.1 The Technique Of *w*-Stacking

A more recent method of accounting for the *w*-term due to non-coplanar arrays and wide field images is the *w*-stacking technique [18]. Note that few sources of literature are available which provide a detailed and comprehensive discussion on the method due to it being a very recent addition to the list of non-coplanar array imaging algorithms, and the majority of the content in this chapter is therefore based on [4]. Despite this, it has thus far proved to be a notable imaging technique that can yield high resolution images at orders of magnitudes faster than older methods, with full polarisation and DDE corrective capabilities. This chapter focuses on detailing the *w*-stacking procedure.

Developed for the same reasons as the *w*-projection, namely wide-field imaging and accounting for non-coplanar baseline effects, *w*-stacking differs from projection in the sense that it does not eliminate the third phase shift term, but rather corrects for it. To derive the result, (3.1.1) can be restated as

$$\mathbf{V}(u,v,w) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\mathbf{I}(l,m)}{\sqrt{1-l^2-m^2}} e^{-2\pi i(ul+vm)} e^{-2\pi iw(\sqrt{1-l^2-m^2}-1)}\,\mathrm{d}l\,\mathrm{d}m, \tag{4.1.1}$$

where it can be inverted, going from *uv* space to *lm* space, to yield the reverse case,

$$\frac{\mathbf{I}(l,m)}{\sqrt{1-l^2-m^2}} = e^{2\pi iw(\sqrt{1-l^2-m^2}-1)} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbf{V}(u,v,w) e^{2\pi i(ul+vm)}\,\mathrm{d}u\,\mathrm{d}v. \tag{4.1.2}$$

With visibility being a 3D function of variables *u*, *v* and *w*, the integral in (4.1.2) must be adjusted to incorporate the *w* spatial component. Integration on both sides across the full range of *w* gives

$$\frac{\mathbf{I}(l,m)(w_{max}-w_{min})}{\sqrt{1-l^2-m^2}} = \int_{w_{min}}^{w_{max}} e^{2\pi iw(\sqrt{1-l^2-m^2}-1)} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbf{V}(u,v,w) e^{2\pi i(ul+vm)}\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w. \tag{4.1.3}$$

The variables w_{min} and w_{max} are, respectively, the minimum and maximum values for *w* in wavelengths. On the left hand side of (4.1.3), there is a secondary scaling factor that is

introduced due to the extra integral. On the right hand side, the exponential components are separable, making the entire integral comprise of two parts; a 2D inverse Fourier transform and a complex exponential integration over w.

Discretisation of the integration variables modifies the inverse Fourier part to become an inverse FFT, and the w integration to be a summation [4],

$$\frac{\mathbf{I}(l,m)(w_{max}-w_{min})}{\sqrt{1-l^2-m^2}} = \sum_{a=w_{min}}^{w_{max}} e^{2\pi i w_a(\sqrt{1-l^2-m^2}-1)}\ \mathscr{F}^{-1}\{\mathbf{V}(u,v,w_a)\}, \tag{4.1.4}$$

where $\mathscr{F}^{-1}\{\mathbf{V}(u,v,w_a)\}$ is the inverse FFT for gridded visibilities on a specific discretised w_a-plane. It is this result of (4.1.4) that shows how w-stacking works. Having sampled visibility data in Fourier space, the image can be reconstructed by the following procedure:

I. Create a w-key, different to the w-projection w-key, by discretising the range into a user specified number of w-points.

II. Grid $\mathbf{V}(u,v,w)$ samples corresponding to each w-point onto a separate uniform grid.

III. Convert each of the gridded visibility planes to the image domain with the inverse FFT.

IV. Apply the phase correction relating to the w-value for each image.

V. Scale and sum all the images to obtain the final brightness distribution model.

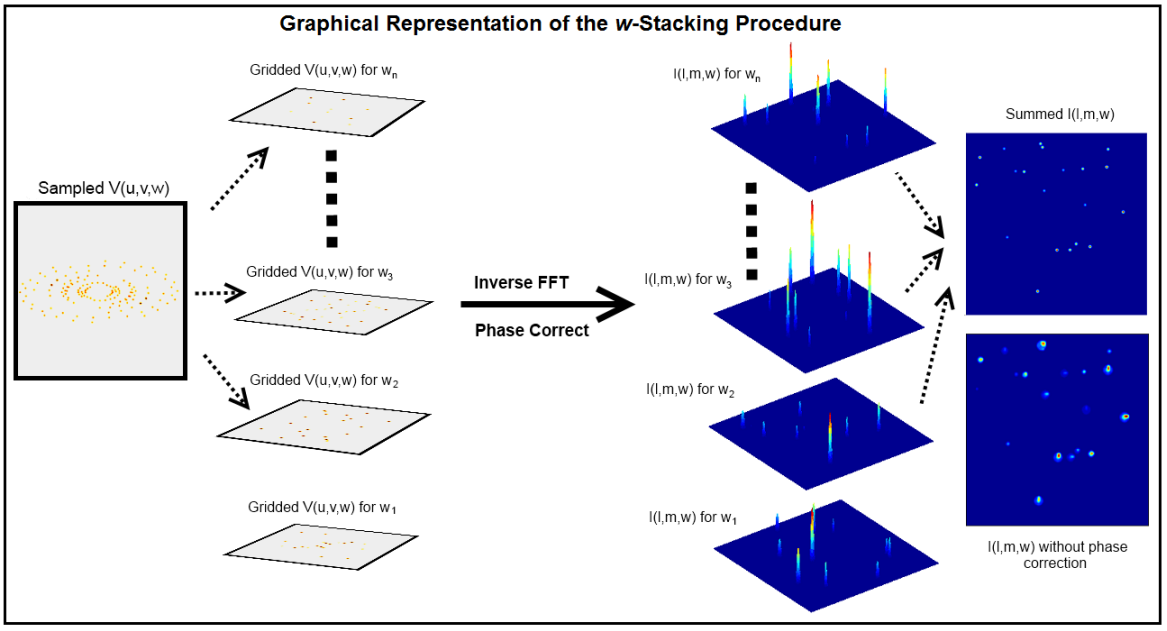


Figure 4.1: Going from left to right, the visibilities are separated onto different girds, based on their w value. Each of these undergo an inverse FFT and the points are phase corrected. Finally, to produce the final image, the individual planes are summed.

Stacking takes the visibilities for a given w-plane and calculates a stack of images from each of these planes. It rectifies the image data by phase shifting the entire inverse transformed plane by $2\pi iw(\sqrt{1-l^2-m^2}-1)$, specific to the w-value associated with the given plane. Then, by overlaying all the image planes, the final image can be reconstructed that has been phase corrected. This is graphically demonstrated in Figure 4.1.

The stacking algorithm can also be reversed for visibility prediction:

I. Create copies of the image depending on the number of w-points.

II. Apply the phase correction relating to the w-value for each individual image layer.

III. Individually FFT each layer in the image domain to obtain the visibility, pertaining to the w value.

IV. Sample the desired visibility plane corresponding to the required visibility point.

4.2 Limitations and Benefits of w-Stacking

As with most imaging methods, there are advantages and disadvantages which make one technique better suited than the other, depending on the application. Stacking is appealing as each image layer is independent from another, until the summation step, making it ideal for exploiting parallelism. Each gridded visibility plane, specific to a discretised w, can be inverse transformed and phase corrected.

An additional consideration is the number of layers to use. Discretising w into a user-defined range, as discussed in section 3.4 regarding the w-key, can save computation at the cost of image fidelity. Note that the w-key discussed in this section is different to that of the w-projection w-key as there is an empirically determined formula to calculate the number of layers for stacking, which is discussed below. Using too few planes results in aliasing artifacts and ghost sources [4], as the visibility plane can become undersampled and not meet the Nyquist criteria. Too many planes will increase the amount of processing, with decreasing computational improvement as the number of layers approach the number of unique w points. Through empirical means, it is suggested that the number of planes for stacking be

$$N_{layers} \gg 2\pi(w_{max}-w_{min})\,(1-\sqrt{1-l_{max}^2-m_{max}^2}), \tag{4.2.1}$$

where the separation between the w-values is uniform and l_{max} and m_{max} are the maximum l and m values, respectively [4]. For the case of undersampling w and not stacking enough layers, the final image will contain aliasing artifacts that will appear as ghost sources. This issue is shown in Figure 4.2, where the minimum layers for the simulated FoV was determined to be 178. Expression (4.2.1) states that the number of layers must be greater than this. The undersampling clearly introduces aliasing components.

For cases where the brightness of the sources in the image are known, a threshold could be set in the CLEANing procedure to remove aliasing components, a topic of discussion in Chapter 5. If the threshold were set to be slightly less than the weakest source, CLEANing would remove

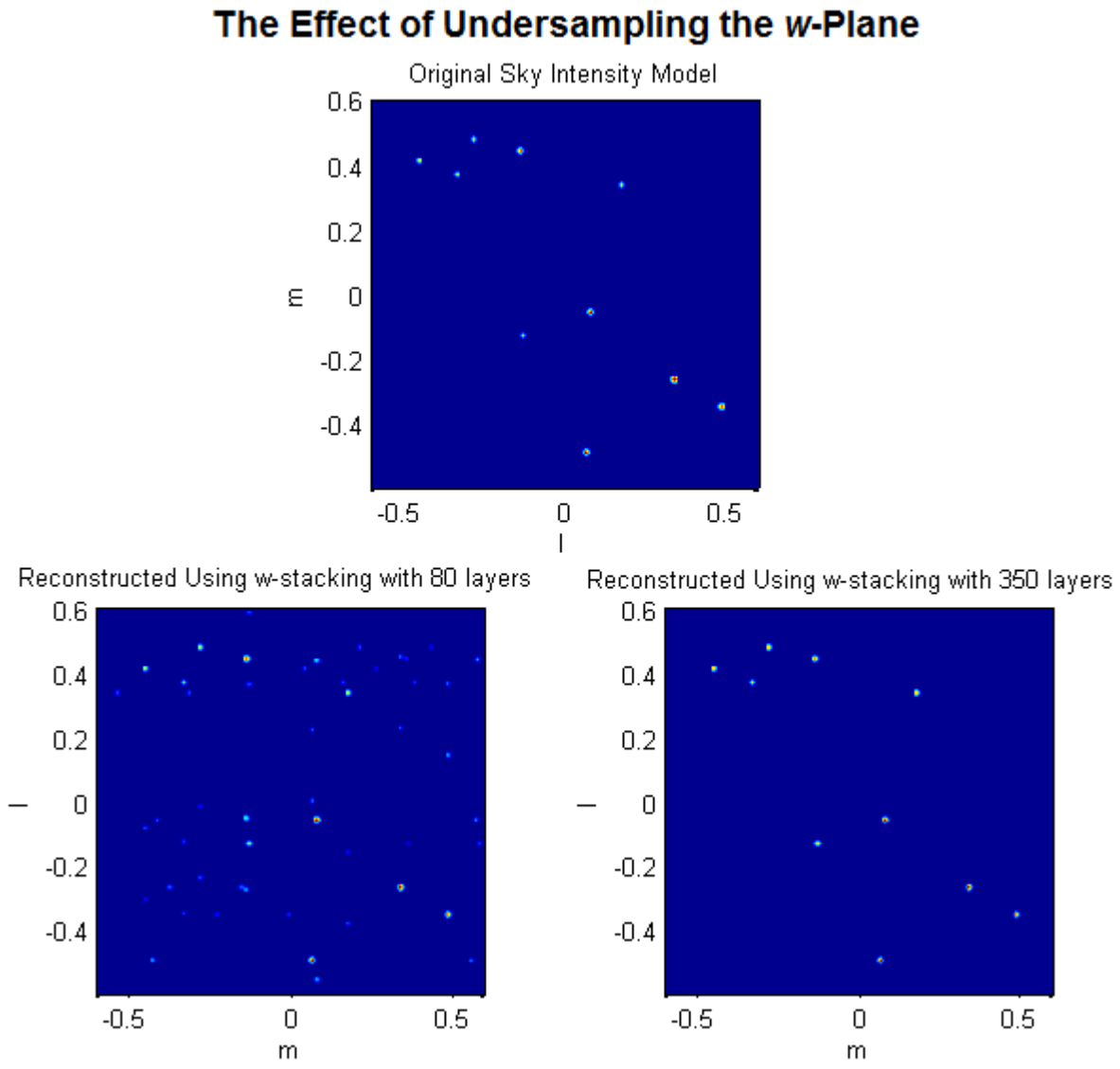


Figure 4.2: Top) The original image reconstructed using the full visibility model. Bottom Left) The image reconstructed by stacking 80 layers. The minimum layers required should be much greater than 178. Extra 'sources' are shown in the image as a result of this undersampling. Bottom Right) The image reconstructed by stacking 350 layers. The aliasing components have been removed, and the original sources are identifiable.

any sources from the image that do not meet the bounds of the threshold. This method is not ideal though, as images that have a large range of source brightness magnitudes, will have aliasing components. If the aliased components that are folded into the image are larger than the magnitudes of any weaker sources, they could potentially result in being mistaken for actual sources, which will appear in the final image. While this method could allow one to use fewer w-layers, as there is a trade-off between how many layers to use before aliasing data starts to become ambiguous to that of actual source data, it could be useful if the image contains a small range of source brightness magnitudes.

For both *w*-stacking and *w*-projection, aliasing components can be identified by using different cell sizes for gridding. For a fixed FoV, if the *uv*-data is gridded onto a grid and an intensity map is produced from this data, then using the same *uv*-data and gridding onto a slightly finer grid to produce another intensity map where the true sources will appear at the same locations. The aliased components however, will appear to move about the FoV depending on how the *uv*-data is gridded. Doing this makes distinguishing aliased components from actual sources possible; however, in practice, gridding is often the most dominating computational step so producing the same intensity map using various gridding techniques is expensive. This is potentially a handy approach to identifying which sources exist or are a result of spectral folding [4].

4.3 In Summary of *w*-Stacking

In summary of the chapter, the stacking algorithm is an alternative approach to accounting for the non-coplanar baseline effects. The number of image layers that are phase corrected directly influences the image data fidelity as the layers provide a trade-off between image accuracy and computational efficiency. The image can become more prone to aliasing artifacts for too few layers, while high resolution and large images can make the algorithm computationally expensive if too many layers are needing to be phase corrected. In spite of these drawbacks, it is evident that stacking can effectively shift the data to be coplanar. For the image reconstruction in Chapter 6, the exact model, as illustrated by (4.1.4), will be used where a 2D FFT is performed on each image layer.

Unavoidably, both stacking and projection cannot remove the effects of the PSF and incomplete sampling from the images. These effects can be removed by using a popular algorithm called CLEAN — the topic of the next chapter.

