Radio Interferometric Studies of Cool Evolved Stellar Outflows

A dissertation submitted to the University of Dublin
for the degree of Doctor of Philosophy

Eamon O'Gorman

Supervisor: Dr. Graham M. Harper

Trinity College Dublin, September 2013

SCHOOL OF PHYSICS
UNIVERSITY OF DUBLIN
TRINITY COLLEGE

Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work.

I agree to deposit this thesis in the University's open access institutional repository or allow the library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

Name: Your Name

Signature: .... **Date:** ....

Summary

You should write a nice summary here...

A dedication if you wish...

Acknowledgements

Some sincere acknowledgements...

List of Publications

Refereed

1. Richards, A. M. S., Davis, R. J., Decin, L., Etoka, S., Harper, G. M., Lim, J. J., Garrington, S. T., Gray, M. D., McDonald, I., **O'Gorman, E.**, Wittkowski, M.
   "e-MERLIN resolves Betelgeuse at wavelength 5 cm"
   Monthly Notices of the Royal Astronomical Society Letters, 432, L61 (2013)

2. **O'Gorman, E.**, Harper, G. M., Brown, J. M., Brown, A., Redfield, S., Richter, M. J., and Requena-Torres, M. A.
   "CARMA CO(J = 2 - 1) Observations of the Circumstellar Envelope of Betelgeuse"
   The Astronomical Journal, 144, 36 (2012)

3. Sada, P. V., Deming, D., Jennings, D. E., Jackson, B. K., Hamilton, C. M., Fraine, J., Peterson, S. W., Haase, F., Bays, K., Lunsford, A., and **O'Gorman, E.**
   "Extrasolar Planet Transits Observed at Kitt Peak National Observatory"
   Publications of the Astronomical Society of the Pacific, 124, 212 (2012)

4. Sada, P. V., Deming, D., Jackson, B. K., Jennings, D. E., Peterson, S. W., Haase, F., Bays, K., **O'Gorman, E.**, and Lundsford, A.
   "Recent Transits of the Super-Earth Exoplanet GJ 1214b"
   The Astrophysical Journal Letters, 720, L215 (2010)

Non-Refereed

1. **O'Gorman, E.**, & Harper, G. M.
   "What is Heating Arcturus' Wind?",
   Proceedings of the 16th Cambridge Workshop on Cool Stars, Stellar Systems and the Sun. Astronomical Society of the Pacific Conference Series, 448, 691 (2011)

Contents

List of Figures

List of Tables

1
Data Analysis

The complex visibilities outputted from the correlator of a radio interferometer are far from ideal and many additional steps of processing are required before they can be of scientific use. The imperfection of the synthesis radio telescopes (e.g., surface accuracy, receiver noise, gain stability, etc.), the adjustments to the signal (e.g., filter bandpass, etc.), hardware or software failures, poor atmospheric conditions, and the presence of RFI are some of the many sources of visibility corruption that must be accounted for before they can be Fourrier transformed to get the sky brightness distribution. This chapter describes the three main steps involved in reducing any standard radio interferometric data set, namely: data examination and flagging, data calibration, and imaging. In each step we give relevant examples of the data reduction techniques used in O'Gorman *et al.* (2012) and O'Gorman *et al.* (2013). A general work flow chart is given in Figure 1.1 which highlights the standard procedure required to go from raw visibilities to image analysis and summarizes what will be discussed in this chapter.
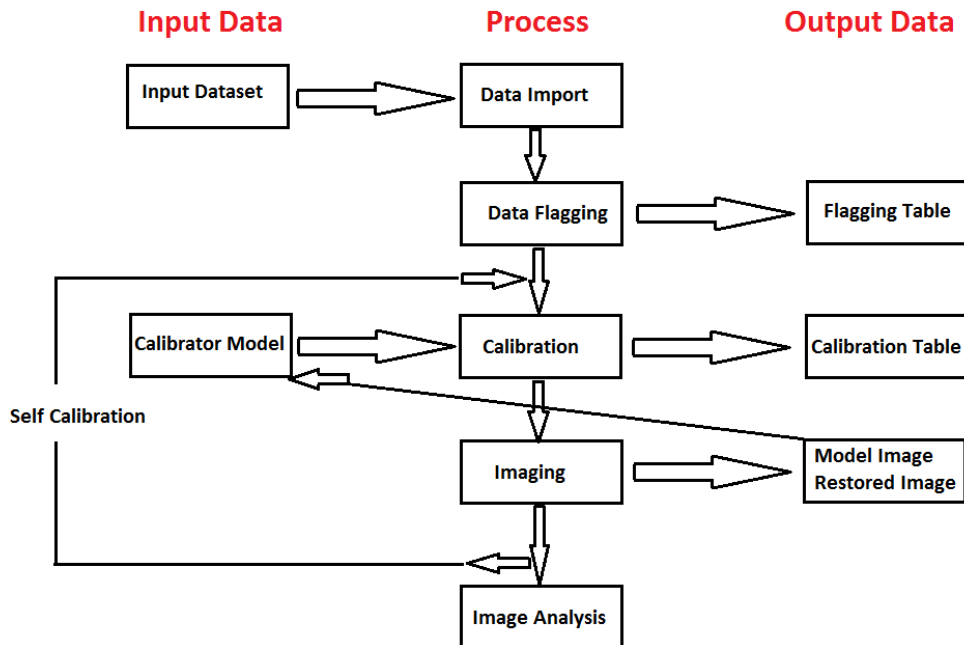
Figure 1.1: Work flow chart highlighting the general procedure required to go from the raw visibilities outputted by the correlator to a final radio image that can be used for scientific analysis. (Image adapted from the CASA cookbook, NRAO)

1.1 Data Examination and Flagging

The Common Astronomy Software Application (CASA; McMullin *et al.*, 2007) package was used to flag, calibrate, and image the main data sets used in this thesis. CASA is operated through a Python interface and uses a suite of astronomical data reduction tools which have been developed to meet the processing requirements of the large data sets from the Karl G. Jansky VLA and ALMA. It can also be used to process data from practically all other modern radio synthesis arrays. For synthesis data to be processed in CASA, it must be in a "measurement set" format. VLA data is easily transferable into this format using the *importevla* task within CASA. CARMA data files however come in *Miriad* format and need to be first converted into Flexible Image Transport System (i.e., FITS) format within the Miriad (Sault *et al.*, 1995) data reduction package. During this process the raw data are smoothed by a Hanning filter (combining adjacent frequency chan-

nels with weights 0.25, 0.5, and 0.25) to dampen ringing in the bandpass. Once in FITS format, the data can then be imported into CASA using the *importuvfits* task.

