Full Mueller Imaging

Direction Dependent Corrections in Polarimetric Radio Imaging

Preshanth Jagannathan

Dec 2017

Acknowledgements

This thesis is a result of a four-year journey that has taken me across three continents. It would not have been possible but for the support of my supervisors, family and colleagues. At this juncture, I would like to thank a few of them that have been indispensable to my journey. I would like to thank Russ Taylor who has been patient and supportive of me. He has found time to answer my questions and provide guidance no matter the continent or timezone either of us was in. Without him, this thesis and I might not have made it to this juncture.

Another instrumental figure in this journey is Sanjay Bhatnagar, my co-supervisor. He has been kind enough to take me under his wing and guided me through the worlds of radio interferometry and algorithm developments. My interactions with Sanjay have been illuminating to me both scientifically and personally and I am very fortunate to have a friend and mentor in him. He has also spent the better part of the last year poring over these pages to make it more user friendly for the reader. Sanjay, thank you for everything, I am excited about getting back to Socorro and start working with you again.

Many people at NRAO have contributed towards this thesis undertaking, I would especially like to thank my collaborators Urvashi Rau, and Walter Brisken for taking the time to answer my myriad questions no matter how trivial. Kumar Golap and Urvashi Rau have been an encyclopedia of CASA knowledge and without them, my work would have progressed at a much much slower pace. I would like to thank Vivek Dhawan, Rick Perley and Peter Napier for the many holography and primary beam related discussions which have helped in improving my understanding of the instrument. My thanks to Steve Myers for playing devil's advocate throughout, and for raising my morale with his witticisms. Thanks, Huib Intema for the many hours of discussions on interferometry and imaging and for the opportunity to work and learn with him on the TGSS ADR project. I would also like to thank Erik Bryer and James Robnett for their invaluable help with deploying my code on HPC architecture.

My time in NRAO Socorro as a Reber fellow has been rewarding both professionally

and personally. I would like to thank NRAO for supporting me financially for the last three years through the Reber fellowship program. I am especially thankful for the coffee group in Room 300, for the much-needed sustenance in terms of caffeine and company.

My parents and my sister have been a source of consistent support and encouragement, and I want to thank them for it. Ramya, you have been the source of my strength and sanity these past years of my PhD. As my wife and best friend you have helped shore up my moral when I needed it the most and for that I thank you.

All software written for this project has made use of open-source development packages and code within CASA. I would also like to acknowledge the Python software development community, in particular, NumPy, SciPy, Matplotlib and, Seaborn packages which have been used extensively in this thesis.

Contents

1. Introduction

1.1 Motivation

"A very large part of space-time must be investigated, if reliable results are to be obtained."

—Alan Turing

Magnetic fields pervade the universe, spanning a multitude of scales from the dipolar field on Earth, to the largest gravitationally bound structures such as galaxy clusters [1]. The magnetic fields play a vital role in the evolution of these astronomical systems. In addition to the multitude of scales, magnetic fields are present in different astronomical systems of varying strengths. The strongest observed astronomical magnetic fields are in neutron stars with a field strength of $\approx 10^{15}$ G [2], far higher than any man-made fields till date. In stark contrast magnetic fields in the interstellar medium while ubiquitous are only a few μG in field strength. Many fundamental processes in astrophysics have magnetism at their heart, be it cosmic ray particle acceleration, star formation, or the launch of radio galaxy jets, pulsars, etc. One key fundamental process that allows us to detect and characterize cosmic magnetic fields with radio astronomy is the polarization of synchrotron radiation.

Synchrotron radiation is intrinsically polarized broadband continuum radiation emitted by relativistic charged particles accelerated by the presence of magnetic fields. The emissivity of the synchrotron radiation is tied to the magnetic field strength B and the spectral index α (defined such that the flux density $S \propto \nu^{-\alpha}$) such that $\varepsilon \propto B^{1+\alpha}$. Radio continuum observations measuring the total intensity and accurate spectral index measurements are a direct probe of the source magnetic field. While total intensity measurements provide the magnetic field strength, measurements of the fractional polarization in the emission are related to the degree of order in the source magnetic field, over the volume contained

in the angular resolution of the observation. Polarization angle (also known as the electric vector polarization angle) is another measure that can be derived from the linearly polarized emission. The polarization angle of the source is related to the direction of the source magnetic field. The observed polarization angle is affected by Faraday rotation, a phenomenon by which intervening plasmas between the source and the telescope rotate the plane of polarization. Faraday depth ϕ of the source as defined by [3] is,

$$\phi = 812 \int_{L}^{0} n_e B_{||} dl \tag{1.1}$$

where L is the path length along the line of sight in kpc, n_e is the electron density of the plasma in cm^{-3} and $B_{||}$ is the magnetic field component parallel to the line of sight. Faraday rotation can be thought of in terms of refractive optics such that the magnetized plasma has a different refractive index for each of the orthogonal components of the incident polarizations of the radio waves. This would result in the slowing down of one of the polarization inputs with respect to the other, which manifests as a rotation in the observed polarization angle of the radio waves (refer Chen [4] for a detailed derivation).

As shown in Eq.1.1, the Faraday rotation of the incident angle of polarization due to an intervening magnetized plasma is a function of a) the electron density of the intervening plasma and b) the strength and direction of the magnetic field component along the line of sight. These factors collectively define the strength of the RM signal that can be observed. In cases where both the Faraday depth along the line of sight and the electron density of the intervening plasma are available the magnetic field structure along the sight can be gleaned. In practice, because ϕ is an integrated quantity, it is difficult to measure the electron density of various intervening plasmas along the line of sight which means it is not possible to dissociate various components of Faraday depth ϕ in terms of plasma magnetic fields at a known distance between the source and the telescope. In the limiting case of L being the distance of the source the Faraday depth reduces to the source rotation measure (RM) given by,

$$\Omega = \Omega_0 + RM\lambda^2 \tag{1.2}$$

where Ω is the measured angle of polarization, Ω_0 is the true angle of polarization of the source and λ is the wavelength of observation. Brentjens and de Bruyn [5] introduced RM synthesis a technique for the reconstruction of linear polarization as a function of Faraday depth from the measured complex polarized intensity as a function of λ^2 despite limited coverage in λ^2 domain (For a detailed overview of the technique and its implementation refer Sec. 3.3.3.)

Large area spectro-polarimetric surveys have produced grids of RM values to background sources allowing us to study the polarization and depolarization properties of radio sources and consequently their magnetic field properties and the RM signatures of the Galactic foregrounds. Taylor et al. [6] derived the largest catalog of polarized sources producing an RM grid of sources from the NRAO VLA sky survey (NVSS) [7] shown in Fig. 1.1a. They obtained a catalog of 37543 polarized radio sources spanning over 80% of the sky. An all-sky background of RM then allows us to study the intervening magneto-ionic media in a statistical fashion as demonstrated by Oppermann et al. [8] in estimating the median galactic RM map, shown in Fig. 1.1b. Using the estimated median galactic RM Oppermann et al. [9] determined the extragalactic contribution arising from

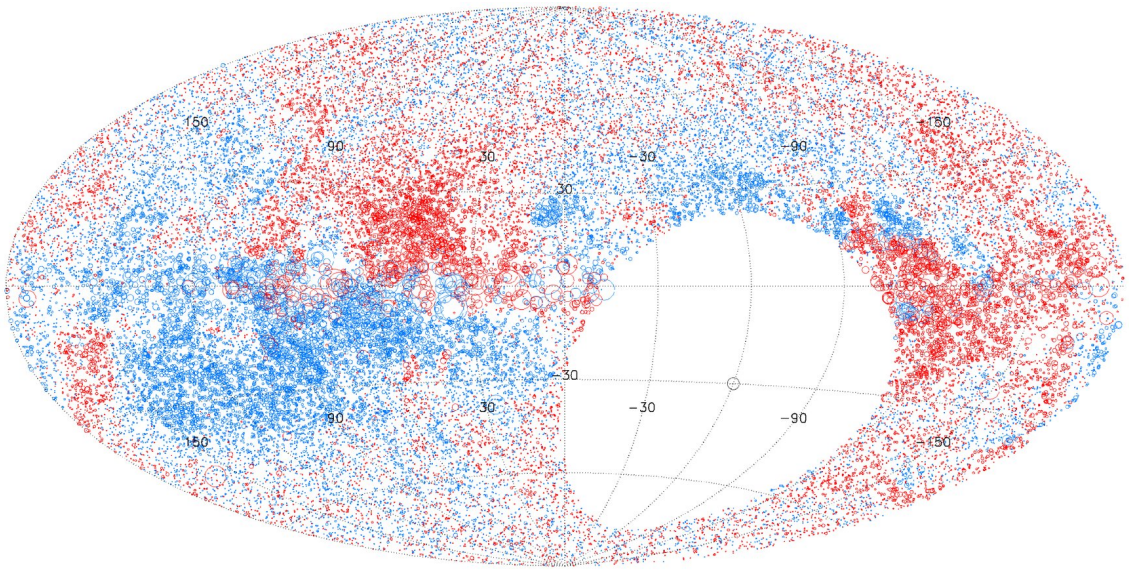
the magneto-ionic media outside our galaxy to a polarized source to be minimal, ≤ 7 rad/m^2. The background RM grid of sources has been used to probe the magnetic field structure of intervening sources from coronal mass ejection of the sun (Kooi et al. [10]), to nebulae(Savage et al. [11]; Purcell et al. [12]; Costa et al. [13]), to our own Galactic Halo (Mao et al. [14]), to galaxies (Stil [15]), and clusters of galaxies (Lawler and Dennison [16]).

Magnetic fields in galaxies also manifest a variety of depolarization mechanisms (differential Faraday rotation, internal Faraday dispersion) (Sokoloff et al. [17]). These depolarization mechanisms have been utilized to study the large scale magnetic fields of nearby galaxies (Heald et al. [18];Stil et al. [19]). In the cases where the radio source displays significant depolarization in addition to polarized emission, it is essential to understand $\Pi(\lambda^2)$, i.e. the fractional polarization (Π) as a function of λ^2 [20]. They show that a source with two spatially unresolved RM components can produce an apparent linear behavior of the polarization position angle as a function of λ^2 a degeneracy broken when examining the fractional polarization.

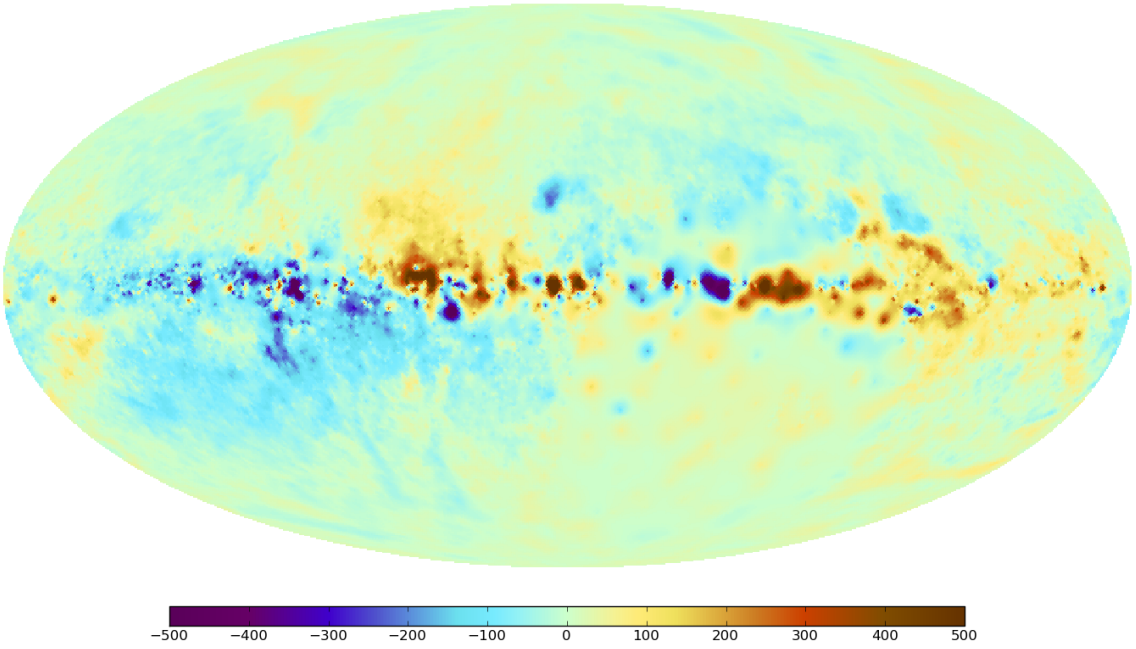
Beyond nearby galaxies, deep radio observations have enabled the study of polarized emission and magnetic field evolution across cosmic time. Observations by Mesa et al. [21] and Taylor et al. [22] indicate an increase in fractional polarization with decreasing flux density for sources down to $500\mu Jy$ in polarized intensity. This suggests a change in either the magnetic field structure or the Faraday screen surrounding these objects, or the emergence of a different class of objects [21]. Deeper observations down to the $10\mu Jy$ by Hales et al. [23] and Rudnick and Owen [24] observe no trend of increasing fractional polarization with flux density. Since the primary limitation of deep fields is the sample size of the sources used in determining the nature of the ensemble a next generation survey that is both deep and wide is required to truly determine the nature of the faintly polarized population.

The next generation of radio polarimetric surveys are a part of a new era of wide-band, wide-area, full Stokes continuum surveys on parabolic antenna interferometer arrays. These surveys will collect data over very large instantaneous bandwidths in the GHz frequency regime. From past surveys [25], it is known that at the minimum a quarter of the observed polarized sources will display complex behavior in Faraday depth (Law et al. [25];O'Sullivan et al. [26]). To accurately reconstruct the true nature of these polarized sources either through RM synthesis [5] or model fitting [26], it is essential to produce high-fidelity, high dynamic range images in all polarization states of the incoming radiation over the full field-of-view of the array antennas. While interferometric arrays record data in all polarizations the recorded data has inter-mixing of polarization states. Reconstruction of pure polarization states of the incoming electromagnetic waves to a high fidelity is an imaging requirement for all next-generation instruments.

Imaging algorithms to date have focused on achieving high fidelity and high dynamic range imaging primarily only in total power. In the absence of high fidelity reconstruction of polarization states across the entire field of view, polarimetry is restricted to sources very close to the antenna pointing centers. While this is a time consuming means of studying individual sources of interest, it limits our ability to survey and map the polarization state of the entire sky efficiently. It is therefore vital to extend the existing algorithmic framework to account for wide-field polarization effects during radio interferometric image reconstruction. The process of extending the existing algorithmic framework to achieve



(a) RM across the sky derived from the NVSS by Taylor et al. [6]. The red and blue colors indicate positive ($B_{||}$ pointing towards observer) and negative ($B_{||}$ pointing away from observer) rotation measures. The size of the circle represents the magnitude of the rotation measure detected.



(b) Galactic RM estimated for the whole sky by Oppermann et al. [8] derived from Taylor et al. [6] primarily.

Figure 1.1: The sky in rotation measure

high fidelity polarimetric imaging forms the crux of this thesis as I will explore in the coming chapters.

1.1.1 Goals

The three primary goals of this dissertation are listed below.

- Evaluate the amount of error incurred in wide-field wide-band imaging using parabolic antenna arrays when not accounting for the direction dependent polarization response of the antenna. These simulations will establish the spatial and spectral variation of the polarization error across wide bandwidths and the entire field of view of the antenna.
- Through simulations evaluate the full Stokes Karl G. Jansky Very Large Array (VLA) antenna primary beam across the wide-bandwidth receivers to understand the spectral behavior of the antenna. Effectively model any frequency-dependent variation of the antenna primary beam in terms of physically motivated parameters utilizing a geometric optics framework. To determine the model parameters for the measured antenna primary beams so that they maybe implemented within the framework of interferometric imaging algorithms that perform corrections on the measured visibilities.
- Implement a full Stokes A-Projection algorithm and determine the viability of the projection algorithmic framework in the reconstruction of the true polarization of the incident radiation. Demonstrate the extent to which the true sky polarization can be imaged across the entire field of view for all frequencies.

1.2 Overview of the thesis

In the forthcoming sections of this chapter, I introduce the basic concepts of polarization. Beginning with the origin of polarized emission in the sky we lay out the framework for the description of the polarization state of electromagnetic waves.

In chapter 2, I introduce the mathematical framework of the interferometric response to incident polarized electromagnetic waves. I review the Jones and Mueller matrix formulations of the interferometric response and the direction independent calibration and imaging framework that enables the transformation of observed visibility data to images of the sky.

In chapter 3 I examine the polarized antenna response as a direction dependent effect and the errors that are introduced into the polarization off-axis due to polarization flux leakage. Through simulations, I demonstrate the effect of polarization flux leakage on rotation measure synthesis.

In chapter 4 I study the spectral variation of the antenna response. I model the frequency dependent behavior through a parameterized geometric optics model of the antenna. I examine the efficacy of the parameterized geometric optics model in capturing the observed spectral dependence of the primary beams as spectral variations of the parameters within the geometric optics model.

In chapter 5 I introduce the full-Mueller A-Projection algorithm and demonstrate the levels to which one can reconstruct the polarization state of the incident electromagnetic waves across the field of view of the antenna. I implement the full-Mueller A-Projection

algorithm within the Common Astronomy Software Application framework and show its performance on measured data from the VLA.

In chapter 6 I examine the performance of the implementation of the full-Mueller A-Projection algorithm, its limitations and the scope of improvements that will be explored in the future.

1.3 Radio Polarimetry

One of the primary mechanisms of radio emission is synchrotron radiation. Synchrotron radiation is the emission of electromagnetic waves by ultra-relativistic electrons that are accelerated by a magnetic field into gyrations. Synchrotron radiation is observed from our galaxy, supernova remnants and external radio galaxies, particularly at the centimeter and meter wavelengths. It is the dominant source of cosmic radiation as low frequency

If we consider a charged particle of mass m_0 and charge ze, moving at a relativistic velocity $\vec{v}$, and Lorentz factor γ in a uniform magnetic field $\vec{B}$, its equation of motion can be written as,

$$\gamma m_0 \frac{d\vec{v}}{dt} = ze(\vec{v} \times \vec{B}). \tag{1.3}$$

One can separate the components of the particle velocity into a parallel and perpendicular component. With only the perpendicular component being altered by the magnetic field, causing the particle to undergo gyrations around the magnetic field lines. The parallel component of the velocity couples to enable a helical motion in the most general case. This accelerated charged particle emits in along the direction of acceleration which is perpendicular to both the velocity vector v and the direction of the magnetic field B. If we consider the case of the charged particle to be an electron we can determine its energy loss rate as,

$$-\frac{dE}{dt} = \frac{\gamma^4 e^2}{6\pi\varepsilon_0 c^3} a_\perp = \frac{e^4 B^2}{6\pi\varepsilon_0 c m_e^2} \frac{v^2}{c^2} \gamma^2 sin^2\theta \tag{1.4}$$

where m_e is the mass of the electron and ε_0 is the emissivity of free space, $a_\perp$ is the perpendicular acceleration vector. The energy loss rate $-\frac{dE}{dt}$ in the ultra-relativistic limit $v \to c$ is converted into synchrotron radiation that is emitted across a wide spectrum of frequencies as electromagnetic waves from high energy X-rays to radio waves. The radiation that is emitted is highly polarized (for a detailed derivation of the polarized nature of the emitted radiation refer Pacholczyk [27]).

Since the incoming electromagnetic wave is the only vector observable there are streams of astronomy dedicated to observing and analyzing the different aspects of the vector. The electromagnetic wave can be completely characterized by knowing the amplitude (photometry), frequency (spectroscopy), the direction of its poynting vector (imaging) and the orientation of the electric and magnetic field vectors with respect to a reference frame (polarimetry).

1.3.1 Polarization ellipse & the Poincaré sphere

Since Maxwell's equations impose the condition that the direction of either field cannot be along the direction of propagation we can determine the polarization state of the wave by

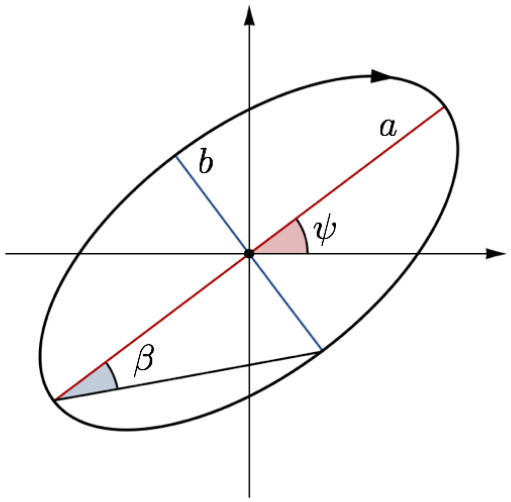


Figure 1.2: The polarization ellipse is shown in the xy plane. Where a and b are the semi-major and semi-minor axes of the ellipse. ψ is the inclination of the semi-major axis of the ellipse with respect to the local x coordinate axis.

looking at either the electric or magnetic field vectors. In radio astronomy, only the voltages induced by the incident electric field on an antenna is measured and so the description henceforth will be focused on the electric field vector of an electromagnetic wave. Let us consider a Cartesian coordinate system that is oriented such that the z-axis is pointed along the line of sight. Then the electric field vector is oscillating in the xz and yz planes while propagating towards us along the z-axis. The equation of electromagnetic waves can then be written as

$$\vec{E} = E_x\hat{x} + E_y\hat{y} = E_{0x}e^{i(2\pi\nu t+\delta_x)}\hat{x} + E_{0y}e^{i(2\pi\nu t+\delta_y)}\hat{y}, \tag{1.5}$$

where $\vec{E}$ is the electric field vector, E_{0x} and E_{0y} are the amplitudes of the electric field components E_x and E_y along the x and y axes. $\hat{x}$ and $\hat{y}$ are the unit vectors along the x and y axes. δ_x and δ_y are the phases of the electric field components. The electric field vector is moving helically in three dimensions whose projection on the xy plane corresponds to the polarization ellipse. The ellipse equation can then be written as,

$$\frac{E_x^2}{E_{x0}^2} + \frac{E_y^2}{E_{y0}^2} - \frac{2E_xE_y\eta}{E_{x0}E_{y0}} = sin^2(\eta), \tag{1.6}$$

where $\eta = \delta_x - \delta_y$ is the phase difference between electric field phases along the x and y axes. Some special cases of the polarization ellipse are of particular interest, $E_{x0} = E_{y0}$ and $\eta = 0$ is when the wave is perfectly linearly polarized. Similarly, when $E_{x0} = E_{y0}$ and $\eta = \pm pi/2$ we get a circularly polarized wave, when the phase is positive the electric field vector rotates in a clockwise direction also known as right circular polarization (RCP) and when the phase is negative the electric field vector rotates in an anti-clockwise direction known as the left circular polarization (LCP). A more compact representation of

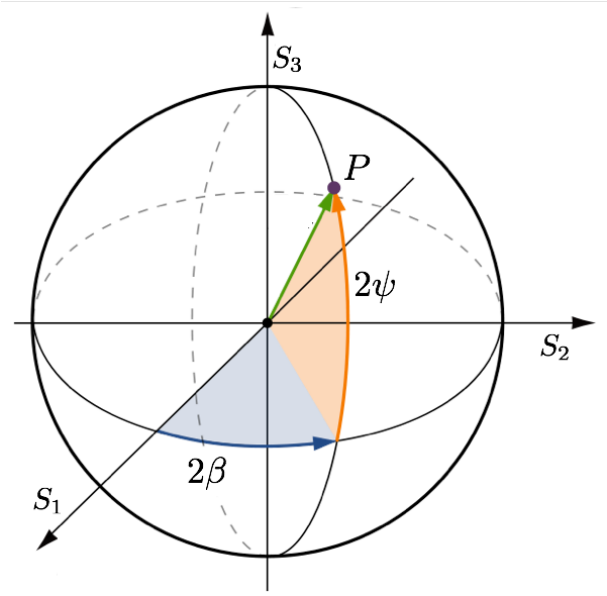


Figure 1.3: The Poincaré sphere of a polarization vector $P(2\psi, 2\beta)$ is shown. The Cartesian axes, S_1, S_2 and S_3, form the Stokes parameter axes, as the projection of the P along each of the axes gives rise to the polarized Stokes parameters.

the polarization state of the electric field vector can be had in the polar coordinates of the polarization ellipse. In the polar form, the polarization ellipse is constrained only by the ratio of the semi-major (a) and semi-minor axes (b) given by $sin(2\beta)$, and the inclination angle of the electric field vector with respect to the x-axis of the coordinate frame $tan(2\psi)$ as shown in Fig. 1.2.

$$tan(2\psi) = \frac{2E_{x0}E_{y0}cos\eta}{E_{x0}^2 - E_{y0}^2} \tag{1.7}$$

$$sin(2\beta) = \frac{2E_{x0}E_{y0}sin\eta}{E_{x0}^2 + E_{y0}^2} \tag{1.8}$$

Henri Poincaré demonstrated that any polarization state can then be represented on a sphere utilizing the ellipticity parameter (2β) as the longitude, the azimuth angle of the polarization ellipse (2ψ) as the latitude, and the amplitude of the electric field vector given by $P = (E_{x0}^2 + E_{y0}^2)$ as the radius of the sphere as shown in fig 1.3. The sphere provides a compact representation of the different polarization states and consequently the different basis in which the electric field vector is represented. The states are defined as follows:

- Linear Polarization : When $2\beta = 0$, and $2\psi = 0$, the electric field vector is horizontally polarized. When$2\beta = 0$ and $2\psi = 180^\circ$ the electric field vector is vertically polarized.
- Circular Polarization : When $2\psi = 90$ for all β the vector is considered to be circularly polarized.
- Elliptical Polarization : All other states of polarization barring the special cases of linear and circular polarization with arbitrary values of ψ and β constitute elliptical polarization.

In the Poincaré sphere the radius vector from the origin to the surface corresponds to the

eigenvector for that particular polarization form. Since the vectors in opposite directions i.e. $\pm\beta$ are orthogonal in theory infinitely many polarization bases exist in which we can choose to define the polarization state of the electric field vector. Any polarization state can then be reconstructed using the projection of the polarization vector along the three axes shown in Fig. 1.3. The projection of the polarization vectors along the x, y and z-axis given by:

$$S_1 = P\,cos\,2\psi\,cos\,2\beta = 2E_{x0}E_{y0}cos\eta \tag{1.9}$$
$$S_2 = P\,sin\,2\psi\,cos\,2\beta = 2E_{x0}E_{y0}sin\eta \tag{1.10}$$
$$S_3 = P\,sin\,2\beta = E_{x0}^2 - E_{y0}^2 \tag{1.11}$$
$$P = \sqrt{S_1^2 + S_2^2 + S_3^3} \tag{1.12}$$

The Cartesian coordinates S_1, S_2 and, S_3 of the Poincaré sphere give rise to the Stokes parameters.

1.3.2 Stokes Parameters

Stokes parameters defined by George Stokes in 1852 and introduced to astronomy by Chandrasekhar is a notation for describing the state of the measured electromagnetic waves. Consider a measurement of integration time t for a monochromatic radiation of frequency ν, its Stokes parameters can be written as

$$I = S_0 = \langle E_{x0}^2 + E_{y0}^2 \rangle_{\nu,t} \tag{1.13}$$
$$Q = S_1 = \langle E_{x0}^2 - E_{y0}^2 \rangle_{\nu,t} \tag{1.14}$$
$$U = S_2 = \langle 2E_{x0}E_{y0}cos\eta \rangle_{\nu,t} \tag{1.15}$$
$$V = S_3 = \langle 2E_{x0}E_{y0}sin\eta \rangle_{\nu,t} \tag{1.16}$$

If the electric field fluctuations are not coherent over the integration time scales then it would result in complete depolarization, i.e. $Q = U = V = 0$. The degree of polarization then changes with the degree of coherence of the incident electromagnetic wave ensemble over its integration time, which leads to the identity

$$I \geq \sqrt{Q^2 + U^2 + V^2} \tag{1.17}$$

The degree of polarization of the incident electromagnetic wave is given by

$$\Pi = \frac{P}{I} = \frac{\sqrt{Q^2 + U^2 + V^2}}{I} \tag{1.18}$$

the total degree of polarization includes both linear and circular polarizations. Most astronomical sources barring pulsars do not show significant amounts of circular polarizations (i.e. $V = 0$) and hence it is convenient to define the degree of linear polarization as a separate parameter Π_L

$$P = \Pi_L\, I\, e^{i\psi} \tag{1.19}$$

where Π_L is the fractional linear polarization and ψ the electric vector polarization angle. Π_L and ψ are defined in terms of the observed Stokes parameters as,

$$\Pi_L = \frac{\sqrt{Q^2+U^2}}{I} \tag{1.20}$$

$$\psi = \frac{1}{2}\tan^{-1}\left(\frac{U}{Q}\right) \tag{1.21}$$

high fidelity measurement of the complex polarization vector P of the incident electromagnetic wave is one of the primary goals of radio polarimetry.

1.3.3 Jones Calculus

While the Stokes parameters are convenient to describe the state of the ensemble of measured electromagnetic waves, it is not a convenient form to describe the transformations that the incident electromagnetic waves have undergone. The Jones vector $\langle E\rangle_{\nu,t}$ is defined as

$$\langle E\rangle_{\nu,t} = \begin{pmatrix}\langle E_{x0}e^{i\eta_x}\rangle \\ \langle E_{y0}e^{i\eta_y}\rangle\end{pmatrix}_{\nu,t} \tag{1.22}$$

The Jones vector is defined by the amplitude and phase of the electric field vector in the x and y directions. Any transformation of the Jones vector can then be represented as a 2×2 Jones matrix. The Jones matrices are then operators that describe the transformations of the Jones vector such as polarization mixing, birefringence etc. In subsequent chapters of this thesis we will revisit the role of the Jones matrix in radio interferometry

1.3.4 Mueller Calculus

The Mueller calculus is a means of manipulating the Stokes vectors themselves to represent transformations of the incident electromagnetic waves. Introduced by Hans Mueller in 1943, it extended the Jones Calculus formalism to include the most general description of an optical transformation of the incident electromagnetic wave. The measured Stokes vector $\vec{S'}$ is the transformation of the emitted Stokes vector $\vec{S}$ by the 4×4 Mueller matrix M.

$$\begin{pmatrix}I'\\Q'\\U'\\V'\end{pmatrix} = \begin{pmatrix}M_{00} & M_{01} & M_{02} & M_{03}\\ M_{10} & M_{11} & M_{12} & M_{13}\\ M_{20} & M_{21} & M_{22} & M_{23}\\ M_{30} & M_{31} & M_{32} & M_{33}\end{pmatrix}\begin{pmatrix}I\\Q\\U\\V\end{pmatrix} \tag{1.23}$$

The Mueller matrix is a transform matrix of the incident polarized electromagnetic waves it does not retain phase information of incident electric field, unlike the Jones matrix which works directly on electric fields themselves. In the case of polarized electromagnetic waves the Jones matrices can be transformed into the Mueller matrix by means of the tensor product operation, $M = J\otimes J^*$, where J^* is the complex conjugate of the Jones matrix J. It is also worthwhile noting that the basis of representation of the incident electromagnetic waves determines the basis of the Mueller matrix that is applied to it. Any orthogonal basis of representation of the Stokes vectors such as circular (R,L, right and left circularly

polarized), or linear basis (X,Y) are equivalent to each other. For a more detailed overview of the matrix transformation from one basis to another is can be found in the Appendix.

Armed with the Jones and Mueller calculi I will apply it in the context of the measurement equation of a radio interferometer, in the next chapter. I will also explore direction independent calibration and the direction dependent nature of polarization leakage as it pertains to the antenna response function also called the antenna primary beam.