Chapter 5

Deconvolution Using the CLEAN Algorithm

5.1 Motivation for CLEANing

One of the fundamental algorithms used in producing a more accurate model of the true sky brightness is the CLEAN algorithm. The algorithm was developed by Högbom in the 1970's as a non-linear method to replace the two-dimensional deconvolution of the dirty beam from the observed dirty map data [5].

Consider the two-dimensional sampling function, $\mathbf{S}(u,v)$, that is applied to the true sky visibility data, $\mathbf{V}(u,v)$. This sampling function, or the transfer function of the synthesis array, is dependent on the physical layout of the antenna system and observation parameters. The sampling is discrete so $\mathbf{S}(u,v)$ is essentially the accumulative affect of the various snapshots of the true visibility throughout the course of the observation. The less sparse $\mathbf{S}(u,v)$ is, the more coverage one will achieve in the uv-plane, in turn recording more of the visibility data.

$\mathbf{S}(u,v)$ can be seen as containing a number of zero terms which correspond to the sampling gaps in the uv-plane. Applying the inverse Fourier relation to $\mathbf{S}(u,v)$ results in the dirty beam, $\mathbf{B}(l,m)$, where these zero terms generate unwanted sidelobes and distortions. Using convolution theory, when applying the two-dimensional Fourier relation between the sampled visibilities, $\mathbf{V}(u,v)\mathbf{S}(u,v)$, it corresponds to a convolution in the image domain between the true sky intensity, $\mathbf{I}(l,m)$ and the dirty beam, $\mathbf{B}(l,m)$. The result of this convolution is the dirty map, $\mathbf{D}(l,m)$, given by

$$\mathbf{V}(u,v)\,\mathbf{S}(u,v) \xrightarrow{\mathscr{F}^{-1}} \mathbf{I}(l,m) * \mathbf{B}(l,m) = \mathbf{D}(l,m). \tag{5.1.1}$$

It is now clear that the ideal sampling function would be one containing no zeros, as this would fill the uv-plane entirely, and the corresponding dirty beam would be in the shape of a 2D Dirac delta function,

$$\mathbf{S}(u,v) = 1 \xrightarrow{\mathscr{F}^{-1}} \mathbf{B}(l,m) = \delta(l,m). \tag{5.1.2}$$

Without any aliasing caused by unfilled sampling terms in $\mathbf{S}(u,v)$, it would mean that the dirty map, $\mathbf{D}(l,m)$, is actually the true sky intensity, $\mathbf{I}(l,m)$, as shown when substituting (5.1.1) into (5.1.2),

$$\mathbf{V}(u,v)\,\mathbf{S}(u,v) \xrightarrow{\mathscr{F}^{-1}} \mathbf{I}(l,m) * \delta(l,m) = \mathbf{I}(l,m) = \mathbf{D}(l,m). \tag{5.1.3}$$

This ideal case is only a theoretical construct and is physically unrealisable in practice. Full *uv*-coverage cannot be achieved due to physical limitations. These limitations include the layout of the antenna network, the physical properties of the antennas, variations in baselines as a result of non-coplanar effects and observation parameters; all of which determine the shape of $\mathbf{S}(u,v)$.

The CLEAN algorithm was developed in order to correct distortions introduced by convolution with $\mathbf{B}(l,m)$, producing the dirty map [22]. $\mathbf{B}(l,m)$ is the PSF — an indication of the amount a single point will be spread or distorted in the image domain due to sparse sampling. The observed sky intensity distribution is the true sky intensity that has been convolved with the dirty beam, among other observation errors and noise, to create the the dirty map. CLEAN aims to remove the effects of the dirty beam to create a much more accurate representation of the true visibility data. The reasons why normal deconvolution cannot be applied is better illustrated in the *uv*-domain,

$$\mathbf{V}(u,v)\,[\mathbf{S}(u,v)\mathbf{S}^{-1}(u,v)] \xrightarrow{\mathscr{F}^{-1}} \mathbf{I}(l,m) * [\mathbf{B}(l,m) * \mathbf{B}^{-1}(l,m)] \tag{5.1.4}$$

$$\mathbf{V}(u,v)\mathbf{I}_d \xrightarrow{\mathscr{F}^{-1}} \mathbf{I}(l,m) * \mathbf{D}_n = \mathbf{I}(l,m), \tag{5.1.5}$$

where the noise has been omitted for the sake of simplicity. Assume $\mathbf{S}(u,v)$ contains no zero elements making it invertible for mathematical purposes, to illustrate the deconvolution process, $\mathbf{D}_\mathrm{n}$ is the Dirac Identity [23] which is similar to the identity matrix but applicable to convolution with matrices and $\mathbf{I}_d$ is the Identity Matrix.

While seemingly appealing, equations (5.1.4) and (5.1.5) are spurious representations of the actual model. Applying the Fourier Transform to the dirty map, $\mathbf{D}(l,\ m)$ will result in the sampled visibility function, $\mathbf{V}(u,v)\mathbf{S}(u,v)$. The equivalent operation to deconvolution in the image domain is inverse multiplication, or element-wise 'division' per se, in the *uv*-domain.

With the afore-mentioned in mind, obtaining $\mathbf{I}(l,m)$ requires a different method of deconvolution as $\mathbf{S}(u,v)$ and $\mathbf{B}(l,m)$ comprise of many zero terms, making these matrices non-invertible or non-singular. Additionally, one cannot discard the noise in the model as the introduction of noise into the system makes the deconvolution process a non-linear operation. If a 2D deconvolution could be performed to separate $\mathbf{B}(l,m)$ from $\mathbf{I}(l,m)$, the solution would not be unique [24]. This is why iterative subtraction must be performed. This is the motivation behind the development of the CLEAN algorithm [22].

5.2 The CLEANing Procedure

CLEAN is an algorithm that iteratively subtracts the effects of brighter sources and the dirty beam from the dirty map until other sources become more visible — this being analogous to deconvolution. Removing these effects allow a more accurate representation of the true sky intensity to be obtained. The original CLEAN algorithm works under the assumption that $\mathbf{I}(l,m)$ contains no extended source emissions and the sources are all considered to be time-invariant [5]. Other variations of CLEAN have been developed for more sophisticated deconvolution techniques as, with the growing advances in imaging and instrument capabilities, the original CLEAN algorithm assumptions introduce limitations to recovering high fidelity images [25].

The original model, based on the assumption that the dirty map contains only point sources, models the dirty map as a combination of scaled and translated dirty beams. These dirty beams are translated to the locations of the sources in the brightness distribution and scaled by the source magnitude. It is the accumulation of the dirty beams that produces the final dirty map that is observed. The objective is to iteratively translate and subtract these dirty beams from the dirty map until a acceptable intensity map is obtained; mostly free of the detrimental effects of the PSF.

The algorithm works as follows [5]:

I. Search through the dirty map, $\mathbf{D}(l,m)$, until the maximum absolute energy peak, $|\mathbf{I}_p|$ is found.

II. Once $|\mathbf{I}_p|$ is located, record both the location and magnitude of the peak. Subtract the dirty beam, which has been translated to the location of $|\mathbf{I}_p|$, from the entire dirty map, which has been scaled by a loop gain, γ. The resulting dirty map is the residual map, $\mathbf{R}_x(l,m)$, which is normalised to $\gamma|\mathbf{I}_p|$.

III. Repeat Step II, replacing the dirty map by the updated residual map, $\mathbf{R}_{x+1}(l,m)$, each time. Terminate this process when $|\mathbf{I}_p|$ is sufficiently small.

IV. Use the recorded locations of the identified peaks in Step II to generate a cleaned map; replacing all removed components with appropriately scaled and translated clean beams.

This process describes the entire Högbom CLEAN model. The residual maps are updated as

$$\mathbf{R}_{x+1}(l,m) = \mathbf{R}_x(l,m) - \hat{\mathbf{B}}(l,m), \tag{5.2.1}$$

where $\hat{\mathbf{B}}(l,m)$ is the translated and scaled dirty beam and $\mathbf{R}_x(l,m)$ is the residual map at iteration x. Step II is repeated until the energy in the current residual map is sufficiently small or, a predetermined maximum number of iterations is reached to limit inefficient computation or, if other stopping criteria that may be specified are satisfied — this is discussed in detail in Section 5.3.1.

The loop gain, γ, modifies the actual residual map scaling to "peel" off layers of data at a time from the image, rather than do massive subtractions. This is analogous to subtracting scaled variations of the dirty beam. Large beam subtractions will potentially cause larger errors which

could result in the dirty map being contaminated with more noise from sidelobes. In practice, it is impossible to completely remove this error, unless the dirty map contains no noise and the cumulative subtractions of $\mathbf{B}(l,m)$ result in a perfect fit to the original observed source.

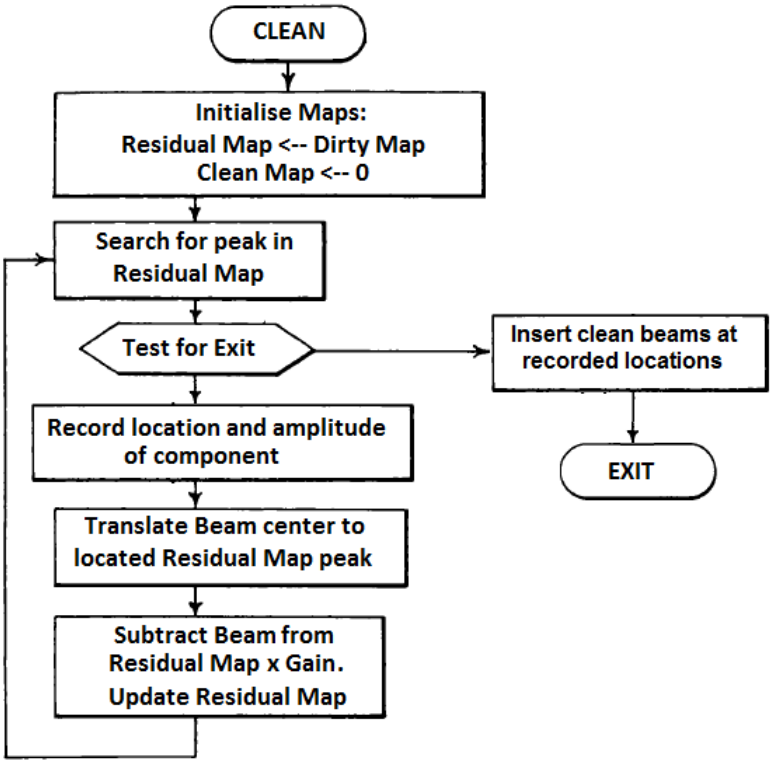


Figure 5.1: A flow chart of the CLEAN algorithm. Graphics adapted from [25]

There are other CLEAN algorithm variations which improve the original CLEANing procedure such as the Clark variation [26]; most of which consist of a major and minor cycle. The major and minor cycles vary depending on the CLEAN model used but, the general procedure for the minor cycle consists of the CLEAN process described and shown in Figure 5.1. The major cycle involves converting the dirty beam or residual maps to the visibility domain in order to subtract the unwanted PSF data in the Fourier space, rather than working in the image domain. This domain transformation may cause additional aliasing. The combination of these errors may introduce ghost artifacts into the clean map, which is discussed in 5.3.5.

Once the stopping criteria of the algorithm has been reached, the final step is to insert a clean beam at all the recorded peak energy locations. The inserted clean beams have equal magnitude to that of the original source intensity at the given location. This intensity map, containing the inserted clean beams at the original source locations, is the clean map. This clean map is a more desirable representation of the true sky brightness function as it does not include the detrimental effects of a dirty beam such as unwanted sidelobes; the sidelobes are dependent on the array configuration and *uv*-coverage. The clean map is only accurate if $\mathbf{I}(l,m)$ contains only point sources as the clean beam is selected to be Gaussian in shape. This gives the desirable result of having low sidelobes for minimal source interference and noise [25].

Once a CLEAN iteration has been completed, the entire process is repeated for the new, updated residual map, as the beam subtraction may not completely remove the entire source. A spike may still be located at a subtracted peak, just of lower intensity, as only a fraction of it was removed. Sources in the updated residual map may become more prominent after $\hat{\mathbf{B}}(l,m)$ is subtracted, as the effect of brighter sources could mask weaker sources. This is shown in Figure 5.2 where the CLEAN algorithm was used in CASA on the Circinus X-1 binary system. The measurement set data was obtained from the CASA archives. The supernova radiation dominates the map and masks other sources contained in the image but, after a single iteration, other sources become more prominent as seen on the left. Multi-scale CLEAN was used for this simulation as the measurement set contained extended emissions.

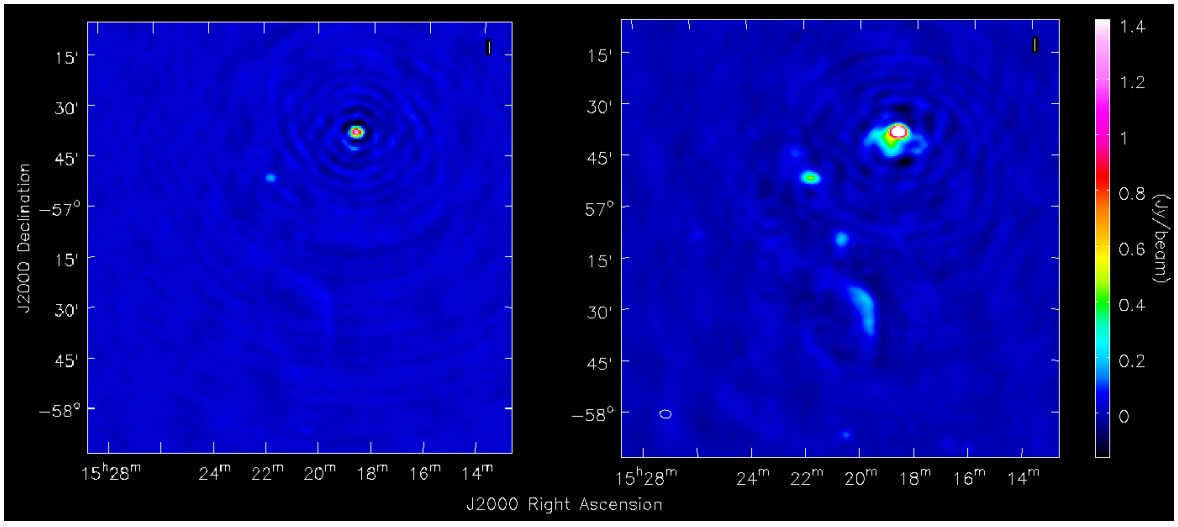


Figure 5.2: Left) A Dirty Map image of the binary star system, Circinus X-1. All sources are masked by the dominant supernova and its sidelobes, located in the top right portion of the map. Right) After one iteration of CLEAN, more sources, including a jet of gas, become more prominent as the supernova is CLEANed. The image was CLEANed using CASA and the measurement set data was obtained from CASA archives.