Once the data has successfully been imported into CASA the *listobs* task can be used to get a summary of the data set allowing the user to make sure the observing track contains the requested sources at the correct times. At this point it is also a good idea to check any observing logs which are created on site during the observation by the array operators. These logs usually contain important information about the specific track such as non-operational antennas, unavailable receivers, weather conditions etc., and such data will need to be treated appropriately during calibration or flagged at this stage. The *plotms* task, which is a GUI-style plotter, can then be used to obtain X-Y plots of visibility data. A good visual overview of the observation track is obtained by plotting all the source visibility amplitudes as a function of time. Averaging the data over channels or baselines often alerts the user of obvious bad data, which can be manually flagged through the *plotms* interface or through the *flagdata* task. An example of such a plot is shown in Figure 1.2 where the visibility amplitudes of three sources in a 1.3 cm VLA observing track for Aldebaran are displayed against time. The relatively weak target is shown between interleaving scans of the stronger phase calibrator, while the flux calibrator is the scan at the end of the observing track and is the strongest source in this observation. The data has been averaged over all channels and over a time period of x s. Some relatively low visibility amplitudes can clearly be seen for all scans of the phase calibrator which can be traced to a single poor performing antenna.

Another important way to represent the data at this stage of the flagging process is to plot the visibility amplitude of each of the sources as a function of $u - v$ distance or baseline length (i.e., $\sqrt{(u^2 + v^2)}$). The amplitude distribution should be relatively constant as a function of $u-v$ distance for the phase calibrator (i.e., a point source) and will fall off with increasing baseline length if the flux calibrator is resolved. Plotting the data in this format often results in extreme points at certain baselines which can then often be traced back to poorly behaving antennas or baselines. There is also some flagging which can be carried out based in part on a priori information. Antennas in very compact configurations can
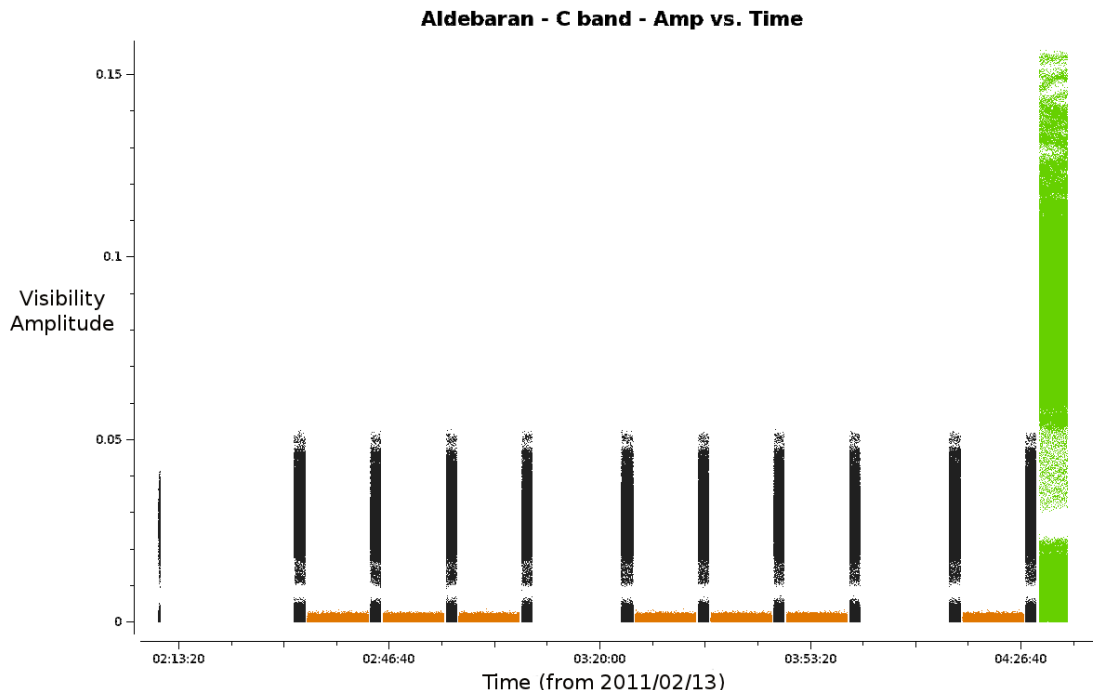
Figure 1.2: Data examination of a 6 cm VLA data set for Aldebaran. A good visual overview of the observation track is obtained by plotting all the source visibility amplitudes as a function of time. Averaging the data over channels or baselines sometimes allows rogue data to stand out. Here the black, orange, and green data points represent the phase calibrator, the science source, and the flux calibrator. The left most black data points are part of the dummy scan which has to be flagged. The low data points of the phase calibrator are data which needs further investigation. The absence of data at certain times is just when observations at 3 cm where being taken.

partially block the incoming RF signal to other antennas which is often referred to as antenna shadowing. This was not a problem in our VLA B configuration observations but some data obtained in the more compact CARMA configurations had to be flagged as a result of this shadowing. At the time our VLA observations took place, every observing track (i.e., scheduling block) needed to commence with at least a one minute *dummy scan* to facilitate the correlator setup. We subsequently flagged these scans as they contained no useful scientific data. Other a priori flagging was to remove any visibilities with zero amplitudes and to flag the edge channels in the CARMA data sets. Finally visual inspection of each

4

scan was carried out to determine if data at the beginning or end of these scans needed to be flagged. This processes is often referred to as quacking in radio interferometry analysis.

