THÈSE DE DOCTORAT

de l'Université de recherche Paris Sciences et Lettres

Préparée à l'Observatoire de Paris & UCKAR

Development and Application of New Algorithms for LOFAR & SKA Survey Images

École doctorale n°127

ASTRONOMIE ET ASTROPHYSIQUE D'ILE-DE-FRANCE

Spécialité INTERFÉROMÉTRIE RADIO & ALGORITHMIE

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

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

COMPOSITION DU JURY :

M Pierre Kervella, Président
LESIA

M Neal Jackson, Rapporteur
University of Manchester

M Patrick Charlot, Rapporteur
Laboratoire d'Astrophysique de Bordeaux

Mme Maajke Mevius
ASTRON, Membre du jury

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

M Cyril Tasse, Directeur de Thèse
GEPI

M Oleg Smirnov, Directeur de Thèse
SKA-SA

M Philippe Zarka, Directeur de Thèse
LESIA

Contents

Page

Chapter 1

Foreword

write introduction to the thesis

Chapter 2

Introduction

Contents

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 we will rely on throughout this manuscript, the Radio Interferometer's Measurement Equation, is described in greater details in Chapter 3. We will begin by discussing radio interferometry itself. We will then discuss the so-called *imaging problem*, and end by discussing the *calibration problem.* 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[1] 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.*

2.2 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.

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

[1]And, famously, attempting to observe "intelligent" sources (cf Enriquez et al. 2017)

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 2.5.

So far, the astronomer will have worked entirely with visibilities. However, generally speaking, the aim will be to get information on the object's 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 2.3.5.

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 Chapter 3.

2.3 Interferometry: Bypassing the Diffraction Limit

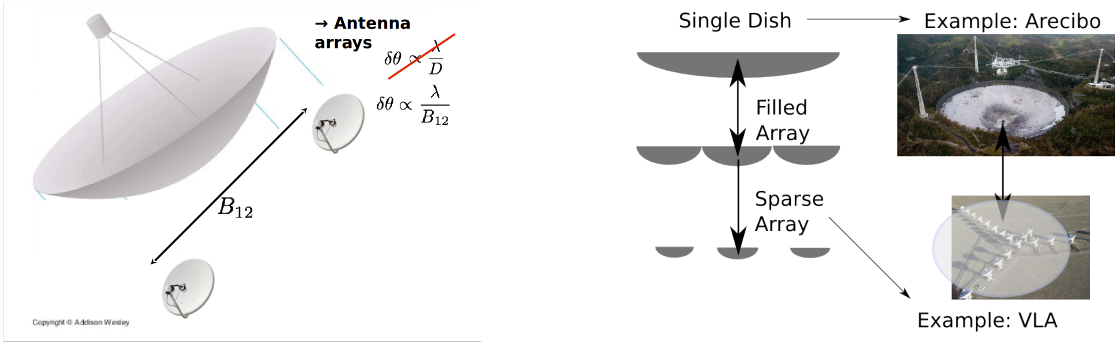
There are two quantities of interest to all astronomers: sensitivity and angular resolution[2]. An instrument's sensitivity is a function of its collecting area[3]. Resolution, for well-designed instruments[4], is limited by diffraction. This introduces specific issues in the radio domain. Radio waves have very long wavelengths - often comparable to meters, rather than the $\sim 100 - 1000$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.

[2]Other kinds of resolutions - 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.



(a) *A pair of dishes can surpass the resolution limit of its components.*

(b) *With enough pairs of dishes, it is possible to synthesise a much larger dish.*

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 observations follow a very familiar pattern: the elements of the interferometric array (which can be single dishes or phased arrays[5] 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 2.3.4). They must be calibrated, by comparing the observations of a calibrator source and modelled visibilities. Calibration is described in much greater detail in **??**. 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.

2.3.1 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.