5.3 Finer Details: Problems and Improvements

5.3.1 Stopping Criteria

The original CLEAN stopping criteria computes the peak absolute energy contained in the residual map and the absolute average energy [5],

$$\epsilon = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b}(|\mathbf{R}(l,m)_{i,j}|)}{ab} \qquad \text{where } a = \text{rows}, b = \text{columns}, \tag{5.3.1}$$

which is calculated by dividing the sum of the total absolute energy by the number of elements as seen in Equation (6.2.2).

The algorithm terminates if the maximum absolute peak located in the residual map, exceeds the average absolute energy in the map by some user defined factor. This threshold factor is calculated as

$$\vartheta = \frac{|\mathbf{R}(l,m)_{max}|}{\epsilon}, \tag{5.3.2}$$

and Högbom [5] proposes that CLEAN stop when $\vartheta \leq 4$. Too low a factor would possibly allow for more sources to be identified at the risk of over-CLEANing, whereas too high a factor will result in stopping the CLEANing process too early and not resolving some of the sources.

This stopping criteria is simple; however, if the dirty image contains primarily weak sources but some dominating strong sources, the total average energy will increase due to the large source magnitudes. This will make the threshold factor higher. Essentially the large sources 'skew' the threshold factor value and can be seen as outliers, when the map contains primarily much weaker sources. This may be problematic as the strong sources will be removed from the map, which will greatly decrease the energy in the map. The energy may be below the high threshold factor so weaker sources that lie below the threshold will not be identified. This will result in an under-CLEANed map.

The resolution of dirty maps and the structures of the sources contained in the map, call for different stopping criteria depending on the type of data. In keeping with generality, an extra threshold check can be put into place that use basic probability theory. The 3σ Rule can be applied as the noise in the system is presumed to be of a normal Gaussian distribution. The 3σ Rule states that most data points will lie within 3 standard deviations of the mean [27]. The residual map has been satisfactorily CLEANed and the cycles will terminate when

$$|\mathbf{R}_x(l,m)|_{max} \leq \mu + 3\sigma, \tag{5.3.3}$$

where μ is the mean and σ is the standard deviation of the energy contained in the residual map. Note that the mean and standard deviation are calculated for a two-dimensional matrix, not a vector.

Equation (5.3.3) states that the process must terminate when the maximum value of the current residual map is within three times the standard deviation from the mean; where μ will converge towards the noise level as peaks are removed from the map. If the map still contains a large peak, it will not be within the bounds of (5.3.3), and the CLEANing process will continue until all sources have been removed, leaving only noise and weak sources behind.

5.3.2 Point Source Approximation

CLEAN works on the assumption that all sources located in the dirty map are point sources, with structures that are independent of time. Problems arise when a dirty map contains sources that are complex in shape, such as large sources that are irregular in structure. The exact definition of a large source is not explicitly stated in literature, but one can consider the source to be large when the point occupies more than one pixel or grid point in the image, resulting in incorrect image reconstruction.

Consider a jet of cosmic radiation as shown in the left of Figure 5.3. In the CLEANing process the peak of the source is found; however, the surrounding intensities are all of similar magnitude. The algorithm has only identified the maximum peak. An approximation for the dirty beam is now calculated, convolved with the peak that has been mis-identified as a point source, and then subtracted from the residual map. Only a section of the jet will be removed.

The sub-sequent cycles will continue to find various 'point-sources' until the entire jet has been CLEANed away. However, the cleaned map will not reflect this complex jet structure, but rather show a concentration of point sources of various amplitudes. Tracing a perimeter along the group of point sources may closely resemble the outline of the jet shape, however the finer details of the jet structure are lost as it is represented by a combination of point sources. The assumption that the sources are all point sources makes distinguishing between a cluster of point sources and a complex structured source difficult, as shown in Figure 5.3.

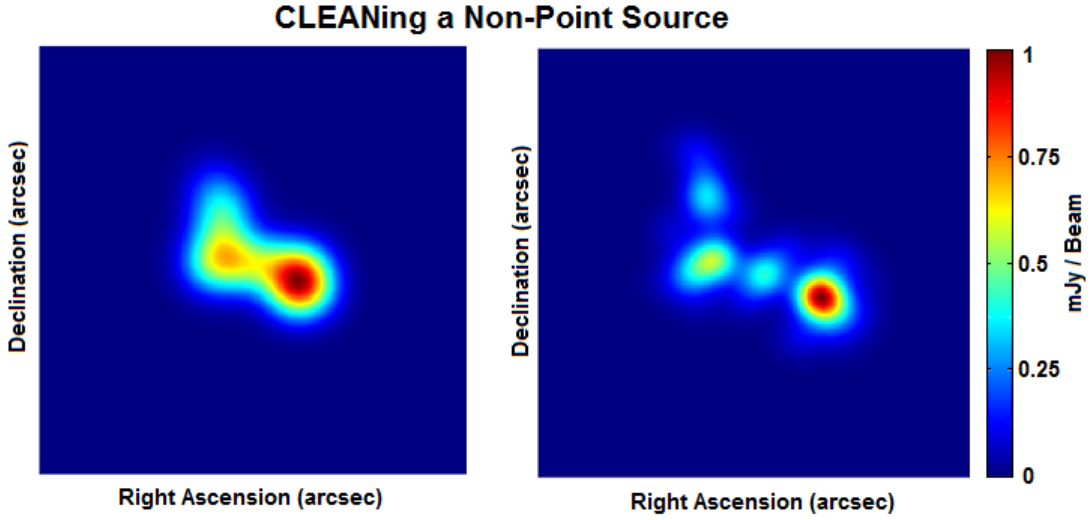


Figure 5.3: Left) A simulated model of source of complex structure. Right) The effect of CLEANing a source that is not a point source. CLEAN reconstructs the jet-like structure using a combination of point sources.

One improvement that accounts for extended structures, is the concept of contour trimming which is utilised by the Multi-Scale CLEAN variant [28]. The PSF was originally subtracted from the gain-modified residual map. The multi-scale model recognises that the source under question may not be a point source and adjusts the dirty beam accordingly using contour trimming. Once a peak is located, as opposed to taking the individual peak value, contour trimming assesses the surrounding data and groups all brightness values that are above the contour trim together [25]. The contour trim value is determined by the maximum peak value and the trim parameter,

$$\mathbf{T}_r = t \times \mathbf{I}_p, \tag{5.3.4}$$

where $\mathbf{T}_r$ indicates the contour value for which any data points that are above this value will be grouped together, and t is the trim parameter set by the user.

The algorithm identifies various 'islands' of brightness, as opposed to single point components and adjusts the dirty beam width and height to subtract a more fitted dirty beam, essentially viewing an extended source as a large component, rather than a group of point source components. This helps to identify extended source structures and is essentially a method of flagging the source.

The issue with contour trimming is that it requires another Fourier transform to the visibility domain, which is performed on the group of components before it is converted back to the image domain and subtracted from the residual map [29]. This additional transform operation is required in order to get the appropriate scaling for the dirty beam, as a convolution with a point source will result in a beam that has the peak value of the point brightness, whereas a convolution with a group of components will have a much more accumulative affect; the dirty beam approximation will be weighted by each component in the contour. The dirty beam must be adjusted so that it takes the value of the maximum peak in the residual map or else the beam will be too large, even after the loop gain has been incorporated. As seen in Figure 5.4, the structure for the same jet-like source seen in Figure 5.3 is more identifiable, as groups of components were identified. Scaling the dirty beam accordingly has not been implemented; however, the overall structure of the jet is more identifiable.

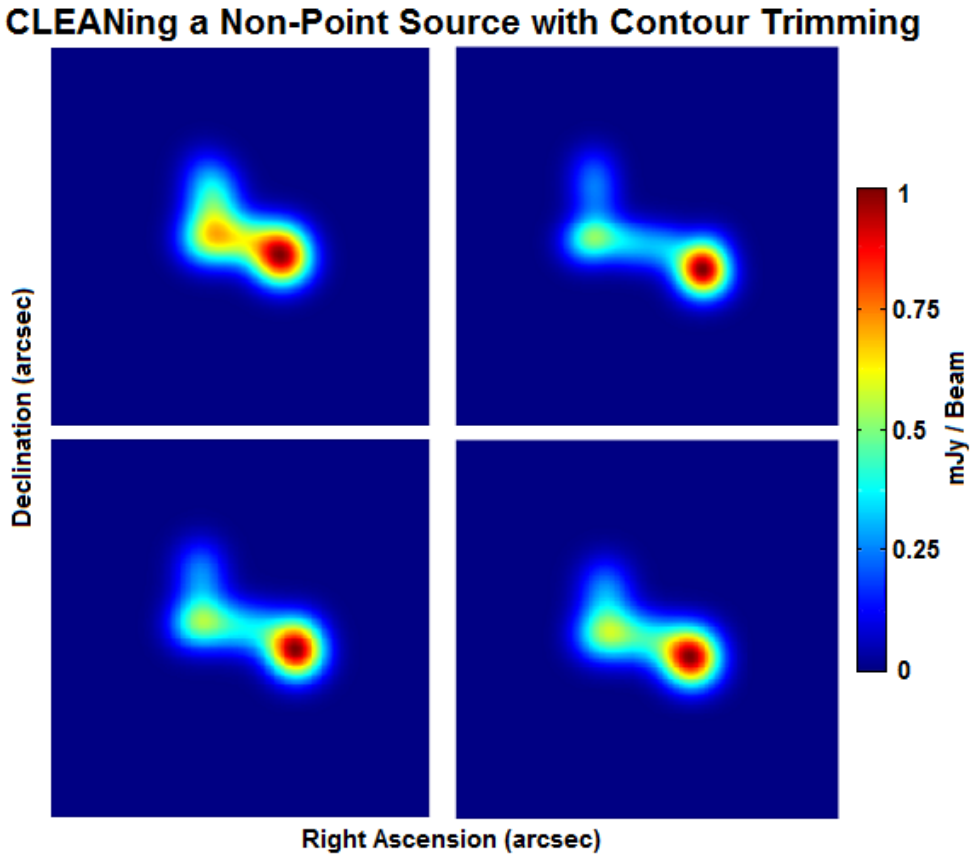


Figure 5.4: The effect of CLEANing an extended emission source using contour trimming, making the structure of the source more evident. Top Left) Simulated model of source of complex structure. Top Right) CLEAN reconstruction of jet-like structure using a more fitted dirty beam with contour trim equal to 3% of peak. Bottom Left) Contour trim equal to 6% of peak. Bottom Right) Contour trim equal to 9% of peak.

Changing the contour trim parameter value affects the image. If *t* is high, then only immediate values around the identified peak will be used to determine the dirty beam that is subtracted. This will result in a large source being reconstructed by many point sources, as previously discussed. A small *t* will group more components together, resulting in a large dirty beam subtraction, in terms of width. For *t* being too small, the dirty beam will be too wide, removing too much of the source data and distorting the actual shape of the source. As with much of the CLEANing procedure, the fidelity of the image will always be questionable, so the ideal trim contour is up to the discretion of the user but can be used to give indication that an image contains extended source structures.

5.3.3 Loop Gain

The loop gain, γ, is dependent upon the maximum value of the residual map that is being CLEANed. Each time a CLEAN cycle finishes, the residual map is updated and the loop gain is adjusted. The loop gain ensures that not too much of the dirty beam is removed at once to avoid over-CLEANing. The ideal case would be to have an infinitesimally small loop gain, skimming off tiny amounts off of the peak until the residual map energy becomes zero. However, a tiny loop gain implies a computationally inefficient number of iterations.

The need for an adjustable loop gain becomes clear — the closer the peak value becomes to the average energy in the residual map, the less the loop gain should be. Presuming a CLEANing procedure has just commenced, the gain can be high, subtracting larger portions of the original dirty beam, as it is presumed that there are definite peaks in the residual map. Then, as the peaks become lower, the loop gain adjusts to smaller values to remove less which ensures a finer trimming, until the stopping criteria is reached [5]. The implemented CLEAN algorithm uses a simple gain calculation, using normalised values which produces accurate CLEANed maps and is computationally affordable, with the loop gain set as

$$\gamma = 0.1 \qquad \text{for } \Gamma < 4, \tag{5.3.5}$$

$$\gamma = 0.5 \qquad \text{for } 4 \leq \Gamma \leq 10, \tag{5.3.6}$$

$$\gamma = 1 \qquad \text{for } \Gamma > 10. \tag{5.3.7}$$

where Γ is calculated as $\frac{|\mathbf{R}(l,m)_{Nmax}|}{|\mathbf{B}(l,m)_{Nmax}|}$, the ratio between the normalised residual map and the normalised dirty beam maxima.

5.3.4 Noise

Noise affects the algorithm's ability to locate weak sources as it introduces confusion into the model as to what may or may not be a source. As with most practical systems, noise is unavoidable, especially considering the level of sensitivity and the quality of measurements of modern day interferometric systems — although having increasingly lower noise than past systems. Physical system parameters, environmental effects, aliasing and numerical errors all contaminate the actual observed image, and it is impossible to have an interferometer network completely devoid of noise; however, antenna designs and advanced electronic can be used to minimise the noise to acceptable levels that do not result in unusable or misleading data, where sources are indistinguishable.

For standard Gaussian noise, among other noise models, the total energy in the dirty map is increased as the noise contributes to the energy contained in the map. If weak sources are present, the threshold for stopping may terminate the process before the weak sources are located and CLEANed. This was the motivation behind adding an additional stopping criteria shown in (5.3.3).

The goal is to get closer to the base noise level of a map, as the stopping criteria proposed by Högbom ends CLEANing before weaker sources are identified. The ideal case would be to completely CLEAN the residual map to the noise level. Additionally, the remaining residuals are added to the final cleaned map for a more accurate representation of the brightness distribution, as the clean map contains no noise. The issue of noise only affects the weaker sources that are close to the noise level. In the presence of noisy data the CLEANing of a map is more effective in the first few iterations as the dominant sources are CLEANed. The brighter sources are identified and removed, revealing weaker sources.

5.3.5 Ghost Sources

Ghost sources — sources that appear in the observation data that are not there — can become an issue during the CLEANing process. Distinguishing ghost sources from actual sources is problematic as there are few image catalogues for weak sources to compare to when trying to identify new sources. The most common reason for ghost sources appearing during CLEANing is misidentified ‘source’ peaks due to spikes in observation data, or aliasing components caused by the sampling process.

