

THÈSE DE DOCTORAT

de l'Université de recherche Paris Sciences et Lettres

Préparée à l'Observatoire de Paris & UCKAR

Statistical Analysis of the Radio-Interferometric Measurement Equation,
a derived adaptive weighting scheme, and applications to LOFAR-VLBI
observation of the Extended Groth Strip

École doctorale n°127

ASTRONOMIE ET ASTROPHYSIQUE D'ILE-DE-FRANCE

Spécialité INTERFÉROMÉTRIE RADIO & ALGORITHMIQUE

COMPOSITION DU JURY :



Soutenue par **ETIENNE Bonnassieux**
le 20 Septembre 2018

Dirigée par **Oleg Smirnov,**
Cyril Tasse et **Philippe Zarka**

Prof Pierre Kervella, Président
LESLIA

Prof Neal Jackson, Rapporteur
University of Manchester

Prof Patrick Charlot, Rapporteur
Laboratoire d'Astrophysique de Bordeaux

Dr Maaike Mevius
ASTRON, Membre du jury

Dr Chiara Ferrari, Membre du jury
Observatoire de la Côte d'Azur

Dr Cyril Tasse, Directeur de Thèse
GEPI

Prof Oleg Smirnov, Directeur de Thèse
SKA-SA

Dr Philippe Zarka, Directeur de Thèse
LESLIA

Contents

	Page
1 Thesis Overview: Astrophysics & Techniques	5
1.1 Astrophysical Interest	5
1.2 The Low- ν Sky: Emission Mechanisms	8
1.2.1 Synchrotron Radiation	9
1.2.2 Free-Free Radiation	11
1.3 LOFAR: The Low-Frequency Array	13
1.4 This Thesis	16
1.4.1 Algorithmic Developments	17
1.4.2 Imaging the Full Primary Beam	17
1.4.3 3C295 and the Extended Groth Strip	19
2 Techniques of Interferometry	21
2.1 A Brief Introduction to Radio Astronomy	23
2.2 Interferometry: Bypassing the Diffraction Limit	23
2.2.1 The Baseline	25
2.2.2 The Visibility	25
2.2.3 The uv -plane	27
2.2.4 The Point-Spread Function	27
2.2.5 From Dirty to Clean: Deconvolving the PSF	30
2.3 Unifying Calibration & Imaging	32
2.4 Weighting Schemes	35
2.5 An Observer's Perspective of Radio Interferometry	37
2.6 Calibration Methods in Radio Interferometry	39
2.6.1 Generational Analysis	39
2.7 The RIME Formalism	44
2.8 Setting up the RIME: a single point source	44
2.9 Expanding our Jones Chain	45
2.9.1 Geometric Effects: the Phase Delay Jones Matrix	45
2.9.2 The Coherency Matrix	46
2.10 Multiple Point Sources	47
2.11 Diffuse Emission: the Full-sky RIME	47
2.11.1 Deriving the Fully-Sky RIME	47
2.11.2 Recovering the van Cittert-Zernike Theorem	48
2.12 Time-dependent Jones matrices, Smearing, and Decorrelation	49
2.12.1 Time-dependence of Jones matrices	49
2.12.2 Smearing and Decorrelation	50

2.13	Noise	51
3	Analysing the Variance of Gain Solutions	53
3.1	Introduction	53
3.2	Building the Noise Map	56
3.2.1	The Cov-Cov Relationship in the $\delta u \delta v$ plane	56
3.2.2	Statistical Analysis	59
3.2.3	Noise Map Simulations	62
3.3	Adaptive Quality-based Weighting Schemes	67
3.3.1	Optimising sensitivity	67
3.3.2	Minimising Calibration Artefacts	68
3.4	Estimating the Covariance Matrix	68
3.4.1	Baseline-based Estimation	69
3.4.2	Antenna-based Estimation	70
3.5	Applying the Correction to Simulated Data	72
3.6	Applying the Correction to Real Data	73
3.7	Discussion	76
4	Imaging the full LOFAR field of view	79
4.1	Aims & Methodology	79
4.2	Data Reduction	80
4.2.1	Data & Observation Properties	80
4.2.2	Reducing the Data	80
4.3	Effect of Quality-based Weighting Scheme	84
4.4	Zoo of Radio Sources	88
4.4.1	Selected Sources	88
4.4.2	Images & Cross-matching	89
4.5	Discussion	103
5	Imaging 3C295 with International LOFAR	105
5.1	Aims & Methodology	105
5.2	Data Reduction	106
5.2.1	Our RIME	106
5.2.2	Calibration Strategy	107
5.2.3	Testing the impact of initial model choice	108
5.2.4	Frequency-dependent Scale Error	110
5.2.5	Examining gain solutions	112
5.3	Overlays & Spectral Analaysis	124
5.4	Discussion	127
5.5	Testing Directional Gain Errors - Aims & Methodology	129
5.6	Testing Directional Gain Errors - LBCS Sources in the Primary Beam	132
6	Conclusion & Future Work	141
6.1	Algorithmic Work	142
6.1.1	The Cov-Cov Relationship	142
6.1.2	Mapping the variance of images made with interferometric data	143
6.1.3	Quality-based Weighting Schemes	144
6.2	Imaging the Extended Groth strip	145

CONTENTS

1

6.2.1	Imaging the LOFAR primary beam	146
6.2.2	Imaging the EGS with LOFAR-VLBI	146
6.2.3	Future Work	149
Bibliography		154

Foreword

J.R.R Tolkien wrote, in his *Mythopoeia*, that “He sees no stars who does not see them first, of living silver made that sudden burst, to flame like flowers beneath the ancient song”¹. In his defense of myth-making, he formulates the argument that the attribution of meaning is an act of creation - that “trees are not ‘trees’ until so named and seen” - and that this capacity for creation defines the human creature. The scientific endeavour, in this context, can be understood as a social expression of a fundamental feature of humanity, and from this endeavour flows much understanding. This thesis, one thread among many, focuses on the study of astronomical objects as seen by the radio waves they emit.

What are radio waves? Electromagnetic waves were theorised by James Clerk Maxwell (Maxwell 1864) in his great theoretical contribution to modern physics, their speed matching the speed of light as measured by Ole Christensen Rømer and, later, James Bradley. It was not until Heinrich Rudolf Hertz’s 1887 experiment that these waves were measured in a laboratory, leading to the dawn of radio communications - and, later, radio astronomy. The link between radio waves and light was one of association: light is known to behave as a wave (Young double-slit experiment), with the same propagation speed as electromagnetic radiation. Light “proper” is also known to exist beyond the optical regime: Herschel’s experiment shows that when diffracted through a prism, sunlight warms even those parts of a desk which are not observed to be lit (first evidence of infrared light). The link between optical light and unseen electromagnetic radiation is then an easy step to make, and one confirmed through countless technological applications (e.g. optical fiber to name but one). And as soon as this link is established, a question immediately comes to the mind of the astronomer: what does the sky, our Universe, look like to the radio “eye”?

Radio astronomy has a short but storied history: from Karl Jansky’s serendipitous observation of the centre of the Milky Way, which outshines our Sun in the radio regime, in 1933, to Grote Reber’s hand-built back-yard radio antenna in 1937, which successfully detected radio emission from the Milky Way itself², to such monumental projects as the Square Kilometer Array and its multiple pathfinders, it has led to countless discoveries and the opening of a truly new window on the Universe. The work presented in this thesis is a contribution to this discipline - the culmination of three years of study, which is a rather short time to get a firm grasp of radio interferometry both in theory and in practice. The need for robust, automated methods - which are improving daily, thanks

¹See here for the full text.

²https://www.nrao.edu/whatisra/hist_reber.shtml

to the tireless labour of the scientists in the field - is becoming ever stronger as the SKA approaches, looming large on the horizon; but even today, in the precursor era of LOFAR, MeerKAT and other pathfinders, it is keenly felt. When I started my doctorate, the sheer scale of the task at hand felt overwhelming - to actually be able to contribute to its resolution seemed daunting indeed!

Thankfully, as the saying goes, no society sets for itself material goals which it cannot achieve. This thesis took place at an exciting time for radio interferometry: at the start of my doctorate, the LOFAR international stations were - to my knowledge - only beginning to be used, and even then, only tentatively; MeerKAT had not yet shown its first light; the techniques used throughout my work were still being developed. At the time of writing, great strides have been made. One of the greatest technical challenges of LOFAR - imaging using the international stations - is starting to become reality. This technical challenge is the key problem that this thesis set out to address. While we only achieved partial success so far, it is a testament to the difficulty of the task that it is not yet truly resolved.

One of the major results of this thesis is a model of a bright resolved source near a famous extragalactic field: properly modeling this source not only allows the use of international LOFAR stations, but also grants deeper access to the extragalactic field itself, which is otherwise polluted by the 3C source's sidelobes. This result was only achieved thanks to the other major result of this thesis: the development of a theoretical framework with which to better understand the effect of calibration errors on images made from interferometric data, and an algorithm to strongly mitigate them.

The structure of this manuscript is as follows: we begin with an introduction to radio interferometry, LOFAR, and the emission mechanisms which dominate for our field of interest. These introductions are primarily intended to give a brief overview of the technical aspects of the data reduced in this thesis. We follow with an overview of the Measurement Equation formalism, which underpins our theoretical work. This is the keystone of this thesis.

We then show the theoretical work that was developed as part of the research work done during the doctorate - which was published in *Astronomy & Astrophysics*. Its practical application - a quality-based weighting scheme - is used throughout our data reduction. This data reduction is the next topic of this thesis: we contextualise the scientific interest of the data we reduce, and explain both the methods and the results we achieve.

Chapter 1

Thesis Overview: Astrophysics & Techniques

Contents

1.1	Astrophysical Interest	5
1.2	The Low- ν Sky: Emission Mechanisms	8
1.2.1	Synchrotron Radiation	9
1.2.2	Free-Free Radiation	11
1.3	LOFAR: The LOW-Frequency Array	13
1.4	This Thesis	16
1.4.1	Algorithmic Developments	17
1.4.2	Imaging the Full Primary Beam	17
1.4.3	3C295 and the Extended Groth Strip	19

1.1 Astrophysical Interest

Radio astronomy is a young but extremely dynamic field with a large range of science cases. From the study of the epoch of reionisation, which involves looking for polarised, redshifted 21cm line emission to study the emergence of the first luminous sources¹, to that of transient radio sources such as pulsars and exoplanets, going through the study of ultra-high-energy cosmic rays, solar science, space weather science, the study of cosmic magnetism, and the creation of deep extragalactic surveys, radio-astronomy opens up an extremely fertile window into our Universe.

In Fig. 1.1, for example, much structure is immediately visible: supernova remnants in the galactic plane (the “bubbles”), gas following galactic magnetic field lines (the filamentary structure), and emission from diffuse gas structure around the core of the Milky Way. Other sources are also visible, however: some unresolved sources, and very

¹See <https://www.haystack.mit.edu/ast/science/epoch/> for more information - the ultimate goal of this topic is to learn more about the process of structure formation in the universe: what exactly are the evolutionary links between the very smooth matter distribution revealed by CMB studies and the highly structured universe of galaxies and clusters thereof which we know today?

many AGN sources. This is one particularly striking example of what radio telescopes allow us to see of astrophysical sources, but is only one among many.

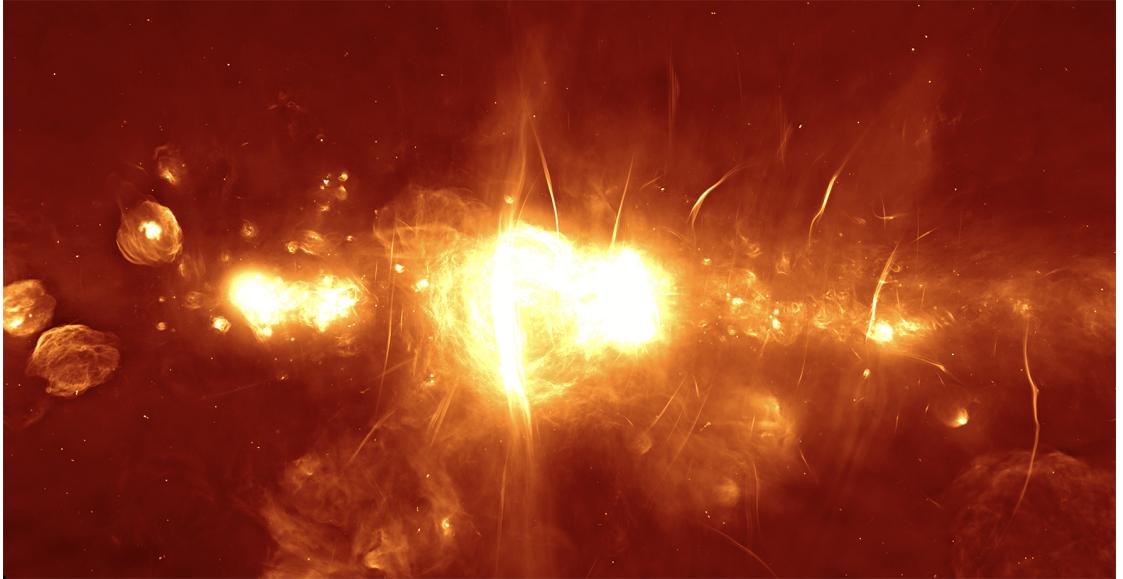


Figure 1.1. *MeerKAT image of the galactic centre, made by SARA0. This is a low-resolution version of the true image.*

The work presented here falls under the auspices of the deep extragalactic surveys science project. The underlying scientific motivation behind this Key Science Project or KSP is the study of, among other things, radio-active galactic nuclei, or radio AGNs. A radio AGN is a type of radio galaxy whose emission is dominated by its nucleus. This nucleus emission is non-thermal, and assumed to be emitted from both collimated jets springing from the accretion of matter onto a galaxy's supermassive black hole and from a hot accretion disk surrounding said black hole. AGNs are then commonly split into different types, based on spectral line behaviour (as in Seyfert 1943) or morphological characteristics (as in Fanaroff & Riley 1974).

So-called unified AGN models, which aim to explain the underlying physical object which is assumed to lie behind all observed examples of AGN, were developed to explain the diversity in observed AGN types (e.g. Antonucci & Miller 1985; Tasse et al. 2008; Netzer 2015; Morabito et al. 2017). These models would include a central engine, usually a supermassive black hole with hot accretion (innermost accretion regions estimated to be around 10^7 K by Pringle & Rees 1972), emitting thermally; an optically thick disk torus, typically of size 10 to 100pc, blocking light emitted from the central engine along the equatorial plane, and the cavity inside which would be the region from which broad emission lines would emit; and finally, ionisation cones in the direction of the poles, where the emission from the central engine would collimate into large-scale jets (which would also, at the scale of a few tens of parsecs, create the region of narrow spectral line emission).

The interaction between AGN jets and the surrounding intergalactic medium can result in structure such as that shown in Fig. 1.2 below: a radio galaxy in the centre of two extremely radio-bright jets, with lots of diffuse emission from their lobes. These lobes show where the accretion-accelerated particles begin to slow down - and therefore emit - due to contact with the intergalactic medium. Hence, their sometimes peculiar shape: this medium can be complex and inhomogeneous (if the radio galaxies lie in a cluster, for example) which can lead to the sort of jet structure shown in this image.

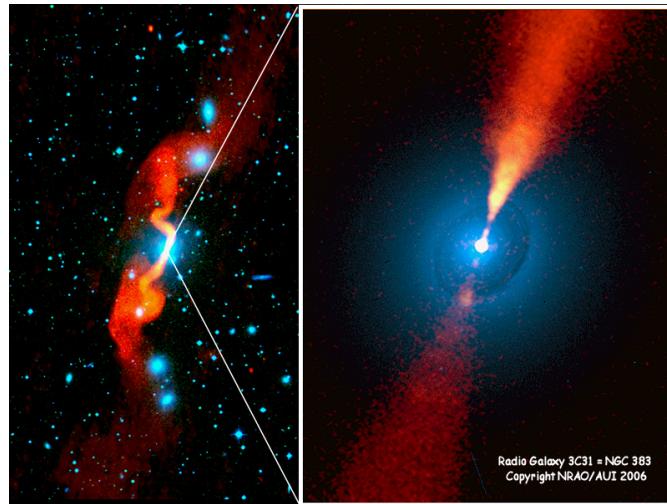


Figure 1.2. Radio optical montage of 3C31, by Alan Bridle:
<https://www.cv.nrao.edu/~abridle/3c31xopt.htm>.

Of course, not all galaxies have active nuclei, and those which do also have other forms of radio emission: for example, free-free radiation from the interstellar medium, ionised by light from newborn stars, will often contribute to radio-galactic flux. Some galaxies with active nuclei do not have strong jet emission: these are called radio-quiet AGNs. We can thus identify two components to a radio AGN: the host galaxy and its jets. Peak star formation is thought to have occurred around $z \sim 2$, and so the study of AGN activity around this period is of strong scientific interest. Could the jets from active AGNs have contributed to the decline in star formation? Are there large populations of older AGNs, and if so, how old? These are the astrophysical questions which motivate the study of AGN sources (Best 2007; Xu et al. 2018; Ueda et al. 2018), which is the driving scientific concern which motivates the present doctorate.

1.2 The Low- ν Sky: Emission Mechanisms

This section aims to bring a brief introduction to the emission mechanisms which dominate at low frequencies, and thus determine the physics accessible to extragalactic astronomers working in this band. Specifically, it will briefly describe free-free radiation (which is emitted from ionised plasma outflows, and thus a tracer of various physical objects e.g. young stellar objects - cf. Coughlan et al. (2017) and references therein - and starburst regions - cf. Varenius et al. (2015)) and synchrotron radiation, which is a tracer of more violent and energetic processes (and thus associated with supernova remnants, AGN, or radio halos - cf. Cassano et al. (2010)).

Of course, each individual galaxy will have differing contributions from different mechanisms - in practice, when studying galactic populations, the overall flux at a given frequency will simply be summed up for each galaxy and become one point in a Spectral Energy Distribution, or SED. However, when studying individual objects, it can be critical to understand which emission mechanism dominates at what frequencies. For example, the radio and far-infrared spectrum for nearby M82 are shown in Fig. 1.3².

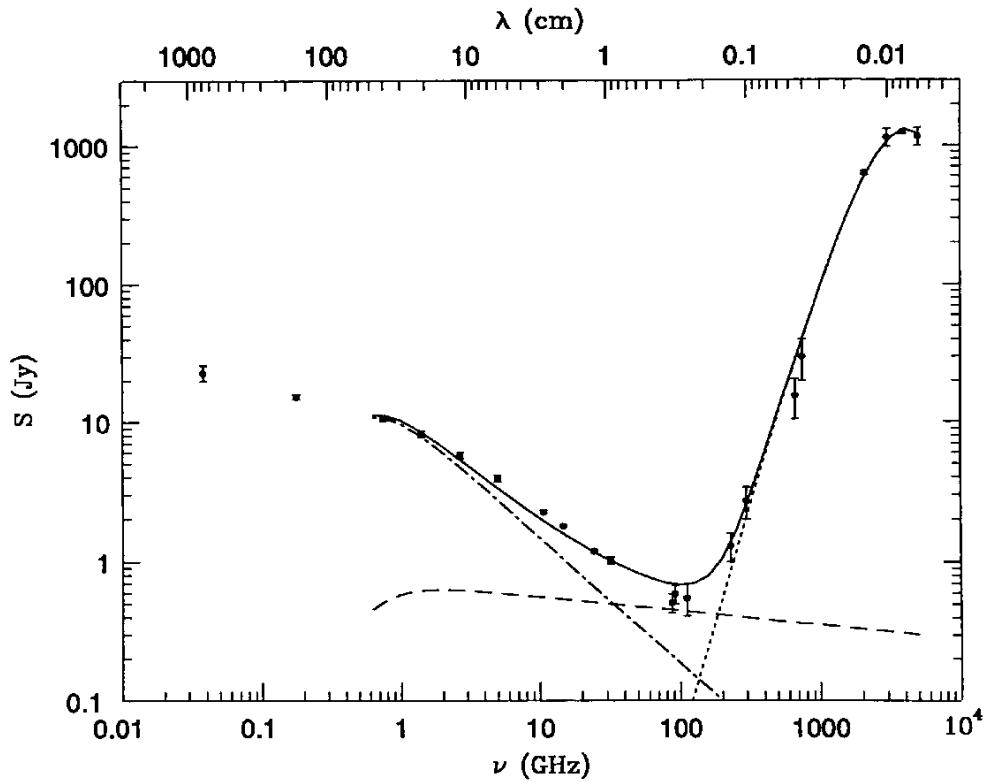


Figure 1.3. Radio and far-infrared spectrum for galaxy M82, as estimated by the NRAO online course (<https://www.cv.nrao.edu/course/astr534/FreeFreeEmission.html>). The flat curve corresponds to free-free emission, while synchrotron radiation and thermal dust emission dominate at low and high frequencies respectively.

²Figure and work taken from the NRAO website. For further information, see here

1.2.1 Synchrotron Radiation

Synchrotron radiation is the dominant mode of emission for extragalactic sources at low wavelengths. As such, it will be covered in greater detail than the other mechanisms above, which are important in other bands (and therefore relevant for multi-spectral analysis) but much less relevant to LOFAR observations. This section draws heavily from Garret Cotter's high-energy astrophysics lectures³, as well as Alan Loh's doctoral thesis (Loh 2016) and the NRAO online course⁴.

Synchrotron radiation (or “magnetobremsstrahlung”) occurs when the trajectory of a relativistic charged particle is guided by a magnetic field. As the German name suggests, it is the magnetic equivalent of free-free radiation. It is the tracer of very energetic processes, such as the interaction between AGN jets and the intergalactic medium. In the low-energy regime, it is also known as cyclotron radiation, after the device in which it was first measured in a laboratory.

The synchrotron power spectrum of a population of electrons with exactly the same energy is a function of its Lorentz factor and its emission angle⁵ (Longair 1994). The typical radiation spectrum for a single electron is shown in Fig. 1.4 as a function of the critical frequency, which is itself a function of the emission angle θ , the Lorentz factor γ and the magnetic field strength B : $\nu_c = 3/2 \frac{\gamma^2 e B}{2\pi m_e c} \sin(\theta)$. e and m_e are the electron charge and mass, respectively.

Of course, in reality, we know that electron populations in astrophysical sources do not all share exactly the same energy when emitting. The effect of the electron energy distribution must therefore be taken into account. Because the population in question is energised through non-thermal processes, its energy distribution is not Maxwellian, and is generally written as a power law: $S(E) \propto E^\alpha \rightarrow S_\nu \propto \nu^{-\alpha}$, where α is the object's spectral index. Note that the index of the energy distribution (which relates the number of electrons at a given energy) is different from the synchrotron index, with the two related by the following equation: $\alpha = -(\gamma - 1)/2$. The question now becomes: what α is appropriate for what kind of object? Does every source have the same α at all frequencies? Do all synchrotron sources have the same spectral index at a given frequency, and if not, what does this tell us?

To answer these questions, we must recall that so far, we have assumed that every photon emitted through synchrotron radiation reaches us. This is not necessarily the case in practice: it is possible for a given photon to be scattered (i.e. absorbed & re-emitted) by surrounding high-energy electrons. The likelihood of this occurring at a given frequency is a function of the absorption cross-section, and is a complex quantity to calculate (see Longair (1994) pp258-260). Suffice to say that at longer wavelengths (lower frequencies) absorption is more effective, and at shorter wavelengths (higher frequencies)

³<http://www-astro.physics.ox.ac.uk/~garret/teaching/>

⁴<https://www.cv.nrao.edu/course/astr534/SelfAbsorption.html>

⁵The emission angle is the angle between the electron's total velocity (including z-axis) and its radial velocity.

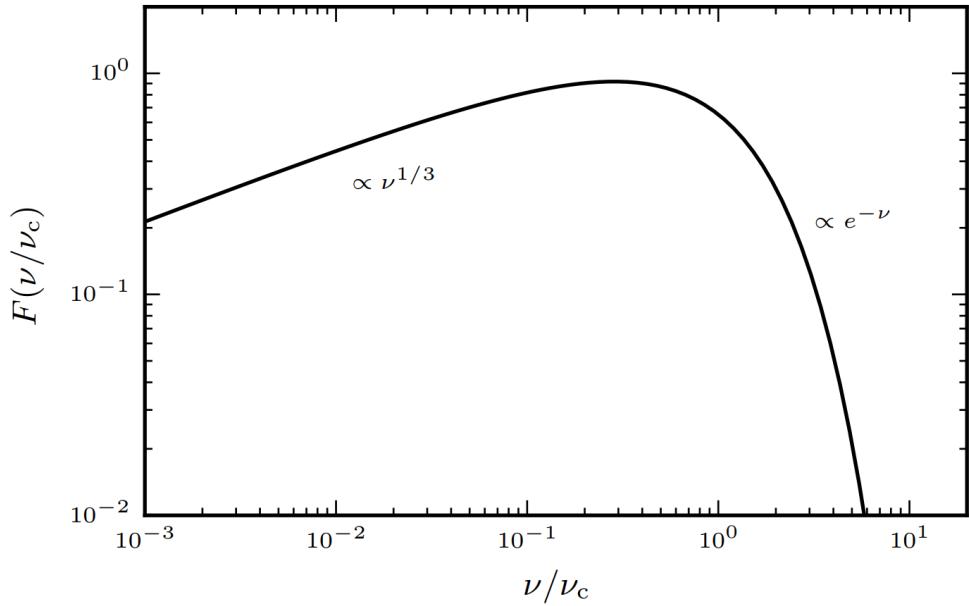


Figure 1.4. Characteristic spectrum of synchrotron emission for a population of single-energy electrons. Arbitrary unit scale. Source: NRAO online course:
<https://www.cv.nrao.edu/~sransom/web/Ch5.html>

the underlying electron energy power-law distribution begins to dominate. An example is shown in Fig. 1.5, taken from Tasse et al. (2007).

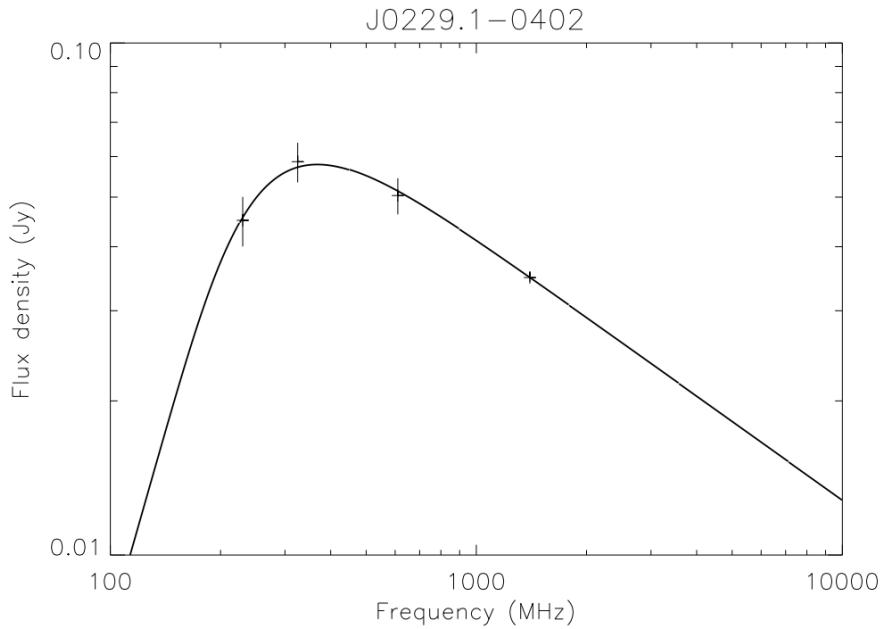


Figure 1.5. Power spectrum for J0229.1-0402. Source: Tasse et al. (2007).

In the optically-thick regime, we know that the most efficient radiation mechanism is simple black-body radiation, and so the power spectrum at a given frequency can be defined in terms of an “electron temperature”. In the Rayleigh-Jeans approximation (valid as we are explicitly in the low-frequency regime), this gives $S_\nu \propto \nu^{5/2}$, or $\alpha = 5/2$. In the optically-thin regime, meanwhile, we recover a mixture of the underlying power-law distribution ($S_\nu \propto \nu^{-\alpha}$) and black-body radiation (which has a frequency-dependent α). Empirically, we generally observe $\alpha \sim 0.7$.

1.2.2 Free-Free Radiation

Free-free or “bremsstrahlung” (“braking”) radiation occurs when the trajectory of a charged particle is deflected by an electric field. This non-thermal emission mechanism is the dominant mechanism in HII regions (which contain ionised hydrogen), where star formation has previously taken place. It is called free-free emission because it is produced by free electrons deviating off ionised hydrogen (i.e. protons) without being captured. When the underlying electron distribution is Maxwellian, it is referred to as thermal Bremsstrahlung.

This emission’s spectrum is heavily dependent on a number of factors: including frequency, temperature, and critically, free-free opacity τ_ν , itself a function of electron density. It is characterised by a knee in its spectrum, occurring where $\tau_\nu \sim 1$. This knee delineates two regions with different spectral indices; $\alpha \sim -0.1$ at higher frequencies, and $\alpha \leq 2$ at lower frequencies. This gives a characteristic shape, shown in Fig. 1.6.

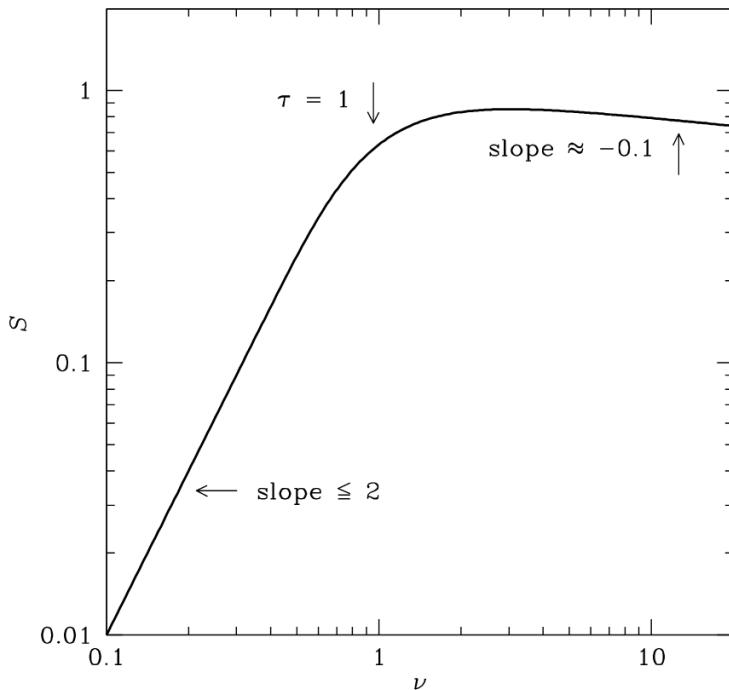


Figure 1.6. Characteristic spectrum of free-free radiation. Arbitrary unit scale. Source: NRAO online course:
<https://www.cv.nrao.edu/course/astr534/FreeFreeEmission.html>.

As we can see, this spectrum falls off sharply with decreasing frequency. As such, while it is expected to dominate over thermal emission in the absence of synchrotron radiation, it is synchrotron emission that is expected to dominate overall - if present - at the low frequencies of our observations, which are done with LOFAR. Free-free emission acts as a tracer for star formation and other "gentler" physical processes detected at low radio frequencies.

1.3 LOFAR: The LOw-Frequency Array

Our observations have all been made with the LOw Frequency Array LOFAR (van Haarlem et al. 2013). In this section, we describe its technical properties and its current state of the art. In particular, the distinction between “Dutch” LOFAR and “international” LOFAR - and the technical problems associated with each - will be made explicit in this section.

LOFAR is a SKA pathfinder instrument, which means that it serves not only as a cutting-edge instrument in its own right, but does so with the explicit aim of serving as testing grounds for technologies & techniques which could be usefully implemented in the SKA. It is in this context, for example, that trailblazers such as NenuFAR (Zarka et al. 2012), a low-frequency extension of LOFAR, are tested. LOFAR is an interferometric array, with each of its elements made up of antennas combined into a phased array, themselves forming an interferometer. It thus consists of antennas which are combined to form stations, which are themselves distributed throughout the Netherlands and Europe.

There are two bands to LOFAR, which are known as LOFAR-HBA (High-Band Antennas) and LOFAR LBA (Low-Band Antennas). Figs. 1.7a and 1.8 show the layout of the 6 innermost core stations and the French international LOFAR station, respectively.

We see the presence, in both cases, of two very different antenna types. One of these antenna types is not a single antenna, but rather a phased array: 16 antenna dipoles distributed in a 4×4 array. International LOFAR stations include 96 such tiles, distributed in a “filled-disk” pattern, where the disk fill ratio is determined by the size of the tile. Dutch “remote” stations have 48 such tiles spread into 2 disks (HBA_INNER and HBA_OUTER), while “core” station HBA tiles are split into two 24-tile crosses. This pattern is shown in Fig. 1.9a. The antennas from each tile are combined into a phased array with a single “tile beam”, and all tile beams are themselves combined as a phased array into a station beam⁶.

LBA antennas, use a very simple - and cost-effective - design. A typical LBA antenna is shown in Fig. 1.10. 96 such antennas are spread in a semi-random pattern in each LOFAR station⁷. In Dutch stations, 48 of these are used at all times. They are combined as a single phased array.

LOFAR-HBA is sensitive to higher frequencies, from 120 MHz to 240 MHz⁸. It has a smaller field of view than LOFAR-LBA, but a better angular resolution.

⁶Reference: ASTRON technical description: <https://www.astron.nl/radio-observatory/astronomers/technical-information/antennae/antennae-description>

⁷Reference: LOFAR.org website: <http://www.lofar.org/about-lofar/system/lofar-numbers/lofar-numbers>

⁸Reference: LOFAR.org website: <http://www.lofar.org/about-lofar/system/lofar-numbers/lofar-numbers>

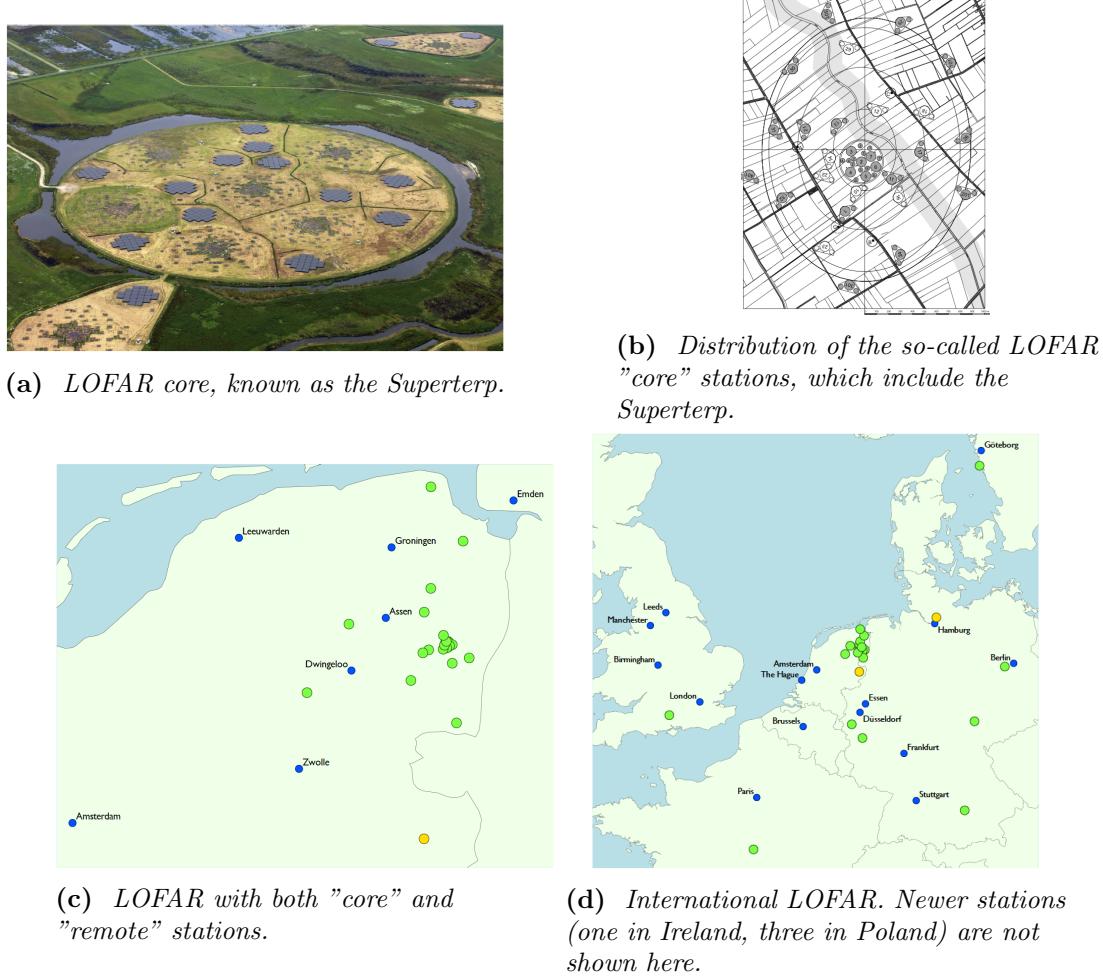
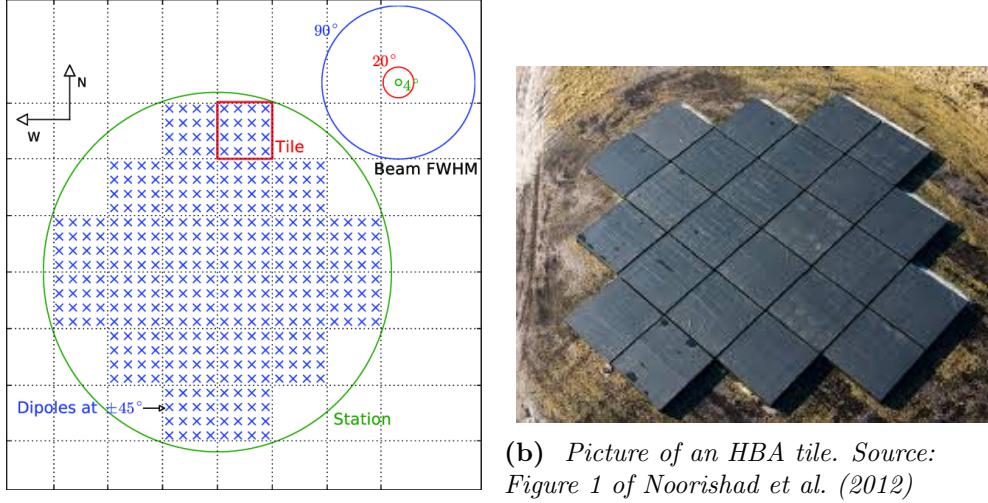


Figure 1.7. Geographic location and distribution of LOFAR stations, explicitly showing what is meant by Superterp, core, remote and international stations. All images from the official ASTRON website: <https://www.astron.nl/radio-observatory/astronomers/users/technical-information/lofar-array-configuration/lofar-array-conf>



Figure 1.8. Layout of FR606, the French LOFAR station at Nancay. At bottom left are the HBA tiles, bottom right the LBA dipoles, and at the top are some of the NenuFAR tiles.

The dipole design frees observers from the need to physically point antennas at all: the final station pointing is achieved by digitally introducing delays in the recorded phase



(a) Layout of a single HBA tile. Source: Sarod Yatawatta researchgate profile

Figure 1.9. HBA tile layout and image. Note that both cases show a 24-tile HBA array: each core station includes two such arrays. International stations include a single larger tile array.



Figure 1.10. Picture of a single LBA antenna. Source: LOFAR technology website.

before averaging the data. In this sense, the pointing is achieved in a similar way as individual HBA tile pointing, albeit . The LBA antennas are receptive to signals emitted in the 30-80 MHz frequency range.

At the time of writing, use of the LBA data is still relatively new, as its calibration is a very tricky problem. For similar reasons, international LOFAR (i.e. the full LOFAR array) has not been used, at the time of writing, to create wide-field survey images. A large part of the work described in this manuscript consists of reaching a point where full use can be made of international LOFAR, in a streamlined and repeatable way.

1.4 This Thesis

The work done in this thesis can be summarised as the development of a new algorithm for mitigating calibration artefacts in radio-interferometric images, and its application to a peculiar famous extragalactic field. Because our data volumes put radio interferometry in the domain of “Big Data” (there are \sim 245TB of LOFAR data centred on the EGS, for example), there is a very real need for fast, reliable methods with which to improve interferometric images. This algorithm follows from an improved understanding of the noise properties of interferometric images, and can be understood as a way to better constrain the flux distribution in images by minimising the presence of artefacts. This, in turn, helps self-calibration convergence by improving the image resulting from each imaging step of the self-calibration process.

We then outline the application of this work, along with other modern tools, to the creation of a large, deep and high-resolution image of a rich patch of sky. While this approach is potentially extremely rewarding in terms of the science one can do with its resulting maps, it is also uniquely complicated. There exist many LOFAR HBA (\sim 120-240 MHz) sky surveys (cf. Williams et al. 2016; Shimwell et al. 2017), but they do not yet make use of international LOFAR, and are thus limited in resolution ($\sim 5''$). And while there are surveys made at LOFAR frequencies (\sim 150MHz) with other instruments (e.g. Intema et al. 2017), they do not match the resolution attainable with international LOFAR. The aim of our image is then to show that the international LOFAR stations can be used to create a radio image that matches HST/ACS resolution - which would allow for significant cross-matching and source analysis, since there is HST data available for the field being imaged, the Extended Groth Strip (EGS). This would not be a true survey, but more of a proof of concept: the application of new tools and the development of an appropriate strategy to image a large patch of the sky with the full LOFAR array. This image would have better resolution than FIRST (see Helfand et al. 2015, and references therein), which has a $4''$ resolution, relatively low sensitivity, and no short baselines (thereby missing diffuse emission and extended structure) at \sim 1.4 GHz, while still being sensitive to diffuse emission due to the use of the LOFAR core stations.

Making this image is possible only because LOFAR observations of the Extended Groth Strip includes an extremely bright source, 3C295, in their field of view. This is both a blessing and a curse: unless this source is properly modeled and subtracted, it will pollute images of the EGS. However, because it is so bright, it can serve as a good calibrator source for the LOFAR international stations - if a good enough model of 3C295 can be made and used to calibrate these stations, then high-resolution images of the EGS becomes attainable. This would give access to a large population of radio AGNs at high resolution, observed at a large range of frequencies, and many of which lie at redshifts near the peak of star formation; this improvement of our low-frequency understanding of these objects could, in turn, lead to an improved understanding of these objects more generally.

To really be able to perform wide-field images using the full LOFAR array, we cannot rely on the happy circumstances we find ourselves in for our applications project: a much more thorough and reliable approach is necessary. The work of the LOFAR long-

baseline team (e.g. Jackson et al. 2016) shows that kind of approach: in our project, we test specific strategies and approaches which are complimentary to their work, but cannot be applied more generally without their results.

1.4.1 Algorithmic Developments

A significant portion of the work done as part of this thesis was on investigating the relationship between calibration gain errors and calibration artefacts in the image-plane. What started as an attempt to properly write a weighting scheme to down-weigh poorly-calibrated visibilities⁹ eventually became a formal relationship between these quantities, showing both the behaviour of image-plane variance as a function of direction but also of pixel covariance over the full image. This work is shown in much greater detail in Chapter 3.

The key finding of our algorithmic work is a fundamental relationship between the covariance matrices of residual visibilities and the covariance of pixels in the image-plane: the “Cov-Cov relationship” between visibility covariance and image-plane covariance. The pixel variance and covariance in the image-plane are determined by a “noise-PSF”, convolved with every pixel in the sky. This also means that thermal noise in interferometric images is correlated between pixels, and that this correlation is given exactly by the PSF: in the limit of an empty, non-deconvolved sky, we therefore properly model the thermal noise properties of its interferometric images. This noise-PSF is the product of the Fourier transforms of the calibration gain covariance matrix with each cell mapped not from uv space to lm coordinates but rather between their respective covariance spaces - from a new differential Fourier plane (henceforth “ $(\delta u \delta v)$ -plane) to the image-plane covariance space $\delta l \delta m$. This image-plane covariance space describes the variance in each pixel and the covariance between pixels¹⁰. It describes the expected calibration artefacts and noise level around each source, does not vary as a function of direction, and is convolved with each source in the field to yield the final error map. Because all unwanted (in our case, unphysical) signal can be thought of as noise, we will refer to the pixel variance map as the “noise-map”.

Based on this theoretical finding, we develop a weighting scheme algorithm with which to improve the noise properties of interferometric images by down-weighting either visibilities with high variance in their calibration residuals, or visibilities which are strongly correlated to others in their calibration residuals. We implement a robust version of the first weighting scheme, which relies on an antenna-based estimate of the visibility variance.

1.4.2 Imaging the Full Primary Beam

In this section, we briefly describe previous observations of the Extended Groth Strip (EGS), an extragalactic field with a rich multi-wavelength coverage described in Table

⁹i.e. to justify weighting visibilities by a function inversely proportional to the variance of the associated residual visibilities.

¹⁰The noise-PSF also relates δw to δn , as shown in our matrix formalism, but we do not explicitly reference this in the text for the sake of brevity, as visibility space is usually referred to as “the UV-plane” in literature, rather than “the UVW-space”.

1.1. It has been long observed as part of the All-Wavelength Extended Groth Strip International Survey collaboration (AEGIS, Davis et al. (2007)), which later became part of the CANDELS collaboration (Grogin et al. (2011)). The field is centred at $\alpha = 14^h 17^m$, $\delta = +52^\circ 30'$, placing it between the tail of Ursa Major and Draco. Its size is $0.7' \times 0.1'$. It has notably been the subject of very deep Hubble Space Telescope observations, which the international LOFAR could attempt to match.

Telescope	Band	Resolution	Area
Chandra	X-ray	$0.5'' - 6.0''$	$17' \times 120'$
GALEX	UV	$5.5''$	1.25° diameter
HST/ACS	Optical	$0.1''$	$10' \times 67'$
HST/NICMOS	Optical	$0.35''$	0.0128 deg^2
Megacam	Optical	$1.0''$	1 deg^2
IRAC	IR	$2.0''$	$10' \times 120'$
Spitzer	IR	$5.9'' - 19''$	$10' \times 90'$
VLA	Radio	$1.2'' - 4.2''$	$30' \times 80'$

Table 1.1. Table recapitulating the multi-wavelength coverage of the Extended Groth Strip, with observations performed as part of the AEGIS (later, CANDELS) collaborations.

Our data consist of an observation centred on the EGS. The LOFAR beam (i.e. how much of the sky LOFAR is sensitive to at a given time) is much larger than the EGS. As such, our measured visibilities includes contributions not just of sources in the EGS, but also from those all around it - including a very bright 3C¹¹ source, which will be discussed at length below. Due to technical limitations (i.e. available memory), it is not feasible to image the full LOFAR primary beam at once at international LOFAR's $0.1''$ resolution. As such, we begin by imaging at a lower resolution, and thus do not use the international stations at first. This allows us to create a low-resolution model of all the sources in the primary beam before proceeding to international station calibration, giving us a sanity check for core and remote station gains. It also allows us to subtract out-of-field sources when imaging the EGS at high resolution.

Initial direction-independent calibration is done using Prefactor, the same pipeline as for the Lofar Two-metre Sky Survey, followed by direction-dependent calibration with the killMS-DDF facet-based pipeline (LOTSS, Shimwell et al. 2017).

This work, and its results, are shown in Chapter 4. Note that the AEGIS observations summarised in Table 1.1 do not cover the full LOFAR primary beam. The sources observed at low resolution are therefore matched with Sloan Digital Sky Survey (SDSS, York et al. (2000)) images. Overlays are shown to highlight the wealth of interesting sources, and their optical counterparts, which can be picked up as a byproduct of high-resolution maps of the radio sky. We select 12 particularly interesting sources and show the results of source matching for these candidates. Overlays of the same SDSS images with NVSS (NRAO VLA Sky Survey, Condon et al. (1998)) postage stamps are also shown for comparison. Many of the chosen sources do not lie within the Extended

¹¹Third Cambridge Catalogue of Radio Sources, a 1959 survey of the Northern radio sky

Groth Strip, and are therefore only imaged at low resolution, but high-resolution follow-up images and observations could be made should the need arise.

1.4.3 3C295 and the Extended Groth Strip

The brightest source in the primary beam of our observation is 3C295. It is one of the brighter sources in the Northern radio sky. It lies less than a degree away from the centre of the Extended Groth Strip, which has two important consequences: first, that it can potentially be used as a calibrator source for LOFAR; second, that unless well-modeled and subtracted at the observing instrument's maximum resolution, its sidelobes will pollute the EGS such that very little scientifically useful information might be recovered from a given observation. If we wish to image the EGS with international LOFAR, we therefore cannot rely on a low-resolution model: small errors in the exact location of 3C295's flux could corrupt most of the very long-baseline visibilities, thus making them useless. A high-resolution model of 3C295 is necessary for us to proceed.

The first step towards acquiring a deep, high-resolution image of the Extended Groth Strip therefore of creating a good, high-resolution model of 3C295 at LOFAR frequencies and for long baselines. Creating this model is one of the key results of this thesis. Starting from a high-resolution VLA model of 3C295, subbands selected across the LOFAR bandwidth are self-calibrated in iteratively greater numbers until a given noise threshold is reached. The flux scale is then boot-strapped based on the results of Scaife & Heald (2012). This ensures that 3C295 will be adequately subtracted from all corrected visibilities - including international baselines - and thus that images made using these visibilities will not be polluted by sidelobes from 3C295. The work done towards this goal is shown in Chapter 5.

Before launching the EGS imaging run using all calibrated datasets, we measure the impact of direction-dependent effects on known sources in the field. Two direction-dependent effects are expected to dominate: decorrelation and differential gains. The first is modeled by the algorithms used in this thesis, whereas the second is a function of the gain variability as a function of distance from the calibrator. Eight calibrator sources, lying at various distances and directions from the calibrator source, are examined. They are summarised in Table 5.2, which is reproduced below: note that number 6 is actually 3C295.

If directional gains are measured to have little or no impact, then direction-dependent calibration of the international stations will not be required. This is the ideal scenario, as it does not require the use of fringe fitting (Pearson & Readhead 1984). If they are measured to have a significant impact, then direction-dependent calibration will be required for international LOFAR. We expect the angular scale of differential gain evolution to be of the order of a degree, and so most of the EGS should be relatively unaffected.

#	RA [hms]	Dec [dms]	Dist. from EGS [deg.]	Dist. from 3C295 [deg.]
1	14:30:18.72	52:17:29.80	2.041	2.904
2	14:19:44.44	54:23:04.58	1.928	2.517
3	14:21:20.05	53:03:46.00	0.864	1.743
4	14:21:09.41	51:22:32.46	1.294	1.728
5	14:11:50.32	52:49:02.66	0.844	0.619
6	14:11:20.23	52:12:04.30	0.915	0.000
7	14:08:07.00	52:55:11.36	1.409	0.869
8	14:08:09.76	52:44:46.56	1.354	0.680

Table 1.2. Table recapitulating the positions of all 8 chosen calibrator sources, along with their distance from the observation phase centre (EGS phase centre) and calibrator phase centre (3C295), respectively. Number 6 is 3C295.

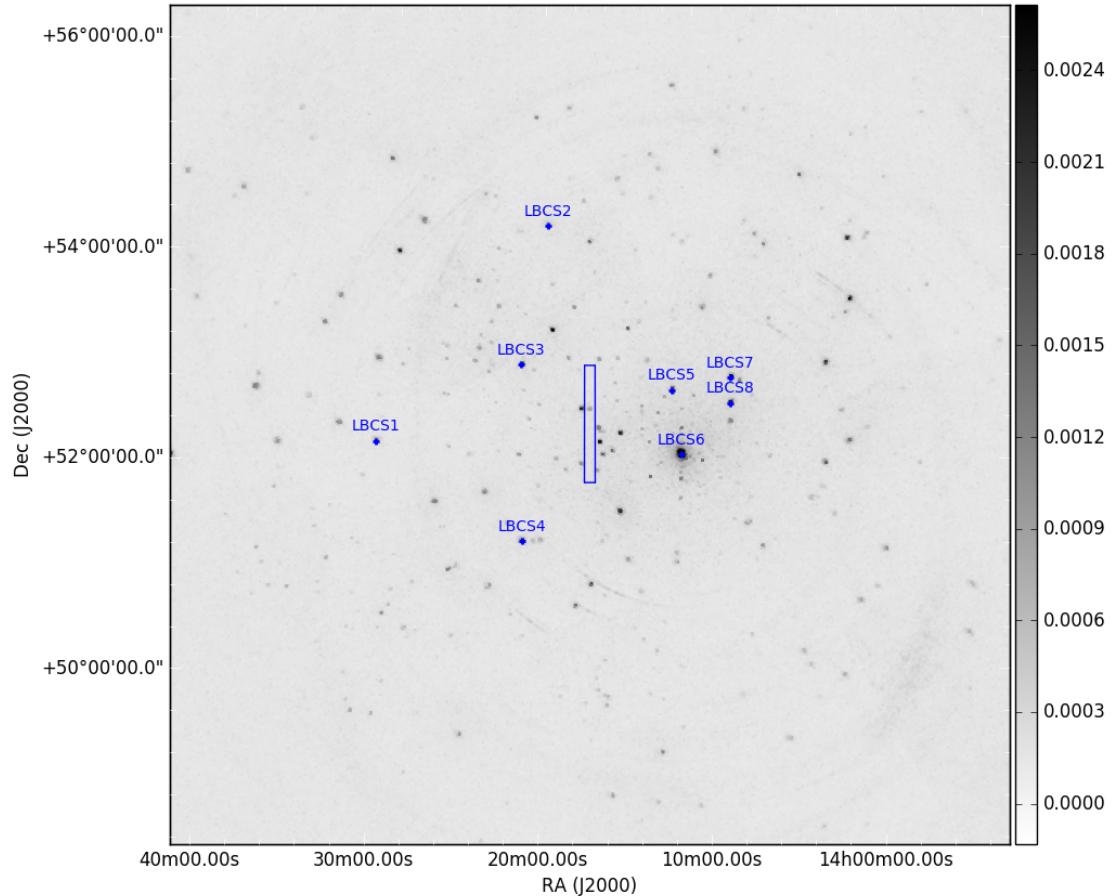


Figure 1.11. Position of the calibrator sources around the EGS, overlaid on the widefield EGS image. LBCS6 is 3C295.

Chapter 2

Techniques of Interferometry

Contents

2.1	A Brief Introduction to Radio Astronomy	23
2.2	Interferometry: Bypassing the Diffraction Limit	23
2.2.1	The Baseline	25
2.2.2	The Visibility	25
2.2.3	The uv -plane	27
2.2.4	The Point-Spread Function	27
2.2.5	From Dirty to Clean: Deconvolving the PSF	30
2.3	Unifying Calibration & Imaging	32
2.4	Weighting Schemes	35
2.5	An Observer's Perspective of Radio Interferometry	37
2.6	Calibration Methods in Radio Interferometry	39
2.6.1	Generational Analysis	39
2.7	The RIME Formalism	44
2.8	Setting up the RIME: a single point source	44
2.9	Expanding our Jones Chain	45
2.9.1	Geometric Effects: the Phase Delay Jones Matrix	45
2.9.2	The Coherency Matrix	46
2.10	Multiple Point Sources	47
2.11	Diffuse Emission: the Full-sky RIME	47
2.11.1	Deriving the Fully-Sky RIME	47
2.11.2	Recovering the van Cittert-Zernike Theorem	48
2.12	Time-dependent Jones matrices, Smearing, and Decorrelation	49
2.12.1	Time-dependence of Jones matrices	49
2.12.2	Smearing and Decorrelation	50
2.13	Noise	51

In this section, we contextualise the work done as part of this PhD by discussing the difficulties introduced by the combination of sparse *uv*-coverage and weak *a priori* constraints on the sky brightness distribution. The conceptual framework which we rely on throughout this manuscript, the Radio Interferometer’s Measurement Equation, is described in greater detail in Section 2.7. We will begin by discussing radio interferometry itself. We will then discuss the so-called *imaging problem*, and end by discussing the *calibration problem*. Note that both of these are artificial subsets of the general inverse problem of radio interferometry. While in practice calibration is done before imaging, it is conceptually helpful to begin with imaging. Until we reach our introduction to radio interferometric calibration, we will therefore assume that calibration has been successfully carried out, and that we are working on the *corrected visibilities*, i.e. gain-corrected visibilities.

We start by linking interferometry to more concrete concepts: specifically, we will give a (very!) brief introduction to radio antennas and their characteristics. We will then use this introduction to show the advantages and drawbacks of radio interferometry, along with the concrete technical problems that the method introduces.

2.1 A Brief Introduction to Radio Astronomy

Radio astronomy consists of observing the electromagnetic emission of astrophysical sources at very long wavelengths by means of radio antennas. These antennas measure a voltage proportional to variations in the electromagnetic field in all the directions they are sensitive to. Since astrophysical signals are weak, a good sensitivity is necessary. Achieving good sensitivity with radio antennas means having a very large collecting area - in this respect, they behave exactly the same way as optical telescopes. Similarly, when well-designed, they are diffraction-limited - this means that, once again like optical telescopes, their resolution is limited by their diameter.

However, because radio frequencies are so much lower, achieving a resolution comparable to e.g. the HST requires extremely large dishes. While there exist telescopes, both old and new, which work on this principle (from Arecibo Observatory, shown in Fig. 2.1a, to the upcoming FAST telescope in China, shown in Fig. 2.1b), the associated technical difficulties (pointing is complex business, as are the associated optics requirements; maintenance costs are high, etc...) have made the single-dish approach prohibitive.



(a) *Arecibo telescope, in Puerto Rico.*



(b) *FAST telescope, in China*

Figure 2.1. Examples of large single-dish radio telescopes.

One way to move past this technical limitation is to resort to interferometry: this is the topic of this section.

2.2 Interferometry: Bypassing the Diffraction Limit

There are two quantities of interest to all astronomers: sensitivity and angular resolution². An instrument's sensitivity is a function of its collecting area³. Resolution, for well-designed instruments⁴, is limited by diffraction in the absence of atmospheric effects. This introduces specific issues in the radio domain. Radio waves have very long

¹ And, famously, attempting to observe “intelligent” sources (cf Enriquez et al. 2017)

² Other kinds of resolution - in time and frequency, for example - are also extremely important to many astronomers, but are not what concern us here.

³ It is also a function of technological factors, of course, but *ceteris paribus*, a more sensitive telescope means a telescope with a wider collecting area

⁴ By this, we mean that we assume that an instrument is also designed to optimise resolution.

wavelengths - often comparable to meters, rather than the $\sim 100 - 1000\text{nm}$ wavelengths of optical light. Achieving a resolution comparable to those of optical telescopes would thus require making telescopes with apertures tens of millions of times larger than those already titanic instruments!

In practice, this technical constraint on resolution is one that astronomers are very interested in overcoming. This technical limitation therefore demands a technical solution. In practice, this solution consists of recourse to interferometric techniques. Indeed, interferometry can be thought of as the construction of a "sparse" dish, as illustrated in Fig. 2.2.