2. Radio Interferometry

In order to understand the Full-Mueller A-Projection algorithm, it is vital that we understand the modern formulation of the measurement equation. With that goal in mind, I will revisit some fundamental concepts of interferometry and the framework upon which the simulations and algorithms of the forthcoming chapters are built. The notation used in this chapter follows Hamaker et al. [28] and Rau et al. [29]. The derivation of spatial coherence of an electric field follows chapter 1 of Taylor et al. [30].

2.0.1 Spatial Coherence Function

To image a source with an interferometer the source of interest must be an incoherent emitter, i.e the there is no spatial correlation in the emission from one region of the source to the other. If you consider a source that is a monochromatic emitter of electromagnetic waves and the emission is constant over time, we can write the emission from a direction $\vec{r}$ as the vector $\varepsilon(\vec{r},t) = \varepsilon_\nu(\vec{r})e^{-2\pi i\nu t}$, where $\varepsilon_\nu(\vec{r})$ is the complex amplitude of the electric field vector. Consider the electric field of a quasi-monochromatic EM-wave originating from the source at $\vec{R}$ and is incident on the detector location at $\vec{r}$. Under the assumptions that, the intervening space between the source and detector is empty we can relate the complex amplitude of the electric field at the detector $E_\nu(\vec{r})$ to the EM-wave of frequency ν originating at the source of strength $\varepsilon_\nu(\vec{r})$ through Huygen's principle.

$$E_\nu(\vec{r}) = \int \varepsilon_\nu(\vec{R}) \frac{e^{2\pi i\nu|\vec{R}-\vec{r}|/c}}{|\vec{R}-\vec{r}|} dS \tag{2.1}$$

where S is the projected shape of the source on the two-dimensional celestial sphere and dS is the elemental area of the projected source shape. The spatial coherence of the electric fields incident at two locations $\vec{r}_1$ and $\vec{r}_2$ on the detector originating from two locations $\vec{R}_1$

and $\vec{R}_2$ within the source S can be expressed as follows.

$$\langle E_\nu(\vec{r}_1)E_\nu^*(\vec{r}_2)\rangle = \langle \int\int \varepsilon_\nu(\vec{R}_1)\varepsilon_\nu^*(\vec{R}_2)\frac{e^{2\pi i\nu|\vec{R}_1-\vec{r}_1|/c}}{|\vec{R}_1-\vec{r}_1|}\frac{e^{2\pi i\nu|\vec{R}_2-\vec{r}_2|/c}}{|\vec{R}_2-\vec{r}_2|}dS_1dS_2\rangle \tag{2.2}$$

If the source is an incoherent emitter $\langle \varepsilon_\nu(\vec{r}_1)\varepsilon_\nu^*(\vec{r}_2)\rangle = \langle|\varepsilon_\nu|^2\rangle\delta(\vec{r}_1-\vec{r}_2) = \langle|\varepsilon_\nu(\vec{R})|^2\rangle$, where $\vec{R}_1 = \vec{R}_2 = \vec{R}$. If the distance of the source vastly exceeds the size of detector then $\vec{R} >> \vec{r}_1, \vec{r}_2$ then $\frac{1}{|\vec{R}-\vec{r}_1|}, \frac{1}{|\vec{R}-\vec{r}_2|} \to \frac{1}{|\vec{R}|}$. Each source element dS can be written in terms of the solid angle subtended on the celestial sphere by the source, as seen by the detector as $|R|^2 d\Omega$. If $\hat{s}$ is the unit vector along $\vec{R}$ and we define $I_\nu(s) = \langle|\varepsilon_\nu(s)|^2\rangle$ as the observed intensity of the source at the detector, Eq. 2.2 reduces to

$$\langle E_\nu(\vec{r}_1)E_\nu^*(\vec{r}_2)\rangle = \int I_\nu(s)e^{-2\pi i\nu s.|\vec{r}_1-\vec{r}_2|/c}d\Omega \tag{2.3}$$

The time-averaged spatial coherence function is called the visibility function $V(\vec{r}_1 - \vec{r}_2, \nu) = \langle E_\nu(\vec{r}_1)E_\nu^*(\vec{r}_2)\rangle$. The spatial coherence function or visibility is only a function of the relative separation of the detectors or the baseline vector in an interferometer. Only when the source is an incoherent emitter does the coherence depend on $\vec{r}_1 - \vec{r}_2$ rather than r_1 and r_2.

Coordinate System

To relate the measurement of Eq. 2.3 to an interferometer on the ground we need to recast the equations in a convenient coordinate system. We define a terrestrial coordinate system in Cartesian coordinates x, y, z, that describes the physical location of the antennas of the interferometer, z-axis points to the north celestial pole and the xy plane is the equatorial plane of the earth. Next, let the source of interest be on the surface of the Celestial sphere described by the 3D directional cosines, l, m, n. The unit vector $\hat{s}_0$, is the vector from the origin of the lmn coordinate system perpendicular to the lm plane. The source on the sky has a declination δ_0 and an hour angle H in the xyz coordinate system of the interferometer.

The coordinate system on the celestial sphere be in terms of direction cosines l, m, n, i.e. $I(\hat{s}) = I(l, m, n)$, where $n = \sqrt{1-l^2-m^2}$. So at $\hat{s_0}$, the phase reference center $l, m = 0$ & $n = 1$. All directions other than the phase reference center are of the form $\hat{s} = \hat{s}_0 + \hat{s}_\sigma$, where $\hat{s}_\sigma$ is a vector perpendicular to $\hat{s}_0$ in the plane tangential to the celestial sphere at $\hat{s}_0$. Eq. 2.3 describes the measurement of the intensity distribution in a plane that is tangent to the celestial sphere in the direction of $\hat{s}_0$. Finally, we define the aperture coordinate system represented by u, v, w and measured in units of wavelength $\lambda = c/\nu$. The u, v, w coordinates are defined such that the uv plane called the aperture plane is oriented perpendicular to the source $\hat{s}_0$ such that the w axis points in the direction of $\hat{s}_0$. The u, v, w aperture coordinates are related to x, y, z, the Cartesian coordinates of the antenna through a rotation coordinate transformation.

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \frac{1}{\lambda}\begin{pmatrix} sinH & cosH & 0 \\ -sin\delta_0 cosH & sin\delta_0 sinH & cos\delta_0 \\ cos\delta_0 cosH & -cos\delta_0 sinH & sin\delta_0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{2.4}$$

The antennas move in the uvw coordinate system with changing hour angle H, as the interferometer tracks the source $\hat{s}_0$. Each antenna then traces an ellipse in the uv aperture

plane, this phenomenon is called the earth rotation aperture synthesis. A baseline vector of two antennas given by the 3D vector $\vec{b}(u,v,w) = \vec{r}_1 - \vec{r}_2$. The non zero w value means we have to introduce a time delay $\tau = \vec{b}\hat{s_0}/\nu = (w_1 - w_2)/\nu$ to ensure that the pair of antennas are sampling the same wavefront in the direction $\hat{s_0}$. A point of note is that w is a function of (δ, H) and hence is a direction dependent effect. The delay τ corrects for only the direction of $\hat{s_0}$. Directions in the sky away from the phase center ($\hat{s_0}$) still possess a non zero $w-$ term that needs to be corrected during imaging. Finally, we have a finite number of antennas sampling the aperture plane in uv denoted by $S(u,v)$, the sampling function

$$S(u,v) = \mathbf{W}(u,v)\sum_n \delta(u-u_n)\delta(v-v_n) \tag{2.5}$$

where n is the index of baselines. The sampling is an instantaneous delta function in u,v of the 3D baseline vector. The instantaneous sampling changes with earth rotation and the cumulative sampling $S(u,v)$, represents the spatial frequencies sampled by the interferometer. We can then recast Eq. 2.3 in terms of coordinate system we have created for the aperture and the sky planes.

$$V^{obs}(u,v,w) = S(u,v,w)\int\int I(l,m,n)\frac{e^{-2\pi i(ul+vm+w(n-1))}}{n}dldm \tag{2.6}$$

where $V^{obs}(u,v,w) = V(u,v,w)\cdot S(u,v)$, is the visibilities sampled by the interferometer discretized per baseline. For a coplanar array ($w=0$) or when imaging a small region of the sky such that $l^2+m^2 \simeq 1$ resembles Eq. 2.6 reduces to

$$V^{obs}(u,v) \approx S(u,v,w)\int\int I(l,m)e^{-2\pi i(ul+vm)}dldm \tag{2.7}$$

An interferometer then measures the spatial frequencies corresponding to the Fourier transform of the sky brightness distribution as measured by a pair of antennas. This is known as the van Cittert-Zernike theorem (Thompson et al. [31]), or the measurement equation and is at the heart of interferometry. The image made using the data from all the baselines is a convolution of the Fourier transform of the sampling function ($\mathscr{F}S$) and the sky brightness distribution ($\vec{I}^{sky}$). The effects of the sampling function S are removed as a part of imaging as discussed in subsequent sections.

2.1 Matrix formulation of the measurement equation

For full Stokes imaging, we need to include the full polarization description of the emission from the sky and the full polarization response of the interferometer. Consider a quasi-monochromatic electromagnetic wave emitted by a single point source in the sky, as it propagates towards us the observer, given by ε. We can then write the measured voltage at the feed of antenna a in an interferometric baseline as, $\vec{V} = \mathbf{J}_a \cdot \vec{\varepsilon}$. In terms of the orthogonal feed basis p and q of the antenna a.

$$\vec{V} = \mathbf{J}_a \cdot \vec{\varepsilon} = \mathbf{J}_a \begin{pmatrix} \hat{e}_p \\ \hat{e}_q \end{pmatrix} \tag{2.8}$$

where $\mathbf{J}_a$ is the Jones matrix that contains all the linear transformation occurring affecting the incident EM wave. The feed basis p,q can be replaced with any orthogonal basis such

as the linear X,Y, or circular R,L basis, without any loss of generality. The transformations that the signal might be subjected to along its propagation path can be completely characterized in terms of a 2×2 Jones matrix under the assumptions that the intervening transformations are linear. The transformation falls under two categories, direction independent (DI, such as receiver gains, leakages, etc) and direction dependent (DD gains, such as ionospheric delays, primary beam) We will revisit the effect of the Jones matrices on the signal in the discussion on calibration and imaging.

Coherence Matrix

From the vCZ theorem in Eq. 2.6, the measured visibility is the cross-correlation of the electric field from two antennas. In practice, we sample the incident EM waves at the feed of the antennas at a finite time resolution. As with all signal sampling, we need the sampling rate to be the Nyquist rate or better. The Nyquist rate for sampling for a signal of bandwidth $\Delta\nu$ is given by $1/2\Delta\nu$. The correlation of the signal streams from individual antennas is carried out at the correlator. Correlators cross-correlate the signal streams from the difference antenna baselines mindful of the geometric delay τ. In a FX correlator, the time sampled signals are Fourier transformed (the F portion of correlation) into discrete frequency samples (channelization). The discrete frequencies are multiplied together (X portion of correlation) at the highest time sampling rate after correcting for delay τ, which is a per antenna delay applied prior to the multiplication. This high time resolution data stream is averaged down to the integration times t_{int} and recorded. In a XF correlator the signal from the antennas are first cross-correlated. This is done by introducing a time delay in one of the signal streams (also known as a "lag", hence XF correlators are called lag correlators) and accumulated and averaged over the integration time t_{int}. The time integrated "lagged" cross-correlated signal is then Fourier transformed (F) to obtain the frequency information from the "lag" correlated and accumulated signal. The output post correlation is then a set of visibilities which are discrete samples of the incident electric field.

$$\vec{V}_{ab}^{obs} = \langle \vec{\mathcal{E}}_a^{obs} \otimes \vec{\mathcal{E}}_b^{obs} \rangle_{t_{int}} = \begin{pmatrix} \langle e_a^p e_b^p * \rangle \\ \langle e_a^p e_b^q * \rangle \\ \langle e_a^q e_b^p * \rangle \\ \langle e_a^q e_b^q * \rangle \end{pmatrix}^{obs} \tag{2.9}$$

where a and b are the antennas in a baseline and $\otimes$ is the Kronecker or tensor product of the two vectors. From the van Cittert-Zernike theorem we can write the true sky coherence vector as, $\vec{V}_{ab}^{sky} = \mathscr{F}\vec{I}^{sky}$ where $\mathscr{F}$ is the Fourier transform operator and $\vec{I}^{sky}$ the true sky brightness distribution. Since the Jones matrices encode the transformation of the EM waves during propagation we can utilize 2.8 to write the observed coherence vector $\vec{V}_{ab}^{obs}$ in terms of the true coherence vector $\vec{V}_{ab}^{sky}$ and Jones matrices.

$$\vec{V}_{ab}^{obs} = (\mathbf{J}_a \circledast \mathbf{J}_b^*) \star (\vec{V}_{ab}^{sky} \cdot S) \tag{2.10}$$

where $\circledast$ is the outer convolution operator[1], and S is the sampling function of the interferometer. The reason why the outer convolution operator is required comes from Parse-

[1] In theory, we can write the correlation equation in its entirety as $\vec{V}_{ab}^{obs} = (\mathbf{J}_a \star \mathcal{E}_a^{sky} \circledast \mathbf{J}_b^* \star \mathcal{E}_b^{sky})$, can be written as $(\mathbf{J}_a \circledast \mathbf{J}_b^*) \star (\mathcal{E}_a \circledast \mathcal{E}_b^*)$. The condition for the source being an incoherent emitter reduces $\mathcal{E}_a \circledast \mathcal{E}_b^*$ to $\mathcal{E}_a \otimes \mathcal{E}_b^*$.

val's theorem or the Fourier convolution theorem, which states that $f(l,m)\cdot g(l,m) \xrightarrow{\mathscr{F}} f(u,v)\star g(u,v)$. If we utilize the Jones calculus we can represent the observed apparent sky brightness $\vec{I}^{app}$ as:

$$\vec{I}^{app}(l,m) = \mathbf{M}(l,m)\cdot\vec{I}^{sky} = \mathbf{J}_a(l,m)\otimes\mathbf{J}_b^*(l,m)\cdot(\varepsilon_a\otimes\varepsilon_b^*)^{sky} \tag{2.11}$$

where $\mathbf{M}(l,m)$ is the Mueller matrix altering the polarization state of the incident sky brightness $\vec{I}^{sky}$. Interferometric measurement, of the sky brightness, turns the $\otimes$ in Eq. 2.11 to $\circledast$ in Eq. 2.10. Alternately the measurement equation in matrix form including the corruptions along the signal path is,

$$\vec{V}_{ab}^{obs} \approx \mathbf{M}_{ab}^{DI}\int\int \mathbf{M}_{ab}^{dd}\vec{I}^{sky}(l,m)e^{-2\pi i(ul+vm)}dldm \tag{2.12}$$

where $M_{ab}^{DI} = \mathbf{J}_a\otimes\mathbf{J}_b^*$ are the wide-band descriptions of the DI gains which are typically factored into narrowband gains G, wide-band gains B and polarization leakage D. $\mathbf{M}_{ab}^{dd} = \mathbf{J}_a(l,m)\otimes\mathbf{J}_b^*(l,m)$, are direction dependent(DD) gains which cannot be factored out of the integral, DD effects include eg. antenna response function, differential ionospheric delay. Correction for DI gains is carried out prior to imaging by dividing $\vec{V}_{ab}^{obs}$ by M_{ab}^{DI}. Corrections for M_{ab}^{dd} cannot be corrected for in the same fashion as M_{ab}^{DI} and can only be corrected during imaging as shown by Bhatnagar et al. [32] and Bhatnagar et al. [33].

The Fourier transform relationship between the sky brightness distribution and the measured spatial coherence function allows us to move the Fourier transform of the DD effects out of the integral and write the Eq. 2.10 in terms of a Fourier convolution.

$$\vec{V}_{ab}^{obs}(u,v) = \mathbf{M}_{ab}^{DI}\cdot\left\{\mathbf{A}_{ab}(u,v)\star\int\int\vec{I}^{sky}(l,m)e^{-2\pi i(ul+vm)}dldm\right\} \tag{2.13}$$

where the convolution is a consequence of Parseval's theorem and $\mathbf{A}_{ab} = \mathscr{F}M_{ab}^{dd}$ is the direction dependent aperture matrix. The aperture matrix is a function of time, frequency, and polarization. If we consider that the emission on the sky is discrete and can be pixelized, then we can write the direction integrals in eq 2.13 as a sum.

$$\vec{V}_{ab}^{obs} = \mathbf{M}_{ab}\sum\mathbf{A}_{ab}\star\mathscr{F}\vec{I}^{sky} \tag{2.14}$$

which gives us the discretized measurement equation that underlies most modern interferometric software.

2.2 Calibration

The goal of imaging in interferometry is to accurately reconstruct the sky brightness distribution from the measured visibilities in Eq. 2.14. In order to do that it is necessary to remove the effects introduced in the signal due to the instrument and the atmosphere. From the Eq. 2.14 we know that there are two categories of corruptions present in the measurement, the direction independent gains $\mathbf{M}_{ab}$, and the direction dependent gains $\mathbf{A}_{ab}$. We can break down $\mathbf{M}_{ab}$ into Jones matrices that represent the DI gains per antenna as $\mathbf{M}_{ab} = \mathbf{J}_a\otimes\mathbf{J}_b^*$[2]. Jones DI gains are obtained in stages beginning with the receiver gain

[2] As the gains here are direction independent, they are characterized per 4 unknowns for each antenna for a total of $4N_{ant}$ unknowns to be solved for. The system of equations contains $4N_{ant}(N_{ant}-1)/2$ equations. Also since direction independent gains are only 4 complex numbers per antenna, the $\circledast$ operator simply reduces to an outer product $\otimes$.

solution as a function of time (gain **G**), followed by the bandwidth response of the receiver (bandpass, **B**), and finally the polarization leakage from one polarization to other (**D**). Each of these matrices is a complex 2×2 matrix from which the cumulative Jones matrix is computed as

$$\mathbf{J}_a = \mathbf{G}_a\mathbf{B}_a\mathbf{D}_a = \begin{pmatrix} g_a^p(t) & 0 \\ 0 & g_a^q(t) \end{pmatrix} \begin{pmatrix} g_a^p(\nu) & 0 \\ 0 & g_a^q(\nu) \end{pmatrix} \begin{pmatrix} 1 & d_a^p(\nu,t) \\ -d_a^q(\nu,t) & 1 \end{pmatrix} \tag{2.15}$$

For the sake of simplicity, we can consider a simplified form of the measurement equation containing only direction independent gains, utilizing $\vec{I}^{app} = M_{ab}^{dd}\vec{I}^{sky}$.

$$\vec{V}_{ab}^{obs} = \mathbf{M}_{ab}\mathscr{F}\vec{I}^{app} \tag{2.16}$$

the unknowns in this equations are $\vec{I}^{app}$ and the direction dependent gains $\mathbf{M}_{ab}^{DI}$. For our unknown target source we do not know $\vec{I}^{app}$ and so we can utilize a source we know $\vec{I}^{app}$. For a point source at the phase center $\vec{V}_{ab}^{obs} = \mathbf{J}_a \otimes \mathbf{J}_b^* . |\vec{I}^{app}|$, the visibility is completely characterized by the amplitude of the source, with zero phase, which allows us to compute $\mathbf{J}_a^{-1}$ which when applied to $\vec{V}_{ab}^{obs}$ yields calibrated visibilities. It is important to note that we do not necessarily need a point source to obtain the gain solutions. It is enough if we know the source model *a priori*.

Self Calibration

If we can break down the per baseline gains given by the Mueller matrix $\mathbf{M}_{ab}$ into its antenna Jones vectors, $\mathbf{M}_{ab} = \mathbf{J}_a \otimes \mathbf{J}_b^*$ we now have a system of equations that we can solve using a weighted least square as given by,

$$\chi^2 = \sum_{ab} |\vec{V}_{ab}^{obs} - (\mathbf{J}_a \otimes \mathbf{J}_b^*)\vec{V}_{ab}^{model}| \tag{2.17}$$

where the initial $\vec{V}_{ab}^{model}$ is unknown. Self-calibration [34] is the process of computing the gains of the source of interest utilizing an approximate initial model for $\vec{V}_{ab}^{model}$, generated by imaging $\vec{V}_{ab}^{obs}$. The gain solutions are obtained from Eq. 2.17 and applied before the source is re-imaged to update $\vec{V}_{ab}^{model}$ this process is repeated until χ^2 minimum is achieved.

Both calibration and self-calibration described here utilize only antenna based solutions, which are better constrained in comparison to baseline based solutions which solve for $\mathbf{M}_{ab}$ as a whole, which are prone to lower signal to noise in the solved model and a greater reliance on the initial source model when an *a priori* model is used. Another point of note is that we have only succeeded in calibrating out the DI gains of the instrument by reconstructing $\mathscr{F}\vec{I}^{app}$, any source away from the pointing center still has DD gain corruptions which need to be corrected for before we can retrieve the true sky brightness across the entire field of view given by $\vec{I}^{sky}$.

2.3 Imaging

The symbol conventions in the imaging section of the chapter will closely follow Rau et al. [29]. The derivation of imaging as a generalized minimization problem follows Cornwell [35], Cornwell [36] and Cornwell [37].

2.3.1 Imaging through Optimization

For the remainder of the section, all operators acting on vectors are shown in bold font and the vectors themselves are $\vec{V}^{corr}$ the visibility vector and is related to the sky brightness vector $\vec{I}^{sky}$. The instantaneous visibility vector $\vec{V}^{corr}$ is a $n \times 1$ column vector where each column element contains four Stokes quantities, so the visibility vector for all Stokes parameters is then $4n \times 1$ column vector, where n is the number of baselines of the interferometer. The image vector $\vec{I}^{sky}$ is also a column vector containing m pixels per Stokes visibility to give a $4m \times 1$ image vector for all Stokes parameters. The measurement equation relating $\vec{V}^{corr}$ to $\vec{I}^{sky}$ including direction dependent terms can be written as:

$$\vec{V}^{corr} = \mathbf{WS}\mathscr{F}\mathbf{M}^{dd}\vec{I}^{sky} \tag{2.18}$$

where $\mathbf{M}^{dd}$ is the direction dependent(DD) Mueller matrix acting on the sky brightness vector $\vec{I}^{sky}$ is an operator of shape $4m \times 4m$, $\mathscr{F}$ is the Fourier transform operator[3] of shape $4m \times 4m$, $\mathbf{S}$ the aperture plane sampling or the UV coverage whose operator matrix has the shape $4n \times 4m$ as it samples the m pixel discretized sky image in terms of n baselines and $\mathbf{W}$ is the nominal set of noise based visibility weights of shape $4n \times 4n$. The description of imaging in the section below ignores the effects of the DD terms. We, therefore, set $\mathbf{M}^{dd} = \mathbb{1}$ to describe the general imaging framework and we will revisit DD imaging in later chapters taking into account calibration for $\mathbf{M}^{dd}$.

Since the operators $\mathbf{W}, \mathbf{S}, \mathscr{F}$ are all linear operators, determining the sky brightness $\vec{I}^{sky}$ is an optimization problem. Let $\vec{I}^{sky}_M$ be the model of sky brightness that to be determined. We can define the χ^2 statistic and use it to estimate the error of the fit of the sky model ($\vec{I}^{sky}_M$) given the data (V^{corr}). We can write the χ^2 as ,

$$\chi^2|_k = \vec{R}^{\dagger}|_k \mathbf{W} \vec{R}|_k \tag{2.19}$$

where the residual vector per baseline $a-b$ (of shape $n \times n$) is given by $\vec{R}|_k = \vec{V}^{corr} - \mathbf{WS}\mathscr{F}\vec{I}^{app}_M|_k$, $|_k$ defines the evaluation of the χ^2 computation, and $\vec{I}^{app}_M = \mathbf{M}^{dd}\vec{I}^{sky}_M$ with the parameters at iteration k. $\vec{I}^{app}_M$, is the apparent sky brightness model that is the product of the corruption due to DD gains and the true sky brightness distribution model $\vec{I}^{sky}_M$. We can now attempt to iteratively minimize the χ^2 to estimate the apparent sky brightness $\vec{I}^{app}_M$. The normal equation for the χ^2 can be obtained by equating the derivate of χ^2 to zero.

$$\left.\frac{\partial \chi^2}{\partial \vec{I}^{app}_M}\right|_k = -2\vec{R}^{\dagger}\mathbf{W}\left.\frac{\partial R}{\partial \vec{I}^{sky}_M}\right|_k = -2\mathscr{F}^{\dagger}\mathbf{S}^{\dagger}\mathbf{W}(\vec{V}^{corr} - \mathbf{S}\mathscr{F}\vec{I}^{app}_M|_k) \tag{2.20}$$

Equating Eq: 2.20 to zero, we get

$$\mathscr{F}^{\dagger}\mathbf{S}^{\dagger}\mathbf{WS}\mathscr{F}\vec{I}^{app}_M = \mathscr{F}^{\dagger}\mathbf{S}^{\dagger}\mathbf{W}\vec{V}^{corr} \tag{2.21}$$

as the normal equation. In optimization theory, the matrix operator $\mathscr{F}^{\dagger}\mathbf{S}^{\dagger}\mathbf{WS}\mathscr{F}$ is the Hessian matrix while $\mathscr{F}^{\dagger}\mathbf{S}^{\dagger}\mathbf{W}\vec{V}^{corr}$ is the raw calibrated visibility data transformed into the image domain. The Hessian matrix in radio astronomy represents the point spread function (PSF) of the interferometric array and the derivative of χ^2 represents the residual image. The finite spatial sampling of the aperture plane by the interferometer has two

[3] if $\mathscr{F}$ is the Fourier transform operator, then $\mathscr{F}^{\dagger}$ is the inverse Fourier transform such that $\mathscr{F}^{\dagger}\mathscr{F}X = X$

consequences for the equation, a) It renders the Hessian matrix singular and rules out a simple inversion of the matrix to solve the normal equation. b) It introduces an impulse response function or point spread function into the measurement.

$\vec{I}^{psf}$ is the PSF or the response function in the direction of the image produced by a point source on the sky in that direction. From Eq. 2.18 the PSF of the image can be derived by replacing $\vec{I}^{app}$ with a point source of unit source flux density placed at the phase center $\vec{V}^{corr}_{ab}$ then becomes $\mathbb{1}$ and

$$\vec{I}^{psf} = \mathscr{F}^{\dagger}\mathbf{S}^{\dagger}\mathbf{W}^{\dagger}\mathbb{1}. \tag{2.22}$$

The PSF shape is determined by the weighted UV coverage, and the width of the PSF and consequently, maximum resolution of the interferometric array is $\theta = \frac{1}{B_{max}}$, where B_{max} is the maximum baseline length in wavelengths. The finite sampling means $\vec{I}^{psf}$ as shown in Eq. 2.22 has positive and negative side-lobes due to the missing spacings between the shortest and longest spatial frequencies, and a sharp cut-off beyond the largest sampled spatial frequencies of the interferometer.

Since the Hessian in Eq. 2.21 is not invertible we will have to solve the normal equation using an iterative Newton-Raphson approach to obtain I^{obs}_M.

1. Initialize any prior image model we might have as I^{sky}_M, and in the absence of *a priori* information an empty image model. The residual vector image is defined as

$$\vec{I}^{res}|_k = \mathscr{F}^{\dagger}\mathbf{S}^{\dagger}\mathbf{W}\vec{R}|_k = \mathscr{F}^{\dagger}\mathbf{S}^{\dagger}\mathbf{W}(\vec{V}^{corr} - \mathbf{S}\mathscr{F}\vec{I}^{app}_M|_k) \tag{2.23}$$

   For $n = 0$ and $\vec{I}^{app}_M|_k = 0$, then $\vec{I}^{dirty} = \vec{I}^{res}|_{k=0}$, is the raw image, also called the dirty image of the sky obtained by taking a Fourier transform of the calibrated visibilities. When $\vec{I}^{sky}_M \neq 0$ the computation of the model visibilities as, $\vec{V}^{mod} = \mathbf{S}\mathscr{F}\vec{I}^{app}_M|_k$ is called the *predict* step or *forward* transform used to determine the residuals visibility vector $\vec{V}^{res}|_k = \vec{V}^{corr} - \vec{V}^{mod}|_k$. Equivalently the computation of the residual image $\vec{I}^{res} = \mathscr{F}^{\dagger}\mathbf{S}\mathbf{W}\vec{V}^{res}$ is called the *reverse* transform or the imaging step. The operation of updating the model image and computing the residuals is known as the ***major cycle***. To perform the *forward* and *reverse* transform efficiently Fast Fourier Transform (FFT) is utilized. FFT requires that the data lie on a regular grid We will revisit *gridding* in sec. 2.3.2.

2. We now update the $I^{obs}_M|_k$ based on the derivative of $\frac{\partial\chi^2}{\partial I^{obs}_M}$, as

$$\vec{I}^{app}_M|_k = \vec{I}^{app}_M|_{k-1} + \mathscr{T}\left(\left.\frac{\partial\chi^2}{\partial\vec{I}^M}\right|_{k-1}\right) \tag{2.24}$$

   where $\mathscr{T}$ is the deconvolution operation on the residual image as $\left.\frac{1}{2}\frac{\partial\chi^2}{\partial\vec{I}^M}\right|_{k-1} = I^{res}$. The operator $\mathscr{T}$ deconvolves, I^{psf} from the residual I^{res} to build up our estimate of the sky model and can be written as,

$$\vec{I}^{app}_M|^{i+1}_k = \vec{I}^{app}_M|^{i}_k + \gamma\mathscr{T}(\vec{I}^{res}|_k, \vec{I}^{psf}) \tag{2.25}$$

   where i is the index of iteration for the deconvolution process and γ is the loop gain of the deconvolution process. The model $I^{obs}_M|_k$ is built up by adding γI^{mod}, where

$\vec{I}^{mod}$ is the deconvolved model flux. The loop gain factor γ in a general Newton-Raphson gradient minimization is the second derivate of χ^2 with respect of $\vec{I}^M$, in radio astronomy, however, the second derivative is highly unstable hence a small step size (usually ~ 0.1) is chosen to prevent the image model from diverging from the estimate obtained in the previous iteration. The deconvolution loop is terminated when the peak in $\vec{I}^{res}$ is comparable to the first side-lobe level of the $\vec{I}^{psf}$. Finally the updated model $\vec{I}_M^{app}|_k$ is reconciled with the data via Eq. 2.23, as an *update* step in the optimization process. The process of deconvolution to build up the sky model $\vec{I}_M^{app}|_k$ is called the ***minor cycle***.

3. The optimization is carried out by looping back to step 1 and repeating the process until $\vec{I}^{res}$ is noise-like. Once the stopping criterion is reached the final model $\vec{I}_M^{app}$ is restored by smoothing with the *clean* beam. The *clean* beam is obtained by fitting a 2D Gaussian to the main-lobe of $\vec{I}^{psf}$. The final residual image contains any residual signal that has not been modeled during deconvolution along with the measurement noise. To produce a representative image of the sky the final residual image $\vec{I}^{res}$ is then added to the smoothed final model image $\vec{I}_M^{app}$.

Most modern-day interferometric algorithms working follows the general framework described above. The choice of the operator $\mathscr{T}$ in Eq. 2.25 is a source of distinction between the different variants of the optimization algorithm. CLEAN algorithm and its variants are iterative minimization algorithms following the steepest gradient to find the χ^2 minima. For a more detailed overview of the different deconvolution algorithms refer to Rau et al. [29], section III.

The resulting sky brightness model vector $\vec{I}_M^{app}$ still contains direction dependent effects that are not corrected in the imaging framework described above. The imaging framework described above does not yet correct for DD effects. DD effects during the ***major cycle***, as a convolutional *gridding* correction, was proposed by Bhatnagar et al. [32] and expanded for imaging of wide-band data in Bhatnagar et al. [33]. Before delving into the class of algorithms known as Projection algorithms that correct for DD effects, let us examine the *gridding* and *de-gridding* steps used in imaging.