A peak may be identified that is not actually a source because it is a local peak in relation to the immediate surrounding data values. This is more common in extremely noisy dirty maps. The location of the misidentified source will be recorded and placed onto the clean map as CLEAN will misinterpret this brightness peak as a valid source. Aliasing further contributes to this problem. If the *uv*-data has been under-sampled, it is quite possible that the higher frequency image components have been folded back into the dirty map. These aliased components will be seen as possible sources by the CLEAN algorithm as they appear to be brightness spikes and are generally larger than the noise level.

Loop gain affects how much of an identified peak is subtracted and proportionally placed into the clean map. If loop gain is too high and a dirty beam is subtracted that contains high valued sidelobes, these subtracted sidelobes may result in pockets of ’negative brightness’, close to the actual source identified. ’Negative brightness’ is not a physical, but a mathematical term, as negative brightness values may appear in the residual map when too much brightness is subtracted. The CLEANing process works with the absolute data of the dirty map and these pockets, now positive valued, may be interpreted as a source peak in later CLEAN cycles. The issue of negative brightness can be avoided by making all negative brightness values zero, however, aliased components can still be misinterpreted as sources. Nulling the negative values is used in this implementation of CLEAN, however this runs the risk of over cleaning the images.

5.4 Variations of CLEAN

The algorithm has since undergone much development to extend it to be more adaptable to handle various types of sources. Various models are now widely used in the radio astronomy industry to deconvolve the unwanted dirty beam data to obtain a more accurate representation of the true sky brightness model. Among these variations, the most popular are the Clark [25], Cotton-Schwab [12] and Multi-Scale CLEAN [28] algorithms.

The Clark model is ideal for parallelism and CLEANing of large images [25]. The non-linear part of CLEAN is locating and recording the peak values of a source in the residual map. Clark introduces the concept of CLEAN windows, boxes that divide the map into sections that can be CLEANed separately. Each box can be CLEANed and the final image can be reconstructed by placing the images together. Furthermore, only the largest residuals are searched for as the Clark CLEANing method assumes that the sidelobe response from smaller sources is acceptably small and the larger source sidelobes dominate the distortion in the image. Only points having an intensity greater than the highest sidelobe are CLEANed.

During the subtraction phase, a portion of the dirty beam is subtracted, not the entire beam and as with Multi-Scale CLEAN, the identified components are first Fourier Transformed and multiplied with the sampling function in the uv-domain, and then transformed to the lm-domain for subtraction. This corrects many errors that are introduced in the minor cycle.

Clark is one CLEAN variant; the other popular variation is the Cotton-Schwab algorithm [12]. The algorithm works with ungridded visibility data, so aliasing and gridding distortions are removed when the inverse Fourier transform is applied. The full phase equation is also implemented, which includes the w-term, making it adaptable to non-coplanar baseline effects. The drawback is that the data must be gridded again once the CLEANing is complete and this can be computationally expensive.

Furthermore, these algorithms can implement CLEANing at multiple frequencies. Doing various CLEANing processes at various frequencies, and then overlaying the CLEANed images is called Multi-Frequency Analysis and allows for much more complex evaluations on structures [30]. Each frequency band contains different information regarding a source, so individually CLEANing in a certain band will allow more detailed images to be reconstructed. The most appropriate CLEANing procedure is dependent on the type of data being analysed and the performance requirements of the analysis.

5.5 Conclusion to CLEAN

Both Clark and Cotton-Schwab are prevalent alternatives that extend the CLEAN algorithm to be more versatile for larger and more complex images. Various techniques like Multi-Frequency and Multi-Scale analysis, among others, can be used to improve the CLEAN model and create higher fidelity images. While the original CLEAN model has a few drawbacks, one of the major issues being that of using the point source approximation, it still remains a computationally efficient method for non-linearly deconvolving unwanted data from an image.

It is clear that CLEAN is the fundamental deconvolution algorithm and forms one of the basic methods for other imaging and deconvolution packages used in radio astronomy today. Alternative methods include the Maximum Entropy Method [31] or Compressive Sensing [15], both of which take different approaches in acquiring the true sky brightness model.

In general, imaging can be greatly improved by combining non-coplanar baseline methods and CLEAN, as will be shown in the following chapter. The standard CLEAN model, with contour trimming and the 3σ stopping criteria will be utilised in Chapter 6. This combination of non-coplanar techniques and CLEAN is essentially what happens in practice when reconstructing an image — although more refined CLEANing techniques are used.

Chapter 6

Simulations And Results

6.1 Simulation Setup Overview

The objective of this section is to comparatively evaluate both the *w*-projection and *w*-stacking techniques, by simulating the entire process of recovering a sky model from its sampled visibilities. Furthermore, to mimic real world data as accurately as possible, the sampled visibility model must encompass issues that would typically be found in real world practical applications such as noise and avoid aliasing. Each step in constructing the sky model and the corresponding sampled visibilities as well as using CLEAN and the imaging methods is explained below. The results were implemented on a desktop PC with Windows 7 64-bit OS, with 8192MB RAM and a Intel Core i5-3340 processor with four 3.1GHz CPU's.

6.1.1 The Simulated Sky Model

A 256×256 image was generated in Matlab. The model primarily contains 10 point sources which vary between 0.1mJy and 5mJy in magnitude, with two exceptions: a large disk-shaped source with a peak magnitude of 10mJy which is located slightly off center from the phase reference center, and a jet-like fume located at the bottom left section of the map, as seen in Figure 6.1. Both of these exceptions are included to test the capabilities of the CLEAN algorithm. The jet-like structure was included to test the contour trimming function as it should allow this structure to be identified as one whole source, rather than a combination of point sources whereas the large, bright source was included so that weaker sources are masked when viewed by an array, not originally appearing in the model until the image is CLEANed, making the weaker sources more prominent.

In practice, noise in the interferometer system further adds to the visibility and image data. To encompass all noise that may be experienced by a real world system, Gaussian noise was added to the dirty map. The magnitude of the noise added into the image was 5% of the total average energy in the intensity model; this was 0.009mJy for this model. While this is the most basic of models for noise, it was included to represent atmospheric effects, electronic noise in hardware components and instrumental noise from the antennas. Lastly, a Tukey window with an α=0.25 was applied to the sky model to suppress any aliasing components that may be folded into the image, when the sampled non-coplanar visibilities and associated dirty map are calculated. The models in Figure 6.1 are the benchmark for the image that needs to be reconstructed; the closer the reconstructed image is to Figure 6.1, the better.

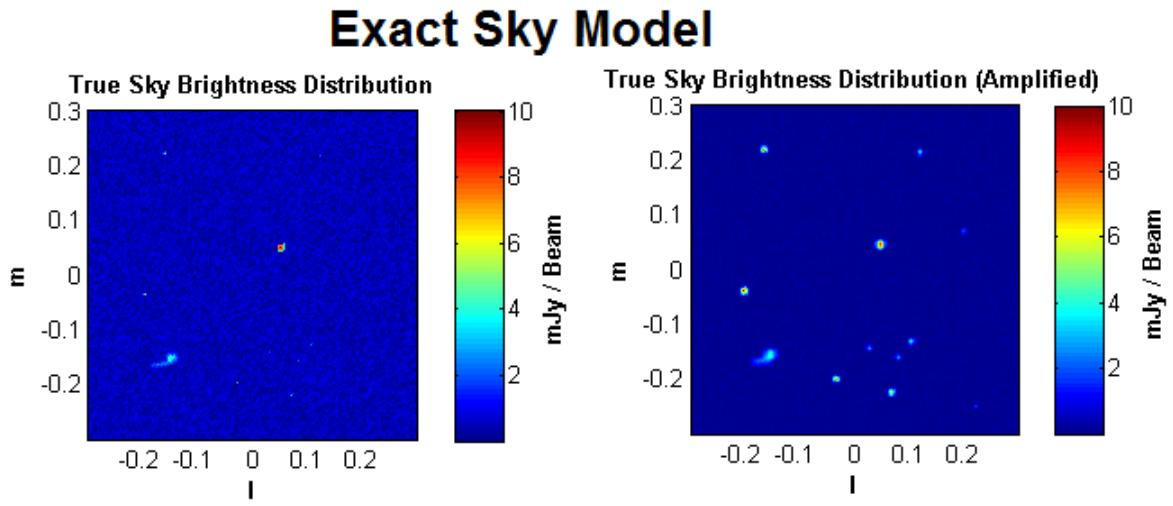


Figure 6.1: Left) The simulated, ideal sky model to be analysed. Right) For the sake of illustration, the large, bright source located near the center is damped to better show the weaker sources.

6.1.2 KAT7 and *uv*-Coverage

In practice, an array configuration would be used to observe the sky. To replicate this, the KAT7 telescope was chosen and simulated for the observation as there are few antennas, limiting the amount of *uvw*-coordinates. Additionally, KAT7 was chosen to avoid memory performance issues when calculating the visibilities. The coordinates of each of the antennas were used to determine the baseline projections for the *uv*-coverage. For the observation parameters a frequency of 1.4GHz was selected as this lies in the middle of the receiver band and the declination was set to -25 degrees. The coordinate data for each antenna was taken from the SKA-SA public directory.

Antenna	North Offset (m)	East Offset (m)	Altitude Offset (m)	Adjusted Altitude Offset (m)
1	-9.1	25.1	0.05	0.05
2	26.38	90.28	-0.23	-1.3
3	26.89	3.99	0.00	0.00
4	25.49	-21.6	0.02	0.02
5	-2.59	-38.27	0.39	0.39
6	-79.7	-61.60	0.70	5.55
7	75.75	-87.99	0.14	0.14

Table 6.1: Antenna offset in meters from the center of the KAT7 array configuration. The array is located at Longitude 30 °43' 17.34", Latitude 21 ° 24' 38.46" and Altitude 1038m. Coordinates were obtained from the SKA public site: http://public.ska.ac.za/kat-7

The array height variation is minimal, with the greatest height deviation being 0.93m between Antennas 2 and 6. With sufficiently small variation in w, this system is considered to be approximately coplanar however, to fully demonstrate the effect of the w-term and both

w-projection and stacking techniques, a much greater variation is needed. The antenna height offsets were altered to give a maximum deviation of 6.85m, which corresponds to a w_{max} = 32 at an observation frequency of 1.4GHz. The coordinates used with the original East-West, North-South data and the altered heights can be seen in Table 6.1.

The *uv*-coverage affects the dirty beam shape; with more coverage in the *uv*-plane, the more ideal the PSF shape will become. Greater *uv*-coverage can be achieved by changing the observation parameters, as the array antenna positions are fixed. Longer observations that have greater integration time, with more frequent sampling intervals will result in a denser sampled *uv*-plane. Figure 6.2 shows the effect of filling the *uv*-plane and the ensuing PSF.

For the simulated KAT7, the most ideal case would be to have an extremely long observation that utilises all of the time that the sources are visible. Additionally short sample intervals should be used. This observation requirements are impractical as both hardware and software limitations constrain the observation capabilities; large observations require immense amounts of data storage and processing. To mimic practical cases, an observation time of 4 hours with 10 minute intervals was selected. As the array only contains 7 antennas, the *uv*-coverage will contain many gaps which will lead to a noisy dirty beam; however, this is not an issue as the images can be improved using the CLEAN algorithm, as shown in Section 6.2.4.

6.1.3 Gridding

Gridding the visibility data is often one of the most computationally expensive steps in the imaging process. It involves making use of the fact that the sampled visibilities are discrete and finite, and if the sampled positions are arranged on a uniform grid, then the FFT can be used to improve the computation of the sky brightness [32]. As this is not vital in showing the effects of the *w*-projection and stacking techniques, the sky model and the corresponding visibilities were arranged onto a uniform rectangular grid from the start using (2.3.1), where each pixel of the intensity map, corresponds to a pixel in the visibility function. This is equivalent to having an ungridded visibility function and a sampling function whose points align with the pixels of the visibility function so the visibility function can be element-wise multiplied by the sampling function, which will result in a gridded, sampled visibility function.

This corresponds to the pillbox method as described by [32] which is equivalent to summing the visibility data in a given cell and assigning this value to the closest rectangular grid point, where each cell is a pixel. This approach is the most efficient in terms of computation, however it is the least effective in deterring aliasing components [32]. For this reason the sources generated in the image are set to lie within a buffer of the edges of the image to avoid additional aliasing in the dirty map. This buffer was set to be 5% of the total image size.

When gridding onto a uniform grid, the finer the grid, the more accurately the *uv*-data will represent the original visibilities obtained from the array configuration. For a grid that has larger cells, the amount of *uv*-data that lies within a cell will increase. This is not desirable as the pillbox method will sum these points together, essentially decreasing the amount of points that describe the visibility data. This will negatively effect the quality of the dirty map as it will be more distorted and noisy.

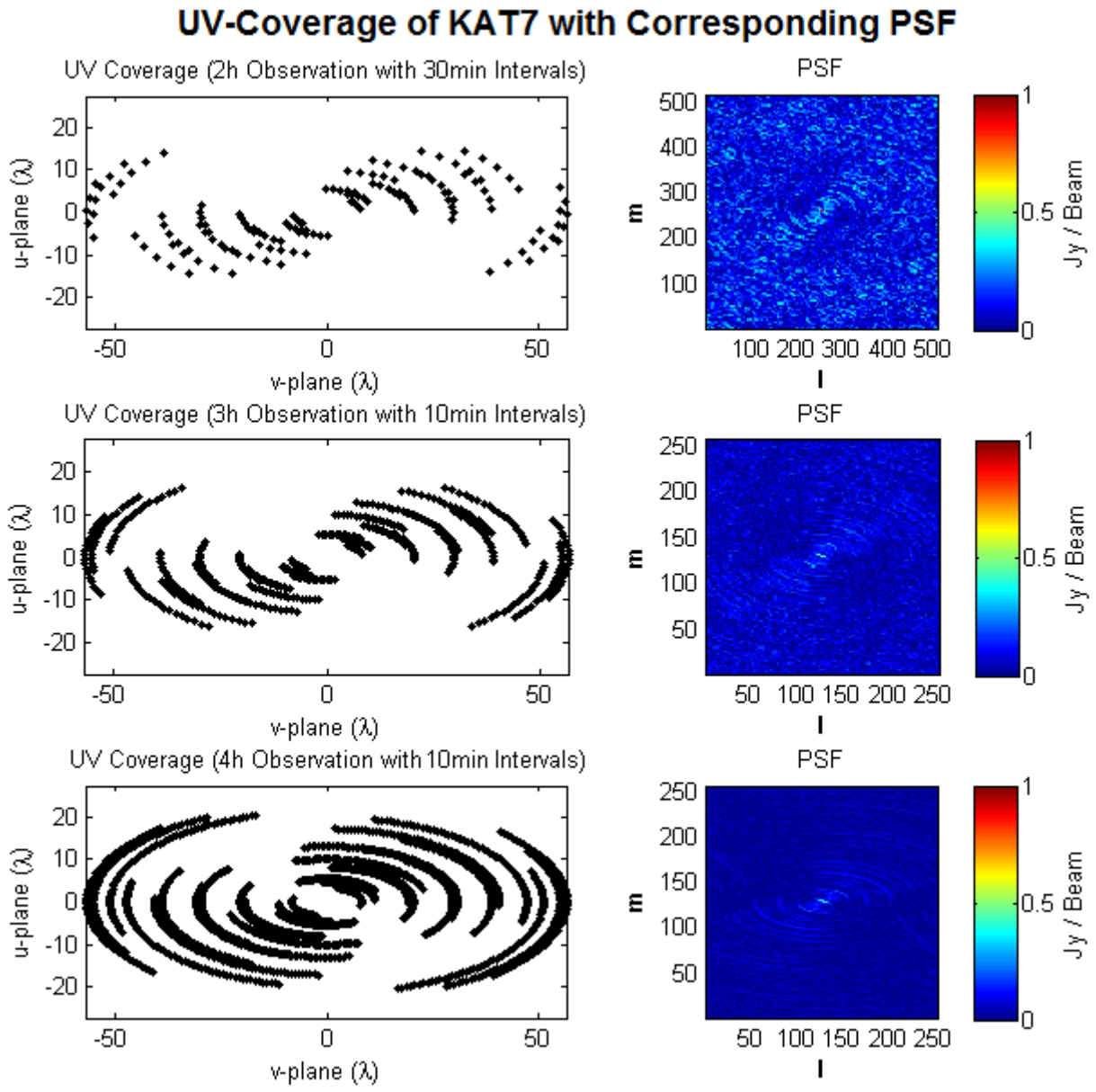


Figure 6.2: Various KAT7 uv-coverage plots for different observation times are shown in the left column. The right shows the corresponding point spread functions. Increased coverage results in a less distorted dirty beam as the sampling function is more dense.