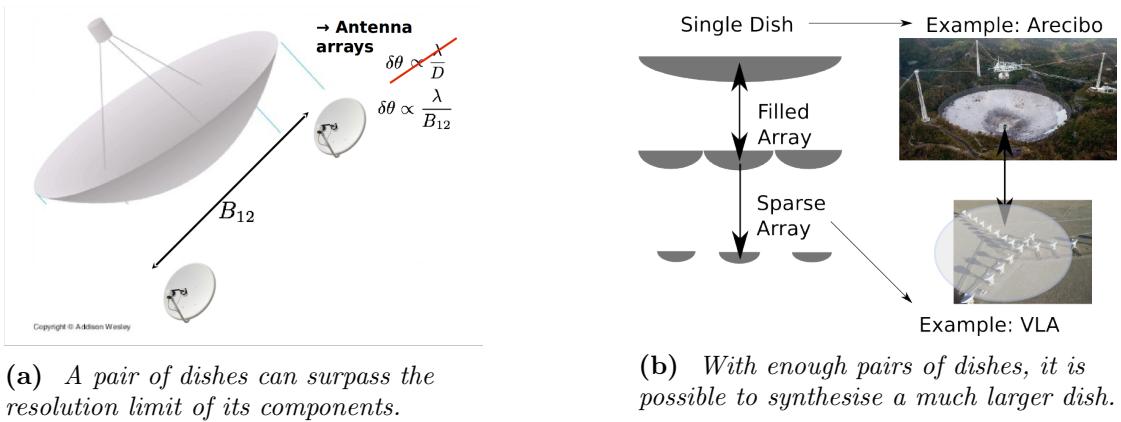


Figure 2.2. Illustration of the underlying principle of interferometry. The 27 dishes of the VLA can be thought of as "synthesising" a similar circular dish as Arecibo. This idea is the reason why radio interferometry is historically known as "aperture synthesis" in the literature of radio astronomy. Fig. 2.2a is copyrighted by Addison Wesley.

Interferometric measurements are made as follows: the elements of the interferometric array (which can be single dishes or phased arrays⁵ of smaller antennas) measure a voltage. This voltage is correlated between antennas, as will be described shortly. Each of these correlations correspond to a single Fourier mode of the sky, as per the Zernike-van Cittert theorem (cf. Section 2.2.4). They must be calibrated, by comparing the observations of a calibrator source and modelled visibilities. Calibration is described in much greater detail in Section 2.7. Then, the calibrated visibilities are used to reconstruct an image of the sky, which often requires deconvolution due to the sparsity of interferometric arrays.

The resolution improvement of interferometers does thus not come for free. To better understand the cost of interferometry, we will cover calibration after explaining the particularities of interferometers. We will begin with the properties of their core component: the baseline.

⁵Described in Section 2.2.2

2.2.1 The Baseline

To define the baseline, we must begin by considering the geometric properties of an interferometric array. For now, let us assume that we are observing the sky above the array, a practice known as drift-scanning. A baseline then consists of the vector subtraction of the positions, in 3-dimensional space, of its two constituent antennas. Note that each antenna pair therefore has 2 corresponding baselines, since for each pair of antennas A and B we create baselines AB and BA. These distance vectors are then divided by the observing wavelength to give a dimensionless set of coordinates, known as (u, v, w) . These coordinates define the baseline entirely.

2.2.2 The Visibility

We have defined what a baseline corresponds to: a vector coordinate in (u, v, w) -space. To each baseline we associate a measurement, which we call the *visibility*. The visibility

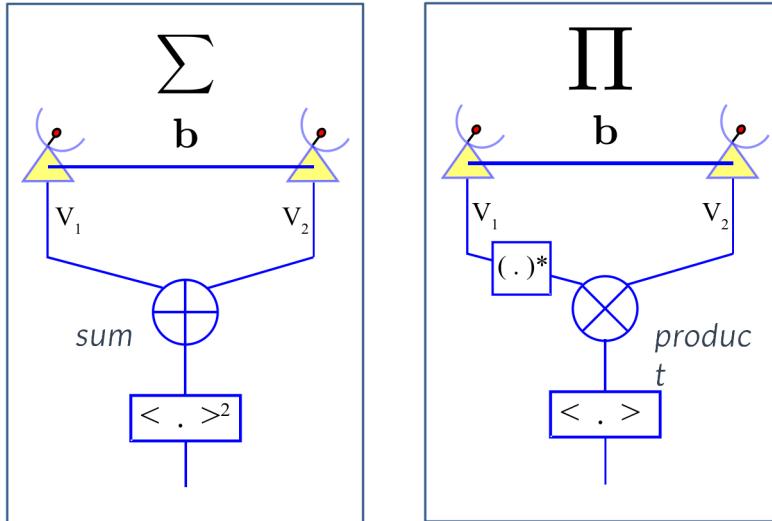


Figure 2.3. There are two ways to combine the voltages from two antennas into a visibility: they are sum-correlation and π -correlation. In this manuscript, we will only concern ourselves with the latter. Image credit: Julien Girard

associated with baseline \mathbf{b}_{AB} is created by taking the voltage V_A measured by antenna A, multiply it by the complex conjugate V_B^* of the voltage measured by antenna B, average over the correlator dump time (i.e. the time over which the measurement is made). This scalar quantity is associated to the baseline position vector \mathbf{b}_{AB} . In other words:

$$\mathbf{b}_{AB} = \frac{\mathbf{x}_B - \mathbf{x}_A}{\lambda_{\text{obs}}} \quad (2.1)$$

$$V_{AB} = \langle V_A V_B^* \rangle_{\delta t, \delta \nu} \quad (2.2)$$

where $\langle \dots \rangle_i$ denotes an averaging over quantity i . We see that a visibility is a complex vector quantity. We also see that $\mathbf{b}_{AB} = \mathbf{b}_{BA}^*$: the information of the visibility associated with baseline BA is contained in the visibility associated with baseline AB. This means that in practice, only half of the visibilities ever need be stored. What does the visibility measurement correspond to?

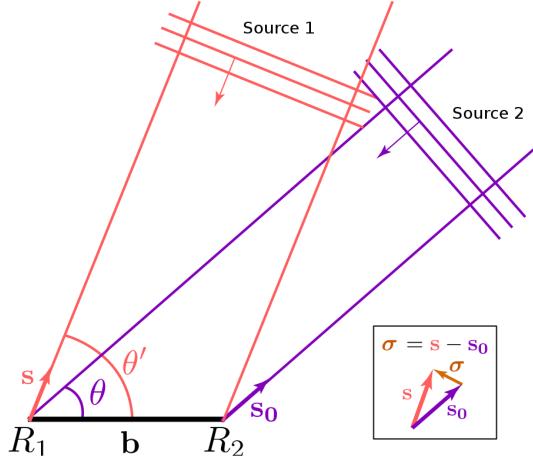


Figure 2.4. Here, we assume that there are only two sources in the sky whose signal can be measured by our antennas. The final visibility is the sum of the visibilities associated with each individual source. Image credit: Julien Girard (Girard 2012)

The signals from different astrophysical do not interfere. They are measured additively in both antennas. Provided that the signal from each sources is coherent when observed by the dishes, the correlation between the voltages measured by antennas A and B will simply be the sum of the voltage correlations associated with individual sources. This is shown in Fig. 2.3: the radio signal from two different sources, at two different locations, is picked up simultaneously by both radio antennas. Because only the correlation between both antennas is measured, the signals do not interfere: if the signal from each source gives a voltage V_p^s for antenna p , then the additive property of electric fields means that the total visibility measured by baseline pq can be written as:

$$V_{pq} = \left\langle \left(\sum_s V_p^s \right) \left(\sum_s (V_q^s)^* \right) \right\rangle_{\Delta t} \quad (2.3)$$

$$= \left\langle \sum_{s,s'} V_p^s (V_q^{s'})^* \right\rangle_{\Delta t} \quad (2.4)$$

$$= \underbrace{\left\langle \sum_s V_p^s (V_q^s)^* \right\rangle_{\Delta t}}_{=0} + \underbrace{\left\langle \sum_{s,s' \neq s} V_p^s (V_q^{s'})^* \right\rangle_{\Delta t}}_{=0} \quad (2.5)$$

where the fact that signal from different sources is not coherent means that we average out cross-contributions (i.e. signal from $s \neq s'$).

Note that in Fig. 2.4, neither source is at the zenith. This introduces the notion of the *effective baseline* which is shorter than the *physical baseline* when measuring a signal from a source that is not at the zenith. Sources in different positions in the sky will be measured by an interferometric array, but the same array will have different effective baselines for different sources. To point an interferometric array in a specific direction in the sky, one must introduce a phase delay in each antenna (or, for fundamental interferometer elements such as the LOFAR tiles or NenuFAR, by playing with the cable length between each dish and the correlator). This is done by three means: delay tracking

(which is exact only for the delay tracking centre⁶, and corresponds to playing with the cable length between antennas in the interferometric array), antenna pointing⁷, and fringe stopping⁸. This means that there are actually three different “field centres”, which are usually (but not always) set to be the same: the pointing (direction of maximum antenna sensitivity), the delay (which can be introduced in cable length) and the phase tracking (which consists of applying a known signal to the measured signal and recording the resulting heterodyne wave so as to keep a reasonable amount of post-correlation data without losing information on sources outside of the phase delay centre). The position of sources are then given in terms of the cardinal angle σ : this is the difference between the position of a source and the phase centre, which is where the array is pointing. For example, in Fig. 2.4, if we point antennas and introduce phase delays such that we are pointing the array towards Source 2, the position of Source 1 will be σ .

2.2.3 The *uv*-plane

For the sake of brevity, radio astronomers tend to talk of a *uv*-plane rather than *uvw*-space to describe visibility space. The set of *uvw*-values for all the baselines of an interferometric array is known as its *uv*-coverage, and defines the array’s properties entirely. For the VLA, for example, the instantaneous *uv*-coverage when observing the zenith will be as shown in Fig. 2.5.

Individual antennas of an array can be pointed mechanically, and so the loss of antenna sensitivity in the direction of interest (and therefore the loss of net interferometric array sensitivity) can be minimised. But what happens to the array itself? It is useful here to go back to the illustration of Fig. 2.2b. Think of each dish in the array representing a “filled” segment of a massive but empty dish. By projecting our observation in a given direction, this dish goes from circular to elliptical.

2.2.4 The Point-Spread Function

We have seen that the purpose an interferometric array is to overcome the diffraction limit of single-dish antennas. We have described visibilities, which are the quantities measured by an interferometric array. What remains is to describe how these measurements are related to the sky brightness distribution.

Assuming that all the antennas in an array are equivalent and perfectly calibrated, the van Cittert-Zernike theorem (van Cittert (1934), covered in Thompson et al. (2001)) allows us to equate a visibility with a single Fourier mode of the plane tangent to the sky where the instrument is pointed. This direction is called the *phase centre*, because phase shifts are introduced between antenna voltages before averaging so as to point

⁶Source: http://www.astron.nl/eris2013/Documents/2_laing_fundamentals_interferometry.pdf, <https://science.nrao.edu/science/meetings/2016/15th-synthesis-imaging-workshop/SISSL5Advanced.pdf>

⁷This maximises antenna sensitivity in a given direction, but is not done with LOFAR

⁸Fringe stopping consists of down-converting the signal received by antennas by mixing with a local oscillator. The geometric delay compensation is usually applied on this down-converted signal. Source: http://www.gmrt.ncra.tifr.res.in/gmrt_hpage/Users/doc/WEBLF/LFRA/node74.html

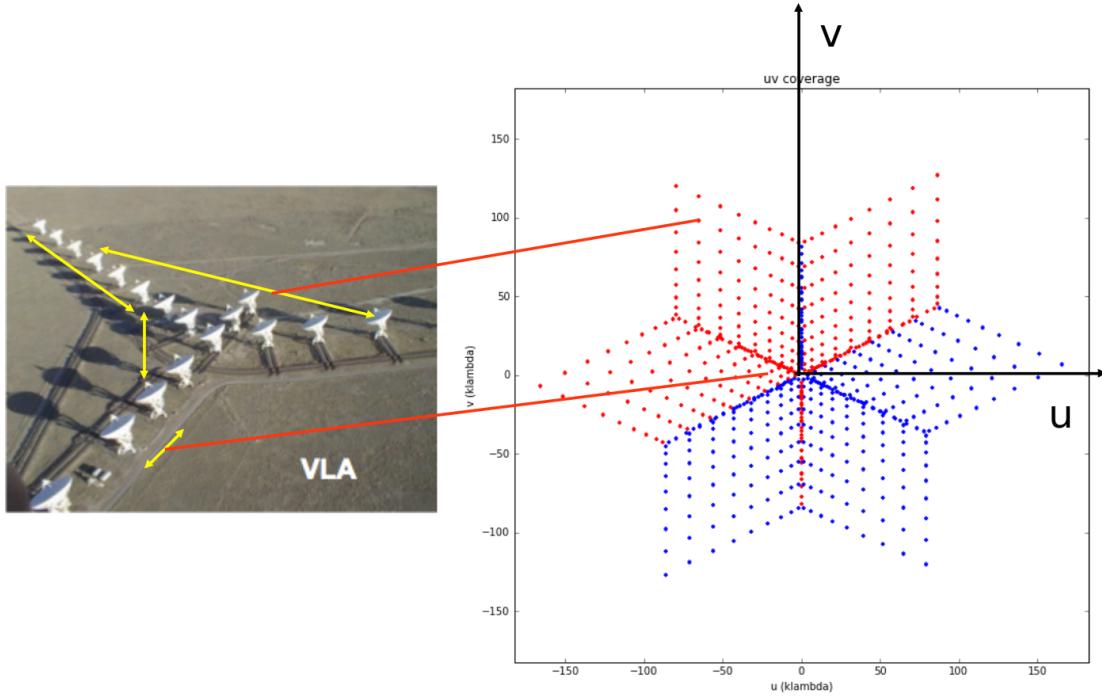


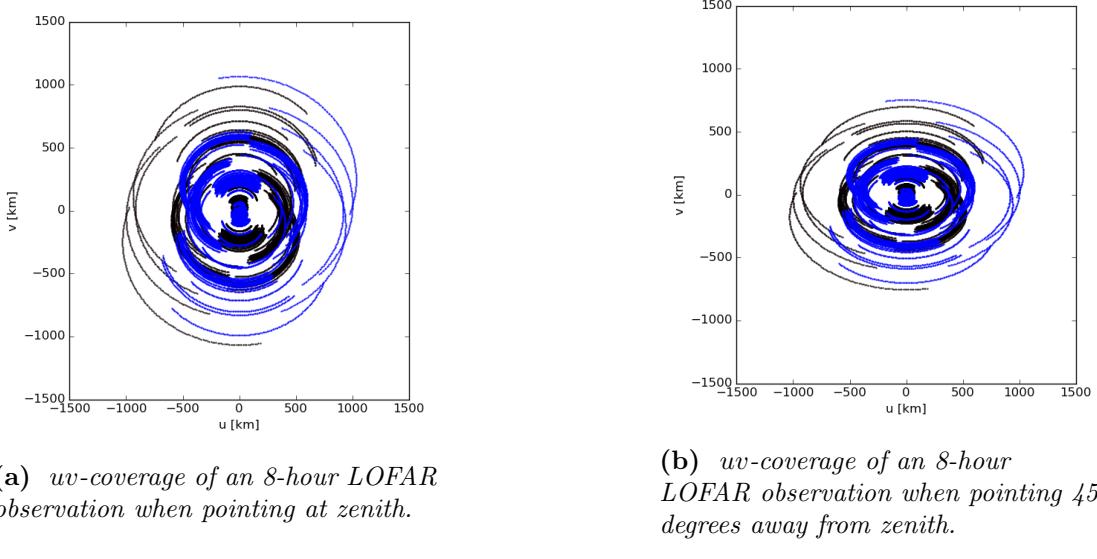
Figure 2.5. The VLA contains 27 radio dishes placed as shown above. Each antenna pair between those 27 gives two single baselines, here, one red and one blue. Image credit: Julien Girard

each visibility in this direction. This behaviour is encoded in the Fourier kernel, which is written as:

$$k_p = 2\pi i((u_p l + v_p m + w_p(n - 1))) \quad (2.6)$$

The Point-Spread Function (PSF) of an instrument is its point source response: it will determine the angular resolution limit due to the instrument. We know that the Fourier transform of a single point source with unit flux at phase centre is a constant. If we measured the full (infinite) visibility space, then we could reconstitute a point source perfectly. In practice, however, we only ever measure a subset of the visibility space: an interferometric array's PSF is thus the inverse Fourier transform of the array's *uv*-coverage, as these are the only points for which we have information on visibility space. The more elements in the array, the better the *uv*-coverage and thus the better the PSF. An interferometer's PSF will always be worse than an antenna dish of equivalent diameter, the UV-coverage of which is the inverse Fourier transform of a disc. This is because an interferometer behaves like a *sparse* dish: the sparsity of our coverage manifests as very large point-spread function. The more baselines are available, the less sparse the UV-coverage, and so the better the PSF. This is shown in Fig. 2.7.

Mathematically, the *uv*-coverage of an interferometer can be treated as a series of delta functions centred on the antenna positions. To go from visibility space to image



(a) *uv-coverage of an 8-hour LOFAR observation when pointing at zenith.*

(b) *uv-coverage of an 8-hour LOFAR observation when pointing 45 degrees away from zenith.*

Figure 2.6. *Effect of array pointing on uv-coverage. By pointing the array 45 degrees in the v -axis, the array's uv-coverage (and thus maximum resolution) is decreased along the v -axis. Two colours are used, because $V_{AB} = V_{BA}$, and so only half the visibilities are stored in practice.*

space, we apply the inverse Fourier transform, giving us:

$$\mathbf{S}_{\text{obs}} = \mathcal{F}^* \left[\mathbf{B} \sum_i^n \mathbf{x}_i \right] \quad (2.7)$$

where \mathbf{S}_{obs} is the observed sky brightness distribution, \mathbf{B} the Fourier transform of the true sky brightness distribution \mathbf{S} , and n is the number of antennas in our array. \mathcal{F}^* is the inverse Fourier transform. By using the convolution theorem, we immediately see that

$$\mathbf{S}_{\text{obs}} = \underbrace{\mathcal{F}^* [\mathbf{B}] * \mathcal{F}^* \left[\sum_i^n \mathbf{x}_i \right]}_{= \mathbf{S} \quad \text{the PSF}} \quad (2.8)$$

$$(2.9)$$

and that our image is the convolution of the “true”, underlying sky brightness distribution, and our array’s PSF. As $n \rightarrow \infty$, the uv -plane fills, and the inverse Fourier transform of this unit plane becomes 1. This illustrates the description above.

This PSF can be improved by assuming that the sky is constant over periods of time (supersynthesis) or constant over fractions of the total bandwidth. This improves UV-coverage by a factor $N_t \times N_\nu$, the number of measurements in time and frequency respectively. The effect of supersynthesis on the PSF is shown in Fig. 2.8.

Of course, even with these techniques, the PSF is never perfect, and flux from very bright sources will contaminate the rest of the field. The source in the last image of Fig. 2.8, for example, only appears point-like by contrast to the others. If we wish to recover faint emission - and when interested in deep extragalactic fields, that is exactly what we

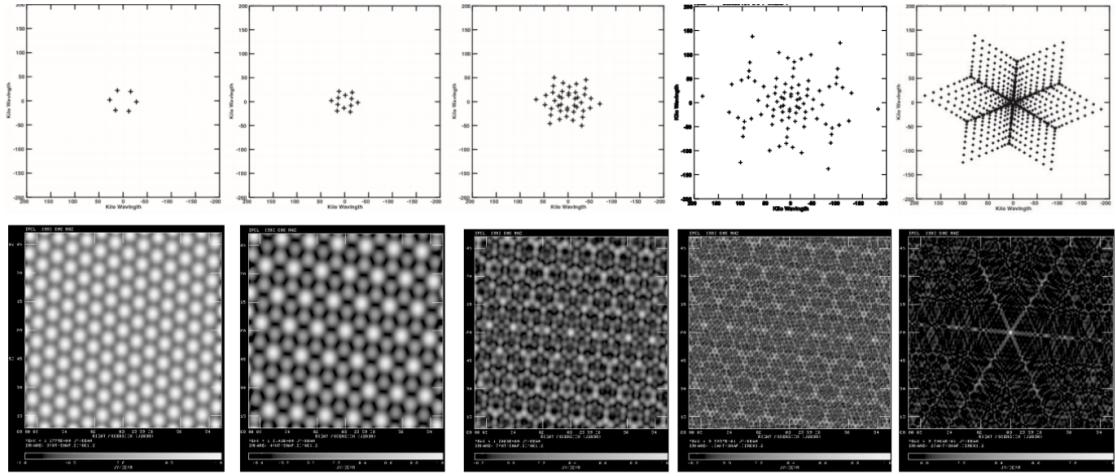


Figure 2.7. Plot of the PSF associated to more and more complete UV-coverage (here, consisting of more and more elements of the VLA). Image credit: Rick Perley, 3GC4 lecture.

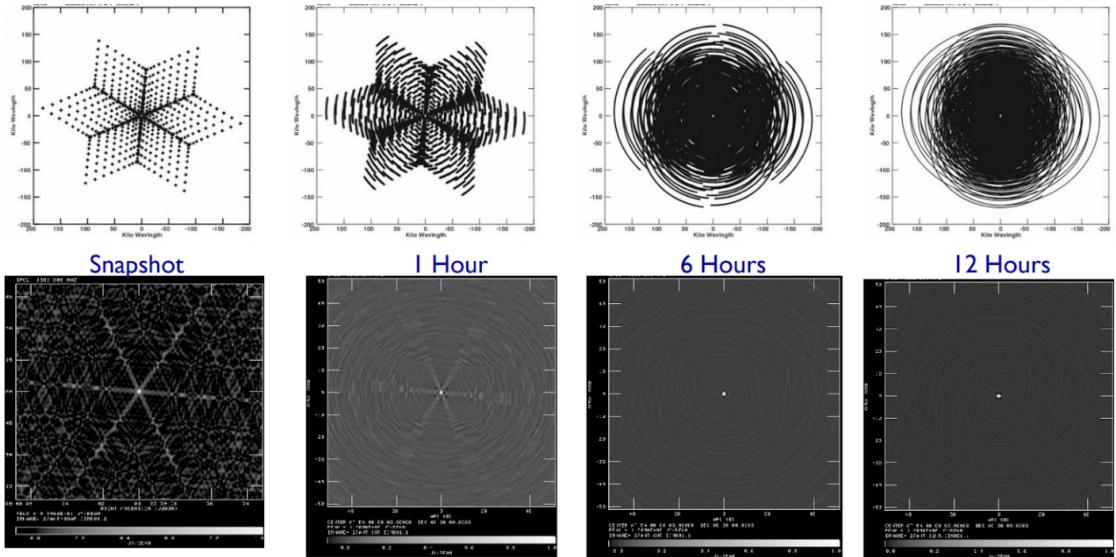


Figure 2.8. Effect of supersynthesis on the UV-coverage of the full VLA. Image credit: Rick Perley, 3GC4 lecture.

wish to do - then we need to ensure that the effect of the PSF in our images is mitigated, for the brighter sources at least, lest their sidelobes (tails of the PSF convolved with the source) drown out signal from fainter sources. This means *deconvolving* the PSF from radio interferometric images.

2.2.5 From Dirty to Clean: Deconvolving the PSF

Raw images made from radio interferometric data consist of the underlying flux distribution convolved to the array's PSF. The quantity of interest to astronomers is the underlying flux distribution. To recover that information from images, deconvolution is needed.

In practice, performing a full deconvolution on large images (easily over 25 million pixels in total) is not computationally viable. Scientists therefore resort to algorithms and techniques to accelerate imaging and deconvolution. For example, the use of Direct Fourier Transforms (DFT), which are extremely slow, is avoided when going from visibility space to image space - Fast Fourier Transforms (FFTs) are preferred. Similarly, one seeks to avoid having to perform full deconvolution on the dirty image, since many pixels in the image do not actually contain astrophysical signal. In this chapter, we will briefly cover the most common method of deconvolution used in radio astronomy.

We use dynamic range (DR) as a quality metric in radio images. DR is the ratio of the flux of the brightest source in the field to the flux of the faintest source in the field detectable above the noise. It is therefore a function of both source brightness and noise, and so we define dynamic range as:

$$DR = \frac{S_{max}}{\max(S_{min}, \sigma)} \quad (2.10)$$

where S_{max} is the flux of the brightest source in the image, S_{min} the flux of the faintest source detectable above the noise, and σ is the noise in the image. The higher the DR, the deeper we image the sky, and thus the better the image. There are two big constraints to reckon with: firstly, we do not want to deconvolve the PSF from all pixels in the image, but only from those within which large amounts of underlying flux fall. Secondly, if we only deconvolve some pixels rather than performing a full deconvolution, then care must be taken not to treat artefacts in the field (which can be created from overlapping PSF sidelobes from neighbouring sources, even with perfect calibration) as true flux. Doing so would lead to an attempt to deconvolve a PSF from a pixel while assuming an incorrect underlying flux value for this pixel and thus result in the introduction of further artefacts in the image.

The dominant family of deconvolution algorithms used in radio astronomy at the time of writing is CLEAN⁹ (see Högbom 1974; Pearson & Readhead 1984), which is understood in compressed sensing theory as a matching-pursuit algorithm optimising L2 (i.e. a least-squares fit) with the addition of an L1 regularisation (i.e. the total flux in the model is minimised). The various CLEAN algorithms seek out the brightest pixel(s) in the image, potentially applying a mask to the image first in cases where some *a priori* knowledge on flux distribution is available. It then deconvolves those pixels up to a predefined threshold (either a fraction of the initial pixel value or some factor of the estimated noise value). It does this some predefined number of times. Then, it collects the aggregate *model* (the flux estimate for each deconvolved pixel up to this point), convolves the model with the PSF, and subtracts the result from the dirty image. It then begins to choose the brightest pixels again, and continues to clean until it reaches a stopping condition (usually either a flux value or a maximum number of iterations).

In Fig. 2.9, we show the effect of deconvolution on a dirty image of 3C295 (a very bright & complex source which is the subject of much of this thesis' work). Since this source includes diffuse emission, standard CLEAN is insufficient (since CLEAN assumes

⁹The other main family of deconvolution algorithms, Maximum Entropy Methods (MEM), are not very widely-used at the time of writing, but remain an active area of research (cf. Chael et al. (2018)).

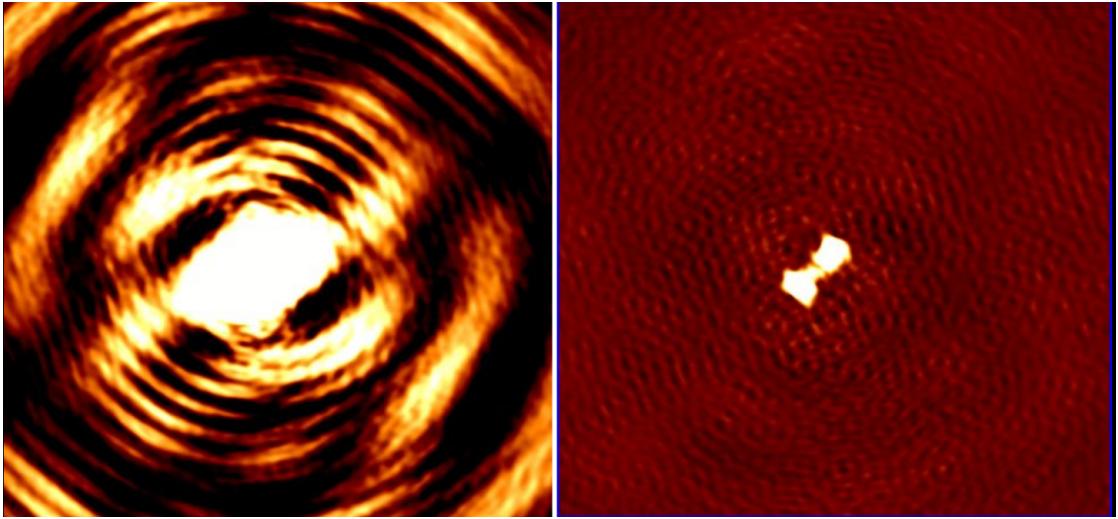


Figure 2.9. Effect of deconvolution on a dirty image of 3C295. To the left, the dirty image; to the right, the deconvolved image (same size, resolution & flux scale, made with the same data). As we can see, the latter is much more scientifically interesting.

a sparse distribution of point-like sources). Instead, the DDFacet software package's novel SSD/GA algorithm (see Tasse et al. (2017)), which deconvolves islands consisting of sets of pixels at once, was used. This allowed for a far greater dynamic range than standard CLEAN could achieve, while also more properly recovering diffuse emission: when imaging a resolved source, this is absolutely necessary.

The problem of imaging radio interferometric data is still a very active field of research today. For more information on this topic, the NRAO website¹⁰ is an excellent start for a general introduction to this topic, while Tasse et al. (2017) provides an outstanding example of cutting-edge work at the time of writing.

2.3 Unifying Calibration & Imaging

The previous section mentions that good deconvolution is dependent on good calibration: poor calibration introduces artefacts in images, which can in turn bias deconvolution. A wrong model of the sky brightness distribution, in turn, will bias calibration, introducing artefacts in images. This feedback of one inverse problem onto the other is a strong indication that both are, in fact, different facets of the same underlying inverse problem: going from measured visibilities to a sky brightness distribution. Indeed, one could - conceptually, at least - solve for both problems at once: find both the sky brightness distribution and the antenna gains which best fit the measured visibilities. In practice, this is not practical (though it is feasible); not only is there no proof that the overall calibration & imaging problem is convex, but even if it is, it is also poorly conditioned: multiple sets of sky brightness distributions & antenna gains are equally viable for a given set of measured visibilities. This means that, even if the problem is convex, it takes an extremely long time to converge properly. Examples of convexity

¹⁰See the CLEAN section of the NRAO website.

and conditioning are shown in Fig. 2.10.

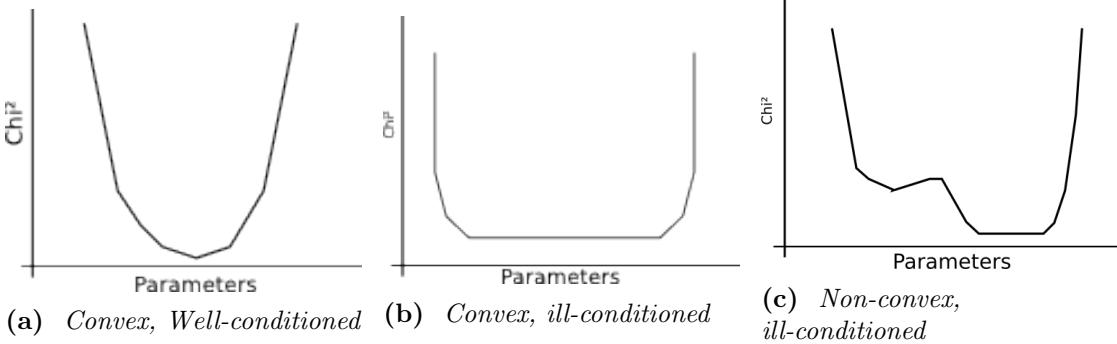


Figure 2.10. Examples of poor conditioning and non-convexity, plotted in terms of residual χ^2 values vs parameter values. Note that all the parameters of our problem are collapsed onto a single axis and treated as a scalar: this is for illustrative purposes. Note that the local minimum of the non-convex example is well-conditioned while its global minimum is ill-conditioned.

Bad conditioning and non-convexity can be alleviated through regularisation: by specifying a well-chosen prior (e.g. for interferometry, using a good sky model for calibration), the underlying distribution of the solutions can be constrained to fall in a narrow, convex region. This is shown in Fig. 2.11

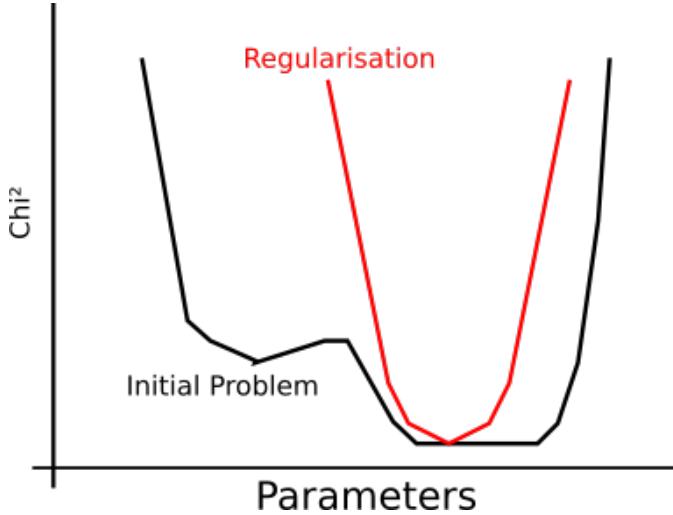


Figure 2.11. Plot showing the effect of regularisation on an ill-conditioned, non-convex problem. The χ^2 fit is constrained to lie in a convex, well-conditioned region: it falls within the global minimum, at at one point only rather than somewhere in a valley.

Because the imaging problem (i.e. the problem of deconvolution) is linear, it can only be convex - though it is not necessarily well-conditioned. As for the calibration problem, experience strongly indicates that it is nearly convex *and* ill-conditioned. For calibration convexity, Cyril Tasse (pers. comm.) has found that the Hessian matrix of the gain solutions is positive semidefinite, i.e. that all its eigenvalues are positive

and non-zero except for one. This behaviour likely explains why self-calibration is, in the practical experience of radio-interferometric practitioners, neither well-conditioned nor convex: if the problem is poorly regularised, then the same initial model may give different end results, and to truly verify the convergence property of self-calibration would require both a full Monte Carlo simulation and enough time to allow the very slowly-converging procedures to finish converging. It is not a practical thing to test rigorously.

Because the net calibration & imaging problem has so many degrees of freedom, it is convenient to break it into two parts: calibration and imaging. By doing this, instead of exploring a massive set of hyperparameters (e.g. relative flux distribution at a given frequency, overall spectral index, the 8 values of a complex 2×2 matrix per antenna per time and frequency interval, ...), a smaller subset of these hyperparameters is explored while all else is kept constant. By ensuring that these subsets each give reasonable results individually, it becomes possible to solve the full problem. This practice of imposing specific priors to each subset of the problem is a form of regularisation: it can turn a poorly-conditioned problem into a better-conditioned one, provided that the priors are well-chosen.

In effect, this means going from exploring the sky-gains parameter space all at once to exploring it along one “axis¹¹” at a time. This is illustrated in Fig. 2.12.

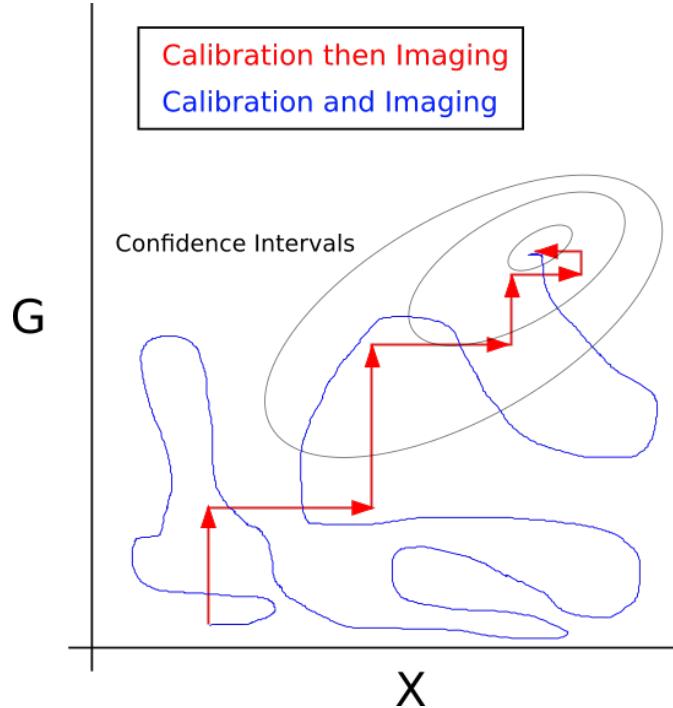


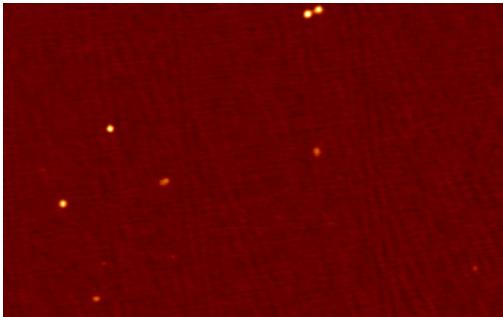
Figure 2.12. Plot showing the exploration of sky-gains (X - G) parameter space. In red, this space is explored one axis at a time (this is “self-calibration”); in blue, the full space explored at once. Here, we show both procedures converging to the best confidence value: in practice, this is rarely the case.

¹¹which includes all the hyperparameters associated with that subset of the problem

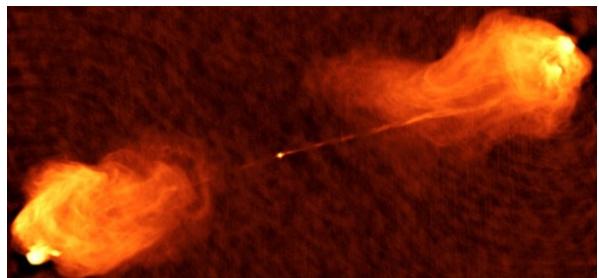
In this context, a good deconvolution algorithm is one that is able to minimise how much it is affected by calibration errors while maximising how much true underlying flux it can recover into a model, while a good calibration algorithm is one that is able to minimise how much it is affected by sky model errors while maximising how much true underlying gain structure it can recover.

2.4 Weighting Schemes

We will finish our introduction to the inverse problem of interferometry by discussing one method by which the conditioning of deconvolution can be improved. Consider two extreme cases: a field containing a few point sources (Fig. 2.13a), and a field containing a diffuse and turbulent source, with very complex flux distribution at all scales (Fig. 2.13b). Deconvolving the PSF from the first image using CLEAN algorithms will be child’s play, but doing the same with the second will be extremely complex. The first case is well-conditioned, and the second to ill-conditioned.



(a) Part of the LOFAR field near the Extended Groth Strip as seen by LOFAR.



(b) Cygnus A as seen by the VLA. Image courtesy of NRAO/AUI

Figure 2.13. Two extreme examples of good and poor conditioning. Fig. 2.13a shows a field with a few point-like sources. Fig. 2.13b shows a very complex & diffuse structure.

The better the conditioning, the easier the deconvolution and the closer to the ground truth its result. As such, when the true underlying structure of a source is poorly-known, getting the best conditioning possible for deconvolution becomes a key concern. Thankfully, because the actual measurements taken with interferometers are the Fourier modes of the sky brightness distribution, they can be weighted when reconstructing the images to fulfill specific requirements. In particular, Briggs weighting (Briggs 1995) gives a sliding parameter between *natural* weighting, where each visibility is weighed equivalently (this optimises noise in the field) and *uniform* weighting, which gives each visibility a weight inversely proportional to *uv*-plane density. Uniform weighting optimises for resolution at the cost of image signal-to-noise.

Fig. 2.14 shows images done with the same data and the same imaging parameters bar weighting scheme choice. We can see that different weighting schemes give radically different deconvolved images. Images made with uniform (Fig. 2.14b) or Briggs (Fig. 2.14c) weighing are much sharper, with uniform weighting (Fig. 2.14b) giving such

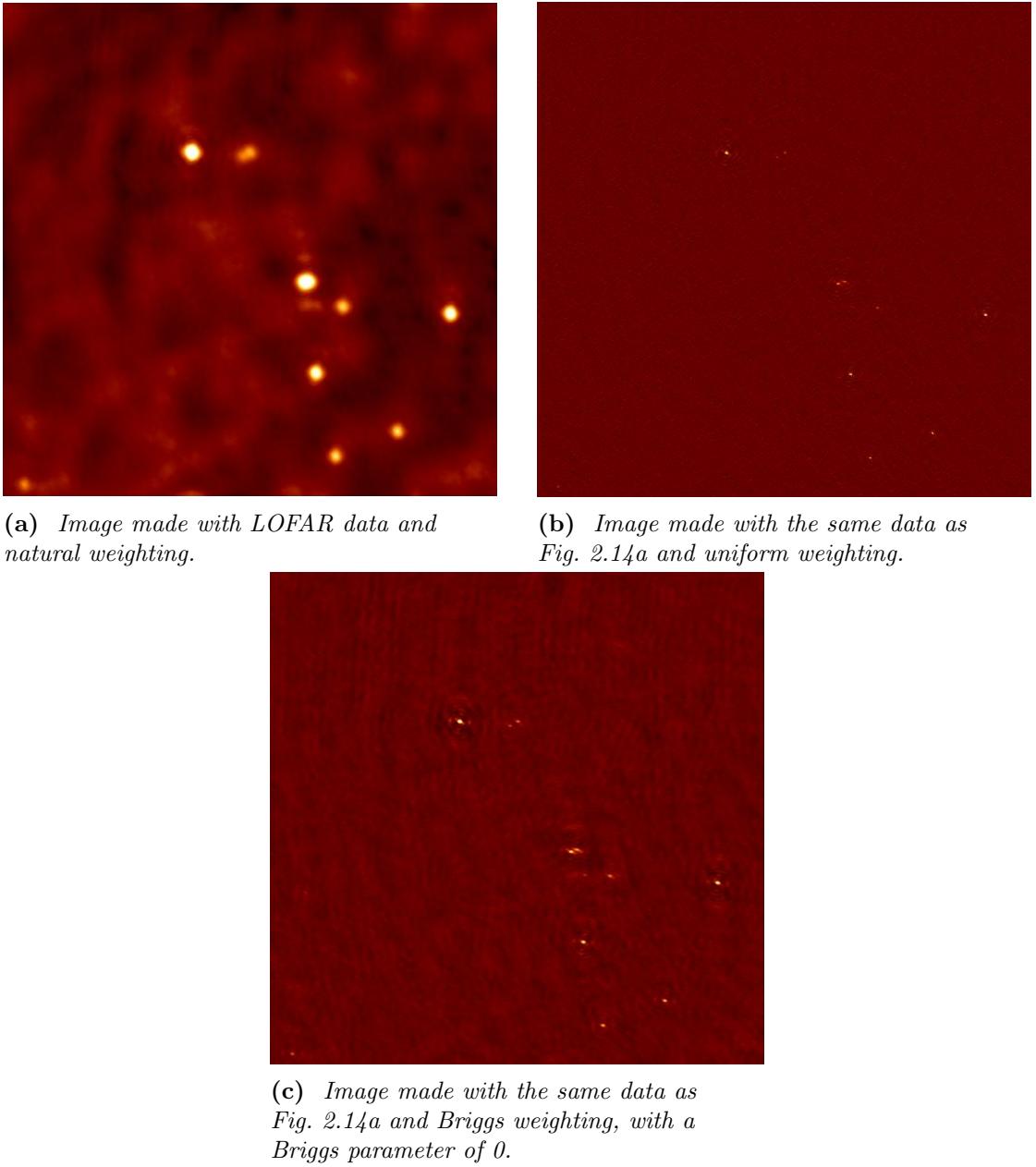


Figure 2.14. Impact of different weighting schemes on interferometric image reconstruction. All the images shown above are deconvolved.

resolution that the sources are nearly invisible in the field (though they are present - the contrast was not upped so as to show all three images on the same flux scale).

We close this section by noting that data flagging can be considered a form of weighting scheme. This consists of assigning a null weight to those data points considered too corrupted by various instrumental or atmospheric effects (typically, Radio Frequency Inter-

terference - RFI - cf. Offringa (cf. 2010) for an example) to be scientifically useful. Those weights which are deemed too corrupted are assigned a weight of zero. This corresponds to simply dropping those data points which would corrupt the image reconstruction by introducing unphysically strong fringes in the image.

Flagging is also useful to remove corrected visibilities for which near-zero gains are found: the associated corrected visibilities will consist of some number divided by a near-zero number, for which numerical errors can introduce very large errors. “Clipping” these visibilities after calibration helps improve the final images.

2.5 An Observer’s Perspective of Radio Interferometry

Astronomers who request observation time on interferometric arrays are generally not experts in the theory and operation of these arrays. This is a simple consequence of scientific division of labour. This thesis is written from the perspective of an “expert” user: one who specialises in understanding interferometric arrays and reducing their data, rather than their subsequent analysis. However, it can be very helpful to contextualise this perspective in the experience of other scientists, which is what this section aims to do. This is simply to highlight the conceptual structure of a radio interferometric observation, and where the concepts outlined in this chapter fall in the overall pattern. In other words, it is best understood as a guide to see where a given section’s topic slots into the overall production of an image made from visibilities.

LOFAR data can be acquired through observation or via the Long-Term Archive. The worst of this data will already be flagged. Users can thus generally proceed directly to calibration. Because there is no guarantee of sufficient signal-to-noise in the “target field” (the object that the astronomer is actually interested in), it is standard practice to observe a calibrator source for 5 minutes before and after each observation. Such sources typically need to be bright and unresolved for the interferometer being used. Because they are bright and at phase centre for the 5-minute observation, the SNR when solving for calibration solutions for these 5 minutes will tend to be quite good. These solutions will then be interpolated between the 5 minutes onto the 8-hour observation. Calibration will be described in Section 2.6.

So far, the astronomer will have worked entirely with visibilities. However, generally speaking, the aim will be to get information on the object’s spatial brightness distribution at the observing frequency. This will require *imaging*: going from visibilities to the image-plane. This can be done through a variety of tools, but will generally involve deconvolution. Imaging will be the subject of the rest of this chapter, and deconvolution will be described in more detail in Section 2.2.5.

After obtaining the initial image, it may be desirable to perform a few rounds of self-calibration (cf. Brogan et al. 2018) on the target field. This consists of extracting a model from the initial image and using it as a new calibration model, then re-imaging. Doing this will usually dramatically improve the final image.

The framework for interferometric calibration is complex, and understanding it requires a good grasp on interferometry itself. For now, we will assume that calibration has been performed perfectly, and that the gain-corrections have been applied to the voltage measured by our antennas. We therefore work with the true signal from astrophysical sources until Section 2.7.

2.6 Calibration Methods in Radio Interferometry

In this section, we will discuss the implementation of interferometric array calibration. Our analysis is based on the RIME formalism, described in Section 2.7. One key metric of calibration quality is the *dynamic range*, mentioned in Sec. 2.2.5. High dynamic ranges mean that a high contrast has been obtained, and fainter sources can be reached. Here, because the difference between thermal noise and artefacts becomes relevant, we redefine dynamic range as follows

$$DR = \frac{S_{max}}{S_{min, \max(\sigma_{thermal}, \sigma_{artefacts})}} \quad (2.11)$$

where S_{max} is the flux of the brightest source, S_{min} the flux of the faintest source detectable above the noise, $\sigma_{thermal}$ the thermal noise in the image, and $\sigma_{artefacts}$ the noise associated with calibration artefacts. This gives a definition for dynamic range useful even in fields with a single source. We can thus use it as a metric for calibration quality specifically.

The distinction between these two noise sources is crucial; one can never go ‘below noise’ for a given observation, no matter the quality of calibration. Astronomers typically observe for longer periods of time in order to reduce $\sigma_{thermal}$ in their images, but this will not reduce the artefacts caused by poor calibration solutions. Uncorrected Direction-Dependent Effects¹² will not go away on their own, no matter how long the integration time. Similarly, there is a limit to how much improving calibration will improve the final image: eventually, more data is required to drive noise down.

We can distinguish three ‘generations’ of calibration methods, of increasing complexity (Noordam & Smirnov 2010). We will describe them in terms of the RIME, showing how each generation increases in generality to account for more exotic effects. They are referred to interchangeably as ‘nth-generation calibration’ or ‘nGC’ methods.

2.6.1 Generational Analysis

First-Generation: Open-Loop Calibration

First-generation calibration methods (1GC methods) consist of open-loop calibration. This relies entirely on instrument stability, and thus imposes significant design constraints on radio telescopes. It consists of briefly observing an external calibrator before and after each observation run to find calibration solutions for those calibrator observations, and then interpolate between them over the “target” observation. In RIME terms (see Section 2.7), this consists of solving for a very basic form of \mathbf{G}_p :

$$\mathbf{G}_p = a_p \mathbf{I} \quad (2.12)$$

where a_p is a complex constant solved for during open-loop calibration. By interpolating between the two calibrator observations, it is possible to have a linear time-dependent gain estimate, though it is not much more complex than the above.

¹²See Section 2.10

While values for a_p and b_p can in theory be found for both autocorrelation and both crosscorrelations (e.g. XY and YX for linear receptors), low signal-to-noise means that in practice, a single set of values is solved for per antenna¹³. With these techniques, one can achieve dynamic ranges of about 100:1 ((Noordam & Smirnov 2010)).

Second-Generation: Self-Calibration

Second-generation calibration methods (2GC methods) are defined by *self-calibration*¹⁴, commonly referred to as *self-cal* ((Pearson & Readhead 1984)). As described in Section 2.11.2, this method can only be deployed if the brightness matrix of the sky is the same for all baselines¹⁵ (i.e. that we are not affected by direction-dependent effects).

The first instance of self-calibration (and adaptive optics in radio interferometry) is a paper published in the era of 1GC by Jennison (Jennison 1958), expanding on his PhD work, which was published in 1951). With sufficient signal-to-noise, he showed that *phase closure* could be calculated and errors due to the atmosphere thus mitigated.

Self-calibration and adaptive optics are essentially the same thing, albeit meeting different constraints. Most notably, interferometric data is digital, allowing radio astronomers to perform their adaptive optics correction after the observation.

If amplitude and phase gains can be written as antenna-dependent, then each antenna-based error is estimated N times (once per each baseline which includes this antenna). By estimating, and correcting for, these errors, one can use a simple source model to infer an improved one. This is why the method is referred to as self-cal; by calibrating ‘on’ a good calibrator source (bright, compact and unresolved), one can drastically improve one’s source model, along with one’s calibration solutions.

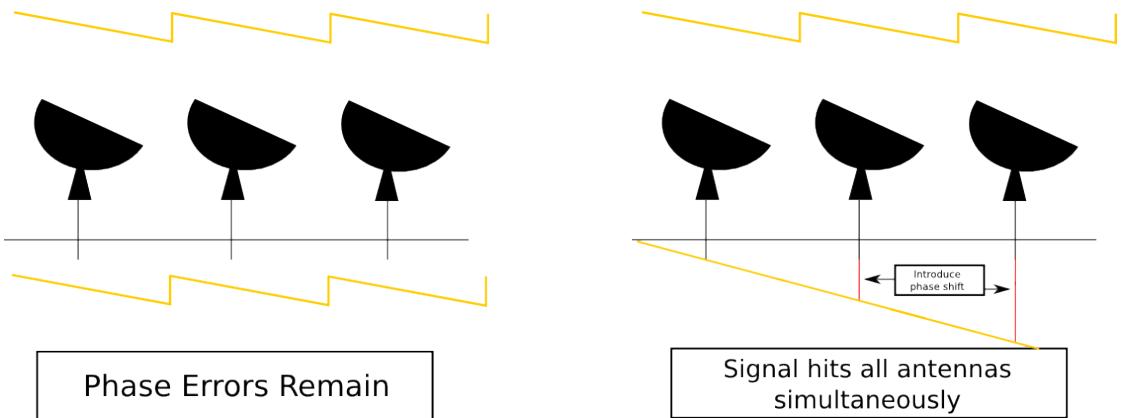


Figure 2.15. Representation of self-calibration as adaptive optics

¹³This reduces calibration to solving only for the intensity gains: the data can then only be used for *intensity mapping* (e.g. (Jennison 1957)). This practice therefore precludes polarimetry.

¹⁴On the discovery of self-calibration and its evolution in parallel to adaptive optics, see the chapter titled ”The Almost Serendipitous Discovery of Self-Calibration” in ”(Kellerman & Sheets 1984).

¹⁵This is not to say that the brightness matrix must contain only point sources, but rather that moving our interferometer 500m to the East should not change its measured visibilities

By extending this idea to VLBI (see e.g. (Rogers et al. 1974)), *amplitude closure* was introduced to the field along with phase closure (as described in Rogers et al. 1983). These quantities are immune to antenna-based effects.

The great advantage of radio interferometry over optics, however, is that we can *iterate* over progressively improved source models. This is because interferometric data is digital, and because we record phase information.

In practice, of course, self-cal will be limited by noise. More precisely, it will be limited by two factors: sensitivity, in the form of the signal-to-noise ratio (henceforth SNR), and the atmosphere (as sufficient SNR is needed per coherence time, which is a function of atmosphere activity). For sufficiently bright sources, with high SNR, self-calibration can easily improve dynamic range by a factor of 10 (Kellerman & Sheets (1984), p. 154) in a single iteration.

Third-Generation: Direction-Dependent Effects

Third-generation calibration (3GC) is an extension of 2GC calibration which takes direction-dependent effects into account. At the time of writing, this is the cutting edge in radio interferometric calibration.

In this section, I will discuss the *facet-based* approach taken by Tasse et al. (2017). The approach cannot be divorced from imaging, for the simple reason that solving for and applying Jones matrices¹⁶ in different directions requires image-plane knowledge of the sky brightness distribution \mathbf{B} .

When performing first and second-generation calibration, one solves for a set of Jones matrices or parameters thereof. Direction-independent calibration then consists of applying these calibration solutions to the visibilities directly. However, while these solutions will be correct in the direction of the calibrator source, they are increasingly likely to be wrong as a function of distance from said calibrator source. The difference between the true gains in a given direction and the gains estimated from the calibrator source are called *differential gains*.

Why then not solve for a set of calibration solutions for each source in the calibration model, iteratively adding new sources to this model as they become above the noise as calibration artefacts - introduced by poor calibration - decrease? There are two main reasons for this: firstly, the SNR when calibrating using very faint source models will be poor. This means that gains will likely be extremely noisy, and applying them will result in increased calibration artefacts in the final image. Secondly, even by optimising for signal-to-noise and only using a few of the brightest sources in the field, the problem quickly becomes intractably large in terms of computing power required, provided that standard complex differentiation is used (see van Weeren et al. 2016). Wirtinger differentiation (see Tasse 2014; Smirnov & Tasse 2015) does alleviate this somewhat by allowing the Hessian matrix to be written as a block-diagonal matrix, but is not implemented in all standard direction-dependent calibration software at the time of writing.

¹⁶See Section 2.7 for a full discussion on calibration and Jones matrices

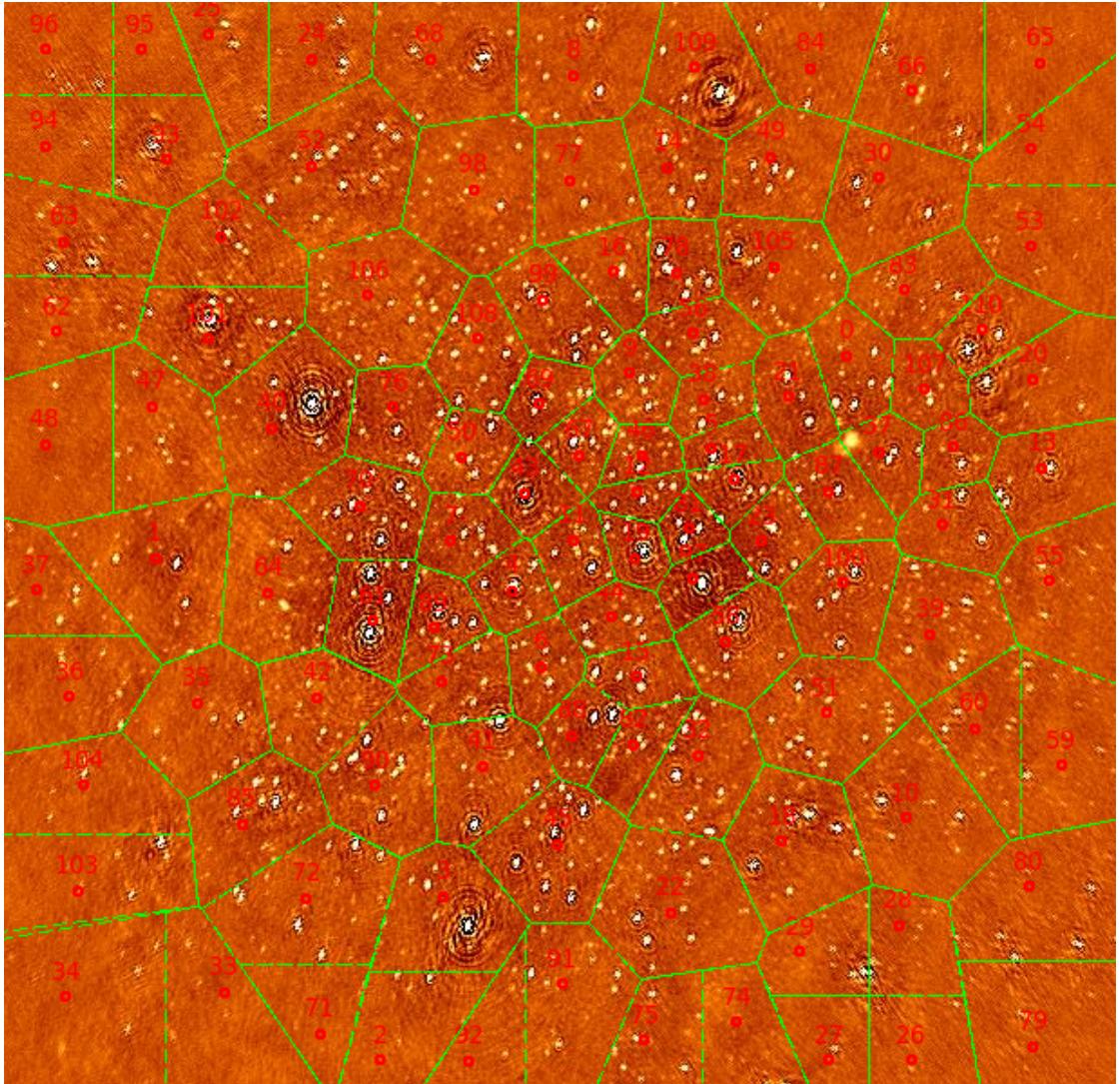
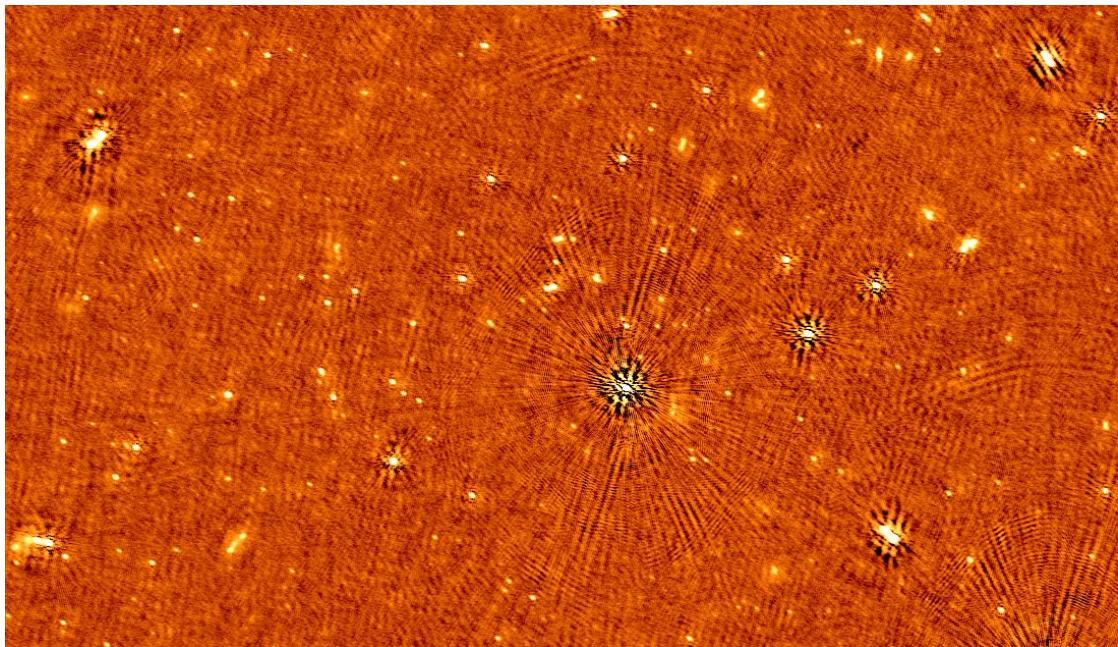
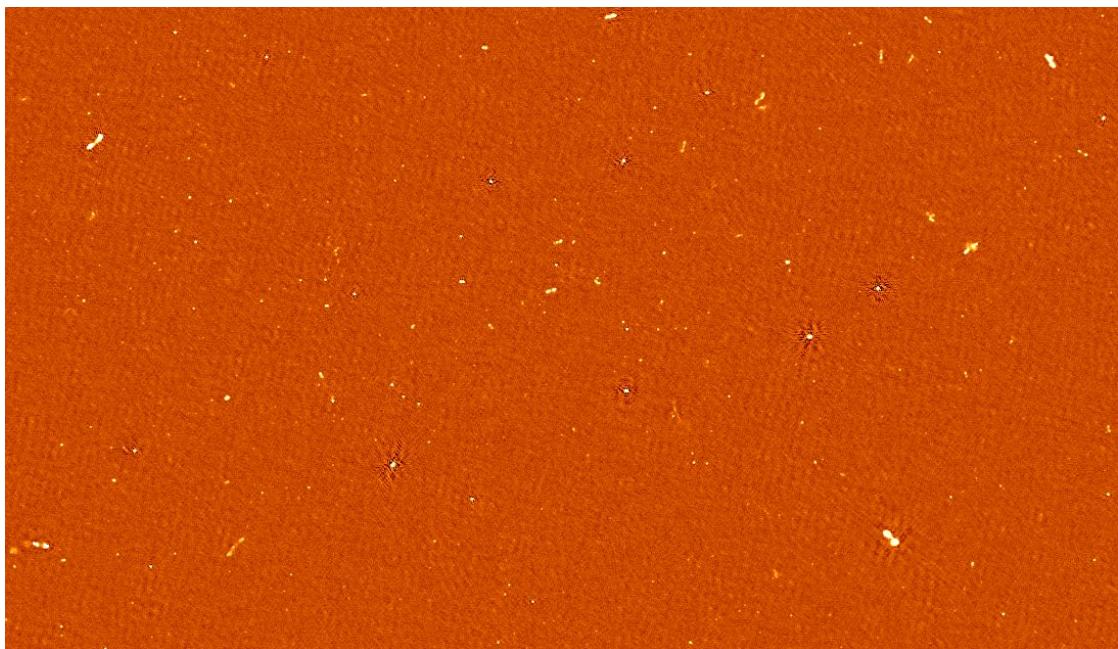


Figure 2.16. Voronoi facets, as implemented in DDFacet. Image is made before application of direction-dependent correction. This is an image of the Bootes field, made by Cyril Tasse.