2.3.2 Gridding and De-gridding

The major cycle of the CLEAN algorithm performs Fourier transform ($\mathscr{F}$ and $\mathscr{F}^\dagger$) operations on the visibility data in the course of the *forward* and *reverse* steps as described in the previous section. These are intensive operations for which the Discrete Fourier Transform (DFT) would take a prohibitive amount of time to complete. A faster numerical operation called the Fast Fourier Transform (FFT) takes only a fraction of the time taken by DFT. FFT requires that the data needs to be resampled on a regularly sampled *uv* grid, an operation called "gridding" in radio astronomy literature. Gridding is then the process of interpolating the weighted visibility data $V^w(u,v)$ to $V^R(u_k,v_k)$, where u_k and v_k are the centers of the regular grid. This interpolation is done by means of a convolution operation wherein a convolution function C_F is convolved with the weighted visibilities and it is resampled at each (u_k,v_k). In theory, we can use this method of interpolation to span all the k grid points, in practice, we only allow a finite support for the convolution function. The support of the convolution function is defined as the number of *uv* grid cells it spans when centered at $V^w(u,v)$. The support sized consequently defines the scale of this convolutional

gridding. We can write the gridded visibilities $\vec{V}^R$ as:

$$V^R(u_k, v_k) = \text{Ш}(u/\Delta u, v\Delta v)(C_F \star V^w(u,v)) \tag{2.26}$$

where Ш is the resampling function known as the 'sha' function (chapter 5, Bracewell [38])and Δu and Δv the size of the uv-grid cell. The Ш function has the unique property of being discrete delta functions in the aperture plane and in the image sky plane, i.e

$$(\mathscr{F}\text{Ш}(u/\Delta u, v\Delta v))(l,m) = \Delta u \Delta v \text{Ш}(l/\Delta l, m\Delta m). \tag{2.27}$$

$\Delta\theta_l, \Delta\theta_m$ is the angular span of each pixel respectively, and let the image be made of N_l and N_m. If N_s is the sampling required to reconstruct the resolution element or PSF of the image, then $N_s \geq 2$(from Nyquist sampling theorem). The size of the *uv* grid cell in units of wavelength in terms of the pixels on the sky and the PSF sampling is

$$\Delta u = \frac{1}{N_s N_l \Delta\theta_l} \quad \Delta v = \frac{1}{N_s N_m \Delta\theta_m}. \tag{2.28}$$

In practice, the PSF is sampled by at least 3 pixels along l and m ($N_s \geq 3$). From the resampled visibilities in Eq. 2.26 we can compute the residual in Eq. 2.23 using the gridded visibilities as

$$\vec{I}^{res}|_k = \vec{I}_{CF}^{-1} \mathscr{F}^\dagger \mathbf{G}_{CF}^\dagger \mathbf{S}^\dagger \mathbf{W} \vec{V}^{res}|_k \tag{2.29}$$

where $\mathbf{G}_{CF}^\dagger$ is the convolutional resampling operator, and $\vec{I}_{CF} = \mathscr{F}^\dagger \mathbf{G}_{CF}$ is the normalization for the convolutional operator in the image plane. An appropriate choice of the functional form to avoid aliasing would result in an image normalization factor $\vec{I}_{CF}$ that drops off very rapidly beyond the edges of the image. The prolate spheroidal function fits the requirements of an anti-aliasing convolution function with an image normalization factor that suppresses signal at $> 10^-3$ of the peak beyond the region of interest [39]. The convolutional gridding can be represented by a diagonal matrix operator[4] $\mathbf{P_s}$ which can be written as $\mathbf{G}_{CF}^{ps} = (\mathscr{F}^\dagger \mathbf{P_s})\mathscr{F}$, and $\vec{I}_{PS} = \mathscr{F}^\dagger \mathbf{P_s}$ is the corresponding image plane normalization of the $\mathbf{P_s}$ operator. In the case of the prolate spheroidal function a real-valued convolution function, $\mathbf{G}_{CF}^{ps\dagger} = \mathbf{G}_{CF}^{ps}$, trivially. For DD gain corrections, however, the convolution functions are complex where the order of operators are significant and is discussed in the next sub-section.

The *reverse* transform follows the Eq. 2.29 while the *forward* transform to model visibilities required to compute the residual vector $\vec{R}$ is modified to include the gridding operator as:

$$\vec{V}^{mod}|_k = \mathbf{S}\mathbf{G}_{CF}^{ps} \mathscr{F} \vec{I}_{CF}^{-1} \vec{I}_M^{app}|_k \tag{2.30}$$

where $\vec{I}_{CF} = \vec{I}_{PS}$.

It is worthwhile remembering that the visibility data being gridded here has the embedded effects of $\mathbf{M}^{dd}$ as shown in Eq. 2.18. Convolutional gridding with an appropriate convolution function offers the means to correct for DD gains. In the following section, I will examine a broad class of imaging algorithms that account for DD imaging during the convolutional gridding step as a part of the imaging major cycle.

[4]The convolutional matrix operator is known as a Toeplitz matrix. Any discrete convolution operation can be represented in terms of a Toeplitz matrix product. The structure of the Toeplitz matrix T in terms of the elements is such that $T_{i,j} = T_{i+1,j+1} = t_{i-j}$, where i, j are the row and column indices respectively and t is an arbitrary element of the Toeplitz matrix T.

2.4 Direction Dependent Gains and Projection Algorithms

The Eq. 2.18 have corruptions due to DD gains in the form of $\mathbf{M}^{dd}$. The gains can be introduced during the signal transmission (differential ionospheric delay) or can be introduced into the measurement due to instrumental effects (non-coplanarity of the array, antenna far field voltage pattern(VP)) are multiplicative in the image plane (as in Eq. 2.18), or a convolution in the aperture plane. $\mathbf{M}^{dd}$ is the DD gains as a function of time, frequency and, polarization, and is the 4×4 Mueller matrix per direction on the sky. The DD gains in the aperture plane $\mathbf{G}^{dd} = \mathscr{F}\mathbf{M}^{dd}\mathscr{F}^{\dagger}$ can be written as a convolution operator. From Eq. 2.18, the observed visibilities including DD gains in the aperture plane in terms of the true sky coherence function can be written as:

$$\vec{V}^{obs} = \mathbf{W}\mathbf{S}\mathbf{G}^{dd}\vec{V}^{sky}. \tag{2.31}$$

If the DD gains $\mathbf{G}^{dd}$ acting on the true sky coherence matrix has finite support in aperture (restricted to few uv cells per visibility), and the $\mathbf{G}^{dd}$ matrix is approximately unitary then the reverse step in the major cycle to get the true model visibilities including DD gains is given:

$$\vec{V}^{mod}|_k = \mathbf{S}\mathbf{G}^{ps}_{CF}\mathbf{G}^{dd}_{CF}\mathscr{F}I^{-1}_{dd}\vec{I}^{-1}_{CF}\vec{I}^{app}_{M}|_k. \tag{2.32}$$

where $\vec{I}^{app}_{M} = \mathbf{M}^{dd}\vec{I}^{sky}_{M}$ is the apparent sky brightness model in terms of the DD Mueller matrix and the true sky brightness model. Algorithms that perform convolutional gridding corrections for DD gains relying on the approximately unitary[5] nature of the DD gains in the aperture plane are called *projection* algorithms. We can rewrite the residual image from Eq. 2.32 to include the DD gridding convolution function $\mathbf{G}^{dd\dagger}_{CF}$ as:

$$\vec{I}^{res}|_k = \vec{I}^{-1}_{dd}\vec{I}^{-1}_{CF}\mathscr{F}^{\dagger}\mathbf{G}^{ps}_{CF}\mathbf{G}^{dd\dagger}_{CF}\mathbf{S}^{\dagger}\mathbf{W}\vec{V}^{res}|_k. \tag{2.33}$$

where $\mathbf{G}^{dd\dagger}_{CF}\mathbf{G}^{dd}_{CF} \approx C\cdot\mathbb{1}$, is a time and frequency independent constant, and I^{-1}_{dd} is the weighted image plane normalization of the DD convolution function, given by $I_{dd} = \mathscr{F}^{\dagger}\mathbf{G}^{dd\dagger}_{CF}\mathbf{W}\mathbf{G}^{dd}_{CF}\mathscr{F}$ and $\vec{V}^{res}|_{k=0} = \vec{V}^{obs}$. Eq. 2.33 is the forward step producing the residual image for PSF deconvolution. In the forthcoming section, we will explore the two forms of the $\mathbf{G}^{dd}_{CF}$, the effect of non-coplanar arrays and the role of the antenna far field power pattern (also known as the primary beam (PB)).

2.4.1 W-Projection

Let us revisit the van Cittert-Zernike equation given in Eq. 2.6. We made the simplifying assumption that for a coplanar array $w = 0$. The additional term $e^{-2\pi iw(\sqrt{1-l^2-m^2}-1)}$ known as the w-term is a significant DD effect that increases as we move away from the phase center. The w-term is significant when imaging very wide fields of view, where the flat sky approximation is not valid $\sqrt{1-l^2-m^2} \neq 0$, or when the baselines are non-coplanar, i.e $w \neq 0$. The appropriate choice of $\mathbf{G}^{dd}_{CF}$ during gridding can ensure that the gridded visibilities and the residual image, are free from the effects of the w-term in the measurement. The choice of $\mathbf{G}^{dd}_{CF} = e^{-2\pi iw(\sqrt{1-l^2-m^2}-1)}$ in the *reverse* step and its

[5] It can be shown that for non conserving operators represented by a normal (not unitary) matrix operator can also be projected out in theory.

conjugate during the forward step allows for the removal of the DD effects of the *w*-term. This convolutional gridding method of removing *w*-term was proposed by Cornwell et al. [40] and has been shown to be significantly faster computationally than the faceted imaging approach. Variants of this basic approach (*w*-stacking, [41], *w*-snapshots [42], etc) are optimizations of the w-projection algorithm for arrays with large *w*-term (MWA, LOFAR, etc).

2.4.2 Antenna Primary Beam

The antenna far field power pattern is also a DD effect. The DD Mueller Matrix ($\mathbf{M}_{dd}$) from Eq. 2.18, varies as a function of time, frequency and polarization. The form and choice of the convolution function ($\mathbf{G}_{CF}^{dd}$) allows us to account for the full polarization response of the antenna PB is the full polarization baseline aperture illumination pattern (AIP). $\mathbf{A}_a$ and $\mathbf{A}_b$ are the AIP of antennas a and b then, $\mathbf{A}_a = \mathscr{F}\mathbf{J}_a$ and $\mathbf{A}_b = \mathbf{J}_b$, describes the Fourier transform relationship between the DD Jones matrix (the antenna PB in this case). Similarly, the baseline aperture illumination function $\mathbf{A}_{ab} = \mathbf{A}_a \circledast \mathbf{A}_b^* = \mathscr{F}\mathbf{M}_{dd} = \mathscr{F}\mathbf{J}_a \otimes \mathscr{F}\mathbf{J}_b$, where $\circledast$ is the outer convolution operator.

Bhatnagar et al. [32] proposed the approximately unitary nature of the antenna AIP ($\mathbf{A}_a$). Bhatnagar et al. [32] used the approximately unitary nature of the antenna AIP to correct for the time and direction dependence in Stokes -I and -V of the VLA antenna for narrowbandwidths. By choosing the gridding convolution function($\mathbf{G}_{CF}^{dd} = adj(\mathbf{A}) \approx \mathbf{A}^\dagger$) such that the predicted model visibilities in the reverse step as in Eq. 2.32 is recast as

$$\vec{V}^{mod}|_k = \mathbf{S}\mathbf{G}_{CF}^{ps}\mathbf{A}^\dagger \mathscr{F} I_{dd}^{-1}\vec{I}^{app} \tag{2.34}$$

where $\mathbf{A}^\dagger = \mathbf{A}_a^\dagger \circledast \mathbf{A}_b^\dagger$ is the adjoint (and the complex conjugate) of the baseline AIP $\mathbf{A}$. The accumulated model visibilities are now independent of the DD $\mathbf{A}$ term and is used in forward step to compute the residual image for deconvolution as,

$$\vec{I}^{res}|_k = \vec{I}_{dd}^{-1}\vec{I}_{CF}^{-1}\mathscr{F}^\dagger \mathbf{G}_{CF}^{ps}\mathbf{A}^\dagger\mathbf{S}^\dagger\mathbf{W}\vec{V}^{res}|_k, \tag{2.35}$$

where $\vec{V}^{res}|_k = \vec{V}^{obs} - \vec{V}^{mod}|_k$, and,

$$\vec{I}_{dd}^{-1} = det(\mathscr{F}^\dagger \sum_t \mathbf{A}) \tag{2.36}$$

is the image domain normalization is the Fourier transform of the determinant of the antenna AIP $\mathbf{A}$ averaged over all time. The $\mathbf{A}$-projection framework depends on the individual antenna AIP, and is applicable even in the case of heterogeneous arrays involving different types of antennas in the array such as ALMA.

$\mathbf{A}$-Projection was extended to account for the frequency dependence of the antenna AIP by Bhatnagar et al. [33]. They defined $\mathbf{A}^\dagger(\nu^*)$ as the wide-band gridding operator used in Eq. 2.34, where ν^* is called the conjugate frequency, and is chosen such that

$$\mathscr{F}^\dagger\mathbf{A}(\nu*)\mathscr{F}^\dagger\mathbf{A}(\nu) = \mathscr{F}^\dagger\mathbf{A}(\nu_{ref}) \tag{2.37}$$

where, ν_{ref} is the reference frequency at which the wide-band continuum image is being produced. This allows for the use of wide-band deconvolution algorithms in the construction of the image model such as the MTMFS, (Rau and Cornwell [43]). The wide-band A-Projection algorithm till date utilizes only two of the sixteen terms in $\mathbf{A}$ matrix.

This thesis extends the algorithmic framework described above to account for the full polarization response of the antenna AIP in time and frequency. In chapter 3 we demonstrate the errors induced due to the uncorrected PB polarization response. In chapter 4 we examine the frequency dependence of the antenna AIP and show that modeling the AIP by means of ray-tracing offers a means of generating the convolution function($\mathbf{G}_{CF}^{dd}$). In chapter 5 we demonstrate the approximately unitary nature of the $\mathbf{A}$ matrix and use it to correct for the full Mueller DD gains.

3. Direction Dependent Effects

3.1 Introduction

The PB response of radio antennas varies with both direction and frequency. For altitude-azimuth mounted antennas the sky brightness distribution rotates with respect to the antenna primary beam as a function of the antenna parallactic angle. Consequently, for long integration observations during which the parallactic angle changes, the response of the array to a radio source includes an instrumental component that varies with time, frequency and polarization. These variations corrupt the wide-field image by introducing image artifacts that are not removed by standard image deconvolution approaches, thereby limiting dynamic range and fidelity. The latter effect has been shown to be particularly important for the polarization response (Jagannathan et al. [44]).

In this chapter, I examine the effects of wide-band DD errors on full-Stokes imaging performance based on ray-trace models of the JVLA L-band full-Stokes PB response. In chapter 5, I will assess the efficacy of the full-Mueller wide-band A-Projection algorithm in correcting for the DD effects. In this chapter using simulations of the sky brightness distribution, I demonstrate the limits of imaging fidelity that can be achieved with classical calibration and imaging, and the levels at which the full-Mueller A-Projection algorithm becomes necessary to enable high-fidelity and -dynamic full-Stokes imaging of wide-band observations.

3.2 Primary Beam as a Direction Dependent Effect

The continuous visibility full-polarization vector-field $\vec{V}_{ab} = \mathscr{F}\vec{I}$ is sampled by the interferometer via $\mathbf{A}_{ab}$ as in Eq. 2.31 to give $\vec{V}_{ab}^{obs}$. $\mathbf{A}_{ab}$ terms introduce the effects of the antenna far-field pattern in the observed data, which if ignored, limits the imaging performance of the instrument (Rau et al. [45]). The $\mathbf{A}_{ab}$ varies with frequency, time (due to relative rotation of the sky with for Alt-Az mount antennas as, or as time-dependent antenna

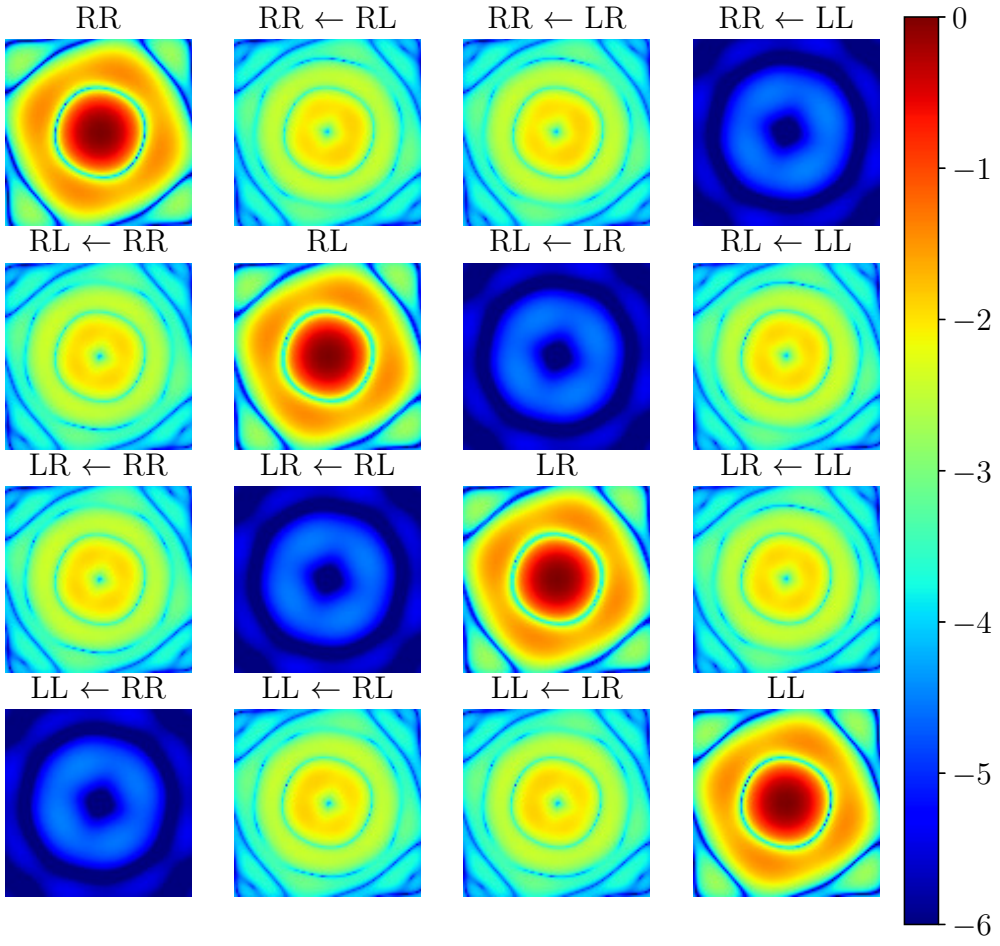


Figure 3.1: Amplitude of the image domain complex Mueller Matrix for an arbitrary baseline $a-b$ at the VLA at 1.5GHz. The color scale is a logarithmic scale spanning nine orders of magnitude. The Mueller elements are computed using the normalized Jones matrix for the diagonal elements to have a peak value of unity along the pointing direction. The direction of the arrows have been reversed from [44] to be more in line with the vector notation in [28].

pointing errors, structural deformations, etc.) and polarization. The dominant variation of $\mathbf{A}_{ab}$ with time for an Alt-Az mounted antenna is given by the change in parallactic angle χ, defined as the angle between the source's hour circle and the great circle passing through the source and the zenith as viewed by the antenna. For a source at declination δ the parallactic angle changes with hour angle h, as

$$\chi = \tan^{-1}\left(\frac{\cos L \cdot \sin h}{\sin L \cdot \cos\delta - \cos L \cdot \sin\delta \cdot \cos h}\right) \tag{3.1}$$

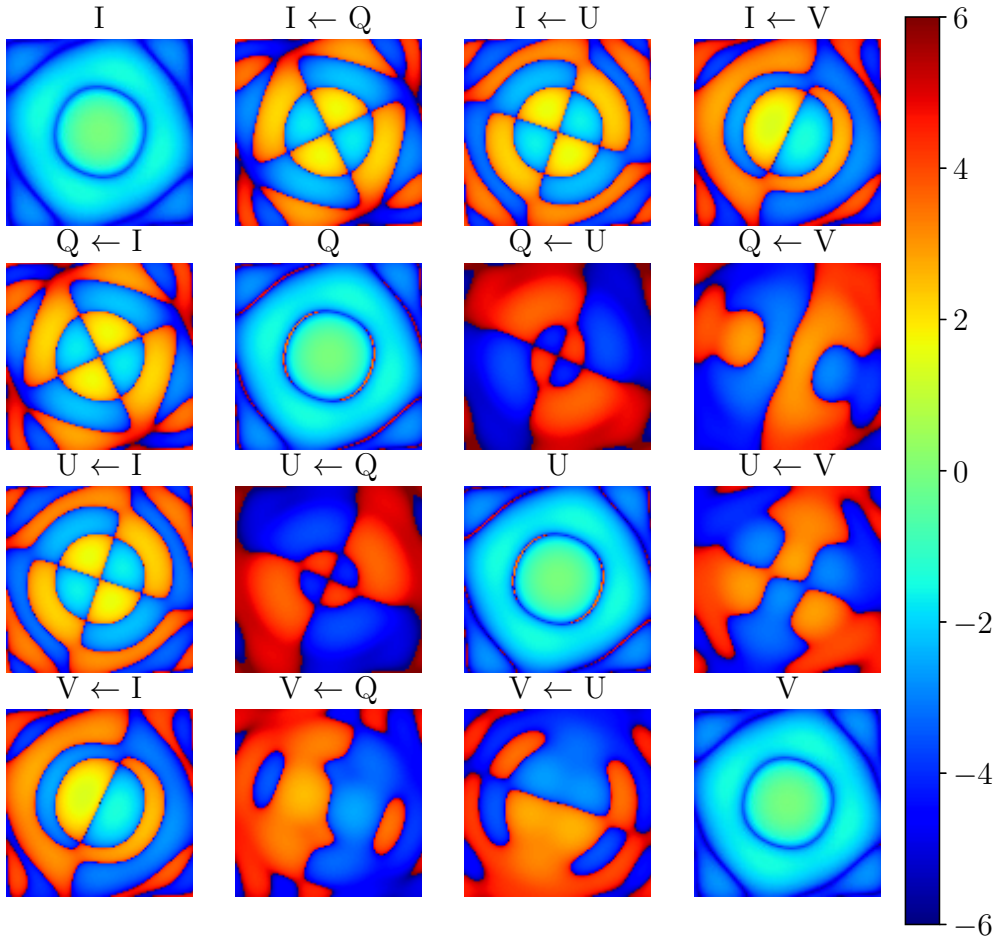


Figure 3.2: The direction dependent Mueller in Stokes basis at 1.5 GHz. The images show $sign(M_{ab}) \times (log_{10}(abs\,(M_{ab}))$ and the color scales is shown as the exponent of 10. The leading order terms are along the diagonal. The `I<-V` and `V<-I` term (due to polarization squint) are ~ 2 orders of magnitude lower than the diagonal. The `I<-Q` and `I<-U` terms (due to off-axis beam polarization) are $\sim 2.5-3$ orders of magnitude lower than the diagonal terms.

where, L is the Geographical latitude of the antenna. The WB-AWP algorithm corrects for this parallactic angle-, direction- and frequency-dependence of $\mathbf{A}$ for Stokes-I and -V imaging.

In the sections below I characterize the time and frequency-dependent errors introduced into Stokes Q and U for circularly polarized feeds from an analysis including the effects of the off-diagonal terms of the Mueller matrix. I perform a suite of simulations that characterize the effects on a polarized and unpolarized off-axis source. I show that corrections for the time- and frequency-dependent PB in all polarization including the effects of polarization leakages will be necessary (e.g. using a full-Stokes version of the

WB-AWP algorithm) for accurate reconstruction of the true full-Stokes flux density vector in wide-field, wide-band imaging. In chapter 5 I will lay out the framework and workings of the full-Stokes WB-AWP algorithm that corrects for the entire Mueller matrix during image reconstruction.

The measurement equation in the image-plane can be cast in Stokes basis as,

$$\vec{I}^{M} = \sum_{k} \mathbf{M}^{k} \cdot \vec{I} \tag{3.2}$$

where k is an index over time, frequency, and baseline. The vector $\vec{I}$ is the true incident full-Stokes sky brightness distribution. $\vec{I}^{M}$ is the measured apparent sky brightness distribution. $\mathbf{M}^{k}$ encodes the response of the interferometer to the incident polarization vector. Expanding $\mathbf{M}^{k}$ in Eq. 3.2 shows explicitly the dependence on the incident polarization.

$$\vec{I}^{M} = \sum_{k} \begin{pmatrix} M^{k}_{\mathtt{II}} I + M^{k}_{\mathtt{IQ}} Q + M^{k}_{\mathtt{IU}} U + M^{k}_{\mathtt{IV}} V \\ M^{k}_{\mathtt{QI}} I + M^{k}_{\mathtt{QQ}} Q + M^{k}_{\mathtt{QU}} U + M^{k}_{\mathtt{QV}} V \\ M^{k}_{\mathtt{UI}} I + M^{k}_{\mathtt{UQ}} Q + M^{k}_{\mathtt{UU}} U + M^{k}_{\mathtt{UV}} V \\ M^{k}_{\mathtt{VI}} I + M^{k}_{\mathtt{VQ}} Q + M^{k}_{\mathtt{VU}} U + M^{k}_{\mathtt{VV}} V \end{pmatrix} \tag{3.3}$$

The diagonal of the Mueller matrix encodes the response of the interferometer pair to each individual Stokes parameter. The off-diagonal elements encode the leakage between Stokes parameters. M^{k}_{IQ} and M^{k}_{IU} encodes the amount of flux leaking from I into Q and U respectively. Magnitude of leakage i.e. the fraction of flux leakage is $\sim 5 \times 10^{-2}$ of I for the VLA at L-band at the half power point (refer, Fig. 3.2), and grows with increasing distance from the beam center. The typical linear polarized intensity of astrophysical sources is at the level of a few percent of Stokes-I flux. Hence, the flux leakage results in a fractional error of up to 100% in the wide-band full-Stokes measurements of typical astrophysical signals. It is noteworthy that the Mueller matrix element M^{k}_{QI} representing the mutual coupling between Stokes Q and I, has exactly the same magnitude as M^{k}_{IQ}, within calibration errors. However, the amount of flux leakage is modulated by the intensity of the Stokes parameters. The amount of flux leaking from Stokes Q into I is $M^{k}_{IQ} \cdot Q$. For a typical linear polarization of a few %, the magnitude of flux leaking from Stokes Q into I is a $10^{-4} - 10^{-5}$, which will be of a concern only for very high dynamic range imaging. However, in cases of highly polarized emission such as may be seen in extended radio jets of some radio galaxies, leakage from Q to I of order 10^{-3} may occur in uncorrected images.

One can isolate the effects of leakage by recasting Eq. 3.2 into a sum of the diagonal and off-diagonal elements,

$$\vec{I}^{M} = \sum_{a,\mathtt{V},t} \left(M_{aa} \cdot I_{a} + \sum_{b, a \neq b} M_{ab} I_{b} \right) \tag{3.4}$$

The first term is the direction-dependent PB responses in each Stokes parameter. The second term isolates the direction-dependent leakage between Stokes flux leakage, which for linear polarization is dominated by the direction-dependent Stokes I leakage. Note that for the time-variable component, all terms can be ignored and the second term is represented as a constant *only* for very short "snapshot" observations with Alt-Az antennas.

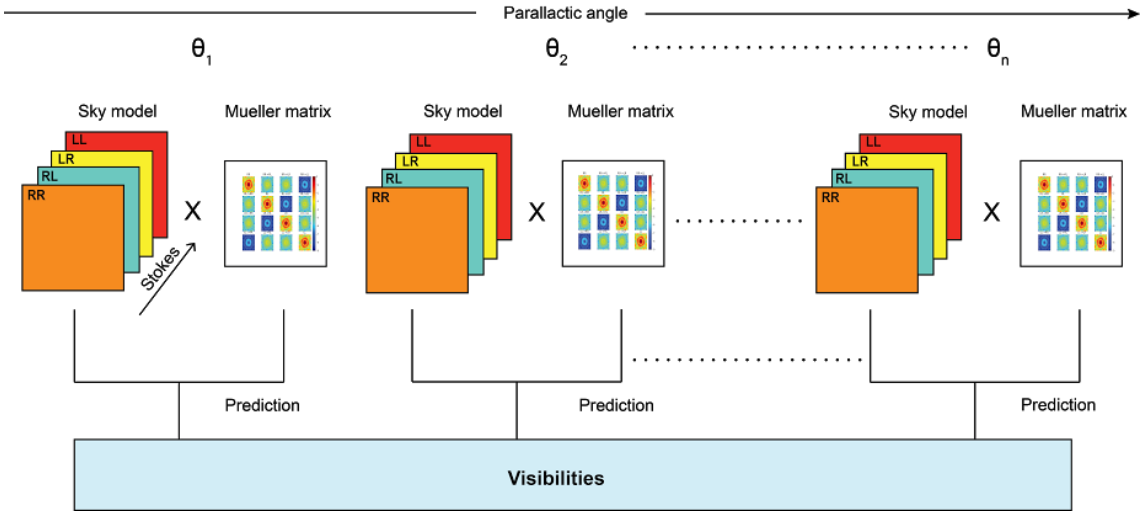


Figure 3.3: The imaging simulations followed the methodology outlined in the flow diagram. The sky model image was multiplied by the Fourier transform of the Mueller matrix for a change in parallactic angle of 5°

This approach was used for the NVSS (Condon et al. [7]). Observations with equatorial mount antennas or antennas with the third axis of motion to maintain a fixed parallactic angle (McConnell et al. [46]) allow for a simple correction in the form of a direction-dependent flux subtraction post imaging. This technique was used to good effect for the Canadian Galactic Plane Survey (Taylor et al. [47]) using the equatorial mount antennas of the Dominion Radio Astronomy Observatory synthesis radio telescope. For observations covering a large parallactic range with Alt-Az antennas, these effects must be corrected as part of the deconvolution and imaging stage.

3.3 Simulations

To examine the characteristics of the errors in polarized signals and their effect on astrophysical measurements I carried out a suite of simulations. In each case, I consider a simple point source located at an off-axis point in the PB. This simple model allows for an intuitive understanding and interpretation of the resulting PB effects. Three cases were considered. In the first case I consider the effect on an unpolarized point source. I then consider two polarized sources, one with zero Faraday Rotation Measure (RM) and the second with an RM of 100 rad m^{-2}. These cases allow us to explore the effects of the frequency dependence of the Mueller terms on Faraday Rotation synthesis. The simulations in all these cases utilized the ray-traced PBs using the implementation in the CASA R&D code base.

I simulated observations over an hour angle range of -4 hr to $+4$ hr and spanning 1 GHz bandwidth in two frequency resolutions, the first with eight 128 MHz-wide spectral windows across the VLA L-Band (1-2 GHz) to understand continuum polarization imaging fidelity. The second with 64 spectral windows with a bandwidth of 16 MHz each to understand the effects of antenna PB on Faraday rotation synthesis. The sources were at Declination $+70°$. An image domain sky model was created following Eq. 3.2. The sky model was then used to create visibilities given by the Eq. 2.31 for the VLA in C-

Configuration. The multiplication of the Mueller Matrix with the sky models was carried out in a standalone Python routine and the prediction from the image plane to model VLA visibilities was carried out within the CASA framework. A block diagram of the simulation framework is provided in Fig. 3.3.