No RFI was present in any of the CARMA data sets and in any of the VLA data sets at wavelengths $\lesssim 3\,\text{cm}$ (i.e., X, K, Ka, and Q bands). For the 2011 long wavelength ($> 3\,\text{cm}$) data the two sub-bands were centered at relatively RFI free regions of the bandpass and only a very small amount of RFI had to be manually flagged. The 2012 wide-band data was taken at $10\,\text{cm}$ (i.e., S band) and $20\,\text{cm}$ (i.e., L band) and many of the sub-band were severely contained with RFI especially at $20\,\text{cm}$. In Figure 1.4 we plot the visibility amplitude of the flux calibrator 3C286 against frequency for the 2012 wide-band data set at $20\,\text{cm}$. The upper panel shows the raw data before any RFI has been removed. Some of the sub-bands were so contaminated with RFI that they had to be completely flagged. The remaining sub-bands were initially Hanning smoothed to suppress Gibbs ringing. This action spreads the single-channel RFI into three channels, but importantly removes the effects of some of the worst RFI from a number of channels and allows as much good data to be retained as possible. The *testautoflag* task was then used to conservatively flag RFI from all sources and any remaining RFI was manually flagged. The final result for the flux calibrator is shown in the bottom panel of Figure 1.4 with only some residual RFI still remaining, particularly between 1 and $1.2\,\text{GHz}$.

1.2 Calibration

The role of calibration is to correct the measured visibilities $V'(u, v)$ to approximate as closely as possible the true visibilities $V(u, v)$. As discussed in Chapter **??**, the true visibilities are related to the sky brightness via the Fourier transform:

$$V_{ij}(t) = \int A(l, m) I(l, m) \exp[-i2\pi(u_{ij}(t)l + v_{ij}(t)m)] dl dm \qquad (1.1)$$

where i, j represent the discrete sampling of the antennas i and j at time t. The term $u_{ij}(t)l + v_{ij}(t)m$ is the geometric phase difference produced by the geometric path length difference between antenna i and antenna j from the source (or part
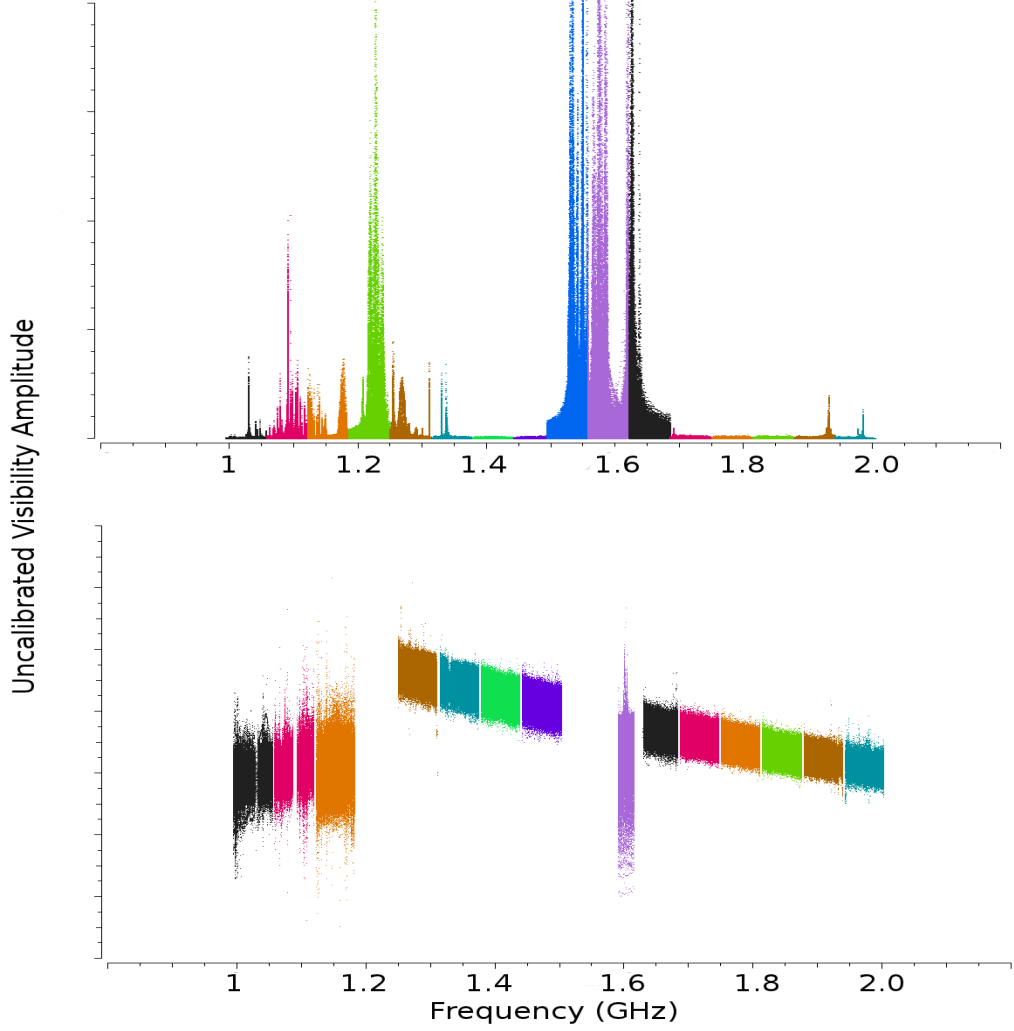
Figure 1.3: Eliminating RFI from the 20 cm wide-band data set. *Top panel:* Raw visibility amplitudes showing the presence of high levels of RFI in many sub-bands. *Bottom panel:* Post flagging visibility amplitudes. Some of the sub-bands were so severely contaminated with RFI that they had to be completely flagged. The data is still uncalibrated at this stage and the gain as a function of frequency is clearly present.

of) at location (l, m) relative to the phase center. The relationship between the measured visibility and the true visibility on a baseline between antennas i and j may be expressed as

$$V'_{ij} = J_{ij} V_{ij} \qquad (1.2)$$

6

where J_{ij} represents the accumulation of all corruptions affecting baseline ij. This equation is known as the Hamaker-Bregman-Sault Measurement Equation (Hamaker *et al.*, 1996). The most important of the effects contained in J_{ij} are antenna-based and arise from the measurable physical properties of individual antenna elements or the measurable physical conditions in the atmosphere above them. Thus, an array of N antennas forming $N(N-1)/2$ baselines can usually be adequately calibrated through the determination of only N factors.