The crucial insight of a faceting approach is then to split the image into a set of *facets*, solving for a set of calibration solutions for each of these facets. These solutions are then applied to the visibilities when mapping them onto the dirty image. The faceting in DDF is shown in Fig. 2.16. The facet distribution shown in Fig. 2.16 encourages an even flux distribution in multiple facets, and ensures that no facet is so small that too little signal would be available within. As we can see from Fig. 2.17, direction-dependent calibration can drastically improve the quality of wide-field images. It currently represents the state of the art in interferometric calibration.



(a) Image of the Boötes field, made without direction-dependent calibration. Image by Cyril Tasse.



(b) Image of the same Boötes field, made with direction-dependent calibration. Image by Cyril Tasse.

Figure 2.17. Difference between imaging made with (Fig. 2.17b) and without (Fig. 2.17a) third-generation calibration.

2.7 The RIME Formalism

In this section, we introduce the mathematical framework which forms the basis for our algorithmic work. While radio interferometry has historically been a complex business (Thompson et al. 2001), modern radio interferometry is based on the Radio Interferometer’s Measurement Equation, a powerful and elegant formalism which underpins modern calibration and imaging algorithms. In this section, I will give a simplified account of this formalism as described in Smirnov (2011a) and its companion papers¹⁷. This will form the theoretical basis on which further sections will expand. I cannot overstate the importance of the RIME as the foundation of modern calibration and imaging algorithms.

For a mathematically rigorous description of the RIME, the references used in this work (particularly Smirnov (2011c) and Smirnov (2011d)) are obviously the first place to look. We frame the RIME in the continuity of previous theoretical frameworks of radio interferometry, while also describing it in the context of contemporary software packages and algorithmic tools.

2.8 Setting up the RIME: a single point source

The Radio Interferometer’s Measurement Equation, or RIME, is a formalism which allows for an elegant and efficient formulation of the physical processes which affect signal propagation, from astrophysical effects to instrumental effects. Its fundamental underlying hypothesis is *linearity*: that transformations along the signal are linear with respect to the basis chosen to represent our signal. In other words, this means assuming that propagation effects are separable by antenna.

Consider a single quasi-monochromatic point source in the sky. Its signal at a point in time and space can be described by a complex 2-vector \mathbf{e} – this assumes plane waves. We can then represent physical processes which affect this signal’s propagation using *Jones matrices*. In other words:

$$\mathbf{v} = \mathbf{J} \mathbf{e} \quad (2.13)$$

where \mathbf{v} is the voltage measured by our antenna, and \mathbf{J} the Jones matrix describing the *net* propagation effects - atmospheric, astrophysical, and instrumental - which affect our source’s signal. \mathbf{J} can be written as a series of matrix products, each individual matrix describing a physical phenomenon. They are then called a *Jones chain*, and the resulting \mathbf{J} is often called the *total Jones matrix*.

A visibility is the correlation of voltage from two antennas. In other words,

$$\mathbf{V}_{pq} = 2\langle \mathbf{v}_p (\mathbf{v}_q)^H \rangle \quad (2.14)$$

$$\mathbf{V}_{pq} = 2\langle \mathbf{J}_p \mathbf{e} (\mathbf{J}_q \mathbf{e})^H \rangle \quad (2.15)$$

$$\mathbf{V}_{pq} = 2\langle \mathbf{J}_p \mathbf{e} \mathbf{e}^H \mathbf{J}_q^H \rangle \quad (2.16)$$

¹⁷Smirnov (2011b), Smirnov (2011c) and Smirnov (2011d)

where \mathbf{J}_p corresponds to the signal propagation chain of antenna p , and \mathbf{J}_q for antenna q . Note that a factor of 2 has been introduced here, as a matter of convention. This is due to the definition of the Stokes parameter in relation to our \mathbf{e} outer product.

For \mathbf{e} to be a 2-vector, rather than a 3-vector, we must choose to use xyz as our coordinates, with z the direction of propagation of our signal. Hamaker et al. (1996) then show that the following relation holds:

$$2\mathbf{e}\mathbf{e}^H = 2 \begin{bmatrix} \langle e_x e_x^* \rangle & \langle e_x e_y^* \rangle \\ \langle e_y e_x^* \rangle & \langle e_y e_y^* \rangle \end{bmatrix} = \begin{bmatrix} I + Q & U + iV \\ U - iV & I - Q \end{bmatrix} = \mathbf{B} \quad (2.17)$$

We can thus write:

$$\mathbf{V}_{pq} = \mathbf{J}_p \mathbf{B} \mathbf{J}_q^H \quad (2.18)$$

This gives us an elegant formulation of the relationship between the signal emitted by an astrophysical source and its corresponding measured visibility. The Stokes parameters¹⁸ of our source can be directly calculated from our measured visibility, provided \mathbf{J}_p and \mathbf{J}_q are known. Of course, in practice, they are not, and must therefore be modelled. To do this, we must expand each Jones matrix into a corresponding Jones chain.

2.9 Expanding our Jones Chain

Having written our RIME as a matrix multiplication problem, we can now begin to differentiate between different types of Jones matrices. \mathbf{J}_p describes the *aggregate* effects which occur over the course of the signal's propagation to our antenna p . It can be split into a Jones chain of n separate matrices, corresponding to different effects:

$$\mathbf{J}_{sp}^H = \prod_1^n \mathbf{J}_{nsp}^H = \mathbf{J}_{1sp}^H \mathbf{J}_{2sp}^H \dots \mathbf{J}_{nsp}^H \quad (2.19)$$

$$\mathbf{J}_{sp} = \mathbf{J}_{nsp} \dots \mathbf{J}_{2sp} \mathbf{J}_{1sp} \quad (2.20)$$

where s is a particular direction in the sky: these Jones matrices are therefore direction-dependent. Since the Jones terms are applied sequentially to the brightness matrix, starting with \mathbf{J}_{1sp} , the terms to the right of our decomposition of \mathbf{J}_{sp} can be said to occur 'at the source', i.e. before those to the left (said to occur 'at the antenna').

Our aim is thus to find the Jones chain which most concisely describes the different *types* of effects affecting our signal. In order of increasing complexity, there are three: so-called *geometric* effects, *direction-independent* effects, and *direction-dependent* effects.

2.9.1 Geometric Effects: the Phase Delay Jones Matrix

The physical quantity measured by an interferometric array, the *visibility*, is not a function of the absolute phase of our signal, but rather a function of the *phase difference*

¹⁸Stokes parameters determine the polarisation of light; we will not get into more details in this manuscript, as it is hardly relevant to the topic at hand. Suffice to say that the 2-vector formalism helps account for the polarised nature of light.

between our measured voltages \mathbf{v}_p and \mathbf{v}_q . In other words, it is a function of the phase difference on *baseline* pq , which consists of antennas p and q . The direction which minimises this difference is called the *phase centre*¹⁹.

We shall use the conventional coordinate system and notations (Smirnov 2011a; Thompson et al. 2001, etc). We thus have a z axis pointing towards the phase centre, and an antenna p at coordinates $\mathbf{u}_p = (u_p, v_p, w_p)$. The phase difference between p and phase centre (where, by definition, $\mathbf{u} = 0$) for a signal arriving from direction $\boldsymbol{\sigma}$ is then, assuming a small field of view:

$$k_p = 2\pi i((u_p l + v_p m + w_p(n - 1))) \quad (2.21)$$

where $n = \sqrt{1 - l^2 - m^2}$. Here, l, m, n are the direction cosines²⁰ of $\boldsymbol{\sigma}$, and \mathbf{u} is defined in units of wavelength²¹.

Having done this, we can now introduce a scalar *K-Jones* matrix to represent the effect of this phase delay. Since it is scalar, it will commute with all other Jones matrices (see Smirnov 2011a). It can be written as:

$$\mathbf{K}_p = e^{-ik_p} = e^{-2\pi i(u_p l + v_p m + w_p(n - 1))} \quad (2.22)$$

2.9.2 The Coherency Matrix

The commutation property of \mathbf{K}_p with other Jones matrices allow us to always place it at the rightmost end of our Jones chain. We can thus write our Jones chain as follows:

$$\mathbf{J}_p = \mathbf{G}_p \mathbf{K}_p \quad (2.23)$$

$$(2.24)$$

Here, we have decoupled the nominal phase delay term (\mathbf{K}_p) from all ‘corrupting’ effects (\mathbf{G}_p). We can take this one step further by defining the *source coherency*, \mathbf{X}_{pq} , as follows:

$$\mathbf{X}_{pq} = \mathbf{K}_p \mathbf{B}_{pq} \mathbf{K}_p^H \quad (2.25)$$

$$\mathbf{V}_{pq} = \mathbf{G}_p \mathbf{X}_{pq} \mathbf{G}_q^H \quad (2.26)$$

This is useful on a conceptual level: the process of *calibration* is now to find a best estimate for \mathbf{G}_p , which contain all non-‘geometric’ physical effects on the signal’s propagation from an astrophysical source to our voltage measurement.

On a practical level, using the coherency matrix \mathbf{X} allows us to write our RIME more concisely, by absorbing the geometric effects into the core of our RIME. While not strictly necessary, it can be a useful thing to do.

¹⁹Bear in mind the limitations of this, described in Section 2.2.2.

²⁰The $n - 1$ term in the exponential is due to the fact that, at phase centre, $k_p = 0$ for all \mathbf{u} and $n = \sqrt{1 - 0 - 0} = 1$.

²¹This means that, for a given baseline pq , \mathbf{u} will vary as a function of frequency.

2.10 Multiple Point Sources

So far, \mathbf{B} represents the signal from a single point source. However, the formalism can be extended to multiple point sources. This is because electric fields – and therefore radio signals – are additive, and that signal from different sources is not coherent (see Eq. (2.5)), which means that visibilities are additive. If our sky contains s point sources, then our RIME becomes:

$$\mathbf{V}_{pq} = \sum_s \mathbf{J}_{sp} \mathbf{B}_s \mathbf{J}_{sq}^H \quad (2.27)$$

We now expand our Jones chain as before. Some, but not all, elements in the chain will only depend on the antenna. These will be the same for all sources. We can thus write our Jones chain as:

$$\mathbf{J}_{sp} = \mathbf{G}_p \mathbf{E}_{sp} \mathbf{K}_{sp} \quad (2.28)$$

Note that we have kept the source-dependent effects (i.e. E_{sp} those occurring ‘at the source’) are to the right of our chain, and the antenna-dependent (G_p) effects to the left. K_{sp} is a scalar and can thus be moved at will; it is therefore placed at the far right to recover the coherency matrix if necessary.

Our multiple-source RIME can thus be written in the following form:

$$\mathbf{V}_{pq} = \mathbf{G}_p \left(\sum_s \mathbf{E}_{sp} \mathbf{K}_{sp} \mathbf{B}_s \mathbf{K}_{sq}^H \mathbf{E}_{sq}^H \right) \mathbf{G}_q^H \quad (2.29)$$

This ‘onion’ form of the RIME is usually sufficient to describe signal propagation in radio-interferometric arrays. We find the three types of Jones matrices promised earlier: the geometric effects accounted for by \mathbf{K} , the direction-dependent effects modeled into \mathbf{E} , and the direction-independent effects (antenna-based effects) modeled into \mathbf{G} .

There are still limitations to this RIME; notably, it assumes that the sky consists entirely of point sources (absence of diffuse emission), and that Jones matrices are time-independent. These are of course unphysical hypotheses, though they are a very good first approximation. We shall now address these two points.

2.11 Diffuse Emission: the Full-sky RIME

2.11.1 Deriving the Fully-Sky RIME

If we want to be able to take diffuse emission properly into account, we must go from our set of discrete point sources \mathbf{B}_s and integrate the underlying continuous brightness distribution $\mathbf{B}(\sigma)$ over all possible directions. This gives us the following expression:

$$\mathbf{V}_{pq} = \int_{4\pi} \mathbf{J}_p(\sigma) \mathbf{B}(\sigma) \mathbf{J}_q^H(\sigma) d\Omega \quad (2.30)$$

Here, effects such as our array's beam lie in the Jones matrices. By performing the analysis shown in Sect. 3.1 of Thompson et al. (2001), we can rewrite this expression in terms of a sine projection of the sphere onto the (l, m) plane tangential at field centre. The integral then becomes:

$$\mathbf{V}_{pq} = \int_{l,m} \mathbf{J}_p(\boldsymbol{\sigma}) \mathbf{B}(\boldsymbol{\sigma}) \mathbf{J}_q^H(\boldsymbol{\sigma}) \frac{dldm}{n} \quad (2.31)$$

Having written this, and using \mathbf{l} as shorthand (l, m) , we can once again decompose our Jones matrix into a Jones chain containing a direction-independent term \mathbf{G} , a direction-dependent term \mathbf{E} , and the phase term \mathbf{K} . Substituting this into our integral, we get the following expression:

$$\mathbf{V}_{pq} = \mathbf{G}_p \left(\int_{\mathbf{l}} \frac{1}{n} \mathbf{E}_p(\mathbf{l}) \mathbf{B} \mathbf{E}_q(\mathbf{l}) \mathbf{K}_p(\mathbf{l}) \mathbf{K}_q^H(\mathbf{l}) dldm \right) \mathbf{G}_q^H \quad (2.32)$$

$$= \mathbf{G}_p \left(\int_{\mathbf{l}} \frac{1}{n} \mathbf{E}_p(\mathbf{l}) \mathbf{B} \mathbf{E}_q(\mathbf{l}) e^{-2\pi i(u_{pq}l + v_{pq}m + w_{pq}(n-1))} dldm \right) \mathbf{G}_q^H \quad (2.33)$$

where

$$u_{pq} = u_p - u_q \quad v_{pq} = v_p - v_q \quad w_{pq} = w_p - w_q \quad (2.34)$$

This is a three-dimensional Fourier transform. It can be reduced to a two-dimensional Fourier transform by writing the n -dependence in the exponential (and normalisation) as a Jones matrix, thus removing the *non-coplanarity* terms from the explicit RIME. We can thus define the following:

$$\mathbf{W}_p = \frac{1}{\sqrt{n}} e^{-2\pi i w_p(n-1)} \quad (2.35)$$

$$B_{pq} = \mathbf{E}_p \mathbf{W}_p \mathbf{B} \mathbf{W}_q^H \mathbf{E}_q^h \quad (2.36)$$

and write our RIME in its final form:

$$\mathbf{V}_{pq} = \mathbf{G}_p \left(\int_{\mathbf{l}} \mathbf{B}_{pq}(\mathbf{l}) e^{-2\pi i(u_{pq}l + v_{pq}m + w_{pq}(n-1))} dldm \right) \mathbf{G}_q^H \quad (2.37)$$

Note that \mathbf{B} was the actual sky brightness distribution as a function of direction, but now \mathbf{B}_{pq} is the sky brightness *as seen by baseline pq* – in the presence of direction-dependent effects, this is not the same for all baselines.

2.11.2 Recovering the van Cittert-Zernike Theorem

Earlier calibration models and algorithms rely on the hypothesis that all baselines see the same sky. This is the *fundamental premise of self-calibration*. The topic of self-cal is covered in Section 2.6.1. This premise only holds when all Direction-Dependent Effects (henceforth referred to as DDEs) are identical for all baselines (and therefore for all

antennas). This is a necessary condition for the apparent sky \mathbf{B}_{pq} to be the same for all baselines²², as it allows us to write:

$$\mathbf{E}_p(\mathbf{l}) \mathbf{W}_p(\mathbf{l}) = \mathbf{E}(\mathbf{l}) \mathbf{W}(\mathbf{l}) \quad (2.38)$$

$$\mathbf{B}_{pq}(\mathbf{l}) = \mathbf{B}_{app}(\mathbf{l}) = \mathbf{E}(\mathbf{l}) \mathbf{W}(\mathbf{l}) \mathbf{B}(\mathbf{l}) \mathbf{W}^H(\mathbf{l}) \mathbf{E}^H(\mathbf{l}) \quad (2.39)$$

If this condition is met, we can then rewrite the full-sky RIME as:

$$\mathbf{V}_{pq} = \mathbf{G}_p \mathbf{X}_{pq} \mathbf{G}_q^H \quad (2.40)$$

where $\mathbf{X}_{pq} = \mathbf{X}(u_{pq}, v_{pq}) = \mathbf{X}(\mathbf{u})$. Comparing Eqs. (2.25) and (2.37) directly shows that the coherency matrix is simply the two-dimensional Fourier transform of $\mathbf{B}_{app}(\mathbf{l})$. We can thus call $\mathbf{X}(\mathbf{u})$ the *sky coherency*, in keeping with our nomenclature for the RIME of a point source.

Note that deriving the van Cittert-Zernike theorem ((van Cittert 1934), covered in (Thompson et al. 2001)) from the RIME was simply of matter of treating phase as a Jones matrix in its own right²³. A significant consequence of this is that DDEs can be incorporated within the RIME formalism.

2.12 Time-dependent Jones matrices, Smearing, and Decorrelation

2.12.1 Time-dependence of Jones matrices

Until this point, we have not considered any kind of sky variability in time. In effect, Eq. (2.40) describes the snapshot taken by an interferometer for a single measurement. Limiting ourselves to this scenario, however, strongly limits the sampling of the *uv*-plane which can be performed. Indeed, interferometers generally rely on the Earth’s rotation to rotate their baselines pq , thus increasing *uv*-coverage at no additional cost. This technique is known as *super-synthesis* (Joardar et al. (2010)²⁴) For Eq. (2.40) to hold throughout an observation, we must therefore assume that B_{app} remains constant in time over the course of the observation. In effect, this means assuming that

$$\mathbf{E}_p(t, \mathbf{l}) = \mathbf{E}_p(\mathbf{l}) = \mathbf{E}(\mathbf{l}) \quad \forall (t, p) \quad (2.41)$$

The above means that direction-dependent effects must be time-independent for Eq. (2.40) to hold. We can thus split direction-dependent effects into two categories: *trivial* DDEs, which satisfy this condition and are a simple multiplicative correction, and *non-trivial* DDEs, which are not. An example of a trivial DDE is the primary gain; and example of a non-trivial (i.e. time-variable) DDE is ionospheric turbulence.

²² \mathbf{B}_{pq} would then correspond to the ‘true’ sky attenuated by the “power beam”, which corresponds to our instrument’s directional sensitivity in the traditional view

²³Noordam, J.E. 1996, AIPS++, Note 185: The Measurement Equation of a Generic Radio Telescope

²⁴<http://www.jpier.org/PIERB/pierb22/05.10032105.pdf>

2.12.2 Smearing and Decorrelation

Of course, our Jones matrices are not the only time-dependent element of the RIME formalism. As mentioned above, to image with interferometric arrays, radio astronomers rely on the Earth's rotation to better sample *uv*-space. This means that \mathbf{u} becomes a *time-dependent* variable.

Furthermore, if we choose to define \mathbf{u} in terms of wavelength, \mathbf{u}_p becomes dependent in frequency. Our time-independent, frequency-independent RIME thus becomes only valid for infinitesimal increments in time and frequency.

In practice, measurements made by any interferometric array are actually averaged over the integration time of the measurement, as well as the frequency channel size. A better formulation is therefore to explicitly write the RIME as an integration over time and frequency bins:

$$\langle \mathbf{V}_{pw} \rangle = \frac{1}{\Delta t \Delta \nu} \int_{t_0}^{t_1} \int_{\nu_0}^{\nu_1} \mathbf{V}_{pq}(t, \nu) d\nu dt \quad (2.42)$$

$$= \frac{1}{\Delta t \Delta \nu} \int_{t_0}^{t_1} \int_{\nu_0}^{\nu_1} \mathbf{J}_p(t, \nu) \mathbf{B} \mathbf{J}_q^H(t, \nu) \quad (2.43)$$

$\mathbf{J}(t, \nu)$ describes the propagation of an electromagnetic signal through various physical processes. This signal has a complex phase, which is variable in both time and frequency. Since it is a rotating vector, integrating over any time or frequency band will always result in a *net loss* in amplitude in the measured $\langle \mathbf{V} \rangle$. This mechanism is commonly referred to as *time/bandwidth width decorrelation* or *smearing* (Smirnov 2011a).

The general effect is referred to as *decorrelation*²⁵, and the specific case of decoherence caused by the \mathbf{K} term is referred to as *smearing*.

Smearing increases with baseline length (\mathbf{u}_{pq}) and distance from phase centre ((l, m)). The noise amplitude, however, does not decrease. This means that smearing results in a decrease in sensitivity, which becomes significant when attempting to do wide-field, high-resolution images (see Section 5.6).

In the context of the RIME, smearing is an element-by-element operation, i.e. different elements do not affect each other. Treating smearing within the RIME is thus a trivial expansion of the scalar equations.

Assuming that Δt and $\Delta \nu$ are small enough that the amplitude of \mathbf{V}_{pq} remains constant over the integration band (i.e. that \mathbf{V}_{pq} is well-sampled by our instrument), then a useful first-order approximation is:

$$\langle \mathbf{V}_{pw} \rangle \simeq \text{sinc} \frac{\Delta \Psi}{2} \text{ sinc} \frac{\Delta \Phi}{2} \mathbf{V}_{pq}(t_{mid}, \nu_{mid}) \quad (2.44)$$

²⁵Decorrelation can equivalently be referred to as “decoherence”.

where

$$t_{mid} = (t_0 + t_1)/2 \quad (2.45)$$

$$\nu_{mid} = (\nu_0 + \nu_1)/2 \quad (2.46)$$

$$\Delta\Psi = \arg \mathbf{V}_{pq}(t_1, \nu_{mid}) - \arg \mathbf{V}_{pq}(t_0, \nu_{mid}) \quad (2.47)$$

$$\Delta\Phi = \arg \mathbf{V}_{pq}(t_{mid}, \nu_1) - \arg \mathbf{V}_{pq}(t_{mid}, \nu_0) \quad (2.48)$$

This equation is convenient to determine the impact of decorrelation for a source at a given distance from phase centre. Note, however, that it can only - by definition - be valid for a single source at a time. If using a RIME such as Eq. (2.29), then it can be calculated for each source individually.

2.13 Noise

The RIME presented thus far has been for an ideal case, free of noise in our measured visibility. In practice, each complex visibility will be affected by uncorrelated Gaussian noise²⁶, which means that a RIME of the same form as Eq. (2.29) would have to be formulated as follows:

$$\mathbf{V}_{pq} = \mathbf{G}_p \left(\sum_s \mathbf{E}_{sp} \mathbf{K}_{sp} \mathbf{B}_s \mathbf{K}_{sq}^H \mathbf{E}_{sq}^H \right) \mathbf{G}_q^H + \sigma_{pq} \quad (2.49)$$

where σ_{pq} is a (2, 2) matrix with real and complex parts in each component. A similar addition can be made for any of the RIMEs presented here.

This noise sets a hard limit on the sensitivity achieved for a given observation, and ‘reaching the noise’ becomes the gold standard of calibration (Smirnov 2011b). One must then take σ_{pq} into account properly when solving for the Jones matrices.

²⁶Detailed treatment given in (Thompson et al. 2001), section 6.2, which shows how to obtain the ideal radiometer equation

Chapter 3

Analysing the Variance of Gain Solutions

1

Contents

3.1	Introduction	53
3.2	Building the Noise Map	56
3.2.1	The Cov-Cov Relationship in the $\delta u \delta v$ plane	56
3.2.2	Statistical Analysis	59
3.2.3	Noise Map Simulations	62
3.3	Adaptive Quality-based Weighting Schemes	67
3.3.1	Optimising sensitivity	67
3.3.2	Minimising Calibration Artefacts	68
3.4	Estimating the Covariance Matrix	68
3.4.1	Baseline-based Estimation	69
3.4.2	Antenna-based Estimation	70
3.5	Applying the Correction to Simulated Data	72
3.6	Applying the Correction to Real Data	73
3.7	Discussion	76

3.1 Introduction

Interferometers sample Fourier modes of the sky brightness distribution corrupted by instrumental and atmospheric effects rather than measuring the sky brightness directly. This introduces two problems² for astronomers to invert: calibration and imaging. Both of these problems are ill-conditioned.

¹This section is based on Bonnassieux et al. (2017), which was published as part of this PhD.

²Which are simply an artificial split of a single original problem; see Section 2.3 for more details.

The problem of imaging consists of correcting for the incomplete *uv*-coverage of any given interferometer by deconvolving the instrument’s Point-Spread Function (PSF) from images. Its poor conditioning comes from our limited a priori knowledge of the sky brightness distribution, combined with large gaps in our *uv*-coverage, which prevents us from placing strong constraints on image deconvolution. It can be better-conditioned in different ways, including through the use of weighting schemes (see Briggs 1995; Yatawatta 2014, and references therein) to improve image fidelity at the start of deconvolution. When inverting the imaging problem, we often assume that the sky is stable within the domain (i.e. is constant in time and frequency). There are exceptions, such as wide-band deconvolution algorithms (e.g. Rau & Cornwell 2011) that explicitly take into account the sky’s frequency-dependence, but still assume that the sky brightness distribution does not vary with time.

The problem of calibration is what concerns us in this section. It consists of estimating and correcting for instrumental errors (which includes effects such as antenna pointing errors, but also the phase-delays caused by ionospheric activity, troposphere, etc). Calibration consists of solving for gain estimates, where a gain models the relationship between the electromagnetic field of an astrophysical source and the voltage that an antenna measures for this source. Because measurements are noisy, calibration often involves some fine-tuning of solution intervals, to ensure that the solutions are well-constrained while the solution intervals stay as small as signal-to-noise allows. The calibration inverse problem involves three competing statistical effects: thermal noise in the measurements, true gain variability, and sky model incompleteness. If gain solutions are sought individually for each measurement, then (unless the observed field is *exceptionally* bright) calibration estimates will be dominated by thermal noise, and will not adequately describe the actual gains. Similarly, if a single gain estimate is fitted to too many measurements, the intrinsic gain variability will be “averaged out”; for example, a choice of time and frequency interval that is too large will cause the solver to estimate a constant gain while the underlying function varies quickly, thereby missing much of the gain structure. This will introduce error which will be correlated in time and frequency. This occurs, for example, when solving for ionospheric phase delays: in the most extreme case, where the solution interval is significantly larger than the scale of ionospheric fluctuations, its varying phase can average out to zero over the interval in time and frequency. Finally, if the model being fitted is incomplete, unmodeled physical flux will likely be absorbed unpredictably into both the gain solutions and the residual visibilities: this absorption of physical flux into gain solutions is known as source suppression (see Grobler et al. 2014; Kazemi & Yatawatta 2013, and references therein).

In practice, it is reasonable to assume that there exist a scale at which gain variation is acceptably low: we can then reduce the noise of our gain estimates by finding a single gain solution for a small number of measurements, assuming that the underlying gain variation is very small and stable over short intervals. This is *generally* a valid hypothesis, but the specific value for the variation scale can be contentious. Indeed, while the noise level can often be treated as constant throughout an observation, the gain variability itself is generally not constant: there will be time periods where the gains will tend to remain constant for longer, and others where variability will be very quick. This means that, for any choice of calibration interval, some gain estimates will

be better than others, and almost all could have been improved (at a cost to others) by a different choice of time (and frequency) intervals.

Since we have measurements which are better-calibrated than others (in that better estimates for their gains were obtained through chance alone), we could, in principle, take inspiration from “lucky imaging” (an optical-domain method for making good images: for more details, see Fried 1978, and references therein) to weigh our visibilities according to their calibration quality. Those weights would in effect be an improvement of currently existing methods such as clipping noisy residual visibilities: in the extreme case where all visibilities are equally-well calibrated except a few which are extremely noisy, it should be equivalent to clipping. Otherwise, the weights should show at least a slight improvement over clipping.

The key finding of our algorithmic work is a fundamental relationship between the covariance matrices of residual visibilities and the map of the covariance in the image-plane: the “Cov-Cov relationship” between visibility covariance to image-plane covariance. We show that the pixel statistics in the image-plane are determined by a “noise-PSF”, convolved with each source in the sky (modeled or not). This noise-PSF is the product of the Fourier transforms of the gain covariance matrix with each cell mapped not from uv space to lm coordinates but rather between their respective covariance spaces - from a new differential Fourier plane (henceforth “ $(\delta u \delta v)$ ”-plane) to the image-plane covariance space $\delta l \delta m$. This image-plane covariance space describes the variance in each pixel and the covariance between pixels³. It describes the expected calibration artefacts and thermal noise around each source, does not vary as a function of direction, and is convolved with each source in the field to yield the final error map. Because all unwanted (in our case, unphysical) signal can be thought of as noise, we will refer to the pixel variance map as the “noise-map”.

The notion of a $(\delta u \delta v)$ -plane arises organically from the framework of radio interferometry: we are associating a *correlation* between visibilities to *coordinates* in covariance space, just as we associate the visibilities themselves to the uv -domain. The $(\delta u \delta v)$ -plane is the natural domain of these correlations. As previously stated, even if all sources in the field are perfectly known and modeled, a poor choice of calibration interval can introduce correlated noise in the residuals, which would then introduce larger variance near sources in the noise-map. Conversely, if calibration is perfect, the noise-map should be completely flat (i.e. same variance for all pixels), as there would be no noise-correlation between pixels.

The main result of this section is a new adaptive, quality-based weighting scheme based on this insight. Using the Cov-Cov relationship, we can create a new weighting scheme by estimating the residual visibility covariance matrix in a given observation. By weighting visibilities so as to change their covariance matrix, one can change the shape of the noise-PSF and thus improve the final image: this manifests as either decreased noise or decreased calibration artefacts. Note that this weighting is applied after calibration,

³The noise-PSF also relates δw to δn , as shown in the matrix formalism, but this is not explicitly referenced in the text since visibility space is usually referred to as “the UV-plane” in literature, rather than “the UVW-space”.

but before image deconvolution: applying it will therefore not only improve the residual noise in the image, and thus the sensitivity achievable with a given pipeline, but will also improve deconvolution by minimising calibration artefacts in the field: it should thus effectively remove spurious, unphysical emission from final data products. Estimating the covariance matrix is the main difficulty of our framework: we do not know the underlying covariance matrix, and the conditioning of our estimation thereof is limited by the number of measurements within each solution interval. As such, we have no guarantee that our estimate of the corrected visibility covariance matrix is accurate. This problem can be alleviated, for example by estimating the covariance matrix for the antenna gains themselves, and use it to build the visibility covariance matrix: this effectively improves conditioning (cf Sec. 3.4).

The remainder of this chapter is split into four main sections. In Section 3.2, we derive the Cov-Cov relationship. With its newfound insights, we propose quality-based weighting schemes with which to improve radio interferometric images in Section 3.3. We follow in Section 3.4 by showing how to estimate, from real data, the covariance matrix from which the quality-based weights are derived. Our approach seems to give good results. Finally, we close the chapter on a discussion of the applicability and limitations of the quality-based weighting scheme.

3.2 Building the Noise Map

In this section, we derive our first fundamental result: the Cov-Cov relationship, Eq. 3.25, which describes how the statistics of residual visibilities (and thus the antenna calibration solutions, henceforth “gains”) relate to the statistics of the image plane, i.e. of images made using the associated visibilities. The dimensions of the matrices (denoted by boldface capital letters) and vectors (denoted by boldface lowercase letters) used in this work are given in Table 3.1, along with the scalar numbers used to denote specific dimensions. All other variables are scalars.

3.2.1 The Cov-Cov Relationship in the $\delta u \delta v$ plane

Let us begin by defining visibility gains. Using the Radio Interferometry Measurement Equation formalism for a sky consisting of a single point source (Hamaker et al. (1996), Smirnov (2011a), and companion papers), we can write the following relation between the sky and the signal as measured by a single baseline at time t and frequency ν :

$$\mathbf{V}_{pq}^{t\nu} = \sum_d \mathbf{K}_{p,t\nu}^d \mathbf{J}_{p,t\nu}^d \mathbf{B}_\nu^d (\mathbf{J}_{q,t\nu}^d)^H (\mathbf{K}_{q,t\nu}^d)^H + \mathbf{N} \quad (3.1)$$

All the quantities above are 2×2 matrices. Eq. 3.1, implies a linear relationship between the coherency matrix \mathbf{B}_ν^d and the visibilities recorded by a given baseline ($\mathbf{V}_{pq}^{t\nu}$), with the addition of a thermal noise matrix \mathbf{N} , which is also of shape 2×2 and contains different complex-valued realisations of the noise in each cell. Since electric fields are additive, the sky coherency matrix can be described as the sum of the contributions from individual sources in directions d in the sky. We also assume that the sky does not vary over time, i.e. that \mathbf{B}_ν^d is not a function of time. The Jones matrices ($\mathbf{J}_{...,t\nu}^d$) contain the antenna gain information in matrix form, while $\mathbf{K}_{...,t\nu}^d$ is the Fourier kernel. Let us limit

Scalars

n_{pix}	Total number of pixels in image-plane & number of cells in uv -grid
n_{ant}	Total number of antennas in the array
n_b	Number of visibilities
b	Index for a single visibility.
	Equivalent to $(pq, t\nu)$
τ	Equivalent to (t, ν)

Vectors

$\tilde{\mathbf{y}}$	Residual image vector, size n_{pix}
ϵ	Vector of ϵ , size n_{pix}
$\tilde{\boldsymbol{\gamma}}$	Contains gain products, size n_b . See Eq. 3.12
$\mathbf{1}$	Vector containing 1 in every cell, size n_b
$\delta\mathbf{u}_{bb'}$	Vector of coordinates in differential Fourier plane, of length 3.
\mathbf{l}_d	Vector of sky coordinates, of length 3.

Matrices

$\mathbf{V}_{pq}^{t\nu}$	Visibility seen by a baseline pq at time and frequency t, ν . Size 2×2 .
$\mathbf{K}_{p,t\nu}^d$	Fourier kernel for direction d , antenna p and one (t, ν) pair. Size 2×2 .
$\mathbf{J}_{p,t\nu}$	Jones matrix for antenna p for one (t, ν) pair.
	Contains the gains information. Size 2×2
\mathbf{B}	Sky brightness distribution matrix, of size 2×2 .
\mathbf{N}	Noise matrix, of size 2×2 . Contains a single realisation n of the thermal noise in each cell.
\mathcal{F}	Fourier transform matrix, of size $n_{\text{pix}} \times n_{\text{pix}}$.
\mathcal{S}_b	Baseline selection matrix, which picks out 1 visibility out of the full set. Size $n_{\text{pix}} \times n_b$
\mathcal{C}_b	$n_{\text{pix}} \times n_{\text{pix}}$ convolution kernel that defines the PSF.
$\mathcal{F}_{bb'}$	Convolution matrix mapping one $\delta u \delta v$ to $\delta l \delta m$. The set of all $\mathcal{F}_{bb'}$ determines the noise-PSF. Size $n_{\text{pix}} \times n_{\text{pix}}$.

Table 3.1. Table recapitulating the meaning and dimensions of vectors and matrices used in Sec. 3.2.1. Only scalars which give matrix dimensions or indices are given here.

ourselves to the scalar case, which corresponds to assuming that emission is unpolarised. We assume that $\mathbf{B}^d = s_d \mathbf{I}$, where s is the flux of our single point source and \mathbf{I} is the 2×2 identity matrix. We also assume that $\mathbf{J}_{p,t\nu} = g_{p,t\nu} \mathbf{I}$, where $g_p^{t\nu}$ is the complex-valued gains of antenna p at time t and frequency ν . This means that we assume that the gains are direction-independent, and so $\mathbf{J}_{...,t\nu}^d$ becomes $\mathbf{J}_{...,t\nu}$. Similarly, $\mathbf{K}_{p,t\nu}^d = k_{p,t\nu}^d \mathbf{I}$,

the Fourier kernel in the direction of the source, d . \mathbf{N} has 1 realisation of ϵ in each cell, where:

$$\epsilon \sim \mathcal{N}(0, \sigma) + i\mathcal{N}(0, \sigma) \quad (3.2)$$

where σ is the variance of the thermal noise. Let us denote each (t, ν) pair by τ , and ignore the sky's frequency-dependence. The following scalar formulation is then equivalent to Eq. 3.1:

$$V_{pq}^\tau = \left(\sum_d s_d k_{p,\tau}^d \overline{k_{q,\tau}^d} \right) g_p^\tau \overline{g_q^\tau} + \epsilon \quad (3.3)$$

$$k_{p,\tau}^d = \exp(2\pi i (u_{p,\tau} l_d + v_{p,\tau} m_d + w_{p,\tau} (n_p - 1))) \quad (3.4)$$

Calibration is the process of finding an accurate estimate of g_p^τ for all antennas p , at all times t and frequencies ν . Since we are in a direction-independent regime, the quality of our calibration then determines the statistical properties of the residual visibilities (and the image-plane equivalent, the residual image). The residual visibilities associated with calibration solutions are defined as our measured visibilities minus the gain-corrupted model visibilities. \hat{g}_p^τ then denotes our calibration estimate for g_p^τ . We now begin to limit the generality of our framework by assuming that all sufficiently bright sources have been modeled and subtracted: unmodeled flux is then negligible. We can then write the residual visibilities as:

$$\tilde{r}_{pq}^\tau = \sum_d s_d \left(k_{p,\tau}^d \overline{k_{q,\tau}^d} \right) \left(g_p^\tau \overline{g_q^\tau} - \hat{g}_p^\tau \overline{\hat{g}_q^\tau} \right) + \epsilon \quad (3.5)$$

The flux values in the image-plane pixels⁴ are the Fourier transform of the visibility values mapped onto each pixel. This can be written as follows:

$$\tilde{\mathbf{y}} = \begin{pmatrix} \vdots \\ \sum_{pq} I_{pq,lm}^r \\ \vdots \end{pmatrix} \quad (3.6)$$

$$I_{pq,lm}^r = \sum_\tau \omega_{pq,\tau} \tilde{r}_{pq}^\tau k_{pq,\tau}^{lm} \quad (3.7)$$

where lm are the directional cosine positions of a given pixel, and $k_{pq,\tau}^{lm} = k_{p,\tau}^d \overline{k_{q,\tau}^d}$ the Fourier coefficient mapping a point in Fourier space to a point on the image-plane. $\omega_{pq,\tau}$ is the weight associated to a given visibility.

Let us now write this using a matrix formalism. The contribution of a single visibility $b = (pq, \tau)$ to the image-plane residuals can be written as:

$$\tilde{\mathbf{y}}_b = \mathcal{F}^H \mathcal{S}_b \omega_b (\kappa_b \tilde{\mathbf{y}} + \epsilon) \quad (3.8)$$

$$\tilde{\mathbf{y}} = \sum_b \tilde{\mathbf{y}}_b \quad (3.9)$$

⁴As opposed to the Fourier-plane pixels, which are the elements of the grid onto which the measured visibilities are mapped for imaging.

where ϵ is a vector of the ϵ of Eq. 3.2 and $\tilde{\mathbf{y}}$ is a vector of size n_{pix} , with

$$\kappa_b = \sum_d s_d \kappa_b^d \quad (3.10)$$

$$\kappa_b^d = k_{p,\tau}^d \overline{k_{q,\tau}^d} = k_{pq,\tau}^{lm} \quad (3.11)$$

$$\tilde{\gamma}_b = g_p^\tau \overline{g_q^\tau} - \widehat{g}_p^\tau \overline{\widehat{g}_p^\tau} \quad (3.12)$$

$$\tilde{\gamma} = \begin{pmatrix} \vdots \\ \tilde{\gamma}_b \\ \vdots \end{pmatrix} \quad (3.13)$$

and w_b is the scalar weight associated with each visibility. By default, $w_b = \frac{1}{n_b}$: all visibilities then have the same weight, and $\tilde{\mathbf{y}}$ then becomes the average of all $\tilde{\mathbf{y}}_b$. $\tilde{\gamma}$ is a vector of all $\tilde{\gamma}_b$, and thus of size n_b . \mathcal{F} is the Fourier kernel, of size $n_{\text{pix}} \times n_{\text{pix}}$. \mathbf{S}_b is a matrix of size $n_{\text{pix}} \times n_b$: its purpose is to encode the uv -coverage. Each \mathbf{S}_b contains only a single non-zero cell, different for different \mathbf{S}_b . The height (number of rows) of \mathbf{S}_b is determined by the size of the uv -grid, and its length (number of columns) by the number of visibilities.

The order of operations is thus: each residual visibility ($\kappa_b \tilde{\gamma} + \mathbf{n}$) is assigned some weight w_b and its uv -coordinates are set by \mathbf{S}_b . The inverse Fourier transform (\mathcal{F}^H) is then applied to this grid, and so we recover its image-plane fringe. By averaging over all fringes, we recover the dirty image.

The residual image will thus depend on three quantities: the residual gains, the flux in the image, and the weighting scheme. Let us consider the relationship between the statistics of residual visibilities and the variance at a given point in the corresponding residual image.

3.2.2 Statistical Analysis

In the following analysis, we treat our gain solutions and thermal noise as random variables in order to compute the covariance matrix of our residual image, $\text{Cov}\{\tilde{\mathbf{y}}\}$. The diagonal of this matrix gives the variance for each pixel, while the wings give the covariance between pixels. Using the property that $\text{Cov}\{\mathbf{Ax}\} = \mathbf{A}\text{Cov}\{\mathbf{x}\}\mathbf{A}^H$, we can apply the $\text{Cov}\{\cdot\}$ operator to Eq. 3.9 to write:

$$\text{Cov}\{\tilde{\mathbf{y}}\} = \sum_{bb'} \mathcal{F}^H \underbrace{w_b w_{b'} \kappa_b \overline{\kappa_{b'}}}_{\stackrel{\text{def}}{=} \phi_{bb'}} \mathbf{S}_b \text{Cov}\{\tilde{\gamma}\} \mathbf{S}_{b'}^T \mathcal{F} \quad (3.14)$$

$$+ \sum_b w_b^2 \sigma^2 \underbrace{\mathcal{F}^H \mathbf{S}_b \mathbf{I} \mathbf{S}_b^T \mathcal{F}}_{\stackrel{\text{def}}{=} \mathcal{C}_b} \quad (3.15)$$

$$= \sum_{bb'} \phi_{bb'} \mathcal{F}^H \mathbf{S}_b \text{Cov}\{\tilde{\gamma}\} \mathbf{S}_{b'}^T \mathcal{F} + \sum_b w_b^2 \sigma^2 \mathcal{C}_b \quad (3.16)$$

So far, we have only applied definitions. The net effect of $\mathbf{S}_b \text{Cov}\{\tilde{\gamma}\} \mathbf{S}_{b'}^T$ (dimensions of $n_{\text{pix}} \times n_{\text{pix}}$) is to encode where a given baseline samples the uv -plane, and map *one* cell at matrix coordinates (b, b') from the correlation matrix $\text{Cov}\{\tilde{\gamma}\}$ onto the visibility grid.

\mathbf{S}_b is not the gridding kernel, but rather the sampling matrix, which determines where we have measurements and where we do not. We can thus write that $\mathbf{S}_b \text{Cov}\{\tilde{\gamma}\} \mathbf{S}_{b'}^T = [\text{Cov}\{\tilde{\gamma}\}]_{bb'} \mathbf{S}_b \mathbf{1} \mathbf{1}^T \mathbf{S}_{b'}^T$, where $[\text{Cov}\{\tilde{\gamma}\}]_{bb'}$ is the value from the appropriate cell and $\mathbf{1}$ is the vector-of-ones of appropriate length (here, n_b). This allows us to write:

$$\text{Cov}\{\tilde{\mathbf{y}}\} = \sum_{bb'} \phi_{bb'} [\text{Cov}\{\tilde{\gamma}\}]_{bb'} \mathcal{F}_{bb'} + \sum_b w_b^2 \sigma^2 \mathcal{C}_b \quad (3.17)$$

$$\text{with } \mathcal{F}_{bb'} = \underbrace{(\mathcal{F}_b)^H \mathcal{F}_{b'}}_{\stackrel{\text{def}}{=} \mathbf{1}^T \mathbf{S}_{b'}^T \mathcal{F}} \quad (3.18)$$

Here, \mathcal{C}_b is a Toeplitz matrix, i.e. a convolution matrix, associated with baseline b . The set of all \mathcal{C}_b defines the convolution kernel which characterises the Point-Spread Function (henceforth PSF) associated with a given uv-coverage, of size $n_{\text{pix}} \times n_{\text{pix}}$. $\mathcal{F}_{bb'}$, meanwhile, is not generally Toeplitz. Its cells can be written as:

$$\mathcal{F}_{bb'}[d, d'] = e^{2i\pi(u_b l_d - u_{b'} l_{d'} + v_b m_d - v_{b'} m_{d'} + (n_d - 1)w_b - (n_{d'} - 1)w_{b'})} \quad (3.19)$$

Let us investigate how the sky brightness distribution (i.e. d -dependence) affects the noise-map. We can write the sum over bb' as two sums: one over $b, b' = b$ and one over $b, b' \neq b$. Thus:

$$\begin{aligned} \text{Cov}\{\tilde{\mathbf{y}}\} &= \sum_b (\phi_{bb} [\text{Cov}\{\tilde{\gamma}\}]_{bb} + w_b^2 \sigma^2) \mathcal{C}_b \\ &\quad + \sum_{b, b' \neq b} \phi_{bb'} [\text{Cov}\{\tilde{\gamma}\}]_{bb'} \mathcal{F}_{bb'} \end{aligned} \quad (3.20)$$

Note that the only direction-dependent terms in the above are s_d and κ_b^d , which are both inside $\phi_{bb'}$ (for both $b = b'$ and $b \neq b'$). By making the approximation that the Fourier kernels of different sources are orthogonal (i.e. that $\kappa_b^d \kappa_{b'}^{d'} = (\kappa_b^d)^2 \delta_{bb'}$)⁵ we can write:

$$\phi_{bb'} = w_b w_{b'} \kappa_b \overline{\kappa_{b'}} \quad (3.21)$$

$$= w_b w_{b'} \left(\sum_d s_d \kappa_b^d \right) \left(\sum_{d'} s_{d'} \overline{\kappa_{b'}^{d'}} \right) \quad (3.22)$$

$$\approx \sum_d w_b w_{b'} s_d^2 \kappa_b^d \overline{\kappa_{b'}^d} \quad (3.23)$$

$$\phi_{bb'} \approx \sum_d \phi_{bb}^d \quad (3.24)$$

Note that $\mathcal{F}_{bb} = \mathcal{C}_b$, since those are the coordinates along the diagonal: for these values, the matrix-of-ones at the centre of $\mathcal{F}_{bb'}$ becomes the identity matrix. Note also that $\phi_{bb}^d = w_b^2 s_d^2$, since $\kappa_b^d \overline{\kappa_b^d} = 1$. We can then write Eq. 3.20 as:

$$\begin{aligned} \text{Cov}\{\tilde{\mathbf{y}}\} &= \sum_d \left(\sum_b \phi_{bb}^d \left([\text{Cov}\{\tilde{\gamma}\}]_{bb} + \frac{w_b^2 \sigma^2}{\phi_{bb}} \right) \mathcal{C}_b \right. \\ &\quad \left. + \sum_{b, b' \neq b} \phi_{bb'}^d [\text{Cov}\{\tilde{\gamma}\}]_{bb'} \mathcal{F}_{bb'} \right) \end{aligned} \quad (3.25)$$

where we have now limited our formalism to the case where the sky is dominated by sufficiently-separated point-like sources.

⁵This hypothesis is equivalent to assuming that the sky is dominated by well-separated point sources, where we take “well-separated” to mean that the sources are multiple PSF Full-Width Half-Maximum apart

This is our fundamental result: assuming unpolarised emission coming from well-separated point sources and normally-distributed thermal noise, it gives a direct relationship between the covariance of the residual visibilities and the covariance of the residual image-pixel values. We thus call it the Cov-Cov relationship. It describes the statistical properties of the image-plane as the result of a convolution process changing an average noise level at different points in the image-plane, allowing us to describe the behaviour of variance and covariance in the image. Let us focus on the first

By applying the $\text{Diag}\{\}$ operator (which returns the diagonal of an input matrix as a vector) to both sides of Eq.3.25, we can find an expression for the variance map in the image-plane:

$$\text{Var}\{\tilde{\mathbf{y}}\} = \text{Diag}\{\text{Cov}\{\tilde{\mathbf{y}}\}\} \quad (3.26)$$

$$\begin{aligned} &= \sum_d \left(\sum_b (\phi_{bb}^d [\text{Cov}\{\tilde{\gamma}\}]_{bb} + w_b^2 \sigma^2) \underbrace{\text{Diag}\{\mathcal{C}_b\}}_{=\mathbf{1}} \right. \\ &\quad \left. + \sum_{b,b' \neq b} \phi_{bb'}^d [\text{Cov}\{\tilde{\gamma}\}]_{bb'} \text{Diag}\{\mathcal{F}_{bb'}\} \right) \end{aligned} \quad (3.27)$$

$$\text{where } \mathbf{l}_d = (l_d, m_d, (n_d - 1)) \quad (3.28)$$

$$\boldsymbol{\delta u}_{bb'} = (\delta u_{bb'}, \delta v_{bb'}, \delta w_{bb'}) \quad (3.29)$$

$$= (u_b - u_{b'}, v_b - v_{b'}, w_b - w_{b'}) \quad (3.30)$$

In Eq 3.27, we have:

$$\text{Diag}\{\mathcal{F}_{bb'}\}[d] = e^{2i\pi \mathbf{l}_d \cdot \boldsymbol{\delta u}_{bb'}} \quad (3.31)$$

For $b \neq b'$, the diagonals of $\mathcal{F}_{bb'}$ are the Fourier kernels mapping $\delta u \delta v$ space to $\delta l \delta m$. $\mathcal{F}_{bb'}$ can then be thought of as a Fourier transform. It is not a diagonal matrix. It behaves as a *covariance fringe*, allowing us to extend standard interferometric ideas to covariance space: each fringe can be thought of as a single “spatial filter” applied to the pixel covariance matrix. Just as a given baseline has coordinates in *uv*-space, a given *correlation between baseline residual errors* has coordinates in *uv correlation space*, which we will henceforth refer to as $\delta u \delta v$ -space.

This $\delta u \delta v$ space warrants further discussion: Fig. 3.1 shows, for a given *uv*-track (Fig. 3.1b), both the corresponding $\delta u \delta v$ domain (Fig. 3.1c) and point-spread function (Fig. 3.1a). The symmetric, negative *uv*-track is treated as a separate track, and thus ignored. This means that we do not fully constrain the noise-PSF (since the covariance matrix of the symmetric track is simply the Hermitian of the first), but we do not seek to constrain it in this section, but rather to show that our results hold. We can see that the $\delta u \delta v$ -tracks are symmetrical about the origin. The $\delta u \delta v$ space corresponding to a given *uv*-track can thus be most concisely described as a “filled *uv*-track”, with its boundaries defined by the ends of the *uv*-track. The set of $\mathcal{F}_{bb'}$, each of which maps one value of the covariance matrix to a fringe in the image-plane, would then characterise a PSF equivalent for the noise distribution, which we refer to as the noise-PSF. In our formalism, the only sources of statistical effects in the field are calibration errors and thermal noise. The average variance in all pixels will be given by the diagonal of

the covariance matrix and the thermal noise, provided that it truly follows a normal distribution. The only effects which will cause the variance in the image-plane to vary from one pixel to the next are those mapped onto the covariance fringes, i.e. such position-dependent variance fluctuations will be caused by correlated gain errors, which are spurious signal introduced by erroneous gain estimates. Assuming all sources in the field are point-like and distant, then these variance fluctuations will follow a specific distribution, convolved with every source in the field.

By the same token, while an image made with visibilities containing only Gaussian noise will have a flat variance, the noise-PSF will correlate the noise between pixels – and the noise-correlation will be given exactly by the PSF. Noise in interferometric images is therefore always correlated between pixels – removing this correlation would require a full image deconvolution. In a sense, this is expected: this is what the PSF, in a very fundamental way, encodes.

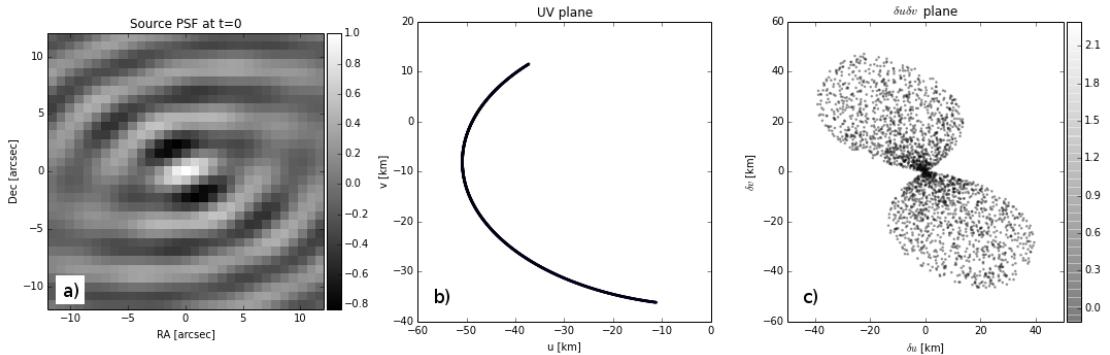


Figure 3.1. Fig. 3.1a shows the PSF image of a simulated 1Jy source at phase centre. Colourbar units are in Jansky. Fig. 3.1b shows the associated UV track, and Fig. 3.1c the corresponding $(\delta u, \delta v)$ tracks. Note that the $\delta u \delta v$ plane does not have a homogeneous point density, but is denser near its origin: here, this is shown by plotting only 1 random point in 10000.

Since the variance fluctuations act as tracers for calibration artefacts, artefacts in the image can be understood as *one realisation* of the variance map, which is characterised by an average level determined by the variance in the gains and thermal noise, and a noise-PSF convolved with the sky brightness distribution. The actual artefacts in the image will still be noisy, as a realisation of the true variance map. For the same reason, in the absence of correlated gain errors, $[\text{Cov}\{\tilde{\gamma}\}]_{b \neq b'}$ are all zero and $\text{Cov}\{\tilde{\gamma}\}$ is a diagonal matrix. We then recover a “flat” noise-map: the variance will be the same in all pixels, as the noise-PSF is absent. In the ideal case, were we to recover the true value of the gains for all times and frequencies, this becomes pure thermal noise.