For the simulation, the *uv*-data obtained from the KAT7 array was used for the sampling function and the *w*-data is taken into account to include the additional phase shift. A grid size of 256×256 was used to grid the visibility data as this coincides with the image pixels.

6.1.4 Beam Shaping: Weighting and Tapering

Depending on the application or data, the use of tapering and weighting can be used to adjust the shape of the dirty beam, allowing different images to be obtained in terms of sensitivity and resolution. Consider the weighting function

$$\mathbf{W}(u,v) = \sum_{k=1}^{N} T_k Q_k \delta(u - u_k, v - v_k) \qquad \text{for } N = \text{number of } uv\text{-elements}, \tag{6.1.1}$$

where T_k is the effect of tapering and Q_k is the effect of weighting the densities of the *uv*-data). T_k is the tapering function which tapers the data on the outer edges of the *uv*-plane, while Q_k weights the data values in the *uv*-plane. It is the combination of these that results in the overall weighting [22].

For this simulation, Q_k is considered to be unity, corresponding to natural weighting. The simulation does not specifically require high sensitivity or high resolution; however, the natural weighting scheme was chosen as it minimises the noise in the dirty map and makes the system more sensitive so point sources are more identifiable [22].

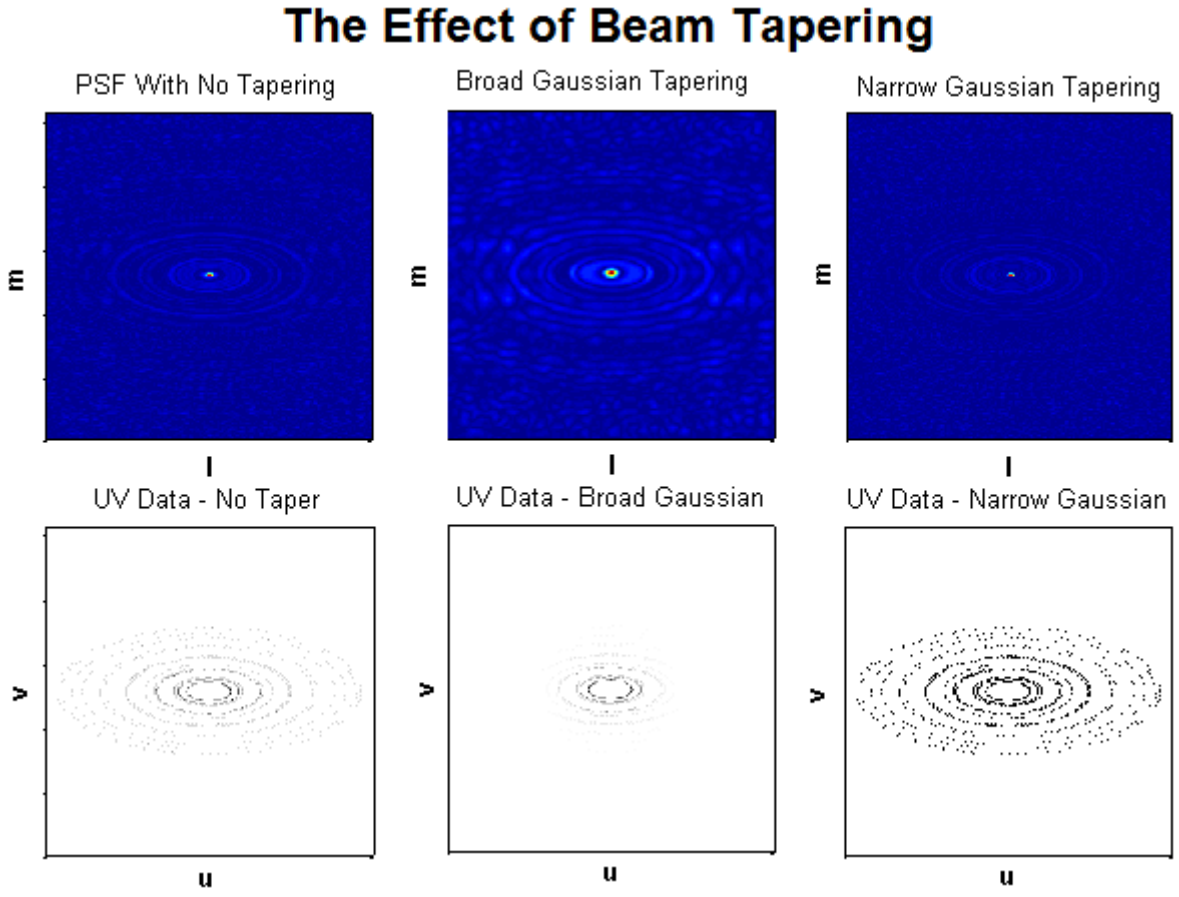


Figure 6.3: The top images show the dirty beams and the effect of tapering using both a broad and narrow Gaussian taper. The uv-data below shows the effect on the uv points and their weightings. A Broad Gaussian taper in the image domain corresponds to a form of natural weighting of the uv-data points , where as the narrow taper corresponds to uniform weighting.

In the image domain, tapering of the PSF can be employed to suppress the sidelobes of the dirty beam, which is equivalent to adjusting the *uv*-data in a similar manner to that of weighting. A Gaussian function remains a Gaussian function if Fourier transformed. Applying a circularly

symmetrical two-dimensional Gaussian taper to the dirty beam of varying width essentially weights the *uv*-points. The decay rate of the sides are inversely proportional; a fast decaying Gaussian taper in the image domain translates to a broad Gaussian taper in the *uv*-domain. As shown in Figure 6.3, tapering the the dirty beam by a broad Gaussian is equivalent to natural weighting or multiplying the *uv*-data by a fast decaying Gaussian.

The weighting function in (6.1.1) was implemented by converting the unweighted *uv*-data to the image domain to obtain the dirty beam. A broad Gaussian taper was then applied to the dirty beam to slightly suppress the sidelobes and improve the SNR and the sensitivity. The Gaussian taper is defined as,

$$T_k(u,v) = e^{\frac{-(u^2+v^2)}{\sigma^2}}, \tag{6.1.2}$$

where σ adjusts the width of the Gaussian taper.

Finally, the dirty beam was then transformed back to the visibility domain so the *uv*-points that were used to determine the sampled visibilities were naturally weighted. The aim is to make the system sensitive to compact sources, although this comes at the cost of resolving the large extended source structures. In spite of this, the extended sources should still be easily detectable but their shapes may be warped due to less resolution. This is not a concern as the structure of the extended sources will not be clearly defined due to CLEAN, but they will be flagged.

6.1.5 Generating the Dirty Map

The objective is to generate a realistic dirty map model that takes the dirty beam, the *w*-effect and the FoV into account. For simulation, two approaches can be taken as shown in Figure 6.4. Process 1, as indicated by the red arrows, involves convolving the ideal sky model with the dirty beam. The resulting dirty map contains intensity data where each source is convolved with the dirty beam, however, the sources will not include any phase shift effects caused by the combined effects of *w* and *n*. To include these effects, another approach was taken.

Using the full visibility equation,

$$\mathbf{V}(u,v,w) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\mathbf{A}(l,m)\mathbf{I}(l,m)}{\sqrt{1-l^2-m^2}} e^{[-2\pi i(ul+vm+w(\sqrt{1-l^2-m^2})-1)]}\,\mathrm{d}l\,\mathrm{d}m, \tag{6.1.3}$$

where *w* and *n* are included in the exponential term, will give the full visibility matrix, where the effects of non-coplanarity and FoV are included. This can then be sampled with a given *uv*-coverage to yield the exact sampled visibility. The data is gridded and then Fast Fourier Transformed to acquire the dirty map data. This process in indicated by the green arrows. This process is limited by the gridding procedure and the number of sampling points in both the *uvw* and the *lmn* spaces, where the amount of noise and aliasing in the image data decreases for finer spacing between between *u* and *l*, *v* and *m*, and *w* and *n*.

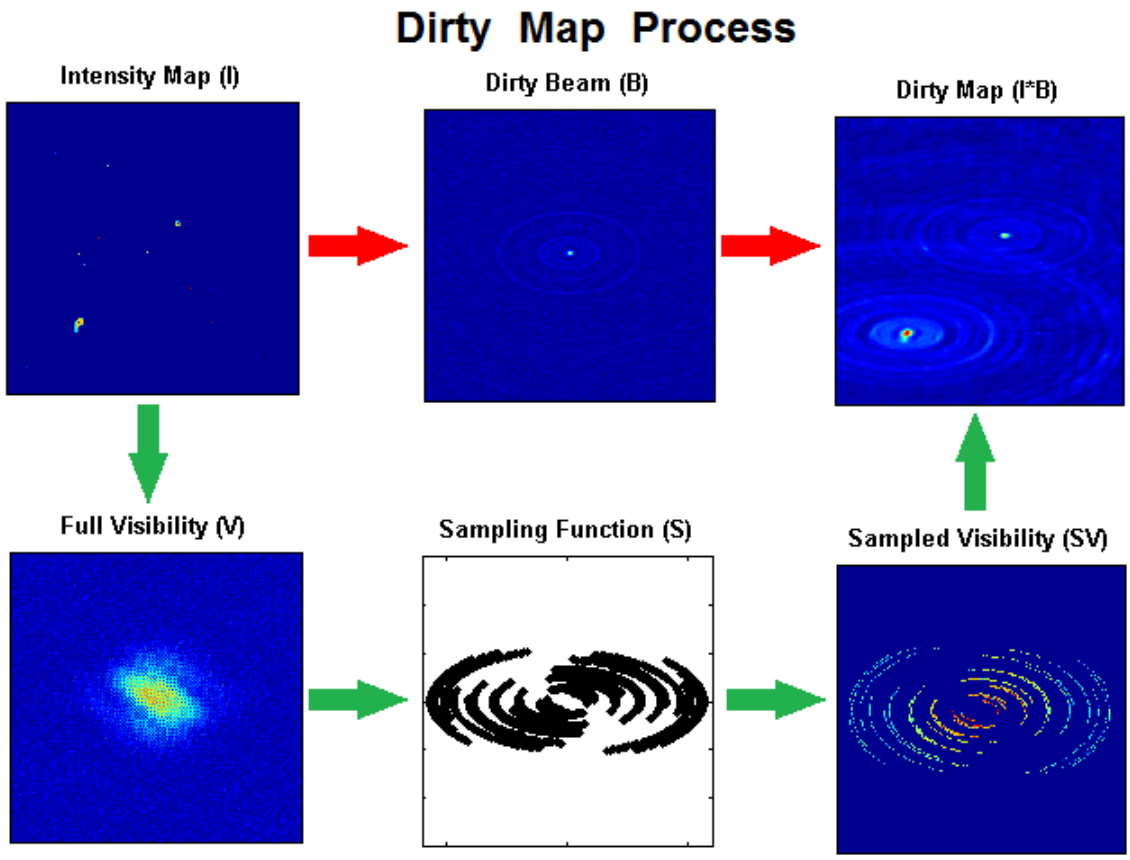


Figure 6.4: Various stages of data, illustrating how the dirty map was generated. The red arrow indicates the process of convolving the dirty beam with the ideal sky model, which ignores non-coplanar and larger FoV effects. The second process, which was implemented, is indicated by the green arrows. This process IFFT's the ideal sky image using the full 3D model, which is then sampled and converted back to the image domain. This process introduces more noise into the final map as it is heavily dependent on sampling accuracy.

6.2 Reconstructing The Sky Model

The idea behind projection and stacking is to remove the phase shift introduced by the w-term and directional cosine, n, due to the non-coplanar array and the FoV, respectively. In the visibility plane, this means one wants to go from $\mathbf{V}(u,v,w)$, which incorporates the phase shift, to $\mathbf{V}(u,v,w=0)$, which can then be transformed into the image domain to give the desired sky model image. This desired model should, in theory, be a good representation of the dirty map obtained by a coplanar array as the non-coplanar effects have been removed. For this section, the coplanar dirty map refers to the dirty map image produced with w=0, which in the Fourier domain correlates to $\mathbf{V}(u,v,w=0)$. The non-coplanar dirty map is used to describe the dirty map that contains the additional phase shift, associated with $\mathbf{V}(u,v,w)$.

The full visibility function was determined by using the exact w-values and performing the full 3D transform in (2.1.6). This visibility function and the corresponding dirty map provides the data set that w-projection and w-stacking must make coplanar. To differentiate between the coplanar sampled visibilities, $\mathbf{V}(u,v,w=0)$, that were calculated from (2.1.10), the visibilities obtained by w-projection and w-stacking will be indicated by $\mathbf{V}_p(u,v,w=0)$ and $\mathbf{V}_s(u,v,w=0)$, respectively. By performing either projection or stacking,

$$\mathbf{V}_p(u,v,w=0) \approx \mathbf{V}_s(u,v,w=0) \approx \mathbf{V}(u,v,w=0). \tag{6.2.1}$$

The methods are not exact as they are influenced by sampling, approximations and the FoV so there may be some deviation from that of the actual $\mathbf{V}(u,v,w{=}0)$. Once $\mathbf{V}(u,v,w{=}0)$, $\mathbf{V}_p(u,v,w{=}0)$ and $\mathbf{V}_s(u,v,w{=}0)$ are computed, they are transformed to the image domain. The coplanar dirty map was set as the benchmark image.