For the purpose of the work presented in this thesis, the Measurement Equation can be written as

$$V'_{ij}(u, v, \nu) = b_{ij}(t)[B_i(\nu, t)B_j^*(\nu, t)](t)g_i(t)g_j(t)V_{ij}(u, v, \nu)e^{i[\theta_i(t) - \theta_j(t)]} \qquad (1.3)$$

where

- g_i and θ_i are the amplitude and phase portions of the complex gain which are usually determined separately in the calibration process and may change over the observation period with temperature, atmospheric conditions, etc.

- B_i is the complex bandpass, the instrumental response as a function of frequency, ν and may also vary over time.

- $b_{ij}(t)$ is the baseline term and is important shortly after a configuration change when antenna positions may not be known well.

The general calibration strategy is then to derive a series of scaling factors from both the phase and flux calibrators, which are then collectively applied to the science target in the final stage of calibration. A general workflow diagram describing the main steps involved in calibration process are summarized in Figure 1.4. We now discuss each of these steps while placing emphases on our CARMA and VLA data.

1.2.1 Prior Calibration

Our 2011 VLA data were acquired just after an array re-configuration which meant that the positions of some of the antennas were not accurately known at the time of observations. This resulted in some data points having inaccurate
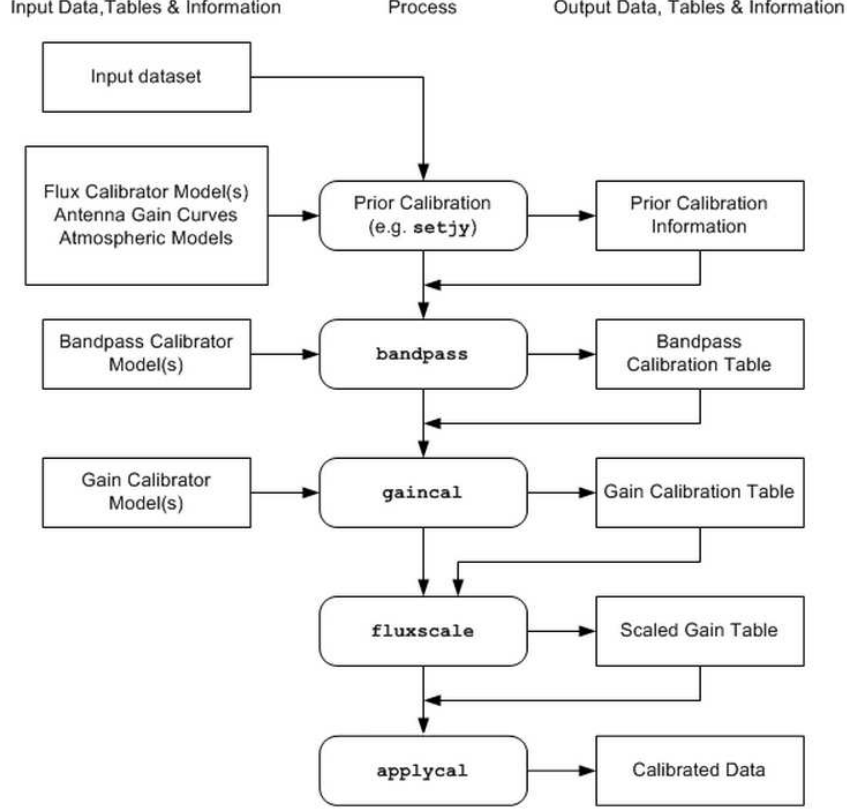
Figure 1.4: A workflow diagram outlining the main steps involved in calibrating radio interferometric data. Each of these steps are discussed in the text in relation to our CARMA and VLA observations.

$u - v$ coordinates. During the course of observations in each configuration, the exact position of all antennas become known and so the $u - v$ data could be calibrated to account for the discrepancies in the $u - v$ coordinates. At the VLA site, atmospheric opacites become significant at frequencies $\gtrsim 20\,\mathrm{GHz}$ as shown in Figure 1.4. Therefore, atmospheric opacity corrections were applied to the high frequency data sets. The adjustment values were based on the average of a seasonal model (based on many years of measurements) and information from the weather station obtained during the observations. For the CARMA data, the opacity at $1.3\,\mathrm{mm}$ is measured by a tipper (White, 2009). The tipper reflects radiation from the blank sky at different inclination onto a radiometer. The voltage from the radiometer can then be plotted against inclination to allow the

opacity to be calculated.