3.3.1 An Off-Axis Point Source as viewed by a single interferometric baseline

I examined the effect of the PB off-axis through two simulations. The first is an unpolarized point source of 1 Jy flux density located at the half power of the PB at the reference frequency of 1.5 GHz. The second simulation is of a polarized point source at the same location with 1 Jy flux density in total intensity and a frequency-independent fractional linear polarization (Π_L) of 5% with polarization position angle (EVPA) of 22.5 degrees.

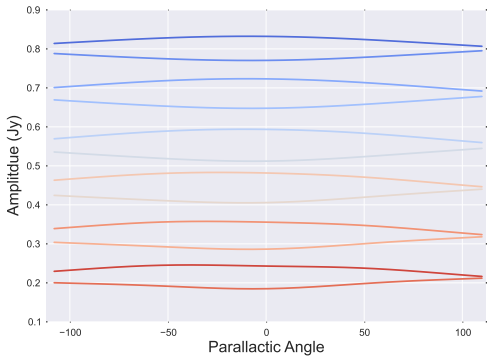


Figure 3.4: The visibility amplitude of RR (convex curves) and LL (concave curves) of an unpolarized source situated at the half-power point at the reference frequency 1.5 GHz. The RR and LL show different amplitudes due to the squint of the VLA antennas. The color axis represents 6 spectral windows, spanning from 1 GHz (blue) to 2 GHz (red) in steps of 0.2 GHz

.

Fig 3.4 shows the RR and LL amplitudes for the unpolarized source at the half power point at 1.5 GHz as a function of parallactic angle for a single baseline. The traces show data at six frequencies in steps of 0.2 GHz from 1 GHz in blue to 2 GHz in red. The differences in amplitude with frequency are due to the changing width of the PB across the large fractional bandwidth. The source at 0.5 PB gain at 1.5 GHz is at 0.8 PB gain at 1.0 GHz and 0.2 PB gain at 2.0 GHz. This introduces a steep false spectral index that is corrected for in *WB A-Projection*. The similar but oppositely curved shapes of the RR and LL signals at each frequency arises from beam squint caused by the RR and LL beams having different pointing centers in the sky from the off-axis feed geometry of the VLA antennas. The effect of the squint is maximum at the half power of the PB and reduces in magnitude as the source moves away from the half-power point in either direction.

Fig. 3.5 shows the amplitude of the cross-hand (RL, LR) visibilities for a baseline for the unpolarized source. The polarization signal arises from the leakage of flux from the total intensity of the parallel hands RR and LL into the cross-hands, RL and LR. The leakage signal is a strong function of both parallactic angle and frequency. Fig. 3.6 shows the same quantities for the 5% polarized source placed at the same beam position. These traces show the combined effect of Stokes *I* leakage and the other Mueller matrix terms that are

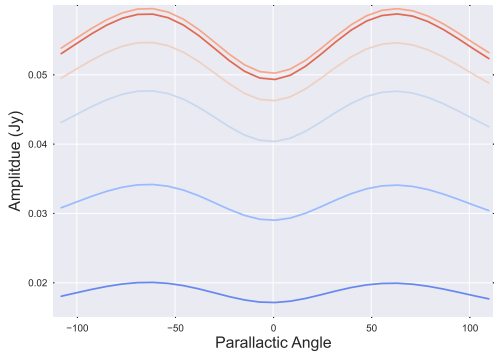


Figure 3.5: Polarized flux density in RL and LR as a function of parallactic angle for the unpolarized source at the half-power point at 1.5 GHz. Colors are as in Fig 3.4. The signal arise from angle-dependent leakage from the source total intensity signal and also has strong frequency dependence as the source moves out through the PB with increasing frequency. RL and LR are both complex-valued with the same amplitude. These curves therefore overlap exactly

.

dependent on Q and U. The combined effects create a strong modulation of the polarized signal from less than 1% to over 7%.

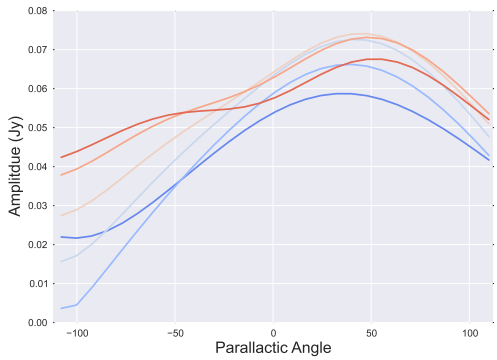


Figure 3.6: Polarized flux in RL and LR as a function of parallactic angle for the 5% polarized source at the same location of the unpolarized source in Fig.3.5. Colors are as in Fig. 3.4. RL and LR are both complex-valued with the same amplitude. These curves therefore overlap exactly

.

3.3.2 Polarization Fidelity and Observing Interval

The reconstructed flux density full-Stokes vector $\vec{I}^M$ during imaging results from a summation over time as given in Eq. 3.4. Over the course of an observation with total interval Δt, the parallactic angle will swing through a corresponding interval in $\Delta\chi$. To show the effects of the duration of an observing on the error incurred in polarimetric flux density, I extended the simulations over 24 hours and examine net polarization (Eq. 3.4) as a function of the parallactic angle range $\Delta\chi$. Figs. 3.7, 3.8, 3.9 show the observed fractional Q and

U intensity (Q^M/I^M, U^M/I^M) and the total fractional polarization Π_L^M as a function of the parallactic angle interval $\Delta\chi$, for both the unpolarized and polarized sources at the half-power point of the PB at 1.5 GHz. In the plots, the families of solid lines represent the response to the unpolarized source, while the dotted lines represent the polarized source with true $\Pi_L = 0.05$. For very small $\Delta\chi$, the error in fractional polarization can be very large. At the high-frequency end of the band, the fraction leakage is > 0.25 for $\Delta\chi < 50^\circ$, and remains high until $\Delta\chi$ approaches 180°. At $\Delta\chi = 180^\circ$ the net instrumental polarization sums to zero at all frequencies, and the true fractional polarization in all Stokes is retrieved. However, for typical observing intervals from a few minutes to a few hours, where the parallactic angle range is much less than 180°, instrumental polarization will be significantly larger than the true polarized signals over most of the PB.

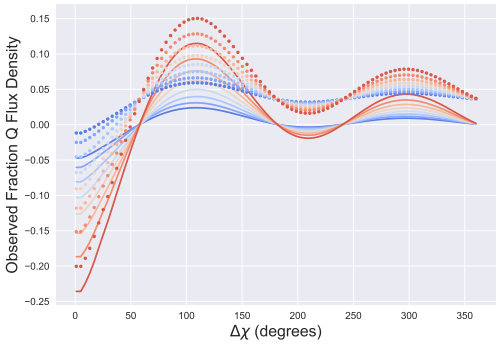


Figure 3.7: The amplitude of fractional polarization (Q^M/I^M) as a function of the interval of parallactic angle during observations. The solid lines represent the unpolarized point source at the half power point(at 1.5 GHz). The dotted lines represent the polarized point source of input Stokes Q/I of 0.0353. Colors are as in Fig. 3.4

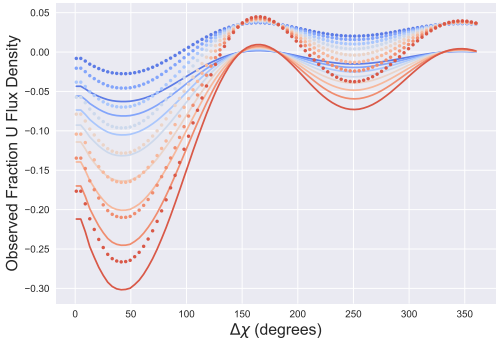


Figure 3.8: As in Fig. 3.7 but for U^M/I^M.

3.3.3 Effect on Rotation Measure Synthesis

At frequencies of a few GHz and below, the incident linear EVPA of a polarized source may be a complicated function of frequency due to Faraday effects either intrinsic to the source or due to propagation effects through the intervening Galactic or extragalactic media. In the

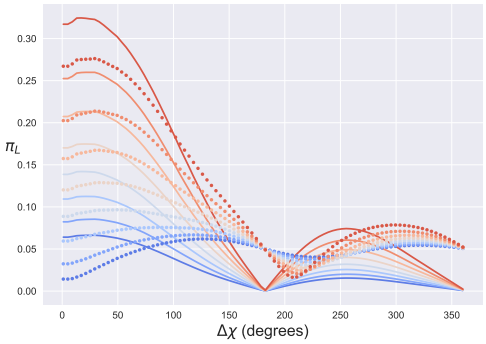


Figure 3.9: As in Fig. 3.7 but for Π_L

simplest scenario of an intervening Faraday screen the rotation of the EVPA is proportional to the square of the wavelength. Burn [3] introduced the Faraday Dispersion Function (FDF) a measure of the polarization as a function of Faraday depth ϕ to measure the distribution of Faraday components effecting the polarization vector. This technique and associated analysis is becoming widely used as a powerful means to study the astrophysics of the Faraday effects in radio sources and intervening space. Brentjens and de Bruyn [5] introduced a window function and demonstrated that for a band-limited interferometric measurement the FDF can be approximated by

$$\begin{aligned} \widetilde{F}(\phi) &\approx K \sum_{i=1}^{N} W_i P_i(\lambda^2) e^{-2i\phi(\lambda_i^2 - \lambda_0^2)} \\ &\approx F(\phi) \star R(\phi) \end{aligned} \tag{3.5}$$

where i is the channel of observation within a finite bandwidth, W_i is the weights per channel, K the inverse of the summed weights and λ_0 is the mean observing wavelength. Eq. 3.5 is a Fourier transform of the weighted complex polarized intensity. The observed FDF is a convolution of the true FDF with the Rotation Measure Spread Function (RMSF), or the Faraday depth point spread function given by

$$R(\phi) \approx K \sum_{i=1}^{N} W_i e^{-2i\phi(\lambda_i^2 - \lambda_0^2)} \tag{3.6}$$

The FWHM of the main lobe of the RMSF is a metric of the “resolution" in Faraday space and is approximately given by

$$\Delta\phi \approx \frac{2\sqrt{3}}{\Delta\lambda^2} \quad \mathrm{rad\,m^{-2}} \tag{3.7}$$

with the wavelength λ in meters. For a band spanning $1-2$ GHz, $\Delta\phi \simeq 50$ rad m^{-2}. The shape and range of the Faraday dispersion function of the instrumental leakage will be determined by $\Delta\phi$ and by the complexity of the wavelength dependence of the time-averaged M_{IQ}^k and M_{IU}^k relative to M_{II}^k. A smooth variation with frequency will produce an instrumental signal largely confined to low values of RM.

FDF of the 1-2 GHz off-axis response for the point source in the three simulations are plotted in Figs 3.10, 3.11, and 3.12. For these plots the colors now indicate increasing

observing intervals of the parallactic angle $\Delta\chi$, ranging from $\Delta\chi = 36^\circ$ (blue) to $\Delta\chi = 180^\circ$ (red).

The FDF of the net Stokes I leakage into Q and U from the unpolarized source at the half-power point of the beam is shown in Fig. 3.10. The spectrum shows a central peak around RM = 0 rad m^{-2}. The "side-lobes" are largely from the RMTF determined by the overall bandwidth and the channelization. The amplitude of the central peak decreases with increasing $\Delta\chi$ as the net instrument response reduces. The signal vanishes entirely at $\Delta\chi = 180^\circ$, as expected. Fig 3.11 shows the FDF of P^M (the measured polarized intensity) of the source with $\Pi_L = 0.05$ and RM = 0 rad m^{-2}. At small $\Delta\chi$ the true polarized signal is almost entirely removed. The trace at $\Delta\chi = 36^\circ$ shows a very weak maximum shift slightly to negative RM. As $\Delta\chi$ get larger P^M increases and approaches a better approximation of the true polarization, reaching $P^M = P$ at $\Delta\chi = 180^\circ$. This result demonstrates that for sources with low RM the instrumental polarization for observations with parallactic angle intervals less than $\sim$100° will significantly corrupt the true FDF signal.

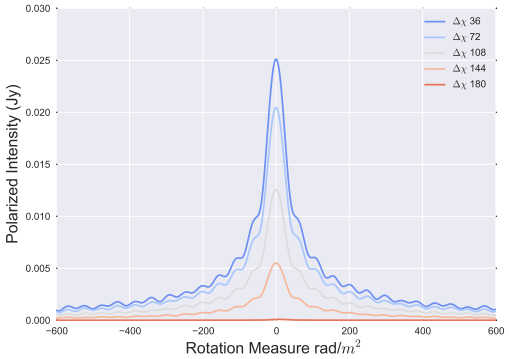


Figure 3.10: The FDF of the instrumental leakage for the unpolarized source at the PB half power point at 1.5 GHz. In this plot the color now denotes increasing parallactic angle intervals spanned from 36° (blue) to 180° (red). The instrumental signal is maximum close to a Faraday depth of zero. The signal amplitude decreases for observations with increasing $\Delta\chi$.

Fig. 3.12 shows FDF for the case of a polarized source with RM of 100 rad m^{-2}. The signal from the source is present at close to the correct amplitude and at the correct RM for all parallactic angle intervals. For shorter intervals ($\Delta\chi < 100^\circ$) there are strong spurious signals at RM=0 of similar amplitude to the source signal, and both the peak polarization and the RM of the peak of the signals from the polarized source differ from the true value by about 5%. As in the other cases, as $\Delta\chi$ approaches 180° the instrumental effect vanishes and the signal from the source approaches its true value.

Faraday rotation synthesis thus offers the possibility to distinguish instrumental effects from true polarization even in data uncorrected for off-axis effects for sources with large RM that are well separated in Faraday depth from the instrumental signal. For the VLA at L-band RM significantly larger than about 50 rad m^{-2} is required. While RM values of this magnitude are not uncommon in the plane of the Milky Way Galaxy due to the dense Galactic magneto-ionic medium (Brown et al. [48]), at higher galactic latitudes such large RM values are the exception rather than the rule (Taylor et al. [6]). Moreover, broadband polarimetry reveals complex FDF with signals over a range of RM in a significant fraction

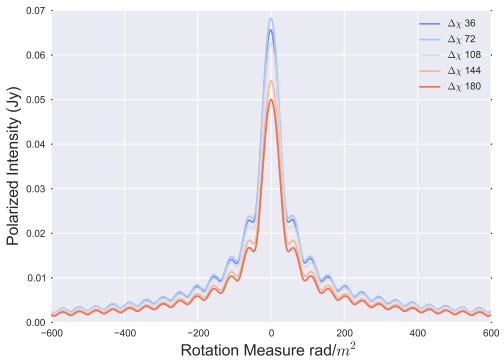


Figure 3.11: Plotted is the linear polarized intensity of the polarized source (Π_L=0.05, $\psi = 22.5$, RM=0 rad m^{-2}) at PB the power point at 1.5 GHz, as a function of rotation measure. The color axis denotes the different parallactic angle intervals spanned from 0 to 180 degrees. The spread in the RM around zero arises from the addition of the beam RM to the intrinsic source polarization.

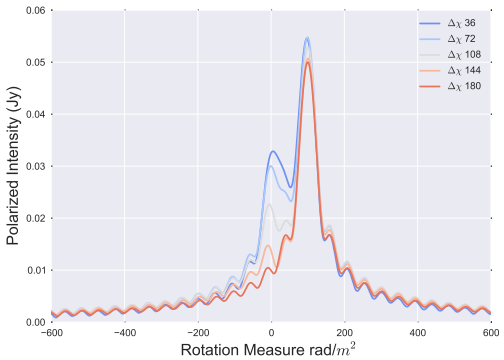


Figure 3.12: Plotted is the linear polarized intensity of a polarized source (P=0.05, $\psi = 22.5$, RM=100) at half power at 1.5 GHz, as a function of rotation measure. The color axis denotes the different parallactic angle intervals spanned from 0 to 180 degrees.

of sources, arising from internal Faraday effects (O'Sullivan et al. [26]). Upcoming wide-area surveys will require high precision polarimetry and Faraday Rotation synthesis over the full range of RM. Wide-band, off-axis polarization corrections, therefore, will be essential.

3.3.4 Effects of Primary Beam Squash

The effects of the antenna PB on polarimetric imaging described in the sections above is based on ray traced antenna models that account only for the leading order terms in phase (eg. *squint*). Measured PB from holography (Perley [49]) of the VLA antennas shows that the PB response is antenna dependent with different antennas displaying beam squash in addition to well known beam *squint* of the VLA. While the PB *squint* is a linear term in phase, which represents a displacement of the beam centers of `R` and `L` beams, PB *squash* is a quadratic term in phase (Heiles et al. [50]). *Squash* is symmetric and is the superposition of *defocus* and *coma*, along with the addition of polarized emission from the

sub-reflector feed legs. All of these effects alter the measured linear polarization response of the antenna.

The most noticeable effect is the redistribution of power from the main lobe to the side-lobes and changes to the peaks of the two opposing polarization lobes in Stokes Q and U. If the quadrapolar symmetry is altered, the flux leakage from Stokes I leaking into the linear polarizations would not reduce to zero even when $\Delta\chi = 180^\circ$. Fig. 3.13, shows the leakage as a fraction of Stokes I for different intervals of parallactic angle, for an ideal beam and for a *squashed* beam. The ideal antenna Stokes Q, U and fractional polarization are plotted in solid lines in blue, green and red, while the squashed beam Stokes Q, U are plotted in dotted blue, green and red respectively. The squashed beam shows higher levels of leakage for smaller parallactic angle intervals and does not average to zero $\Delta\chi = 180^\circ$. The residual flux at $\Delta\chi = 180^\circ$ is $\sim 0.5\%$ of Stokes I.

The *squashed* beam thus exacerbates the problem of leakage correction and removes the symmetry that averages out the leakage over long integrations. The *A-Projection* algorithm which can naturally account for the *squash* term with an appropriate aperture models offers the ideal solution for achieving noise-limited high fidelity imaging in polarization over the full PB.

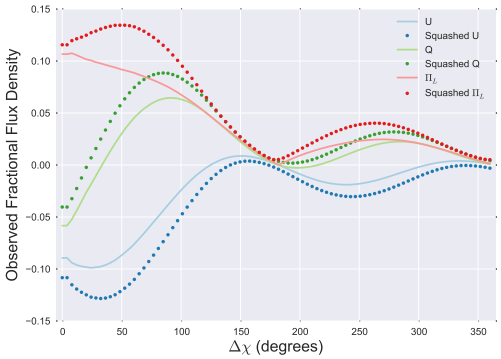


Figure 3.13: Plotted is the leakage flux at the half power of the PB at 1.5GHz, as a fraction of Stokes I for different parallactic angle intervals $\Delta\chi$. The dotted(squashed) and solid(*ideal*) in red represent leakage into to polarized intensity Π_L, the blue lines represent the Stokes `U` and the green lines the Stokes Q.

3.4 Conclusion

The next generation of radio polarimetric surveys utilizing parabolic dishes, spanning wide bandwidths and wide-area in all Stokes parameters cannot achieve high fidelity in polarimetric imaging using conventional imaging algorithms. The effects of direction-dependent gain and polarization leakage in parabolic dishes with off-axis feed location is shown to create spurious polarized signals and significantly corrupt the polarization response at levels similar to or exceeding the true sky polarization. The off-axis effects from Stokes I leakage can be ameliorated to some degree by observations over long averaging time intervals that provide a range of parallactic angle $\Delta\chi > 100^\circ$. The net effect of the off-diagonal terms of the Mueller matrix averages to zero for $\Delta\chi = 180^\circ$. However note that due to the time variability of the PSF, time averaging will lead to deconvolution

errors during image reconstruction. Observations that do not span such long time intervals or cover exactly symmetric hour angles about the meridian, the polarization leakage does not vanish. I have also shown that for realistic beam models with measurable second-order errors (beam *squash*) averaging of polarization leakage is not a viable solution.

The off-axis instrumental polarization response of the antenna in Faraday space is shown to be close to RM of 0 rad m^{-2}. For typical RM signatures of astrophysical sources at GHz frequencies the off-axis leakage response alters the Faraday dispersion spectrum, introducing spurious signals similar in scale to the sky polarization and corrupting the true sky RM signature. Our simulations show that achieving polarimetric imaging simultaneously over wide-fields and wide-bands for next-generation deep and wide surveys will not be possible through conventional imaging methods that do not include DD corrections.

4. Antenna Aperture Illumination

The aperture illumination pattern (AIP) of an antenna determines the directional gain and sensitivity of the antenna to the sky brightness distribution. For an interferometric baseline consisting of a pair of antennas, the outer convolution of the two AIPs determines the Mueller matrix. The Mueller matrix encodes the mixing of the input polarization signals, including the effects of the off-axis leakage of one polarization product into another. It also largely determines the imaging performance. Accurate knowledge of the antenna AIP is essential for high-fidelity imaging performance of a radio interferometric array.

The current and next generation of interferometric arrays are outfitted with dual-polarization, wide-bandwidth receivers having high fractional bandwidths (Total bandwidth/center frequency), e.g. in the case of the VLA, by as much as 66–75% (L, S, and C bands). The directional properties of the AIP in each polarization will change significantly across the band. In addition to smoothly varying geometric frequency scaling of the AIP, effects can arise due to, imperfect optical alignments or standing waves between optical elements of the antenna (e.g. Popping and Braun [51]). Standard calibration and imaging algorithms that do not account for the directional and frequency dependence of the antenna AIP lead to errors whose magnitude increases with distance from the antenna pointing direction and are particularly significant for polarization imaging (Jagannathan et al. [52]). With a known AIP, the direction-dependent errors can be corrected over the field-of-view using the *A-Projection* algorithm (Bhatnagar et al. [32, 33]).

The AIP can be modeled to the first order by a simple ray-tracing geometric model. *Geometric* models of aperture illumination deliver sufficient accuracy in the regime where the incident wavelength of electro-magnetic waves is much smaller than the blocking antenna structures in the optical path. At its highest operation frequencies of 10s of GHz, the VLA falls in this *geometric* regime. However at GHz frequencies and lower purely geometric approaches are insufficient, higher order effects from diffraction and, scattering significantly affect and alter the AIP. These effects introduce higher order frequency-dependent terms. Full Electromagnetic (EM) simulations of the antenna in principle

allow for the accurate modeling of the AIP including higher order effects (Brisken [53]). However such approaches are computationally expensive (even more for high spectral resolution simulation across wide bandwidths) and are limited by the accuracy of antenna models and illumination patterns which are given as an initial input. Results from such EM simulations often do not accurately reflect the real AIPs and are difficult (and expensive) to perturb to fit the measured AIPs. In this chapter, I describe a hybrid method, called the A-solver, that uses holographic measurements in combination with low-order parametric modeling based on the physical structure of the antennas to efficiently create a high spectral resolution model of the full-polarization AIP over the very wide-bandwidth of the VLA.

4.1 A-Solver

4.1.1 Physical Modeling of the AIP

Approaches to computer models of the AIP or the antenna voltage pattern (VP) can be broadly classified as Physical Modeling or Phenomenological Modeling. While a detailed discussion of these styles of modeling is beyond the scope of this discussion, I would like to mention here that Physical Modeling (e.g. simulators using Physical Optics (PO) or full-EM simulators) minimizes the required degrees-of-freedom in the model and follows the physics of the problem, both of which have significant numerical and computational advantages. Physical Modeling also leads to a fundamental understanding of the instrument. Phenomenological modeling on the other hand[1] ignores physics and models individual effects as free parameters.

A simple model of the far-field radiation pattern can be computed using geometric optics (GO). This works well for smooth surfaces away from edges and structures that are much smaller than the wavelength of radiation, where diffractive effects become important. For observations above several GHz with the VLA, diffractive effects are expected to be small (e.g. see Bhatnagar et al. [32] where instrumental Stokes-V is modeled using a GO simulator for the VP). An adaptation of a GO simulator [53] exists in CASA, and referred to as the "CASSBEAM" simulator. The code though VLA centric is general and can be used to model Cassegrain antennas in general. The simulator takes as input a parametric description of the structure of the antenna. The VLA antennas are shaped Cassegrain system, with a nearly parabolic primary and a hyperbolic secondary designed to attain a more uniform illumination of the antenna aperture. The shaped aperture alters the side-lobe levels of the antenna far-field, and the side-lobe azimuthal symmetry is altered by the presence of the quadrupod legs holding up the secondary. The general shape of the main lobe and the side lobes is also altered by the central blockage due to the sub-reflector.

The set of parameters used in our work to describe the structure and optics of a VLA antenna are shown in Fig. 4.1 and are listed in Table 4.1. The shape of the secondary reflector is not part of the model and is computed on-the-fly during ray tracing by enforcing the optical path length of the rays to be a constant, from the time of the first incidence. The algorithm computes the changes in the electric field for the different reflections, following the rays to the feed where the electric fields in a natural linear polarization basis are

[1] An extreme example of this in radio interferometric imaging is the Peeling approach where each component of the sky brightness distribution is modeled separately, parameterized to separately account for each effect, like polarization squint, antenna pointing offset, the dependence of beam shapes with time, frequency and polarization, amongst others.

transformed into circular basis having been multiplied by the feed illumination function and the feed illumination taper function. CASSBEAM computes all the elements of the direction-dependent antenna Jones matrix, the VP for the two orthogonal polarizations along the diagonal and the leakage patterns on the anti-diagonal, including the effects of the off-axis location of the feeds (see Fig. 3.1 and Eq. 3 of Bhatnagar et al. [32]).

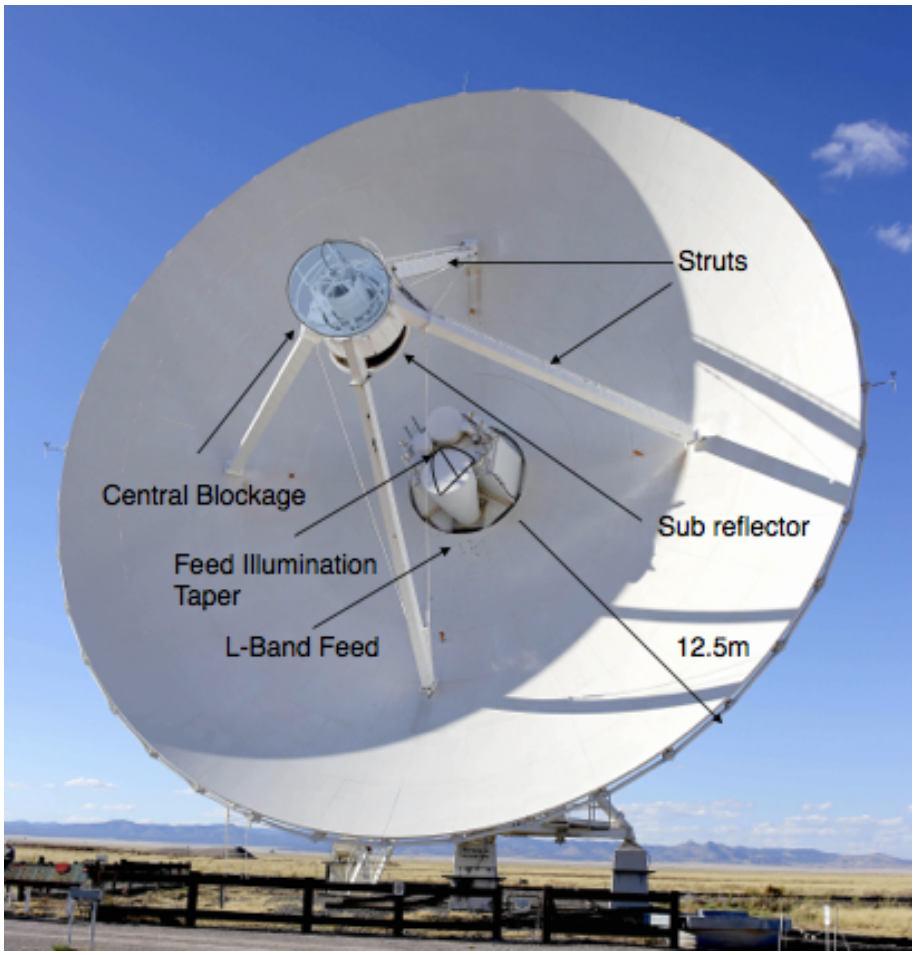


Figure 4.1: A VLA Antenna is shown with the feeds and the obstructing structs, along with the two parameters used to model the AIP accurately a) the apparent central blockage (Rhole), b) the feed illumination taper. The exact physical model parameters used to derive the dish geometry are described in Table. 4.1

This simple *geometric* model of the antenna aperture illumination is insufficient for *A-Projection*. At L, S and C Bands (1-2, 2-4, 4-8 GHz) of the VLA, secondary reflections, diffraction and, scattering play a major role while standing-waves in the optics play a significant role in all the bands. Secondary scatterings involving the feed, sub-reflector, support beams, and struts – all structures of the order of the wavelength of incident electromagnetic waves – alters the PB. The effects on the antenna PB are two-fold. The amount of flux gets redistributed from the main lobe to the side-lobes, and the introduction of higher order frequency-dependent effects in the antenna PB, alters the effective off-axis leakage and the angle of polarization squint.

To refine the AIP model I developed the A-solver approach where the model parameters are perturbed such that the predicted AIP fits the Holographic measurement of the AIP. The

Antenna Parameters in Ray Tracing	L Band Values
Antenna Name	VLA
Sub-reflector height	8.47852
Position of feed in x	-0.10026
Position of feed in y	0.97019
Position of feed in z	1.67640
Sub-reflector Angle	9.26
Width of strut legs	0.27
Strut legs distance from vertex	7.55
Height of strut legs above vertex	10.93876
Radius of central hole	2.0
Radius of the Antenna	12.5
Band reference frequency	1.5
Feed taper polynomial	10.0, 2.0
Order of feed taper polynomial	2

Table 4.1: All measurements of length are in meters. All angle measures have units of degrees. Polynomial coefficients are unit less quantities. All dimensions provided here are from (Brisken [53]).

latter is a measure of the real AIP, which usually is significantly different from idealized AIP.

4.2 Measuring the PB through holography

Holography directly measures the antenna VP, E_i, using the signals from strong unpolarized, compact calibrator radio source at a grid of positions (l, m) over the AIP. This replaces $I(\vec{s}, \nu)$ in Eq. 2.7 with an approximation of a Kronecker delta function in $\vec{s}$ at each (l, m). Typically a subset of the array antennas, whose VP is measured, scan the source, while the rest of the antennas are used as reference antennas and are pointed towards the source (i.e. the sources is placed at $l = m = 0$). The reference antennas provide the reference signal, with respect to which the signal from the scanning antennas is measured, and when projected in the antenna Az-El plane, gives a sampled map of the complex VP, $E_i(l,m)$, for all of the scanning antennas.

An important limitation of this method is the low signal in areas of the AIP with low directional gain. An accurate aperture model would require the holography measurement to sample beyond the first side-lobes in a dense grid with high signal-to-noise ratio. Since our interests lie in the Fourier transform of the antenna VP and not the VP itself, truncation of the measured VP after the first side lobe gives rise to errors (due to aliasing) in $\mathbf{A}_i = \mathscr{F}[E_i]$.

$\mathbf{A}_i^{M^\dagger}$ is applied as a convolutional correction while gridding the observed visibility data onto a regular grid as described in Eq. 2.29. Here are two kinds of oversampling required to represent the convolution function (CF) in its appropriate digital form for gridding. For computational efficiency reasons, the CF used for gridding in general (not just for projection algorithms) is a look-up table. To minimize quantization errors, the

CF look-up table needs to be oversampled by a factor represented by the symbol O_{ap} (typically $O_{ap} \geq 20$). Holographic measurements are the measurements of the antenna VP (E_i) itself. To minimize aliasing as well as to measure the various features of the antenna VP accurately, the holographic measurements are also oversampled. However, due to practical limitations, a much smaller oversampling factor ($1.5\times$) was used and found to be sufficient for these antenna VP measurements. Since the holographic oversampling factor in the antenna VP measurement is much less than the oversampling factor needed for gridding (O_{ap}), a parameterized model of the antenna AIP is required.