[5] Described in Section 2.3.2

2.3.2 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

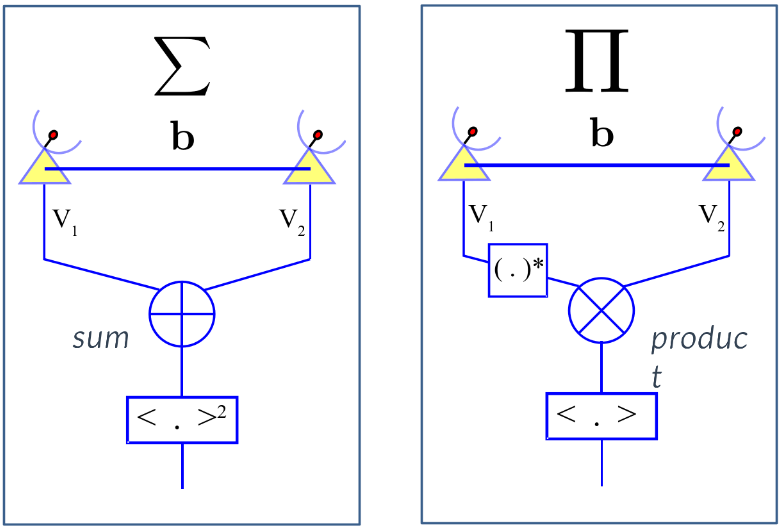


Figure 2.3. *There are two ways to combine the voltages from two antennas into a visibility: they are sum-correlation and π-correlation. In this manuscript, we will only concern ourselves with the latter. Image credit: Julien Girard*

associated with baseline $\mathbf{b}_{AB}$ is created by taking the voltage measured by antenna A, multiply it by the complex conjugate 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:

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

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

where $\langle \cdots \rangle_i$ 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?

Different astrophysical sources are located very far apart, and so their signals do not interfere. They are measured additively in both antennas. Provided that the signal from both 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.

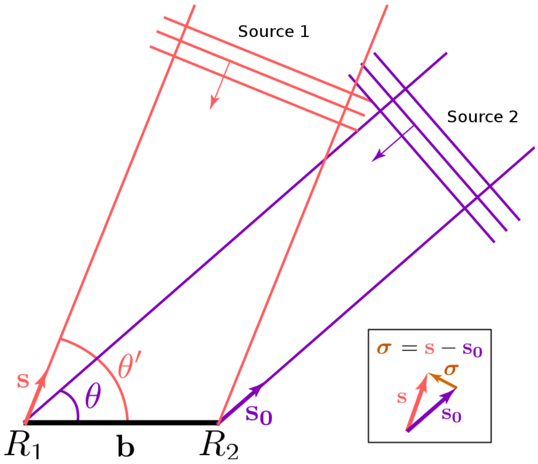


Figure 2.4. *Here, we assume that there are only two sources in the sky whose signal can be measured by our antennas. The final visibility is the sum of the visibilities associated with each individual source. Image credit: Julien Girard add reference to julien's phd*

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 digitally 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 ensures that the constituent antennas' sensitivities are maximised in the direction of interest. 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.

2.3.3 The uv-plane

In general, interferometric design is such that the w component of visibilities' (u, v, w) coordinates is negligible (or can be put in a frame of reference where it can usually be approximated as such). Radio astronomers tend to thus talk of a uv-plane rather than uvw-space to describe visibility space. The set of uv-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.