The final a priori calibration step is to provide a flux density value to the flux calibrator via CASA's *setjy* task. The VLA flux calibrators were assigned values using the "Perley-Butler 2010" flux density standard models and with assumed systematic uncertainties of 3% at all frequencies (Perley & Butler, 2013). At the time, no Ka or S-band flux density standard models were available so instead for these we used the K and L-band models, respectively, which were scaled according to their spectral indices. The absolute flux scale for CARMA observations is often determined by observing a planet and using a model of its flux as a function of baseline length. However, no such models were available in the early version of CASA and so the flux calibration was carried out with the quasers, 0530+135 and 3C120. The continuously updated CARMA flux catalog was accessed via the *xplore* GUI to obtain their flux values at each observation. The flux of these objects are more unpredictable and the systematic uncertainties are about 20%.
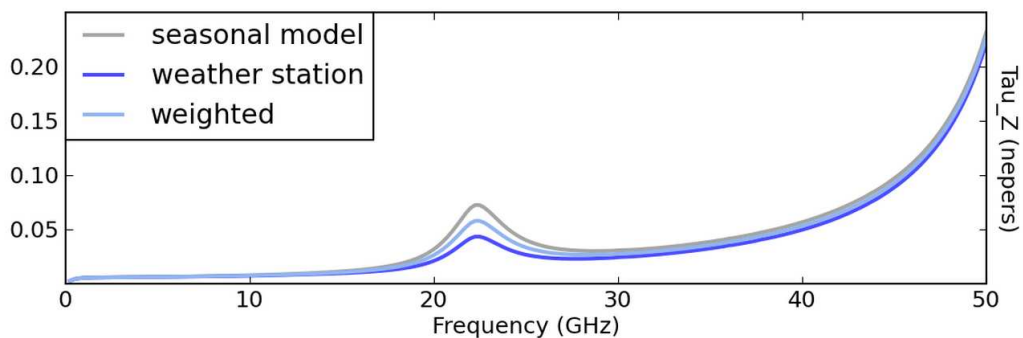

Figure 1.5: Calculation of the atmospheric opacity at the VLA site between $1 - 50\,\mathrm{GHz}$ on 2011 February 11. The adjustment values applied to the data were based on the average of a seasonal model and information from the weather station obtained during the observations.

1.2.2 Bandpass Calibration

The bandpass is the relative gain as a function of frequency and solving for it is the first part of the calibration process. Variation in frequency arises from frequency dependent effects in signal transmission. Such variation is shown in

Figure 1.6 for both phase and amplitude for a single antenna. For the VLA data, the flux calibrators were also used as the bandpass calibrators. Their phases were found to vary significantly over the $5 - 10$ minutes of observation especially at high frequencies; in most cases by a few 10s of degrees. Before a solution to the bandpass could be found, these phase variations needed to be corrected for to prevent decorrelation of the vector averaged bandpass solution. The complex bandpass, B_i could then be solved using the *bandpass* task and applying the antenna positions and phase solutions. The bandpass solutions were then applied to the bandpass calibrator to make sure both the amplitude and phase were then almost constant across channels/frequency.
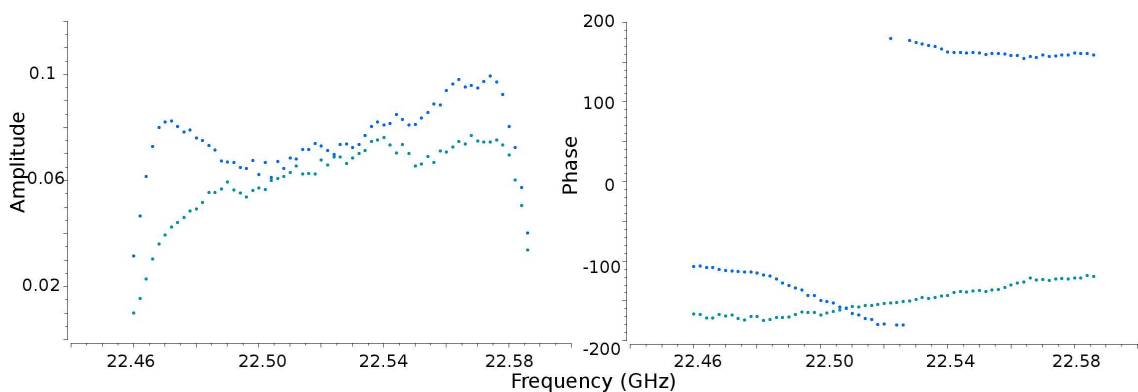

Figure 1.6: One antenna's gain variation as a function of frequency for the flux calibrator 3C138 at $1.3\,\mathrm{cm}$. Bandpass calibration corrects for this variation.

The CARMA data contained three spectral windows (i.e., 468, 62, and 31 MHz in width), each having independent bandpass shapes in both amplitude and phase. For each of these spectral windows, the gain calibrations are typically the same after the full bandpass dependence has been removed. The strategy to calibrate the CARMA data was to initially carry out the bandpass calibration on the wideband data (i.e., $468\,\mathrm{MHz}$) using the same strategy outlined in the previous paragraph and then carry out the gain calibration on this wideband data. Once this had been done, these bandpass independent gain solutions were applied to the narrow band data (i.e., 62, and $31\,\mathrm{MHz}$) while solving for their bandpass.

This narrow band bandpass solution could then be applied to the target and phase calibrator.

1.2.3   Gain Calibration

Once a bandpass solution has been applied, the next step is to derive corrections for the antenna amplitude and phase gains, g_i and θ_i, as a function of time. In order to determine the appropriate antenna-based complex gains for the science target, a phase calibrator which is always a point source and located much closer to the target than the flux calibrator, is regularly observed to minimize differences through the atmosphere. The general procedure then is to solve for these antenna-based gain factors for each scan on all calibrators. The amplitude changes on a much longer timescale than the phase and so they are solved separately. If there is a substantial change in phase over a scan and the un-corrected phases were averaged over this timescale, then the amplitude would be decorrelated. It is therefore important to correctly determine appropriate scan lengths when preparing the observations, especially at high frequencies, where time-dependent gain errors are introduced by the troposphere.

During gain calibration, the relative gain amplitudes and phases for different antennas are determined using the phase calibrator. The *gaincal* task is used to do this. The absolute flux density scale of the phase calibrator is later determined by comparison against the gain amplitudes g_i derived for the flux calibrator. To find the relative phases, a zero phase is determined by selecting a reference antenna for which the phase is defined to be zero. In the first step new solutions of complex gains g_i and θ_i are derived for the flux density calibrator which are corrected for the bandpass shape (here we assume the flux density calibrator has been used as the bandpass calibrator as was the case for our observations). The second and final step requires the determination of the appropriate complex gains from the phase calibrator.