4.2.1 Holography

For the holography data used here, an unpolarized source, 3C147 ($< 0.04\%$ polarized), was scanned in a 35×35 grid in the antenna reference frame with a step size of $\Delta l, \Delta m = 2.5057'$, out to the second null (at 1GHz). In addition, a polarized calibrator 3C286 was observed to provide polarization angle calibration. Half the array was utilized as reference antennas while the other target antennas scanned the array. For more details on the holographic measurement see Perley [49].

The visibility data were imported into AIPS to obtain the antenna grid coordinates (l,m) for each holography scan. The uncalibrated data were exported as UVFITS file and imported into CASA as a Measurement Set. The calibrator fluxes were set using [54] and [55] for both 3C147 and 3C286 in all polarizations across the full bandwidth. On-axis gain, bandpass, frequency-dependent polarization leakage and polarization position angle calibration were carried out and the calibration solutions were applied to the data using the `APPLYCAL` task in CASA. Subsequently, utilizing the CASA toolkit, data from baselines between the target antennas and each of the reference antennas were averaged to improve the signal-to-noise ratio of the measured VPs. Further, the data recorded on a grid point for 10 seconds was averaged. This gave the final set of antenna VP data per channel per holography grid. The VP data were then interpolated onto a 128×128 grid on a per channel basis to create a 1024-channel image cube for each of the target antennas, in polarizations `R` and `L` (the diagonal elements of the DD antenna Jones matrix) and leakage patterns `R`←`L`, `L`←`R` (the anti-diagonal elements of DD antenna Jones matrix).

4.3 Optimizing the model through holography

The ray-tracing AIP simulator code within the CASA imaging R&D code base was modified to accept input parameters from a Python wrapper code. This was wrapped as a parameterized function in Python and utilized as the unknown function to be determined by the Nelder-Mead simplex algorithm (Nelder and Mead [56]), minimizing for the function parameters against each of the individual channel images of the target antenna VP cube produced from the holography data (see Section 4.2.1). The optimization parameters were the apparent blockage (Rhole), the feed illumination taper function (ftaper as a 4^{th} order polynomial), and the antenna pointing offset in R.A. and Dec (xoffset, yoffset)[2]. The simplex algorithm traversed each parameter space independent of the others varying them until a joint minima is found. The residuals before and after the joint minimization are

[2]The antenna pointing offset solved for is the mechanical antenna pointing, different from pointing offsets due to polarization squint. The latter is an optical phenomenon and is naturally included in the computation of the VP due to the physics of off-axis optics and does not need to be solved-for as an independent parameter

shown in the upper and lower panels of Fig. 4.2. The choice of the simplex algorithm over more computationally optimal algorithms arises from the lack of *apriori* knowledge of the gradients of the various minimization parameters, in the seven-dimensional parameter space.

The CASSBEAM code uses OpenMP thread parallelization and was set to launch four threads per process call. This parallelization allowed for the fast production of a new beam model for every convergence iteration. Despite this parallelization, the minimization takes 4 hours per channel per polarization to converge to a solution. So a serial minimization would take $2 \times 1024 \times 4$ hours to derive a channelized solution for an antenna. Since each channel minimization based on the parameterization is independent of the next channel a simple frequency-based parallelization was used to trigger a parallel minimization run of 1024 channels and two polarizations on the Amazon Web Services (AWS) compute cluster. This reduced the compute time down to 6 hours per antenna. The results for three antennas are discussed below. One should note here that the run-time would be unreasonably long if a full Physical Optics simulator like GRASP [3] was used instead.

In order to highlight the efficacy of the parameterized model of $A^M_{\vec{I}}$ as against the *ideal* model of the AIP's Fig. 4.2 shows a comparison between $|E^{Holo}_{\vec{I}} - \mathscr{F}^{\dagger} A^{ideal}_{\vec{I}}|$ (top row) and the $|E^{Holo}_{\vec{I}} - \mathscr{F}^{\dagger} A^{M}_{\vec{I}}|$ (bottom row). $A^{ideal}_{\vec{I}}$ refers to the default aperture illumination produced by the CASSBEAM code, and $A^M_{\vec{I}}$ is derived from the parameter values obtained by the optimization procedure discussed in Sec. 4.3. The first side-lobe is underestimated in the upper panels of Fig. 4.2 by 50% in both polarizations (`L` upper left and `R` upper right). Within the main-lobe of the VP, the residuals in the upper panel show a systematic offset in power within the main-lobe in both polarizations. The offset within the main-lobes that affects both polarizations equally is a sign of mechanical antenna pointing error. In contrast, the lower panel images (lower-left for `L` lower left and lower right for `R` lower right) of the parameterized model residuals shows no sign of side-lobe power discrepancy or residual pointing error. These are residuals for one frequency channel. The optimized residuals show similar improvement across the entire bandwidth at the VLA L-Band for all the optimized antennas.

The optimization procedure solved for the pointing offset of the antenna and then fitted the data for *Rhole* – the *apparent* blockage. Heiles et al. [50] demonstrated that a blocked aperture leads to increased power in the side-lobes. They also show that the size and extent of the VP side-lobes can be effectively shaped by tapering the illumination of the feeds. In line with their finding, I was able to effectively model the first side-lobe power altering the *apparent* blockage in ray-tracing in conjunction with the *ftaper*, feed illumination taper polynomial function utilized in the code. I find that an increased apparent blockage and a sharper tapering function for feed illumination, determined per channel across the entire band allows for the capture of all significant changes in the antenna VP out the first side-lobe. I also note that the trend captured in the optimized parameters correlates with the measured wide-band sensitivity of the VLA L-Band Momjian et al. [57], which suggests that the optimized models correctly estimate the departures from an idealized antenna.

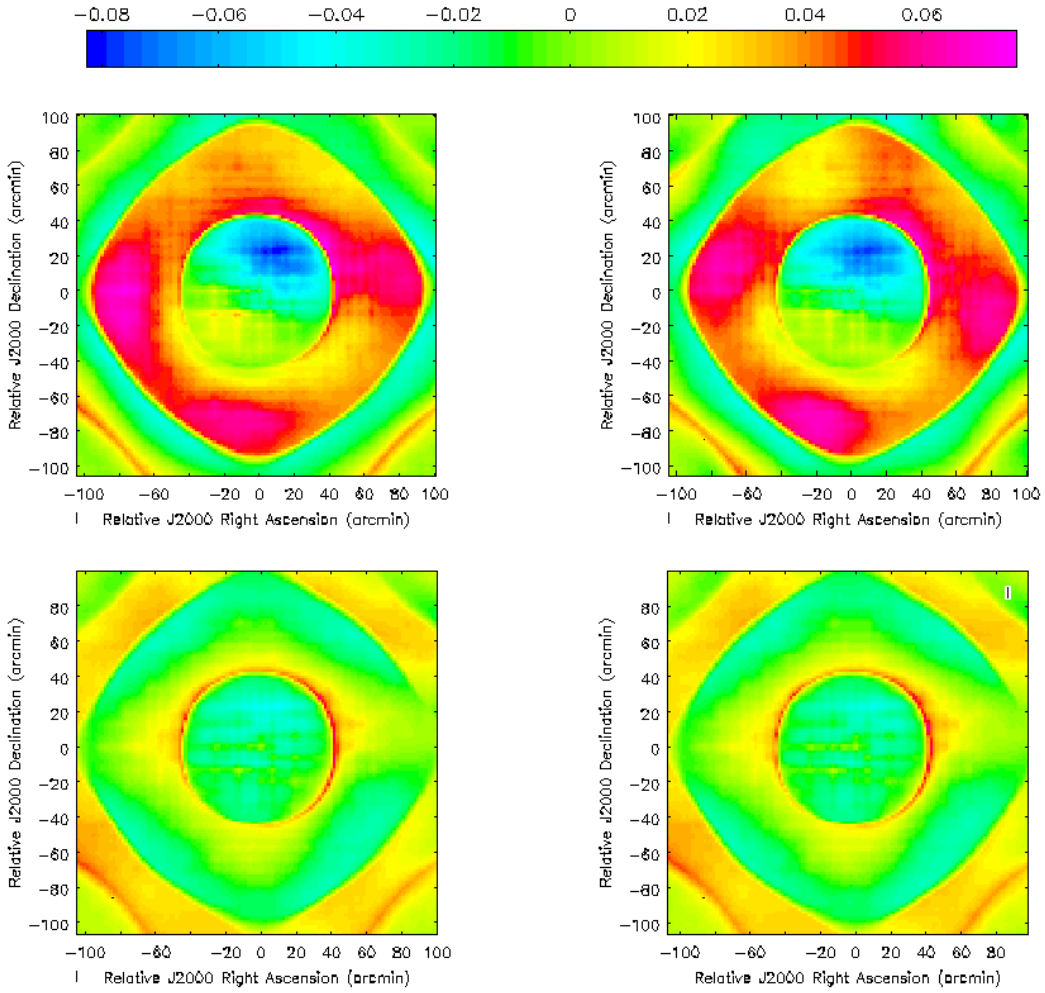


Figure 4.2: The upper panel images shows the residuals (normalized with respect to peak intensity at the beam center) of $|E^{Holo}_{\vec{I}} - \mathscr{F}^{\dagger} A^{ideal}_{\vec{I}}|$ measured at 1.353 GHz. The upper left panel show the residual for the left-circular (L) polarization and the upper right panel for the right-circular (R). Similarly, the lower panels shows $|E^{Holo}_{\vec{I}} - \mathscr{F}^{\dagger} A^{M}_{\vec{I}}|$.

4.3.1 Apparent Central Blockage

The central blockage in an antenna reduces the aperture efficiency and increases the side-lobe levels – an aspect that is alleviated by shaped surface design and off-axis feed geometry of the VLA to improve uniform aperture illumination and increased aperture efficiency (chapter 3, Taylor et al. [30]). The frequency dependence of the VP across the bandwidth, in particular, the presence of a standing wave, altered the first side-lobe and the shape of the polarization properties of the VP for the JVLA antenna across L-Band. Solving for a frequency-dependent apparent blockage allowed us to capture the frequency-dependent variation in the per-channel solutions of the *Rhole* parameter. Plotted in Fig. 4.3 in red, green and blue is the *apparent* blockage parameter for three different antennas derived from the optimization spanning seven spectral windows. The effect of the standing wave is captured in the variation of the *Rhole* parameter with frequency. This frequency-dependent variation of the *Rhole* parameter, in turn, can be fit using a combination of a straight line

[3] http://www.ticra.com/products/software/grasp

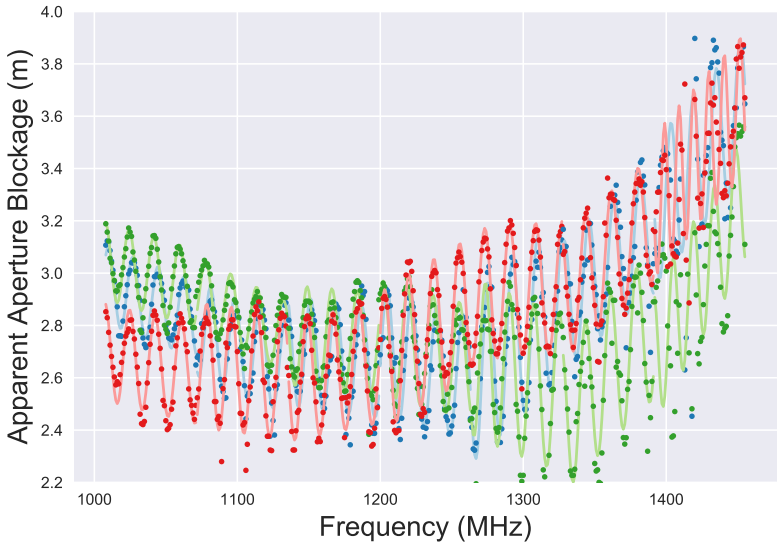


Figure 4.3: Fit to the recovered *apparent* blockage parameter for antennas 6, 10, 12, in red, green and blue respectively, with the lines representing the fit and the points represent the derived apparent blockage data across 448 MHz, of data.

in frequency per spectral window, and a sinusoidal function. The data are fit per spectral window each containing 64 MHz of data utilizing the astropy models package (Astropy Collaboration et al. [58]). The fit reveals that the oscillations in frequency have a period of $\sim$ 17 MHz. The period of this oscillation corresponds to twice the light travel time from the feed to the secondary, consistent with the presence of a standing wave between the antenna secondary and the feed. With this fit – of a line and a sine function – the number of parameters that determine the frequency-dependent behavior of the antenna AIP reduces to five numbers per spectral window. (The data and the fits to the data are available upon request)

4.3.2 Feed Illumination Taper

The VLA receiver feeds lie on a circle around the optical axis. The feeds are illuminated by the sub-reflector and the angular span of the illumination can be altered by tapering the feed illumination pattern. The tapered illumination pattern reduces the amount of radiation received from the edges of the dish, which while marginally reduces aperture efficiency, effectively stops the feed receiving spillover radiation. In addition to reducing the spillover, it also alters the shape and the gain of the PB side-lobe. The parameterized AIP model allowed for the taper function (a 4^{th} order polynomial) to vary along with the central blockage to optimally match the shape and structure of the antenna VP out to the first-lobe. In Fig. 4.4 the normalized amplitude of the feed taper function is plotted against the angular distance from the feed axis for antenna 12, at 1.0, 1.5 and 2.0 GHz in red, green and blue dashed lines respectively. The feed taper function obtained from the optimization is plotted in light red, green and blue solid lines respectively. The taper functions determined from parameter optimization have stronger tapering and a sharper fall off resulting in lesser feed illumination overall to match the VP's determined through

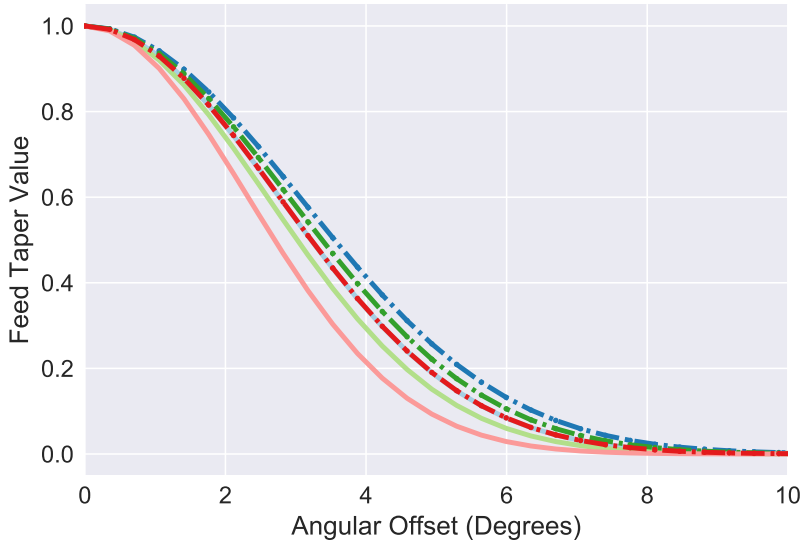


Figure 4.4: The dashed red, green and blue lines show the feed taper function at 1, 1.5 and 2.0 GHz respectively, used to derive $A_{\vec{I}}^{ideal}$. The solid, red, green and blue (overwritten by the dashed red line) lines show the feed taper function at 1, 1.5 and 2.0 GHz respectively, used to derive $A_{\vec{I}}^{M}$ for antenna 12.

holography.

4.3.3 Pointing Offset

Fig. 4.5 plots the per-channel solutions for the pointing offsets for antennas 6, 10 and, 12 in blue green and red respectively in units of the half-power-beam-width (HPBW). Any linear scaling with frequency is therefore removed and all optical effects that scale linearly with frequency should appear as flat curves in this plot. On the other hand, effects like the mechanical pointing offsets, which are not optical effects, should appear in this plot with linear slope as a function of frequency. The mean separation between the R- and L-beams is $\sim 5.7\%$ corresponding to the known polarization squint due to the off-axis optics of the JVLA antenna. The solid lines and the fainter points plotted above the curves showing the R- and L-beam offsets are the mechanical pointing offsets for the three antennas. Antenna 6 shows the largest pointing offset indicated by the line-fit with a slope of $\sim 2.4'$. All antennas show mild variations in the pointing offsets with frequency. Frequency-dependent pointing error over and above the squint can be caused by an uncorrected second-order term in phase across the antenna. The higher order phase terms also affect the $E^{R\leftarrow L}$ and $E^{L\leftarrow R}$ adversely, introducing squash and other higher order distortions. Modeling these higher order phase errors in the off-diagonal Jones matrix is covered in Sec. 4.3.4. Once the pointing offset has been solved for per channel, solutions for the *apparent* blockage (Rhole) and the feed illumination taper polynomial (ftaper) were derived.

4.3.4 Antenna AIP and Imaging

A sub-optimal AIP model, $\mathbf{A}_{ab}^{M}$, will create errors in the image that can be characterized in the residual image. The residual error contribution in a snapshot for a single baseline $i-j$

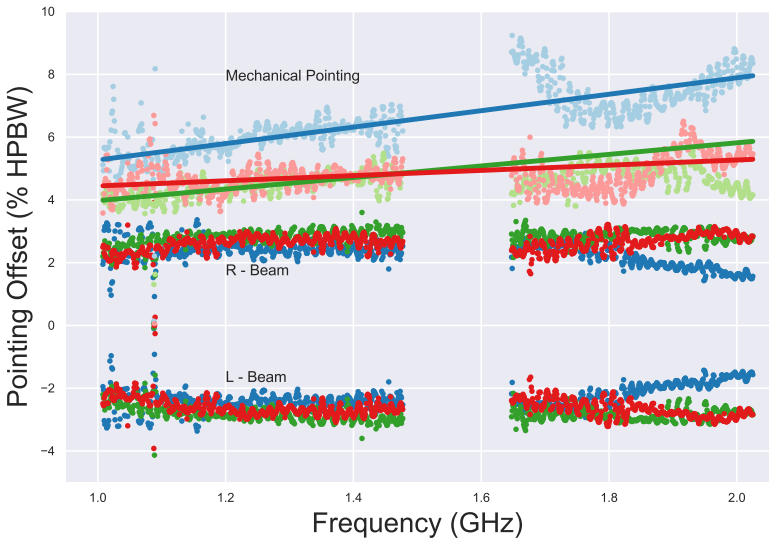


Figure 4.5: Plotted are the pointing offsets for antennas 6, 10, 12, in blue, green and red respectively, for `R` and `L` polarizations of the antennas around zero. The offsets are in percentage of the HPBW of the antenna. The lines represent the actual pointing vectors of the antennas and is obtained by fitting the per channel pointing solutions with a best fit line. The regions with no solutions between 1.4 and 1.65 GHz corresponds to two spectral windows with noisy solutions due to the presence of strong radio frequency interference.

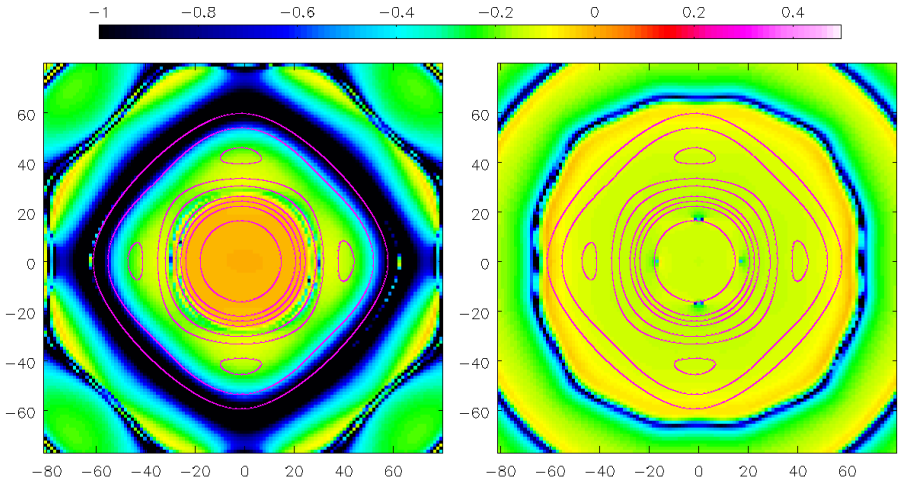


Figure 4.6: Plotted is the fractional change in the antenna PB, $\left(PB^{def} - PB^{True}\right)/PB^{True}$, with magenta contours overlaid of PB^{True} at 80, 50, 10, 5 and 1 % power at 1.448 GHz of antenna 12. The left panel is the fractional change in total intensity, while the panel on the right is the fractional change in linear polarized intensity

can be written as,

$$I^{res} = I^{psf} \star [\Delta M_{ab} \cdot I^{\circ}] \tag{4.1}$$

where I^{res} is the residual image, I^{psf} is the telescope point spread function to be deconvolved, I° is the true sky distribution, $\Delta M_{ab} = \mathscr{F}[A_{ab}^{True}] - \mathscr{F}[A_M]$ is the difference between the true antenna AIP and the model AIP.

Let us consider PB^{True}(or equivalently $\mathscr{F}[A_{ab}^{True}]$) to denote the PB of the antenna AIP with optimized *Rhole* and *ftaper* parameters, and PB^{def} to denote the PB of the antenna AIP with frequency independent *Rhole* and *ftaper* parameters. The left panel of Fig 4.6 then shows the fractional error $(PB^{True} - PB^{def})/PB^{True}$ when using the standard sub-optimal AIP, as against the optimized AIP, for Stokes I at 1.448 GHz of antenna 12. The optimized beam is overlaid as contours in pink. The error within the main lobe of the PB is at the level of several percents, a significant change for high fidelity imaging noise limited wide-field imaging that typically requiring dynamic ranges in excess of 10,000:1. The left panel of Fig. 4.6 also demonstrates that error in flux reconstruction $> 5\%$ start beyond the 0.05 gain position of the PB main lobe and continues to increase to nearly 40 – 60% change across the first side-lobe. At right in Fig 4.6 is the fractional error in polarized intensity $(PB^{True} - PB^{def})/PB^{True}$. The error in the polarized intensity varies between 10 and 20% across the PB out to the first side-lobe.

While the fractional error in the PB gives us the instantaneous error in the residual image for a particular frequency the effect on the total continuum sensitivity offered by the wide bandwidths is gotten by examining the fractional error in the wide-band PB. The instantaneous wide-band PB is defined as $\sum_{\nu_0}^{\nu 1} PB(\nu)$ spanning the range of frequencies from ν_0 to ν_1. The wide-band PB represents the effective forward gain of broadband continuum imaging. The effective wide-band sensitivity extends far beyond the null of the narrowband PB (Bhatnagar et al. [59]). *WB A-Projection* uses the wide-band PB to normalize the image in the final imaging step of the *flat-noise* implementation of the algorithm (see Bhatnagar et al. [33] for more details). Shown in Fig. 4.7 is the fractional error in the wide-band PB at the reference frequency of 1.5 GHz. Overlaid in pink are the contours of instantaneous wide-band $\sum PB^{True}$. The fractional error in the PB means the error in the gain of the PB-corrected image is $\sim 5\%$ at the 0.1 gain of the PB and increases to $\sim 20\%$ at 0.01 PB gain (this includes the first side-lobe). Since every pixel in the wide-band PB image is the sum of the pixel values at all the frequencies, the fractional error beyond 0.1 PB gain is dominated by the lower frequencies (larger beam size) while the error within the 0.1 PB gain being dominated by the higher frequencies.

Off-Diagonal Antenna Jones

Jagannathan et al. [52] demonstrated the effects of beam *squash* on polarimetric imaging. *Squash* is caused by a second-order phase term (Heiles et al. [50]) and in conjunction with other second-order phase terms like defocus and coma, affect polarimetric imaging adversely. Reconstructing the polarized emission from the sky requires the use of all the terms of the antenna Jones matrix. In the prior sections, I have dealt with the frequency dependence of the antenna AIP primarily in the context of the diagonal elements of the antenna Jones matrix. To model the off-diagonal jones elements requires the inclusion of higher order distortions which was done by including a general second-order polynomial in phase in addition to the *Rhole* and *ftaper* parameters.

In Fig. 4.8 the panels on the left represent the real (upper left) and imaginary (lower left) parts of the off-diagonal antenna Jones matrix element $E^{R \leftarrow L}$ of the model. The inclusion of a second-order phase term in the antenna alters the side-lobe flux but does not alter the

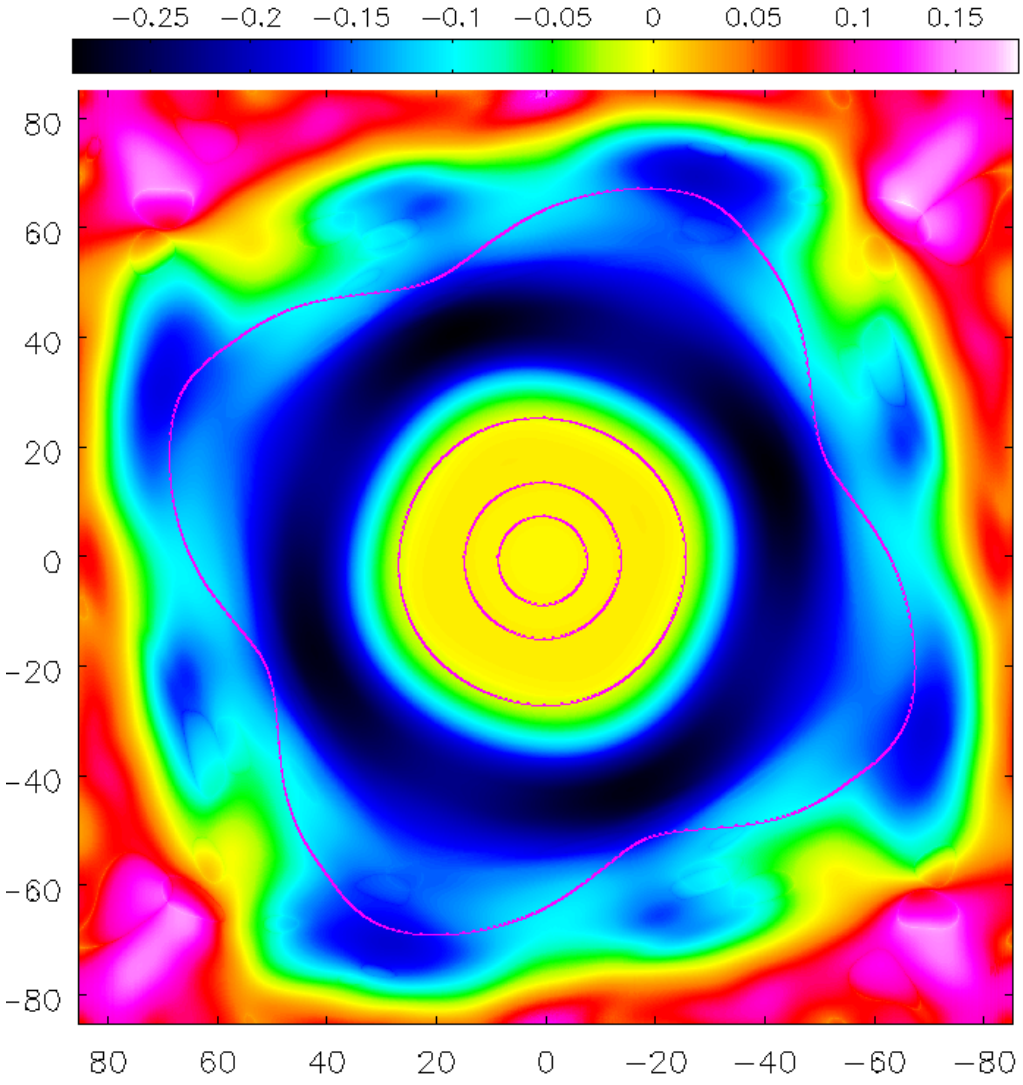


Figure 4.7: Plotted is the fractional change in the antenna wideband PB, $\left(PB^{True} - PB^{def}\right)/PB^{True}$ across 1 GHz of bandwidth, with magenta contours overlaid of PB^{True} at 80, 50, 10, and 1 % power at the reference frequency,1.5 GHz of antenna 12.

general morphology of the clover-leaf pattern as is the case in panels on the right in Fig. 4.8. The real (upper right) and imaginary (lower right) parts of the off-diagonal antenna Jones matrix element $E^{R\leftarrow L}$ of the measured holographic map. The altered morphology of the lobes mimics a rotation of the VP. I, therefore, introduced rotation of the antenna VP as an additional free parameter in the minimization which led to a more realistic model VP shown in the center panels. A rotation of $\approx 18^{\circ}$ gave the least residuals with respect to the holographic data. A similar rotation is quite clearly seen in the polarization squint vector as well for all antennas and at bands in the holographic measurements. The physical origin of this rotation is not yet understood.

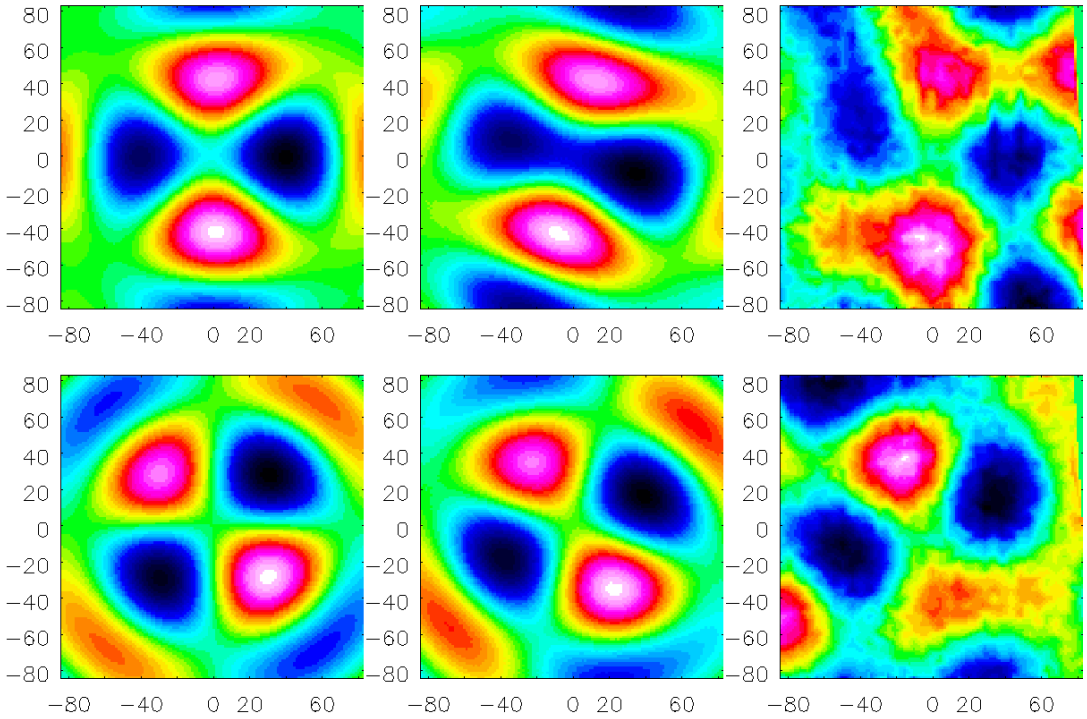


Figure 4.8: The panels of the figure shows the off diagonal antenna Jones matrix element $R \leftarrow L$ of $E_i^M = \mathscr{F}^{\dagger}[\mathbf{A}_i^M]$, with the upper panels show the real portion of E_i^M and the lower panels show the imaginary part of E_i^M at 1.013 GHz of antenna 27. The left most panels (upper and lower) E_i^M that includes an optimized second-order polynomial in phase. The figures in the center of the complex E_i^M, includes an $\approx 18^{\circ}$ rotation in addition to the second-order polynomial in phase. In the two panels on the right the the real and imaginary parts of the measured holographic E_i^M is shown.