3.2.3 Noise Map Simulations

We have shown in Eq. 3.25 that there exists an analytical relationship between residual visibility statistics and image-plane residual statistics. This section gives details of simulations we have performed to support our claims on this “Cov-Cov” relationship. Specifically, we simulate residual visibilities for a single baseline by generating a set

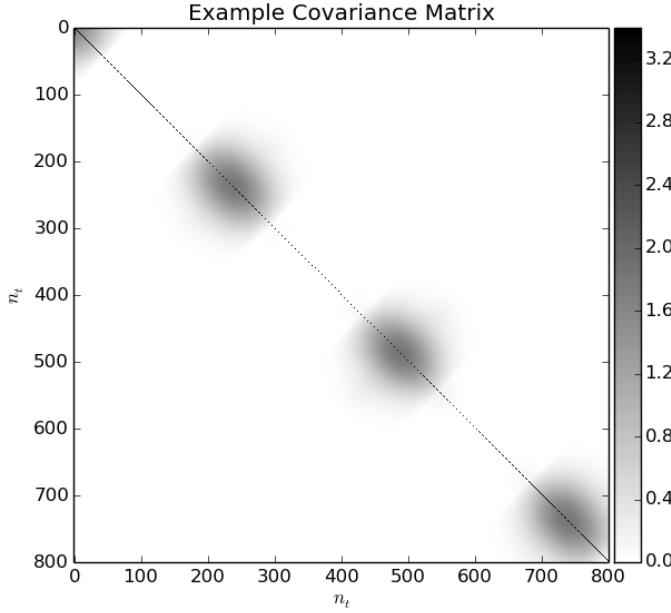


Figure 3.2. Example of a non-stationary covariance matrix, which can be used to simulate $\text{Cov}\{\tilde{\gamma}\}$. The colourbar units are Jy^2 . The correlation scale σ_τ is 40 cells, and the variability period is 500 cells. The matrix is made positive semi-definite (and therefore a covariance matrix) through SVD decomposition. The maximum size of the ‘‘bubbles’’ is determined by σ_τ .

of correlated random numbers with zero mean and a distribution following a specified covariance matrix \mathbf{C} . It contains a periodic function of period T along the diagonal, which is then convolved with a Gaussian of width σ_τ corresponding to the characteristic scale of correlation. The values of these parameters are chosen arbitrarily. A small constant term is added on the diagonal, the net value of which is strictly positive. This simulates a low thermal noise. Finally, singular value decomposition is used to ensure that this matrix is Hermitian positive semi-definite. The net effect is a non-stationary correlation: some residuals are correlated with their neighbours, and uncorrelated with others. An example of this covariance matrix for arbitrary parameter values is shown in Fig. 3.2. We see that, for any given point, correlation is stronger with some neighbours than others (as determined by σ_τ). Samples are drawn as follows: \mathbf{C} is built following user specifications as described above. We then find its matrix square root \mathbf{C}_0 so as to apply it to random numbers generated from a normal distribution. We generate 2000 realisations i of our random variables \mathbf{r}_i :

$$\tilde{\mathbf{y}}_i = \sum_b \mathcal{F}_b \mathbf{r}_i \quad \text{with} \quad \mathbf{r}_i = \mathbf{C}_0 \mathbf{x} \quad (3.32)$$

$$\text{and} \quad \mathbf{x} \leftarrow \mathcal{N}(0, 1) \quad (3.33)$$

$$\tilde{\mathbf{y}} = (\cdots \tilde{\mathbf{y}}_i \cdots) \quad (3.34)$$

where $\tilde{\mathbf{y}}$ is a matrix of dimensions $n_{\text{realis}} \times n_b$. Since \mathbf{x} follows a normal distribution, $\text{Cov}\{\mathbf{x}\} = \mathbf{I}$ and the covariance matrix of each $\tilde{\mathbf{y}}_i$ is, by construction, $\mathbf{C} = \mathbf{C}_0 \mathbf{C}_0^H$. The covariance matrix of $\tilde{\mathbf{y}}$ is thus also \mathbf{C} .

As for the *uv*-track, our simulations read a single one from a specified dataset. In this case, we read an 8-hour *uv*-track for a baseline between two arbitrary LOFAR stations (specifically, CS001HBA0 and RS310HBA) in an observation of the Boötes extragalactic field. The effective baseline length varies between 37.9km and 51.8km. The dataset included 20 channels, each with a spectral width of 97.7 kHz; the central observing frequency is 139 MHz. The temporal resolution is 1 measurement per second.

We compare the *measured* variance map $\mathbf{V}_{\tilde{\mathbf{y}}}^m$, built by measuring the variance across realisations at each pixel in the image-plane, with the *predicted* variance map $\text{Var}\{\tilde{\mathbf{y}}\}$, built using the Cov-Cov relationship (Eq. 3.25). Since we are only interested in the variance map, rather than the covariance between pixels, we compute only the diagonal terms.

$$\mathbf{V}_{\tilde{\mathbf{y}}}^m = \text{Diag}\{\tilde{\mathbf{y}}\tilde{\mathbf{y}}^H\} \quad (3.35)$$

$$\mathbf{V}_{\tilde{\mathbf{y}}}^{pr} = \sum_b [\mathbf{C}]_{bb} \mathbf{I} + \sum_{b,b' \neq b} \text{Diag}\{\mathcal{F}_{bb'}\} [\mathbf{C}]_{bb'} \quad (3.36)$$

where the thermal noise is already incorporated into the diagonal of \mathbf{C} and Eq. 3.36 is merely the diagonal operator applied to the Cov-Cov relationship.

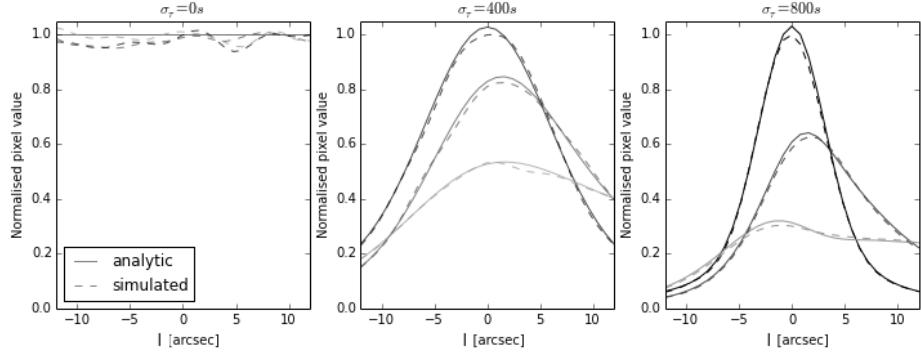


Figure 3.3. The three lines in each figure correspond to three horizontal cross-sections from images in Fig. 3.4. The units on the y-axis are dimensionless $[Jy^2/Jy^2]$. σ_τ is the maximum characteristic error correlation length. In decreasing intensity, they correspond to $m = 0''$, $m = 4''$, and $m = 8''$. The dashed lines correspond to the variance measured with 2000 realisations for each pixel, while the solid line corresponds to the predicted value at that pixel. There are 31 pixels. We do not show cross-sections for negative m due to image symmetry about the origin.

Simulation with a single point source

We model our sky as containing a single 1 Jy point source at phase centre: we thus have $\phi_{bb} = w_b^2$. The source as seen through the set of *uv*-tracks used in our simulation, along with their corresponding $(\delta u, \delta v)$ space, are shown in Fig. 3.1. The *simulated* noise-map is calculated by drawing a large sample ($n_{\text{realis}} = 2000$) of random numbers from the correlated distribution, thereby creating 2000 sets of residual visibilities. By Fourier-transforming the visibilities to the image-plane and taking the variance of the values for each image pixel (i.e. each l, m pair) as per Eq. 3.35, we can estimate $\text{Var}\{\tilde{\mathbf{y}}\}$. The *predicted* noise-map, meanwhile, was found by assigning each cell of \mathbf{C} to the appropriate point in the $(\delta u, \delta v)$ plane and Fourier transforming from this plane into the

image-plane, as per Eq. 3.36. We compare the outcome of simulating a large number n_{realis} of realisations and taking the variance across these realisations for each pixel with mapping the covariance matrix onto the $\delta u \delta v$ -plane and using the Cov-Cov relationship. The results of our simulations are shown side-by-side in Fig. 3.4: the predicted and simulated noise-PSFs match. The peak-normalised predicted noise-PSF is less noisy, as shown in Fig. 3.3 for different correlation scales. This is expected, since it is calculated directly from the underlying distribution, rather than an estimate thereof. As $n_{\text{realis}} \rightarrow \infty$, we expect the two methods to fully converge. As the maximum characteristic correlation length σ_τ increases, the variance becomes ever more sharply peaked.

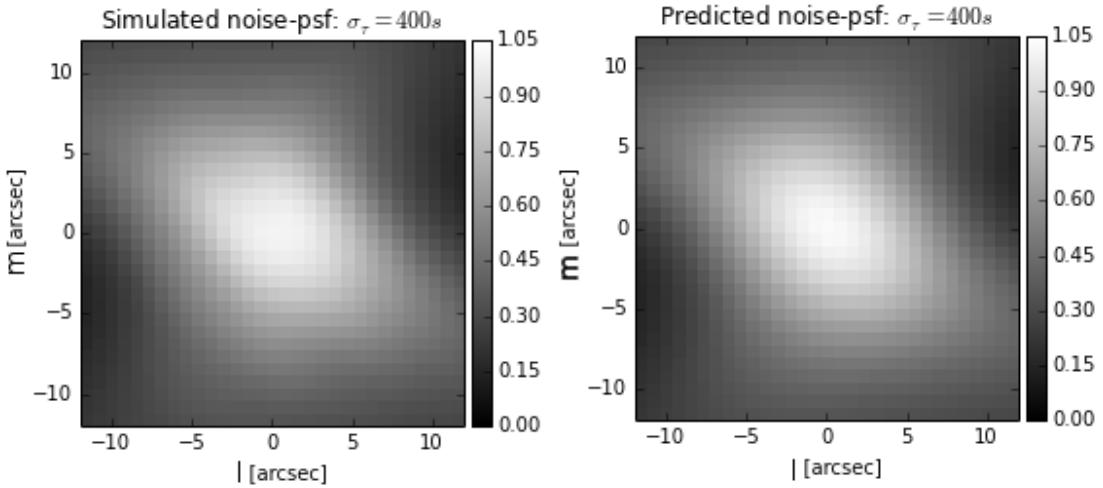


Figure 3.4. Simulated noise-maps, compared with theoretical prediction. The pixel values are normalised by the average value of the covariance matrix: the units of the colourbar are thus dimensionless (Jy^2/Jy^2). These are on the same angular scale as the source shown in Fig. 3.1.

Since our simulated sky consists of a 1Jy source at phase centre, there is only one noise-PSF to modulate the average noise-map, and it lies at phase centre. Let us test our formalism further by considering a model with multiple point sources.

Simulation with 3 point sources

We wish to test our prediction that the noise-map can be described as a convolutional process modulating an average noise level. We thus perform another simulation, this time with three 1Jy point sources. The associated dirty image is shown in Fig. 3.5d.

This dirty image simply consists of performing a direct Fourier transform (i.e. without using a Fast Fourier Transform algorithm) on simulated visibilities corresponding to these three point sources. We now perform a similar test as above on this field. Firstly, we “apply” gain errors to these visibilities by multiplying our model with our residual gain errors. This allows us to find 2000 realisations of residual visibilities, and find the variance for each pixel across these realisations. This gives us the simulated noise-map, shown in Fig. 3.5a. Secondly, we perform a DFT from the differential Fourier plane to the (l, m) plane as before, assigning one cell of $\text{Cov}\{\tilde{\gamma}\}$ to each point of the differential

Fourier plane. This time, however, $\phi_{bb'}^d$ is not simply unity for all points in the differential Fourier plane. Instead, it is calculated for the three-point-source model, and applied for each point. This gives us the predicted noise-map in Fig. 3.5b. Finally, Fig. 3.5c shows the absolute value of the difference between the two images. We see that there is some structure present in these residuals: this is expected, as the PSF of the sources in the dirty images clearly overlap. We are thus not quite in the regime where emission is from well-separated point sources. Nevertheless, our predictions hold to better than 5% accuracy.

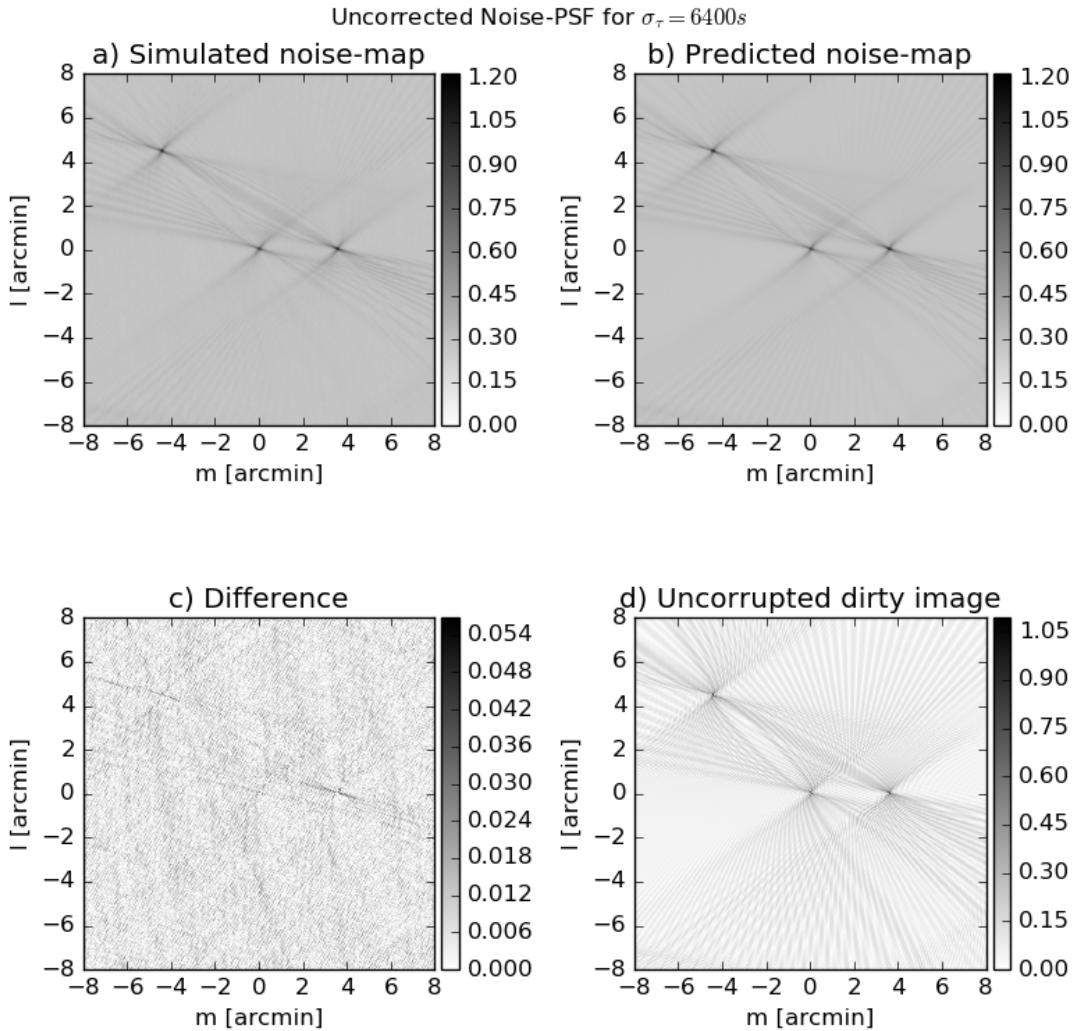


Figure 3.5. Noise-map of sky with correlated gain errors and three point sources. The colourbars of (a), (b) and (c) have dimensionless units, while that of (d) is in Jansky. Note the presence of structure in the residuals (c): these show the limits of our hypothesis that sources are well-separated.

It bears repeating that, for correlated noise, this map can be understood as a distribution map for calibration artefacts: the amount of spurious correlated emission seen

by each baseline will determine the noise-map, and the true image-plane artefacts will then be *one set of realisations* of this *underlying distribution*.

3.3 Adaptive Quality-based Weighting Schemes

As discussed in Section 3.1, some intervals of an observation will have lower gain variability. These will show up in the gain covariance matrix as intervals with lower variance. Similarly, those with larger intrinsic gain variability will have greater error in their gain estimate. By giving greater weights to the former, and lower weights to the latter, we expect to be able to improve image reconstruction. We thus talk of adaptive quality-based weighting, as the weights will adapt based on the calibration quality.

The pixel variance is determined by the visibility covariance matrix, as shown in Eq. 3.27. The diagonal of the visibility covariance matrix will add a flat noise to all pixels, while its wings will determine the calibration artefact distribution, which will be convolved to the sky brightness distribution. We thus have two sources of variance in the image-plane. Minimising the far-field noise (i.e. the variance far from sources) in an image would involve down-weighting noisier calibration intervals while up-weighting the more quiescent ones, without taking noise-correlation between visibilities into account. This is because the far-field noise will be dominated by the diagonal component of the covariance matrix (cf. Eq. 3.27). By the same token, minimising calibration artefacts would involve down-weighting measurements with strongly-correlated noise, and up-weighting the less-correlated. This would not, however, minimise the diagonal component: in fact, it will likely exaggerate its up-weighting and down-weighting. As such, it will increase the constant level of the noise-map, but flatten the noise-PSF's contribution. There are thus two competing types of noise that we seek to minimise: uncorrelated noise, which corresponds to $\delta u \delta v = 0$ (i.e. the diagonal components of the gain covariance matrix), and correlated noise, which corresponds to $\delta u \delta v \neq 0$ (i.e. its wings). Minimising the first will minimise far-field noise without optimally reducing artefacts, while minimising the last will minimise noise near sources at a cost to far-field noise. In the following sub-sections, we will discuss weighting schemes used to accomplish this.

3.3.1 Optimising sensitivity

The Cov-Cov relationship (Eq. 3.25) tells us that, far from any sources, the variance map (Eq. 3.27) is dominated by a constant term: the contribution from thermal noise and the diagonal of the residual visibility covariance matrix. Maximising sensitivity far from sources therefore implies minimising $\text{Diag}\{\text{Cov}\{\tilde{\gamma}\}\}$. This is equivalent to assigning visibilities weights inversely proportional to their variance:

$$w_b = \frac{1}{\text{Var}\{\tilde{\gamma}_b\}} \quad (3.37)$$

For each baseline, those times with larger variance in the residuals will be down-weighted, and those with smaller variance will be up-weighted; this scheme does not require information about the underlying gains, only the error on our solutions. Since we are treating σ_n^2 as a constant for all antennas and all times, those times where

our gains estimate is closer to the true gains will be up-weighted, and those moments where they are farther from the actual gains will be down-weighted: hence the term “adaptive quality-based weighting”. Note that the diagonal of the weighted residuals’ covariance matrix should therefore become constant: this weighting scheme explicitly brings the residuals closer to what is expected in the case of perfect calibration, assuming uncorrelated noise. For the remainder of this paper, we will refer to these weights as *sensitivity-optimal* weighting.

3.3.2 Minimising Calibration Artefacts

Minimising calibration artefacts - i.e. optimising the sensitivity near bright sources - means flattening the noise-map. Since the noise-map can be understood as a noise-PSF convolved with all the modeled sources in the sky modulating the background variance level, it will be flattest when its peak is minimised. From the Cov-Cov relationship (Eq. 3.25), we can see that, at the peak of the noise-PSF (which would be the variance at the pixel where $l = m = 0$), the Fourier kernel is unity: the variance for that pixel is thus the sum of all the cells in the covariance matrix. By accounting for normalisation, we can write the variance at the centre of the noise-PSF as:

$$V(\mathbf{w}) = \frac{\mathbf{w}^T \text{Cov}\{\tilde{\gamma}\} \mathbf{w}}{\mathbf{w}^T \mathbf{1} \mathbf{1}^T \mathbf{w}} \quad (3.38)$$

Our optimality condition is then, after some algebra:

$$0 = \frac{\partial}{\partial \mathbf{w}} (V) \quad (3.39)$$

$$\Leftrightarrow \text{Cov}\{\tilde{\gamma}\} \mathbf{w} = \mathbf{1} \mathbf{1}^T \mathbf{w} (\mathbf{w}^T \mathbf{1} \mathbf{1}^T \mathbf{w})^{-1} \mathbf{w}^T \text{Cov}\{\tilde{\gamma}\} \mathbf{w} \quad (3.40)$$

We find that one \mathbf{w} which satisfies the above is:

$$\mathbf{w} = \text{Cov}\{\tilde{\gamma}\}^{-1} \mathbf{1} \quad (3.41)$$

where $\mathbf{1}$ is a vector of ones. Provided that the hypotheses made so far hold⁶, these weights depend only on calibration quality: badly-calibrated cells will include spurious time-correlated signal introduced by trying to fit the noise n on visibilities. Down-weighting these cells helps suppress artefacts in the field, at the cost of far-field sensitivity. This weighting scheme is thus only a function of the relative quality of calibration at different times, boosting better-calibrated visibilities and suppressing poorly-calibrated visibilities. For the remainder of this paper, we will refer to these weights as *artefact-optimal* weighting.

3.4 Estimating the Covariance Matrix

In our simulations, we have worked from a known covariance matrix and shown that our predictions for the residual image’s behaviour hold. With real data, however, we do not have access to this underlying covariance matrix. Since our weights are extracted from said matrix, estimating it as accurately as possible remains a challenge: this is in turn limited by the number of samples which can be used for each cell.

⁶Specifically, that the residual visibilities do not contain any true signal, i.e. that all sources have been properly subtracted from the visibilities

Each cell in the covariance matrix is built by averaging a number of measurements, or samples. The more samples are available, the better our estimate becomes: once we have more samples than degrees of freedom, we say that our estimation is well-conditioned. Otherwise, it is poorly-conditioned. In this section, we will discuss ways in which we can improve the conditioning of the covariance matrix estimation.

3.4.1 Baseline-based Estimation

One way to improve the conditioning of our covariance matrix estimation is to make the same hypothesis as the calibration algorithm: we can treat the underlying gains as constant within each calibration interval. Provided this interval is known, this allows us to find a single estimate for each interval block of the covariance matrix, turning a $n_b \times n_b$ matrix into a smaller $n_{\text{intervals}} \times n_{\text{intervals}}$ equivalent, where $n_{\text{intervals}}$ is the number of solution intervals used for to find the gain solutions. We then improve our conditioning by a factor of n_{int} , which is the number of samples in a calibration interval. The estimate $\widehat{\text{Cov}\{\tilde{\gamma}\}}$ of the covariance matrix $\text{Cov}\{\tilde{\gamma}\}$ is built by applying the covariance operator:

$$\widehat{\text{Cov}\{\tilde{\gamma}\}} \stackrel{\text{def}}{=} \widehat{\mathbf{C}}_{\tilde{\gamma}} = \frac{1}{n_{\text{int}}} \sum_{i \in n_{\text{int}}} (\tilde{\gamma}_i - \langle \tilde{\gamma} \rangle) (\tilde{\gamma}_i - \langle \tilde{\gamma} \rangle)^H \quad (3.42)$$

where the $\langle \dots \rangle$ operator denotes taking the average over the full vector. If the calibration solver's gain estimates are unbiased (i.e. $E\{\hat{\mathbf{g}}\} = \mathbf{g}$) and the model of the sky is sufficiently complete, this quantity should be zero. Having created $\widehat{\mathbf{C}}_{\tilde{\gamma}}$, which will be of size $n_b \times n_b$, its cells can now be averaged over blocks of $n_{\text{int}} \times n_{\text{int}}$. This allows us to estimate the weights for each baseline and each time.

Mathematically, we retrace the steps of Section 3.2. In the absence of direction-dependent effects, we define the residual visibilities as before, and use them to define the normalised residual visibilities ρ_b :

$$r_b = w_b \kappa_b \tilde{\gamma}_b + \epsilon \quad (3.43)$$

$$\rho_b = \frac{r_b}{k_b} \quad (3.44)$$

We then organise the residuals in cells:

$$\mathbf{r}_{\mathcal{C}} = \begin{pmatrix} \vdots \\ \rho_{b \in \mathcal{C}} \\ \vdots \end{pmatrix} \quad (3.45)$$

$$\mathbf{R} = (\dots \quad \mathbf{r}_{\mathcal{C}} \quad \dots) \quad (3.46)$$

\mathbf{R} corresponds to a matrix containing all the residual visibilities within one calibration cell \mathcal{C} , i.e. for $b \in \mathcal{C}$ where $\mathbf{g}_{\mathcal{C}} = \text{const}$. It is therefore of size $n_{\text{intervals}} \times n_{\mathcal{C}}$, where $n_{\text{intervals}}$ is the number of calibration intervals in the observation. Normalising the residual visibilities by k_b allows us to recover the underlying covariance matrix by multiplying the residual visibility matrix \mathbf{R} with its Hermitian conjugate:

$$\widehat{\mathbf{C}}_{\tilde{\gamma}}[b \in \mathcal{C}, b' \in \mathcal{C}'] = (\mathbf{R}^H \mathbf{R})[\mathcal{C}, \mathcal{C}'] \quad (3.47)$$

Note that we have divided the noise term by the flux model S_b , which can be very small in some cells. As such, care must be taken not to cause the relative thermal noise contribution to explode: those cells where this would occur are dominated by thermal noise, and information on the covariance matrix cannot be recovered from them.

In this framework, we simply treat the index \mathcal{C} as containing all the times and frequencies, for individual baselines, corresponding to a single calibration interval. $\widehat{\mathbf{C}}_{\tilde{\gamma}}$ is then an estimate of the residual visibility covariance matrix.

3.4.2 Antenna-based Estimation

In the subsection above, we assumed that finding one solution per interval will give us strong enough constraints to make the problem of estimating the covariance matrix well-conditioned: this may not be true in all cases. Conditioning may then need to be improved further: in this subsection, we show one way in which this can be done. There are others, e.g. using the rank of the matrix itself to find better-conditioned estimates of the covariance matrix at a lower resolution (i.e. a single estimate for a greater number of cells). They will not be presented in this paper, but are a possible avenue future work on this topic.

In estimating the covariance matrix for each baseline and each calibration cell, we are severely limited by the small number of samples in each cell. One way to overcome this problem is to find estimates for the variance of *antenna* gains, and use these to return to the baseline variances. In this formalism, we extend \mathcal{C} to include all visibilities pointing at a single antenna at a given time. Let us begin by writing an expression for the gain vector, which contains the gains for all antennas and all calibration cells:

$$\hat{\mathbf{g}}_c = \begin{pmatrix} \vdots \\ \hat{g}_p^{\tau \epsilon c} \\ \vdots \end{pmatrix} \quad (3.48)$$

$$\hat{\mathbf{G}} = (\dots \quad \hat{\mathbf{g}}_c \quad \dots) \quad (3.49)$$

and the variance on each antenna in each calibration cell is then:

$$= \mathbb{E}\{\hat{\mathbf{g}}_c \hat{\mathbf{g}}_c^H\} - \underbrace{\mathbb{E}\{\hat{\mathbf{g}}_c\} \mathbb{E}\{\hat{\mathbf{g}}_c\}^H}_{=\mathbf{g}_c \mathbf{g}_c^H} \quad (3.50)$$

As we can see, Eq. 3.50 is simply a vector form of Eq. 3.12. The residual gains of Eq. 3.12 can now be understood as random samples of the covariance between the gains for antennas p and q at a given time, assuming complete skymodel subtraction. We can thus define the variance sample matrix as an *estimate* of the *variance matrix*:

$$\widehat{\mathbf{V}}_{\mathcal{C}} = \widehat{\text{Var}\{\mathbf{g}_c\}} \quad (3.51)$$

$$= \sum_{\tau \in \mathcal{C}} (\hat{\mathbf{g}}_{\tau} \hat{\mathbf{g}}_{\tau}^H - \mathbf{g}_{\tau} \mathbf{g}_{\tau}^H) \quad (3.52)$$

We define the residual matrix as:

$$\mathbf{r}_\tau = \sum_d s_d \mathbf{K}_{d,\tau} (\hat{\mathbf{g}}_\tau \hat{\mathbf{g}}_\tau^H - \mathbf{g}_\tau \mathbf{g}_\tau^H) \mathbf{K}_{d,\tau}^H + \epsilon \quad (3.53)$$

where we explicitly place ourselves in the limits of our formalism, i.e. that we do not have direction-dependent gains. We now see that at the core of Eq. 3.53 lies $\widehat{\mathbf{V}}_{\hat{\mathbf{g}}_\tau}$, where $\sum_{\tau \in \mathcal{C}} \widehat{\mathbf{V}}_{\hat{\mathbf{g}}_\tau} = \widehat{\mathbf{V}}_{\mathcal{C}}$. The K-matrix is defined as follows:

$$\mathbf{K}_{d,\tau} = \begin{pmatrix} k_{p,\tau}^d & 0 \\ 0 & k_{q,\tau}^d \\ \vdots & \ddots \end{pmatrix} \quad (3.54)$$

Since the residual matrix depends on the gains, we define the residual visibility vectors as:

$$\mathbf{r}_{\mathcal{C}} = \begin{pmatrix} \vdots \\ \mathbf{r}_{\tau \in \mathcal{C}} \\ \vdots \end{pmatrix} \quad (3.55)$$

$$\mathbf{R} = (\cdots \mathbf{r}_c \cdots) \quad (3.56)$$

$\mathbf{r}_{\mathcal{C}}$ corresponds to a matrix containing all the residual visibilities within one calibration cell \mathcal{C} , i.e. for $\tau \in \mathcal{C}$ where $\mathbf{g}_{\mathcal{C}} = \text{const}$. Let us define $n_{\mathcal{C}}$ as the number of elements in each calibration cell. The residual variance sample matrix can now be built by multiplying the residual visibility matrix with its Hermitian conjugate:

$$\mathbb{V} = \mathbf{R}^H \mathbf{R} \quad (3.57)$$

Note that we do this because it allows us to turn a single noise realisation ϵ into a statistical quantity σ . We can relate \mathbb{V} to the variance of individual antenna gains:

$$\mathbb{V} = \sum_{\tau \in \mathcal{C}} \left(\sum_{d,d'} s_d \mathbf{K}_{d,\tau} (\widehat{\mathbf{V}}_{\mathcal{C}})^H \mathbf{K}_{d,\tau}^H s_{d'} \mathbf{K}_{d',\tau} (\widehat{\mathbf{V}}_{\mathcal{C}}) \mathbf{K}_{d',\tau}^H + \mathbf{I}\sigma^2 \right) \quad (3.58)$$

To reach this point, in Eq. 3.23, we made the hypothesis that the sky brightness distribution is dominated by emission from well-separated point sources. Applying this hypothesis again here, we can make the approximation that the cross-terms in the sum over d, d' average to zero: $\sum_{d,d' \neq d} \approx 0$. We then have:

$$\mathbb{V} \approx \sum_{\tau} \left(\sum_d s_d^2 \mathbf{K}_{d,\tau} (\widehat{\mathbf{V}}_{\mathcal{C}})^H (\widehat{\mathbf{V}}_{\mathcal{C}}) \mathbf{K}_{d,\tau}^H + \mathbf{I}\sigma^2 \right) \quad (3.59)$$

$$= (\widehat{\mathbf{V}}_{\mathcal{C}})^2 \circ \underbrace{\left(\sum_{\tau} \sum_d s_d^2 \mathbf{k}_{d,\tau} \mathbf{k}_{d,\tau}^H \right)}_{\stackrel{\text{def}}{=} \mathcal{S}} + n_{\mathcal{C}} \mathbf{I}\sigma^2 \quad (3.60)$$

$$\widehat{\mathbf{V}}_{\mathcal{C}} = \sqrt{\mathcal{S}^{-1} (\mathbb{V} - n_{\mathcal{C}} \mathbf{I}\sigma^2)} \quad (3.61)$$

where \circ denotes the Hadamard or entrywise product and $\mathbf{k} = \text{Diag}\{\mathbf{K}\}$. Thus, \mathbb{V} allows us to estimate the variance of each antenna and for each calibration cell by using all the visibilities pointing to that antenna within that calibration cell. With this information, we can then rebuild the baseline-dependent matrix, having improved our sampling by a factor of n_{ant} .

3.5 Applying the Correction to Simulated Data

In this section, we show the impact of our weighting schemes on a noise-map made from arbitrarily strongly-correlated residuals. Here, we assume that our sky contains only a single point source at phase centre: there is thus only a single instance of the noise-PSF, placed at phase centre, to modulate the average variance level. We sample this instance by taking a cross-section from $(l, m) = 0$ to a large l , keeping m constant. The only difference between these cross-sections is the weighting scheme applied: unit weights for all visibilities (“Uncorrected”, blue), sensitivity-optimal weights (green), and artefact-optimal weights (red). We plot both the result predicted by the Cov-Cov relationship (solid line) and the variance estimated across 2000 realisations (dashed line): the result is shown in Fig. 3.6. The two remain in such agreement throughout the cross-section as to be nearly indistinguishable.

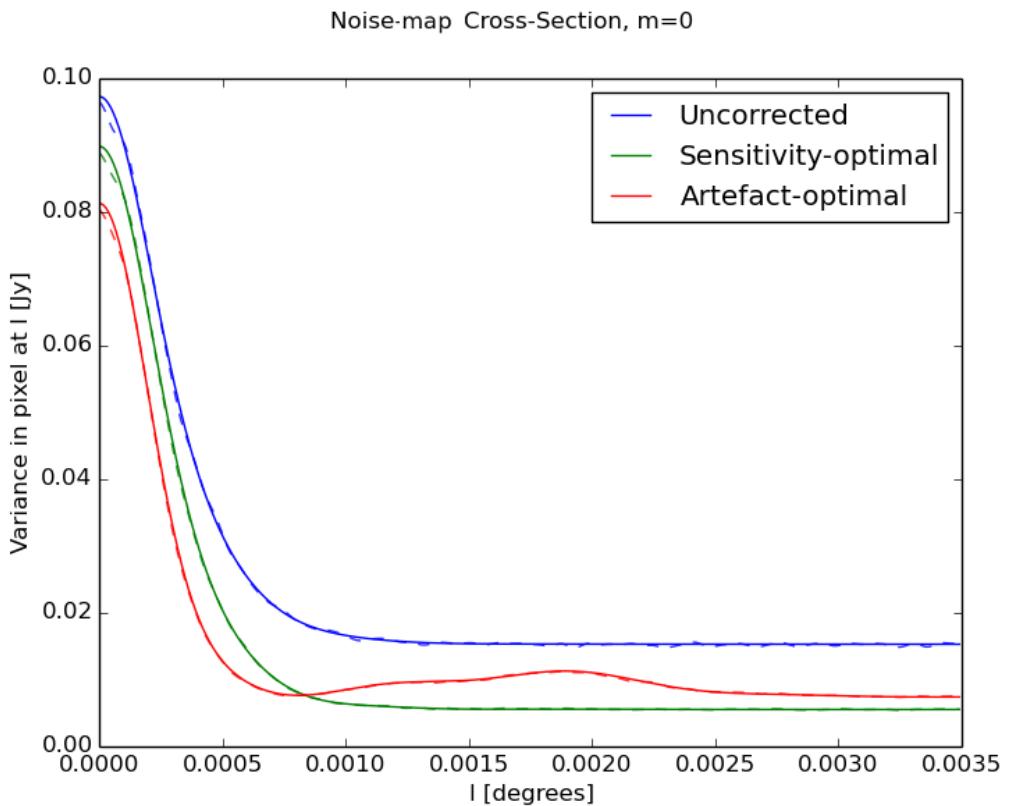


Figure 3.6. The sensitivity-optimal (green) and artefact-optimal (red) weights both give improvements over the unweighted noise-map (blue).

There are a few significant points to note on this figure. Firstly, most of the power in the matrix lies along the diagonal: both weighting schemes thus give good improvements in variance across the map. The artefact-optimal weights, while decreasing the peak further, as expected, also increases the noise far from sources: this is due to the fact that the artefact-optimal weights are in a sense more “selective” than the sensitivity-optimal weights: they up-weigh and down-weigh more severely, and will only result in a

constant covariance matrix if this matrix is zero everywhere outside of the diagonal. In effect, the noise-map becomes flatter, but much broader.

3.6 Applying the Correction to Real Data

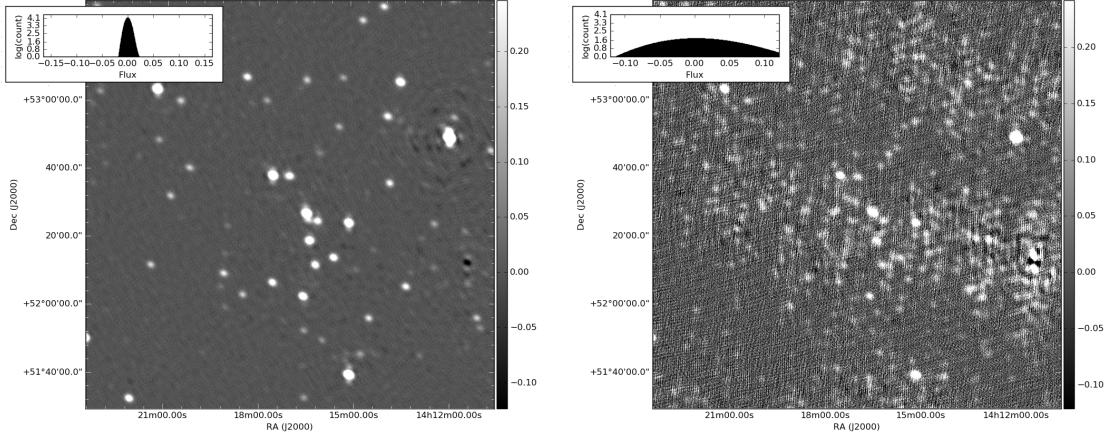
In this section, we show the effect of adaptive quality-based weighting on real data. The dataset used in this section is a single sub-band from an 8-hour LOFAR observation centred on the Extended Groth Strip ($\alpha=14:19:17.84, \delta=52:49:26.49$). The observation was performed on August 28th, 2014. The subband includes 8 channels of width 24.414 kHz each, for a total bandwidth ranging from 150.2 to 150.4 MHz. The data have been averaged in time to 1 data point per second. The data was calibrated using Wirtinger calibration (see Tasse 2014; Smirnov & Tasse 2015, and references therein) and a sky model consisting only of a nearby calibrator source, 3C295. A reference image (a cutout of which is shown in Fig. 3.7a) was made by calibrating the data according to best practice for LOFAR survey data: 1 calibration solution per 8 seconds and per 4 channels. The residual data was then corrected by the gain solutions and imaged using Briggs weighting (robust=0), pixel size of 1.5'', and deconvolved using the default devoncolution algorithm in DDFacet (Tasse et al. 2017).

The data was then time-averaged to create a new, 2.4 GB dataset with 1 data point per 8 seconds. Deliberately poor calibration was then performed on this dataset, solving for 1 calibration solution per 2 minutes (*caeteris paribus*). The resulting corrected residual data was imaged using the same imaging parameters as the reference image, and a cutout of the result is shown in Fig. 3.7b. As expected, the very long calibration intervals introduce calibration artefacts in the image. The brightest sources are still visible, but much of the fainter emission is buried under these artefacts. We are then in a case where our residual visibilities are dominated by calibration error rather than sky model incompleteness.

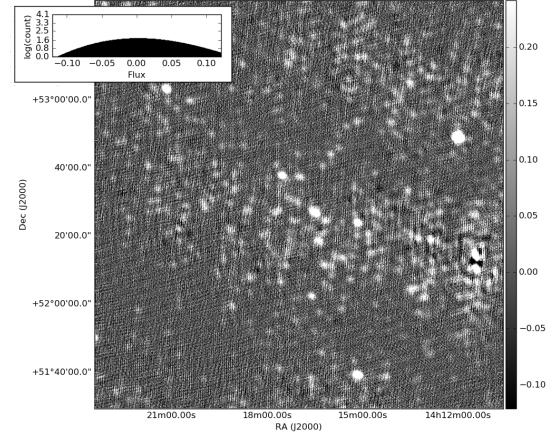
Weights were then calculated based on the badly-calibrated residual visibilities. Fig. 3.7c was made using the same visibilities as Fig. 3.7b and applying baseline-based, sensitivity-optimal weight. Similarly, Fig. 3.7d used the poorly-calibrated residual visibilities with the application of baseline-based, artefact-optimal weighting. These weights are likely to be poorly-conditioned. In both cases, all other parameters were conserved.

Note that applying antenna-based sensitivity-optimal weighting to the badly-calibrated data (not shown here) allows us to recover the reference image with only a very small increase in rms (increased by a factor of 1.14). Further testing on complex field simulations will be required to ascertain the usefulness of artefact-optimal weighting: it is likely that it fails to correct the image fully due to the poor conditioning of the covariance matrix used here.

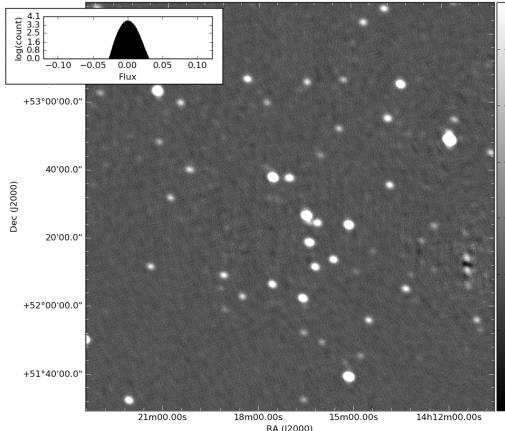
The pixel histograms show us that the weights do not completely mitigate the poor calibration interval choice, but certainly give a dramatic improvement over the unweighted, poorly-calibrated residuals. This is compatible with our statement that the weights give similar residuals in the image with a dramatic improvement in time at some cost in



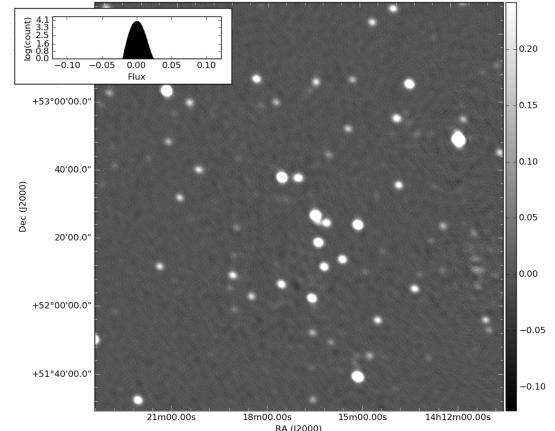
(a) Well-calibrated, unweighted restored image of the sky near the centre of the Extended Groth Strip. Used for comparison with the other images. Units of colourbar are Janskys. This image was made with data calibrated following best practice (solution intervals of 8 seconds, half the bandwidth). $rms=5.87\text{mJy/beam}$



(b) Poorly-calibrated, unweighted restored image of the sky near the centre of the Extended Groth Strip. Units of colourbar are Janskys. This image was made with the same data as for Fig. 3.7a, but averaged in time and calibrated using larger gain solution intervals: 2 minutes and half the bandwidth. $rms=86.4\text{mJy/beam}$



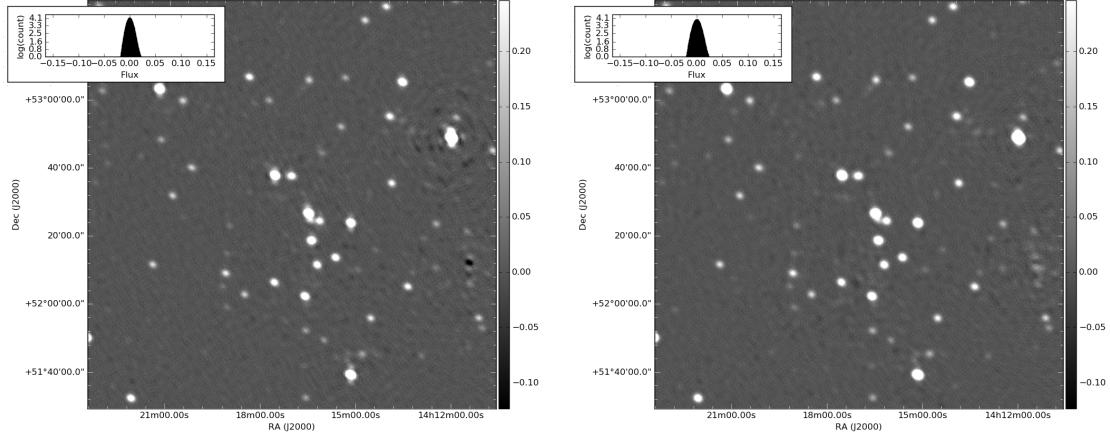
(c) Image made using the same imaging parameters and corrected visibilities as Fig. 3.7b, using sensitivity-optimal weighting. Units of colourbar are Janskys. $rms=9.69\text{mJy/beam}$



(d) Image made using the same imaging parameters and corrected visibilities as Fig. 3.7b, using artefact-optimal weighting. Units of colourbar are Janskys. $rms=15.8\text{mJy/beam}$

Figure 3.7. Restored images of the centre of the Extended Groth Strip, as seen with an 8-hour observation using the full LOFAR array. Fig. 3.7a shows an image of the field made with good calibration intervals. Fig. 3.7b shows an image of the field made with poor calibration intervals. Fig. 3.7c shows image made with the same visibilities and imaging parameters, but with the application of the sensitivity-optimal weighting scheme. Fig. 3.7d, similarly, differs from Fig. 3.7c only in that artefact-optimal weights, rather than sensitivity-optimal weights, were used. The histograms of pixel values in each image have 1000 flux bins ranging from -0.16 Jy to 0.16 Jy. Their ordinates are in log scale. Pixel size is 1.5" in all images.

sensitivity. It is interesting to note that while Fig. 3.7d looks noisier than Fig. 3.7c, its residual flux histogram is actually closer to that of Fig. 3.7a.



(a) Well-calibrated, unweighted restored image of the sky near the centre of the Extended Groth Strip. Used for comparison with the other images. Calibration solution intervals used were 8 seconds, half the bandwidth. $rms=5.87\text{mJy/beam}$

(b) Image made using the same imaging parameters and corrected visibilities as Fig. 3.7b, with the application of antenna-based, sensitivity-optimal weighting. Solution interval of 2 minutes, half the bandwidth. $rms=6.69\text{mJy/beam}$

Figure 3.8. Comparison between the well-calibrated image (i.e. the same image as Fig 3.7a) and antenna-based sensitivity-optimal weights. Units of both colourbars are Jansky. We see that we recover a very similar image, despite the fact that the data used for the weighted image are averaged by a factor of 8 compared to those used for the unweighted image.

As for performance, the weights used for Fig. 3.7d took 8 hours of computing time on a single core⁷, working on a 29 GB dataset, which is not particularly large for LOFAR data. Since the problem is massively parallel, this cost can be alleviated. The main bottleneck is likely due to very poor code optimization. As for the weights used for Fig. 3.8b, they are computed in 1min 6s on the same single core.

⁷Core type: Intel(R) Xeon(R) CPU E5-2660 0 @ 2.20GHz

3.7 Discussion

This chapter began by investigating the use of an algorithm inspired by “lucky imaging” to improve images made using radio interferometric data. By investigating the statistics of residual visibilities, we have made the following findings:

- A relationship between the statistics of residual visibilities and residual pixel values (the “Cov-Cov relationship”).
- A description of the noise-map in the image plane as a constant variance level modulated by a noise-PSF convolved with the sources in the field. This gives the variance in the flux of the image as a function of distance from the sources in the sky for a given calibration.
- An Adaptive Quality-Based weighting scheme, which reduces the noise in the image (and the presence of calibration artefacts) by minimising either the constant noise term or the noise-PSF.

While our results are not a panacea for poor calibration, they show that we can not only improve images made with well-calibrated data, but also mitigate the worst effects of poorly-calibrated visibilities in otherwise well-calibrated datasets. Provided that the gain variability timescale is long enough at certain points of the observation, we can effectively get images of similar quality using both the “standard” best-practice calibration interval for LOFAR survey data (calibration solution interval of 8 seconds) and a significantly larger solution interval of 2 minutes (frequency interval unchanged). Of course, if no such stable interval exists, there will be no good intervals to upweigh, and we will be left only with equally-poor data chunks. This means that, in the right conditions, net pipeline time can be sped up by a factor of nearly three, at a slight cost in sensitivity. This increase will be greater than what could be achieved with existing comparable methods such as “clipping”. However, it does not aim to replace such methods, but rather to act as a complement to them: clipping will be better at flagging a handful of bad visibilities, while this weighting scheme will help push the general visibility statistics (and therefore the image-plane statistics) towards a better distribution.

We emphasize that the adaptive quality-based weighting schemes work because the noise-map describes the *underlying noise distribution*, of which calibration artefacts are *one single realisation*. To fully characterise the artefacts, the correlation between different pixels (i.e. off-diagonal elements of $\text{Cov}\{\tilde{\mathbf{y}}\}$) must be computed; this has not been done in this paper. Nevertheless, lesser constraints on the spatial distribution of artefacts can be found using only the diagonal elements of $\text{Cov}\{\tilde{\mathbf{y}}\}$. The weighting schemes merely seek to minimise this spatial distribution as much as possible: the end result is *fewer artefacts*, which can be distributed across a *much larger area*. This is the source of the dramatic improvement from 3.7b to Fig. 3.7c. We have simply down-weighted those visibilities where spurious signal was introduced by the calibration solutions, and up-weighted those visibilities where such signal was lesser. Since this spurious signal is the source of calibration artefacts, downweighting the associated visibilities reduces it dramatically. The poor improvement from Fig. 3.7b to 3.7d is likely due to limits in the conditioning of our estimation of the covariance matrix.

The work presented here can be improved upon, notably by working on improving the conditioning of our covariance matrix estimate: for real observations, it is impossible to have more than one realization of each gain value for all antennas. By treating each visibility within a calibration interval as a realization of the true distribution, we can better estimate the covariance matrix per baseline, and thus reach a better estimate of the variance in the image-plane. Of course, in practice, we can never access to the true, underlying time-covariance matrix for each baseline. Significant hurdles remain:

- The impact of sky model incompleteness (since calibration requires a sky model) is ignored in this paper; we start by assuming that we have a complete sky model. In practice, of course, acquiring a complete sky model is often a key science goal in and of itself. The impact of this hypothesis therefore ought to be investigated in future work.
- The conditioning of our covariance matrix estimation remains a concern. By using an antenna-based approach, we can improve conditioning by a factor of n_{ant} , but this is only one approach among many. Further work is needed to investigate which method, if any, proves optimal.

Once the conditioning issues of the artefact-optimal weights are resolved, a Briggs-type implementation of the overall adaptive quality-based weighting scheme with a sliding parameter could conceivably be implemented. As things stand, however, there is no good reason to use the currently-existing artefact-optimal weighting scheme over the well-conditioned sensitivity-optimal weighting scheme.

Chapter 4

Imaging the full LOFAR field of view

Contents

4.1	Aims & Methodology	79
4.2	Data Reduction	80
4.2.1	Data & Observation Properties	80
4.2.2	Reducing the Data	80
4.3	Effect of Quality-based Weighting Scheme	84
4.4	Zoo of Radio Sources	88
4.4.1	Selected Sources	88
4.4.2	Images & Cross-matching	89
4.5	Discussion	103

4.1 Aims & Methodology

In this section, we describe the work done to make a wide-field image of the entire LOFAR primary beam, centred on the Extended Groth Strip (EGS). A large-field, image of this field has never yet been made at such low frequencies. Due to technical limitations, the maximum attainable image size (in terms of number of pixels) will limit resolution to about $5''$. As such, we do not use the international stations for this part of the project, as their contribution would be averaged out within a single pixel. They are only used in later parts of the project, when imaging 3C295 and calibrator sources in the field at high resolutions. A more complete sky model can be created from this image, and bright out-of-field sources can be subtracted from high-resolution images of the EGS so that their sidelobe emission do not contaminate the final images. To achieve this, we require direction-dependent calibration. Because we do not use the international LOFAR stations, we do not need a high-resolution model of 3C295.

We aim to achieve a sensitivity of a few hundreds of μJy using visibilities recorded between 110 and 180 MHz. This corresponds to a sensitivity of $\sim 20 \mu\text{Jy}$ at 1.4 GHz for typical synchrotron radio sources, which have spectral indices of $\alpha \sim -0.7$

(where $S_\nu \propto \nu^\alpha$). Reaching this sensitivity over the full LOFAR primary beam requires direction-dependent calibration. Our methodology is therefore to calibrate the full LOFAR bandwidth using the LOFAR Surveys KSP pipeline, which performs a facet-based direction-dependent calibration. We then perform source analysis on the most interesting objects in the field, using ancillary data from NVSS (NRAO VLA Sky Survey, Condon et al. 1998), SDSS (Sloan Digital Sky Survey, York et al. 2000) and WISE (Wright et al. 2010) surveys and cross-referenced using SIMBAD (Wenger et al. 2000). Overlays for those sources will be shown in Section 4.4 in a “radio-source zoo” intended to show the reader what sources a typical LOFAR extragalactic field contains.

4.2 Data Reduction

4.2.1 Data & Observation Properties

The dataset used for this PhD project is part of the LOFAR Surveys KSP, and consists of a number of 8-hour pointings covering as much of the sky visible to LOFAR as possible. We analyse a subset of this dataset, a single one of these pointings, an observation performed on the 28th of August 2014 and centred on the EGS. It is an 8-hour-long observation; there exist another four 8-hour pointings in the same direction, but these data are not reduced in this project. We limit ourselves to analysing only the HBA observation, meaning that we use 365 sub-bands which sample a total bandwidth of 110-182 MHz. We use all Core and Remote LOFAR stations. When imaging 3C295 and other sources in the field with the international stations, we only use those which were online at the time (specifically the German stations 1-5 and 7, along with the Swedish and British stations).

The observation is centred on the EGS, and not on 3C295. This means that our calibrator source is not at phase centre, which can potentially introduce problems. Thankfully, new developments in imaging (Tasse et al. 2017) ensure that the direction-dependent PSF (one of the problems introduced by an off-centre calibrator) is properly modeled and corrected for. This means that direction-independent calibration will be exactly correct in the direction of 3C295, and a good model of this source can be made.

The data was retrieved from the LOFAR Long-Term Archive, and is thus pre-processed and flagged for RFI using the standard tools (NDPPP and AOFlagger; see van Diepen & Dijkema 2018; Offringa 2010, respectively). The next step is therefore to calibrate the data.

4.2.2 Reducing the Data

Reducing LOFAR data is a complex business for a number of reasons. When seeking to make high-fidelity, wide-field images, the need to minimise decorrelation (cf. Section 2.12.2) means that data can only be averaged with care. As visibilities are averaged in time and frequency¹, the relative contribution of signal picked up on longer baselines

¹Provided that simple averaging is performed; the use of baseline-dependent window functions can alleviate this effect to some extent

from sources far from the observation centre is decreased. This translates to a blurring of sources as distance from phase centre increases.

What's more, Jones matrices can vary over short durations or show peculiar spectral behaviour. Even if interferometric data were averaged massively using baseline-dependent window functions, information on this underlying gain behaviour would be lost - this means that averaging data to reduce its size (and associated reduction wall time) is a risky proposition.

As such, we seek to reduce an eight-hour LOFAR HBA observation, averaged down to 1 measurement per second and per 24 kHz². This corresponds to a total of 7 terabytes of raw data: 2920 frequency measurements each second, for 28800 second, of 4 correlations, for each baseline. The observation mode used was HBA_DUAL_INNER, in which all Dutch stations operate with 24 of their 48 tiles, and these substations are then correlated separately with the rest of the array. This was chosen because this configuration does not negatively impact the array density of short baselines (allows for better sensitivity to, and thus recovery of, diffuse emission), nor does it introduce the calibration difficulties that non-uniform beams would cause.

The calibration strategy consists of a first round of direction-independent calibration followed by direction-dependent self-calibration. In effect, we use the following RIME:

$$\mathbf{V}_{pq} = \mathbf{G}_p \left(\sum_s \mathbf{E}_{sp} \mathbf{K}_{sp} \mathbf{B}_s \mathbf{K}_{sq}^H \mathbf{E}_{sq}^H \right) \mathbf{G}_q^H \quad (4.1)$$

which is the same as Eq. (2.29). Here, $\mathbf{K}_{sp} = k_{sp} \mathbf{I}$. \mathbf{K}_{sp} therefore commutes with all other Jones matrices, and we can write: $\mathbf{K}_{sp} \mathbf{K}_{sq}^H = k_{s,pq}$. To account for the faceting explicitly, we can write:

$$\mathbf{V}_{pq} = \mathbf{G}_p \sum_{\Omega} \left(\mathbf{E}_{\Omega p} \left(\sum_{s \in \Omega} \mathbf{B}_s k_{s,pq} \right) \mathbf{E}_{\Omega q}^H \right) \mathbf{G}_q^H \quad (4.2)$$

where Ω indices denote different facets, and s indices denote individual sources. Because we make the approximation that the \mathbf{E} term is constant within a facet, we directly write $\mathbf{E}_{\Omega p}$ rather than \mathbf{E}_{sp} . Each source in a given facet is assumed to have similar direction-dependent gains, and their cumulative signal-to-noise allows for better estimations of this direction-dependent gain than would be possible if solving for each source individually. Thus:

$$\mathbf{V}_{pq} = \mathbf{G}_p \sum_{\Omega} \left(\sum_{s \in \Omega} \mathbf{E}_{\Omega p} \mathbf{B}_s k_{s,pq} \mathbf{E}_{\Omega q}^H \right) \mathbf{G}_q^H \quad (4.3)$$

We begin with direction-independent calibration to find an estimate of $\mathbf{G}_p \quad \forall p$. This first step is done because this part of the Jones chain is the same throughout the sky, and its inverse solution can therefore be applied directly to the visibilities as follows:

$$\mathbf{V}_{pq}^{\text{corr}} = \hat{\mathbf{G}}_p^{-1} \mathbf{V}_{pq} \left(\hat{\mathbf{G}}_q^H \right)^{-1} \quad (4.4)$$

²This is the bandwidth for each channel of our observation.

$$= \hat{\mathbf{G}}_p^{-1} \mathbf{G}_p \sum_{\Omega} \left(\sum_{s \in \Omega} \mathbf{E}_{\Omega p} \mathbf{B}_s k_{\Omega, pq} \mathbf{E}_{sq}^H \right) \mathbf{G}_q^H \left(\hat{\mathbf{G}}_q^H \right)^{-1} \quad (4.5)$$

In so doing, we assume that $\mathbf{E}_{sp} = 1$ quad $\forall s, p$ when solving for $\hat{\mathbf{G}}$. In general, this is of course invalid. For the calibrator source, however, we have $\mathbf{E}_{sp} = 1 \quad \forall p$ by construction. Because we can choose the calibrator source, the presence of a sufficiently bright source (in our case, an unresolved 3C295) will give extremely good estimations of \mathbf{G} . We then have $\hat{\mathbf{G}}_p^{-1} \mathbf{G}_p \sim \mathbf{I}$, and

$$\mathbf{V}_{pq}^{\text{corr}} \sim \sum_{\Omega} \left(\sum_{s \in \Omega} \mathbf{E}_{\Omega p} \mathbf{B}_s k_{s, pq} \mathbf{E}_{\Omega q}^H \right) \quad (4.6)$$

and direction-dependent calibration then consists of solving for $\hat{\mathbf{E}}_{\Omega p} = \hat{\mathbf{G}}_p^{-1} \mathbf{G}_p \mathbf{E}_{\Omega p} \sim \mathbf{E}_{\Omega p}$ for all facets – this means that, if there are errors in the direction-independent gain solutions, they will bias all direction-dependent gain solutions.

Direction-independent calibration is done with Prefactor (van Weeren et al. 2016), using an initial sky model generated using the VLSSr (VLA Low-frequency Sky Survey - redux, Lane et al. 2012), WENSS (WEsterbork Northern Sky Survey Rengelink et al. 1997) and the NVSS (NRAO VLA Sky Survey, Condon et al. 1998), with 3C295 dominating the model. This model also serves as the initial model for the direction-dependent facet-based self-calibration on the corrected visibilities, performed using the killMS-DDF pipeline (Shimwell et al. 2017). The quality-based weighting scheme discussed in Chapter 3 was used to improve self-calibration convergence and the final image.

Direction-dependent calibration is required, as the field of view is large enough that differential ionospheric errors are introduced as a function of direction: the signal from different sources cross different sections of the ionosphere, and the actual gains for these sources are thus different from the gains for the calibrator (see Eq. (4.6)). They must thus be solved for individually. Assuming that the scale of ionospheric change is of the order of a degree (which is what is observed empirically), then a facet-based approach will give a “good enough” result for most of the field. The direction-dependent calibration consisted of a physics-based approach, solving for XX and YY correlations along with a rotation matrix simulating differential Faraday rotation. These gain solutions were then smoothed in time and frequency to provide a time-independent, frequency-dependent amplitude solution for each station. The full LOFAR bandwidth is large enough to allow for Clock/TEC separation, which consists of explicitly modeling the two direction-independent phenomena which dominate in the gain’s spectral structure. The clock difference between stations³ introduces a phase error proportional to ν , while differences in Total Electron Content (TEC) along the line of sight of different stations introduce a phase error inversely proportional to ν . With sufficient spectral coverage, these two effects can therefore be separated and modeled individually, thereby improving the phase solution conditioning.