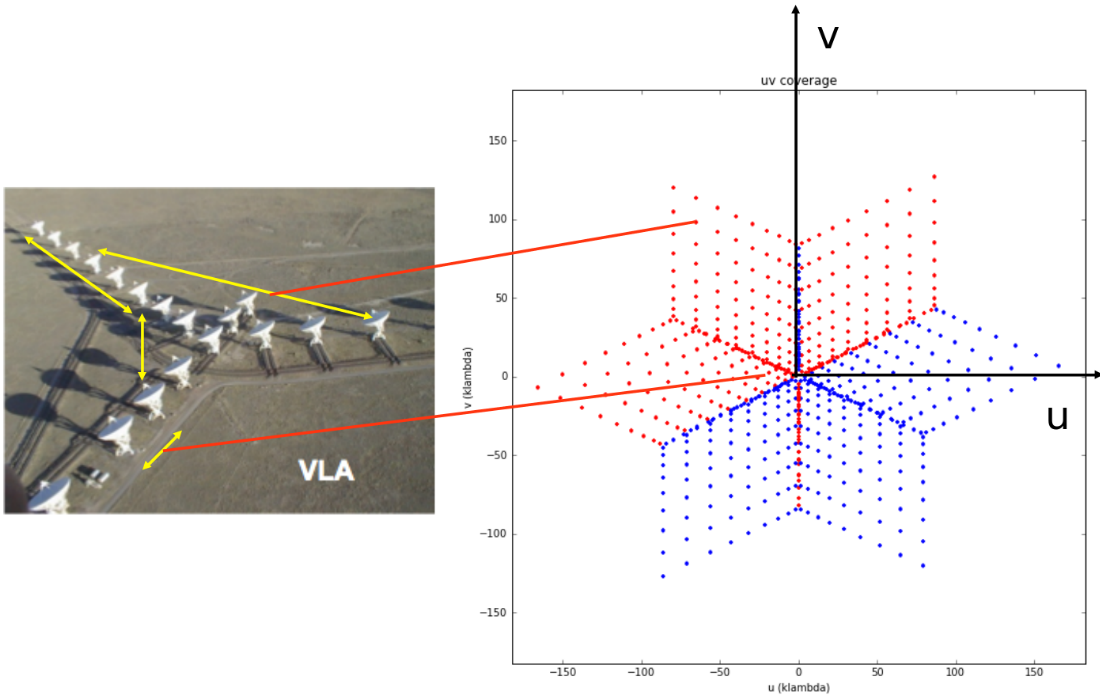


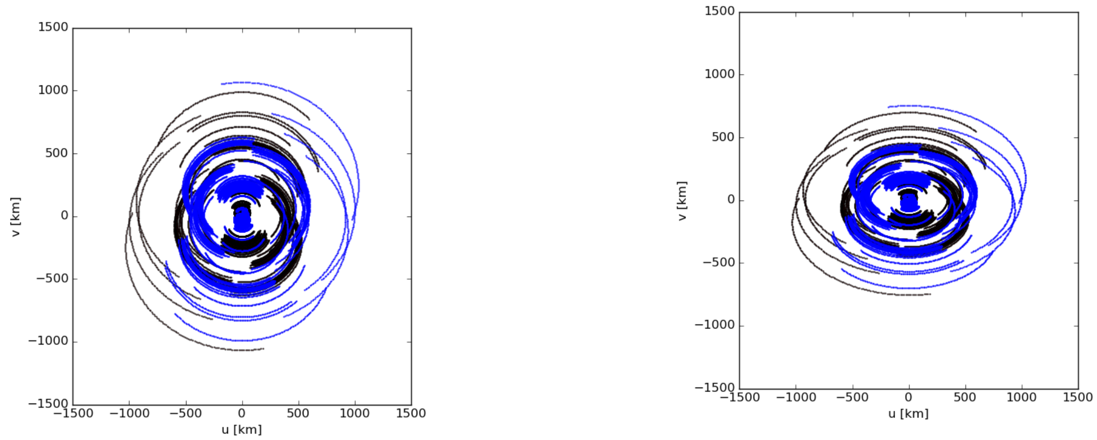
Figure 2.5. *The VLA contains 27 radio dishes placed as shown above. Each antenna pair between those 27 gives two single baselines, here, one red and one blue. Image credit: Julien Girard*

2.3.4 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 each visibility in this direction.

The Point-Spread Function (PSF) of an array is thus the inverse Fourier transform of the array's uv-coverage. The more element 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.



(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.*

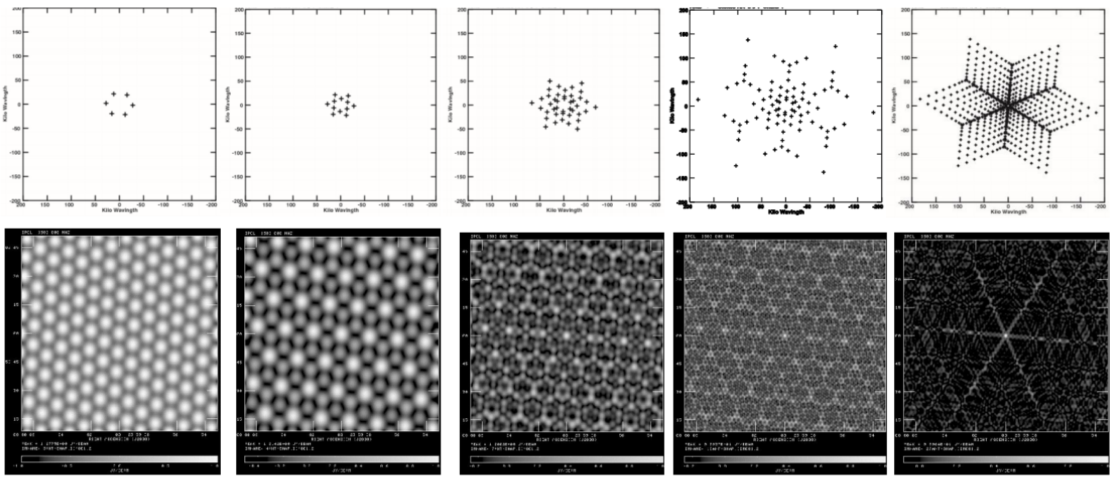


Figure 2.7. *Plot of the PSF associated to more and more complete UV-coverage (here, consisting of more and more elements of the VLA). Image credit: Rick Perley, 3GC4 lecture.crop out garbage on top*