The images produced qualitatively look very similar so there is a need to provide an objective quantification of the error or deviation from the exact intensity data. The energy difference of the images was chosen as a measure of how accurate the two techniques perform. By subtracting the intensity data from the coplanar dirty map and the intensity data from the phase corrected non-coplanar dirty maps, the difference in intensity can be found. For weak sources or sources that are close to the phase reference center, the difference between the two dirty maps is minimal and the residual map is not visually informative and is not an effective means for judging the accuracy of the two methods. By taking the total average energy of the residual map, the error can be quantified. For images with little deviation, the energy content will be similar and the residual map energy will approach zero. For images that have greater deviation, the residual energy will increase. This is calculated as

$$\epsilon = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b}(|\mathbf{R}_{exact}(l,m)_{i,j}| - |\mathbf{R}_{obtained}(l,m)_{i,j}|)}{ab} \qquad \text{where } a = \text{rows}, b = \text{columns}, \tag{6.2.2}$$

where $\mathbf{R}_{exact}$ and $\mathbf{R}_{obtained}$ are the intensity maps of the exact sky model and the obtained sky model, respectively. It must be noted that all images showed were normalised to the peak value of the original sky brightness distribution as some images were scaled differently due to discrete operations.

6.2.1 Projection Results

The sampled visibilities, $\mathbf{V}(u,v,w)$ were convolved with the inverse of the appropriate projection kernel, $\hat{\mathbf{G}}(u,v,w)$, to remove the effect of phase shift. $\hat{\mathbf{G}}(u,v,w)$ was numerically determined from $\mathbf{G}(l,m,w)$. As stated earlier, projection does not perform well in correcting the phase error for wide-field imaging as it is limited to the FoV size, as shown in Section 3.5.1. For a large FoV, the small angle approximation does not apply as the condition $-\pi i w(l^2{+}m^2) \ll 1$ is not met, making the kernel deviate too much from the correct model. The sky model was set to a FoV where $\sqrt{l^2+m^2} \leq 0.3$.

Figure 6.5 shows the benchmark image and the non-coplanar dirty map which includes non-coplanar baseline effects, where the image data is mostly nonsensical. $\mathbf{V}(u,v,w)$ was projected to obtain $\mathbf{V}_p(u,v,w{=}0)$. Theoretically $\mathbf{V}_p(u,v,w{=}0)$ contains the sampled visibilities with the w-term effects removed. Without any practical effects such as aliasing, noise and sampling quality, the result should almost directly match up with the original visibilities. There will be slight differences as the method is based on the small angle approximation which is used to obtain the projection kernel, $\hat{\mathbf{G}}(u,v,w)$. Figure 6.6 illustrates how well w-projection

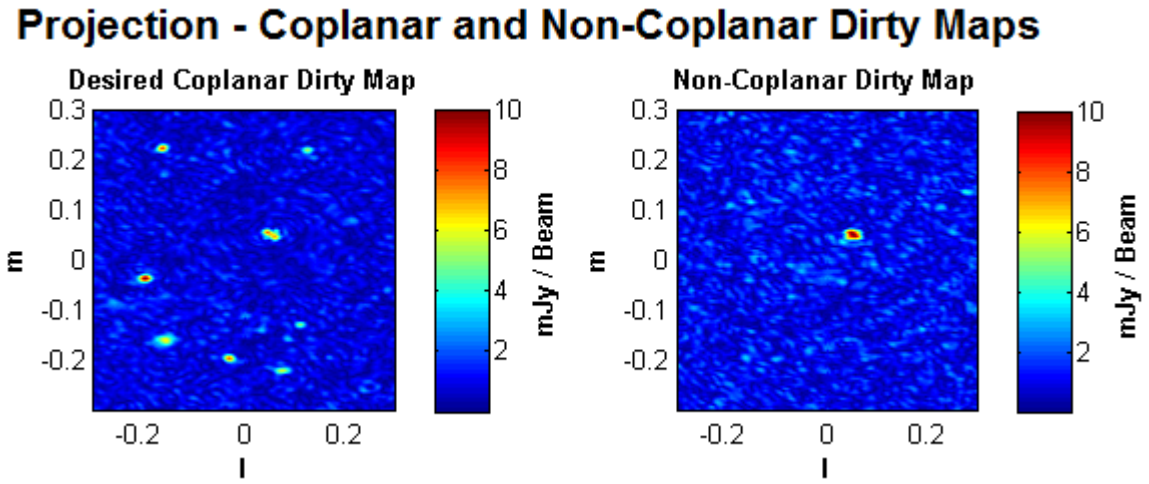


Figure 6.5: Left) The desired coplanar dirty map. Right) The dirty map produced when non-coplanar effects are not accounted for. The sources are completely hidden and indistinguishable due to the additional phase shift, apart from the strong source near the center.

performed. The map on the left is the dirty map obtained from $\mathbf{V}_p(u,v,w{=}0)$, which was Fourier Transformed to produce the dirty map.

The data coincides almost perfectly with only slight differences, as shown in the residual map. The residual map is an indication of how much the data deviates, with the total energy deviation of the normalised maps being 0.00819. This indicates that the two maps are extremely similar as approximately only 0.82% of the data differs. This can further be improved by increasing the number of samples or decreasing the FoV, although exact matching up of the data is not possible, as stated before, due to the small angle approximation. Conversely, analysing a wider FoV or an image where the sources are located closer to the borders of the image as they are more influenced by n, will affect the image reconstruction fidelity. Overall, it is clear that projection works and is a good imaging method to remove the non-coplanar baseline effects to relatively good accuracy.

6.2.2 Stacking Results

As with projection, the same non-coplanar sampled visibilities were determined and the corresponding dirty map was set as the benchmark image to attain with w-stacking. All parameters were kept the same so the two methods are using identical data. The minimum number of layers, as specified by (4.2.1) for this simulated model was calculated to be 110. To avoid aliasing and ghost artifacts appearing in the corrected image, the number of layers was chosen to be 330. By increasing the number of layers, the w-values will require smaller adjustments, which will result in greater accuracy at the cost of computation.

$\mathbf{V}(u,v,w)$ values which shared a w-value were gridded and transformed to the image domain. Each of these layers were phase corrected with the relevant w-value, which produced a variant of the dirty map data. The cumulative data obtained from each of these layers were overlaid and the resulting image can be seen in Figure 6.7. The benchmark and the initial non-coplanar dirty map images are identical to that of Figure 6.5.

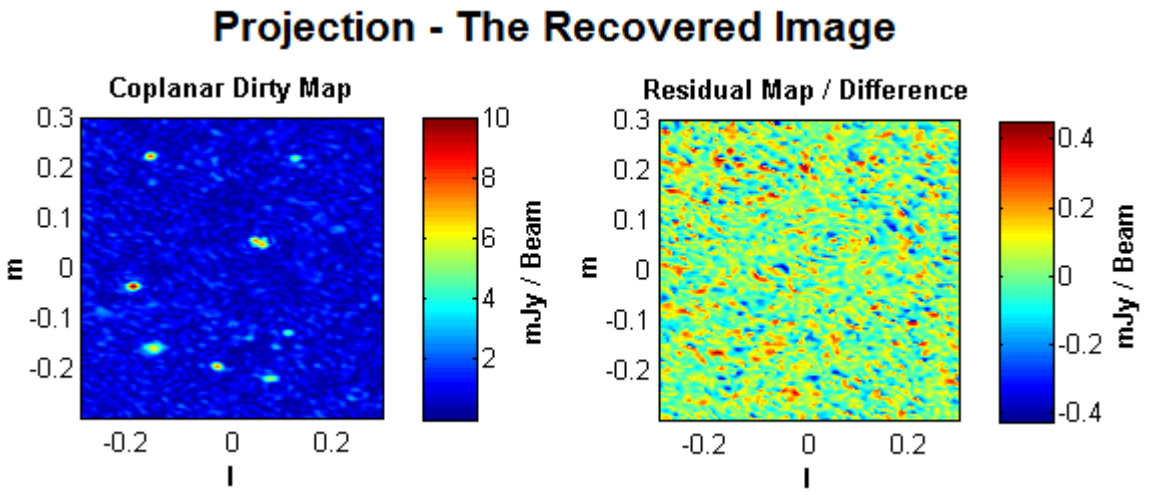


Figure 6.6: Left) The obtained dirty map after performing w-projection. Right) The residual map which shows the difference in obtained dirty map and the desired dirty map. The normalised residual energy difference is 0.82%, indicating most of the brightness distribution matches the desired dirty map.

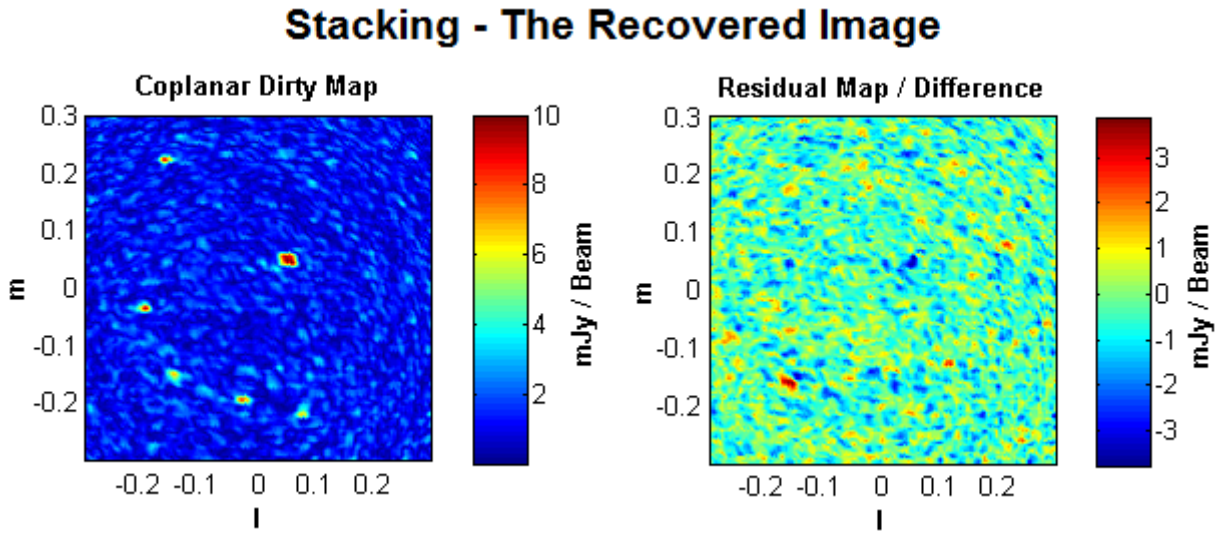


Figure 6.7: The obtained dirty map after performing w-stacking is shown on the left. The residual map is shown on the right which shows the difference in obtained dirty map and the desired dirty map. The normalised residual energy difference is 5.60%.

The obtained dirty map looks extremely similar to that of both the benchmark image and the result of projection. The residual map in Figure 6.7 is also a unclear indication of the difference in the images, so the normalised average residual energy was evaluated, although the scale is ten times higher than that of the projection residual map in Figure 6.6. The energy deviation was measured to be 5.60% which is not as accurate as the *w*-projection results. This accuracy is heavily dependent on the number of layers. For this simulation, the number of unique *w*-values specific to the projected baselines were exactly 900. In practice, most interferometric arrays are approximately coplanar, making the *w*-space sparse as most values of $w \approx 0$, however the simulated KAT7 array was purposefully adjusted to give a large deviation in *w*-values - this is unrealistic but was chosen to show the capabilities of the two algorithms.

While enough layers were used to avoid any aliasing, the algorithm shifted 900 unique *w*-values to 330 discretised values as each sampled *uvw*-point is calculated for each *lmn*-point. While the shifts might be slight, they have a cumulative effect on the data. Decreasing the number of layers to 220 gave a total residual energy deviation of 6.72% where as increasing the number of layers to 440 gave 5.19%. An almost perfect recovery of the original image will be achieved when there is a discretised layer for each unique *w*-value. However, this will be computationally prohibitive.

Comparatively, the two methods each remove non-coplanar baseline effects to great accuracy. While projection was more accurate in recovering the visibilities and the coplanar dirty map, it was limited to the FoV, where as *w*-stacking does not have this issue. Stacking was limited by the computational costs, where the number of layers affected the accuracy. For more realistic *w*-space values and on larger systems where memory and processing power are not constraining factors, the image fidelity can be improved — see also the paper by Offringa et al. [4] for comparison.

6.2.3 Computational Assessment

While many imaging techniques may yield similar or even identical results, the speed and computational complexity of the technique is ultimately what governs which method trumps another, provided that the imaging objective and performance is otherwise equivalent. The processing speed and resource cost of an algorithm are important role players as an algorithm that is too slow is highly undesirable. The computational cost can be quantified in either the time taken to execute an algorithm, or the cost of operations by a processor — both of which are used to compare *w*-stacking and *w*-projection in this section.

Stage	*w*-Stacking	*w*-Projection	3D FFT
Gridding	$\mathrm{N}_v\mathrm{N}_k^2$	$\mathrm{N}_v\mathrm{N}_k^2$	$\mathrm{N}_v\mathrm{N}_k^3$
Fourier Operations	$\mathrm{N}_l\mathrm{N}_p^2\log(\mathrm{N}_p)$	$\mathrm{N}_p^2\log(\mathrm{N}_p)$	$\mathrm{N}_p^2\log(\mathrm{N}_p)$
Corrections	$\mathrm{N}_l\mathrm{N}_p^2$	$\mathrm{N}_v\mathrm{N}_{wk}^2$	N/A

Table 6.2: Computational Costs of w-Stacking, w-Projection and the Direct 3D Fast Fourier Transformation which are subdivided into three stages: the Gridding procedure, Fourier transformation and Correction for the w-term.

As stated by Offringa [4], the computational costs can be split into three stages for both stacking and projection, each with their own operational costs which are shown in Table 6.2, where N_l is the number of layers used, N_p is the number of pixels contained in an image along one dimension, N_v is the number of visibility samples, N_{wk} is the size of the *w*-kernel used for projection and N_k is the number of pixels affected by the anti-aliasing kernel or gridding kernel. It is assumed that the dimensions of the Fourier Transform variables are suitable for radix 2. When considering the computational time cost (CTC), it must be noted that the three stages occur in a different order for each algorithm.