³Each station has a relatively-accurate clock associated with it, save for the Superterp stations which all share a single clock. Like all clocks, they have finite accuracy, and desynchronisation errors thus creep in over time.

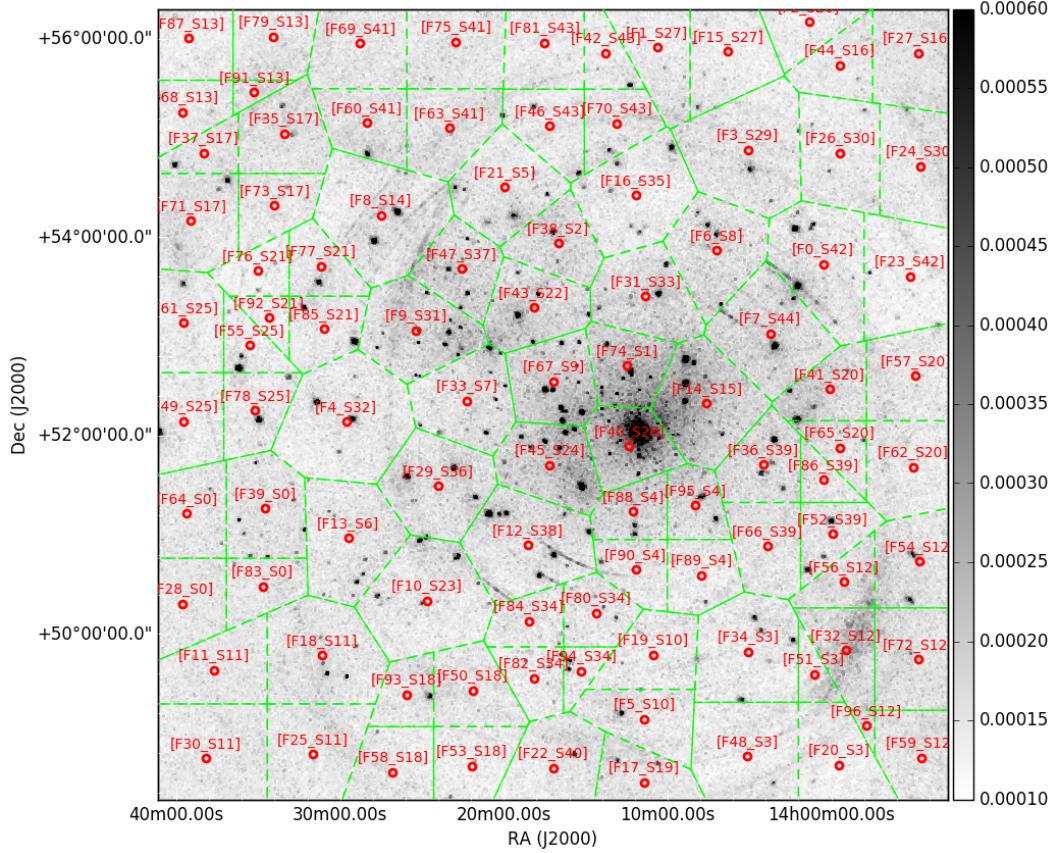


Figure 4.1. Facet distribution for the full direction-dependent calibrated image of the LOFAR primary beam. Image made without international stations. Note that 3C295 is easily found in the image: the dominant calibration artefacts are centred around it.

The direction-dependent calibration was done using 121 facets, spread as shown in Fig. 4.1. Some facets converge faster or better than others, and so artefacts due to direction-dependent effects may remain in certain areas of the image but not others. The final image was imaged with a *uv*-cut, keeping only baselines longer than 100m. This is because shorter baselines were not used during calibration, and so discarded for imaging. The image resolution is 1.5'', and the image was created using both quality-based weighting schemes and Briggs weighting, with a Briggs factor of -0.5. This means that the imaging was skewed slightly in favour of uniform weighting (weighting visibilities by local *uv*-density) rather than natural weighting (same weight for all visibilities). Imaging and calibration were done using the LOFAR beam model used in AWIImager (Tasse et al. 2013), so the image gives integrated flux values rather than apparent flux values. Deconvolution was performed on the full bandwidth⁴ using SSD-clean with an external mask.

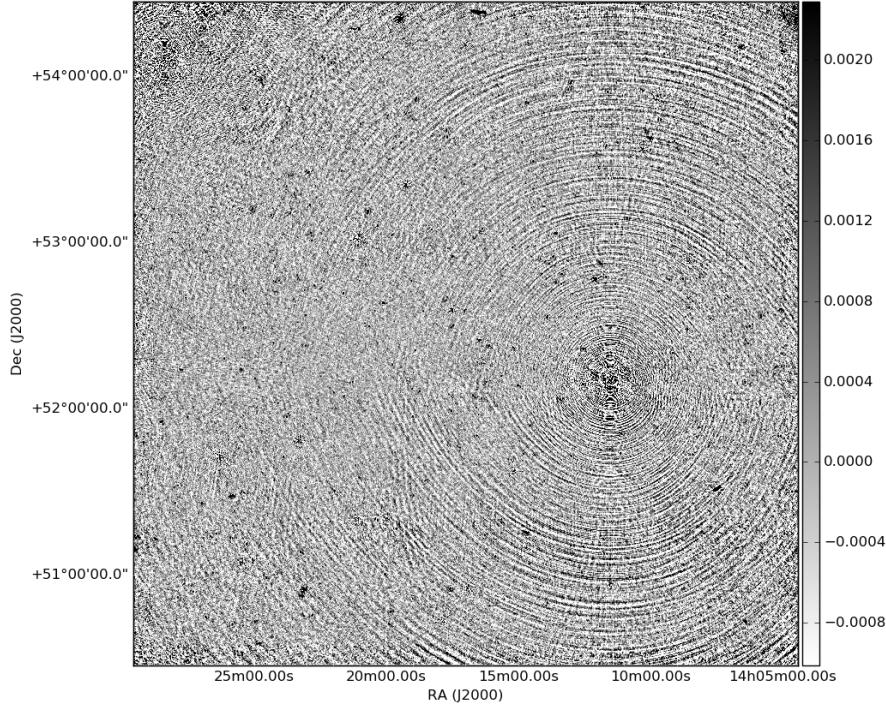
⁴This improves the conditioning of the imaging inverse problem, as it means that bright structure present across the full band (e.g. hotspots) can be deconvolved with the smaller PSF peak of the highest-frequency band.

Direction-dependent self-calibration dramatically improves the noise level in the final image, as shown in Fig. 4.2. Note that not only do artefacts centred on 3C295 (which dominate the noise level in Fig. 4.2a) decrease dramatically, but many of the "diffuse" sources in the field also seem to disappear: this is because their associated artefacts also decrease dramatically, and only physical emission remains.

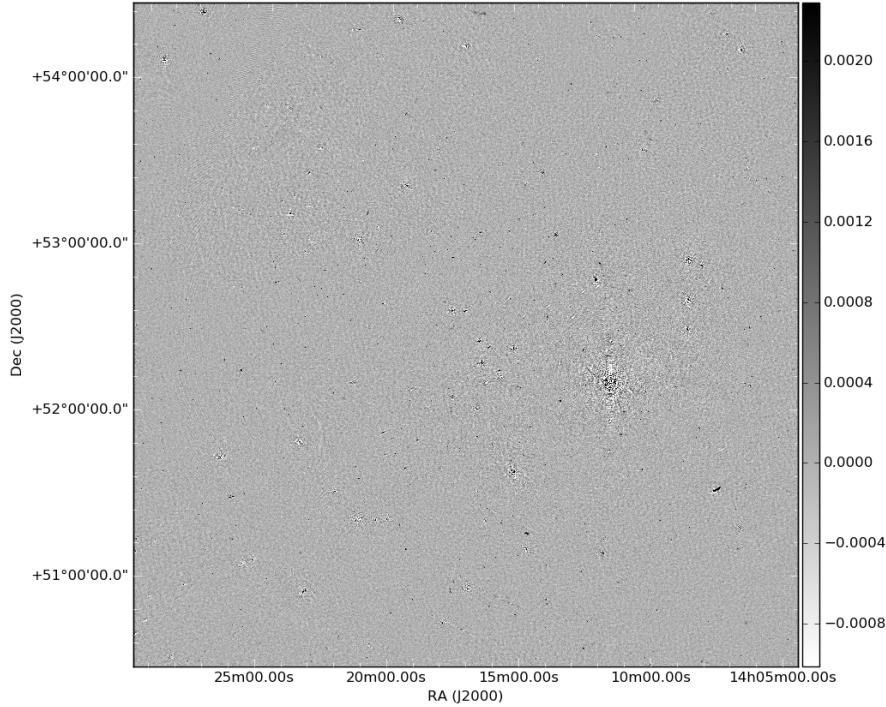
These sources nevertheless remain in the field, as shown in Fig. 4.3. Note that the noise level is not constant throughout the image: this is because the array is not uniformly sensitive to the entire sky, but has a specific beam response. Correcting for this beam response in the image-plane gives integrated flux rather than apparent flux.

4.3 Effect of Quality-based Weighting Scheme

Finally, the quality-based weighting scheme developed in Chapter 3 was applied throughout this procedure. Noise maps made with and without the weighting scheme are shown in Fig. 4.4. A few things are important to note in this comparison. Firstly, the weighting schemes decrease the noise by a factor of ~ 10 : for the noise-map to be visible without weighting schemes, the contrast of the weighted image needed to be pushed way low. For reference, Fig. 4.4b is the same image (albeit without the facets shown explicitly) as Fig. 4.1 - though its levels have been changed to better illustrate the effect of the quality-based weighting scheme. This is why we can't see much at all in that image: it is the contrast between Fig. 4.4b and Fig. 4.4a which is relevant here. The noise distribution is more clearly visible in that figure. Relatedly, by comparing that image and Fig. 4.4a, we note that the unweighted image has a different noise per facet, unlike the weighted image. This is due to the fact that a single visibility will be calibrated better or worse depending on the signal-to-noise of individual facets, constrained by the direction-independent calibration which uses the highest signal-to-noise available. It is no coincidence that the facet containing 3C295 is the only one with a noise even remotely comparable to the weighted image's: it is the facet with the best signal-to-noise. This becomes significant when taking into account the fact that the last 2 hours of this observation are much noisier than the first 6: the quality-based weights will decrease their overall contribution *throughout the image*, whereas not using them will lead to different facets having stronger or lesser contributions from these noisy visibilities - and therefore different noise levels. This is also why fringes can clearly be seen in some facets: a single visibility (or a set of a few visibilities) are particularly badly-calibrated, and the associated fringe dominates in the facet. Linking back to Chapter 3, because we have a gain solution per visibility *and per facet*, we have one noise-PSF per facet; this suggests that the quality-based weighting scheme could likely be extended to a direction-dependent formalism.

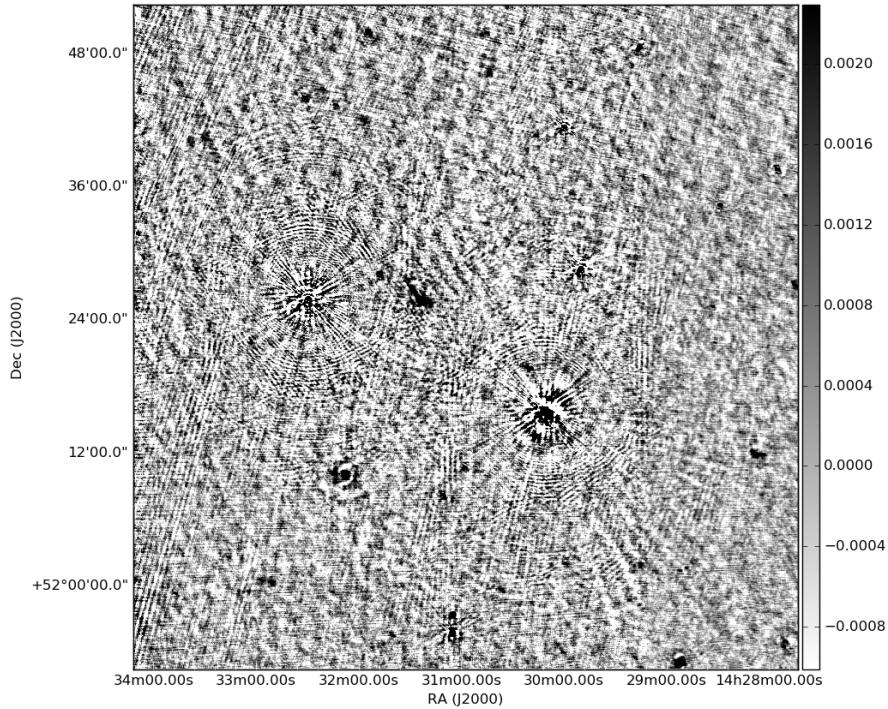


(a) LOFAR wide-field image cutout, calibrated with the initial sky model and no direction-dependent Jones terms.

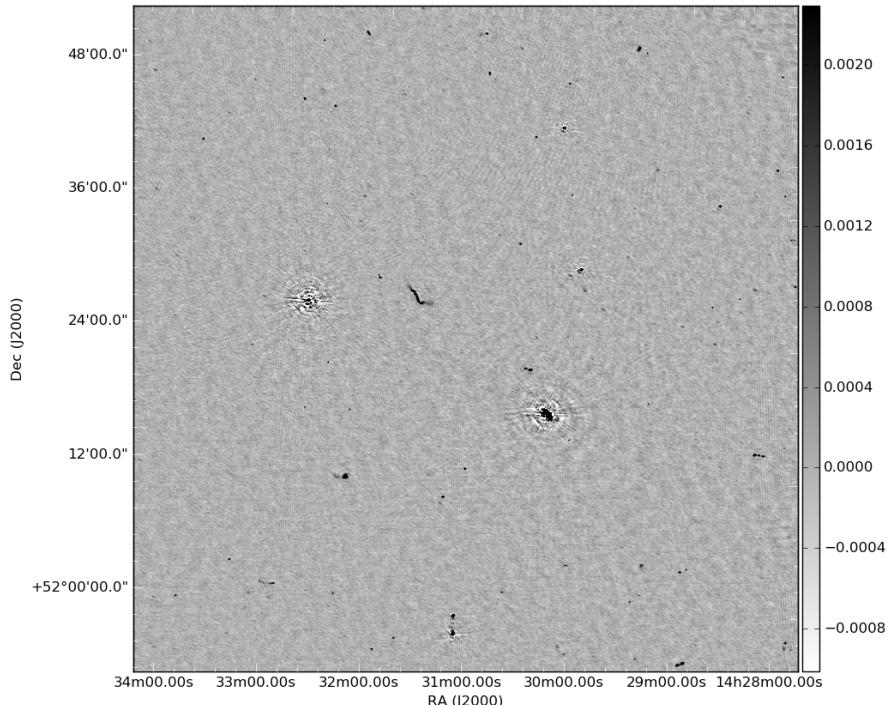


(b) LOFAR wide-field image cutout, self-calibrated from Fig. 4.2a with direction-dependent Jones terms.

Figure 4.2. Wide-field image cutouts of the full primary beam, centred on the EGS, before and after selfcal over multiple facets. Note the massive decrease in artefacts centred on 3C295, but these are not the only artefacts decreased. Fig. 4.3 show zoomed-in cutouts of this field as proof that the calibration artefacts around fainter sources decrease as well as a consequence of direction-dependent calibration.

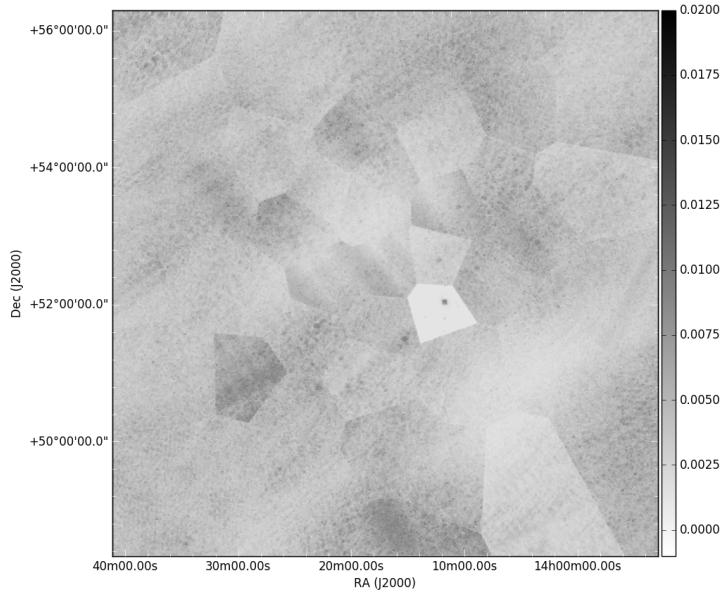


(a) LOFAR wide-field image cutout, calibrated with the initial sky model and no direction-dependent Jones terms.

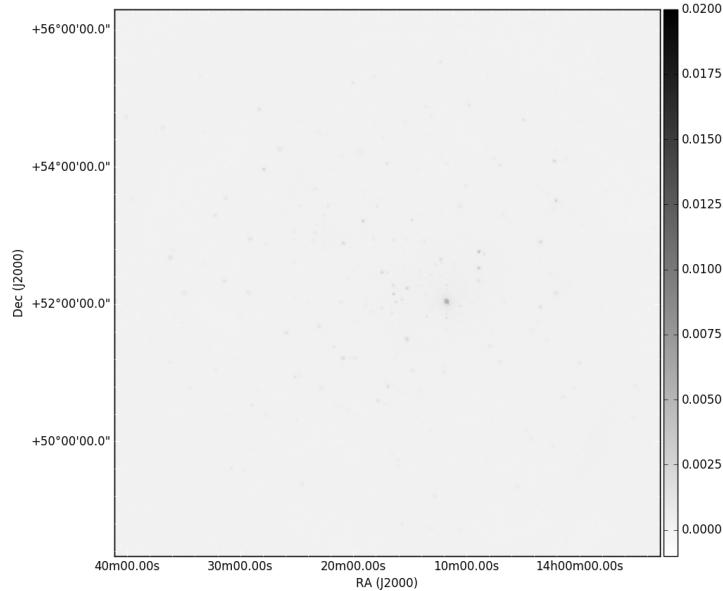


(b) LOFAR wide-field image cutout, self-calibrated from Fig. 4.2a with direction-dependent Jones terms.

Figure 4.3. Cutouts of smaller regions of Fig. 4.2, intended to show that calibration artefacts around fainter sources decrease as well as those around 3C295.



(a) *OFAR wide-field image cutout, self-calibrated from Fig. 4.2a with direction-dependent Jones terms, made without using quality-based weights.*



(b) *LOFAR wide-field image cutout, self-calibrated from Fig. 4.2a with direction-dependent Jones terms, made with quality-based weights.*

Figure 4.4. *Impact of quality-based weighting scheme on the final image.*

4.4 Zoo of Radio Sources

4.4.1 Selected Sources

We select 12 sources in the wide-field image, which is shown in Fig. 4.5. Our selection is primarily based on whether sources have interesting or peculiar diffuse emission, though some are chosen for being particularly classic examples of radio galaxies. We scan the image from south to north and east to west (scanning upwards then shifting to the right in the image). The positions of the brightest sources ($> 1\text{Jy}$) are given in Table 4.1. We insist that this is a very crude scientific analysis, aimed as a showcase for the reader of this manuscript.

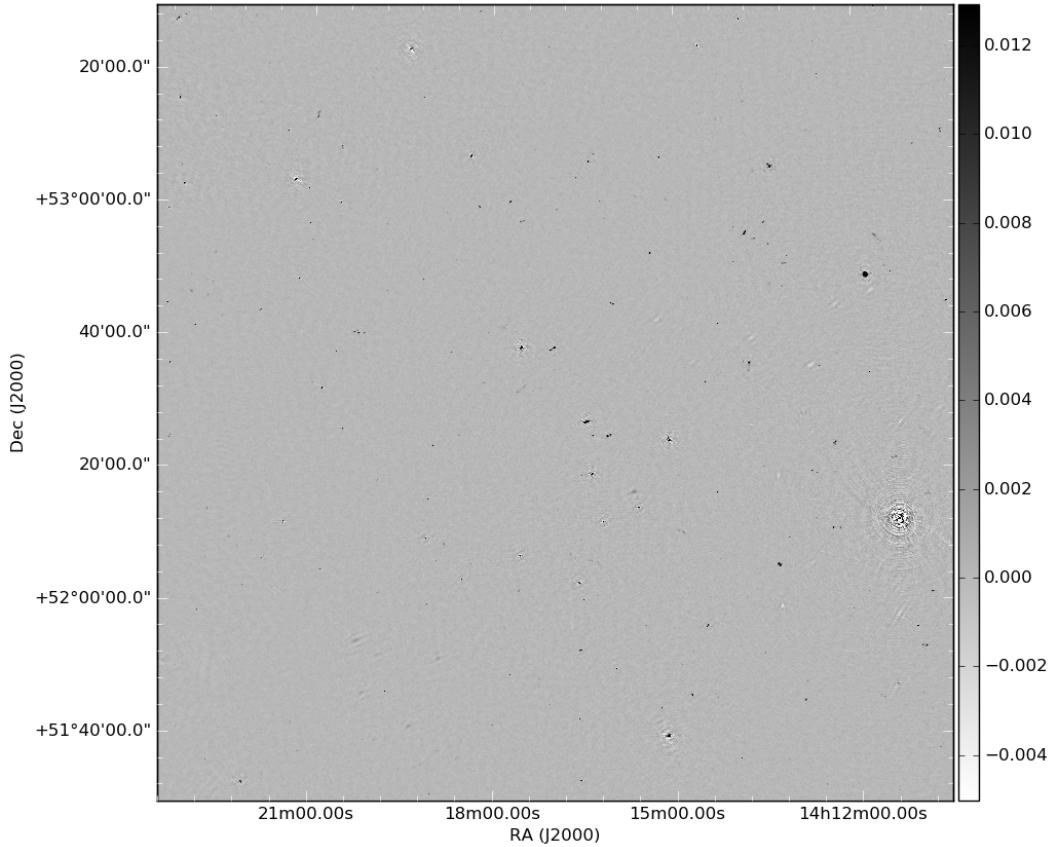


Figure 4.5. Full image of the LOFAR primary beam, centred on the EGS. Image made without international stations and with direction-dependent calibration.

Table 4.1 also gives the likely associations for each chosen EGS source, giving redshift information when available. The analysis here is done by hand using NED (NASA/IPAC Extragalactic Database (NED)), Simbad (operated at CDS, Strasbourg) and the associated catalogues and papers, often accessible through ADS (SAO/NASA Astrophysics Data System (ADS)). Such an artisanal approach is suitable for the analysis of a dozen sources, but quickly becomes intractable. Since source analysis is but a happy byproduct of our primary science aim, a full catalog extraction was not done here, but would be a promising avenue of future work.

Name	RA [hms]	Dec [dms]	Likely Association	LOFAR thumbnail
EGS-1	14:37:39.53	53:36:31.24	NVSS Radio Galaxy	Fig. 4.6a
EGS-2	14:35:27.84	55:07:56.32	Radio Galaxy in Cluster	Fig. 4.7a
EGS-3	14:29:34.13	54:43:46.93	Radio Galaxy in Cluster	Fig. 4.8a
EGS-4	14:31:36.62	52:27:33.75	Radio galaxy, $z = .292$	Fig. 4.9a
EGS-5	14:29:48.89	51:10:30.52	Lobes matched individually	Fig. 4.10a
EGS-6	14:26:04.09	51:29:35.07	One lobe matched individually	Fig. 4.11a
EGS-7	14:17:55.58	50:08:01.75	Radio galaxy, $z = .186$	Fig. 4.12a
EGS-8	14:14:40.42	51:17:41.00	Radio galaxy, z unknown	Fig. 4.13a
EGS-9	14:11:36.43	52:54:25.53	Galaxy cluster, $z = .525$	Fig. 4.14a
EGS-10	14:07:09.90	55:04:22.23	Galaxy cluster, $z = .250$	Fig. 4.15a
EGS-11	14:03:16.00	51:43:35.23	Lobes matched individually	Fig. 4.16a
EGS-12	14:02:43.49	51:03:14.69	Radio Galaxy in Cluster	Fig. 4.17a

Table 4.1. Table recapitulating the names, positions and likely associations of all 12 chosen sources in the primary beam. These sources were primarily chosen because they had peculiar or interesting diffuse emission. The associated LOFAR image thumbnails are also given for reference.

For each source, we show four images: a cutout of the LOFAR image around the source, an SDSS cutout of the same area with the LOFAR image superimposed as an overlay, the same SDSS cutout with an NVSS image overlayed, and finally the same SDSS image again with a WISE image overlay. The first image aims to show why the source was selected. The second shows optical counterparts to LOFAR emission. The third serves to check whether we pick up similar structure and emission in NVSS images as the LOFAR wide-field image. Note that the NVSS images are at a much lower resolution, but can be a useful way to ensure that we are not choosing artefacts or spurious emission as sources to investigate. Finally, the WISE overlay attempts to see whether infrared emission can be associated to these sources. Note that the sensitivity of SDSS images are not homogeneous, as not all patches of the sky have received the same coverage.

4.4.2 Images & Cross-matching

EGS-1

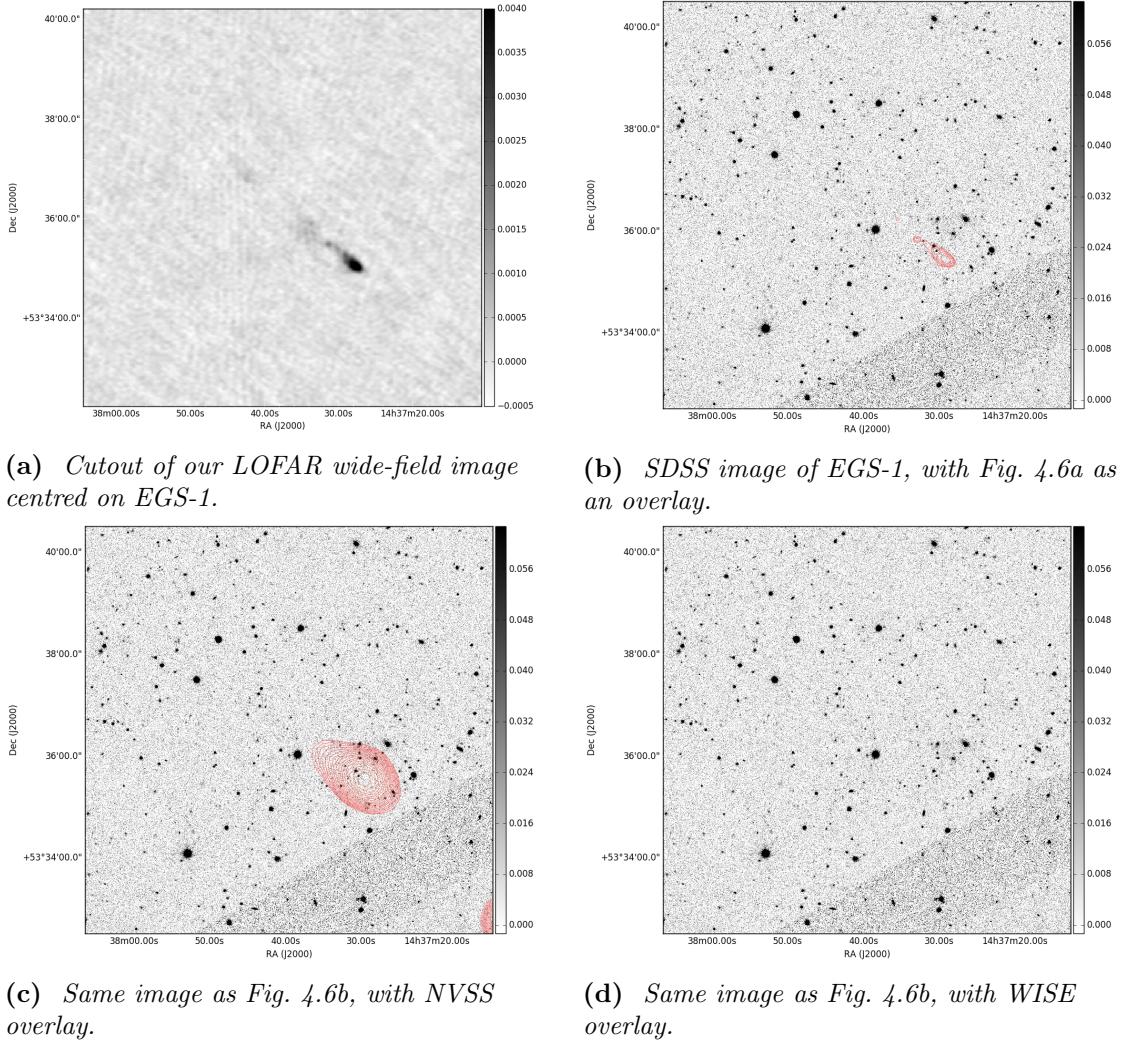


Figure 4.6. Images & Overlays for EGS-1. The overlays start at the 2σ level and end on the image emission peak.

EGS-1 was selected from the wide-field LOFAR image because it showed interesting diffuse structure. It does not seem to match up to any optical source, nor does it match an infrared source. EGS-1 therefore seems to be best identified as a low-frequency counterpart to NVSS J143706+533225. Its shape seems to include two jets, one much brighter than the other, but both visible in Fig. 4.6a. This could include a radio galaxy with one Doppler-boosted jet, or asymmetry in its environment density.

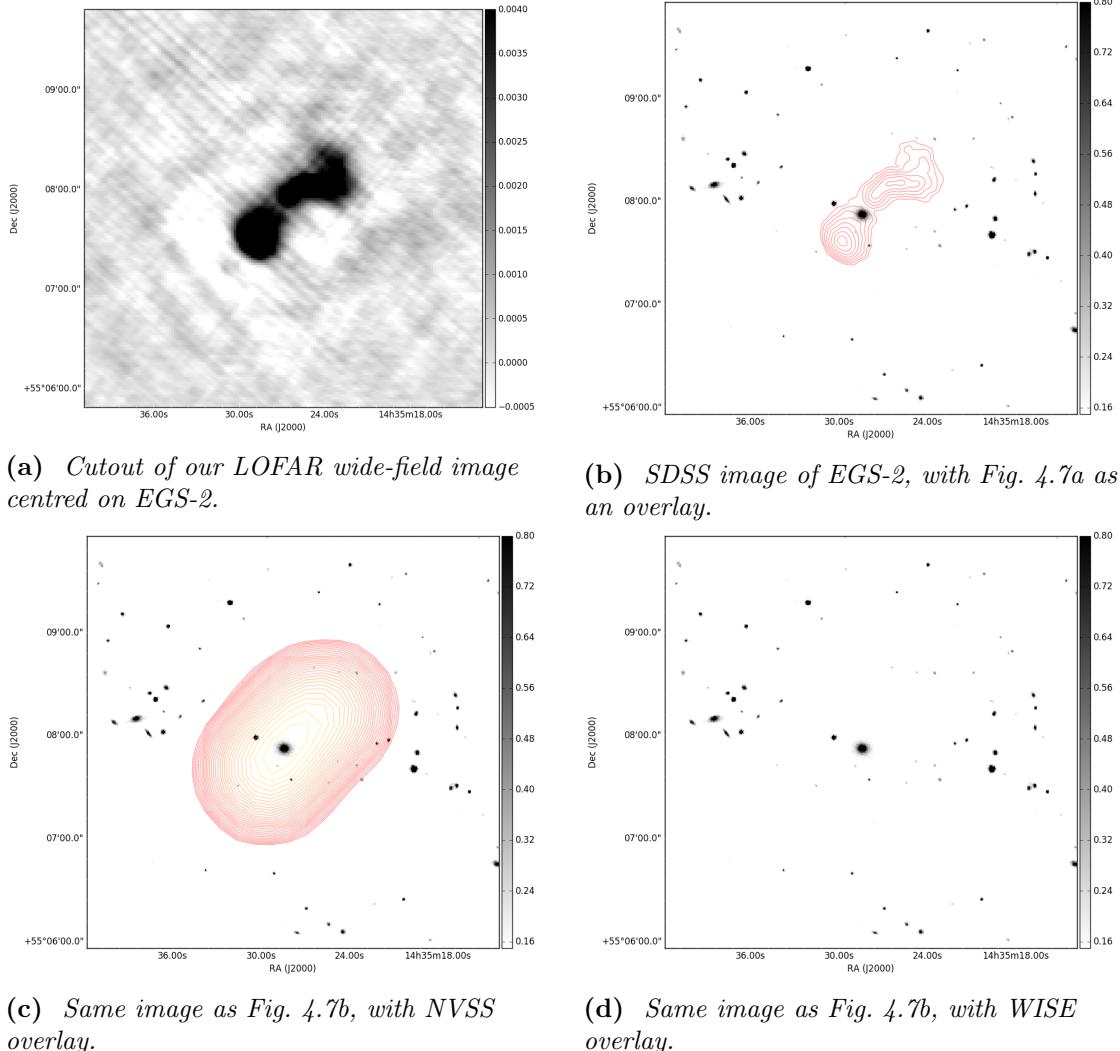
EGS-2

Figure 4.7. Images & Overlays for EGS-2. The overlays start at the 5σ level and end on the image emission peak.

EGS-2 was selected from the wide-field LOFAR image because of the interestingly turbulent morphology of its northern jet, and extended diffuse emission. A bright optical source lies very close to the mid-point of EGS-2. This source - 2MASX J14352846+5507519⁵ - is the brightest galaxy in a cluster. It lies at a redshift of 0.13985 (Ahn & et al. 2013). Furthermore, there is an NVSS source which seems to nicely overlap the EGS-2: NVSS J143527+550756 (Condon et al. 1998). These two sources do not seem to be associated in the literature, though 2MASX J14352846+5507519 is associated with VLSS 1435.4+5507 - a likely counterpart to the NVSS association. These are thus likely to all be the same object.

⁵<http://simbad.u-strasbg.fr/simbad/sim-id?Ident=%40499314&Name=2MASX%20J14352846%2b5507519&submit=submit>

The northern edge of EGS-2's upper jet has interesting morphology: if EGS-2 is indeed associated to the brightest galaxy in the cluster, then this morphology is likely due to differing densities of the intercluster and intracluster media, which the jets will penetrate and interact with in different ways. As ever, because this source lies far from the observation pointing, the sensitivity and calibration are not optimal for analysis: further pointings on EGS-2 might be required to achieve the signal-to-noise required to extract useful information, especially using international LOFAR.

EGS-3

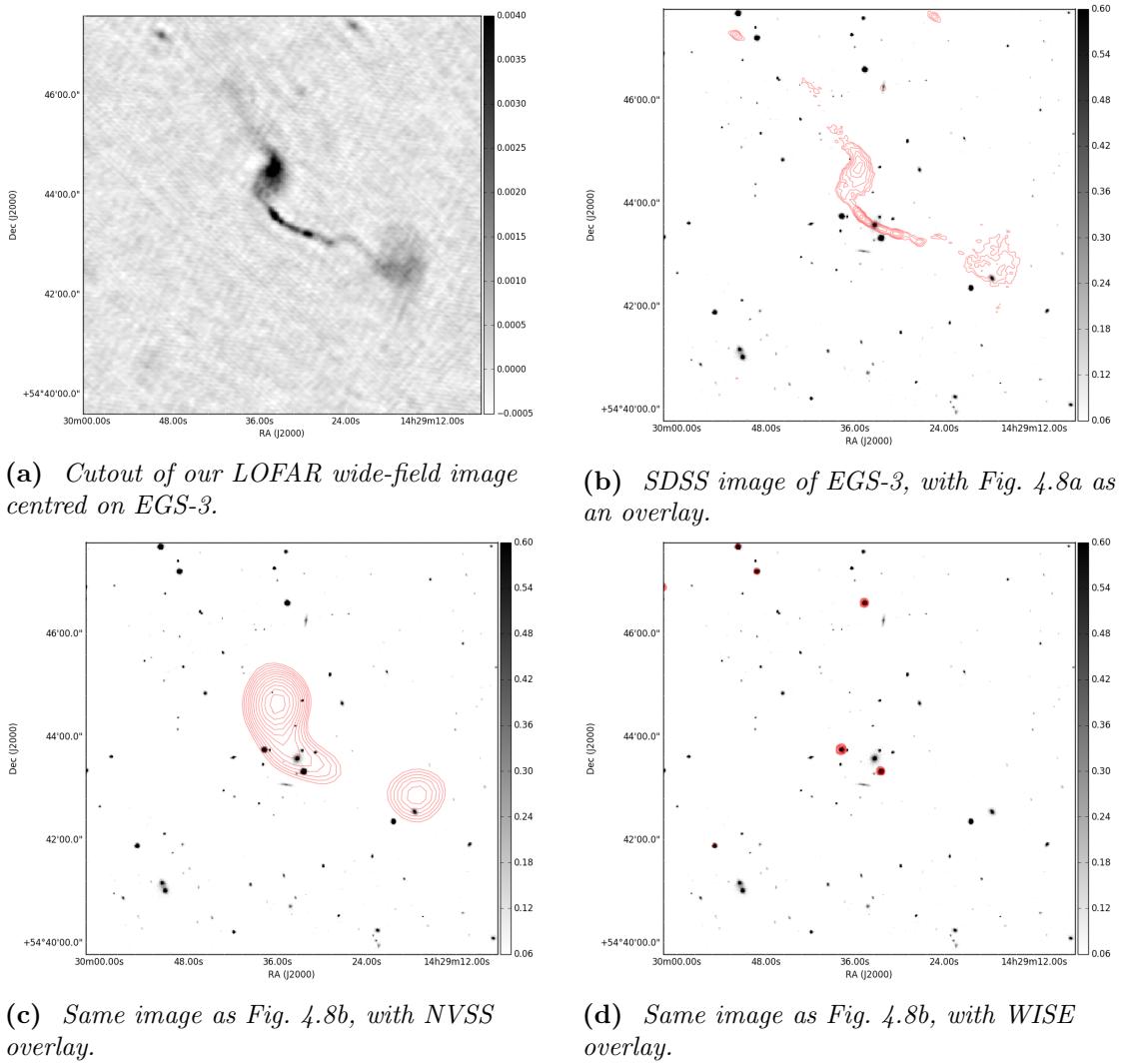


Figure 4.8. Images & Overlays for EGS-3. The overlays start at the 1.5σ level and end on the image emission peak.

EGS-3 is one of the most impressive sources in the primary beam. It has very complex diffuse structure. It almost certainly consists of two jets, though the interactions of these jets with the neighbouring medium must be complex indeed. From Fig. 4.8c, one can

see that the emission can be successfully matched to no less than two individual NVSS sources. In other words, if those NVSS sources are linked to the LOFAR source - and given their spatial distribution, they almost certainly are - then they are two components of a single, deeper structure. The infrared image tells us very little.

Note that an optical source lies along the line linking the two diffuse poles of EGS-3: this source is SDSS J142933.44+544335.2, the brightest galaxy in a cluster. If EGS-3 is a radio galaxy with jet emission, which it almost certainly is (the rough symmetry of the two “cotton balls” on either side of a “rod”-like component strongly indicate that this is a very warped jet), then its optical emission must lie somewhere within the “rod”. That this emission would be associated with violent interactions between jet-accelerated particles and ambient inter-cluster and intra-cluster media seems likely: it would explain the complex, turbulent distribution of emission and the strong “warping” of the jet structure into the whirls and eddies seen in Fig. 4.8a. EGS-3 can therefore be fairly unambiguously associated to a galaxy cluster.

EGS-4

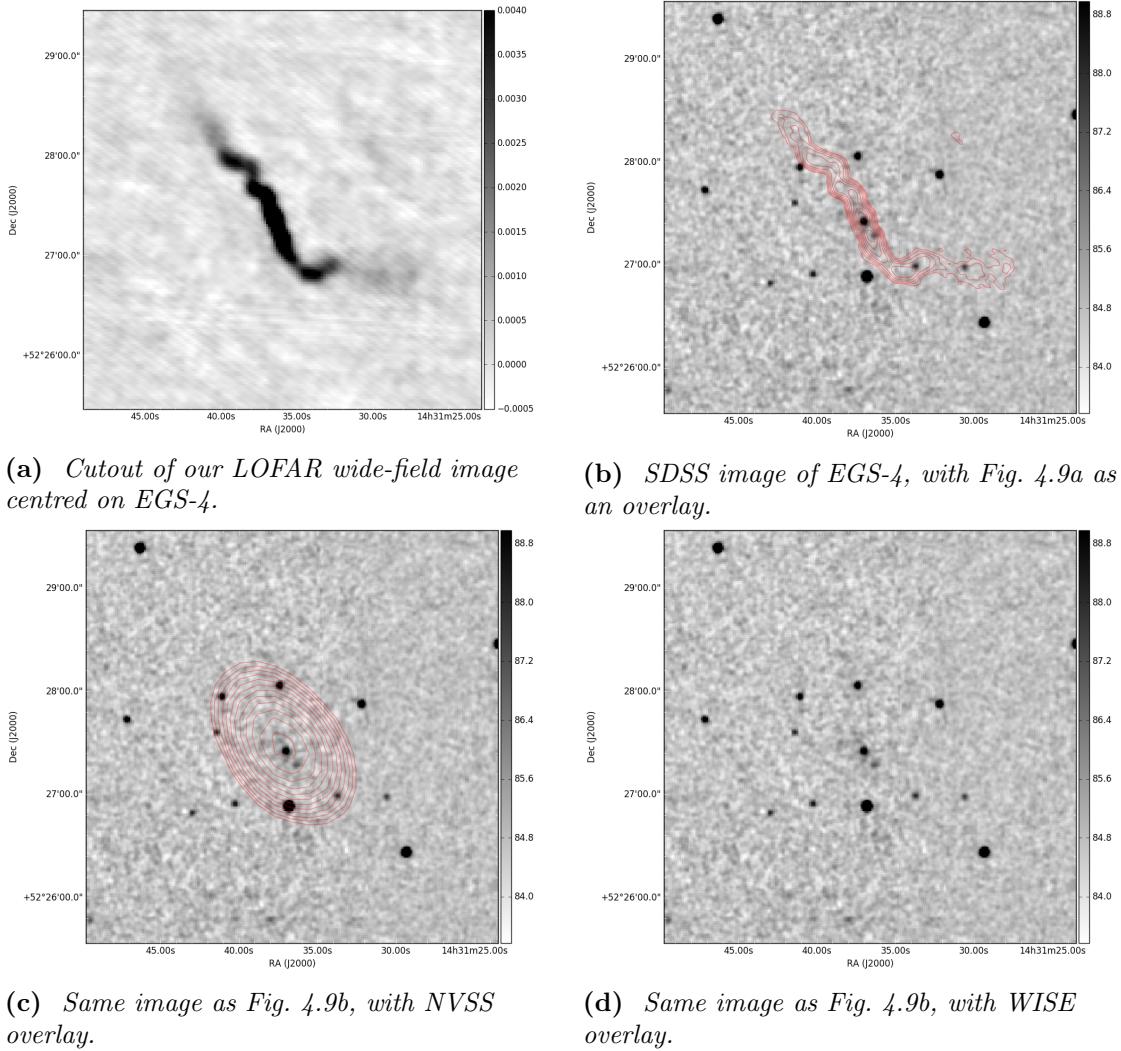


Figure 4.9. Images & Overlays for EGS-4. The overlays start at the 2σ level and end on the image emission peak.

Another impressive source, EGS-4 was selected for its complex diffuse structure. Like EGS-3, the structure of this source seems to indicate that it would be a radio jet galaxy within a galaxy cluster. Its LOFAR structure matches the lower-resolution NVSS structure, which is a strong indication that it is not spurious. By the same argument as for EGS-3, we expect this diffuse emission to have associated optical emission somewhere along its central axis. As we can see from Fig. 4.9c, EGS-4 can be associated to an NVSS source: NVSS J143137+522728. This radio galaxy lies at redshift 0.292 (Lin et al. 2010), and is itself associated to the NVSS source in the middle of EGS-4 and its WISE counterpart. SIMBAD has no further association for this radio galaxy. Further pointings will be required to acquire more information on this source and its environment.

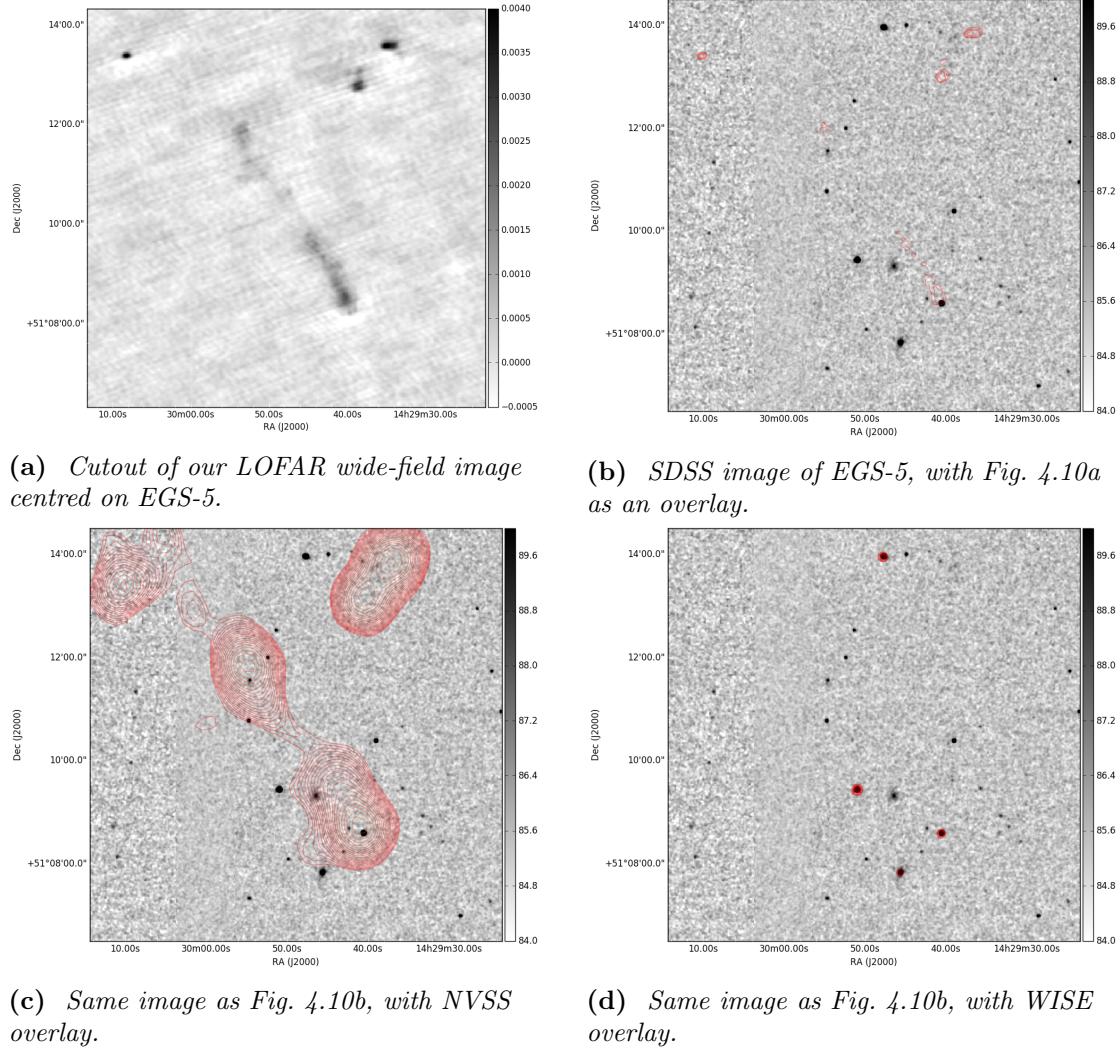
EGS-5

Figure 4.10. Images & Overlays for EGS-5. The overlays start at the 1.2σ level and end on the image emission peak.

EGS-5 was chosen due to its size, and was expected to be a very classic radio galaxy. Unfortunately, it lies too close to the local noise to show up in the overlay map (the lowest level of which lies at 5σ). In fact, its lobes were only matched individually. The southern lobe matches to NVSS J142941+510850 (White et al. 1997). The northern lobe matches to NVSS J142954+511147 (Condon et al. 1998). It is unfortunate that the centre of EGS-5 lies in a region of poor SDSS coverage. Either the original sources are incorrectly identified as individual radio galaxies, and are actually the lobes of a single galaxy, or our analysis of the source is flawed, and we are in fact dealing with two separate sources. The diffuse structure between them in the LOFAR band would seem to indicate that the former case is closer to the truth.

EGS-6

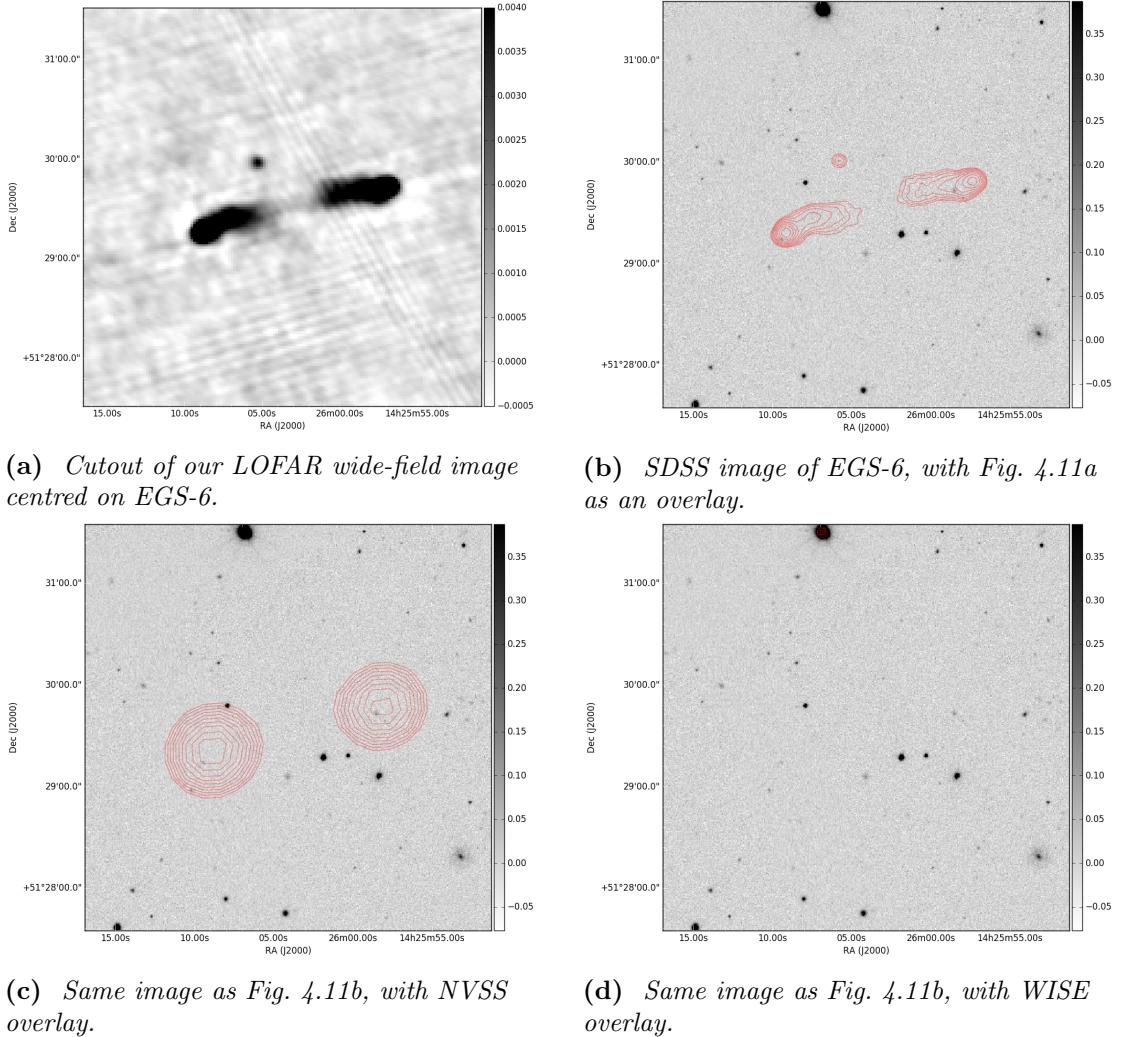


Figure 4.11. Images & Overlays for EGS-6. The overlays start at the 1.5σ level and end on the image emission peak.

EGS-6 was also chosen as it seemed like a very standard, classical radio-strong galaxy. There is no match for the centre of EGS-6. The eastern (leftwards) lobe most likely matches NVSS J142608+512921 (Condon et al. 1998). The western (eastwards) lobe can be matched to SDSS J142558.51+512943.0, but to no NVSS candidate. This is somewhat concerning, as there does seem to be a source in the NVSS overlay at this position, but this source is recorded in neither SIMBAD nor NED.

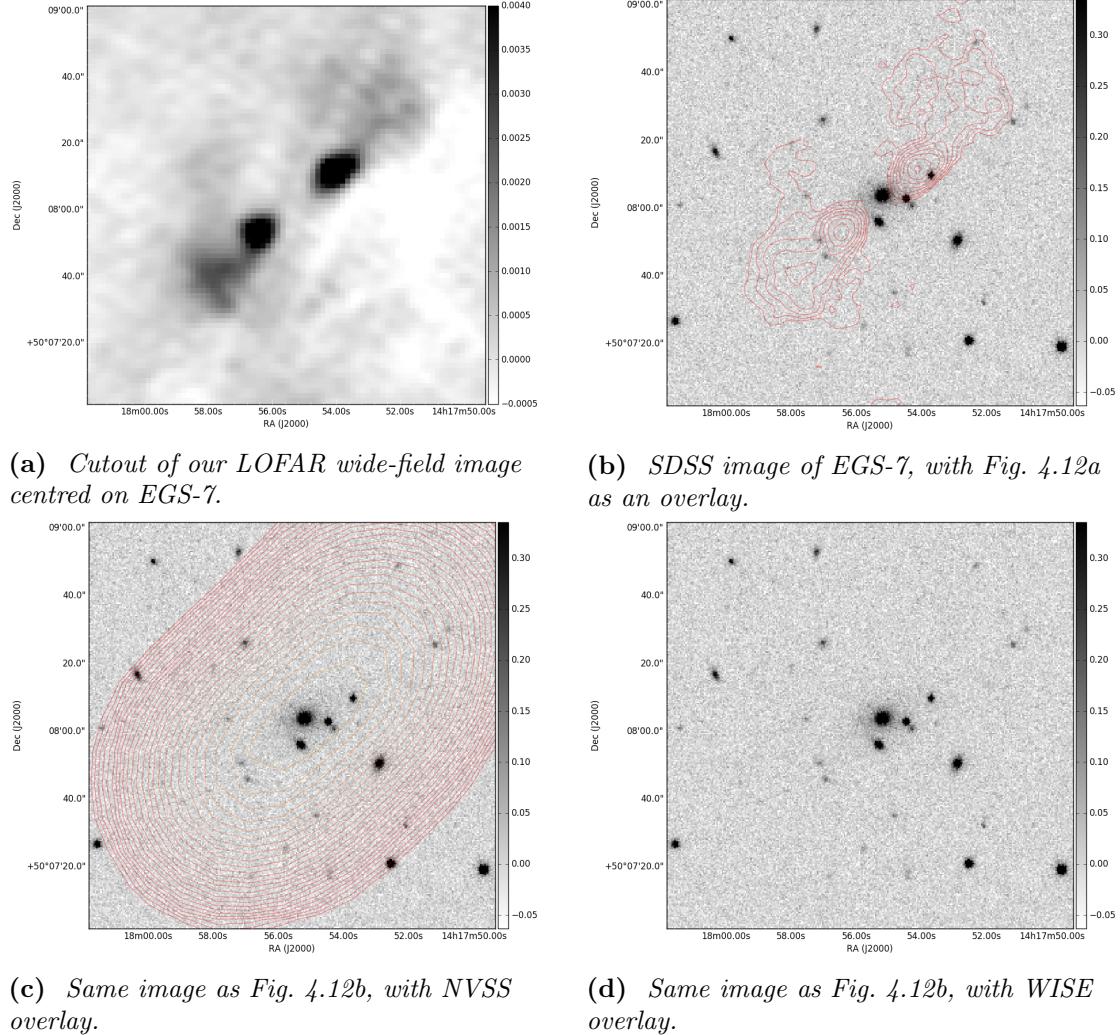
EGS-7

Figure 4.12. Images & Overlays for EGS-7. The overlays start at the 1.5σ level and end on the image emission peak.

This source was chosen because of its very strong lobes and classic shape. It is clearly associated with 2MASX J14175515+5008041, which lies at $z = .186414$. This source was picked up in LBA (van Weeren et al. 2014). Note the resolution improvement over NVSS, even without international LOFAR.

EGS-8

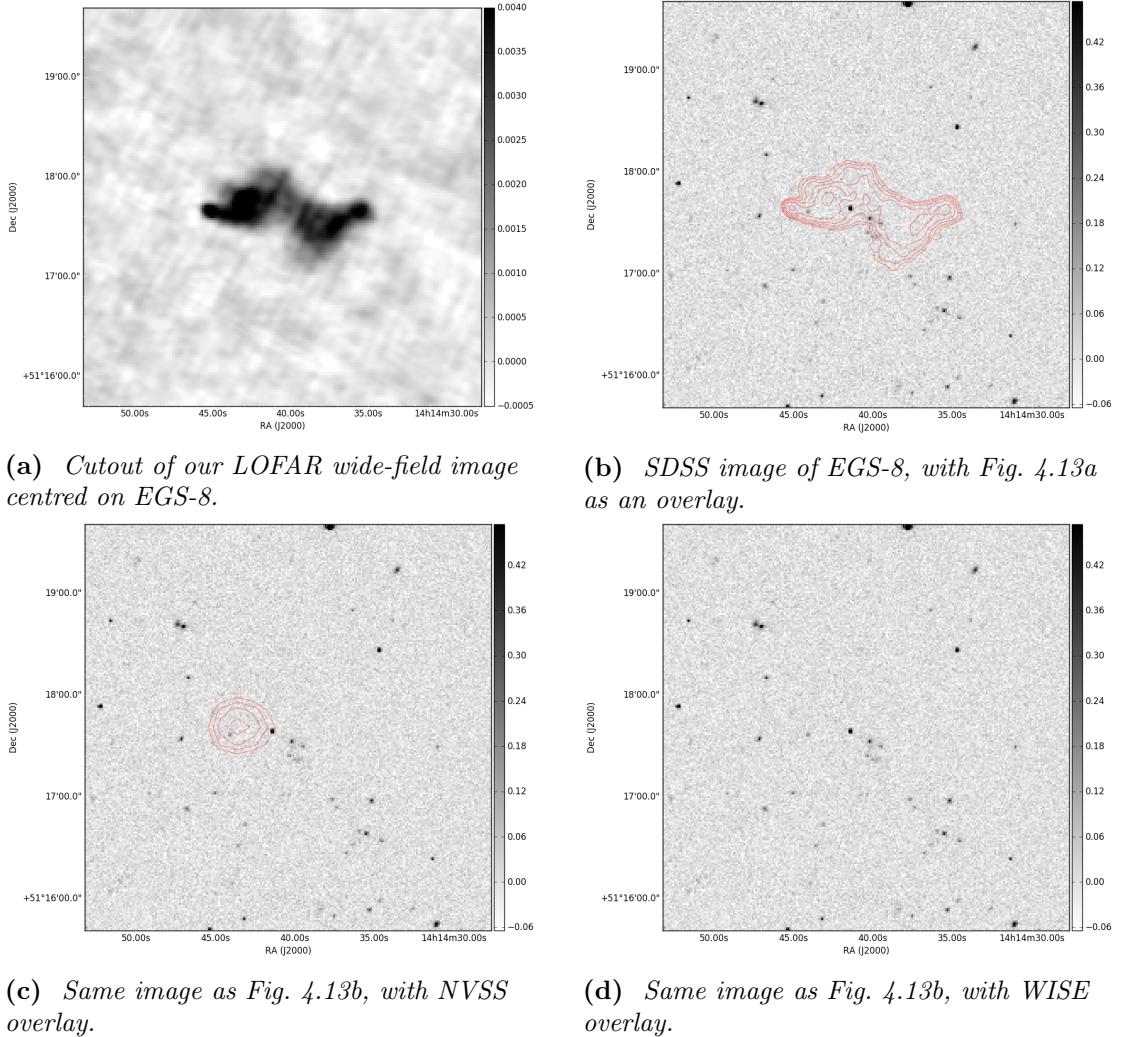


Figure 4.13. Images & Overlays for EGS-8. The overlays start at the 3σ level and end on the image emission peak.