4.3.5 Imaging Simulations

To contrast the difference in full Mueller A-Projection with parameters derived from the A-Solver methodology over the frequency independent model, I performed point source imaging simulations. Eight unpolarized point sources (I=1Jy,Q,U,V=0), were placed at different PB gains across the main-lobe and first side-lobe of the antenna PB. The data were simulated for a total integration time of fifteen minutes, with a bandwidth of 64MHz centered at 1.4GHz to produce a full Mueller predicted measurement set (refer fig . 3.3 for schematic) for the VLA in C-Configuration. The Jones matrix used in the input simulation was produced by taking the median apparent central blockage and feed illumination taper parameters of the three antennas from the A-Solver optimization and provided as input for CASSBEAM.

The measurement set was then imaged with full Mueller A-Projection with the convolution functions produced with a) standard (frequency independent) parameters for the feed illumination taper and central blockage, and b)with the updated (frequency-dependent) parameters. The PB derived from the standard parameters (PB^{def}) and the PB from optimized parameters as PB^{True}. The flux reconstructed for the point sources along with their respective gain locations are plotted in Fig. 4.9. The blue curve (optimized parameters, PB^{True}) shows that I was able to reconstruct the flux in total intensity accurately when utilizing an

accurate frequency-dependent AIP in the full Mueller A-Projection algorithm. The green curve (standard parameters, PB^{def}) shows that when using a frequency independent AIP pattern results in errors that increase from $\sim 2\%$ at the 0.3 gain in the PB to $\sim 9\%$ at the 0.01 gain within the main-lobe. In addition to the six sources in the main-lobe, two more sources were placed in the side-lobes, where the standard parameters over-estimates flux by $\sim 25\%$, as one divides by PB^{True}, which underestimates the power in the side-lobes.

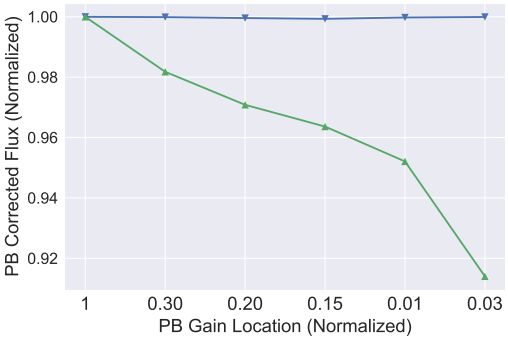


Figure 4.9: Plotted in the figure is the PB corrected point source flux. Plotted in blue are the full Mueller imaged and PB^{True} corrected point source fluxes for the parameterized frequency-dependent model. Plotted in green is the full Mueller imaged and PB^{def} corrected point source fluxes for the frequency independent model.

In Fig. 4.10 is plotted the difference image, $I^{True} - I^{def}$. The point source locations are circled in white. The difference in the imaging is arising entirely from the use of frequency independent convolution functions $\mathbf{A}_i^{M\dagger} = \mathscr{F} PB^{def}$ during gridding. The color scale is chosen to highlight the deconvolution errors introduced when the image dynamic range in consideration is in excess of 10000:1 as discussed in Sec. 4.3.4. The deconvolution errors are more prominent beyond the 5% PB gain mark within the main-lobe. The deconvolution errors denote the loss in fidelity of imaging even though the effects are markedly visible only when the imaging dynamic range is in excess of 10000:1.

The A-Projection framework used for imaging is general enough to include the antenna to antenna variations, in particular, to account for the AIP of heterogeneous arrays such as ALMA. In this chapter, I have modeled the dominant static term of the antenna AIP in terms of the feed illumination taper and the apparent central blockage parameters. While the pointing offset I solve for is used to derive a better fit to the antenna AIP, it is a time varying quantity that cannot be addressed by the A-Solver approach described in this paper. Time-dependent pointing effects can be solved for by means of the pointing self-calibration approach(Bhatnagar and Cornwell [60], Bhatnagar and Cornwell [61]). Time-dependent shape changes that affect the antenna AIP derived from the A-Solver methodology would only affect the highest dynamic range imaging studies ($\geq 10^6$) for homogenous arrays. In which case a coupled shape and pointing self-calibration approach would be required.

4.4 Conclusions

The imaging performance of the *A-Projection* is determined by the knowledge of the AIP. High dynamic range and high fidelity polarimetric imaging across wide-fields requires

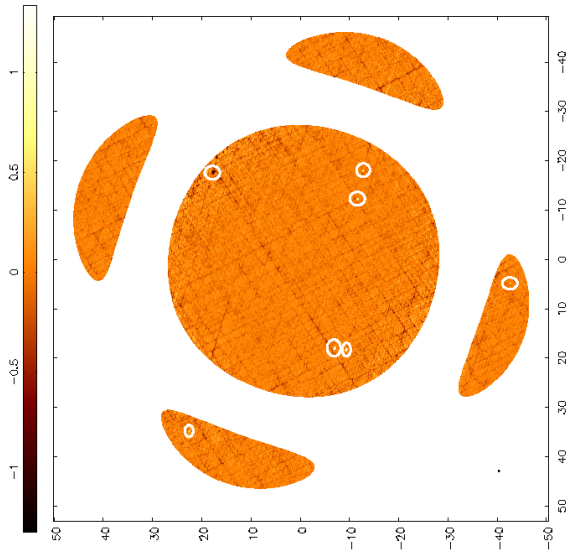


Figure 4.10: The difference image $I^{True} - I^{def}$. The white circles mark the locations of the points sources in the image. The color scales from $-1 \times 10^{-4} Jy/beam$ to $1 \times 10^{-4} Jy/beam$.

an extremely accurate understanding of the antenna AIP across the full bandwidth. The A-Solver approach of solving for the frequency-dependent AIP of antennas based on a parametrized model whose values are determined by comparison to holographic data is a viable approach to obtaining an accurate VP as demonstrated in this thesis. The parameterized model captures the rapid frequency dependence of the AIP including the effects of standing waves. Modeling the central blockage as an apparent blockage in the model allowed for the accurate reconstruction of the amplitude of the VP side-lobe as a function of frequency. The parameterized model of the AIP is a naturally compact representation requiring fewer parameters to capture higher order frequency-dependent effects, than frequency-dependent modeling of antenna VP.

An important point to note is that the product of the two AIPs making the PB for each baseline is, in general, a complex-valued function, and not a purely real function as is assumed when imaging without using the *A-Projection* algorithm (the effective PB with *A-Projection* is $\sqrt{PB^M \cdot PB^{\circ\dagger}}$, which is real-valued at the level the model PB^M accurately models the real PB°) . This could be due to differences between the two AIPs involved and/or non-Hermitian structure of the AIPs due to various EM or antenna structural effects. The PB pattern is already quite complex, and as discussed in Sec. 4.1.1, directly modeling even the real-valued PB is difficult, approximate and needs many more free parameters. In addition to this, modeling of the complex-valued PB also has all the additional numerical complications involved in directly fitting to any complex-valued data. In contrast, the physical modeling approach described in this paper models the PB in the aperture plane. This not only requires a significantly smaller number of parameters, the

parameters themselves are real-valued describing the physics of the optics (here, via the antenna structural parameters). The fitting procedure, therefore, deals with real-valued parameters. This has significant numerical advantages, and computational advantages in the production of parameterized convolution functions for a given frequency.

5. Full Mueller A-Projection

In chapter 2, I have established that there exist a class of algorithms that can account for DD effects during convolutional gridding. In this chapter, I will also outline the algorithmic details of the Full Mueller A-Projection (FM-AWP) algorithm. Further, I will apply the algorithm to simulated data and demonstrate imaging results in the case of representative point sources spread across the antenna PB and a radio field. In both of these cases, the focus is on the comparison of imaging fidelity with FM-AWP as against standard imaging. Finally, I will apply the algorithm to VLA test data at L-Band and discuss the results.

5.1 A-Projection Algorithm

In chapter 2 I briefly introduced the A-Projection algorithm which corrects for the antenna AIP during the imaging major cycle as a convolutional gridding correction. The correction is carried out in two stages during the imaging major cycle. While the algorithm allows for different kinds of image normalization, I discuss the "flat noise" case of normalization which allows for the minor cycle algorithms to carry out deconvolution on an image with uniform noise across the entirety of the image. The algorithm presented in Alg. 1 is the general wide-band FM-AWP form of the algorithm. In the forward step described below, for the narrowband case, the conjugate frequency reduces to the true frequency of the data being gridded, i.e. $\nu^* = \nu$.

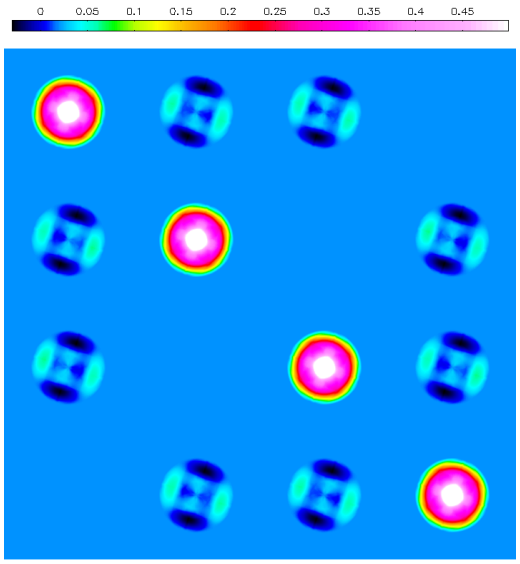
Algorithm 1 : Full Mueller A-Projection

Initialize the full Stokes model and residual images,$\vec{I}^{mod}, \vec{I}^{res} = 0$.
repeat
  Major Cycle
    if convolution function cache exists = False **then**
      Generate convolution function cache via ray-tracing per spectral window
    else
      Load convolution function and proceed with major cycle
    end if
    Predict $\vec{V}^{mod}|_k = \mathbf{A}^M(\nu) \star \mathscr{F}\vec{I}^{mod}|_k$ ▷ De-gridding
    Accumulate $Tr(\overline{\mathbf{A}^M})$ ▷ Normalization
    Compute residual visbilities $\vec{V}^{res}|_k = \vec{V}^{obs} - \vec{V}^{mod}|_k$
    Compute residual image, $\vec{I}^{res} = \mathscr{F}^{\dagger}\mathbf{A}^{M\dagger}(\nu^*) \star \vec{V}^{res}/Tr(\mathscr{F}^{\dagger}\overline{\mathbf{A}^M})$ ▷ Gridding
  Minor Cycle Invoke a deconvolution algorithm of choice
    Updated full Stokes $\vec{I}^{mod}$ derived from $\vec{I}^{res}$
until Convergence ▷ User-defined criterion

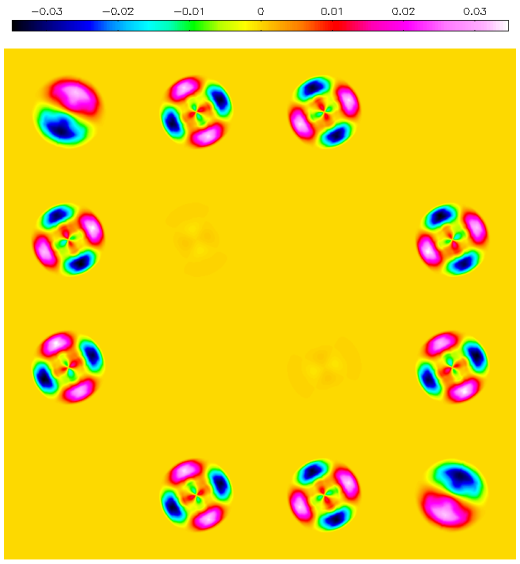
5.1.1 The Full Polarization Antenna AIP

In chapter 3 the DD effects in the image plane introduced by the full polarization Mueller matrix on an interferometer was presented. The aperture plane Mueller matrix i.e. the full polarization aperture illumination pattern (AIP) is the Fourier transform of the image plane Mueller matrix. The individual elements of the aperture plane Mueller matrix are the convolution functions that are used to resample the data onto a regular grid. The matrix shown in Fig. 5.1 is the full polarization convolution functions generated by the FM-AWP algorithm utilizing the ray-tracing methodology described in the previous chapter. The matrix is normalized such that the peak of the Stokes I PB in the image plane is 1. The normalization factor in the aperture plane is the total area of the convolution function, so the normalization of each of the convolution functions is by $2/area(CF_{RR} + CF_{LL}) = 2/area(A_{00} + A_{33})$. If one labels the matrix rows and column indices with i and j then, all matrix terms of the type A_{ii} are the leading order diagonal terms and are centrally illuminated. Elements A_{00} and A_{33} are of the type $J_pJ_p^*$ and $J_qJ_q^*$ and have opposite phases a result of the *squint*. The anti-diagonal terms ($A_{i,3-i}$) are 10^{-6} of the diagonal elements (refer, Fig. 3.1) and are not used in the current implementation. The terms are only of relevance to the highest dynamic range imaging cases while full-polarization corrections are required even for modest dynamic range imaging in polarization. I have, therefore, chosen to implement the FM-AWP imaging without the anti-diagonal elements of the Mueller matrix.

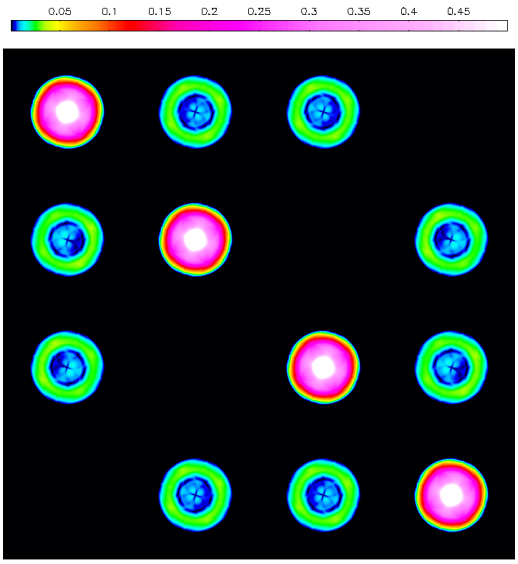
The Mueller matrix is a diagonally dominated entity, with the central illumination exceeding the off-diagonal elements by $\sim 10^2$. While the diagonal elements of the matrix have more power focussed in the central portions of the antenna surface, the off-diagonal elements (denoting leakage) show more power in the periphery of the dish rather than at the center. The AIP in the off-diagonal elements is maximal around the struts that support the secondary reflector indicative of the role of interfering physical structures in the optical ray path.The struts and the bars holding the secondary ensure that the central portion of the



(a) Real part of the normalized full polarization complex AIP

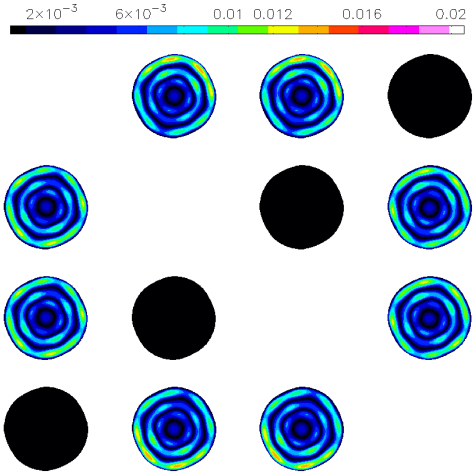


(b) Imaginary part of the normalized full polarization complex AIP

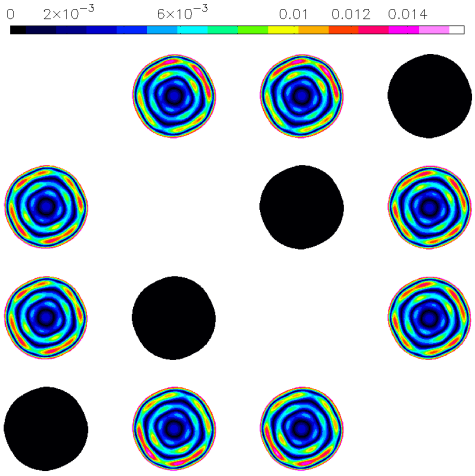


(c) Amplitude of the normalized full polarization complex AIP

Figure 5.1: Complex baseline AIP displayed in matrix form.



(a) Shown in the matrix above is the result of $\mathbf{A}^{M\dagger}(\nu,t)\star\mathbf{A}(\nu,t)$, the result of the convolution function applied during gridding as a part of A-Projection



(b) Shown in the matrix above is the result of $\mathbf{A}^{M\dagger}(\nu^{\star},t)\star\mathbf{A}(\nu,t)$, the result of the wide-band convolution function applied during gridding as a part of A-Projection

Figure 5.2: The inverted matrices post FM-AWP.

aperture for off-axis elements shows a quadrupolar illumination pattern, with alternating sign.

The correction for the full Mueller matrix is implemented as shown in the algorithm. 1, and is applied as a convolutional correction while resampling the data on to a regular grid. The resampling of the data is carried out using a model projection operator, $\mathbf{A}^{M\dagger}_{ab}(\nu)$ in the narrowband implementation of the algorithm and $\mathbf{A}^{M\dagger}_{ab}(\nu^{*})$ in the wide-band implementation. Utilizing the diagonal symmetry of the $\mathbf{A}$ matrix, $\mathbf{A}^{M\dagger}_{ab}$ is constructed and applied as

$$\mathbf{A}^{M\dagger}(\nu)\star\mathbf{A}(\nu)=\begin{pmatrix}A_{33}&A_{32}&A_{31}&A_{30}\\A_{23}&A_{22}&A_{21}&A_{20}\\A_{13}&A_{12}&A_{11}&A_{10}\\A_{03}&A_{02}&A_{01}&A_{00}\end{pmatrix}\star\begin{pmatrix}A_{00}&A_{01}&A_{02}&A_{03}\\A_{10}&A_{11}&A_{12}&A_{13}\\A_{20}&A_{21}&A_{22}&A_{23}\\A_{30}&A_{31}&A_{32}&A_{33}\end{pmatrix} \tag{5.1}$$

The matrix $\mathbf{A}^{M\dagger}$ within the A-Projection framework is constructed under the assumption that

$$(\mathbf{A}^{M\dagger}\star\mathbf{A}^{M})/\mathscr{F}^{\dagger}Tr(\bar{\mathbf{A}}^{M})\simeq\mathbb{1} \tag{5.2}$$

The level to which FM-AWP corrects for the DD effects of the Mueller matrix can be determined by how close to unity the projection operation shown in Eq. 5.2 achieve. The baseline Mueller matrix derived by performing the projection operation is shown for the narrowband (Fig. 5.2a) and wide-band (Fig. 5.2b) cases. For an ideal projection operation the baseline Mueller matrix should be real with negligible (nearly zero) phase. The Mueller matrix is near unity along the diagonal (deviations occur 1×10^{-4} for the narrowband and 3×10^{-4} for the wide-band case, with zero phase in both) and the off-diagonal AIP are non zero particularly around the edges of the aperture and the base of the struts. Despite the residual non zero structure the matrix has minimal phase and is nearly real both on and

off-axis. The residual structure in the off-diagonal elements then provides the extent to which one can correct the antenna AIP using the projection operator of the **A** matrix.

Image plane, azimuthal slices through the first off diagonal element $\mathscr{F}^{\dagger}\mathbf{A}_M^{32} \star \mathbf{A}^{01}$, with $\nu^* = \nu$ are shown in Fig. 5.3 for the narrowband (blue curve) and wide-band(green curve) cases. FM-AWP reduces the off-axis leakage in the narrowband case under 8.4×10^{-5} of its original value across the main-lobe of the antenna PB. In the wide-band case, the suppression is more than the narrowband suppression with the off-axis leakage reduced to under 4×10^{-5} across the main-lobe of the antenna PB. From the plotted images of the off-axis elements post projection, I also conclude that the FM-AWP is effective in removing off-axis polarization flux leakage for all the terms of the Mueller matrix during imaging for polarization dynamic ranges up to 12000, in the narrowband scenario, and up to 25000 in the wide-band scenario.

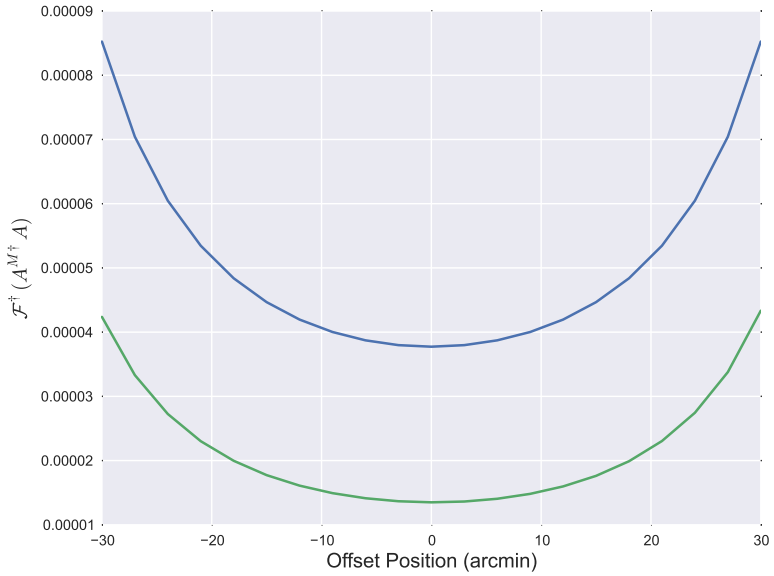


Figure 5.3: Plotted in the figure is a profile through the FM-AWP corrected image of the first off-axis leakage element $\mathscr{F}^{\dagger}\mathbf{A}_M^{32} \star \mathbf{A}^{01}$. The the blue curve is an azimuthal slice through the narrowband leakage residuals post FM-AWP and the green curve is an azimuthal slice through the wide-band correction $\mathscr{F}^{\dagger}\mathbf{A}_M^{32}(\nu^*) \star \mathbf{A}^{01}(\nu)$, post FM-AWP.

5.2 Simulations with point sources

The simulations described in this section are a polarization extension of the simulations work carried out by Rau et al. [45]. Simulations were carried out to validate the FM - AW Projection and contrast it against standard imaging. Deep field simulations were carried out to demonstrate the performance of the algorithm in a situation mimicking real sky. In each instance, the simulations were carried out with unpolarized, and polarized sources to estimate the imaging performance of the FMA-Projection algorithm in correcting flux leaking from one polarization into another. The simulations compare the imaging results based on the metrics of fidelity, dynamic range and nominally the sensitivity. Polarization

flux fidelity is required to accurately reconstruct the true polarization signal from the sky. While the corrupting effects of the antenna PB were examined in chapter 3, here I examine through simulations the flux fidelity when imaging with the FM-AWP algorithm and compared to when FM-AWP is not utilized.

5.2.1 Simulated Deep Field

To simulate a radio deep field with a realistic source distribution a source list was obtained from the SKADS/S-Cubed sky simulation. Sources were chosen with total intensity values ranging from 50 mJy to 5 μJy at 1.4 GHz and spread across 1deg^2. Each source was treated as being polarized with a fractional linear polarization of 0.05 and having a polarization position angle of 45°. A total of 715 sources were placed at random positions within the PB gain value of 0.05 at the reference frequency and Gaussian noise of 1μJy was added. The sky model was used as the input polarized model image. A measurement set including polarization leakage was obtained from the model image following the procedure is shown in Fig. 3.3 spanning 128 MHz centered around the reference frequency. The simulations spanned $\pm 4^h$ in hour angle, equivalent to 218° in parallactic angle.

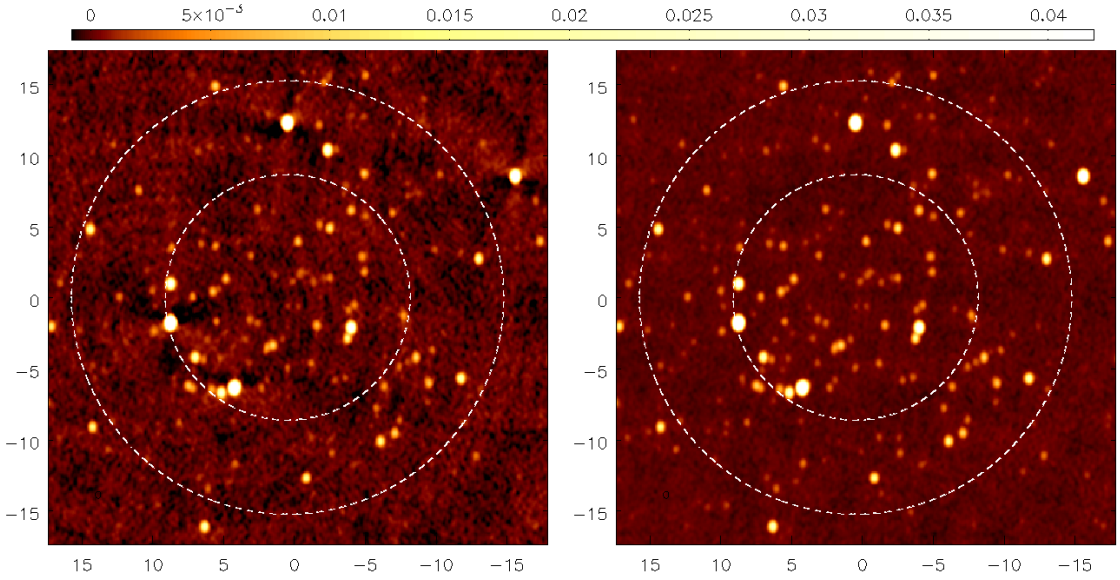


Figure 5.4: The panels shows a comparison of deep field imaging in Stokes I, without FM-AWP (left panel) and with FM-AWP (right Panel). The dashed contours represent the 80% (inner circle) and 50% in the PB.

The measurement set was then imaged with FM-AWP and with standard imaging (no DD correction). The total intensity images obtained are displayed in Fig. 5.4. The image in the left panel of Fig. 5.4 was constructed using standard imaging and shows deconvolution errors around bright sources in the form of spokes (the spokes may also be negative). The RMS noise in the image is $9\mu Jy$ and an imaging dynamic range of $\sim$ 5500 was achieved. The right panel of Fig. 5.4 shows the image obtained through FM-AWP. The bright sources are free from deconvolution errors, fainter sources are visible and the RMS noise in the central region of the image is much lower than in the left panel at $1.4\mu Jy$ corresponding to a dynamic range of $\sim$ 35000. FM-AWP imaging produced an RMS noise much closer to the noise in the input model of 1μJy.

FM-AWP is an extension of the WB-AWP algorithm [33] and corrects for the entire Mueller matrix across all Stokes parameters. A strong DD polarization effect is the *squint* of the VLA antennas, which manifests an instrumental Stokes V signal in the left panel of Fig. 5.5 peaking around the half power point of the PB. In contrast the right panel of Fig. 5.5 shows that the FM-AWP algorithm like its precursor, the WB-AWP, corrects for *squint* and produces as noise-like Stokes V image with an RMS of 1.4μJy. This is an important validation step for the FM-AWP algorithm that it is able to replicate the performance of WB-AWP in Stokes -I and -V.

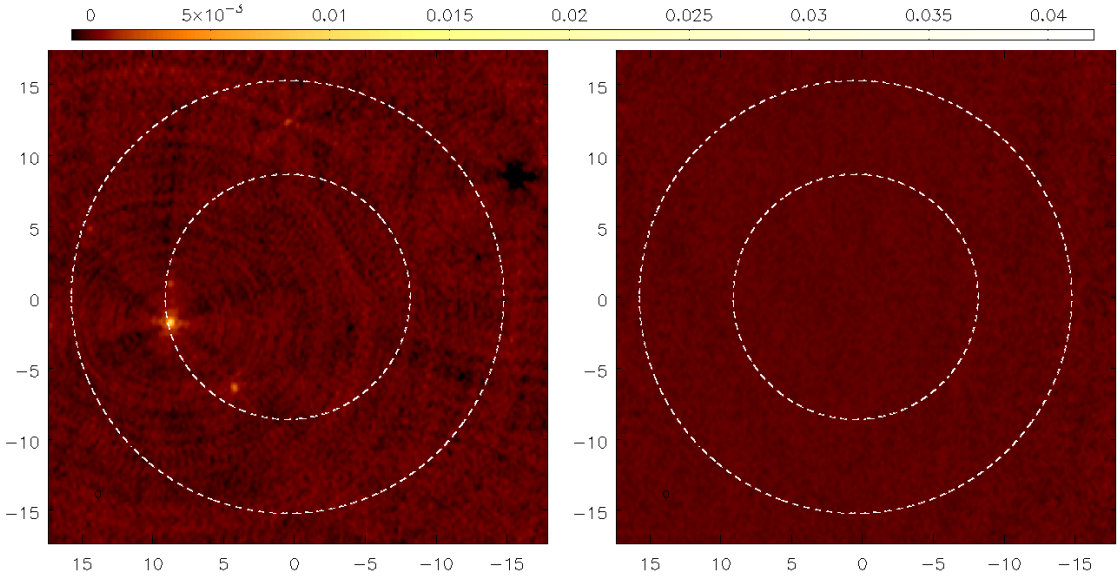


Figure 5.5: The panels shows a comparison of deep field imaging in Stokes V without FM-AWP (left panel) and with FM-AWP (right Panel). The noise in FM-AWP corrected image is a factor of 10 lower and the color scaling for the panel has been saturated to show the noise in the image. The scale bar has units of mJy for the left panel and units of $\times 10^{-}1$ mJy for the right panel. The dashed contours represent the 80% (inner circle) and 50% in the PB.

The left panel of Fig. 5.6 shows the polarized intensity image ($P = \sqrt{Q^2 + U^2}$) without FM-AWP correction. The brightest polarization sources in the image display deconvolution errors (as y shaped spikes either positive or negative) around them with the error increasing for bright sources further out in the beam. The noise in the polarized intensity image follows a Rayleigh distribution unlike the Gaussian distributed noise in the individual Stokes parameters. The RMS noise in the left panel of the polarized intensity image is $12.7\mu Jy$ and the effective dynamic range of about 400. In the right panel of Fig. 5.6 is the FM-AWP corrected polarized intensity image. The brightest sources in polarized intensity show no deconvolution errors around the center and the noise in the image is significantly lower at $1.4\mu Jy$ for an effective dynamic range of about 3500. With the decreased noise floor more sources are visible and are above the noise threshold for deconvolution. These comparisons demonstrate that FM-AWP improves imaging dynamic range in all Stokes parameters.