1.2.4   Flux Scale Calibration and Application of Solutions

The penultimate stage of the calibration process is to use the known flux density of our flux calibrator (whose flux density was set using *setjy* above) to derive the

flux density of the phase calibrator, which was previously assumed to be a point source of 1 Jy located at the phase center. This is achieved using the *fluxscale* task. The final step of the calibration process is to apply the calibration solutions to the actual data, using the task *applycal*. During this process, the calibration solutions are applied to the DATA column in the measurement set, and the results are written in the CORRECTED_DATA column of the measurement set. For the calibrators, the phase and amplitude calibration comes from their own solutions and the bandpass solutions come from the bandpass calibrator. For the science target we again apply the bandpass solution from the bandpass calibrator but the gain solutions come from the close by phase calibrator.
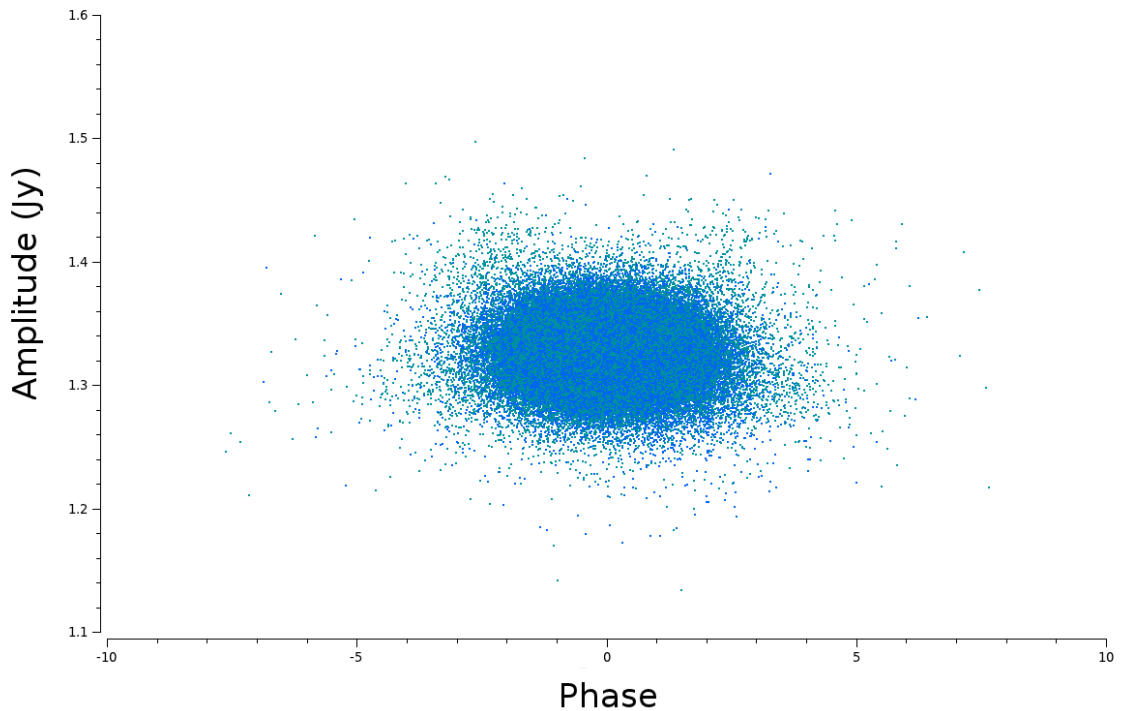

Figure 1.7: A well calibrated source will produce a compact ball of visibilities centered at zero phase and at the amplitude found for that source. Here we plot the calibrated visibilities at 3 cm for J0449+1121; the phase calibrator in the 3 cm data set for Aldebaran.

Once calibration of the data is complete, it is worth spending time inspecting the corrected data to make sure there are no obvious errors in the data. If such

errors are indeed found at this stage, then their cause will need to be flagged and the data must be re-calibrated. Some insightful plots to investigate how successful the calibration was, are: amplitude vs time, amplitude vs $u - v$ distance, and amplitude vs phase. If a point source has being successfully calibrated then it will have an appearance similar to that in Figure 1.7, i.e., a compact ball of visibilities centered at zero phase and at the amplitude found for that source. Once the data has been successfully calibrated, the science data is split off into a separate measurement set using the *split* task. The calibrated visibilities can then be directly analyzed by fitting simple models (e.g. point sources, disks, etc.,) to them or, as is more common, Fourier transformed and CLEANed to create an image of the source.

1.3 Imaging

We have shown in Chapter **??** that the visibility as a function of baseline co-ordinates (u, v) is the Fourier transform of the sky brightness distribution as a function of the sky coordinates (l, m), i.e.,

$$I(l, m) = \int \int V(u, v) e^{2\pi i (ul + vm)} du dv. \tag{1.4}$$

Fourier transforming the calibrated visibilities results in a dirty image which can then be deconvolved to produce a good estimate of the true sky brightness distribution. CASA has a single task called *clean* which carries out both of these operations on the data. in the following two sections we describe how both the VLA and the CRAMA visibility data sets were imaged using this task.

1.3.1 Imaging the VLA Data

The calibrated visibilities were both Fourier transformed and deconvolved using the CASA *clean* task in multi-frequency synthesis imaging mode. This imaging mode separately grids the multiple spectral channels onto the u-v plane and therefore improves the overall u-v coverage. As all science targets were expected to be point sources at all wavelengths, resolution was not paramount and so we used natural weighting for maximum sensitivity. The cell size was chosen so that

the synthesized beam was about five pixels across. The FOV of the VLA at short wavelengths is small (see Table **??**) and at these wavelengths there are less serendipitous background objects. This meant that the science targets were the only objects within a few primary beams of phase center and so it was sufficient to place just one CLEAN circle around the target source. The general procedure was to first use *clean* to create a dirty image (i.e., by setting $niter = 0$) allowing the root mean square noise of this dirty image, σ_{rms}, to be determined. The final CLEAN image was then created by setting *niter* to a very large number and setting the CLEANing threshold to be $\sim 3\sigma_{\mathrm{rms}}$. Setting *niter* to a very large number ensures that this CLEANing threshold is reached.