EGS-8 was chosen because it is a complex diffuse source. Note that the source to its South is contaminated by a lot of artefacts in the LOFAR image - but that there definitely is a source there, as there is a NVSS counterpart. There is no WISE counterpart to EGS-8. It has two likely candidates for counterparts: first, J141440.4+511743 (Marecki & Swoboda 2011), a radio galaxy lying a few arcseconds to the North of the centre of EGS-8. Second, SDSSCGB 5990 (McConnachie et al. 2009), a compact group of galaxies, which lies a few arcseconds to the South-East. The redshift of both objects are unknown, so it is unclear whether EGS-8 could represent the interaction between both sources, but J141440.4+511743 is likely the host galaxy for EGS-8's radio lobes.

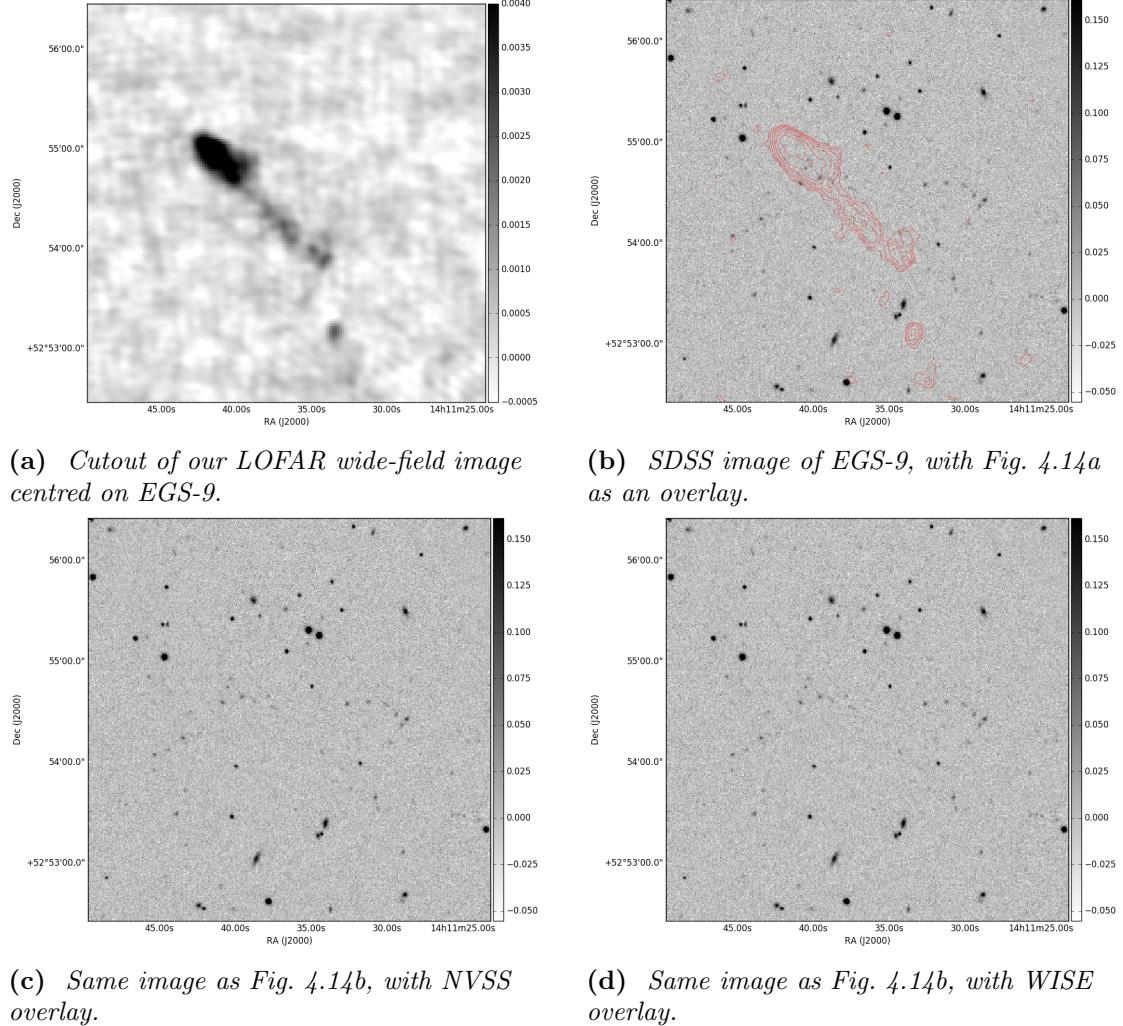
EGS-9

Figure 4.14. Images & Overlays for EGS-9. The overlays start at the 1.5σ level and end on the image emission peak.

This source was chosen because it had one strong lobe and one weak: it was expected to be a good example of a radio galaxy with a single doppler-boosted lobe. There is a likely association to the northern lobe: the galaxy cluster CFHT-W CL J141141.7+525450 (Wen & Han 2011), which lies at $z = .5252$. There is no similarly likely candidate for association with the second lobe, however. It is interesting to note that there is no NVSS equivalent to either lobe.

EGS-10

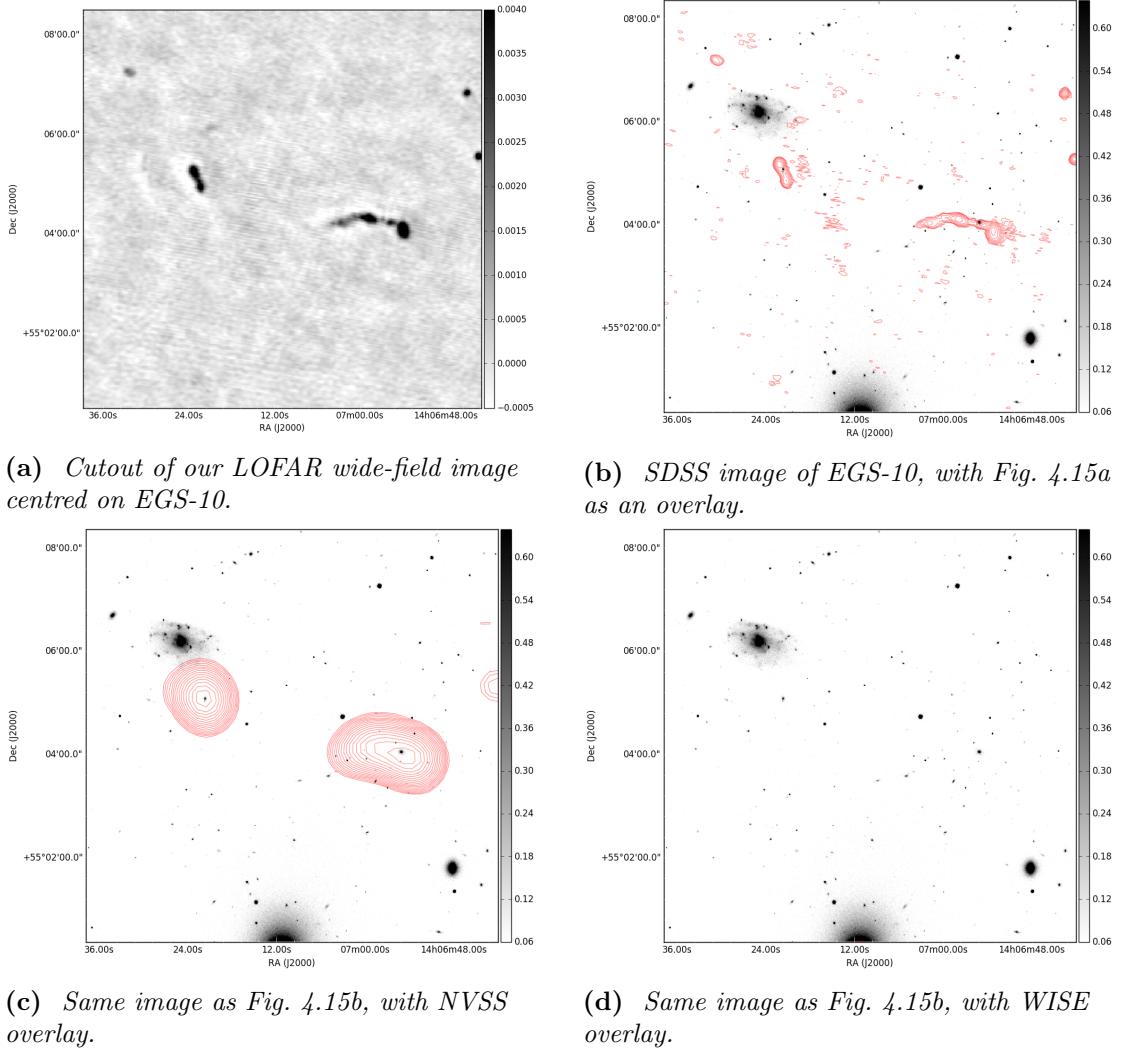


Figure 4.15. Images & Overlays for EGS-10. The overlays start at the 1.8σ level and end on the image emission peak.

EGS-10 was selected because its region includes both interesting diffuse emission (EGS-10-1) and what appears to be a very classic example of a radio galaxy (EGS-10-2). EGS-10-1 has a very strong association candidate in 7C 140513.19+551813.00 (Adelman-McCarthy & et al. 2009), the brightest galaxy in its cluster. That galaxy lies at a redshift of $z = 0.2506(2)$. This association is self-consistent with the association candidates for EGS-3 and EGS-4, which are also very clearly associated with brightest galaxies in their respective galaxy clusters. This seems like a strong indication that such diffuse, turbulent emission is indeed related to the interaction between radio galaxy jets and their surrounding medium. EGS-10-2, meanwhile, is likely associated with SDSS J140721.68+550504.5 (Best & Heckman 2012), which is also a known radio galaxy. This is consistent with the source being picked up in NVSS.

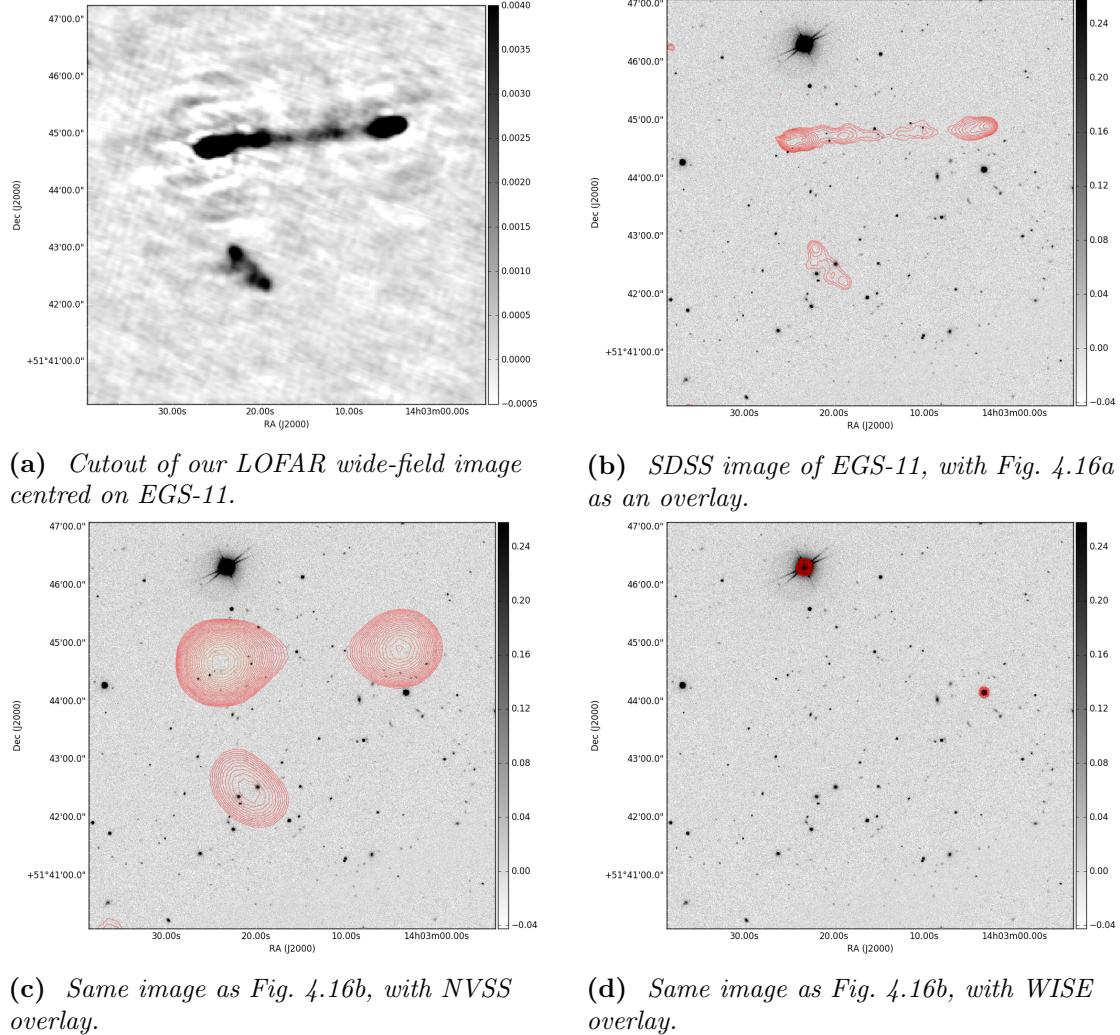
EGS-11

Figure 4.16. Images & Overlays for EGS-11. The overlays start at the 2.5σ level and end on the image emission peak.

This source was chosen because of its size and very straight jets/lobes. While Fig. 4.16a shows that there are some artefacts around this source, the general shape seems to be relatively reliable. Interestingly, we can see that each LOFAR lobe is associated with what appear like individual NVSS sources, while the emission between the lobes in the LOFAR band seems to indicate that the sources are connected. The right-hand lobe (westward) is likely associated to NVSS J140304+514453 (Condon et al. 1998), while the left-hand lobe (eastward) is likely associated with NVSS J140323+514440 (Douglas et al. 1996). Unfortunately, no SDSS counterpart is detected. No counterpart is found at all for the southern source in any band.

EGS-12

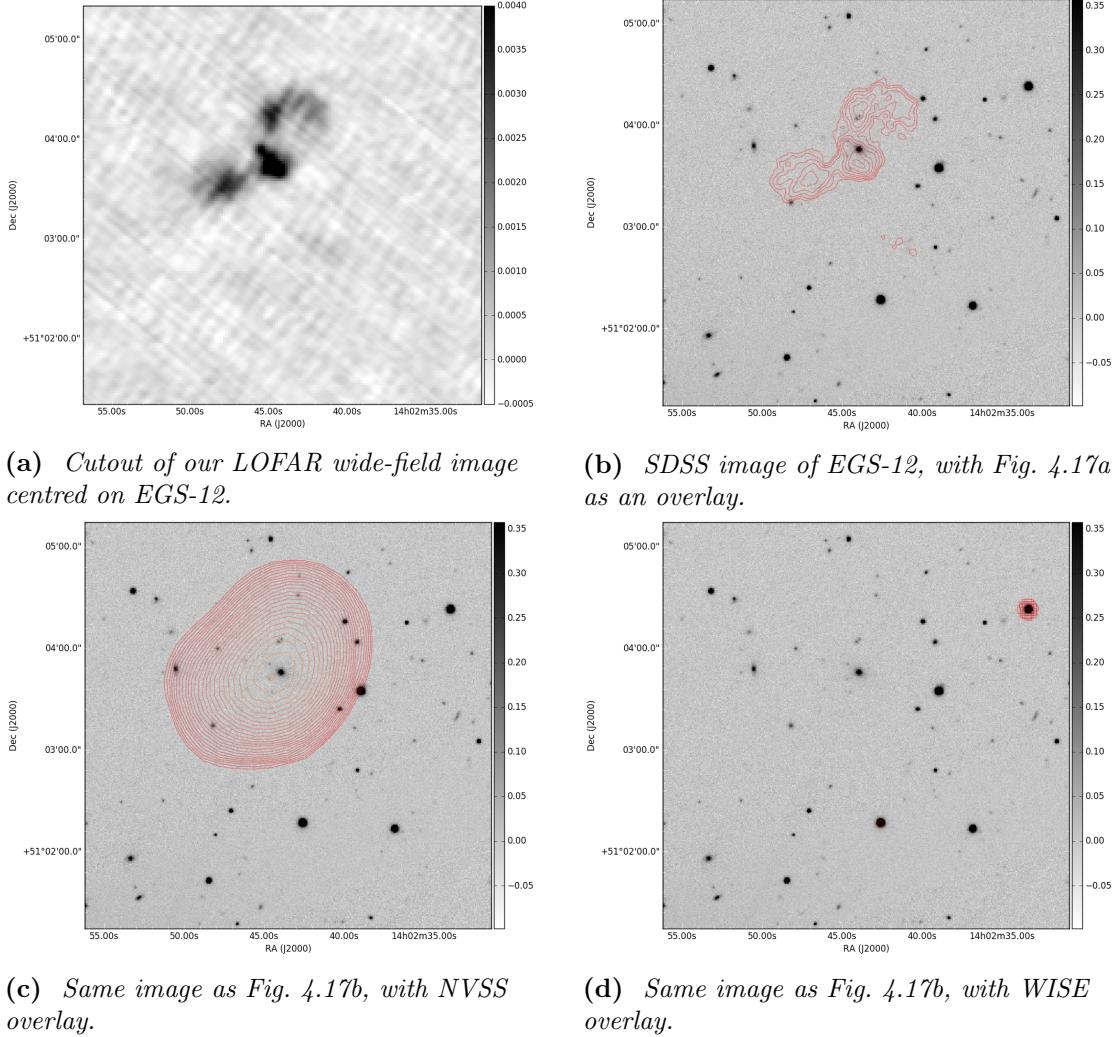


Figure 4.17. Images & Overlays for EGS-12. The overlays start at the 1.5σ level and end on the image emission peak.

This source was chosen for its interesting diffuse emission. Identification was straightforward - assuming that this source is a radio-bright galaxy with jet emission, the central blob of the source, surrounded by two others, must be where its center lies. Within this blob lies a single optical counterpart. This counterpart turns out to be SDSS J140243.96+510345.8 (Adelman-McCarthy & et al. 2009), the brightest galaxy in its cluster. This is consistent with other similar turbulent diffuse structures chosen in this list: EGS-12's diffuse structure is likely a marker of interactions between the jets of a radio-bright galaxy in a cluster and its surrounding inter-galactic and inter-cluster media.

4.5 Discussion

The work done in this section had the primary goal of creating a decent model of the sky seen by LOFAR’s primary beam. To do this, a wide-field, 6"-resolution image of the fully primary beam was made: this image makes no use of the LOFAR international stations due to technical limitations. Using PyBDSM (Mohan & Rafferty 2015), this model was extracted from an image made after a few rounds of direction-dependent self-calibration.

However, the scientific value of this image is not limited to its use in imaging the Extended Groth Strip using international LOFAR. Much of the image will not be recreated using the international stations, in part due to computational constraints (long run time for such a large task) and in part due to technical limitations (performing direction-dependent calibration using international LOFAR would likely require the use of techniques such as fringe-fitting, which are as-of-yet unimplemented in the DDF-kMS pipeline).

Further work on this image is of course both possible and highly desirable: the work done here was artisanal, hand-picking sources and associating them individually to equivalents in other bands. Modern survey work, however, requires automation. Applying Bayesian fitting methods to find most likely counterparts to all the sources extracted in the field could yield useful information, as we would then have access to population statistics. However, doing this would require running yet more self-calibration to remove remaining artefacts, whereas our primary aim is to produce a high-resolution map of the Extended Groth Strip, which has the richest multi-frequency coverage and therefore the most likely to give effective matches to all the sources in the field. We therefore do not proceed to do this work on the low-resolution, wide-field image of the Extended Groth Strip and its neighbourhood, but are likely to return to the data to do this at a later date.

Chapter 5

Imaging 3C295 with International LOFAR

Contents

5.1	Aims & Methodology	105
5.2	Data Reduction	106
5.2.1	Our RIME	106
5.2.2	Calibration Strategy	107
5.2.3	Testing the impact of initial model choice	108
5.2.4	Frequency-dependent Scale Error	110
5.2.5	Examining gain solutions	112
5.3	Overlays & Spectral Analysis	124
5.4	Discussion	127
5.5	Testing Directional Gain Errors - Aims & Methodology	129
5.6	Testing Directional Gain Errors - LBCS Sources in the Primary Beam	132

5.1 Aims & Methodology

Our aim in this section is to create a high-resolution ($0.1''$) model of 3C295, which did not exist at LOFAR HBA frequencies when this project began. This means creating a model that will allow us to find good phase-calibration solutions for LOFAR international baselines, and thus use the international stations to create a high-resolution image of the EGS. Because we only solve for gains in the direction of 3C295, we must then estimate the impact of directional gain errors over the rest of the field.

One major constraint in this project is that 3C295 is fully resolved at these resolutions, which makes deconvolution difficult. As such, we consistently use a *uv*-cut of 100km throughout our imaging in this section, to limit the impact of diffuse emission (which is very hard to deconvolve properly - poor deconvolution introduces artefacts in the restored image) on our images. This choice of *uv*-cut is a compromise between removing diffuse emission in the images and providing sufficiently good conditioning to the calibration

inverse problem¹. Provided that the phases of short-baseline gains do not evolve over the self-calibration procedure, then the information on spatial brightness distribution associated to these spatial fringes is not lost.

Another constraint is that the model needs to be applicable for the full LOFAR bandwidth. We therefore select 6 sub-bands out of the total LOFAR HBA bandwidth, evenly spread throughout the bandwidth as shown in Fig. 5.1. This approach allows us to benefit from the improved conditioning of high-frequency data (since the diffraction limit on angular resolution is proportional to frequency) throughout the LOFAR bandwidth. By starting calibration with a model extracted from a VLA observation of 3C295 with a similar angular resolution, self-calibrating over the initial 6 subbands, then self-calibrating once with 60 subbands (centred on the initial 6) and then the full bandwidth, we expect to achieve good signal-to-noise. Then, we bootstrap the model such that the spectral index of its integrated flux is compatible with Scaife & Heald (2012).

The resulting model is expected to be reliable enough to calibrate our Groth Strip data for the full LOFAR array and the entire HBA bandwidth. It would theoretically be possible to use the direction-dependent gain solutions from Chapter 4 to improve our images, but these are only available for core and remote stations, and not international stations - direction-dependent effects are therefore expected to still arise in the final images due to this. As such, we focus on direction-independent calibration from this point on. Quantifying the impact of the resulting directional gain errors is the subject of Section 5.6.

5.2 Data Reduction

5.2.1 Our RIME

As outlined in Section 2.7, calibration is the process of solving for Jones matrices, which are associated to individual antennas. Here, the only source in our model (the source that dominates, by far, over other sources in the field) is 3C295; we can thus limit ourselves to *direction-independent* calibration, which is equivalent to performing direction-dependent calibration in the direction of 3C295 only. The associated RIME can be written as:

$$\mathbf{V}_{pq}^\nu = \mathbf{G}_p^\nu \left(\sum_{s=3C295} \mathbf{E}_{sp}^\nu \mathbf{K}_{sp}^\nu \mathbf{B} (\mathbf{K}_{sq}^\nu)^H (\mathbf{E}_{sq}^\nu)^H \right) (\mathbf{G}_q^\nu)^H \quad (5.1)$$

$$= \mathbf{G}_p^\nu \mathbf{E}_{3C295,p}^\nu \mathbf{B}_{3C295}^\nu k_{s,pq} (\mathbf{E}_{3C295,q}^\nu \mathbf{G}_q^\nu)^H \quad (5.2)$$

$$= \mathbf{J}_{3C295,p}^\nu \mathbf{B}_{3C295}^\nu k_{s,pq} (\mathbf{J}_{3C295,q}^\nu)^H \quad (5.3)$$

where we mute the time-dependence of all terms, and explicitly model the LOFAR beam during both calibration and imaging, so as to ensure that the true sensitivity of each baseline to this out-of-phase-centre calibrator is taken into account appropriately.

¹The more uncorrupted (i.e. flagged for RFI) visibilities are available, the better the estimation of gain solutions - see Smirnov & Tasse (2015) for the demonstration.

What interests us here is therefore the $\mathbf{J}_{3C295,p}^{\nu}$ for all international stations p and at all times and frequencies t, ν . We improve conditioning by using time and frequency solution intervals of a minute and 8 subbands; at those times where this interval is too large, the quality-based weighting scheme is expected to compensate. Choosing such large intervals also helps self-calibration converge faster. We do not explicitly resort to physics-based solutions to constrain the calibration solutions: we solve for the full 8 parameters of each complex 2×2 Jones matrix (2 parameters per complex quantity, 4 complex quantities).

5.2.2 Calibration Strategy

We self-calibrate over the full LOFAR bandwidth so that the spectral flux distribution (i.e. where the flux of 3C295 lies at different frequencies) can be recovered in greater depth as we introduce more and more of the full bandwidth. We do not begin self-calibration with the full bandwidth, as a single pass of self-calibration with 6 subbands takes about an hour and a half, whereas it takes a full 3 days with the full bandwidth. The full bandwidth is therefore only self-calibrated on once, at the very end. We begin our self-calibration with 6 subbands, chosen across the LOFAR bandwidth, calibrated simultaneously. The chosen subbands, and their position in the total bandwidth, are shown here:

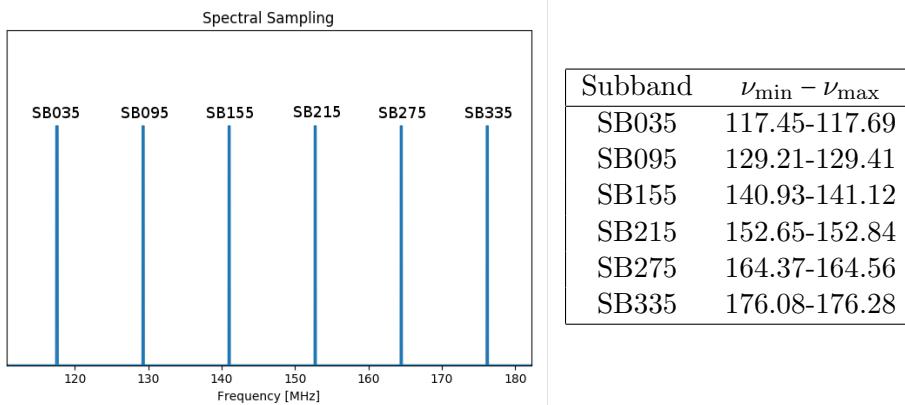


Figure 5.1. Position of subbands chosen across full LOFAR bandwidth

Table 5.1. Frequency bounds for the subbands chosen

Because no high-resolution models exist for 3C295 at our observing frequencies, which span quite a large bandwidth, we start by making a model from a high-resolution image taken at 8.7GHz, shown in Fig. 5.2. We start with a model which is as close to the expected “true” brightness distribution because the self-calibration problem is not convex: while the calibration inverse problem itself is convex (i.e. we expect to converge towards the “true” gains regardless of the initial conditions, given enough SNR), the degeneracy introduced by imaging in the self-calibration procedure means that one could well get stuck in a local minimum.

Because we apply a high-pass *uv*-cut of 1km when calibrating, and a *uv*-cut of 100km when imaging², we need to make sure that we do not introduce spurious structure in the core and remote station gains. Similarly, we can check the quality of our self-calibration by looking at the structure of the international station gains. Gain structure can consist of genuine atmospheric & instrumental effects, but can also include unmodeled flux being absorbed into the gain solutions. At first, this is unavoidable: if flux is resolved by a baseline but not included in the model, then it will be absorbed into the gains. As we improve our model, we expect the gain curves to flatten and become smoother. If this happens for our international station gains, then we have a strong indication that we are in fact recovering physical signal rather than adding artefacts into our model. Similarly, we expect core and remote station gains to remain mostly static: if this is the case, then our strategy does not bias our gain solutions for these stations away from the underlying truth.

5.2.3 Testing the impact of initial model choice

We acquire the initial calibration model by extracting features from a NASA/IPAC Extragalactic Database³ image of 3C295 (see Perley & Taylor 1991), shown in Fig. 5.2. This feature extraction is carried out using PyBDSM (Mohan & Rafferty 2015), changing all extracted Gaussians into points⁴.

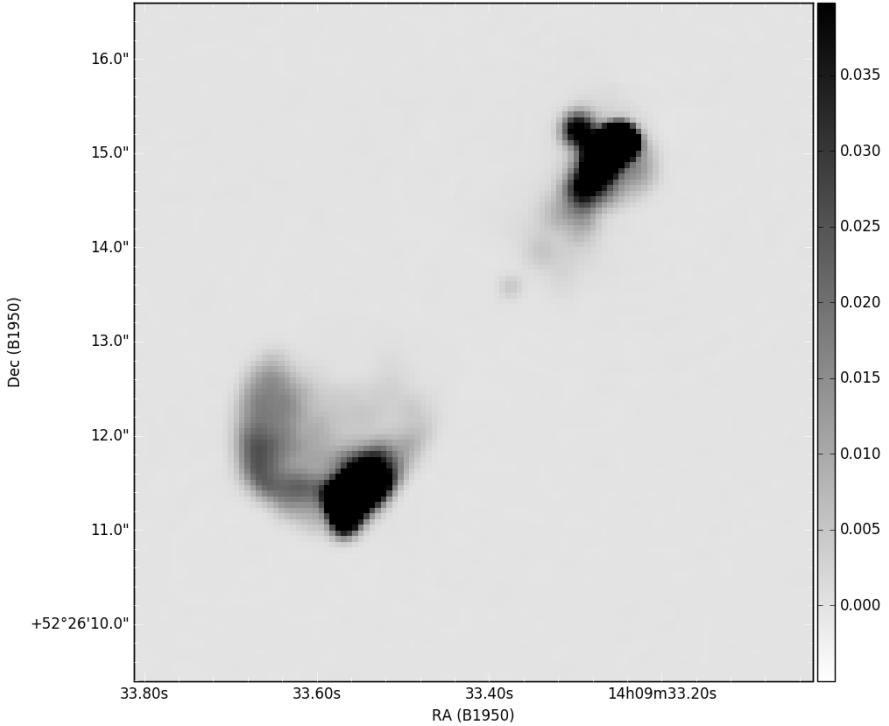


Figure 5.2. VLA observation of 3C295 at 8.7 GHz. Pixel size is $0.2''$.

²i.e. all baselines shorter than 100km are discarded.

³<https://ned.ipac.caltech.edu/>

⁴This practice ensures that wrongly-estimated Gaussians do not end up introducing unphysical bias during calibration.

A good initial test of our hypothesis that we can start from a “good enough” point to reach a global minimum is to self-calibrate on a two-point model. We perform this test, keeping only the two brightest pixels of the model shown above to test whether we converge to the same structure. The result is shown as an overlay in Fig. 5.3: as we can see, we recover the same structure as the full VLA model. We appear to have introduced a small astrometric error, however: this is likely an indicator that using only the peak flux, rather than the full distribution, will bias the “centre of flux” away from its true position – hence introducing an astrometric shift. We therefore start from the full model for our “proper” self-calibration procedure.

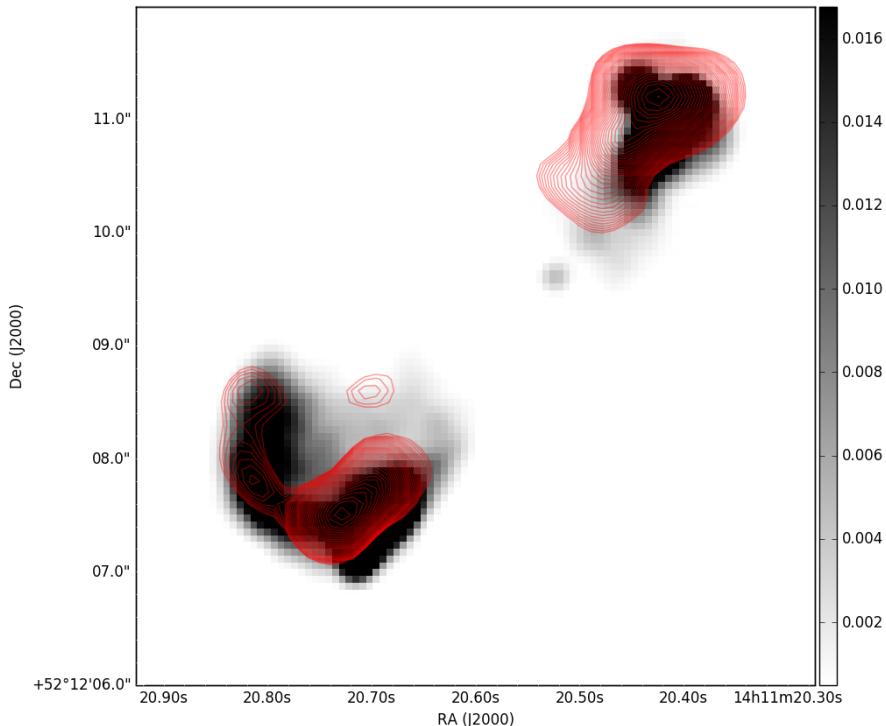


Figure 5.3. Overlay of the model resulting from a self-calibration process starting from a two-point model. The overlays rise exponentially, from the 5σ level to the maximum pixel value in the image. This image is very similar to the LOFAR image made with a single pass of calibration using the full VLA model.

The only exhaustive method to test our hypothesis would be to perform a full Monte Carlo analysis, testing every possible flux distribution for every pixel at every frequency to ensure that we do, in fact, find the model and gains which give the global minimum fit, rather than a local minimum fit. This would be extremely time-consuming, however, and falls outside the scope of this work. We therefore simply try to start as close as possible to where we expect the global minimum to fall, using the full VLA image as a starting model.

5.2.4 Frequency-dependent Scale Error

We begin by using a simple spectral index of $\alpha = -0.8$, which is typical of synchrotron sources at these frequencies. It is not strictly correct, but gives a good first-order approximation of the integrated flux of 3C295 as a function of frequency, which is shown in Fig. 5.4.

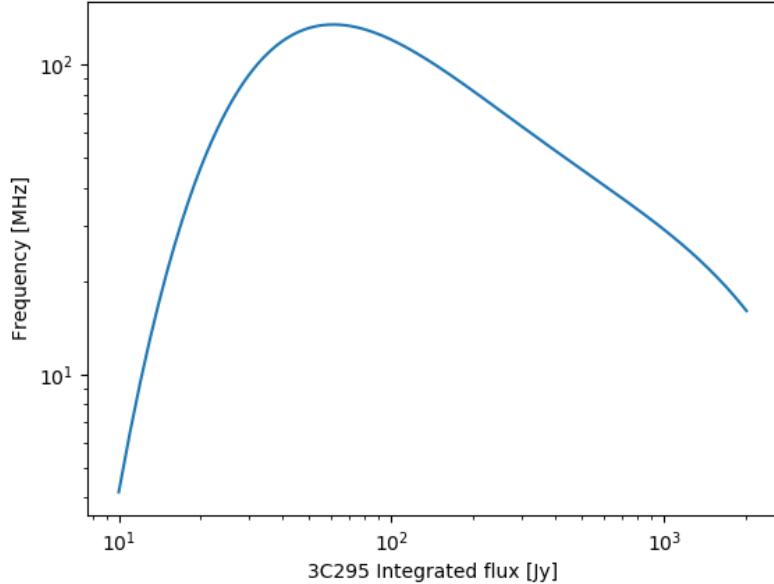


Figure 5.4. Integrated flux of 3C295 as a function of frequency, as fitted by (Scaife & Heald 2012).

In so doing, we introduce a frequency-dependent error factor in our gain solutions due to the fact that the absolute flux scale isn't known *a priori*:

$$\mathbf{V}_{pq}^\nu = \mathbf{J}_{s,p}^\nu \mathbf{B} (\mathbf{J}_{s,q}^\nu)^H \quad (5.4)$$

$$= \mathbf{J}_{s,p} \sqrt{\alpha_\nu} \frac{\mathbf{B}}{\alpha_\nu} (\mathbf{J}_{s,q}^\nu \sqrt{\alpha_\nu})^H \quad (5.5)$$

$$= \hat{\mathbf{J}}_{3C295,p}^\nu \frac{\mathbf{B}}{\alpha_\nu} (\hat{\mathbf{J}}_{s,q}^\nu)^H \quad (5.6)$$

$$\hat{\mathbf{J}}_{3C295,p}^\nu = \sqrt{\alpha_\nu} \mathbf{J}_{3C295,p}^\nu \quad (5.7)$$

where $\hat{\mathbf{J}}_{3C295,p}^\nu$ is our estimation of the Jones matrix $\mathbf{J}_{3C295,p}^\nu$ for time t and frequency ν , and α_ν is a smoothly-varying scalar function of frequency alone. This then introduces a frequency-dependent error factor $1/\sqrt{\alpha_\nu}$ in our visibilities and therefore in our images. This factor is a constant for all visibilities at a given frequency.

Because we solve for the full Jones matrix, we cannot apply phase-only corrections⁵ over multiple self-calibration rounds, as is standard practice. Over self-calibration

⁵Indeed, what is the phase of a 2×2 matrix of complex numbers?

rounds, we expect that α will change, but that it will remain smooth in frequency. The integrated flux of 3C295 over different passes of self-calibration is shown in Fig. 5.5: as this figure shows, the model amplitude can diverge significantly over the self-calibration process.

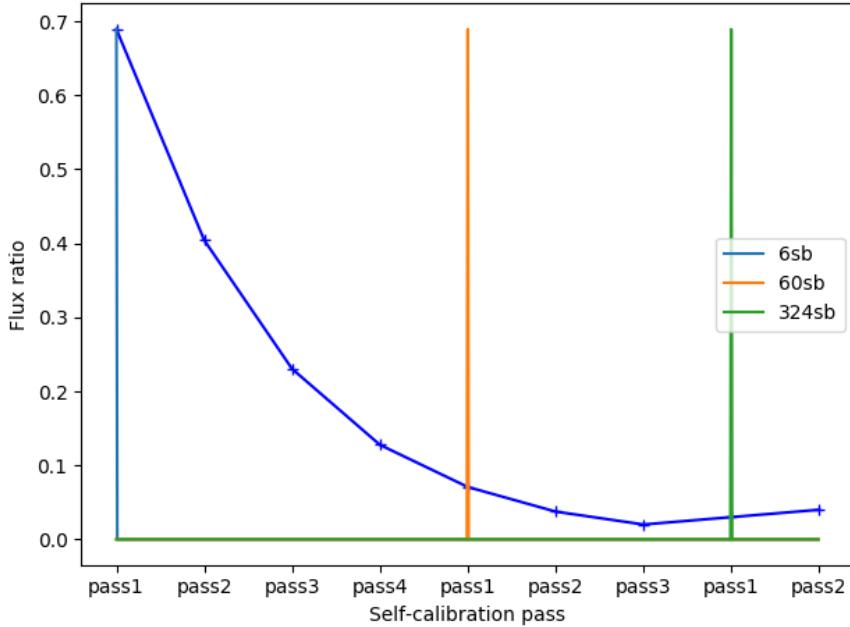


Figure 5.5. Amplitude fluctuation of the models resulting from self-calibration iterations, expressed as a ratio of the model’s integrated flux divided by the Scaife & Heald integrated flux model at the appropriate frequency. Note that the integrated model flux for 324sb-pass1 (the penultimate data point in this plot) was lost due to a careless command: the value shown here is just the average of the two neighbouring values - i.e. those of 60sb-pass3 and 324sb-pass2.

Because α is smooth in frequency, we can apply a frequency-dependent amplitude correction (i.e. “bootstrap”) by relying on the work of Scaife & Heald (see Scaife & Heald 2012). This correction will be done in the imaging plane, since α is a function of uv -position (but not sky position). This is, again, because 3C295 does not lie at phase centre: if it did, the correction could be done equivalently in visibility or image space. By performing it in the image-plane, we rid ourselves of the uv -dependence. The resulting signal-to-noise improvement is shown in Fig. 5.6. It is calculated on the bootstrapped residual images resulting from each round of the self-calibration process - as such, the only thing that changes is the noise, since the signal is corrected to be the appropriate constant. As we can see, the noise falls nicely with self-calibration passes. Tragically, the results from the first full-bandwidth calibration run with the 60-subband model on the full 324 good subbands were lost before this plot could be made, and so only the results from the second full-bandwidth self-calibration pass remain. We can see that it is a small improvement over the 60-subband model.

The large decrease (factor of ~ 2) in RMS between the last 6-subband residual image and the first 60-subband residual image in Fig. 5.6 is significant: we would expect a

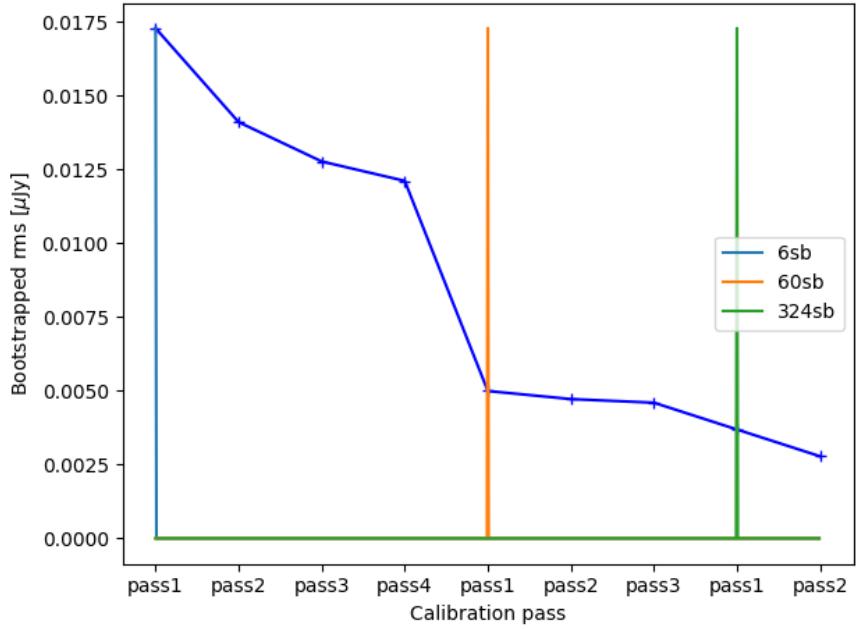


Figure 5.6. Bootstrapped residual RMS for each calibration pass. Note that the data for 324sb-pass1 (the penultimate data in this plot) were lost due to a careless command: the value shown here is just the average of the two neighbouring values.

similar jump when going from 60 to 324 subbands. The fact that we do not is a strong indicator that we are digging into the limits of what our strategy can achieve by the end of the second full-bandwidth self-calibration pass: we are stuck in a local minimum. This is likely due to the fact that our initial sky model (and all subsequent models) only include 3C295: the fact that all the other sources in the primary beam are not included in the model is bound to introduce bias our gain solutions and artefacts in our images.

We want good phase and amplitude gain estimates without concern for the overall scaling factor: this means that we want to find the correct brightness distribution at a given frequency, knowing that the overall scaling factor will be wrong, and correcting the images accordingly afterwards.

5.2.5 Examining gain solutions

Gains for 6 antennas at a given frequency

Let us begin by checking whether our core and remote station gains change over time due to our self-calibration procedure. Here, we look at the gains of SB155, on which we solved for gain solutions using our initial and final models. This subband was included in both the 60-subband self-calibration step and the 6-subband self-calibration step.

As we can see from Fig. 5.7, there is little change in amplitude structure between the first pass of calibration (with the model extracted from Fig. 5.2) and the last pass of calibration (after the successive self-calibration passes). Similarly, there is little apparent change in phase structure (note that the phase wraps from $-\pi$ to π). It is important

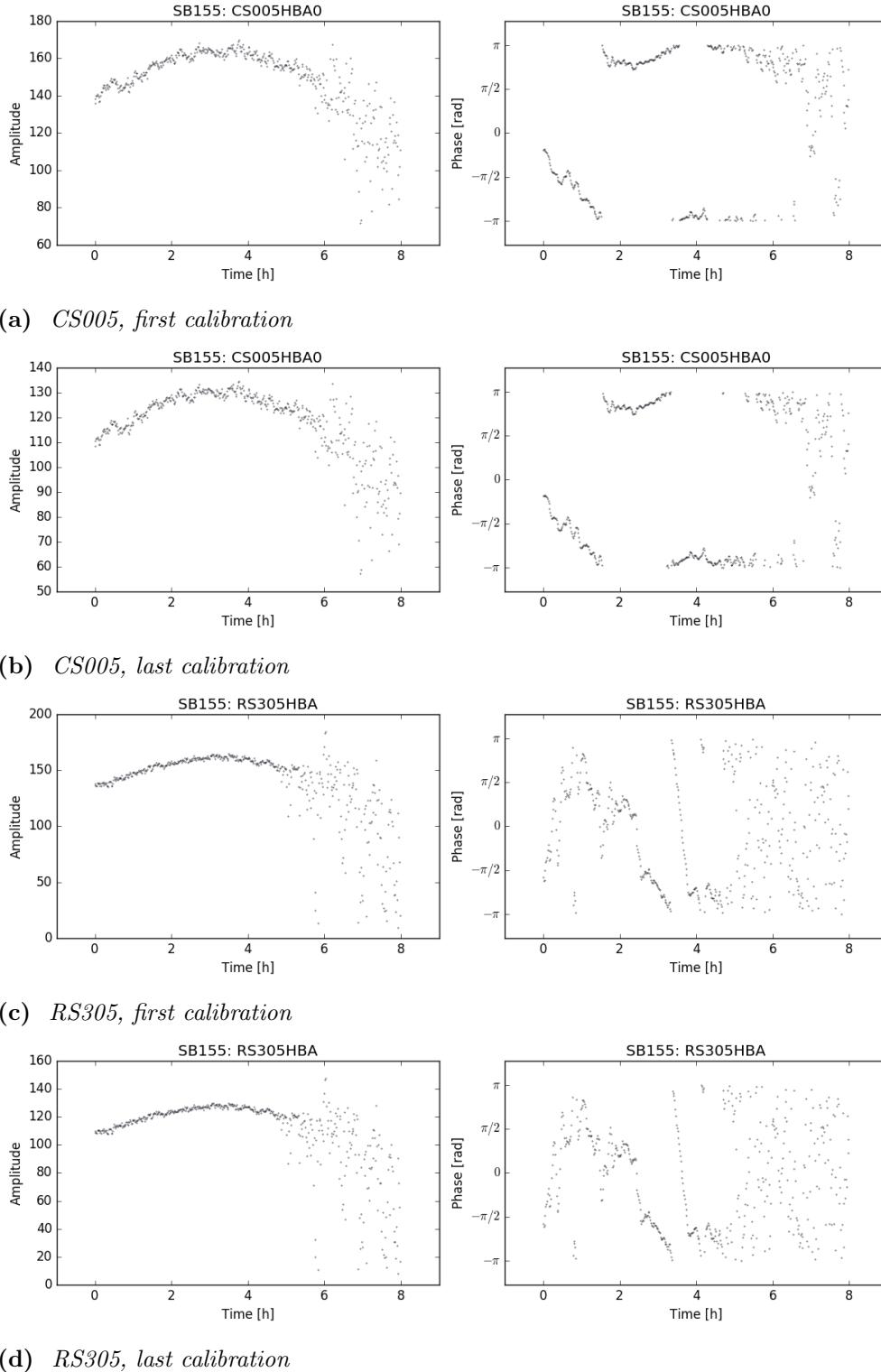


Figure 5.7. Gain amplitude & phase for CS201 of SB155

to note, however, that there is a difference in the *level* of the amplitude curve: in the first pass of calibration, the average amplitude value is generally higher than in the last

pass of calibration. This is expected behaviour, as there is less diffuse emission in the *uv*-cut imaging model of 3C295 than in the initial calibration model. This is equivalent to applying a choice of weighting scheme which down-weights diffuse emission (i.e. a form of uniform weighting).

As for the international station gains, they are shown below for three of the seven international stations used for our observation, in Figs. 5.8 to 5.10. We show only one of the four German stations. As we can see, while structure very clearly remains in some of these amplitude curves (notable in the last two hours, e.g. Fig. 5.8b), some of the structure disappears from these amplitude curves. It should be noted that the last two hours of the observation are generally much noisier than the first 6; the last quarter of these plots show this. Quality-based weighting schemes were found to compensate for this effect quite well, and minimise the associated image deterioration⁶. Note that the phase wraps extremely fast in the international station gain plots: this is because phase is calculated relative to CS001HBA0, the first antenna in our array. Because the international stations are much farther away from this core station than other core or remote stations are, we expect the associated phase solutions to wrap faster.

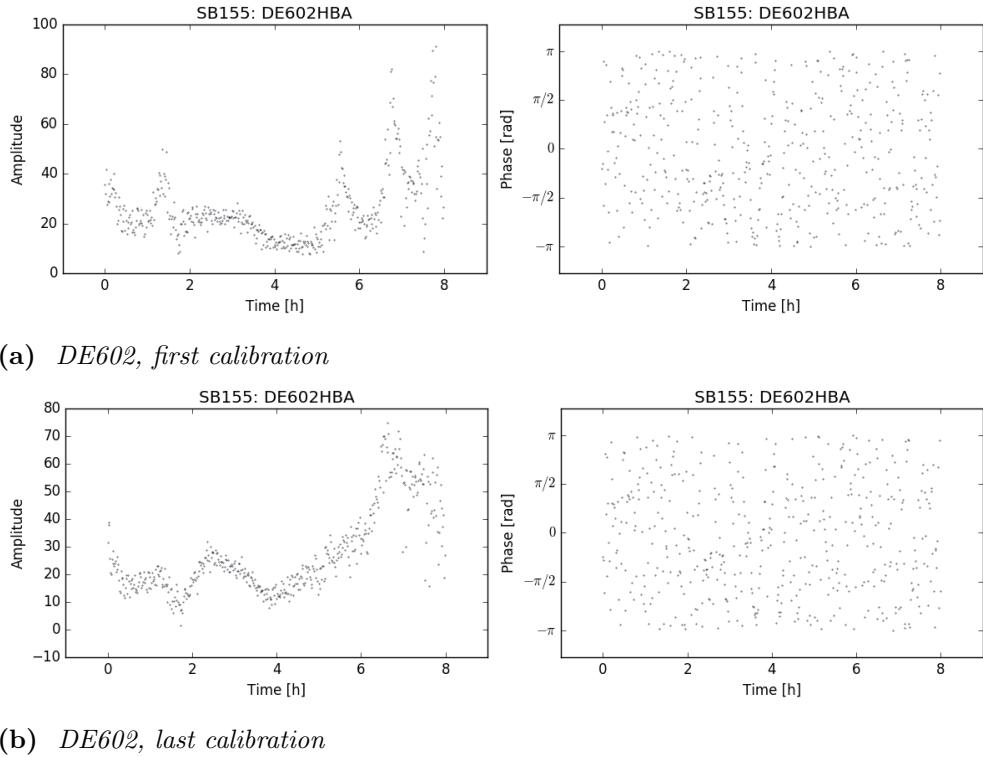


Figure 5.8. Gain amplitude & phase for DE602 of SB155

⁶The evidence for this is in the previous chapter, as the test data is the same as used here; nevertheless, an extra test was carried out in the context of these calibration and imaging tests for the sake of completeness, and found that quality-based weighting schemes resulted in better results than simply flagging the last 2 hours.

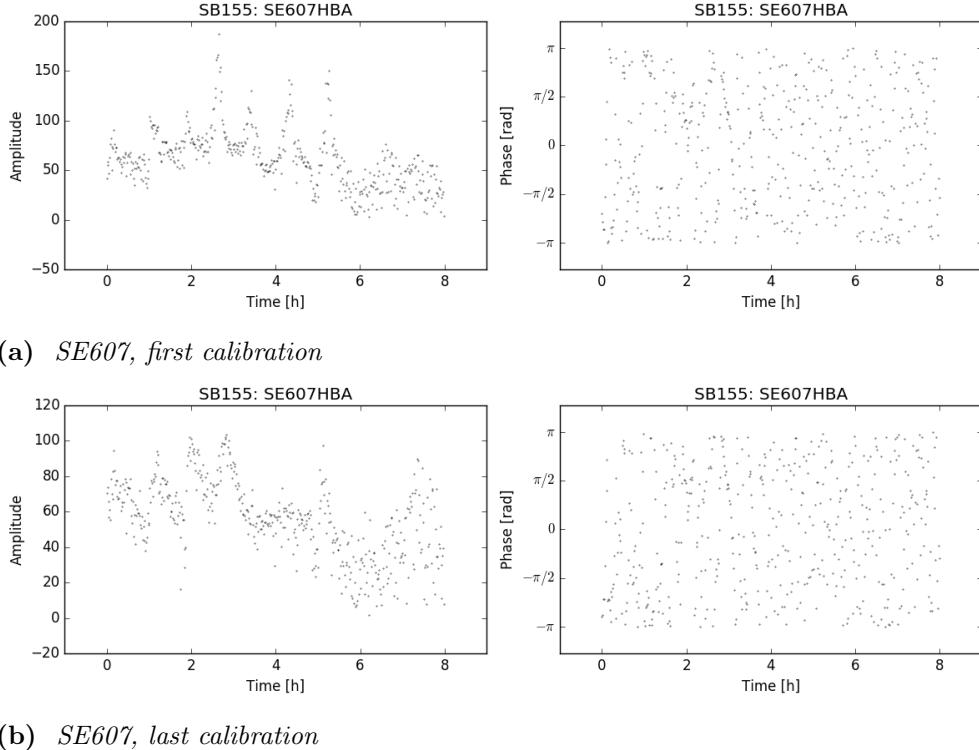


Figure 5.9. Gain amplitude & phase for SE607, an international stations of SB155

Spectral behaviour of gain solutions

So far, we have shown the gain solutions for one subband, out of a total of 324 used subbands. The spectral behaviour of the amplitude and phase solutions can give us insight into whether we are absorbing true flux or calibration artefacts into our model: if our calibration model becomes more accurate to the true underlying flux, we expect e.g. the phase of core station calibration solutions to remain the same (as we are improving high-resolution details of our model) while our international station calibration solutions would become flatter in amplitude. Similarly, we expect to see large-scale frequency-dependent behaviour in the phase solutions of international stations, with smaller fluctuations around this behaviour: the effect of differential TEC, ionospheric turbulence, etc. Indeed, while the core stations (and, to a lesser extent, the remote stations) can be approximated as having the “same ionosphere” above them, this is not the case for international stations: this is a major source of the difficulties in calibrating these stations. As such, for international stations, we expect to see complex structure in the phase gain solutions, characteristic of the differences between the local ionosphere above the core and the international stations. For core stations, we only expect to see structure characteristic of scintillation (regular stripes with no frequency-dependent structure) in the phase gain solutions⁷.

⁷This approach was taken due to time constraints: verifying that our model was compatible with the closure phases and amplitudes would have been a better test to perform, but would also likely have taken much longer to be fruitfully interpreted.

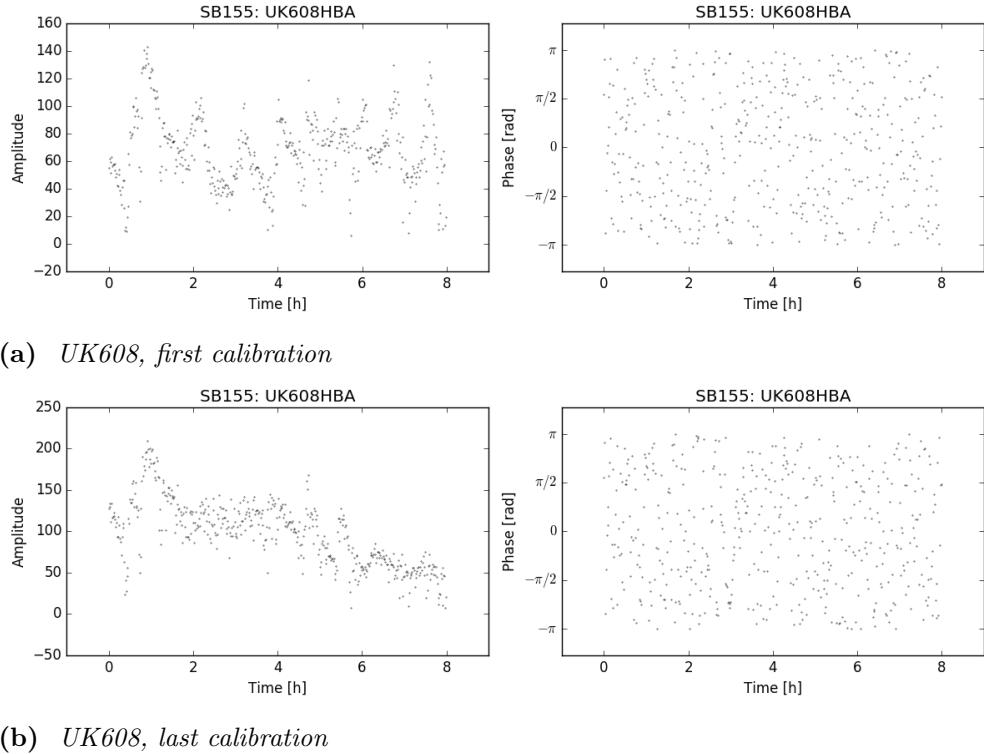


Figure 5.10. Gain amplitude & phase for UK608HBA, an international stations of SB155

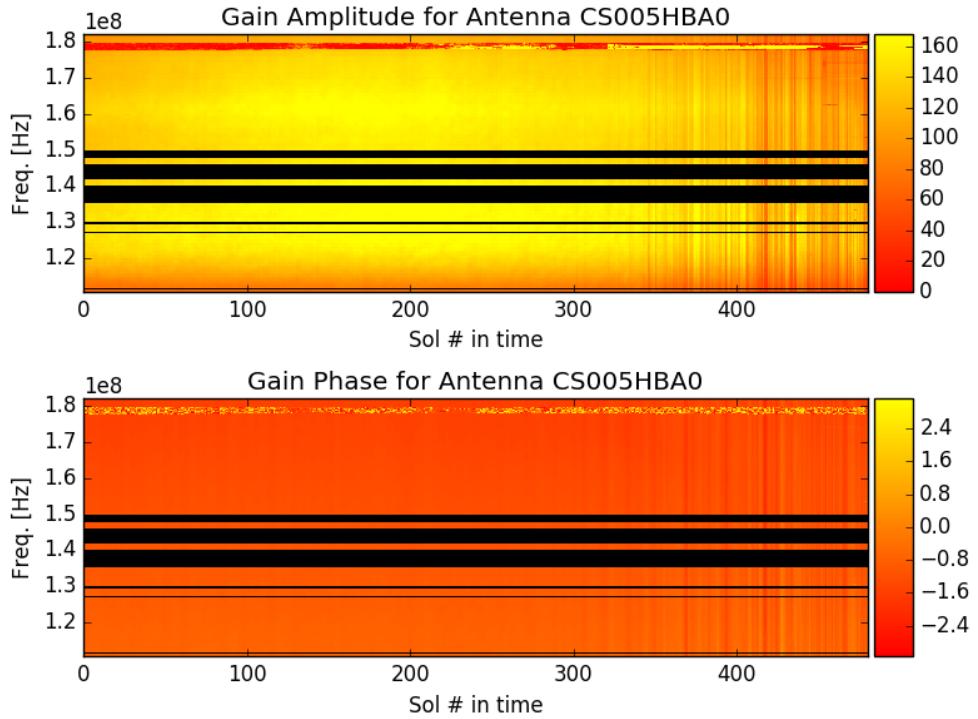
Figs. 5.11 and 5.12 show the gain solution phases and amplitudes for all times and frequencies in our observation. Some patches of the plot are empty (horizontal black bands): this is because some subbands had very poor data, and their use in self-calibration corrupted the final image. They are more numerous for the first calibration pass than the last, as the solutions were not sought for some cases.