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 $N_{\mathrm{t}} \times N_{\nu}$, 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 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

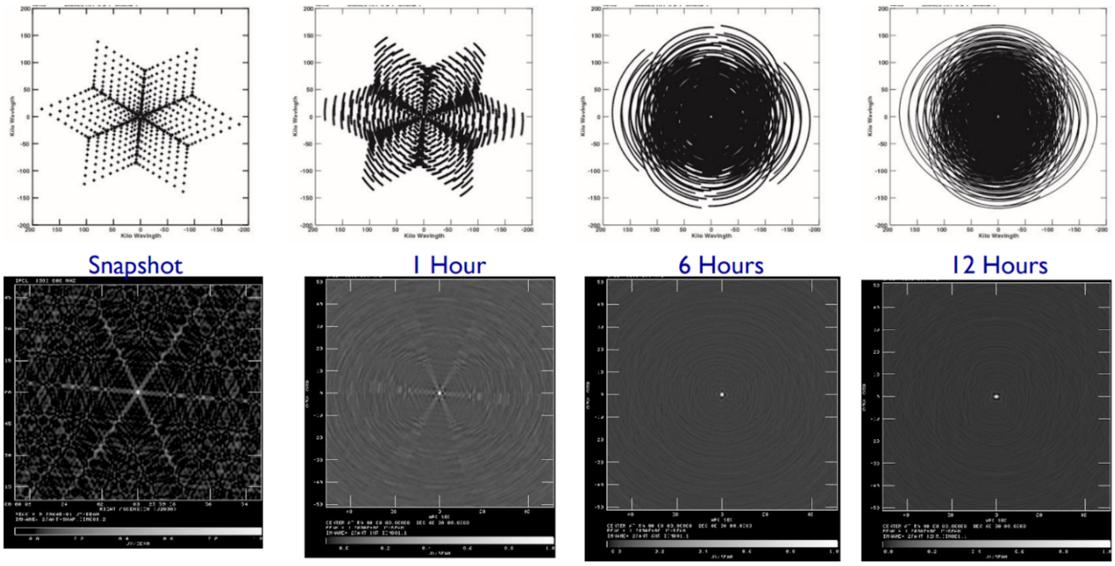


Figure 2.8. *Effect of supersynthesis on the UV-coverage of the full VLA. Image credit: Rick Perley, 3GC4 lecture.*

source) drown out signal from fainter sources. This means *deconvolving* the PSF from radio interferometric images.

2.3.5 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. We define dynamic range as:

$$\mathrm{DR} = \frac{\mathrm{flux}_{\mathrm{brightest}}}{\mathrm{flux}_{\mathrm{faintest}}} \tag{2.3}$$

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) 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 CLEAN[6]. It attempts to optimise for the first requirement while being constrained by the second.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).