For stacking, the data is gridded and then divided into numerous layers, of which all undergo the Fourier transform and w-correction. For projection, each *w*-kernel, specific to a given *w*-value, must be included in the gridding, then the Fourier transform is applied. Finally, for the 3D FFT, no phase corrections need to be performed as the *w*-values are incorporated into the Fourier Transform, where it can be seen as a 2D FFT which then undergoes a 1D FFT, using the projection slice theorem [16]. Gridding however becomes more complex as the third dimension must now be gridded - the visibilities must be gridded onto a discretised cube. Using Table 6.2, the CTC for each of the three algorithms are as follows:

$$CTC_{stacking} = [N_l N_p^2 log(N_p^2)] \ + \ [N_l N_p^2] + N_v N_k^2. \tag{6.2.3}$$

To understand the trend of the operational costs when stacking is scaled, with respect to the operations being performed, (6.2.3) can be simplified by applying the asymptotic boundaries. The operations that require the most processing will dominate and the operations that become negligible can be removed, giving

$$CTC_{stacking} \approx \mathcal{O}[\ N_l N_p^2 log(N_p^2) + N_v N_k^2\ ] \tag{6.2.4}$$

This can process is repeated for projection and the 3D FFT,

$$CTC_{projection} = [N_v N_k^2 \ + \ N_v N_{wk}^2] + N_p^2 log(N_p^2) \tag{6.2.5}$$

$$CTC_{3D} = N_v N_k^3 + N_p^2 log(N_p^2) \tag{6.2.6}$$

where the limiting asymptotic behaviour becomes

$$CTC_{projection} \approx \mathcal{O}[\ N_v N_{wk}^2 N_k^2 + N_p^2 log(N_p^2)\ ] \tag{6.2.7}$$

$$CTC_{3D} \approx \mathcal{O}[\ N_v N_k^3 + N_p^2 log(N_p^2)\ ] \tag{6.2.8}$$

Considering (6.2.8), it contains a cubic term which will become costly for large images. Using variant models of the multi-dimensional FFT could change the complexity costs of using the 3D FFT [33] although this was not investigated in the scope of this thesis. This was only included to give further motivation as to why projection and stacking are used for non-coplanar baseline effects, when the 3D FFT is available as it is more computationally expensive.

Comparing (6.2.4) and (6.2.7) shows that the number of layers, N_l and the size of the projection kernel which is influenced by the size of *w*, N_{wk}, govern which process is more efficient. Stacking will be faster when gridding is the dominating computational step, as the projection kernel needs to be gridded each time and the size of the projection kernel also detrimentally influences the speed of the operation. Projection will be more efficient to use when there are many layers required for stacking as each layer needs to be Fourier transformed, so if the FFT or IFFT is computationally expensive, numerous layers will linearly increase time costs.

The computation of *w*-stacking is directly affected by the number of *w*-planes that are used to discretise *w*. As shown in Figure 6.8, the difference in residual energy between the desired dirty map and the recovered dirty map using stacking is minimised as more layers are used. Ideally,

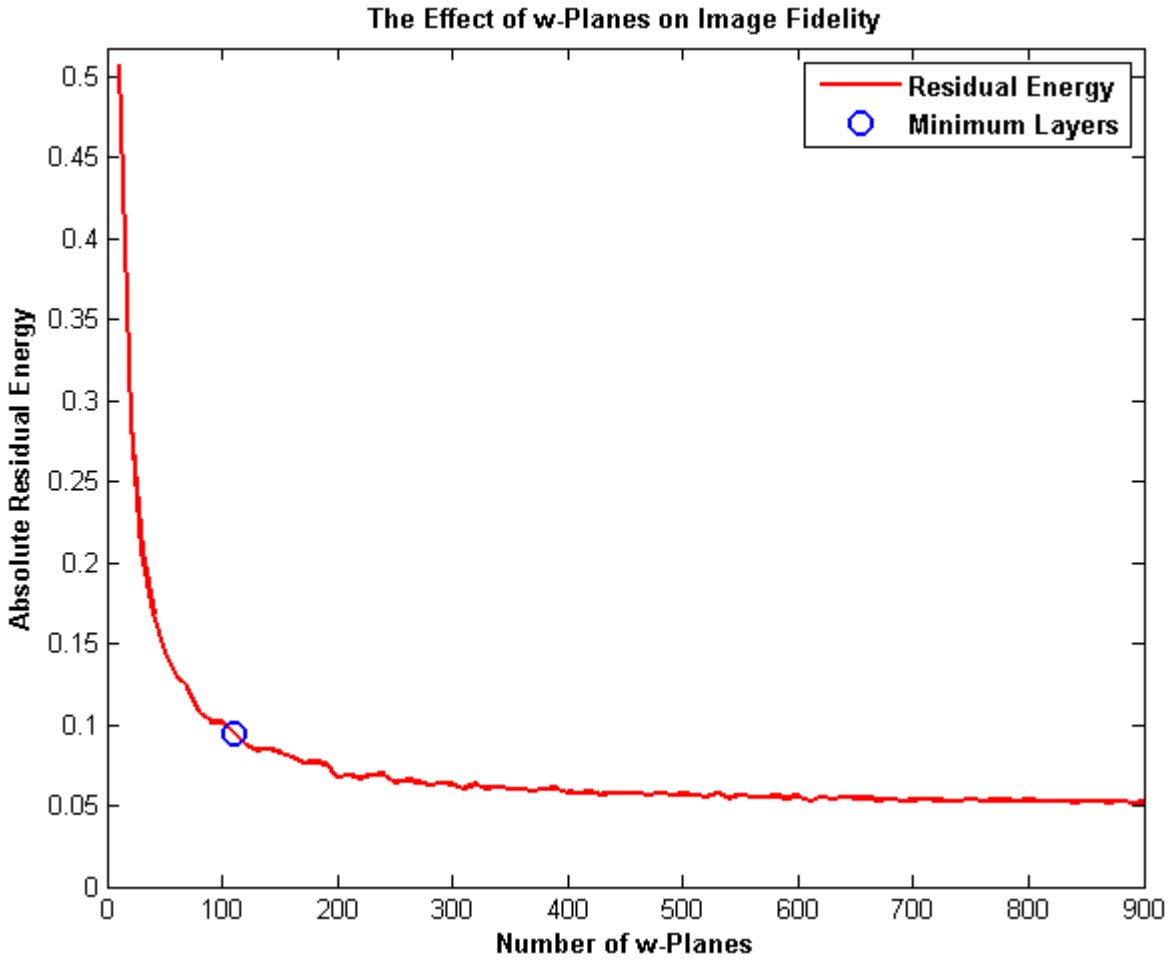


Figure 6.8: The graph illustrates how the number of layers used for w-stacking affect the quality of the recovered image. For increased layer usage, the recovered image approaches that of the desired image. The minimum layers were calculated to be 110.

the desired energy difference is zero but this is not possible due to noise. For this simulation, the minimum layers were determined to be 110. This simulation was highly inefficient and time consuming in order to compute the 900 layers. The graph shows that initially, there is an exponential improvement to the image fidelity, but there is only a slight improvement after, roughly twice the specified number of minimum layers is surpassed. This indicates that using too many layers has no benefits on improving the accuracy of the recovered image, and will only increase computation costs as there are many unnecessary image phase corrections being performed in the stack.

6.2.4 CLEANing The Coplanar Images

The previous sections illustrated how both methods are useful in removing non-coplanar effects and produced almost similar results for their corresponding dirty maps. These dirty maps can be further improved to obtain an intensity model that is free from the distorted effect of the dirty beam. As both coplanar maps associated with $\mathbf{V}_p(u,v,w{=}0)$ and $\mathbf{V}_s(u,v,w{=}0)$ are similar, only the dirty map obtained from the *w*-projection algorithm was used. The slight differences in the two maps are not noticeable and CLEANing each image will once again have almost identical results as the resolution of the images and the PSF is not high. For the Cleaned Map,

the remaining residual data was added to the image to include the energy that was not removed during the CLEANing process. As a last note, the Tukey tapering window with α=0.25 was element-wise divided out of the Cleaned Map to remove the taper that was applied to the original sky model to avoid aliasing of sources close to the edges.

The recovered coplanar dirty map which was produced by *w*-projection was CLEANed as shown in Figure 6.9. The CLEAN algorithm identified all sources apart from one, which is evident when comparing the original sky model and the Cleaned Map. Originally the sky contained a total of ten point sources and two extended sources, but CLEAN was unable to identify one of these point sources. This is due to the added noise in the original image. The missing source magnitude was close to the noise level, so when the CLEAN algorithm removed the effects of the other sources, the energy in the residual map decreased. The CLEANing threshold was reached before the source could be detected as a noticeable peak and the overall CLEANing process was terminated.

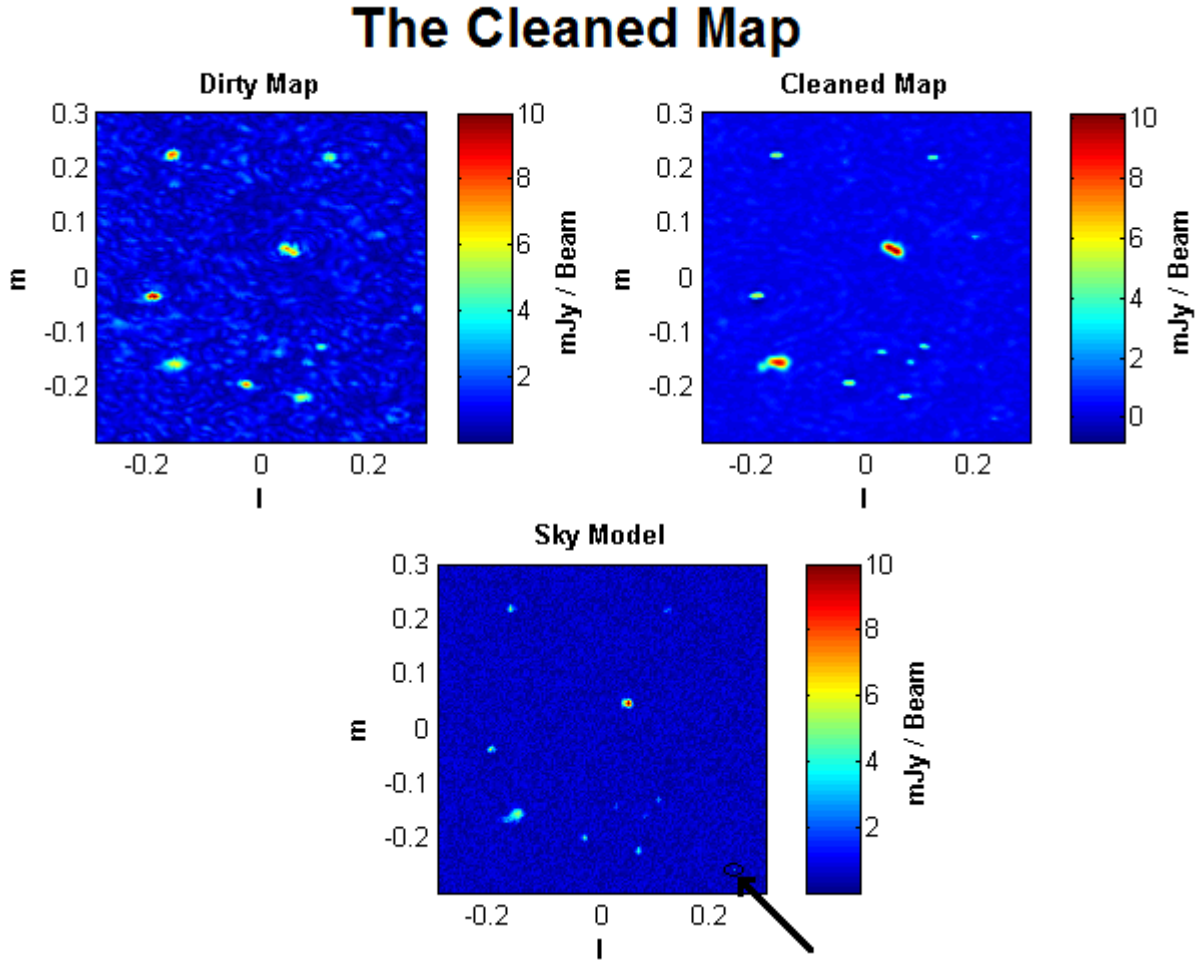


Figure 6.9: Top Left) The dirty map which is the image to be CLEANed. Top Right) The Cleaned Map showing the effect of deconvolving the PSF from the image. Bottom) The true sky brightness distribution. The point source locations align with the Cleaned Map sources and all but one of the sources have been identified. The missing source is indicated by the arrow.

Furthermore, the demonstrated CLEAN algorithm is a variant of the original CLEAN algorithm as it uses contour trimming which makes it slightly more adaptable to extended sources, as explained in Section 5.3.2. The contour trim parameter was set to 0.9, meaning the data surrounding the detected peaks was checked if it was above 90% of the peak value. If any samples were above the contour, they were grouped together and removed from the dirty map.

The removed data was convolved with the PSF to give the appropriate shape of dirty beam that needed to be subtracted. While not too accurate in identifying the shape of the extended sources, the CLEAN variant managed to illustrate some form of the extended sources, rather than depict them as numerous point sources. In practice one would make use of CLEAN models that can handle extended sources; however, this is useful in flagging that an image contains extended sources. Additionally, all sources in the Cleaned Map are slightly bigger than the original image as the Gaussian clean beam occupied 8×8 cells which were convolved with the removed dirty map components. This further limited the amount of detail achievable for the extended sources.

The results show that both the *w*-stacking and *w*-projection algorithms are extremely effective in removing the effects on non-coplanar baselines. Both techniques showed that the additional phase shift can be accounted for, even for large *w*-values, allowing the source data to be less obscured — with the benefit of being able to use the 2D FFT. Projection was more accurate in removing the phase effect than stacking. Conversely, the FoV was restricted as the convolution kernel used for projection must adhere to the small angle approximation requirements. Increasing the FoV would affect the projection performance, while stacking would be unaffected. Other parameters affecting the results include the number of imaging layers, the sparsely sampled *uv*-plane using KAT7, which would be improved for greater *uv*-coverage and the performance capabilities of the machine that the methods were executed on.

Lastly, the standard CLEAN model adequately showed how, through iterative deconvolution, the effects of the dirty beam can be removed, for the most part. This allowed the majority of source to be identified and provided a means of flagging the extended source emissions. No variants of the CLEAN model were considered, which would result in improved image quality. The final chapter concludes both the non-coplanar imaging techniques, the CLEAN algorithm and the obtained results.