As we can see, the behaviour of both phase and amplitude gains remains much the same for core stations, before and after self-calibration. This is a good sign: it means that we do not lose the large-scale structure in our observation as we self-calibrate. We see in Fig. 5.13 that remote stations have much more phase structure as a function of frequency. This structure shows that the ionosphere above the remote Dutch stations is different from that above the core stations: this is an expected result. The fact that the structure does not change significantly before and after the self-calibration run could indicate that the remote station gains, like the core stations gains, do not absorb true flux into their gain solutions during the self-calibration process. It can thus be taken as a sign that our self-calibration is trustworthy for these stations.

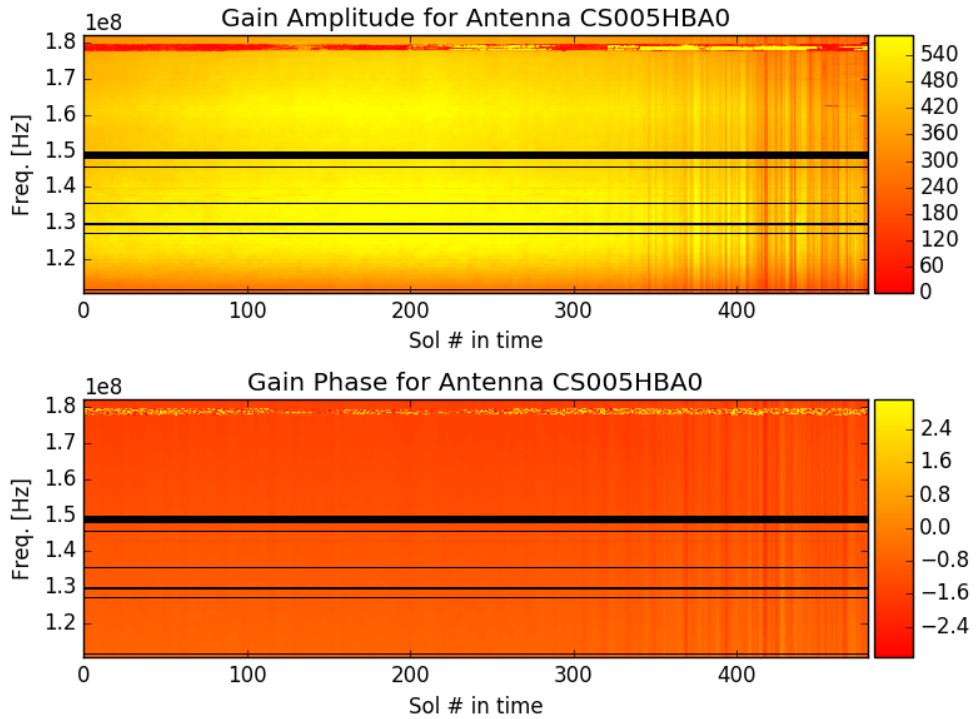
Finally, we examine the gain solutions for international stations. This is where we expect to see the most significant amplitude and phase structure, but also where we expect gain structure to change between the start and end of the self-calibration process. If this is the case, it would mean that the unmodeled astrophysical signal that is initially absorbed into international station gain solutions is introduced into the model over multiple iterations. Notably, we expect a big change in amplitude solutions; phase

solutions are harder to read, as we expect the ionosphere to be drastically different between core and international stations. As such, they will include both true phase structure and the contribution of unmodeled flux.

We see that some of the structure in the amplitude gains shown in Fig. 5.14 goes away over the course of self-calibration, though one of the two arcs present in the plot appears to decrease in amplitude. For Fig. 5.15, we see that the pattern in the amplitude plot clearly changes - the fact that the pattern in Fig. 5.15a does not appear to be either dispersion or scintillation pattern would indicate that it consists of unmodeled flux. While some structure remains in the last 2 hours and low frequencies for this antenna's gain solution amplitudes, the last 2 hours of the observation are generally poor compared to the first six: we expect that the quality-based weighting scheme will minimise the impact of these poor calibration solutions on the final images. In Fig. 5.16, we see that the self-calibration process was clearly less effective than it was for DE602HBA: while some of the structure was minimised (in the first two hours particularly: we go from the peculiar "ellipse" amplitude pattern to something which looks more like scintillation), a lot clearly remains. Our self-calibration process could therefore likely be taken further. However, in our last step, we use 324 out of the full 366 subbands (leaving some out due to poor quality); a single round of self-calibration with this full dataset takes 5000 minutes on the best computing cluster available for this work. Because a single round of self-calibration would, at this point, take as much time as imaging a portion of the EGS due to computational constraints, we treat the self-calibration process as "having converged" at this point: while it could be yet improved, we have reached a noise of 2.78 mJy.



(a) *CS005HBA0, first calibration, full bandwidth*



(b) *CS005HBA0, last calibration, full bandwidth*

Figure 5.11. Plots showing the amplitude and phase solutions for CS005HBA0 at the start and end of self-calibration, respectively. The black areas correspond to regions which were not used for that pass of calibration due to poor data.

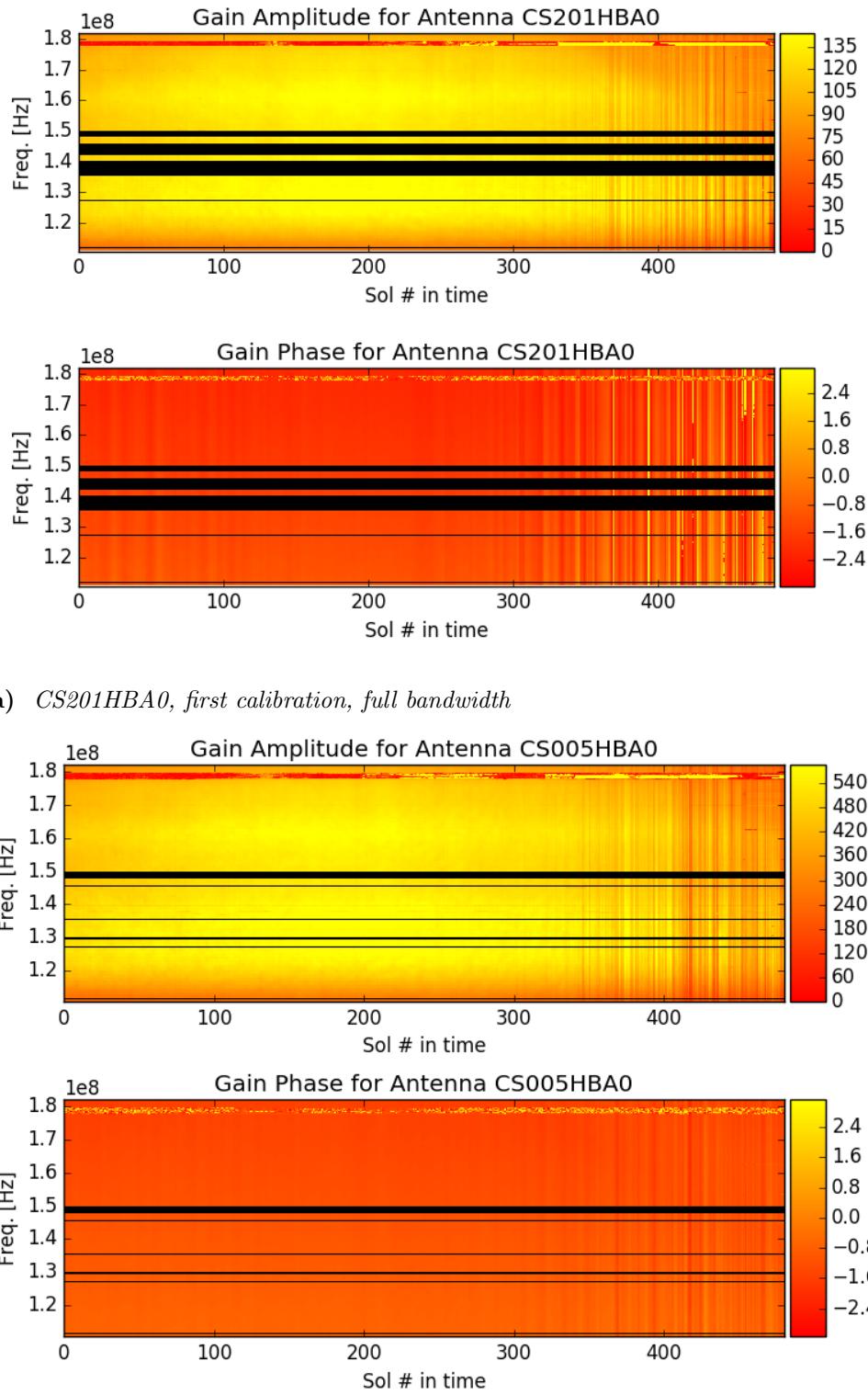
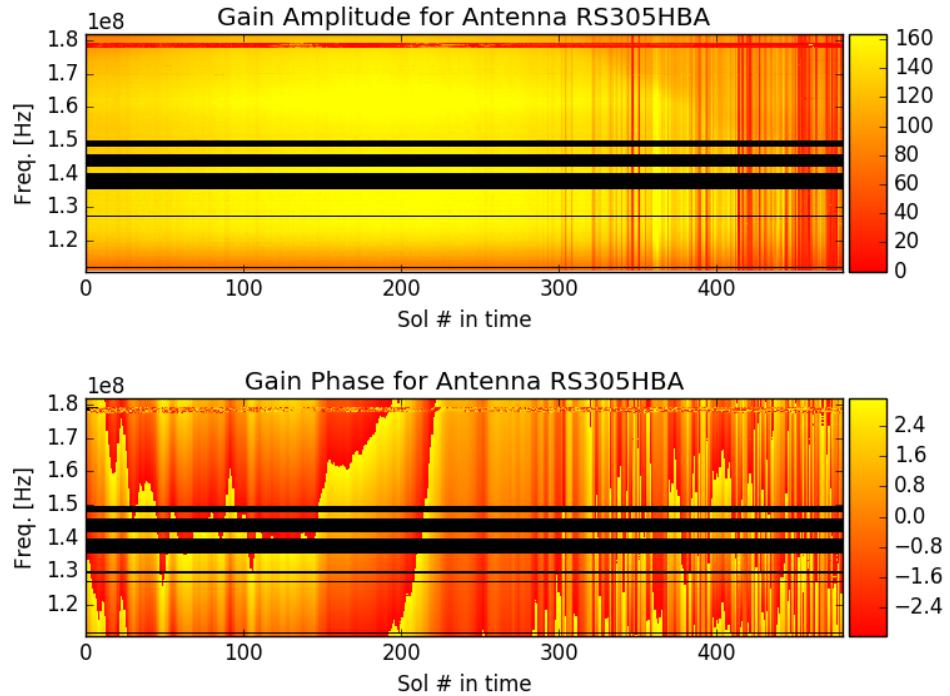
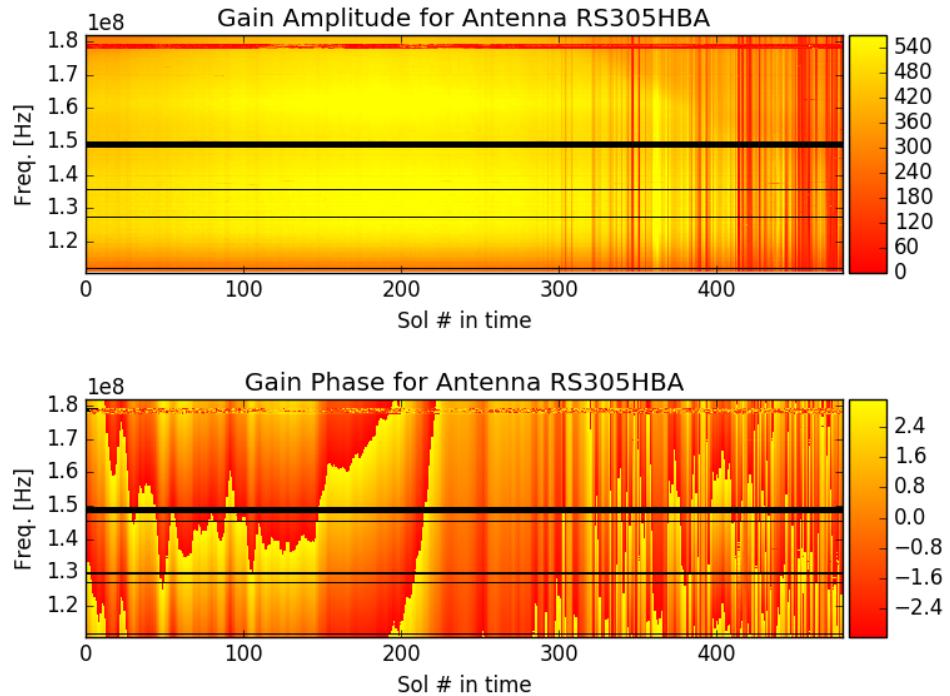


Figure 5.12. Plots showing the amplitude and phase solutions for CS201HBA0 at the start and end of self-calibration, respectively. The black areas correspond to regions which were not used for that pass of calibration due to poor data.



(a) *RS305HBA, first calibration, full bandwidth*



(b) *RS305HBA, last calibration, full bandwidth*

Figure 5.13. Plots showing the amplitude and phase solutions for RS305HBA at the start and end of self-calibration, respectively. The black areas correspond to regions which were not used for that pass of calibration due to poor data.

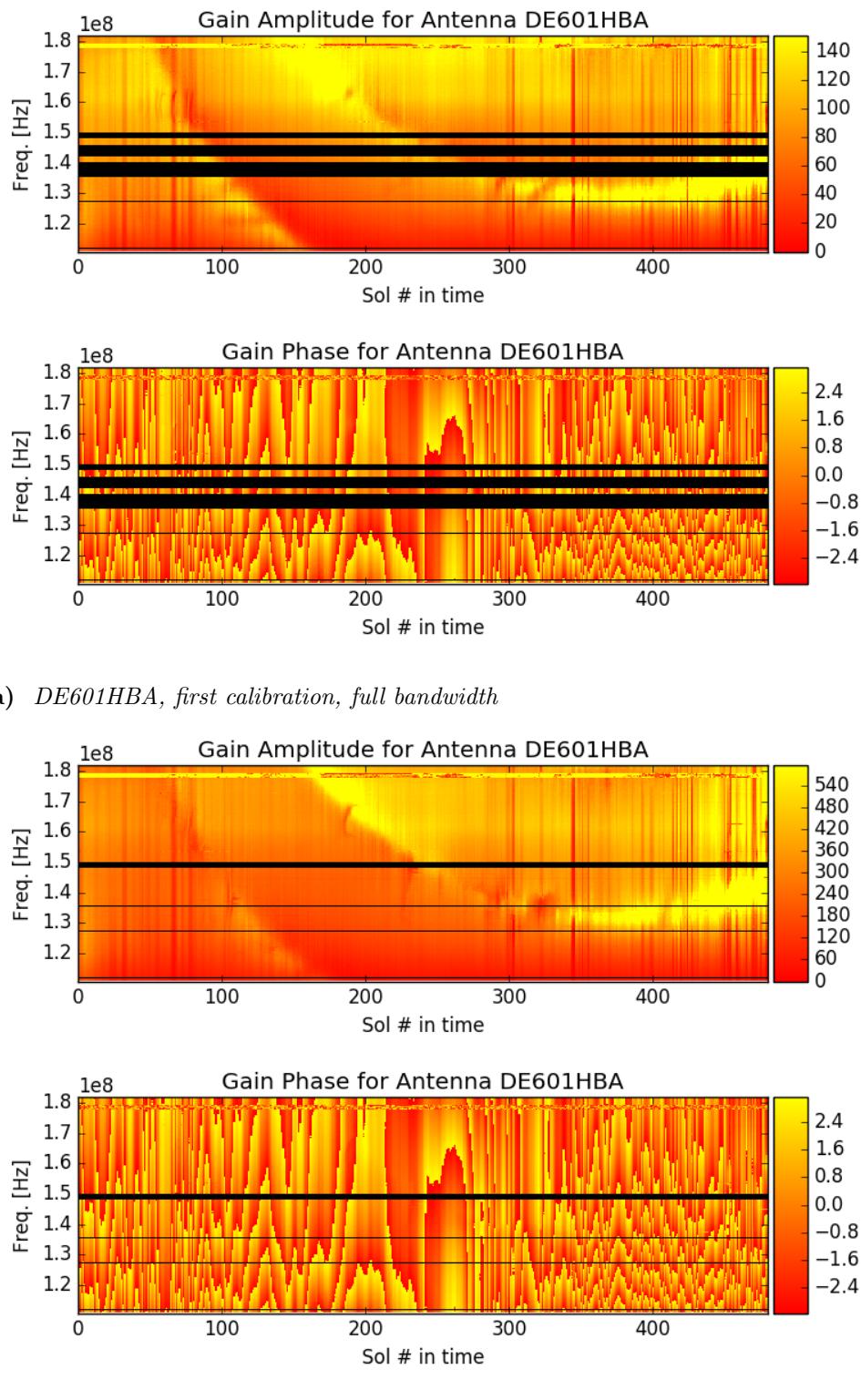
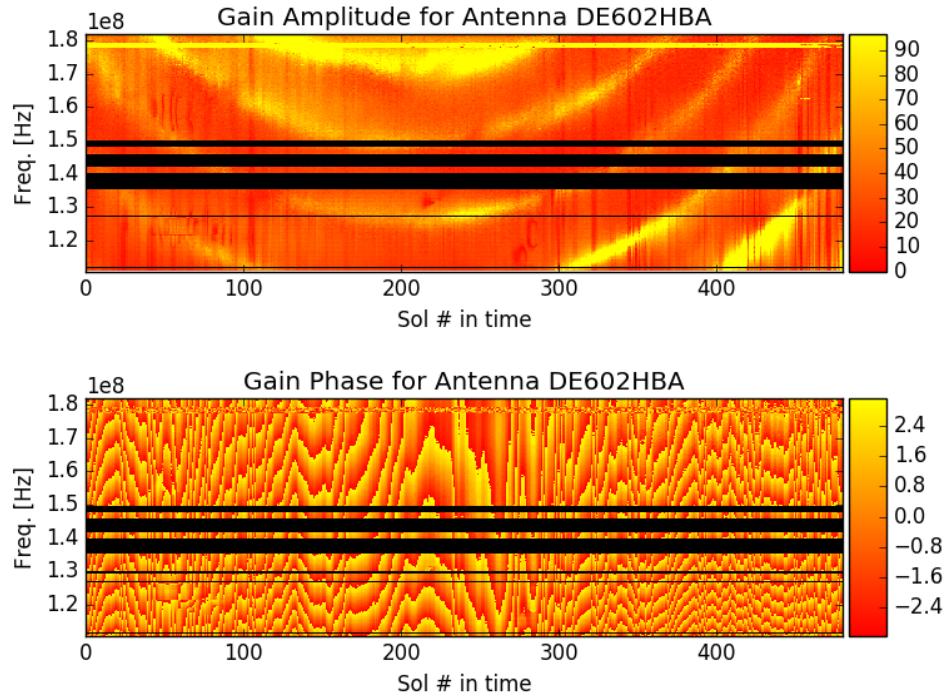
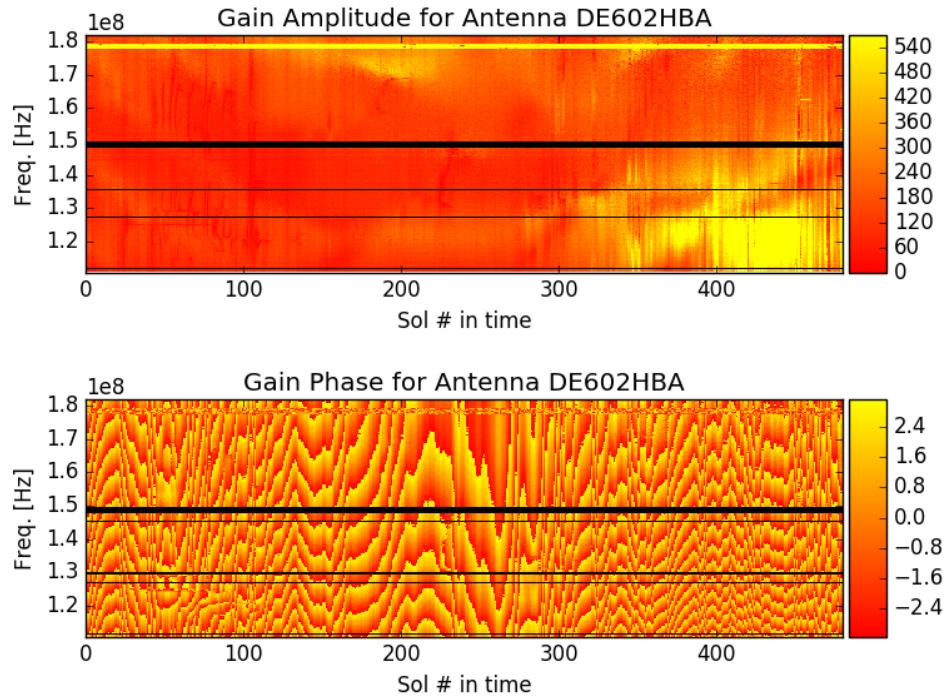


Figure 5.14. Plots showing the amplitude and phase solutions for DE601HBA at the start and end of self-calibration, respectively. The black areas correspond to regions which were not used for that pass of calibration due to poor data.

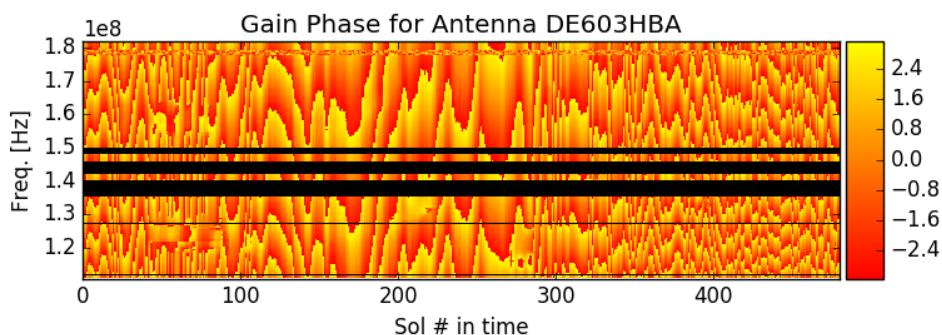
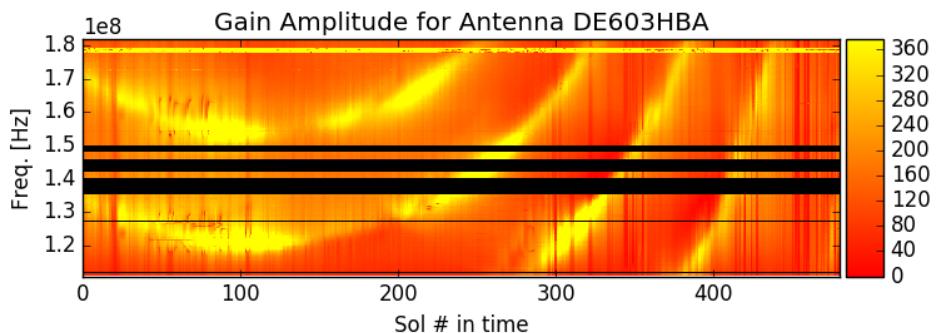


(a) *DE602HBA, first calibration, full bandwidth*

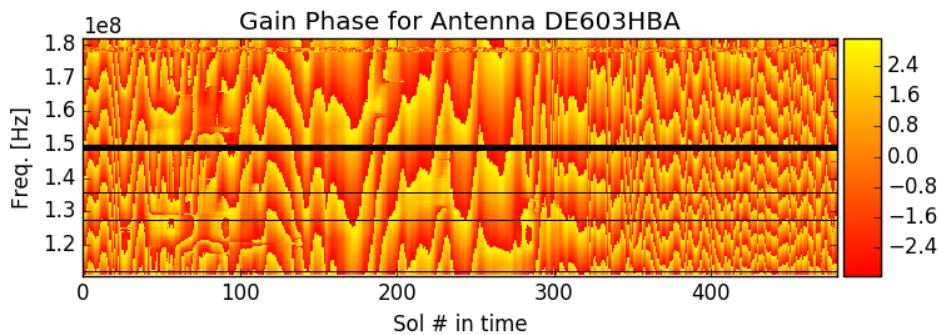
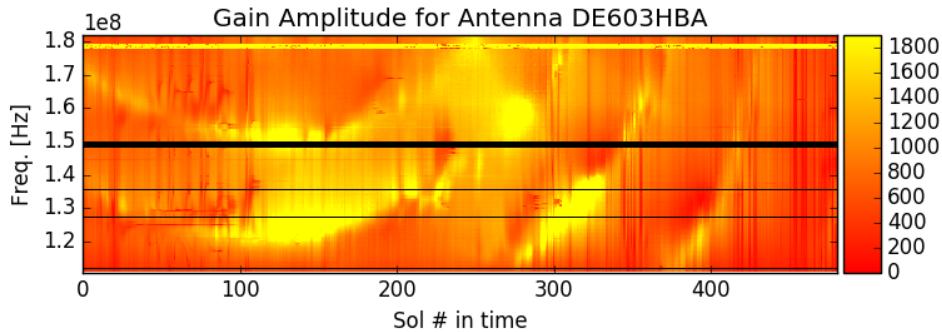


(b) *DE602HBA, last calibration, full bandwidth*

Figure 5.15. *Plots showing the amplitude and phase solutions for DE602HBA at the start and end of self-calibration, respectively. The black areas correspond to regions which were not used for that pass of calibration due to poor data.*



(a) *DE603HBA, first calibration, full bandwidth*



(b) *DE603HBA, last calibration, full bandwidth*

Figure 5.16. Plots showing the amplitude and phase solutions for DE603HBA at the start and end of self-calibration, respectively. The black areas correspond to regions which were not used for that pass of calibration due to poor data.

5.3 Overlays & Spectral Analaysis

The final, high-resolution, multi-frequency, bootstrapped model of 3C295 - which gives a spectral index for each pixel in the model - is shown in Fig. 5.17. An overlay of this model is shown in Fig. 5.18. The brightness distribution seems to have largely stayed within the bounds of the initial model for the southern lobe, but to have shifted eastwards for the northern lobe. We see that artefacts remain in the field, but this is not the case at all frequencies, as shown in Fig. 5.19: three of the six slices (a,b,d) seem to be much better-calibrated than their peers. It should be noted that we both maximise the presence of calibration artefacts *and* minimise the contribution of diffuse emission through our choices of weighting scheme and imaging *uv-cut*, respectively.

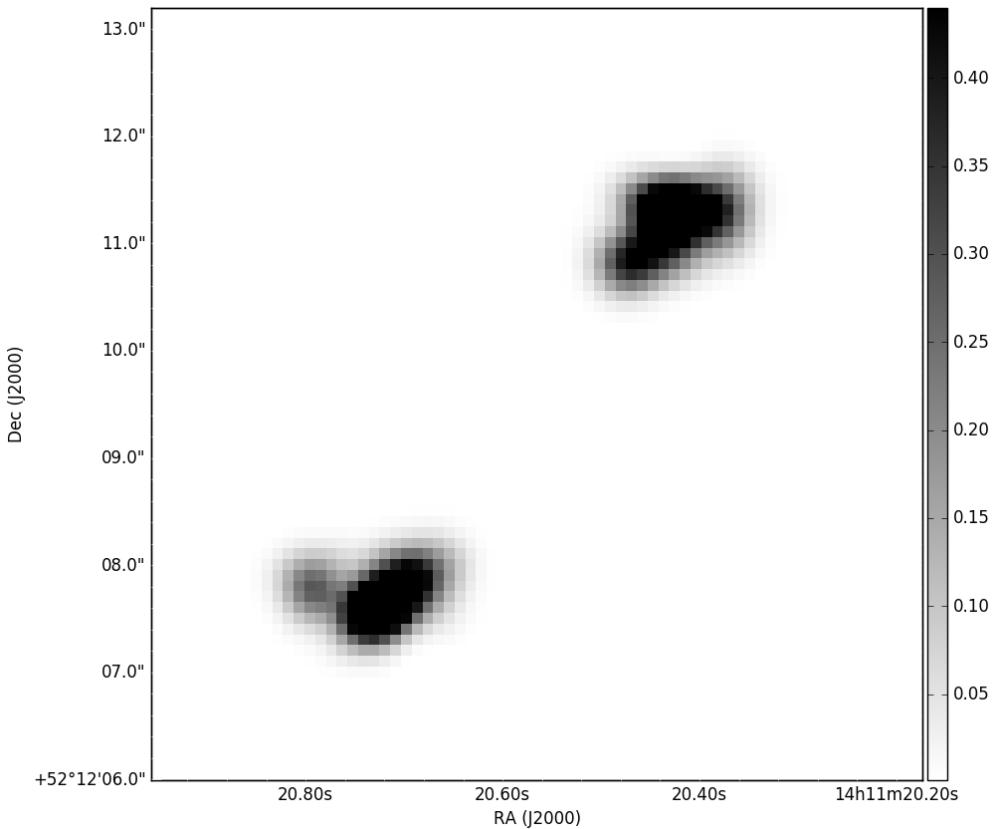


Figure 5.17. Final model of 3C295.

By inspecting individual spectral slices of our multi-spectral model in Fig. 5.19, we can see that the presence of calibration artefacts is not the same at all frequencies. We see a lot of artefacts remain for the slice at 176.17 MHz in particular, and the slice at 140.428 MHz to a lesser extent. Note that the overlay values rise exponentially. The maximum pixel value in the image is \sim 20 Jy, while the noise in the corner of the image is \sim 2 mJy. We achieve a relatively decent dynamic range (DR=10000, which is a factor of 100 lower than what is achievable with self-cal on VLA, WSRT data). We note that the dynamic range here is lower than could be expected for this data. This could be due, among other things, to our choices of *uv-cut* and imaging weights, both of which “artificially”

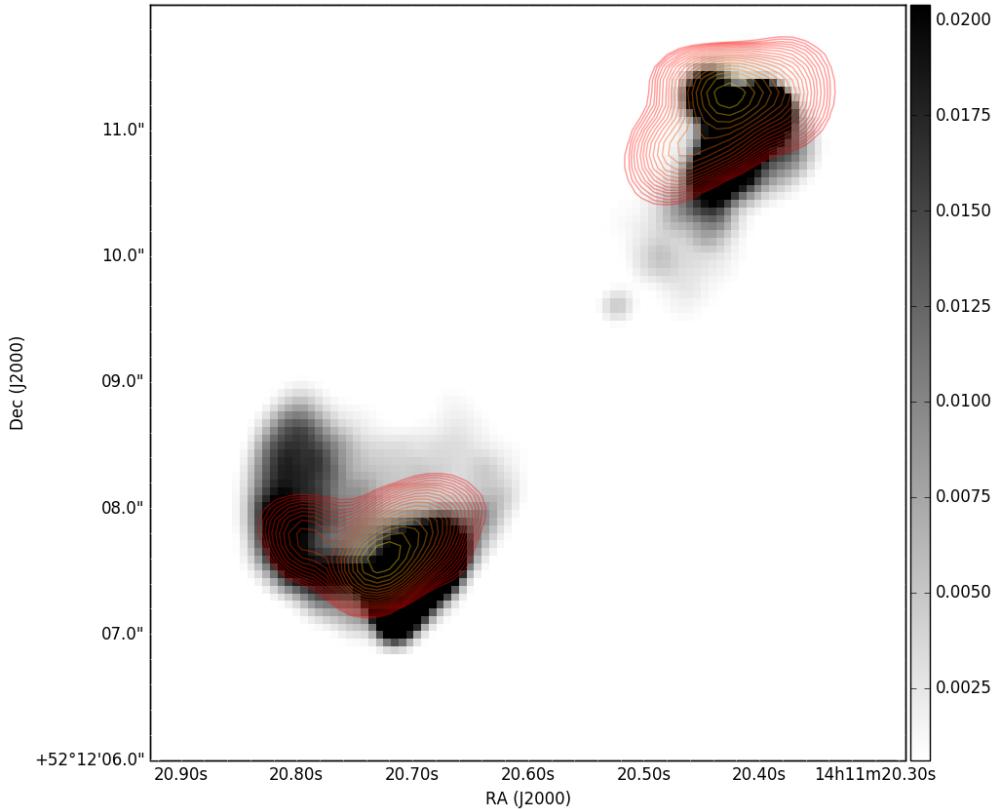


Figure 5.18. Overlay of final model of 3C295 on initial VLA model. This final model is made by adding individual contributions from the spectral slices shown in Fig. 5.19.

increase the rms in the image but are necessary to improve image deconvolution for such a diffuse source. However, considering the still-large presence of artefacts near 3C295, it is likely that we ended up in a local minimum by not modeling all the sources in the field.

It is interesting to note that the rms value in each individual slice is comparable to the rms value in the “collapsed” model (average of all slices), where we would expect a decrease by a factor $\sim \sqrt{6}$ (i.e. about 2.5). The fact that this does not occur is a strong indication that we are in a regime where we are dominated by calibration artefacts, and that the rms is therefore not truly thermal. It is also an indication that we could likely improve the self-calibration strategy: this behaviour could be due to too small a deconvolution mask, for example.

Note that each individual pixel of our model has its own spectral index: after bootstrapping the model, we therefore have a full spectral model of 3C295 across the LOFAR HBA bandwidth, with an integrated flux compatible with the literature and a spectral brightness distribution determined by our observation. This model can be used to perform direction-independent calibration of all future observations of the EGS.

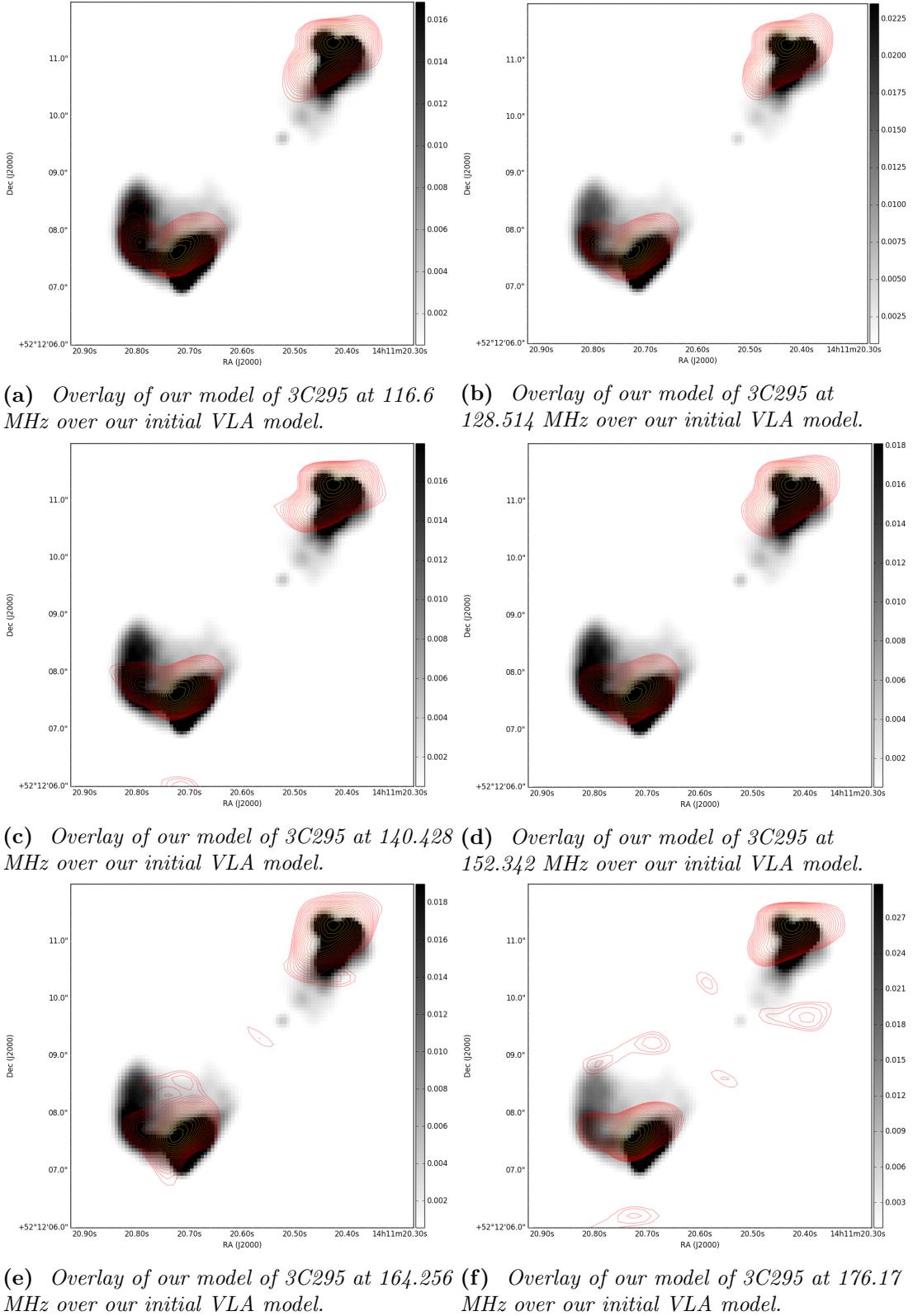


Figure 5.19. Overlays of individual slices of our spectral model of 3C295 over our initial calibration model. The overlays start at 1500σ , and are shown with increments of 1500σ .

5.4 Discussion

The plots of pre-self-calibration gain amplitude and phases as a function of frequency are a strong indicator that the gain structure due to unmodeled flux is dominated by a strong, bright source (ring-like structure in e.g. Fig. 5.16). However, the remaining presence of diffuse ring-like structure, even after the self-calibration process, for the international station gain plots is also a strong indicator that we do not recover all of the true sky brightness distribution in the model, as some of the flux is still clearly absorbed into the gains. The most straightforward way to improve on this would be to repeat the procedure while including a full low-resolution sky model: while 3C295 is by far the brightest source in the field at a recorded \sim 86.9 Jy in the catalog of the sources in our wide-field image (Table 4.1), it is far from the only contributor (the net flux in the field, according to that same catalog, is 531.1 Jy). Since calibration artefacts are introduced by both incomplete sky model and poor calibration inverse problem conditioning, improving one of these two conditions would likely result in a net improvement of the model.

Because of time/computational constraints, we used calibration solution intervals of 8 channels and 1 minute in time and frequency respectively. Given the brightness of 3C295, this was clearly not justified on the basis of signal-to-noise; it did, however, mean that a single pass of self-calibration on one subband took 10 minutes rather than the hour and a half with solution intervals of 4 channels and 8 seconds in time and frequency respectively (factor of 9 improvement in time). With the advent of new tools like cubical (Kenyon et al. 2018a,b), a much faster direction-independent implementation of the algorithm used for killMS⁸ which was in development throughout my PhD, it is likely possible to dramatically speed the calibration part of the self-calibration loop (the imaging time constraint would dominate as more subbands are added). Because of the use of quality-based weighting schemes, it is unlikely that a dramatic improvement would be seen when self-calibrating over a few subbands, as the relative contribution of poorly-calibrated visibilities would be down-weighted by the quality-based weighting scheme: as more and more data is added, however, the relative contribution of the artefacts introduced by the choice of calibration intervals would likely begin to overcome the purely thermal noise in the images, which is expected to decrease as more data is added.

All the same, this project served as a useful exercise in both developing a good working methodology and in testing calibration strategies for international LOFAR datasets. Much of the development which went into this project and led to no directly useful results still proved an extremely helpful context in which to understand how to improve the LOFAR pipeline. In particular, gain smoothing⁹ has shown promising result on LOTSS images, but was not used in this project because of the full-Jones solutions: properly implementing this smoothing in a framework capable of accounting for polarisation could provide the kind of regularisation that our gain solutions need.

⁸In the context of this work, it would literally be an equivalent but faster implementation of exactly the same algorithm.

⁹i.e. the application of a smoothing operation in gain phase and amplitude, justified by the fact that they are expected to be smooth and so jagged behaviour is likely erroneous.

5.5 Testing Directional Gain Errors - Aims & Methodology

Having made both a wide-field image of the EGS without international stations and a high-resolution model of the brightest source in the LOFAR primary beam, 3C295, in order to calibrate the international LOFAR station, we can now begin to proceed to imaging the EGS with the international LOFAR stations. Before we do so, however, we must quantify whether the direction-independent calibration on the international baselines is sufficient when imaging the EGS. We expect to see some direction-dependent effects arising in the neighbourhood of 3C295, and also expect that the SNR in the field will be insufficient to perform direction-dependent calibration on the international stations: are these direction-dependent effect so severe that it is not worth imaging the EGS?

In other words, we seek to estimate whether, in practice, the EGS can, in part or in full, be placed on the same facet as 3C295 - and with what loss in accuracy to directional gain errors. The first direction-dependent effect we must take into account is decorrelation, also known as smearing (see Section 2.12). This effect causes signal loss on the longest baselines for sources far from the phase centre. To do this, we make use of the LOFAR Long-Baseline Calibrator Survey (LBCS, Jackson et al. 2016). This is a survey of suitable compact sources to use as calibrators for international LOFAR baselines: these sources are expected to consist mainly of self-absorbed synchrotron radio cores. The survey's source selection procedure is as follows: it selects all sources identified as single sources in WENSS (Rengelink et al. 1997). Then, based on the work of Moldón et al. (2015), which found that the main predictors for compact structure were a high total WENSS flux density and a flat low-frequency spectrum (in line with the expectation above), a selection was made on the basis of spectral index and WENSS flux density for all these sources.

#	RA [hms]	Dec [dms]	Dist. from EGS [deg.]	Dist. from 3C295 [deg.]
1	14:30:18.72	52:17:29.80	2.041	2.904
2	14:19:44.44	54:23:04.58	1.928	2.517
3	14:21:20.05	53:03:46.00	0.864	1.743
4	14:21:09.41	51:22:32.46	1.294	1.728
5	14:11:50.32	52:49:02.66	0.844	0.619
6	14:11:20.23	52:12:04.30	0.915	0.000
7	14:08:07.00	52:55:11.36	1.409	0.869
8	14:08:09.76	52:44:46.56	1.354	0.680

Table 5.2. Table giving the positions of our 8 LBCS sources and their distance from both the centre of the EGS and 3C295.

A significant preliminary test is therefore to check that we do in fact see any of the LBCS sources. However, this is not the end of the problem: decorrelation causes flux from a source to be “smeared” (hence its other name) around the source. The DDFacet suite can account for this effect by modelling the direction-dependent PSF, i.e. a PSF

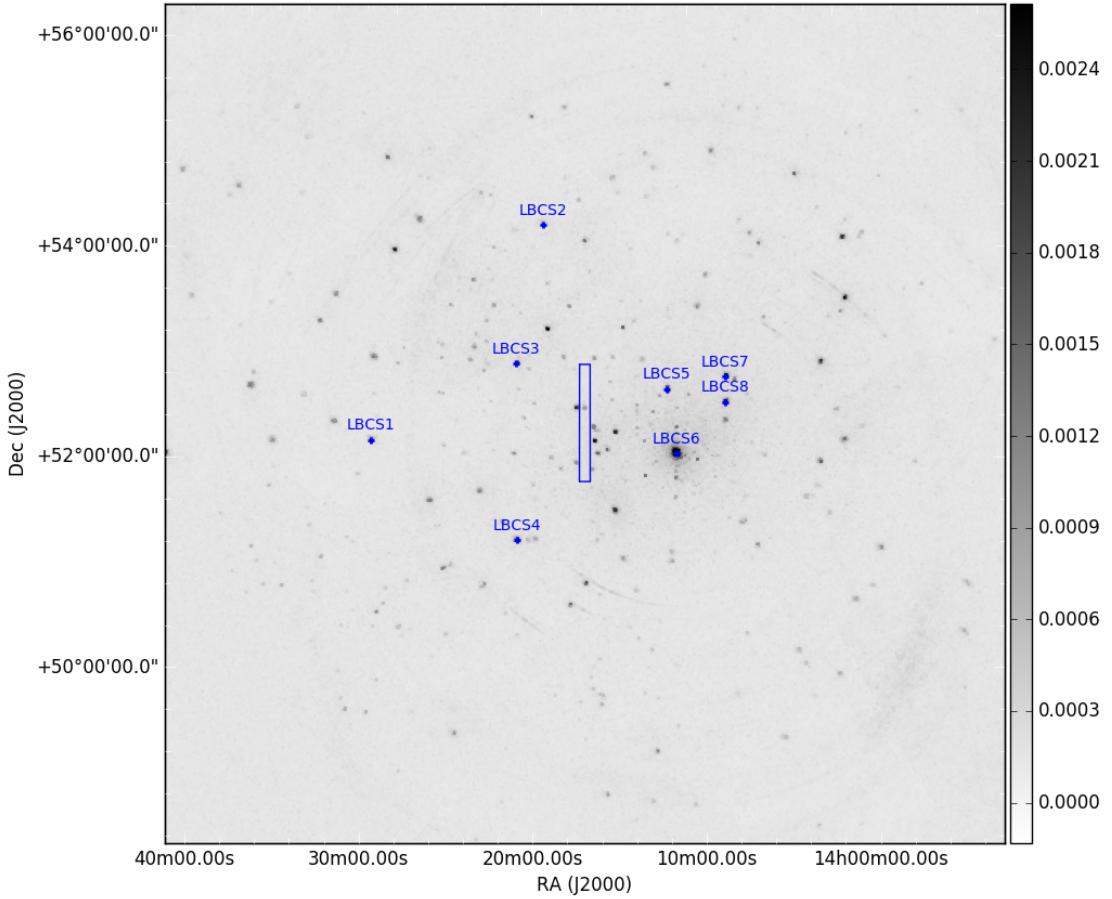


Figure 5.20. Position of the calibrator sources around the EGS, overlaid on the widefield EGS image. LBCS is 3C295.

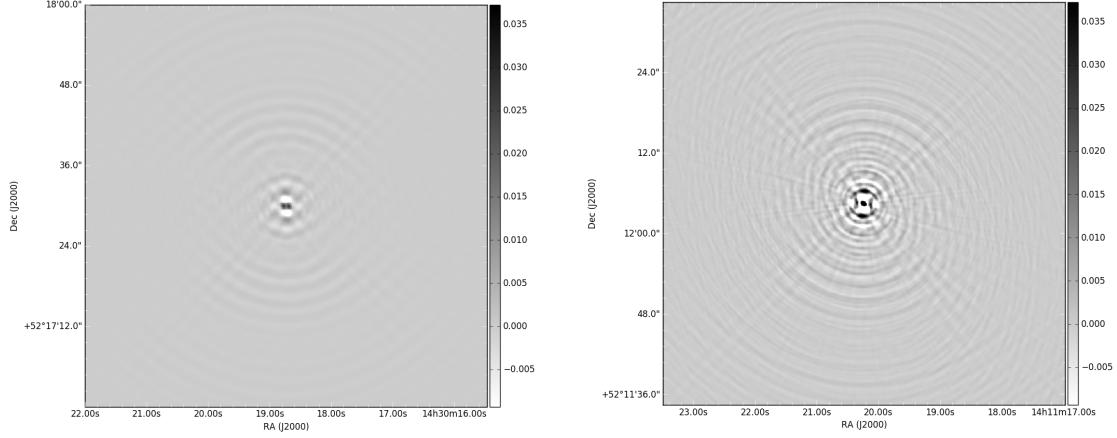
which is smeared appropriately as a function of position¹⁰.

The second source of direction-dependent effects is physical: the signal from different parts of the sky propagates through different parts of the local ionosphere before reaching our antennas.

In the ideal case, our direction-independent calibration of the EGS using 3C295 as a calibrator is good enough, for international baselines, to allow the imaging of the full EGS using international LOFAR, with acceptable upper bounds to decorrelation throughout the field. Of course, this is unlikely to be the case - the aim is then to find out just how far we are from this ideal scenario.

In RIME terms, when performing our direction-independent calibration, we solved for $\hat{\mathbf{J}}_{3C295,p}^\nu$ (see Eq. (5.3)), assuming that it was the dominant source of signal in the sky. By applying the inverse of this gain solution to the visibilities, we image the *corrected*

¹⁰A single smeared PSF is calculated per facet, and this smeared PSF is treated as being constant within the facet: this is a good approximation unless one uses truly massive facets.



(a) Direction-dependent PSF near LBCS1 (b) Direction-dependent PSF near LBCS6

Figure 5.21. Cutouts showing the direction-dependent PSF at two different points in the primary beam of our observation.

visibilities:

$$\mathbf{V}_{pq}^{\text{corr}}(\nu) = \hat{\mathbf{J}}_{3C295,p}^{-1}(\nu) \mathbf{V}_{pq}(\nu) (\hat{\mathbf{J}}_{3C295,q}(\nu))^{-H} \quad (5.8)$$

$$= \hat{\mathbf{J}}_{3C295,p}^{-1}(\nu) \left(\sum_s \mathbf{J}_{s,p}(\nu) \mathbf{B}_{pq}(\nu) \mathbf{J}_{s,q}^H(\nu) \right) \hat{\mathbf{J}}_{3C295,q}^{-H}(\nu) \quad (5.9)$$

$$= \tilde{\mathbf{G}}_p(\nu) \sum_s \left(\sum_s \tilde{\mathbf{E}}_{s,p}(\nu) \mathbf{B}_{pq}(\nu) \tilde{\mathbf{E}}_{s,q}^H(\nu) \right) \tilde{\mathbf{G}}_q^H(\nu) \quad (5.10)$$

where

$$\tilde{\mathbf{G}}_p(\nu) = \hat{\mathbf{G}}_p^{-1}(\nu) \mathbf{G}_p(\nu) \quad (5.11)$$

$$\tilde{\mathbf{E}}_{s,p}(\nu) = \hat{\mathbf{E}}_{3C295,p}^{-1}(\nu) \mathbf{E}_{s,p}(\nu) \quad (5.12)$$

where \mathbf{K} is in \mathbf{J} , and we have muted the time-dependence of all variables for the sake of readability. Provided that our gain solutions are exactly right (i.e. that $\hat{\mathbf{J}}_p^\nu = \mathbf{J}_p^\nu$ for all p, ν) then $\tilde{\mathbf{G}} = \mathbf{I} \forall p, \nu$. If the calibration problem is sufficiently well-conditioned (i.e. good signal-to-noise, many baselines) and the sky model sufficiently complete, then this is a reasonable assumption to make. However, for all $s \neq 3C295$, $\tilde{\mathbf{E}} \neq 1$. This means that, depending on how different the \mathbf{E} terms will be between a given direction and 3C295, direction-dependent calibration artefacts will be introduced in the field.

In practice, both $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{E}}$ will contribute to image deterioration: in our case, we only model part of the sky to which our array is sensitive, and so $\tilde{\mathbf{G}} \neq 1$ differently for various antennas, times and frequencies. Similarly, we do not perform direction-dependent calibration: as distance increases from 3C295, it is increasingly likely that the ionosphere through which non-3C295 sources' signal propagates before reaching our antennas will vary from the ionosphere through which the signal from 3C295 will vary - and because the ionosphere's behaviour is complex, this behaviour is unpredictable as a function of distance or frequency. The departure of both $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{E}}$ from unity as a function of direction is called *directional gain errors*, as it is caused by direction-dependent effects.

Assuming that direction-dependent effects (beam, ionosphere) are relatively smooth and dependent only on distance from 3C295, we attempt to characterise the impact of directional gain errors and their associated image-plane DDEs on various calibrator sources around 3C295 and the EGS.

We select 8 LBCS sources, which are all the “good” sources within 5 degrees of the centre of the EGS. Fig. 5.20 shows the position of these calibrator sources with respect to the EGS. If we compare Fig. 5.20 and Fig. 4.5, we can immediately see that the EGS lies in two separate facets of our direction-dependent calibration, which are the eastern neighbours of the 3C295 facet. Neither of these EGS facets contains a LBCS source (indeed, if they did, our work would have perhaps been easier); however, nearby facets do contain one or multiple LBCS calibrators. If the effect of directional gain errors for these sources is comparable to those for the EGS facets, then we can estimate their impact on an eventual high-resolution image of the EGS.

5.6 Testing Directional Gain Errors - LBCS Sources in the Primary Beam

The LBCS sources were chosen automatically using a routine written by Leah Morabito and shared in a personal communication. This was done in early 2017. Only the brightest reliable sources were kept, provided they had a good reliability factor (the LBCS catalog includes a reliability factor for each source; we have not verified that these reliability factors have not changed since). This left a total of 8 sources, described in Table 5.2. Their position on the sky is shown in Fig. 5.20. They will henceforth be referred to as LBCS1, LBCS2, and so on - note that LBCS6 is actually 3C295.

For each LBCS source, we create an image made using the same data and imaging parameters as our final 3C295 self-calibration pass. This means that we simply use our final restored image of 3C295 for LBCS6, but make a new image (with the same parameters except the direction) for each other LBCS source. We then show overlays of the results onto our wide-field image of the full primary beam. Note that both the international LOFAR images ($\sim .4''$ resolution) and the widefield images ($\sim 20''$) are oversampled: their pixel sizes are $0.1''/\text{pixel}$ and $5''/\text{pixel}$ respectively.

Because we want to show both the impact of DDEs and the positions of the sources, we show two different overlays for each source. One starts at 5σ and the other starts at 15σ ; both increase exponentially up to the maximum value in the image. The first thus gives more information on artefact structure around the sources, while the second gives more information on source structure itself. The σ -level cutoffs were chosen arbitrarily. Note that the σ is calculated using random pixels in the image: it is therefore not the “local RMS” used in e.g. PyBDSM extraction (which will be higher, as it includes artefacts) but rather the “local thermal noise”, so to speak. As such, the minimum of these overlays are not representative of the local 5σ and 15σ levels.

Finally, because 3C295 is such a bright source, the dynamic range around it is extremely high and so the overlay minimum lies at 100σ . We only create a single overlay

for this source (as opposed to 2 for all others), because it is already discussed in detail in Chapter 5. The associated overlay is shown in Fig. 5.22: we see that the low-resolution image shows roughly the same morphology (e.g. northern/southern lobe directions), albeit at much lower resolution. The other overlays are shown below.

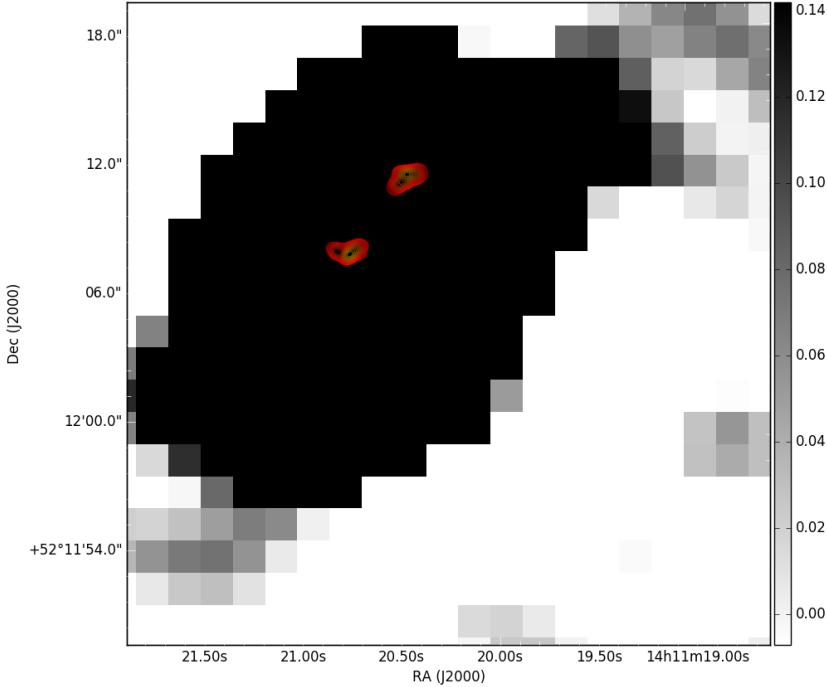


Figure 5.22. Overlay of international LOFAR image of 3C295 over the widefield image.

A few things are immediately apparent: LBCS1 seems to be a binary source, and LBCS3 seems to suffer strongly from directional gain errors (the source structure in stripes seems unlikely to be physical). LBCS4 seems almost unaffected, which would indicate that the portions of the EGS which lie between 3C295 and LBCS4 are not strongly affected by directional gain errors. The fact that LBCS3 seems to be affected would corroborate the notion that the characteristic scale of DDE impacts is of the order of just over a degree.

Of course, it is not enough to eyeball this: we therefore proceed to quantify the impact of directional gain errors for different sources in the field. To do this, we begin by making the approximation that they are a function of distance from the calibrator source only, rather than specifically a function of direction and distance. This is simply because we don't have enough calibrators to do otherwise. Secondly, we treat the effect of directional gain errors as *increasing the noise* around a source compared to what the thermal noise is expected to be around it: this thermal noise is expected to increase as a function of distance from the phase centre because we apply a beam-correction during imaging (i.e. because we do not image apparent flux but integrated flux). Fig. 5.31 shows this effect in practice; note that the noise estimation around LBCS6 (3C295) is biased because we are still quite close to this extremely bright source, and thus likely have a large contribution from artefacts. This explains its outlier behaviour.