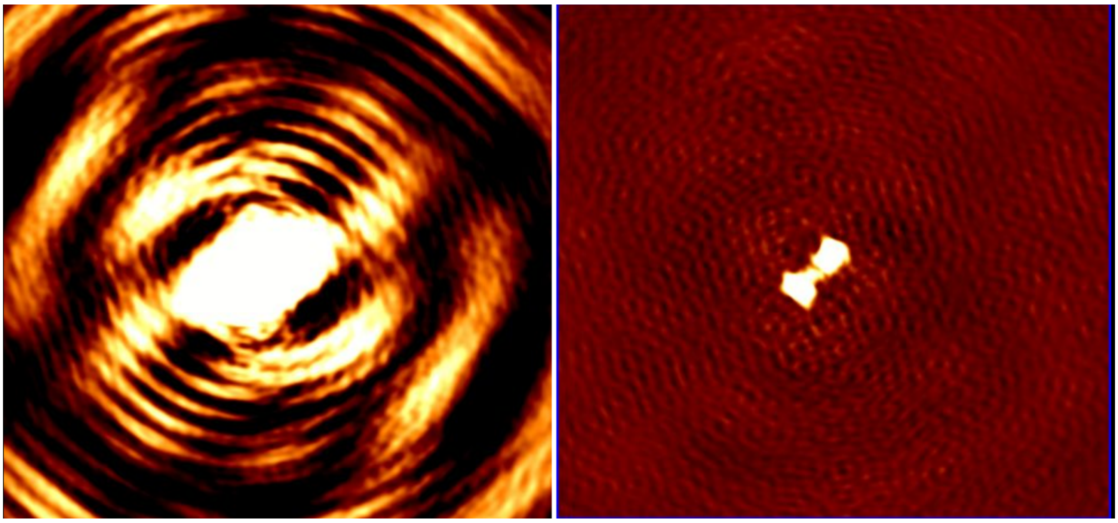


Figure 2.9. *Effect of MetroCLEAN algorithm on an image of 3C295*

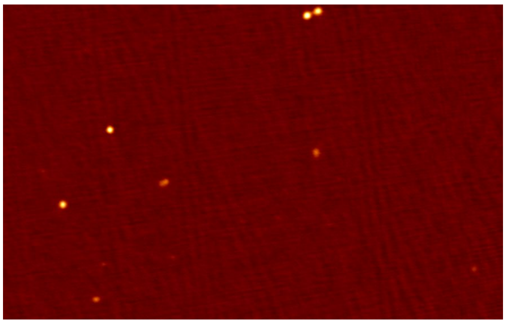
In Fig. 2.9, we show the effect of deconvolution on a dirty image of 3C295 (a very bright & complex source which is the subject of much of this thesis' work). Since this source includes diffuse emission, standard CLEAN is insufficient (since CLEAN assumes a sparse distribution of point-like sources). Instead, the DDFacet software package's novel MetroCLEAN 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.

[6]The 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)).

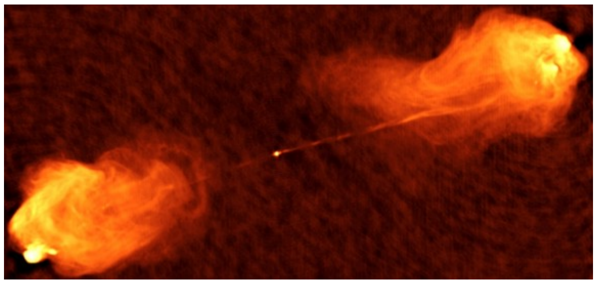
The problem of imaging radio interferometric data, of which deconvolution is but a subset, is still a very active field of research today. For more information on this topic, the NRAO website[7] is an excellent start for a general introduction to this topic, while Tasse et al. (2017) provides an outstanding entry point for cutting-edge work at the time of writing.

2.4 Weighting Schemes

In Section 2.3.5, we discussed the problem of deconvolution in radio interferometric imaging. One consideration that is not raised is the issue of conditioning. Consider two extreme cases: a field containing a few point sources(Fig. 2.10a), and a field containing a diffuse and turbulent source, with very complex flux distribution at all scales (Fig. 2.10b). 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 said to be well-conditioned, and the second to be poorly-conditioned.



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



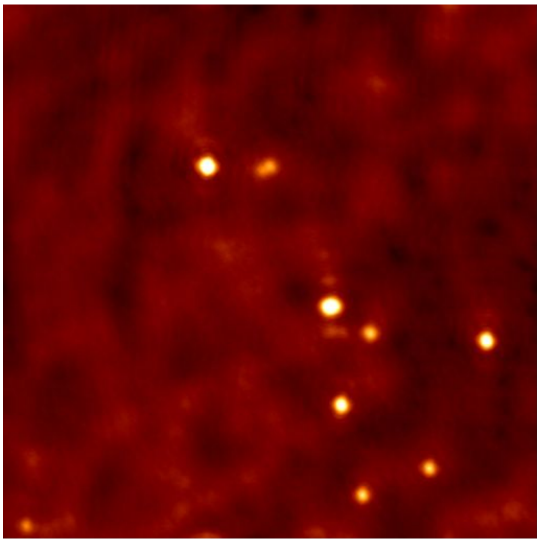
(b) *Cygnus A as seen by the VLA. Image courtesy of NRAO/AUI*

Figure 2.10. *Two extreme examples of good and poor conditioning. Fig. 2.10a shows a field with a few point-like sources. Fig. 2.10b shows a very complex & diffuse structure.*