The last metric of comparison is the fidelity of imaging. To compare the fidelity of the images presented in the Fig. 5.4 and Fig. 5.6 the ratio of input flux density model

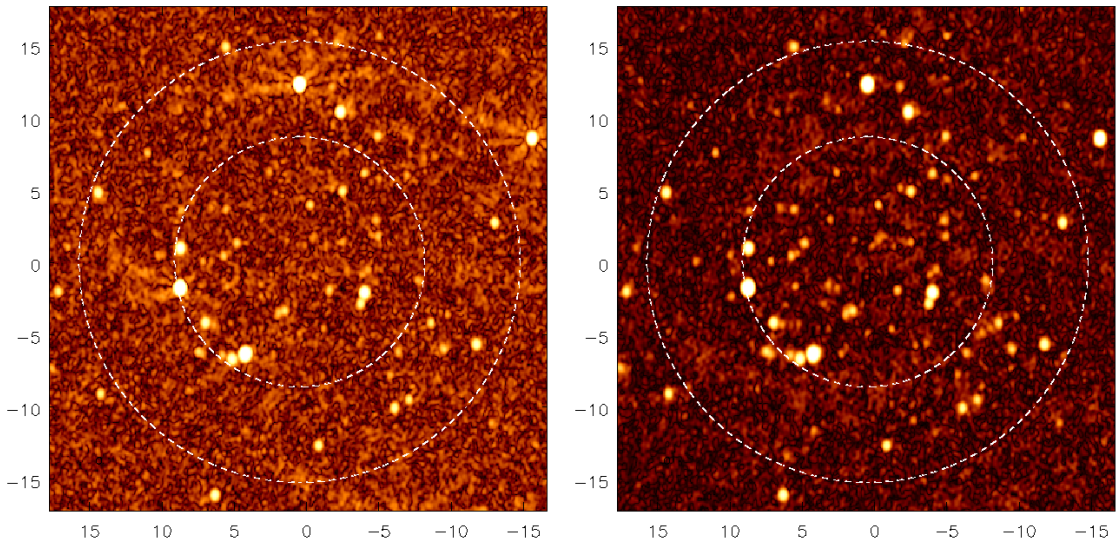


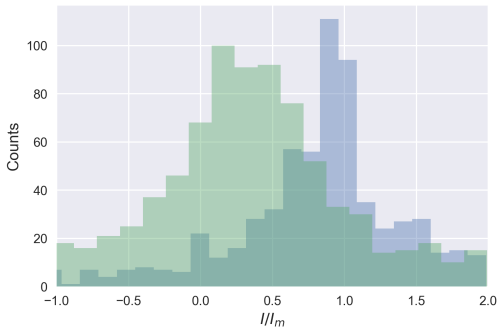
Figure 5.6: The panels provide a comparison of deep field imaging in polarized intensity ($P = \sqrt{Q^2 + U^2}$) without full Mueller A-Projection(left panel) and with full Mueller A-projection(right Panel)

from SKADS (represented by I_m, and P_m, in Stokes -I, and polarized intensity) to the PB corrected imaged flux density from the deep fields (represented by I, and P, in Stokes -I, and polarized intensity) was obtained. A histogram comparison of I/I_m and P/P_m are shown in Fig.5.7a, and Fig. 5.7c respectively. FM-AWP recovered flux values are represented by the blue histogram, while the flux recovered through standard imaging is shown in green. If the primary beam scaled images were restored to the true flux of the source, the histograms of the ratio of the imaged flux to the input model would be Gaussian distributed around unity with the standard deviation corresponding to the imaging noise. The FM-AWP recovered fluxes have a tighter distribution of the histograms (in Stokes -I and polarized intensity) around unity showing that the flux fidelity is better than standard imaging.

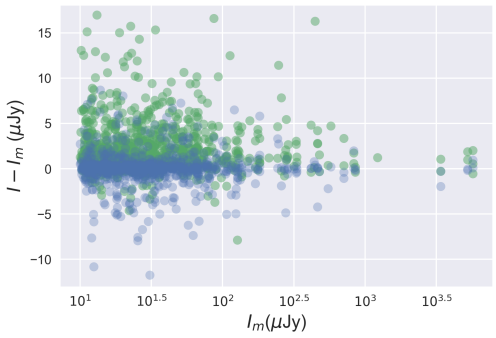
Plotted in Fig. 5.7b and Fig. 5.7d is the plot of the difference between the model and the imaged flux densities versus the model flux density (in logarithmic scale), with FM-AWP imaged values plotted in blue and standard imaging in green. In the case of perfect recovery of imaged flux, all the points would lie on the x-axis at zero, with the spread representative of the standard error in the measurement. The green points representing standard imaging are systematically offset from zero. In addition, the spread of the sources widens considerably more for the fainter sources where fractional error can be over 50% in the case of standard imaging. While bright sources are affected primarily by deconvolution errors, fainter sources are affected by *clean bias*[45] in addition to the deconvolution errors. The systematic shift seen for the majority of sources in the case of standard imaging is due in part to the flux leaking from Stokes -I into Stokes -Q and -U.

When FM-AWP was not utilized for polarimetric imaging it led to increased noise levels and significant deconvolution errors around bright sources, and given the crowded nature of the field, this leads to an error in the deconvolved flux of sources clustered around these bright sources. While FM-AWP reduces the effect of deconvolution errors

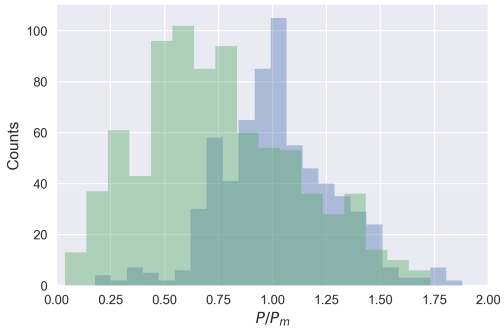
around strong sources as seen in the right panel of Fig. 5.4, it does not eliminate the effect of *clean bias* as seen in the spread of the fainter blue circles in Fig. 5.7b. *Clean bias* is the phenomenon that systematically shifts the intensity of the reconstructed peak values of clustered sources to lower than expected values. *Clean bias* arises from the subtraction of the side-lobe flux of the brighter source from the fainter source particularly, in densely clustered fields. The problem is exacerbated when the initial residual image passed on to deconvolution has uncorrected DD errors. It is also worthwhile remembering that as discussed in chapter 2 that CLEAN (Hogbom) is ideal for reconstructing sparsely populated point source images. While the deep field simulations here are of point sources it is definitely not a sparsely populated.



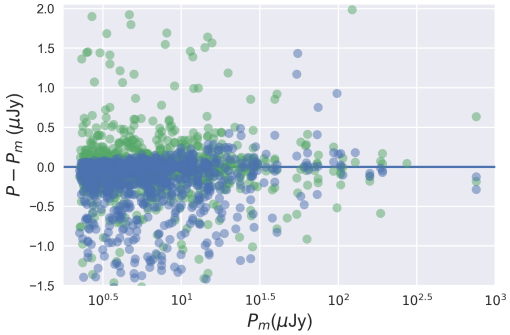
(a) Histogram of the ratio of the reconstructed flux values in Stokes I.



(b) The error in the reconstructed flux density in Stokes $I - I_m$ is plotted against the model flux density I_m.



(c) Histogram of the ratio of the reconstructed flux values in Stokes P.



(d) The error in the reconstructed flux density in Stokes $P - P_m$ is plotted against the model flux density P_m.

Figure 5.7: Histogram of the ratio of the reconstructed flux values, in Stokes I and in linear polarized intensity P are plotted in the panels on the left. The error in the reconstructed flux density in Stokes I and polarized flux density is plotted against the model flux density in the panels on the right. In each of the cases FMAWP imaged sources are plotted in blue and the standard imaged sources in green.

5.3 Full Mueller Imaging of Supernova Remnant G55.7+3.4

G55.7+3.4 is a supernova remnant observed by the enhanced VLA [59], the data are available in the public domain. The observations were designed to produce an imaging test dataset spanning $1-2$ GHz with the supernova remnant occupying the entire PB of the telescope at 2 GHz. Being a complex extended source that is significantly affected by DD effects across the band, it is an ideal reference source to test the new FM-AWP algorithm. The data were taken on August 23^{rd}, 2010 in D-configuration with the WIDAR correlator. There were four non-working antennas which did not have the upgraded wide-band receivers at that time. Two primary flux calibrators 3C147 and 3C286, and a secondary phase calibrator J1925+2106 were observed. J1925+2106 was observed every 30 minutes covering a wide array of parallactic angles over the 8-hour observation.

5.3.1 Calibration

The data were flagged using auto-flagging algorithms implemented in CASA via the **flagdata** task and calibrated using the procedure outlined in the CASA guide[1] to produce the bandpass and gain calibration tables. For wide-band polarization calibration, additional calibration steps are required. The model for the primary calibrator included was frequency-dependent (Perley and Butler [54]). 3C286 is linearly polarized (Perley and Butler [55]) with a known partial polarization spectral index across the band, with an $RM = 0$ and a polarization position angle of 33° and hence is a primary polarization angle calibrator. The **setjy** task within CASA was used to set the flux density model in total intensity, utilizing the spectral index and curvature terms. The task is also capable of setting a frequency-dependent polynomial model in fractional polarization (Π_L) and polarization position angle (ψ) along with the source rotation measure ($RM = 0$). This was used to set the flux density model of the polarization angle calibrator 3C286.

In addition to the total intensity calibration, polarization angle calibration on a per channel basis was carried out using 3C286 as the reference source. The secondary phase calibrator J1925+2106 was used to compute the polarization leakage with no constraints on the source polarization, using the Xf+QU mode in the **polcal** task in CASA. The polarization calibration along with the gain and bandpass calibration tables derived from the CASA guide was applied to the data. Spectral window 5 (centered at 1.576 GHz) was not used for calibration or imaging due to excessive RFI in this part of the frequency band.

5.3.2 Full Mueller Imaging

The data were imaged in all Stokes polarization using W-Projection (major cycle) and MT-MFS (minor cycle), a combination referred to here on as "standard imaging". The data were also imaged using FM-AWP (major cycle) and MT-MFS (minor cycle) in all Stokes polarizations. In both imaging cases 64 w-planes and 2 nterms were used. For the FM-AWP imaging, convolution functions were generated and cached to disk to be used during gridding. For each of the 64 w-planes, 12 convolution functions (as shown in Fig. 5.1 for w=0 are generated. One convolution function for each Mueller matrix element barring the 4 second-order elements) per Mueller matrix are generated, for a total of 768 convolution functions that are cached on disk. During gridding in FM-AWP imaging the

[1] `https://casaguides.nrao.edu/index.php/EVLA_Wide-Band_Wide-Field_Imaging:_G55.7_3.4-CASA4.4`

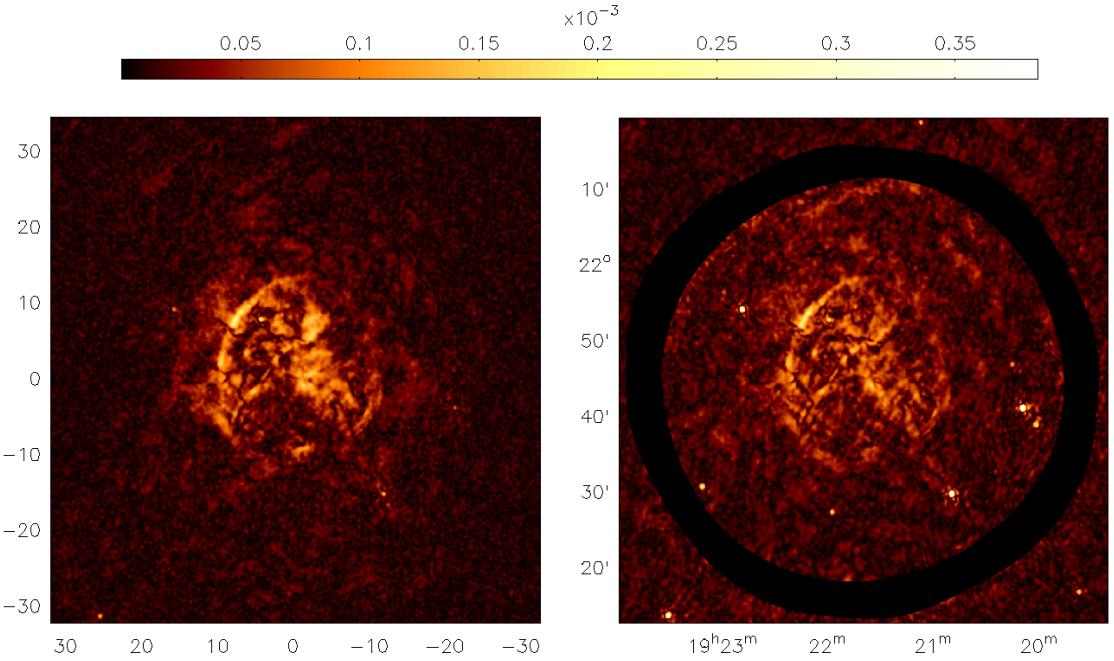


Figure 5.8: The panels provide a comparison of the wide-band polarized intensity image of the supernova remnant through standard imaging (left panel) and with FM-AWP(right Panel). The mask in the FM-AWP image is at a PB gain of 0.01

convolution functions are loaded from the disk and rotated for a change of 5° in parallactic angle. The left panel of Fig. 5.8 is the best possible image that can be produced utilizing only wide-band standard imaging. The right panel of the Fig. 5.8 is the image produced by FM-AWP imaging accounting for the wide-band DD effects of the baseline Mueller matrix. Fig. 5.8 shows the effects of DD corrections on the observed linear polarized flux density as one moves further out in the antenna PB. In particular the right panel of Fig. 5.8 shows bright compact sources towards the edges of the PB that appear fainter in the left panel of the Fig. 5.8. These bright compact sources are of particular interest during RM synthesis.

5.3.3 Rotation Measure Synthesis

In order to perform RM synthesis, the supernova remnant was imaged as a spectral cube with one plane per spectral window of width 128 MHz. The cube contained 5 contiguous images in Stokes -I -Q -U -V spanning spectral windows 2 (1.256 GHz) through 7 (2.02 GHz). Since spectral window 5 was completely flagged due to the presence of excessive RFI its corresponding channel in the spectral cube was restored as an empty image (zero pixel value across all the Stokes planes). The spectral cube in the first instance was produced using the combination of the W-Projection algorithm for the major cycle and Hogbom CLEAN for the minor cycle to gives us a "standard imaged" spectral cube. The spectral cube in the second instance was produced using the combination of narrowband FM-AWP algorithm for the major cycle and Hogbom CLEAN for the minor cycle to provide a DD corrected spectral cube. The two spectral cubes were exported as fits files upon which RM synthesis was performed using **pyrmsynth**[2] [62]. The **pyrmsynth** code is a CLEAN like deconvolution algorithm that deconvolves the rotation measure spread

[2] `https://github.com/mrbell/pyrmsynth`

function (RMSF) from the band-limited Faraday depth function (FDF). The band-limited FDF as given by Eq. 3.5, is a function of the polarized intensity per channel $P_i(\lambda^2)$, where λ is the wavelength of the channel i. DD imaging corrections through FM-AWP alters the P_i per channel for sources lying farther out in the antenna PB. Since the RMSF is nearly identical for the two spectral cubes (RMSF depends only on bandwidth and channelization) and has an HPBW of ≈ 130 rad/m^2. So any difference in the derived Faraday depth function of the two channel cubes is a result of DD imaging corrections.

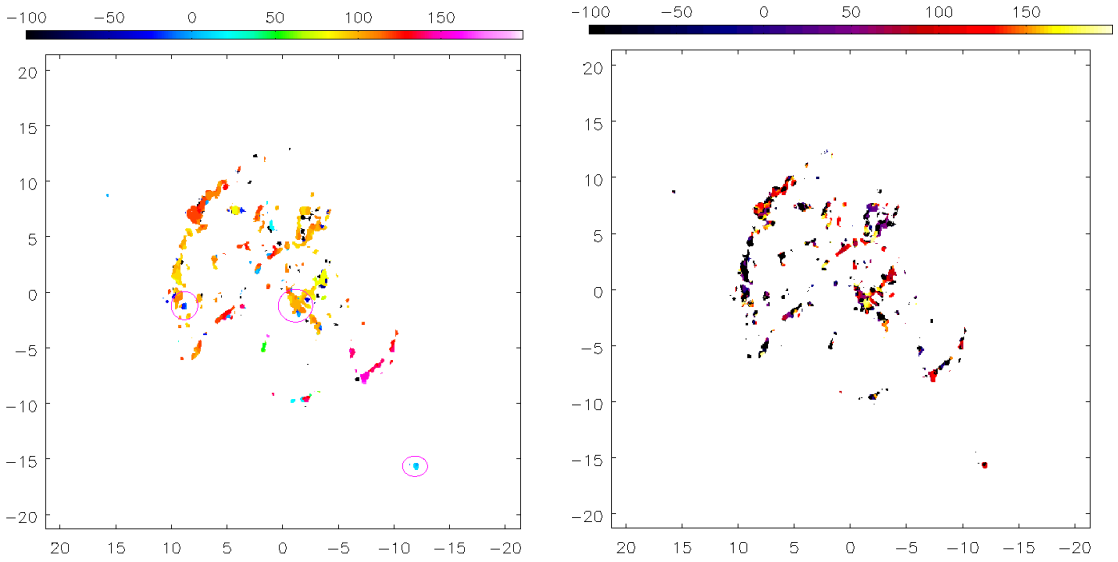
In chapter 3 I show that the Full Mueller Antenna PB needs to be accounted for in order to allow unencumbered deconvolution of RMSF from the FDF. So any uncorrected effects in the antenna PB will result in deconvolution errors in Faraday space that increases with distance from the center of the PB. To quantify the deconvolution error introduced in the cleaned RM map, the **pyrmsynth** code was run on the image cubes produced with and without FM-AWP imaging. Since the sources in this field are fainter than the NVSS detection threshold they do not have a prior RM value in [6]. However, the median RM derived from 60 of the nearest sources around the supernova remnant puts the median RM for the galactic latitude ($l = 55$), and longitude ($b = 3$) at 110 rad/m^2 [3]. In Fig. 5.9a the deconvolved RM of the supernova remnant derived from FM-AWP corrected image has a median RM of 125 rad/m^2. Fig. 5.9b shows the difference in the RM obtained when Faraday synthesis was performed on the image cube with and without FM-AWP. In regions of high signal to noise, the deconvolution of the RMSF from the FDF is similar in both instances at the center of the PB.

For a closer examination of the FDF and the effect FM-AWP has on the data, three lines of sights across the Faraday depth cube are plotted in the Fig 5.10. Plotted in blue is the FDF derived from the FM-AWP corrected cube while plotted in green are the FDF derived from standard imaging. In Fig. 5.10a the data are derived from the central region of the image around the 0.8 gain in the PB at the reference frequency 1.64 GHz, of diffuse extended emission in the supernova remnant. The plotted points differ from each other in polarized intensity by about 5 % where the polarized intensity peaks at an FDF of 120 rad/m^2, in line with the median RM predicted by Taylor et al. [6]. The Fig. 5.10c contains a point source (brightest source in the map, J1920+2129 on the south eastern side of the supernova remnant) located at a PB gain of 0.2 at the reference frequency of 1.64 GHz. The P_i contribution to the FDF in the case of FM-AWP is significantly different from the P_i in the standard imaging cube. This difference in the FDF results in a shifted peak of polarized intensity in the FDF form 120 rad/m^2 in the case of FM-AWP to ≈ -180 rad/m^2 for standard imaging, along with a 15% reduction in the peak polarized flux.

5.4 Conclusions

In this chapter, I introduced the FM-AWP algorithm and demonstrated the existence of the inverse of the **A** matrix in both the case of narrowband and wide-band imaging cases. I have shown that the FM-AWP algorithm is capable of correcting for the effects of full Mueller matrix down to 8.4×10^{-5} of the Stokes I flux in the main-lobe and the first side-lobe in the narrowband case and 4×10^{-5} of the Stokes I flux in the wide-band case. Through simulations of deep radio fields, I have demonstrated that the FM-AWP

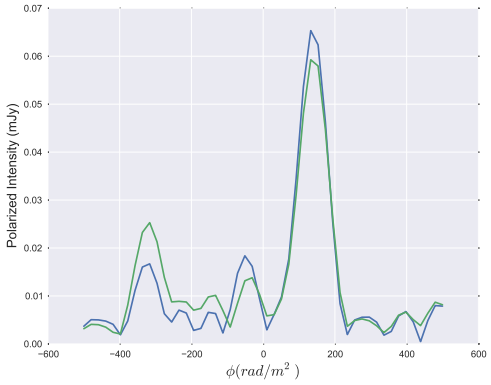
[3] The median RM is a reasonable estimate of the expected RM in the galactic plane as the plasmas probed would be galactic and lying at a similar depth in Faraday space



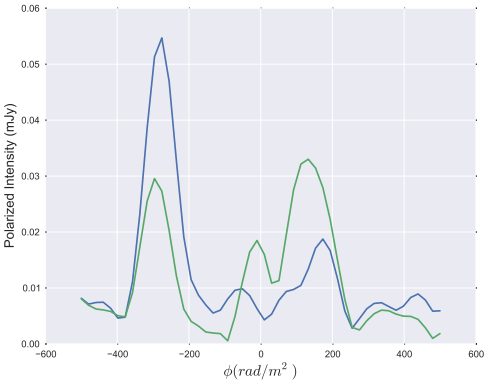
(a) The deconvolved Faraday depth image of the supernova remnant G55.7+3.4 derived from the FM-AWP corrected image cube. The median Faraday depth of the extended structure of the supernova remnant is 110 rad/m^2, inline with the median RM of the region derived from [6].

(b) The difference in the deconvolved Faraday spectrum with and without FM-AWP. The effect on the extended emission is particularly noticeable away from pointing center.

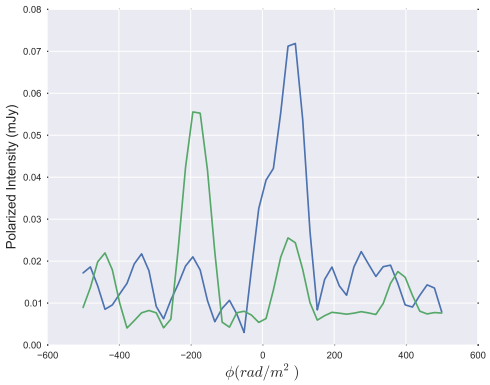
Figure 5.9: Shown in the left panel of the figure is the deconvolved Faraday of the FM-AWP data. The centers of the pink circles denote the regions for which sample RMSF are plotted in Fig. 5.10. Plotted in the right panel of the figure is the difference in the deconvolved spectrum $RM_{fmawp} - RM_{std}$.



(a) The Faraday spectrum of the supernova from around 0.8 PB gain. The FM-AWP corrected cube derived RM spectra is in blue and the standard imaging derived RM spectra is in green.



(b) The Faraday spectrum of the supernova from around 0.5 PB gain. The FM-AWP corrected cube derived RM spectra is in blue and the standard imaging derived RM spectra is in green.



(c) The Faraday spectrum of the compact source south-west of the supernova. The FM-AWP corrected cube derived RM spectra is in blue and the standard imaging derived RM spectra is in green.

Figure 5.10: Faraday depth profiles before and after FM-AWP correction applied during imaging of the supernova remnant G55.7+3.4. The Faraday depth profiles illustrate the change in polarized intensity after FM-AWP correction and the resulting changes in Faraday space.

algorithm outperforms direction independent imaging methods in both dynamic range and polarization flux fidelity. Having shown that the FM-AWP method works on simulated point source, I applied it to the VLA supernova remnant G55.7+3.4 data set. The data clearly show that FM-AWP is better at recovering the true channel polarized intensity (P_i) particularly for point sources farther out in the primary beam. Through Faraday synthesis, I also examined the effect of improper flux reconstruction in polarized intensity for two lines of sights in Faraday depth space. The stark difference for sources located outside the region of 0.8 PB gain is visible in terms of the shift in the location of the peak in polarized intensity and the recovered polarized flux density. I show that FM-AWP imaging produces a Faraday depth measurement for the supernova remnant that is consistent with the RM value predicted by Taylor et al. [6].

6. Algorithms & Computation

"If you were ploughing a field, which would you rather use ? Two strong oxen or 1024 chicken ?"

—Seymour Cray

Algorithmic development and computational advances proceed in tandem and are true for the work in this thesis too. The development of a new algorithm requires conceptualization, followed by iterative development and testing before a viable algorithm is born. The FM-AWP algorithm started as a conceptual extension of the convolution function production within the A-Projection framework. All of the code development was carried out within the Common Astronomy Software Application (CASA), research and development codebase. One of the primary features of CASA that allowed for this work to be possible was the implementation of the parallel **mpicasa**[1]. The computational work required for the completion of the thesis is on two separate fronts, both requiring parallel computation. The first was solving for the ray-traced antenna AIP on a per channel basis for L-Band holography measurements detailed in Sec. 6.1. The second is an implementation of the FM-AWP algorithm within the **CLEAN** task framework within CASA.

6.1 Parallel Optimization & Solving for parameterized AIP

In order to solve for antenna AIP, it was necessary to revisit the ray-tracing code that produces the per antenna Jones matrix. This is carried out by the BeamCalc module within CASA. The BeamCalc module contains the ray-tracing code that constructs the antenna geometry based on physical parameters that are band specific. From the constructed dish geometry the ray-tracing produces the AIP for a given frequency and parallactic angle

[1] `https://safe.nrao.edu/wiki/pub/HPC/CasaParallelization/CASA-%204.3-MPI-Parallel-Processing-Framework-Installation-and-advance-user-guide-v1.2.pdf`

specified. The first step was to test the production of the entire Jones matrix given a test data set. In order to ensure that the code exits after the production of the Jones matrix, (within the full imager framework gridding and the first major cycle would immediately follow normally) a test program in C++ was created. The program instantiates the AWProjectFT module responsible for WB-AWP and ensures that the BeamCalc module contains the requisite imaging parameters such as the coordinate information of the source, the image size, frequency, and resolution. The program wrote to disk both the antenna AIP and image plane PB matrices. Comparison of the ray-traced theoretical AIP to the measured PB from holography revealed that the theoretical AIP was underestimating the gain in the first side-lobe. In order to solve for the side-lobe as a function of frequency, in terms of the input parameters of the antenna geometry in BeamCalc, I modified the calculation of the antenna geometry to accept user-defined inputs for pointing offset, central blockage and the taper polynomial for feed illumination. To optimize these parameters I wrote a wrapper function in Python that accepted user-defined input values and passed them to the standalone C++ code. This allowed us to pass the Python wrapper as a function to a **scipy** minimization routine. The minimization routine traversed the multi-dimensional parameter space providing parameter values to the ray-tracing code until the parameters converged. This minimization routine for a single frequency channel for one antenna takes 4.5 hours of runtime to complete running a single thread with sufficient RAM, with minimal disk I/O. Which meant a serial run for 1024 channels in an antenna would take 192 days of runtime.

As the optimization was compute-limited I moved the execution onto a 16 core machine at NRAO compute cluster. Utilizing Python **multiprocessing**[2] libraries an MPI based Python implementation I parallelized the minimization to run each channel independently allowing us to run 16 simultaneous processes on the machine in parallel. Since the parallelization utilizes sub-processes it is a process parallel implementation that reduced the compute time by a factor if 16. While significant I needed vastly more computation in order to run the optimization for multiple antennas. NRAO data management and systems (DMS) group provided us access to the Amazon web services (AWS) based on a compute grant they had obtained. Utilizing the AWS compute cluster I was able to run 1024 concurrent processes reducing the runtime per-antenna to 6 hours. In this fashion, I completed three antennas at a compute cost of nearly $120 per-antenna with a speed up in the compute time by a factor of 10^3.

The AWS cluster execution of my code was possible because of the support provided in particular by Erik Bryer and James Robnett of the DMS group. They helped me with cluster management and data transfer between the AWS cluster and NRAO, and in setting up a custom Linux environment required for the installation and execution of the CASA based code.

6.2 Full Mueller - AW Projection Implementation

The software implementation of FM-AWP is an extension of the wide-band AW-Projection software. In this section, I will lay out the framework of the code and the associated optimizations that are implemented within CASA that made a working algorithm possible.

The synthesis imaging framework[3] within CASA is encapsulated within a high-level

[2] `https://docs.Python.org/2/library/multiprocessing.html`

[3] `https://casa.nrao.edu/casadocs/casa-5.1.0/global-task-list/task_tclean/about`

task **clean**. It accepts user inputs for data, along with the choice of the iterative major and minor cycle algorithms and their termination criteria. The task then performs synthesis imaging on the input data following the imaging equations laid out in chapter 2. CASA has a base level of C++ code with higher level Python layers made of *tools* which are in turn used to construct *tasks*. This layered model allows for swift testing of new ideas in Python before the final implementation is carried out in C++. In order to implement FM-AWP, I restricted ourselves to the existing A-Projection framework. Once a user invokes A-projection for all Stokes parameters, the **clean** task triggers the C++ module *AWProjectFT* at the base level. FM-AWP is implemented in the code within this module. In the following pseudocode function in italics highlight, the regions of the code that required changes towards the implementation of the FM-AWP algorithm.

Code Outline Full Mueller A-Projection Module Overview

Initialize the full Stokes model and residual images, determine major cycle algorithm

Major Cycle Algorithm : FM-AWP

AWProjectFT

Setup grid and lattice parameters, Stokes parameters

Invoke *PolOuterProduct (POP)*

Define **hybrid** mode Mueller configuration (12 CFs), and initialize *POP::init*
POP::MakePolIndex - Map Stokes to feed parameters
POP::MakeConjPolIndex - Map Stokes to conjugate of feed parameters
POP::MakePolMat - Initialize the Mueller matrix indices (A)
POP::MakeConjPolMat - Initialize Conjugate Mueller indices $(A^{M\dagger})$

Invoke *AWConvFunc(AWCF)* module
AWCF:: findConvFunc - locate cached CF.

AWCF:: makeConvFunc. - calls *VLACalcIlluminationConvFunc (VCICF)*
VCICF: applyPB - Determine frequency, PA, CF size - Invoke *BeamCalc*.
BeamCalc::skyMueller - Ray traced Jones matrix per-antenna, PA, spw.
VCICF::makeMueller - Mueller matrix for the baseline $i - j$, per PA, per spw.
Pass CF and data to the gridder - *AWVisResampler(AWVR)* module
AWVR::accumulateToGrid the forward step to produce gridded visibilities
Pass visiblity buffer(subset) and CF to fortranized gridding loops
AWVR::accumulateFromGrid the reverse step to produce de-gridded visibilities
Pass gridded visiblity buffer(subset) and CF to fortranized de-gridding loops

The progression of the code development was both linear and modular. Initial changes were performed within the **AWConvFunc (AWCF)** module (refer pseudocode shown) to implement the ray-traced AIP, which formed the focus of the chapters 3 and 4. Upon successful testing of the CF's produced by the code, I implemented the **PolOuterProduct**

(POP) module to ensure the polarization mapping between the correlations in the visibility and the imaging Stokes parameters for **hybrid** mode. The hybrid mode excludes the second order anti-diagonal term of the Mueller matrix, thus reducing the total number of terms per row of the Mueller matrix from 4 to 3. As a final step, I altered the **AWVisResampler (AWVR)** gridding and de-gridding module to accept the visibility to Stokes mapping and ensured that it was passed on to the Fortranized loops which perform the gridding based only on the mappings specified.

The CF production in full Mueller A-Projection as implemented in CASA is on a per spectral window basis (typically 16 to 32 spectral windows across a VLA observing band). While higher spectral resolution CF production is possible, the necessary changes in code have not been implemented at the task level of CASA till date. The CFs are produced once at the start of the imaging and are cached by *AWProjectFT*. Parallelization of the CF production is what made the FM-AWP testing a reality. The parallelization of the CF production utilizes the parallel MPI CASA imaging module developed by the CASA high-performance computing group along with the CASA algorithms group. Along with the MPI parallelization, additional changes were carried out to enable smarter cached CF management which allows for only the CF used in the gridding of the subset of data to be read in rather than reading in the entirety of the cache and holding the cache in memory. While these are extremely important updates it must be noted that FM-AWP algorithm development only played a role in testing the implementation of the memory management code. The smart CF cache was developed by the CASA algorithms R&D group independent of the efforts to develop the FM-AWP algorithm.

In a typical A-Projection imaging cycle, 80% of the time is spent in the gridding of the data. The convolution function production is a significantly lesser fraction of the imaging runtime and is a one time cost. With the parallelization, of the gridder through MPI framework in CASA, the gridding cost can now be distributed across multiple processors and nodes efficiently. The parallelization offers a scalable solution for addressing the increased computational complexity of the projection algorithms. With the rise in the availability of computing infrastructure (AWS, Google Cloud), parallel implementations of interferometric imaging algorithms are the way forward.

7. Conclusion

"We can only see a short distance ahead, but we can see plenty there that needs to be done."