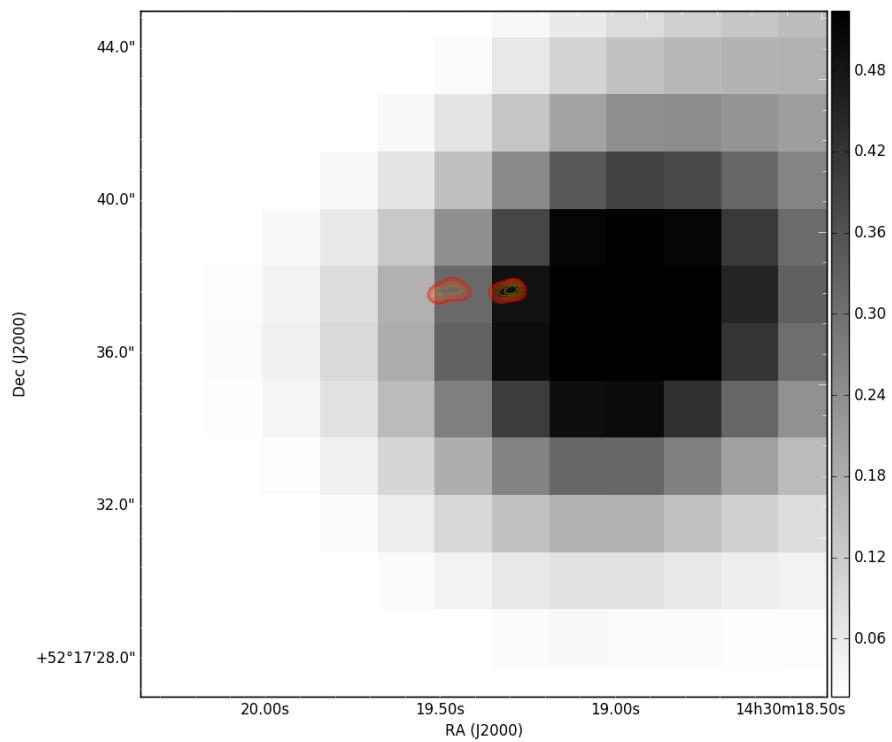


Figure 5.23. 5σ overlay of international LOFAR image of LBCS1 over the widefield image.

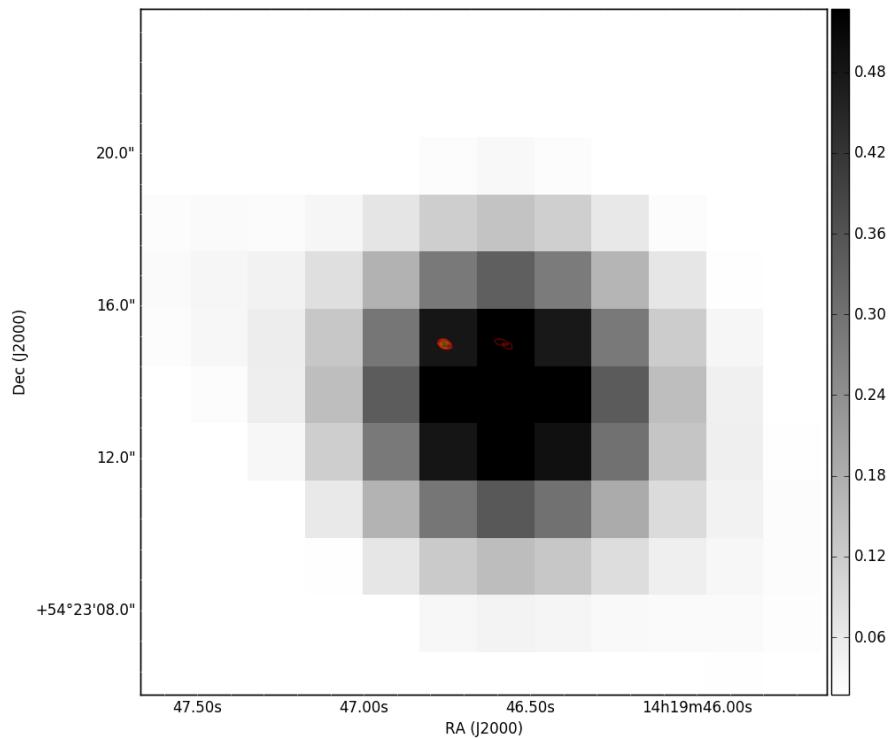


Figure 5.24. 5σ overlay of international LOFAR image of LBCS2 over the widefield image.

5.6. TESTING DIRECTIONAL GAIN ERRORS - LBCS SOURCES IN THE PRIMARY BEAM135

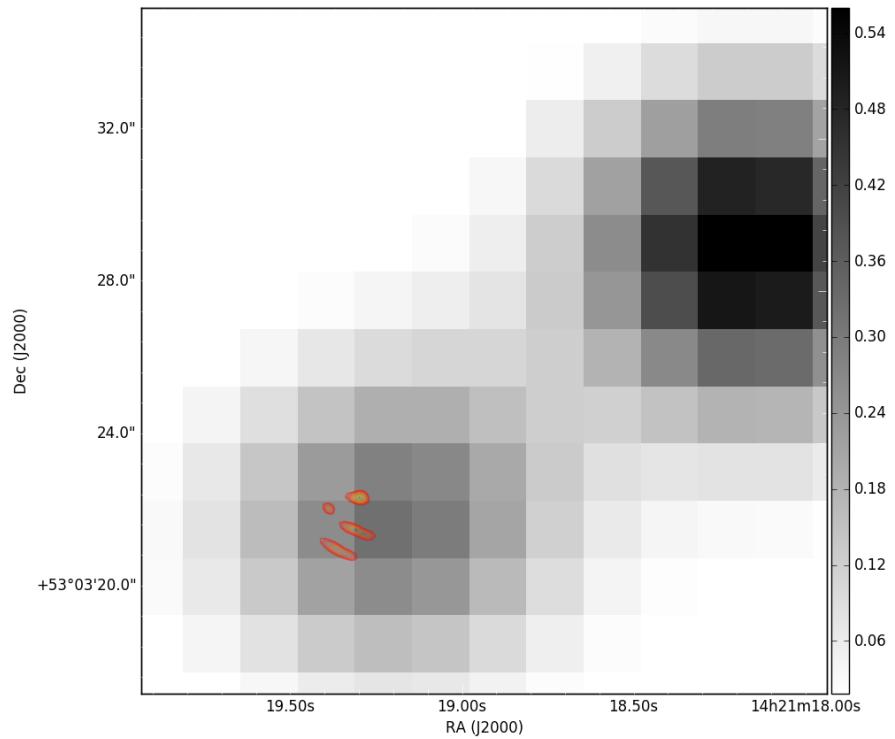


Figure 5.25. 5σ overlay of international LOFAR image of LBCS3 over the widefield image.

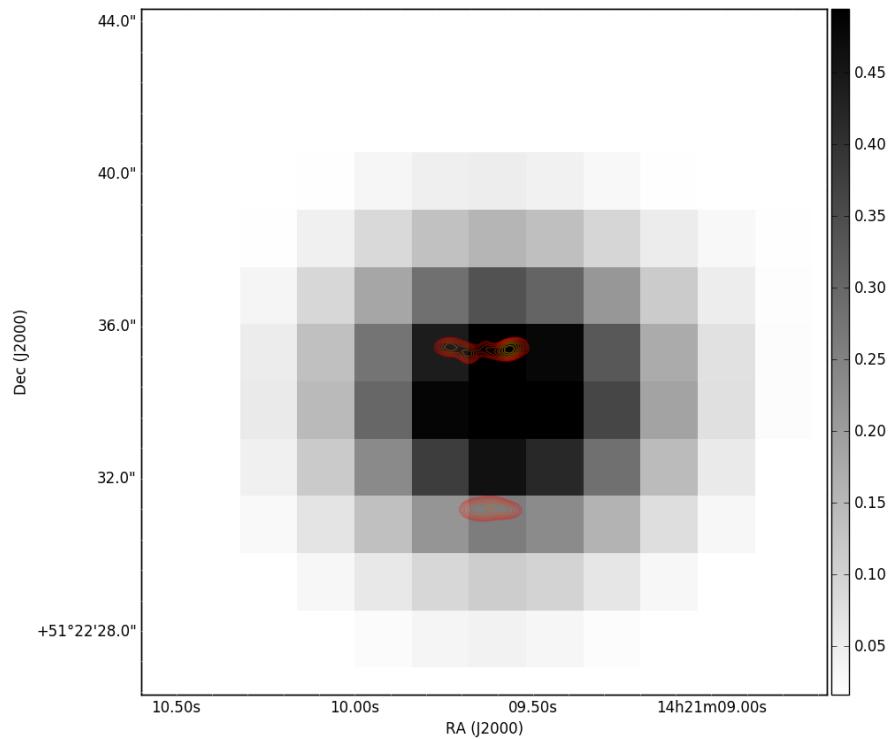


Figure 5.26. 5σ overlay of international LOFAR image of LBCS4 over the widefield image.

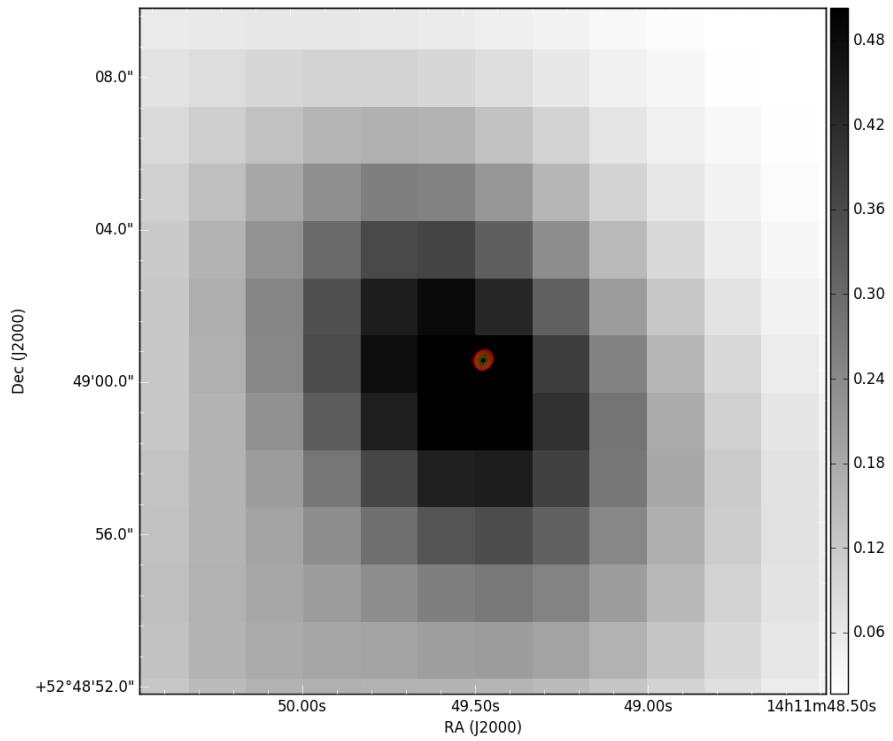


Figure 5.27. 5σ overlay of international LOFAR image of LBCS5 over the widefield image.

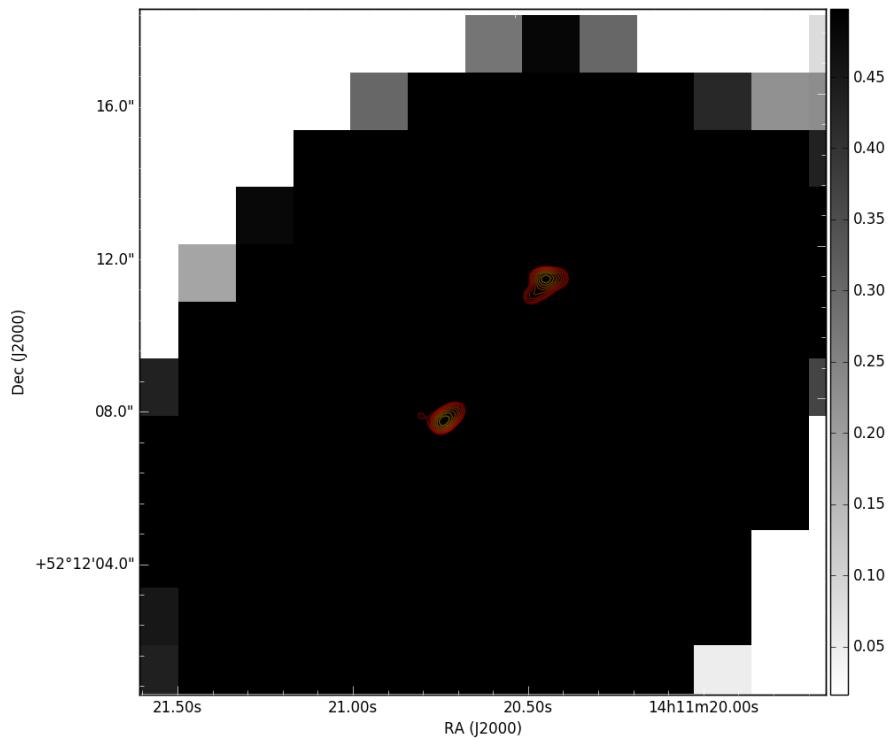


Figure 5.28. 5σ overlay of international LOFAR image of LBCS6 over the widefield image.

5.6. TESTING DIRECTIONAL GAIN ERRORS - LBCS SOURCES IN THE PRIMARY BEAM137

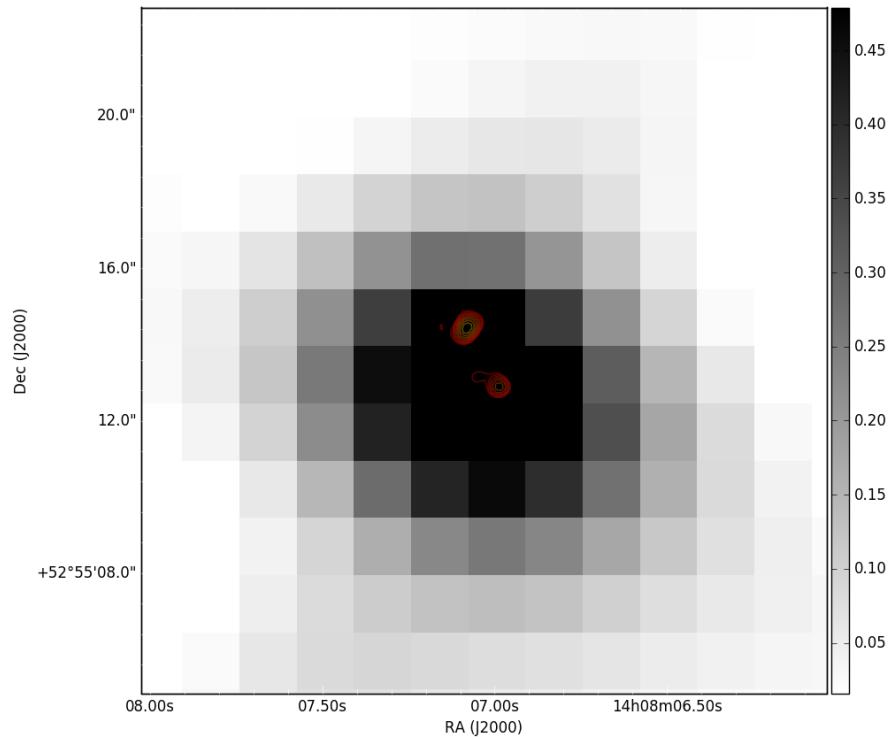


Figure 5.29. 5σ overlay of international LOFAR image of LBCS7 over the widefield image.

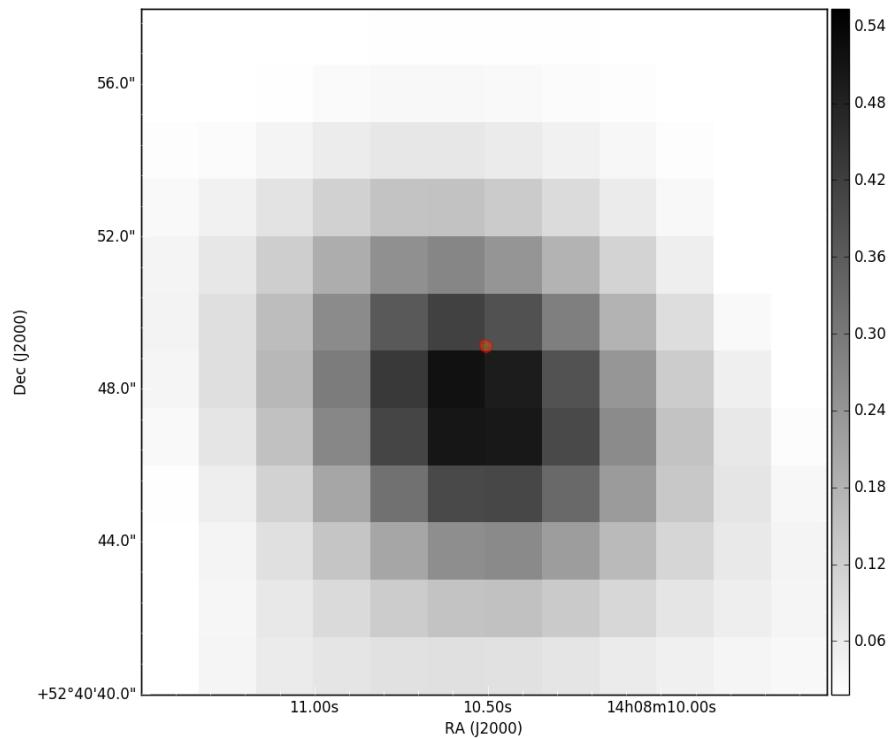


Figure 5.30. 5σ overlay of international LOFAR image of LBCS8 over the widefield image.

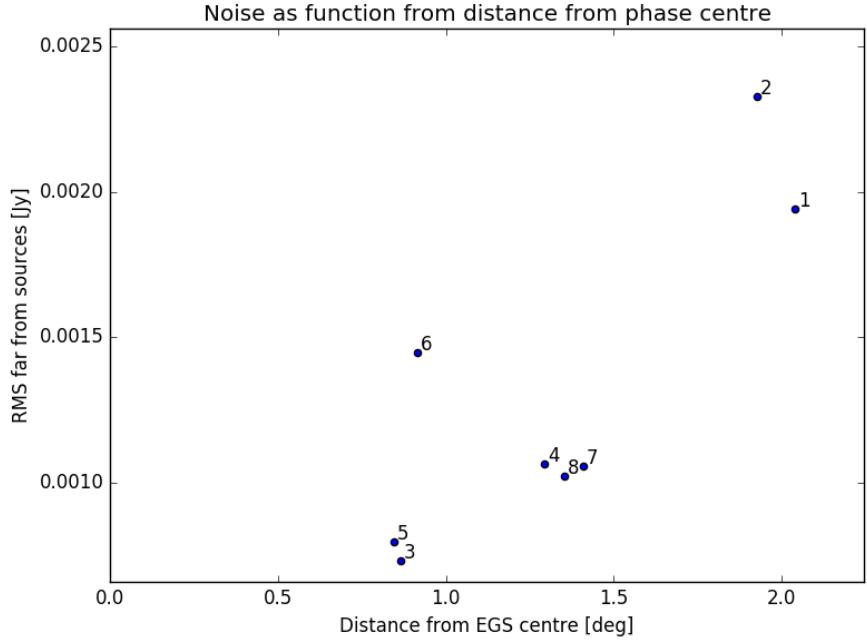


Figure 5.31. Thermal noise near LBCS sources as a function of distance from the phase centre.

To estimate the prevalence of directional gain errors, we then calculate the rms around each LBCS source. This is plotted in Fig. 5.32. The RMS around each source is normalised by the source's integrated flux as calculated on the restored image: this is because we expect brighter images to show stronger artefacts associated to directional gain errors. It should be noted that LBCS8 is not deconvolved at all: as such, its RMS is effectively calculated on a corrupted dirty map of the source, rather than a residual map. This helps explains its outlier behaviour. This outlier aside, the directional gain rms seems fairly flat as a function of distance from 3C295, however - a very good sign, as it indicates that the directional gain errors are comparable across the primary beam. However, we do not take into account the increase in thermal rms as a function of distance from phase centre in this plot. We therefore look at the normalised RMS values for each LBCS source, shown in Fig. 5.33. The relationship remains flat, though LBCS5 now stands as a 3σ outlier. The blue line shows the mean value, excluding LBCS8; it lies at $y = 1.08063981$. In other words, the average RMS induced by directional gain errors, when normalised by both the flux of a source for which it is calculated and the thermal noise, is very close to unity, and does not increase very much as distance from 3C295 increases. While it apparently increases in the direction of LBCS5, the EGS is not in this direction from 3C295, and so this is not a source of concern for our project of imaging the EGS using international LOFAR.

The most encouraging result shown above, however, is for LBCS3 & LBCS4: both of these sources appear to be only very slightly affected by directional gain errors. Since the EGS lies between these two calibrator sources and 3C295, this is extremely positive. Similarly, the fact that LBCS4 resolves into two separate sources (as does LBCS1, for that matter) would indicate that even neighbouring bright sources do resolve properly -

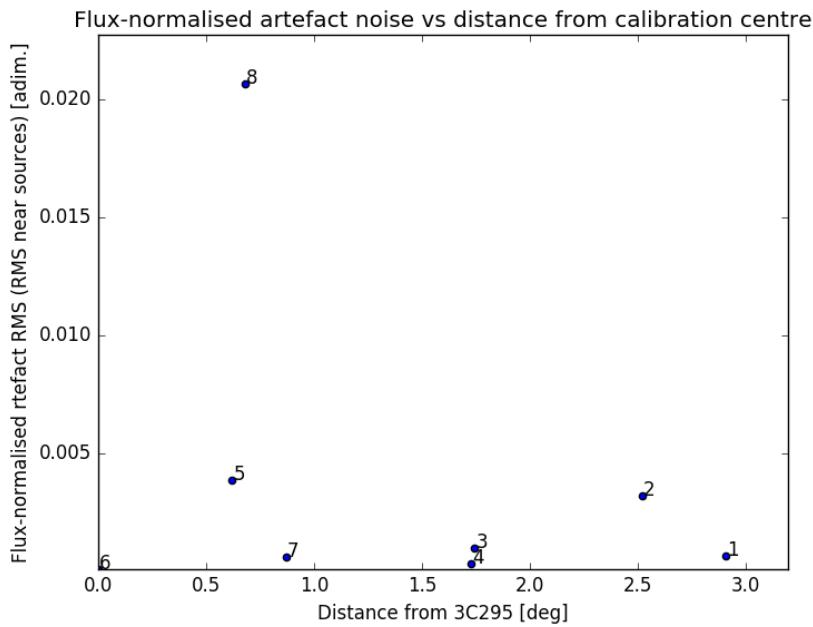


Figure 5.32. DDE noise near LBCS sources as a function of distance from the phase centre. Because 3C295 is so much brighter than all other sources ($\sim 100\text{Jy}$; second-brightest LBCS source is $\sim 2\text{Jy}$), its DDE noise was normalised by its integrated model flux.

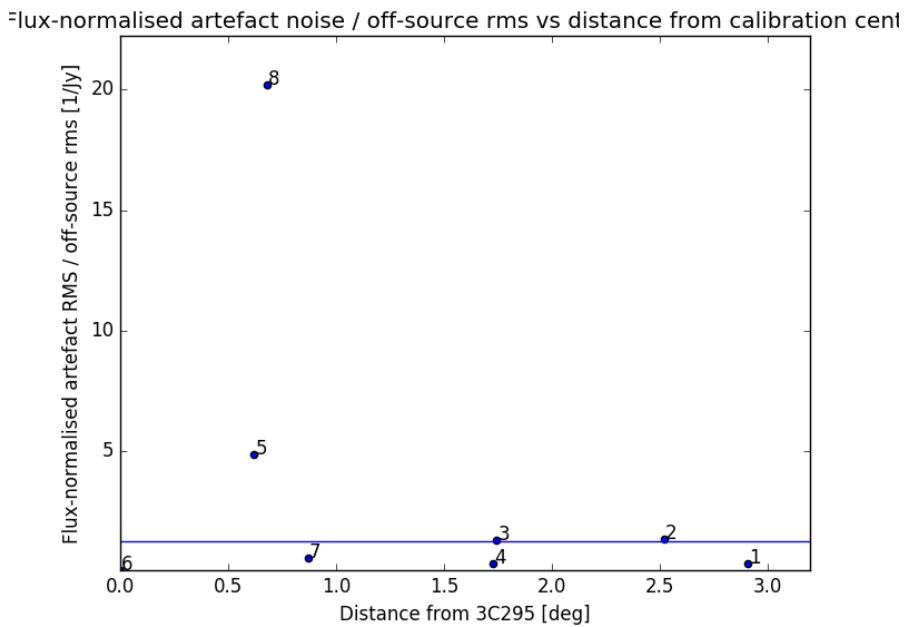


Figure 5.33. Normalised RMS values around each of the LBCS sources in the field as a function of distance from phase centre.

this is not a rigorous test, but it is a fortunate serendipitous observation to be able to make.

Chapter 6

Conclusion & Future Work

Contents

6.1	Algorithmic Work	142
6.1.1	The Cov-Cov Relationship	142
6.1.2	Mapping the variance of images made with interferometric data	143
6.1.3	Quality-based Weighting Schemes	144
6.2	Imaging the Extended Groth strip	145
6.2.1	Imaging the LOFAR primary beam	146
6.2.2	Imaging the EGS with LOFAR-VLBI	146
6.2.3	Future Work	149

We conclude this thesis manuscript with an overview of the results achieved over the course of this doctorate and a discussion of individual items in this larger context. We follow the structure of the manuscript. These conclusions & discussions are intended to supplement the conclusions & discussions at the end of each chapter with a synthetic overview of the results of this doctoral thesis.

The projects undergone as part of this doctoral degree have strong opportunities for future work, both in terms of the algorithmic work & its application. These will be outlined in detail in the relevant section. It is worth immediately pointing out that only preliminary results were obtained for the application to data by the end of the doctorate's three-year period. These preliminary results are shown in this manuscript, and while they are not of science-grade quality (no astrometric corrections, proper flux bootstrapping, strategy likely needs revisiting for best results) they are nevertheless useful markers of the wealth of information available in the EGS.

6.1 Algorithmic Work

One of the aims of this PhD was to develop new interferometric tools & techniques and to apply them to LOFAR data. This was arguably accomplished successfully:

- We found a relationship between the statistics of residual visibilities and residual image-plane pixel values (the “Cov-Cov relationship”).
- From this relationship, we see that **calibration artefacts are directly caused by correlated calibration gain errors**. In the absence of correlated calibration errors, no artefacts are present.
- This allowed for a description of the noise-map in the image plane as a constant variance level modulated by a noise-PSF convolved with the sources in the field.
- This led to the development of an adaptive, quality-based weighting scheme, which reduces the noise in the image (and the presence of calibration artefacts) by minimising either the constant noise term or the noise-PSF.
- One version of this adaptive, quality-based weighting scheme was successfully applied throughout our data reduction & wide-field imaging. Other versions are possible, but at the time of writing suffer from issues due to poor conditioning: the development & deployment of robust versions of these will be the subject of future work.

6.1.1 The Cov-Cov Relationship

The theoretical framework of the cov-cov relationship, which links the variance and covariance in the visibilities to the variance and covariance between pixels in images made with these visibilities, was shown to hold on real data by improving images made with both direction-independent and direction-dependent calibration. This relationship tells us that the variance in images made from interferometric data is the result of two contributions:

$$\begin{aligned} \text{Cov}\{\tilde{\mathbf{y}}\} = \sum_d \left(\sum_b \phi_{bb}^d \left([\text{Cov}\{\tilde{\boldsymbol{\gamma}}\}]_{bb} + \frac{w_b^2 \sigma^2}{\phi_{bb}} \right) \mathbf{c}_b \right. \\ \left. + \sum_{b,b' \neq b} \phi_{bb'}^d [\text{Cov}\{\tilde{\boldsymbol{\gamma}}\}]_{bb'} \mathcal{F}_{bb'} \right) \end{aligned} \quad (6.1)$$

which is a repeat of Eq. (3.25). The first term above, which corresponds to the sum over b , can be thought of as “thermal” (i.e. uncorrelated) noise in the visibilities, giving rise to a constant variance throughout the image. The second, which corresponds to sky brightness distribution absorbed into the gain solutions - i.e. correlated noise in the calibration residual visibilities - gives rise to what we call a noise-PSF: a distinctive shape which is convolved with every source in the field (in the case where the true gains are direction-independent), and which gives the distribution function of which calibration artefacts are a single realisation. It corresponds to the sum over $b, b' \neq b$ above - i.e. to those cells of the visibility covariance matrix which are off-diagonal, or correlations between residual visibilities.

In Eq. (6.1), $\text{Cov}\{\tilde{\mathbf{y}}\}$ is the covariance matrix of the residual image pixel values, and $\text{Cov}\{\tilde{\boldsymbol{\gamma}}\}$ is the covariance matrix of the residual visibilities, which is of size $N_b \times N_b$. $[\text{Cov}\{\tilde{\boldsymbol{\gamma}}\}]_{bb'}$ is a scalar quantity equivalent to the value of the covariance matrix at coordinates b, b' . $\phi_{bb'}^d$ corresponds to the product of the weighted¹ visibilities, associated to a source in direction d ², as seen by baselines b and b' . $\phi_{bb'}^d$ is therefore the square of the magnitude³ of this quantity for a single baseline and direction. σ^2 is the thermal noise in the visibilities, which is assumed to be constant for all of them. Finally, \mathcal{C}_b is the image-space fringe corresponding to a single baseline, while $\mathcal{F}_{bb'}$ is slightly too complex to summarise quickly, but its diagonal maps a single visibility covariance value to a $\delta u \delta v$ fringe, thus determining the artefact distribution around sources in the field.

Taking the diagonal the matrices shown above gives the “var-var” relationship, which gives the variance map of the image as a function of the variance in the visibilities. If the visibility covariance matrix is diagonal (i.e. we have no correlated noise in the visibilities), then the second term will be null, and the variance of each pixel in the image-plane will be a constant value (since $\text{Diag}\{\mathcal{C}_b\} = \mathbf{I} \forall b$). This is an extremely strong prediction, but one that was validated through simulations: calibration artefacts in interferometric images are due *only* to correlated noise in the visibilities. In the absence of such correlated noise, there will be no calibration artefacts in an image.

Note that this relationship also confirms, among other things, that “thermal” noise in interferometric images is actually correlated between pixels, and that this correlation is given exactly by the PSF. This is an expected result, and can only be corrected by fully-deconvolving the PSF from an interferometric image⁴. If this is done, then an image made using visibilities containing only uncorrelated noise will itself have noise uncorrelated between pixels, but will not otherwise.

6.1.2 Mapping the variance of images made with interferometric data

This relationship was strongly tested through simulations, verifying the noise-PSF behaviour of calibration artefacts in Fig. 3.5, reproduced in Fig. 6.1. The visibilities for three point sources were simulated for a single uv -track and frequency, and these were multiplied by multiple realisations of time-correlated residual gains. By taking the DFT of each realisation to create dirty maps of this simulated residual map, and calculating the variance for each pixel across our realisations to find the variance map, we see that we do in fact see a PSF-like behaviour in the noise-map, in agreement with our predictions.

¹e.g. Briggs-weighted, quality-weighted, etc

²i.e. the visibility which the interferometer would record if this was the only source in the sky.

³i.e. the product of the visibility and its own complex conjugate.

⁴Here, when talking about “fully deconvolving the PSF from an interferometric image”, we mean to systematically deconvolve the PSF from each pixel in the image simultaneously. This would be an impractically computationally expensive task, and is therefore not a realistic proposal. Regardless, the point being made here is that the noise of images made using visibilities with purely thermal noise is correlated between pixels - even though the noise is not correlated between visibilities.

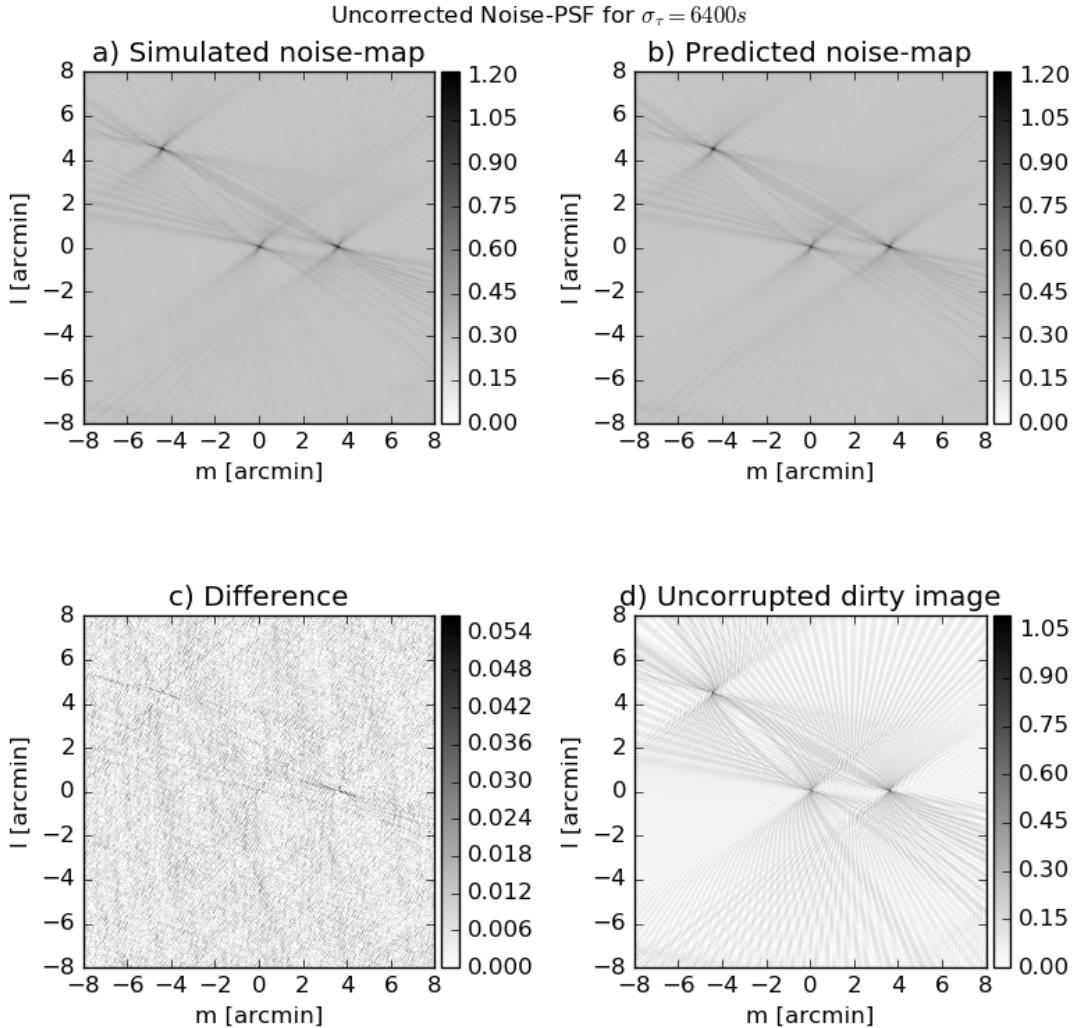
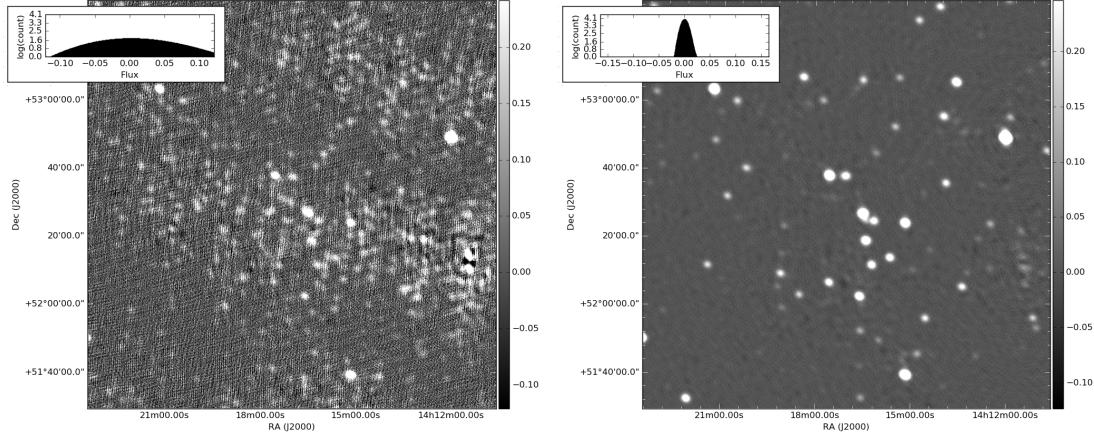


Figure 6.1. Noise-map of sky with correlated gain errors and three point sources. The colourbars of (a), (b) and (c) have dimensionless units, while that of (d) is in Jansky. Note the presence of structure in the residuals (c): these show the limits of our hypothesis that sources are well-separated.

6.1.3 Quality-based Weighting Schemes

The quality-based weighting schemes developed in this section can then be understood as ways to change the noise-distribution: one minimises the constant component of the noise-map, while the other seeks to flatten the noise-PSF. Only the former was implemented for use on real data: the latter suffered from poor conditioning. This weighting scheme consists of down-weighting each visibility by an estimate of the variance of residual.

$$w_b = \frac{1}{\text{Var}\{\tilde{\gamma}_b\}} \quad (6.2)$$



(a) Poorly-calibrated, unweighted restored image of the sky near the centre of the Extended Groth Strip. Units of colourbar are Jy/bm . $rms=86.4mJy/beam$

(b) Image made with the application of antenna-based, sensitivity-optimal weighting. Solution interval of 2 minutes, half the bandwidth. $rms=6.69mJy/beam$

Figure 6.2. Copies of Fig. 3.7b and Fig. 6.2b, used to show the impact of the weighting schemes on real data.

The cov-cov relationship is likely going to be revisited at a future time to be expanded to more complex cases. Much work nevertheless remains: in both improving the conditioning of the covariance matrix estimation, a necessary prerequisite for its successful application, and in attempting to generalise the framework to account for DDEs, more sophisticated interval solutions (e.g. convolution with a Gaussian with characteristic width equal to the solution interval rather than simple binning), and other modern features of calibration solvers. Similarly, the effect of sky model incompleteness has yet to be investigated thoroughly, and simulation work to this effect could bring great insight on this still-open question.

Similarly, the question of properly conditioning our estimates of the visibility covariance matrix remains open. This would be a potentially fertile avenue of research, as being able to estimate the full visibility covariance matrix reliably and swiftly could lead to the ability to implement “artefact-optimal” weights: that is, attempt to completely flatten the noise-PSF, at the cost of increasing the constant, “thermal” noise component of the image’s noise-map.

6.2 Imaging the Extended Groth strip

We applied our newly-developed weighting schemes to the problem of imaging the Extended Groth Strip, a rich extragalactic field near a bright 3C source. In so doing, we followed a four-part strategy, consisting of:

- Imaging the full LOFAR primary beam with the core and remote stations. This requires direction-dependent calibration;
- Creating a high-resolution, frequency-dependent model of 3C295 to use as an international station calibrator source. This required direction-independent self-

calibration with a strong imaging *uv*-cut;

- Preliminary investigation of the impact of directional gain errors from direction-independent calibration using the model of 3C295 obtained from the step above;
- Creating a high-resolution image of the EGS after subtracting all sources from the widefield image & the high-resolution model of 3C295.

Of these four points, only the first three were successfully carried out. We therefore limit our discussion to the first three points, treating the last as the subject for future work.

6.2.1 Imaging the LOFAR primary beam

We begin by creating a decent model of the sky seen by LOFAR’s primary beam. To do this, a wide-field, 6"-resolution image of the fully primary beam was made: this image makes no use of the LOFAR international stations due to technical limitations (memory use). Using PyBDSF (Mohan & Rafferty 2015), this model was extracted from an image made after a few rounds of direction-dependent self-calibration. This direction-dependent self-calibration was done using the killMS/DDF pipeline, which fully models direction-dependent errors due to decorrelation: its application to data is one of the novel contributions of this work.

This model is still in a very preliminary stage: astrometric errors have not been corrected for, nor has integrated flux bootstrapping been performed. As such, this image (and the model extracted from it) is not of science-grade quality. Even so, and preliminary as it is, it is sensitive and wide-field enough to pick up a good number of interesting diffuse sources. 12 of these were selected to form a so-called “primary beam galactic zoo” of sources, with the aim of showing the reader the contents of a typical LOFAR deep extragalactic field. They were then matched to NVSS, NVSS & WISE images to find counterparts: the results are shown in Table 6.1, which is a repeat of Table 4.1.

The true interest of modeling these sources at the wide-field resolution achieved is that will allow us to subtract (most of) the emission associated with them when imaging the EGS at high resolution. This, in turns, will allow us to image the EGS itself without pollution from the sidelobes of nearby bright sources. Note that there is nothing inherently special about this field compared to LOTSS fields (Shimwell et al. 2017): its interest is in what it can allow us to achieve with the EGS.

6.2.2 Imaging the EGS with LOFAR-VLBI

Using 3C295 to self-calibrate the international stations

Over the course of this doctoral thesis, I was able to develop an extensive familiarity with the problem of interferometry. Specifically, the work done on self-calibrating 3C295

Name	RA [hms]	Dec [dms]	Likely Association	LOFAR thumbnail
EGS-1	14:37:39.53	53:36:31.24	NVSS Radio Galaxy	Fig. 4.6a
EGS-2	14:35:27.84	55:07:56.32	Radio Galaxy in Cluster	Fig. 4.7a
EGS-3	14:29:34.13	54:43:46.93	Radio Galaxy in Cluster	Fig. 4.8a
EGS-4	14:31:36.62	52:27:33.75	Radio galaxy, $z = .292$	Fig. 4.9a
EGS-5	14:29:48.89	51:10:30.52	Lobes matched individually	Fig. 4.10a
EGS-6	14:26:04.09	51:29:35.07	One lobe matched individually	Fig. 4.11a
EGS-7	14:17:55.58	50:08:01.75	Radio galaxy, $z = .186$	Fig. 4.12a
EGS-8	14:14:40.42	51:17:41.00	Radio galaxy, z unknown	Fig. 4.13a
EGS-9	14:11:36.43	52:54:25.53	Galaxy cluster, $z = .525$	Fig. 4.14a
EGS-10	14:07:09.90	55:04:22.23	Galaxy cluster, $z = .250$	Fig. 4.15a
EGS-11	14:03:16.00	51:43:35.23	Lobes matched individually	Fig. 4.16a
EGS-12	14:02:43.49	51:03:14.69	Radio Galaxy in Cluster	Fig. 4.17a

Table 6.1. Table recapitulating the names, positions and likely associations of all 12 chosen sources in the primary beam. These sources were primarily chosen because they had peculiar or interesting diffuse emission. The associated LOFAR image thumbnails are also given for reference.

shows that, while we seem to converge towards the same model regardless of initial conditions (2 point sources, full VLA model, 1 point source⁵), this apparent convergence is very slow. As discussed in Section 2.3, the problem of imaging is convex (if not necessarily well-conditioned) but the problem of calibration is not fully convex: this introduces the possibility of a net non-convexity in the problem of interferometry. Regardless of whether my experience was that of non-convexity or of poor conditioning leading to impractically slow convergence time, the solution is to regularise.

One obvious avenue of inquiry likely to yield results would therefore be in the application of amplitude & phase smoothing on our gain solutions: this corresponds to a regularisation of the calibration inverse problem, improving its conditioning by applying a prior which encourages smooth gain solutions. This prior is physically-motivated, in the sense that we expect the true, underlying gain solutions to be smooth, and so jagged structure in the gains can generally be assimilated to source flux being absorbed into the gain solutions. The tools to perform this smoothing are already developed and being tested on LOTTs DR2 by Cyril Tasse, Tim Shimwell, and Martin Hardcastle. They were not applied here because we used full-Jones calibration solutions (which were found to give better gain solutions with our available SNR early on), whereas the DR2 pipeline solves for scalar gain solutions. Once the smoothing framework is generalised to polarisation, however, its application could result in dramatic improvements in international station gain solutions.

This work is complementary to that of the LOFAR long-baseline team, and relies heavily on their research (e.g. Jackson et al. 2016). In essence, we rely on the fact

⁵This initial model was tested, but not followed through to the same extent as the 2-point-source initial model once it became clear that it was both biased away from the ground truth compared to the other two (as we could expect), but that it seemed to converge - albeit very slowly - towards the same ground truth as them.

that 3C295 is a bright, dominant source in the EGS primary beam: assuming that the overall contribution from fainter sources in the field (which peak at $\sim 4\text{Jy}$, to 3C295's $\sim 99\text{Jy}$ at that frequency) is negligible, then self-calibration on 3C295 ought to yield both an improved model of that source **and** good gain solutions for the international LOFAR stations. One point of originality of this work is that we consistently take into account direction-dependent errors due to decorrelation (but not DDEs) - this has only become possible recently, with the appropriate modeling of decorrelation both in imaging (through the direction-dependent PSF) *and* calibration (by modeling the decorrelation of model sources properly when creating model visibilities). This removes one of the sources of bias in solving the overal problem of interferometry, and allows calibrator sources outside phase centre to be used reliably.

As it turns out, this was a bold assumption to make: the model resulting from our self-calibration is still rife with calibration artefacts. Based on our understanding of the cov-cov relationship, many of these can be understood as being due to sky brightness being absorbed into the gain solutions: gain inspection supports this theory, showing unexpected and correlated structure in our international station gain amplitude solutions.

Other avenues remain open to go beyond this issue: using the full low-resolution primary beam sky model from Chapter 4 could be one way to improve our images. However, we would still find ourselves with much unmodeled emission on the baselines including an international station. Another possible avenue of improvement would be through the use the EHT-imager tools: these rely on fitting model data to closure quantities (amplitude & phase), which are unaffected by antenna-based errors. This is currently being tested by the long baseline working group.

All these avenues of research are, of course, only possible because we find ourselves in very specific circumstances: the presence of a very bright resolved source in a famous (and therefore relatively well-known) extragalactic field. As such, we have the freedom to explore methods which supplement those of the LBCS team, who are more strongly constrained in that their aim is to provide a solid, reliable catalog of sources suitable for calibrating the international LOFAR stations. This, in turns, requires a solid, reliable strategy, in ways that our relatively straightforward application to a single field does not.

Quantifying the impact of international station directional gain errors

This work is strongly linked to the self-calibration on 3C295: depending on the quality of both the model and the gain solutions attained (i.e. of how the quality of our solution to the overall inverse problem of radio interferometry), directional gain errors will be more or less present across the primary beam.

We choose 8 nearby sources from the LBCS, a catalogue of compact radio sources in the Northern sky, spread around both the EGS and 3C295. These are reprinted below, in copies of Table 5.2 and Fig. 5.20. Two direction-dependent effects are expected to arise, one as a function of distance from phase centre (i.e. from the EGS) and the other

as a function of distance from the direction-independent calibration source (i.e. from 3C295).

The direction-dependent effect which is function of distance from the EGS, decorrelation, is modeled in the DDF/kMS pipeline. As such, we do not expect this effect to dominate. The other direction-dependent effect, function of distance from 3C295, however, is not modeled: this effect corresponds to directional gain errors. Direction-independent calibration is only exact in the direction of the calibrator source when a single source is present in the field of view; this is because, in reality, the astrophysical signal from each source in the field of view reaches us through a slightly different ionosphere, are in areas of the antenna beam which are of varying sensitivity, etc. Some of these models can be modeled and corrected to some degree (e.g. LOFAR beam), while others cannot at the time of writing (ionospheric effects).

With our final model of 3C295 and gain solutions, flawed as they are, we find that there is little direction-dependent variation in calibration artefact intensity for the various LBCS sources in the field. This is a strong encouragement to continue the project and move on to imaging the EGS using international LOFAR stations.

6.2.3 Future Work

As mentioned in Section 6.1.3, there is much work ahead of us with regards to the algorithmic work developed in this doctoral thesis. Not only can the theoretical framework be expanded to include more complex cases, but those results which it already provides are limited by the conditioning of our estimation of the visibility covariance matrix. There is therefore room for both expanding on and improving existing results.

As for the application to data, it is very encouraging to see that not only are multiple LBCS sources resolved to show physically-sensible morphologies (FRI, FRII, headtail radio galaxies, etc), but in some instances are even resolved to show not just one, but multiple sources with physically-sensible morphology. Many of the LBCS candidates appear to be double or even triple sources. This would indicate that the directional gain errors for many of the LBCS sources are not so strong that they deteriorate the image beyond the recovery of such morphological information; since calibration artifacts, DDEs included, are more present near bright sources, the fact that they are not strong enough to pollute LBCS sources is a promising indicator that even fainter sources in the EGS have recoverable morphological information, given our data and calibration strategy.

However, we were not able to create images of this field before the end of this doctoral degree: creating this image (and performing the associated quality-control and source extraction) is therefore very high indeed on the priority for future work. The preliminary results shown in this manuscript are not of science-grade quality (no astrometric corrections, proper flux bootstrapping, strategy likely needs revisiting for best results), but they are nevertheless useful markers of the wealth of information available in the EGS as seen with the full LOFAR array.

Bibliography

- Adelman-McCarthy, J. K. & et al. 2009, VizieR Online Data Catalog, 2294
- Ahn, C. P. & et al. 2013, VizieR Online Data Catalog, 5139
- Antonucci, R. R. J. & Miller, J. S. 1985, ApJ, 297, 621
- Best, P. N. 2007, New A Rev., 51, 168
- Best, P. N. & Heckman, T. M. 2012, MNRAS, 421, 1569
- Bonnassieux, E., Tasse, C., Smirnov, O., & Zarka, P. 2017, ArXiv e-prints
- Briggs, D. S. 1995, in Bulletin of the American Astronomical Society, Vol. 27, American Astronomical Society Meeting Abstracts, 1444
- Brogan, C. L., Hunter, T. R., & Fomalont, E. B. 2018, ArXiv e-prints
- Cassano, R., Brunetti, G., Röttgering, H. J. A., & Brüggen, M. 2010, A&A, 509, A68
- Chael, A. A., Johnson, M. D., Bouman, K. L., et al. 2018, ApJ, 857, 23
- Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, AJ, 115, 1693
- Coughlan, C. P., Ainsworth, R. E., Eislöffel, J., et al. 2017, ApJ, 834, 206
- Davis, M., Guhathakurta, P., Konidaris, N. P., et al. 2007, ApJ, 660, L1
- Douglas, J. N., Bash, F. N., Bozian, F. A., Torrence, G. W., & Wolfe, C. 1996, AJ, 111, 1945
- Enriquez, J. E., Siemion, A., Falcke, H., et al. 2017, in American Astronomical Society Meeting Abstracts, Vol. 229, American Astronomical Society Meeting Abstracts #229, 116.04
- Fanaroff, B. L. & Riley, J. M. 1974, MNRAS, 167, 31P
- Fried, D. L. 1978, Journal of the Optical Society of America (1917-1983), 68, 1651
- Girard, J. 2012, PhD thesis, École doctorale Astronomie et Astrophysique d'Île-de-France (Meudon, Hauts-de-Seine)
- Grobler, T. L., Nunhokee, C. D., Smirnov, O. M., van Zyl, A. J., & de Bruyn, A. G. 2014, MNRAS, 439, 4030

- Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35
- Hamaker, J. P., Bregman, J. D., & Sault, R. J. 1996, A&AS, 117, 137
- Helfand, D. J., White, R. L., & Becker, R. H. 2015, ApJ, 801, 26
- Högbom, J. A. 1974, A&AS, 15, 417
- Intema, H. T., Jagannathan, P., Mooley, K. P., & Frail, D. A. 2017, A&A, 598, A78
- Jackson, N., Tagore, A., Deller, A., et al. 2016, A&A, 595, A86
- Jennison, R. C. 1957, in IAU Symposium, Vol. 4, Radio astronomy, ed. H. C. van de Hulst, 159
- Jennison, R. C. 1958, MNRAS, 118, 276
- Joardar, S., Bhattacharyya, S., Bhattacharya, A., & Datta, C. 2010, Progress In Electromagnetics Research B, 22, 73–102
- Kazemi, S. & Yatawatta, S. 2013, MNRAS, 435, 597
- Kellerman, K. & Sheets, B. 1984, Serendipitous discoveries in radio astronomy : proceedings of a workshop held at the National Radio Astronomy Observatory, Green Bank, West Virginia on May 4, 5, 6 1983 (NRAO library)
- Kenyon, J. S., Smirnov, O. M., Grobler, T. L., & Perkins, S. J. 2018a, MNRAS
- Kenyon, J. S., Smirnov, O. M., Grobler, T. L., & Perkins, S. J. 2018b, CubiCal: Suite for fast radio interferometric calibration, Astrophysics Source Code Library
- Lane, W. M., Cotton, W. D., Helmboldt, J. F., & Kassim, N. E. 2012, Radio Science, 47, RS0K04
- Lin, Y.-T., Shen, Y., Strauss, M. A., Richards, G. T., & Lunnan, R. 2010, ApJ, 723, 1119
- Loh, A. 2016, PhD thesis, Université Paris Diderot, École doctorale STEP'UP – ED no 560
- Longair, M. S. 1994, High energy astrophysics. Volume 2. Stars, the Galaxy and the interstellar medium.
- Marecki, A. & Swoboda, B. 2011, A&A, 530, A60
- Maxwell, J. 1864, Royal Society
- McConnachie, A. W., Patton, D. R., Ellison, S. L., & Simard, L. 2009, MNRAS, 395, 255
- Mohan, N. & Rafferty, D. 2015, PyBDSF: Python Blob Detection and Source Finder, Astrophysics Source Code Library
- Moldón, J., Deller, A. T., Wucknitz, O., et al. 2015, A&A, 574, A73

- Morabito, L. K., Williams, W. L., Duncan, K. J., et al. 2017, MNRAS, 469, 1883
- Netzer, H. 2015, ARA&A, 53, 365
- Noordam, J. E. & Smirnov, O. M. 2010, A&A, 524, A61
- Noorishad, P., Wijnholds, S. J., van Ardenne, A., & van der Hulst, J. M. 2012, A&A, 545, A108
- Offringa, A. R. 2010, AOFlagger: RFI Software, Astrophysics Source Code Library
- Pearson, T. J. & Readhead, A. C. S. 1984, ARA&A, 22, 97
- Perley, R. A. & Taylor, G. B. 1991, AJ, 101, 1623
- Pringle, J. E. & Rees, M. J. 1972, A&A, 21, 1
- Rau, U. & Cornwell, T. J. 2011, A&A, 532, A71
- Rengelink, R. B., Tang, Y., de Bruyn, A. G., et al. 1997, A&AS, 124, 259
- Rogers, A. E. E., Cappallo, R. J., Hinteregger, H. F., et al. 1983, Science, 219, 51
- Rogers, A. E. E., Hinteregger, H. F., Whitney, A. R., et al. 1974, ApJ, 193, 293
- Scaife, A. M. M. & Heald, G. H. 2012, MNRAS, 423, L30
- Seyfert, C. K. 1943, ApJ, 97, 28
- Shimwell, T. W., Röttgering, H. J. A., Best, P. N., et al. 2017, A&A, 598, A104
- Smirnov, O. M. 2011a, A&A, 527, A106
- Smirnov, O. M. 2011b, A&A, 527, A107
- Smirnov, O. M. 2011c, A&A, 527, A108
- Smirnov, O. M. 2011d, A&A, 531, A159
- Smirnov, O. M. & Tassee, C. 2015, MNRAS, 449, 2668
- Tassee, C. 2014, ArXiv e-prints
- Tassee, C., Best, P. N., Röttgering, H., & Le Borgne, D. 2008, A&A, 490, 893
- Tassee, C., Hugo, B., Mirmont, M., et al. 2017, ArXiv e-prints
- Tassee, C., Röttgering, H. J. A., Best, P. N., et al. 2007, A&A, 471, 1105
- Tassee, C., van der Tol, S., van Zwieten, J., van Diepen, G., & Bhatnagar, S. 2013, A&A, 553, A105
- Thompson, A. R., Moran, J. M., & Swenson, Jr., G. W. 2001, Interferometry and Synthesis in Radio Astronomy, 2nd Edition
- Ueda, Y., Hatsukade, B., Kohno, K., et al. 2018, ApJ, 853, 24

- van Cittert, P. H. 1934, *Physica*, 1, 201
- van Diepen, G. & Dijkema, T. J. 2018, DPPP: Default Pre-Processing Pipeline, Astrophysics Source Code Library
- van Haarlem, M. P., Wise, M. W., Gunst, A. W., et al. 2013, *A&A*, 556, A2
- van Weeren, R. J., Williams, W. L., , C., et al. 2014, *ApJ*, 793, 82
- van Weeren, R. J., Williams, W. L., Hardcastle, M. J., et al. 2016, *ApJS*, 223, 2
- Varenius, E., Conway, J. E., Martí-Vidal, I., et al. 2015, *A&A*, 574, A114
- Wen, Z. L. & Han, J. L. 2011, *ApJ*, 734, 68
- Wenger, M., Ochsenbein, F., Egret, D., et al. 2000, *A&AS*, 143, 9
- White, R. L., Becker, R. H., Helfand, D. J., & Gregg, M. D. 1997, *ApJ*, 475, 479
- Williams, W. L., van Weeren, R. J., Röttgering, H. J. A., et al. 2016, *MNRAS*, 460, 2385
- Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, *AJ*, 140, 1868
- Xu, X., Arav, N., Miller, T., & Benn, C. 2018, ArXiv e-prints
- Yatawatta, S. 2014, *MNRAS*, 444, 790
- York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, *AJ*, 120, 1579
- Zarka, P., Girard, J. N., Tagger, M., & Denis, L. 2012, in SF2A-2012: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics, ed. S. Boissier, P. de Laverny, N. Nardetto, R. Samadi, D. Valls-Gabaud, & H. Wozniak, 687–694

Résumé

Grâce à une analyse statistique de l'Equation de la Mesure des Interféromètres Radio, un schéma de pondération adaptatif est dérivé, basé sur la qualité de calibration des données d'un instrument interférométrique. Ce schéma est utilisé sur une observation d'un champ extragalactique, l'Extended Groth Strip, observation qui contient une source radio-vive (3C295) dans son champ de vue. Cette source est résolue avec LOFAR-VLBI; un modèle de source est créé afin de calibrer les stations LOFAR internationales. Cela permettra d'imager le champ à une résolution comparable à celle du Hubble Space Telescope, dont des données sont disponibles pour ce champ extragalactique.

Abstract

By performing a statistical analysis of the Radio Interferometer's Measurement Equation, we derive adaptive quality-based weighting schemes. These are deployed on an observation of the Extended Groth Strip, which includes a bright 3C source in the field of view. This source, which is resolved for LOFAR-VLBI, is modeled and used as a calibrator source for the Extended Groth Strip. This will allow the field to be imaged with a resolution matching the Hubble Space Telescope's, of which data are available for this field.

Mots Clés

Interférométrie, radio, LOFAR, algorithme, extragalactique

Keywords

Interferometry, radio, LOFAR, algorithms, extragalactic