At long VLA wavelengths the primary beam becomes larger and the background sky sources become brighter. The new wide bandwidth capabilities of the VLA means that it is quite sensitive even far from phase center. For example, at $\sim 20\,\mathrm{cm}$ (L-band), the HPBW of the primary beam is $\sim 30'$ and yet the primary beam gain is as much as 10% around 1° away. In Figure 1.9 we show a wide field view of our VLA 20 cm image clearly showing many serendipitous sources out to and beyond the HPBW of the primary beam, all of which needed to be individually CLEANed to reduce their sidelobe contamination of the final image. For this reason the image sizes were usually set to a few times the size of the primary beam (if not too computationally expensive) so that any nearby strong serendipitous sources could be CLEANed. These images were again CLEANed interactively, and also took sky curvature into account. CLEANing was carried out interactively with CLEAN circles first placed around the strongest sources in the image and then placed around weaker sources as they appeared in the residual image.
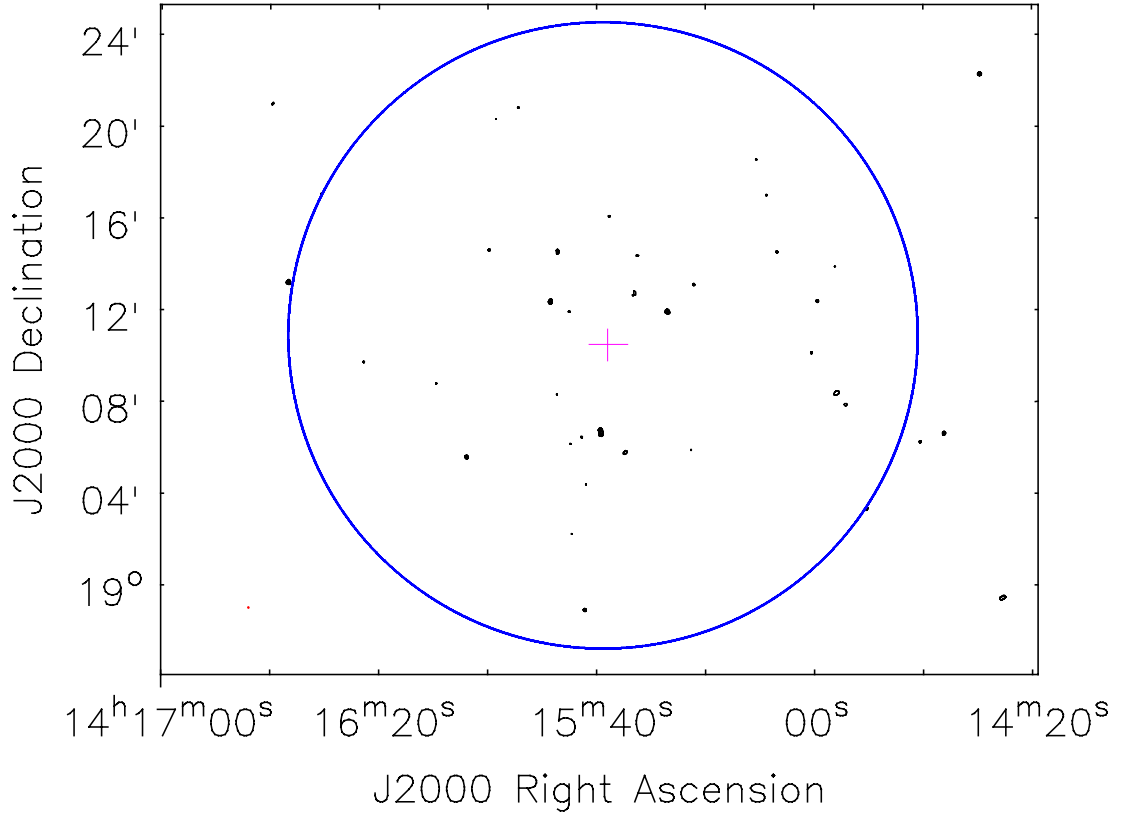
1.3.2   Imaging the CARMA Data