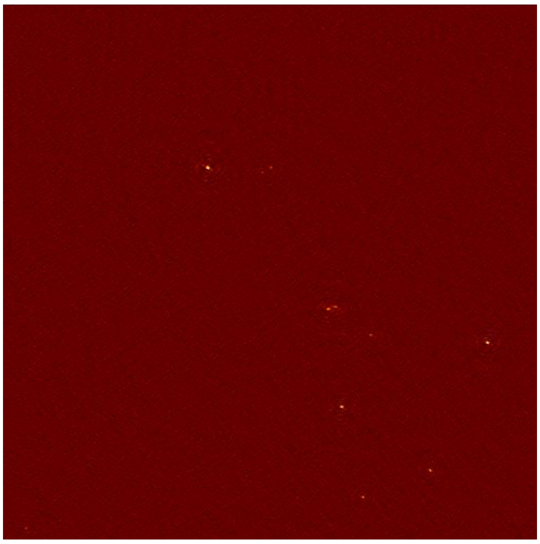
The better the conditioning, the easier the deconvolution and the closer to the ground truth its result. As such, when the true underlying structure of a source is poorly-known, getting the best conditioning possible for deconvolution becomes a key concern. Thankfully, because the actual measurements 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. 2.11 clearly shows that imaging done with the same data but different weighting schemes give radically different deconvolved images. Images made with uniform

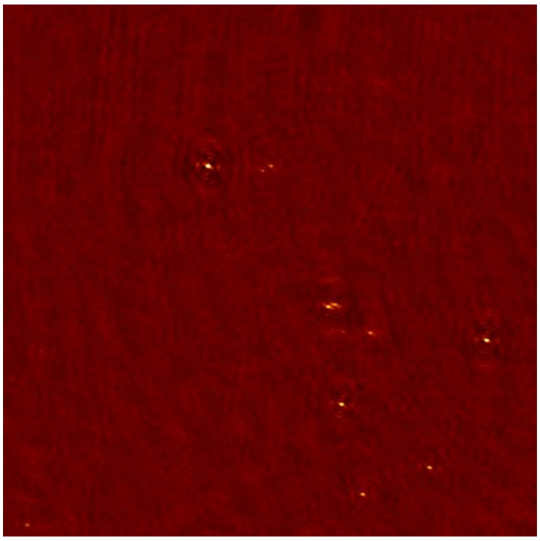
[7]See the CLEAN section of the NRAO website.



(a) *Image made with LOFAR data and natural weighting.*



(b) *Image made with the same data as Fig. 2.11a and uniform weighting.*



(c) *Image made with the same data as Fig. 2.11a and Briggs weighting, with a Briggs parameter of 0.*

Figure 2.11. *Impact of different weighting schemes on interferometric image reconstruction. All the images shown above are deconvolved.*

(Fig. 2.11b) or Briggs (Fig. 2.11c) weighing are much sharper, with uniform weighting (Fig. 2.11b) giving such 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).

Note that the weighting schemes shown here rely only on a choice between normalising visibilities by local *uv*-density vs. signal-to-noise, with Briggs weighting providing a formal method for going from one regime to another or optimising between both. Other weighting schemes exist - one can attempt to optimise data size vs decorrelation[8], for example (see Atemkeng et al. 2016). This can help ensure a homogeneous conditioning of the deconvolution problem in the image, by ensuring that the PSF is as peaked as possible throughout the field of interest. Part of the work of this thesis was in devising a weighting scheme dependent on calibration quality, to drastically improve deconvolution conditioning and final image quality. This weighting scheme is described in much greater detail in Chapter 4.

We close this section by noting that data flagging (cf. Offringa 2010) 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 Interference - RFI) to be scientifically useful. Those weights which are not deemed too corrupted are assigned a unit weight. 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.

[8] As data is averaged, decorrelation becomes more and more important in the field. See Chapter 3 for more details.

2.5 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 Chapter 3. One key metric of calibration quality is the *dynamic range*, mentioned in Sec. 2.3.5. High dynamic ranges mean that a high contrast has been obtained, and fainter sources can be reached. Here, we redefine dynamic range as follows

$$DR = \frac{flux_{max}}{\max(\sigma_{thermal}, \sigma_{artefacts})} \tag{2.4}$$

where $flux_{max}$ is the flux of the brightest source, $\sigma_{thermal}$ 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 Effects[9] 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.

There are three 'generations' of calibration methods, of increasing complexity. 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.

2.5.1 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 a calibrator before and after each observation run to find a gain factor and offset error[10]

Phase calibration in the 1GC era 'proper' was not necessary, as engineers were capable of ensuring adequate phase stability in contemporary interferometers. Phase was thus calculated relative to a fixed frame of reference, usually the central antenna of a 3-antenna array.