—Alan Turing

In line with the thesis goals outlined in chapter 1, I have developed a Projection algorithm (FM-AWP) that corrects for the effects of the Mueller matrix during the course of the imaging major cycle. I demonstrated the effects of the instrumental Mueller matrix on polarimetric imaging with particular emphasis on Faraday synthesis in chapter 3. I developed a frequency-dependent high spectral resolution model of the antenna AIP to be used within the FM-AWP as a part of the convolution function production in chapter 4. I demonstrated the workings of the FM-AWP on VLA data in chapter 5. In the following sections 7.1, I will summarize the work required to achieve FM-AWP algorithm and its current capabilities and limitations and examine the road ahead for further algorithmic development in Sec. 7.2

7.1 Full Mueller A-Projection & Polarimetric Imaging

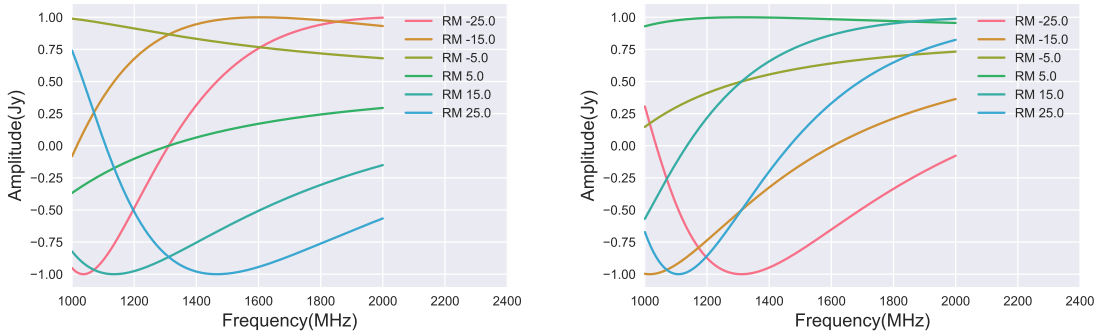
The initial stage of the FM-AWP algorithm was to evaluate the effect of the Mueller matrix on imaging. I characterized the error induced in polarized flux due to the time-varying Mueller matrix through simulation of point sources distributed across the field of view of the antenna PB for the VLA. The tests showed that flux leakage from Stokes -I into -Q and -U is significant up to 10% of the total intensity. Furthermore, I show that the leakage changes with the changing PA for an alt-az mounted telescope such as the VLA. Long integration observations of $\Delta PA > 100^{\circ}$ leads to a reduction in the polarization flux leakage due to averaging. However, flux leakage vanishes only for a perfectly symmetrical hour angle coverages around the meridian or for observations with $\Delta PA = 180^{\circ}$. I also

demonstrated that for realistic beam models with measurable second-order effects, such as beam *squash*, averaging of polarization leakage does not go to zero for $\Delta PA = 180^\circ$. I also examined the role of off-axis polarization leakage in the antenna PB in Faraday space and showed that it manifests as a peaked signal near RM of 0 rad m^{-2}. For realistic extragalactic sources with a Faraday space signatures ranging ± 100 rad m^{-2}, the off-axis polarization leakage acts as a distorting signal in the Faraday space. The distorted signal causes confusion for deconvolution algorithms such as RM Clean. FM-AWP algorithm corrects for these off-axis effects provided a model of the antenna AIP as a function of time, frequency and polarization are available.

In the second stage of the FM-AWP algorithm development, I considered ray-traced antenna AIP models that can be supplied as convolution functions to the algorithm to correct for the effects of the full Mueller matrix. The models did not reflect the frequency-dependent structure (such as a standing wave) across the bandwidth of L, S and C bands seen in Holographic measurements. In order to bridge the difference between the measured PB and the PB derived from the model AIP, I carried out parameter optimization of physically motivated antenna parameters. The parameters being optimized were the antenna pointing off-set, tapering of the feed illumination and the radius of the central aperture blockage. The parameters capture the frequency-dependent variations of the antenna PB per channel out to the first-side lobe. The optimized parameters were fit as a function of frequency using a straight line and sine function for the central blockage parameter. The fit captured the presence of a standing wave 17 MHz in frequency arising from a secondary reflection between the feed and the antenna secondary. The aperture illumination parameters were fit also with straight lines. This resulted in a highly compressed representation of the frequency-dependent behavior of the AIP (and consequently the PB). I also performed point source simulations to demonstrate that the frequency dependence of the antenna PB is required for high fidelity high dynamic range imaging for dynamic ranges of 100000:1 or higher.

As a culmination of the algorithmic work, I introduced the FM-AWP algorithm. It utilizes the unitary nature of the Jones matrix to project out the effects of the full Mueller matrix. I demonstrated the existence of the normalizable projection operator $\mathbf{A}^{M\dagger}(\nu^*) \star \mathbf{A}(\nu) = Tr(A)^2 \mathbb{1}$, for the narrowband and broadband instances of the algorithm. FM-AWP in the narrowband case reduces the off-diagonal elements ($\mathscr{F}^\dagger \mathbf{A}_{ij}, i \neq j$) of the Mueller matrix down to 4×10^{-5} in the narrowband implementation and 8.4×10^{-5} in the broadband implementation. Through point source simulations of radio deep fields, I showed that for crowded fields with many clustered sources the FM-AWP algorithm outperforms standard imaging (Cotton-Schwab CLEAN). Sources of known polarized flux were simulated to create a measurement set that was imaged using FM-AWP and standard imaging. The ratio of the recovered flux to the input flux was plotted as histograms in the case of Stokes -Q, -U and in polarized intensity. Imaging with FM-AWP led to tightly distributed histograms around unity while standard imaging resulted in extremely widely distributed denoting poor flux fidelity in terms of imaging. I then applied the FM-AWP algorithm to VLA data of the supernova remnant G55.7+3.4. FM-AWP reduced the depolarization due to the time-variable primary beam and removed flux leaking from Stoke -I into Stokes -Q and -U as demonstrated by the polarized intensity images.

In order to perform Faraday synthesis, the supernova remnant was imaged per spectral window to produce an image cube with both FM-AWP and standard imaging. The resulting



(a) Stokes -Q (b) Stokes -U

Figure 7.1: Shown in the two panels of the figure are the plots of Stokes -Q (7.1a) and -U (7.1b) as a function of frequency spanning the 1-2 GHz. The curves represent different RM's for a 1Jy source of polarization position angle 30°.

image cubes were processed with **pyrmsynth** to produce an RM map of the supernova remnant. The RM maps showed significant differences particularly in the outer regions of the supernova remnant and the compact source J1920+2129 to the southeast of the supernova remnant. The Faraday depth profiles determined from FM-AWP and standard imaging were nearly identical with both profiles peaking in Faraday space around 120 rad/m^2. The Faraday depth profiles for the supernova remnant close to 0.5 PB gain (at 1.5GHz) and the compact source J1920+2129 at 0.3 PB gain (at 1.5 GHz) show a 15% decrease in the magnitude of the peak of the Faraday depth profile and the location of the peak differs by 180 rad/m^2 in the case of FM-AWP as against standard imaging. Imaging with FM-AWP results in an RM in line with the predictions of the median Galactic RM map from [6].

7.2 Future Work

Next generation spectro-polarimetric surveys are mapping the radio sky with the intent of performing Faraday tomography. With the FM-AWP algorithm, I demonstrated that the Full Mueller matrix effects on polarization leakage can be removed during the imaging major cycle. The image produced by the major cycle is passed to the minor cycle for PSF deconvolution. For polarimetric observations needing RM synthesis, an image cube is produced containing images at multiple frequencies. Cube imaging is limited in sensitivity due to the smaller bandwidth utilized in producing each slice in the cube. Thus the sensitivity achievable in polarimetry and RM synthesis consequently is many-fold less than the sensitivity in Stokes I imaging. In order to achieve the full broadband sensitivity, the deconvolution algorithm must model the behavior of source across the imaging bandwidth.

Fig. 7.1 shows the spectral behavior Stokes -Q (Fig. 7.1a) and -U (Fig. 7.1b) of a synchrotron emitting source of spectral index $\alpha = 0.7$ across the L band of the VLA. The source is a compact source of flux density 1Jy, fractional polarization $\Pi = 0.05$ and polarization position angle $\psi = 30°$. The MT-MFS algorithm [43] offers a means of modeling the spectral behavior of the sky. The MT-MFS algorithm models the spectral behavior of the sky in terms of a frequency-dependent Taylor series polynomial. Each

term of the frequency-dependent polynomial (nterms ≥ 2) are gridded and deconvolved separately (for details refer Rau and Cornwell [43]) and the resultant images are used to reconstruct the spectral behavior of the source in total intensity such as the spectral index or spectral curvature. The MT-MFS framework extended to polarization would allow for the modeling of the frequency dependence of Stokes -Q and -U in terms of the Taylor polynomial. The fitted polynomials terms beyond the first (intensity in that polarization) can be modeled for different kinds of intervening plasma models (such as Faraday "thin" or Faraday "thick") and compared to the imaged data. This imaging method will allow for a) greater sensitivity (full bandwidth used in imaging), b) resolution (no smoothing required during spectral modeling). c) reduced overall computation (deconvolution is not per channel like in cube imaging but rather per term) and, d) faster Faraday space identification of the source model.

It can be seen that polynomial fits are not the ideal basis for fitting Stokes -Q and -U in the presence of larger rotation measures ($RM \geq |\pm 25|$) from Fig. 7.1. Alternatively, Fourier basis functions(sine and cosine) fitted to the source in wavelength squared space would allow us to model the source spectral behavior in Stokes -Q and -U for all measurable rotation measures. This would also result in an improved linear polarized intensity image along with fits to the Faraday space behavior of each source in the field of view. Such a method will allow for easy classification of sources in wide-field broadband surveys of the sky planned with the next generation of radio telescopes since most of the source in the sky (75% of the detected sources) are Faraday "thin" sources [25]. A polarized MTMFS algorithm is then a means of performing Stokes -Q and -U fitting in a model-independent fashion.

I have demonstrated the FM-AWP imaging method for the VLA, in general, it is applicable to other parabolic dish arrays such as the GMRT, MeerKAT, ASKAP, SKA and the ngVLA. The FM-AWP algorithm in its current form needs rigorous implementation and testing before integration within CASA for use by the wider user community. Imaging with the FM-AWP algorithm for other telescope arrays can be incorporated through the inclusion of their respective antenna AIP. With the antenna AIP incorporated in the FM-AWP code framework, high fidelity wide-field broadband polarimetric imaging for the next-generation of spectro-polarimetic surveys would be possible.

Appendix

Transform Matrix

In Ch. 2 we discussed the measurement equation in radio interferometry in detail. The antenna feeds in the interferometer measure the incident radio waves in some orthogonal basis such as linear or circular basis. The emission from the sky is however, in the Stokes basis. This transformation from the Stokes basis to the feed basis can be represented by a transform matrix $\mathscr{T}$. Let the transform matrix ${}_{pq}\mathscr{T}^{st}$ denote the transformation from the basis st to the basis pq. Then we can represent the transformation from the Stokes (st) to feed basis pq as ${}_{rl}\mathscr{T}^{st}$ in the measurement equation is

$$\vec{V}^{obs}_{ab;pq} = (\mathbf{J}_{a;p,q} \circledast \mathbf{J}^{*}_{b;p,q})_{pq}\mathscr{T}^{st}\vec{V}^{sky}_{ab;st}$$

where $\vec{V}^{sky}_{ab;st}$ the is the coherence vector in the Stokes basis and $\vec{V}^{obs}_{ab;pq}$ is the coherence vector in the feed basis. The transform matrix for circular basis is given by ${}_{rl}\mathscr{T}^{st}$, and for linear basis is given by ${}_{xy}\mathscr{T}^{st}$, where

$${}_{rl}\mathscr{T}^{st} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & i & 0 \\ 0 & 1 & -i & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix}$$

and

$${}_{xy}\mathscr{T}^{st} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & i \\ 0 & 0 & 1 & -i \\ 1 & -1 & 0 & -0 \end{pmatrix}$$

The transform matrix is a unitary matrix , i.e.$\mathscr{T}\mathscr{T}^{\dagger} = \mathbb{1}$, which means we can define the reverse transforms going from feed basis to Stokes basis as ${}_{st}\mathscr{T}^{pq} = {}_{pq}\mathscr{T}^{st\,\dagger}$. The

transform matrix for circular basis is given by ${}_{st}\mathscr{T}^{rl}$, and for linear basis is given by ${}_{st}\mathscr{T}^{xy}$, where

$${}_{st}\mathscr{T}^{rl} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & i & 0 \\ 0 & -i & i & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix}$$

and

$${}_{st}\mathscr{T}^{xy} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -i & i & -0 \end{pmatrix}$$

The transform matrix provides the instantaneous transformation from one basis to another. In the case of an Alt-Az mounted antenna, there is an additional parallactic angle rotation that needs to be included in the transform matrix. If the antenna parallactic angle is χ, the transform matrix due to the parallactic angle rotation for a baseline can be given in terms of the individual antenna parallactic angles as,

$$\mathscr{T}_{\chi} = \begin{pmatrix} e^{i\chi_a} & 0 \\ 0 & e^{-i\chi_a} \end{pmatrix} \otimes \begin{pmatrix} e^{i\chi_b} & 0 \\ 0 & e^{-i\chi_b} \end{pmatrix}^{*} = \begin{pmatrix} e^{i(\chi_a+\chi_b)} & 0 & 0 & 0 \\ 0 & e^{i(\chi_a-\chi_b)} & 0 & 0 \\ 0 & 0 & e^{i(\chi_b-\chi_a)} & 0 \\ 0 & 0 & 0 & e^{-i(\chi_a+\chi_b)} \end{pmatrix}$$

The effective transform matrix in the case of an Alt-Az mounted antenna is then $\mathscr{T}_{\chi} \cdot {}_{pq}\mathscr{T}^{st}$.

Bibliography

Books

[4]F. F. Chen, *Introduction to Plasma Physics and Controlled Fusion* (1984) (cit. on p. 8).

[27]A. G. Pacholczyk, *Radio astrophysics. Nonthermal processes in galactic and extragalactic sources* (1970) (cit. on p. 12).

[30]G. B. Taylor, C. L. Carilli, and R. A. Perley, eds., *Synthesis Imaging in Radio Astronomy II*, Vol. 180, Astronomical Society of the Pacific Conference Series (1999) (cit. on pp. 19, 53).

[31]A. R. Thompson, J. M. Moran, and G. W. Swenson, Jr., *Interferometry and Synthesis in Radio Astronomy, 3rd Edition* (2017) (cit. on p. 21).

[38]R. N. Bracewell, *The Fourier transform and its applications* (2000) (cit. on p. 28).

Articles

[1]C. L. Carilli, and G. B. Taylor, "Cluster Magnetic Fields", Annual Reviews of Astronomy & Astrophysics **40**, 319–348 (2002) (cit. on p. 7).

[2]D. M. Palmer, S. Barthelmy, N. Gehrels, R. M. Kippen, T. Cayton, C. Kouveliotou, D. Eichler, R. A. M. J. Wijers, P. M. Woods, J. Granot, Y. E. Lyubarsky, E. Ramirez-Ruiz, L. Barbier, M. Chester, J. Cummings, E. E. Fenimore, M. H. Finger, B. M. Gaensler, D. Hullinger, H. Krimm, C. B. Markwardt, J. A. Nousek, A. Parsons, S. Patel, T. Sakamoto, G. Sato, M. Suzuki, and J. Tueller, "A giant γ-ray flare from the magnetar SGR 1806 - 20", Nature **434**, 1107–1109 (2005) (cit. on p. 7).

[3]B. J. Burn, "On the depolarization of discrete radio sources by Faraday dispersion", Monthly Notices of the Royal Astronomical Society **133**, 67 (1966) (cit. on pp. 8, 41).

[5]M. A. Brentjens, and A. G. de Bruyn, "Faraday rotation measure synthesis", Astronomy & Astrophysics **441**, 1217–1228 (2005) (cit. on pp. 8, 9, 41).

[6]A. R. Taylor, J. M. Stil, and C. Sunstrum, "A Rotation Measure Image of the Sky", The Astrophysical Journal **702**, 1230–1236 (2009) (cit. on pp. 8, 10, 42, 74, 75, 77, 85).

[7]J. J. Condon, W. D. Cotton, E. W. Greisen, Q. F. Yin, R. A. Perley, G. B. Taylor, and J. J. Broderick, "The NRAO VLA Sky Survey", The Astronomical Journal **115**, 1693–1716 (1998) (cit. on pp. 8, 37).

[8]N. Oppermann, H. Junklewitz, G. Robbers, M. R. Bell, T. A. Enßlin, A. Bonafede, R. Braun, J. C. Brown, T. E. Clarke, I. J. Feain, B. M. Gaensler, A. Hammond, L. Harvey-Smith, G. Heald, M. Johnston-Hollitt, U. Klein, P. P. Kronberg, S. A. Mao, N. M. McClure-Griffiths, S. P. O'Sullivan, L. Pratley, T. Robishaw, S. Roy, D. H. F. M. Schnitzeler, C. Sotomayor-Beltran, J. Stevens, J. M. Stil, C. Sunstrum, A. Tanna, A. R. Taylor, and C. L. Van Eck, "An improved map of the Galactic Faraday sky", Astronomy & Astrophysics **542**, A93, A93 (2012) (cit. on pp. 8, 10).

[9]N. Oppermann, H. Junklewitz, M. Greiner, T. A. Enßlin, T. Akahori, E. Carretti, B. M. Gaensler, A. Goobar, L. Harvey-Smith, M. Johnston-Hollitt, L. Pratley, D. H. F. M. Schnitzeler, J. M. Stil, and V. Vacca, "Estimating extragalactic Faraday rotation", Astronomy & Astrophysics **575**, A118, A118 (2015) (cit. on p. 8).

[10]J. E. Kooi, P. D. Fischer, J. J. Buffo, and S. R. Spangler, "VLA Measurements of Faraday Rotation through Coronal Mass Ejections", Solar Physics **292**, 56, 56 (2017) (cit. on p. 9).

[11]A. H. Savage, S. R. Spangler, and P. D. Fischer, "Probing the Rosette Nebula Stellar Bubble with Faraday Rotation", The Astrophysical Journal **765**, 42, 42 (2013) (cit. on p. 9).

[12]C. R. Purcell, B. M. Gaensler, X. H. Sun, E. Carretti, G. Bernardi, M. Haverkorn, M. J. Kesteven, S. Poppi, D. H. F. M. Schnitzeler, and L. Staveley-Smith, "A Radio-Polarisation and Rotation Measure Study of the Gum Nebula and Its Environment", The Astrophysical Journal **804**, 22, 22 (2015) (cit. on p. 9).

[13]A. H. Costa, S. R. Spangler, J. R. Sink, S. Brown, and S. A. Mao, "Denser Sampling of the Rosette Nebula with Faraday Rotation Measurements: Improved Estimates of Magnetic Fields in H II Regions", The Astrophysical Journal **821**, 92, 92 (2016) (cit. on p. 9).

[14]S. A. Mao, N. M. McClure-Griffiths, B. M. Gaensler, J. C. Brown, C. L. van Eck, M. Haverkorn, P. P. Kronberg, J. M. Stil, A. Shukurov, and A. R. Taylor, "Magnetic Fields in the Milky Way Halo", Highlights of Astronomy **16**, 403–403 (2015) (cit. on p. 9).

[16]J. M. Lawler, and B. Dennison, "On intracluster Faraday rotation. II - Statistical analysis", The Astrophysical Journal **252**, 81–91 (1982) (cit. on p. 9).

[17]D. D. Sokoloff, A. A. Bykov, A. Shukurov, E. M. Berkhuijsen, R. Beck, and A. D. Poezd, "Depolarization and Faraday effects in galaxies", Monthly Notices of the Royal Astronomical Society **299**, 189–206 (1998) (cit. on p. 9).

[18]G. Heald, R. Braun, and R. Edmonds, "The Westerbork SINGS survey. II Polarization, Faraday rotation, and magnetic fields", Astronomy & Astrophysics **503**, 409–435 (2009) (cit. on p. 9).

[19]J. M. Stil, M. Krause, R. Beck, and A. R. Taylor, "The Integrated Polarization of Spiral Galaxy Disks", The Astrophysical Journal **693**, 1392–1403 (2009) (cit. on p. 9).

[20]D. Farnsworth, L. Rudnick, and S. Brown, "Integrated Polarization of Sources at $\lambda \sim 1$ m and New Rotation Measure Ambiguities", The Astronomical Journal **141**, 191, 191 (2011) (cit. on p. 9).

[21]D. Mesa, C. Baccigalupi, G. De Zotti, L. Gregorini, K.-H. Mack, M. Vigotti, and U. Klein, "Polarization properties of extragalactic radio sources and their contribution to microwave polarization fluctuations", Astronomy & Astrophysics **396**, 463–471 (2002) (cit. on p. 9).

[22]A. R. Taylor, J. M. Stil, J. K. Grant, T. L. Landecker, R. Kothes, R. I. Reid, A. D. Gray, D. Scott, P. G. Martin, A. I. Boothroyd, G. Joncas, F. J. Lockman, J. English, A. Sajina, and J. R. Bond, "Radio Polarimetry of the ELAIS N1 Field: Polarized Compact Sources", The Astrophysical Journal **666**, 201–211 (2007) (cit. on p. 9).

[23]C. A. Hales, R. P. Norris, B. M. Gaensler, and E. Middelberg, "ATLAS 1.4 GHz data release 2 - II. Properties of the faint polarized sky", Monthly Notices of the Royal Astronomical Society **440**, 3113–3139 (2014) (cit. on p. 9).

[24]L. Rudnick, and F. N. Owen, "The Distribution of Polarized Radio Sources greater than 15 microJy in GOODS-N", The Astrophysical Journal **785**, 45, 45 (2014) (cit. on p. 9).

[25]C. J. Law, B. M. Gaensler, G. C. Bower, D. C. Backer, A. Bauermeister, S. Croft, R. Forster, C. Gutierrez-Kraybill, L. Harvey-Smith, C. Heiles, C. Hull, G. Keating, D. MacMahon, D. Whysong, P. K. G. Williams, and M. Wright, "Spectropolarimetry with the Allen Telescope Array: Faraday Rotation Toward Bright Polarized Radio Galaxies", The Astrophysical Journal **728**, 57, 57 (2011) (cit. on pp. 9, 86).

[26]S. P. O'Sullivan, S. Brown, T. Robishaw, D. H. F. M. Schnitzeler, N. M. McClure-Griffiths, I. J. Feain, A. R. Taylor, B. M. Gaensler, T. L. Landecker, L. Harvey-Smith, and E. Carretti, "Complex Faraday depth structure of active galactic nuclei as revealed by broad-band radio polarimetry", Monthly Notices of the Royal Astronomical Society **421**, 3300–3315 (2012) (cit. on pp. 9, 43).

[28]J. P. Hamaker, J. D. Bregman, and R. J. Sault, "Understanding radio polarimetry. I. Mathematical foundations.", Astronomy & Astrophysics, Supplements **117**, 137–147 (1996) (cit. on pp. 19, 34).

[29]U. Rau, S. Bhatnagar, M. A. Voronkov, and T. J. Cornwell, "Advances in Calibration and Imaging Techniques in Radio Interferometry", IEEE Proceedings **97**, 1472–1481 (2009) (cit. on pp. 19, 24, 27).

[32]S. Bhatnagar, Cornwell, K. Golap, and J. M. Uson, "Correcting direction-dependent gains in the deconvolution of radio interferometric images", Astronomy & Astrophysics **487**, 419–429 (2008) (cit. on pp. 23, 27, 30, 47–49).

[33]S. Bhatnagar, U. Rau, and K. Golap, "Wide-field wide-band Interferometric Imaging: The WB A-Projection and Hybrid Algorithms", The Astrophysical Journal **770**, 91, 91 (2013) (cit. on pp. 23, 27, 30, 47, 57, 69).

[34]T. J. Cornwell, and P. N. Wilkinson, "A new method for making maps with unstable radio interferometers", Monthly Notices of the Royal Astronomical Society **196**, 1067–1086 (1981) (cit. on p. 24).

[35]T. Cornwell, "Recommendation for the AIPS++ Imaging Model", AIPS++ Note Series **148** (1992) (cit. on p. 24).

[36]T. Cornwell, "The Generic Interferometer I : Overview of Calibration and Imaging", AIPS++ Note Series **183** (1995) (cit. on p. 24).

[37]T. Cornwell, "The Generic Interferometer II: Image Solvers", AIPS++ Note Series **184** (1995) (cit. on p. 24).

[40]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 **2**, 647–657 (2008) (cit. on p. 30).

[41]B. Humphreys, and T. Cornwell, "Analysis of Convolutional Resampling Algorithm Performance", SKA Memo Series **132** (2011) (cit. on p. 30).

[43]U. Rau, and T. J. Cornwell, "A multi-scale multi-frequency deconvolution algorithm for synthesis imaging in radio interferometry", Astronomy & Astrophysics **532**, A71, A71 (2011) (cit. on pp. 30, 85, 86).

[44]P. Jagannathan, S. Bhatnagar, U. Rau, and R. Taylor, "Direction-Dependent Effects In Wide-Field Wideband Full-Stokes Radio Imaging", Astronomical Society of the Pacific Conference Series **495**, edited by A. R. Taylor, and E. Rosolowsky, 379 (2015) (cit. on pp. 33, 34).

[45]U. Rau, S. Bhatnagar, and F. N. Owen, "Deep Wideband Single Pointings and Mosaics in Radio Interferometry: How Accurately Do We Reconstruct Intensities and Spectral Indices of Faint Sources?", The Astronomical Journal **152**, 124, 124 (2016) (cit. on pp. 33, 67, 70).

[46]D. McConnell, J. R. Allison, K. Bannister, M. E. Bell, H. E. Bignall, A. P. Chippendale, P. G. Edwards, L. Harvey-Smith, S. Hegarty, I. Heywood, A. W. Hotan, B. T. Indermuehle, E. Lenc, J. Marvil, A. Popping, W. Raja, J. E. Reynolds, R. J. Sault, P. Serra, M. A. Voronkov, M. Whiting, S. W. Amy, P. Axtens, L. Ball, T. J. Bateman, D. C.-J. Bock, R. Bolton, D. Brodrick, M. Brothers, A. J. Brown, J. D. Bunton, W. Cheng, T. Cornwell, D. DeBoer, I. Feain, R. Gough, N. Gupta, J. C. Guzman, G. A. Hampson, S. Hay, D. B. Hayman, S. Hoyle, B. Humphreys, C. Jacka, C. A. Jackson, S. Jackson, K. Jeganathan, J. Joseph, B. S. Koribalski, M. Leach, E. S. Lensson, A. MacLeod, S. Mackay, M. Marquarding, N. M. McClure-Griffiths, P. Mirtschin, D. Mitchell, S. Neuhold, A. Ng, R. Norris, S. Pearce, R. Y. Qiao, A. E. T. Schinckel, M. Shields, T. W. Shimwell, M. Storey, E. Troup, B. Turner, J. Tuthill, A. Tzioumis, R. M. Wark, T. Westmeier, C. Wilson, and T. Wilson, "The Australian Square Kilometre Array Pathfinder: Performance of the Boolardy Engineering Test Array", Publications of the Astron. Soc. of Australia **33**, e042, e042 (2016) (cit. on p. 37).

[47]A. R. Taylor, S. J. Gibson, M. Peracaula, P. G. Martin, T. L. Landecker, C. M. Brunt, P. E. Dewdney, S. M. Dougherty, A. D. Gray, L. A. Higgs, C. R. Kerton, L. B. G. Knee, R. Kothes, C. R. Purton, B. Uyaniker, B. J. Wallace, A. G. Willis, and D. Durand, "The Canadian Galactic Plane Survey", The Astronomical Journal **125**, 3145–3164 (2003) (cit. on p. 37).

[48]J. C. Brown, A. R. Taylor, R. Wielebinski, and P. Mueller, "On Large-Scale Magnetic Field Reversals in the Outer Galaxy", The Astrophysical Journal Letters **592**, L29–L32 (2003) (cit. on p. 42).

[49]R. Perley, "Jansky Very Large Array Primary Beam Characteristics", VLA Memo **196** (2016) (cit. on pp. 43, 51).

[50]C. Heiles, P. Perillat, M. Nolan, D. Lorimer, R. Bhat, T. Ghosh, E. Howell, M. Lewis, K. O'Neil, C. Salter, and S. Stanimirovic, "All-Stokes Parameterization of the Main Beam and First Sidelobe for the Arecibo Radio Telescope", Publications of the Astron. Soc. of Pacific **113**, 1247–1273 (2001) (cit. on pp. 43, 52, 57).

[51]A. Popping, and R. Braun, "The standing wave phenomenon in radio telescopes. Frequency modulation of the WSRT primary beam", Astronomy & Astrophysics **479**, 903–913 (2008) (cit. on p. 47).

[52]P. Jagannathan, S. Bhatnagar, U. Rau, and A. R. Taylor, "Direction Dependent Corrections in Polarimetric Radio Imaging I : Characterizing the effects of the primary beam on full Stokes imaging", ArXiv e-prints (2017) (cit. on pp. 47, 57).

[53]W. Brisken, "Using Grasp8 to study the VLA beam", VLA Memo **58** (2003) (cit. on pp. 48, 50).

[54]R. A. Perley, and B. J. Butler, "An Accurate Flux Density Scale from 50 MHz to 50 GHz", The Astrophysical Journal Supplements **230**, 7, 7 (2017) (cit. on pp. 51, 72).

[55]R. A. Perley, and B. J. Butler, "Integrated Polarization Properties of 3C48, 3C138, 3C147, and 3C286", The Astrophysical Journal Supplements **206**, 16, 16 (2013) (cit. on pp. 51, 72).

[56]J. A. Nelder, and R. Mead, "A Simplex Method for Function Minimization", The Computer Journal **7**, 308–313 (1965) (cit. on p. 51).

[57]E. Momjian, R. Perley, and R. Hayward, "The Impact of the New Thermal Gap Receiver Assembly on the Sensitivity of the EVLA at L-Band (1?2 GHz)", VLA Memo **165** (2014) (cit. on p. 52).

[58] Astropy Collaboration, T. P. Robitaille, E. J. Tollerud, P. Greenfield, M. Droettboom, E. Bray, T. Aldcroft, M. Davis, A. Ginsburg, A. M. Price-Whelan, W. E. Kerzendorf, A. Conley, N. Crighton, K. Barbary, D. Muna, H. Ferguson, F. Grollier, M. M. Parikh, P. H. Nair, H. M. Unther, C. Deil, J. Woillez, S. Conseil, R. Kramer, J. E. H. Turner, L. Singer, R. Fox, B. A. Weaver, V. Zabalza, Z. I. Edwards, K. Azalee Bostroem, D. J. Burke, A. R. Casey, S. M. Crawford, N. Dencheva, J. Ely, T. Jenness, K. Labrie, P. L. Lim, F. Pierfederici, A. Pontzen, A. Ptak, B. Refsdal, M. Servillat, and O. Streicher, "Astropy: A community Python package for astronomy", Astronomy & Astrophysics **558**, A33, A33 (2013) (cit. on p. 54).

[59]S. Bhatnagar, U. Rau, D. A. Green, and M. P. Rupen, "Expanded Very Large Array Observations of Galactic Supernova Remnants: Wide-field Continuum and Spectral-index Imaging", The Astrophysical Journal Letters **739**, L20, L20 (2011) (cit. on pp. 57, 72).

[60]S. Bhatnagar, and T. Cornwell, "Solving for Antenna Based Pointing Errors", EVLA Memo Series **84** (2004) (cit. on p. 60).

[61] S. Bhatnagar, and T. J. Cornwell, "The Pointing Self Ccalibration Algorithm for Aperture Synthesis Radio Telescopes", The Astronomical Journal (2017) (cit. on p. 60).

[62] M. R. Bell, and T. A. Enßlin, "Faraday synthesis. The synergy of aperture and rotation measure synthesis", Astronomy & Astrophysics **540**, A80, A80 (2012) (cit. on p. 73).

List of Figures