Figure 1.8: Wide field view of the VLA 20 cm image showing the many serendipitous background sources close to Arcturus. CLEAN circles were placed around all these sources during interactive CLEANing. The blue circle marks the FOV (i.e., the HPBW of the primary beam) of the VLA at 20 cm while the pink cross in the center marks the expected position of Arcturus. The contour levels are set at $(0.005, 0.01, 0.05, 0.1, 0.2, 0.4, 0.6, 0.8) \times 80.3 \, \text{mJy}$, where $80.3 \, \text{mJy}$ is the flux of the brightest source in the image.
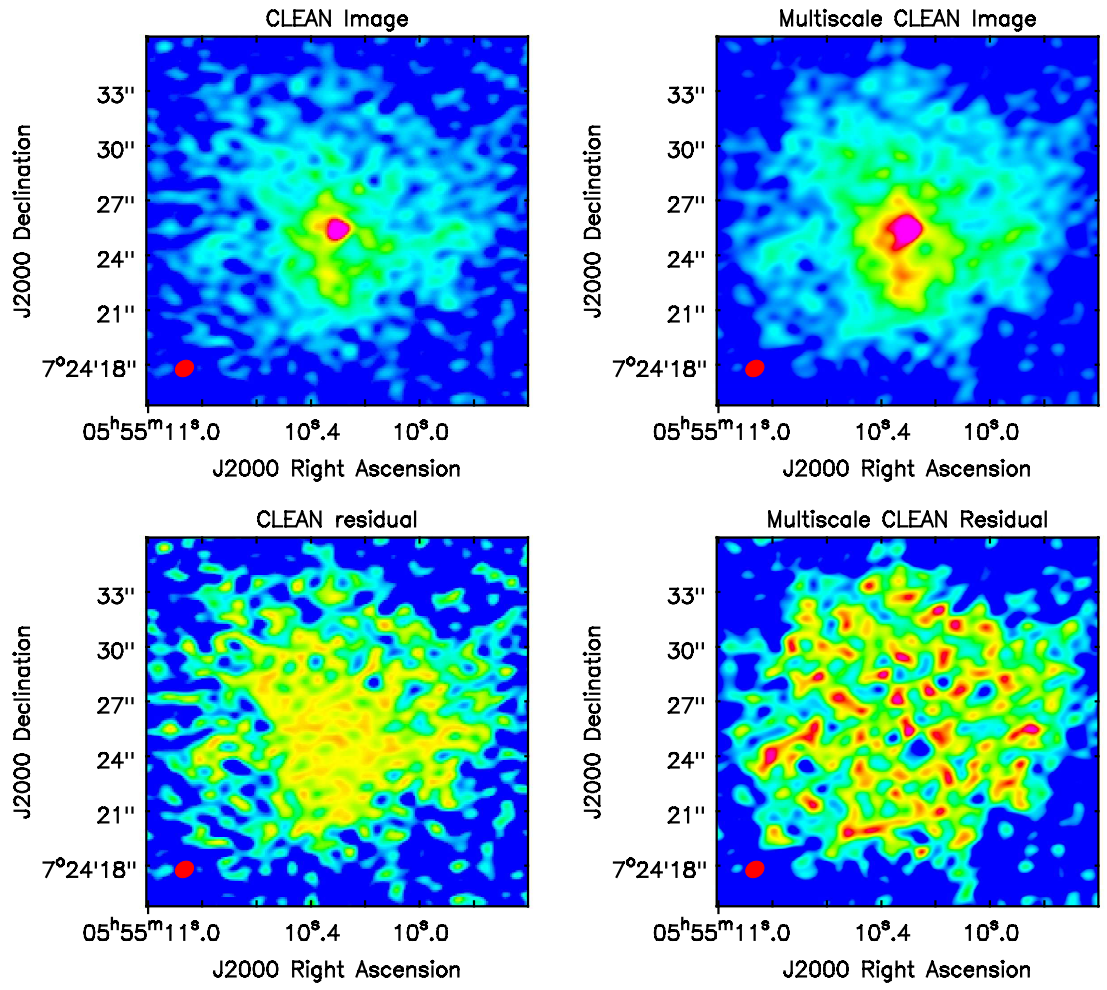
Figure 1.9

A

List of Abbreviations

Table A.1: List of Abbreviations

Abbreviation	Meaning
BIMA	Berkeley Illinois Maryland Association
CARMA	Combined Array for Research in Millimeter-wave Astronomy
CSE	Circumstellar Envelope
DDT	Director's Discretionary Time
e-MERLIN	e-Multi-Element Radio Linked Interferometer Network
FOV	Field of View
GREAT	German Receiver for Astronomy at Terahertz Frequencies
HPBW	Half Power Beamwidth
HST	Hubble Space Telescope
IOTA	Infrared Optical Telescope Array
IR	Infrared
IRAM	Institut de Radioastronomie Millimétrique
IUE	International Ultraviolet Explorer
LSR	Local Standard of Rest
OVRO	Owens Valley Radio Observatory
RFI	Radio Frequency Interference
S/N	signal-to-noise
SOFIA	Stratospheric Observatory for Infrared Astronomy
SMA	Submillimeter Array
UV	Ultraviolet
VLA	Karl G. Jansky Very Large Array
VLBA	Very Long Baseline Array
VLT	Very Large Telescope

References

HAMAKER, J.P., BREGMAN, J.D. & SAULT, R.J. (1996). Understanding radio polarimetry. I. Mathematical foundations. *Astronomy & Astrophysics Supplemental*, **117**, 137–147. (Cited on page 7.)

McMULLIN, J.P., WATERS, B., SCHIEBEL, D., YOUNG, W. & GOLAP, K. (2007). CASA Architecture and Applications. In R.A. Shaw, F. Hill & D.J. Bell, eds., *Astronomical Data Analysis Software and Systems XVI*, vol. 376 of *Astronomical Society of the Pacific Conference Series*, 127. (Cited on page 2.)

O'GORMAN, E., HARPER, G.M., BROWN, J.M., BROWN, A., REDFIELD, S., RICHTER, M.J. & REQUENA-TORRES, M.A. (2012). CARMA CO(J = 2 - 1) Observations of the Circumstellar Envelope of Betelgeuse. *Astronomical Journal*, **144**, 36. (Cited on page 1.)

O'GORMAN, E., HARPER, G.M., BROWN, A., DRAKE, S. & RICHARDS, A.M.S. (2013). Multi-wavelength Radio Continuum Emission Studies of Dust-free Red Giants. *Astronomical Journal*, **144**, 36. (Cited on page 1.)

PERLEY, R.A. & BUTLER, B.J. (2013). An Accurate Flux Density Scale from 1 to 50 GHz. *Astrophysical Journal Supplemental Series*, **204**, 19. (Cited on page 9.)

SAULT, R.J., TEUBEN, P.J. & WRIGHT, M.C.H. (1995). A Retrospective View of MIRIAD. In R.A. Shaw, H.E. Payne & J.J.E. Hayes, eds., *Astronomical Data Analysis Software and Systems IV*, vol. 77 of *Astronomical Society of the Pacific Conference Series*, 433. (Cited on page 2.)

WHITE, S.M. (2009). Single Dish Aperture Effciency Measurements at CARMA. In *CARMA Memorandum Series*, vol. 49 of *CARMA Memorandum Series*. (Cited on page 8.)