In RIME terms, this consists of solving for a very basic form of $\boldsymbol{G}_p$:

$$\boldsymbol{G}_p = \left[a_p t + b_p\right] \tag{2.5}$$

where a_p, b_p are constants solved for during open-loop calibration and t is time.

[9] See Section 3.3

[10] For a concrete example, see the open-loop calibration techniques page of analog.com.

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 antenna[11]. 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*[12], commonly referred to as *self-cal* ((Pearson & Readhead 1984)). As described in Section 3.4.2, this method can only be deployed if the brightness matrix of the sky is the same for all baselines[13] (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 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 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.

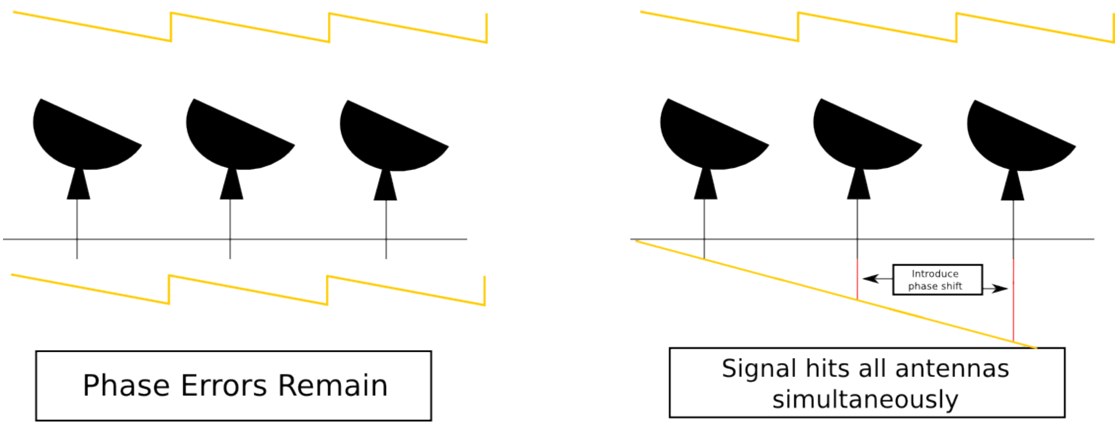


Figure 2.12. *Representation of self-calibration as adaptive optics*

[11] This 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.

[12] On 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).

[13] This 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 sensitivity, in the form of the signal-to-noise ratio (henceforth SNR). 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 matrices[14] in different directions requires image-plane knowledge of the sky brightness distribution **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) 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.