Chapter 7

Conclusion

7.1 A Retrospective Overview

With insight into how non-coplanar baseline effects and wide-field imaging can affect the overall quality of the images, it is clearly evident that there is a need for non-coplanar interferometric imaging techniques. The addition of the extra phase shift in either the image or Fourier domains can produce data that is partially correct if it is not taken into account. In imaging, the phase shift can skew magnitude readings, shift sources to different positions, generate ghost artifacts or alter the shape of the source itself. Non-coplanar effects are problematic and more prominent as the imaging capabilities of radio astronomy and the encompassing fields increase. Meeting these imaging demands is made possisble by the development of larger interferometric systems and more sophisticated technologies and techniques.

Two algorithms discussed in this project, namely *w*-projection and *w*-stacking have shown how to account for the non-coplanarity issue. The *w*-projection algorithm, primarily developed by Cornwell et al. [3], uses a projection kernel which is convolved with the visibilities in the Fourier space to remove the non-coplanar effects. By essentially projecting the *w*-term out of the problem, the non-coplanar visibilities become coplanar and the standard two-dimensional relationships between visibility and the brightness distribution can be preserved. This technique is limited to a semi-wide FoV as the projection kernel works under the small angle approximation and was calculated numerically.

The *w*-stacking algorithm aims to do the same function as *w*-projection, but it accounts for the *w*-term in the image domain. Developed by Cornwell and implemented by Offringa et al. [4], it involves creating multiple layers of the sky brightness distribution which are individually phase-corrected depending on the associated *w*-value. This stack of sky brightness models are then summed, with the overall effect being that the final brightness model is phase corrected for the non-coplanar effects. This is not limited to a FoV size but, the process can become computationally expensive if extremely high image fidelity is required as the accuracy is dependent on the number of images in the stack. The algorithms are both, however, limited to one hemisphere.

Finally, CLEAN is a two-dimensional iterative deconvolution algorithm which is not used to eliminate non-coplanar and wide FoV effects, but rather to improve upon the observed

brightness distribution. It aims to remove the effect of sampling that is introduced by the physical interferometric system, by iteratively subtracting the PSF from the observed intensity. This process was formulated by Högbom and has become standard practice in imaging, although since its original development, the algorithm has fathered many variant models to suit the given data application. The original model was created under the assumption that all sources are point sources; however, this is not usually the case. A variant model that employs contour trimming was implemented to account for some degree of extended source resolution and flagging.

Imaging entails various processes and data stages. The majority of these, and certainly the most important stages, have been illustrated in this project. Apart from non-coplanar baseline effects, full understanding of the entire imaging process has been shown. This includes how one obtains a true sky brightness model from sampled visibilities and how to predict the sampled visibilities from a given sky model. Common imaging steps were used such as gridding, weighting and tapering to improve the desired result, although, with regards to tapering, a buffer on the source locations was included to ensure aliasing did not occur, provided the sampling criteria was met. Concepts of the van Cittert-Zernike Theorem, *uv*-coverage and point spread functions were understood and made use of. Furthermore, typical signal processing theory and imaging issues were discussed and accounted for where possible, such as aliasing and noise.

7.2 Comparison of the Algorithms

Less trivial, was understanding the three main algorithms utilised in this thesis. It was shown for a sky model which was analysed using a simulated KAT7 array configuration, that *w*-projection can remove the effect of *w* to an accuracy of approximately 1% of the original coplanar model. Stacking fared almost as well, only showing a deviation of approximately 5% in the recovered image data compared to the original coplanar model. The accuracy of these algorithms is dependent on all observation parameters and is subject to change for each different simulation.

Additionally, there is a trade-off between image fidelity and the computational costs of each algorithm. Depending on the dominating process of either gridding or the inverse Fourier transforming, one algorithm is preferable to the other. For expensive convolution computational costs for the projection kernel, *w*-stacking is the preferred method. If the inverse Fourier transform is too expensive, then *w*-projection is more efficient in obtaining the coplanar visibilities.

Overall it was shown that these two imaging techniques are accurate algorithms that can be used to correct non-coplanar baseline effects, with the benefit of a limited wide-field image of $\sqrt{l^2 + m^2} \leq 0.3$ for projection, and full wide-field imaging capability for stacking. Non-coplanar visibilities that result in heavily distorted intensity maps can be corrected to give the coplanar models which allows the Fast Fourier Transform to be used. It was also shown how CLEAN can deconvolve the effects of the point spread function to predict the 2D visibility models that have reduced detrimental effects due to sampling in the *uv*-plane.

The entire process of recovering the true brightness distribution sky model was demonstrated: Generating a true sky brightness model; applying most of the detrimental effects of gridding, aliasing, system noise and non-coplanar effects to predict the visibilities and the associated dirty map as seen by an interferometric array configuration; using the two non-coplanar imaging methods to correct the visibility data and corresponding dirty maps to the 2D coplanar case; assessing areas where aliasing and noise have an influence on the data and taking precautions to avoid them; and finally CLEANing the images to recover the true sky brightness model.

7.3 Future Work Recommendations

The algorithms have been proven to account for non-coplanar baseline effects. Future recommendations for *w*-projection would be to determine the convolution kernel, avoiding the small angle approximation. The kernel, $\hat{\mathbf{G}}(u,v,w)$, was determined numerically in this thesis and was limited to the performance capabilities of the machine. If $\hat{\mathbf{G}}(u,v,w)$ could be determined exactly without the small angle restrictions, the errors due to the deviation from the phase tracking center would be removed. This would allow *w*-projection to be used for full wide-field imaging.

The *w*-stacking algorithm already has full wide-field imaging capabilities. As the goal was to compare projection and stacking, the same limited FoV was used which did not get to test the wide-field stacking capabilities. Additionally, some of the issues encountered included performance issues specific to the machine, whose specifications are defined in the start of Chapter 6. The processing issues severely hampered the sampling and gridding procedures which directly influenced the resolution and overall quality of the images that could be produced. It would be handy to push the performance boundaries, with demanding gridding and FFT requirements, for both the *w*-projection and *w*-stacking techniques. If done on more advanced machines, or using a GPU implementation, this would allow a detailed computational analysis of the two methods. Various wide-field observations and a general performance analysis were implemented by [4], however more sources would be ideal as the algorithm has only recently been developed.

Lastly, Offringa et al. [4] suggests that a combination of both projection and stacking could be used to beneficial effect. This was not investigated, only a theoretical motivation was provided. Projection could be used to correct for small *w*-values prior to the FFT, as the required kernel performance support will be small. Subsequent to the FFT, stacking could be performed on the larger *w*-values to correct for the remaining phase shifts. The combined algorithms could be optimised to find an effective method that allows small kernels to be utilised while simultaneously reducing memory requirements for faster and more efficient imaging results.

Bibliography

[1] J. Kraus, *Radio Astronomy*, 2nd ed. Ohio: Cygnus-Quasar Books, 1986.

[2] A. Thompson, J. Moran, and G. Swenson, *Interferometry and synthesis in radio astronomy*, 2nd ed. New York: Wiley-VCH, 2001.

[3] T. J. Cornwell, K. Golap, and S. Bhatnagar, "The Noncoplanar Baselines Effect in Radio Interferometry: The W-Projection Algorithm," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 5, pp. 647–657, Oct. 2008. [Online]. Available: http://arxiv.org/abs/0807.4161

[4] A. Offringa, B. McKinley, and et al., "WSCLEAN: an implementation of a fast, generic wide-field imager for radio astronomy," *Monthly Notices of the Royal Astronomical Society*, vol. 14, no. July, pp. 1–14, Jul. 2014. [Online]. Available: http://arxiv.org/abs/1407.1943

[5] J. Högbom, "Aperture synthesis with a non-regular distribution of interferometer baselines," *Astronomy and Astrophysics*, vol. 15, p. 417, 1974. [Online]. Available: http://articles.adsabs.harvard.edu/full/1974A%26AS...15..417H

[6] K. Chiste, "Astronomy: The Constellations and Coordinate Systems," 2014. [Online]. Available: http://www.dorthonion.com/drcmcm/ASTRONOMY/Lessons/Lectures/02_Constellations.php

[7] A. Thompson, "Chapter 2. Fundamentals of Radio Astronomy," *Synthesis Imaging in Radio Astronomy II*, no. 180, pp. 11–36, 1999. [Online]. Available: http://adsabs.harvard.edu/full/1999ASPC..180...11T

[8] R. Booth, W. de Blok, J. Jonas, and B. Fanaroff, "An open invitation to the Astronomical Community to propose Key Project Science with the South African Square Kilometre Array Precursor," *Instrumentation and Methods for Astrophysics*, 2009. [Online]. Available: http://arxiv.org/abs/0910.2935

[9] R. Perley, "Imaging with non-coplanar arrays," *Synthesis Imaging in Radio Astronomy II*, vol. 180, pp. 383–400, 1999. [Online]. Available: http://adsabs.harvard.edu/full/1999ASPC..180..383P

[10] T. Cornwell and S. Bhatnagar, "Wide field imaging problems in radio astronomy," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, pp. 861–864, 2005. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1416440

[11] T. J. Cornwell and R. Perley, "Radio Interferometric Imaging of Very Large Fields," *Astronomy and Astrophysics*, vol. 261, pp. 353–364, 1992.

[12] U. Rau, S. Bhatnagar, M. Voronkov, and T. Cornwell, "Advances in Calibration and Imaging Techniques in Radio Interferometry," *Proceedings of the IEEE*, vol. 97, no. 8, pp. 1472–1481, Aug. 2009. [Online]. Available: http://arxiv.org/abs/0902.0817

[13] T. J. Cornwell, K. Golap, and S. Bhatnagar, "EVLA Memorandum 67 w-Projection: A new algorithm for non-coplanar baselines," Tech. Rep., 2003.

[14] L. Wolz, J. D. McEwen, F. B. Abdalla, R. E. Carrillo, and Y. Wiaux, "Revisiting the spread spectrum effect in radio interferometric imaging: a sparse variant of the w-projection algorithm," *Monthly Notices of the Royal Astronomical Society*, vol. 12, Jul. 2013. [Online]. Available: http://arxiv.org/abs/1307.3424

[15] T. Cornwell, F. Li, and F. de Hoog, "The Application of Compressive Sensing to Radio Astronomy: Deconvolution," *Astronomy and Astrophysics*, 2011. [Online]. Available: http://arxiv.org/abs/1106.1711

[16] B. Osgood, "The Fourier Transform and its applications," 1986. [Online]. Available: https://archive.org/details/ost-engineering-book-fall-07

[17] R. Frater and I. Docherty, "On the reduction of three dimensional interferometer measurements," *Astronomy and Astrophysics*, no. 84, pp. 75–77, 1980. [Online]. Available: http://adsabs.harvard.edu/full/1980A%26A....84...75F

[18] B. Humphreys and T. Cornwell, "SKA Memo 132: Analysis of Convolutional Resampling Algorithm Performance Analysis," Tech. Rep., 2011. [Online]. Available: http://www.skatelescope.org/uploaded/59116_132_Memo_Humphreys.pdf

[19] Y. Wiaux, G. Puy, Y. Boursier, and P. Vandergheynst, "Spread spectrum for imaging techniques in radio interferometry," *Monthly Notices of the Royal Astronomical Society*, vol. 10, no. 2009, Jul. 2009. [Online]. Available: http://arxiv.org/abs/0907.0944

[20] R. Ramirez, *The FFT: Fundamentals and Concepts*, 1st ed. New Jersey: Prentice Hall PTR, 1985.

[21] P. Bloomfield, *Fourier Analysis of Time Series: An Introduction*, 2nd ed. New York: John Wiley & Sons, 2004.

[22] D. Briggs, "High fidelity deconvolution of moderately resolved sources," Ph.D. dissertation, The New Mexico Institute of Mining and Technology, 1995. [Online]. Available: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:High+fidelity+deconvolution+of+moderately+resolved+sources#0

[23] A. Kiliçman and Z. Al-Zhour, "Note on the Numerical Solutions of the General Matrix Convolution Equations by Using the Iterative Methods and Box Convolution Product," *Abstract and Applied Analysis*, vol. 2010, pp. 1–16, 2010. [Online]. Available: http://www.hindawi.com/journals/aaa/2010/106192/

[24] U. Schwartz, "Mathematical-statistical description of the iterative beam removing technique," *Astron. Astrophys*, vol. 65, pp. 345–356, 1978. [Online]. Available: http://adsabs.harvard.edu/abs/1978A%26A....65..345S

[25] D. Steer, P. Dewdney, and M. Ito, "Enhancements to the deconvolution algorithm CLEAN," *Astronomy and Astrophysics*, vol. 137, pp. 159–165, 1984. [Online]. Available: http://adsabs.harvard.edu/full/1984A%26A...137..159S

[26] B. Clark, "An Efficient Implementation of the Algorithm CLEAN," *Astronomy and Astrophysics*, vol. 89, pp. 377–378, 1980. [Online]. Available: http://adsabs.harvard.edu/abs/1980A%26A....89..377C

[27] P. Peebles, J. Read, and P. Read, *Probability, random variables, and random signal principles*, 4th ed. India: McGraw-Hill, 1987.

[28] J. W. Rich, W. J. G. de Blok, T. J. Cornwell, E. Brinks, F. Walter, and et al., "Multi-Scale CLEAN: a Comparison of Its Performance Against Classical CLEAN on Galaxies Using Things," *The Astronomical Journal*, vol. 136, no. 6, pp. 2897–2920, Dec. 2008. [Online]. Available: http://arxiv.org/abs/0810.2825

[29] T. J. Cornwell, "Multiscale CLEAN Deconvolution of Radio Synthesis Images," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 5, pp. 793–801, Oct. 2008. [Online]. Available: http://arxiv.org/abs/0806.2228

[30] R. Sault and M. Wieringa, "Multi-frequency synthesis techniques in radio interferometric imaging." *Astronomy and Astrophysics*, vol. 108, pp. 585–594, 1994. [Online]. Available: http://adsabs.harvard.edu/full/1994A%26AS..108..585S

[31] D. Briggs, R. Braun, and T. Cornwell, "8. Deconvolution," *Synthesis Imaging in Radio Astronomy II*, vol. 180, pp. 151–170, 1999. [Online]. Available: http://adsabs.harvard.edu/full/1999ASPC..180..383P

[32] D. Briggs, F. Schwab, and R. Sramek, "7. Imaging," *Synthesis Imaging in Radio Astronomy II*, vol. 180, pp. 127–149, 1999. [Online]. Available: http://adsabs.harvard.edu/full/1999ASPC..180..383P

[33] A. Guessoum and R. Mersereau, "Fast algorithms for the multidimensional discrete Fourier transform," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 34, pp. 937–943, 1986. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1164883&isnumber=26199