[14] See **??** for a full discussion on calibration and Jones matrices

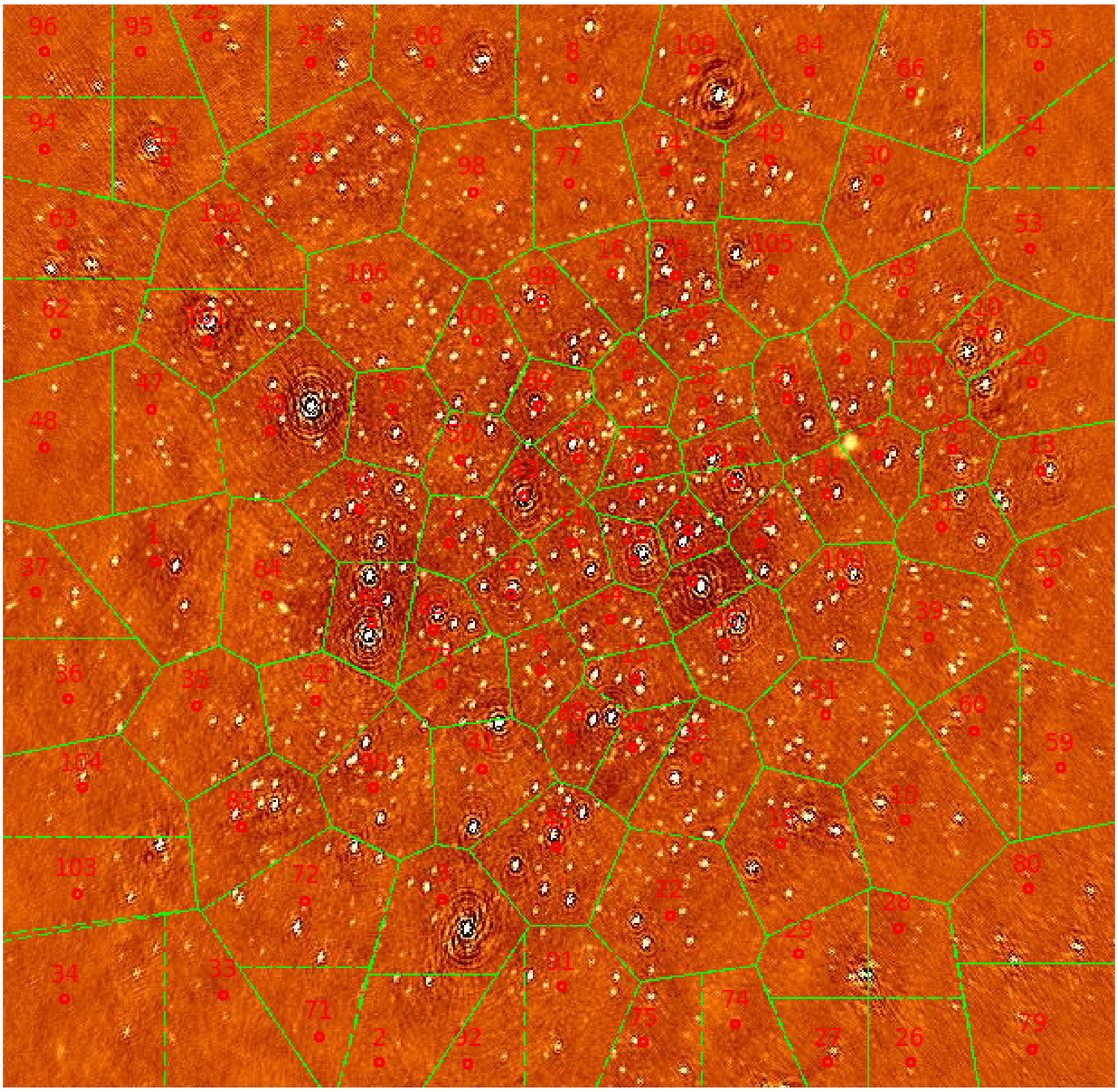
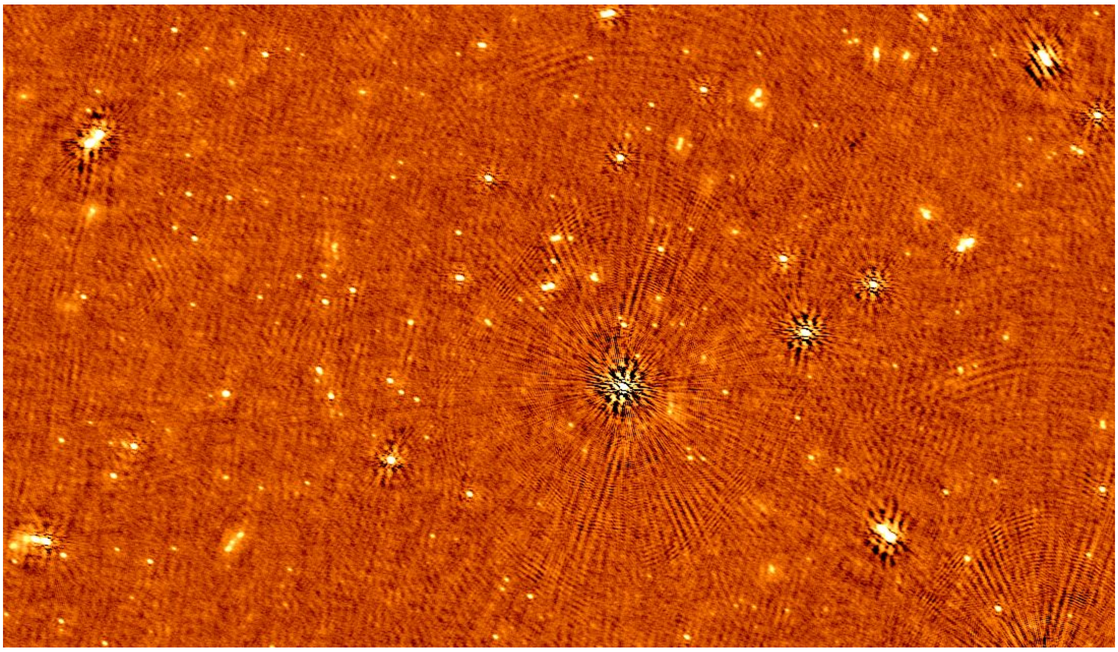
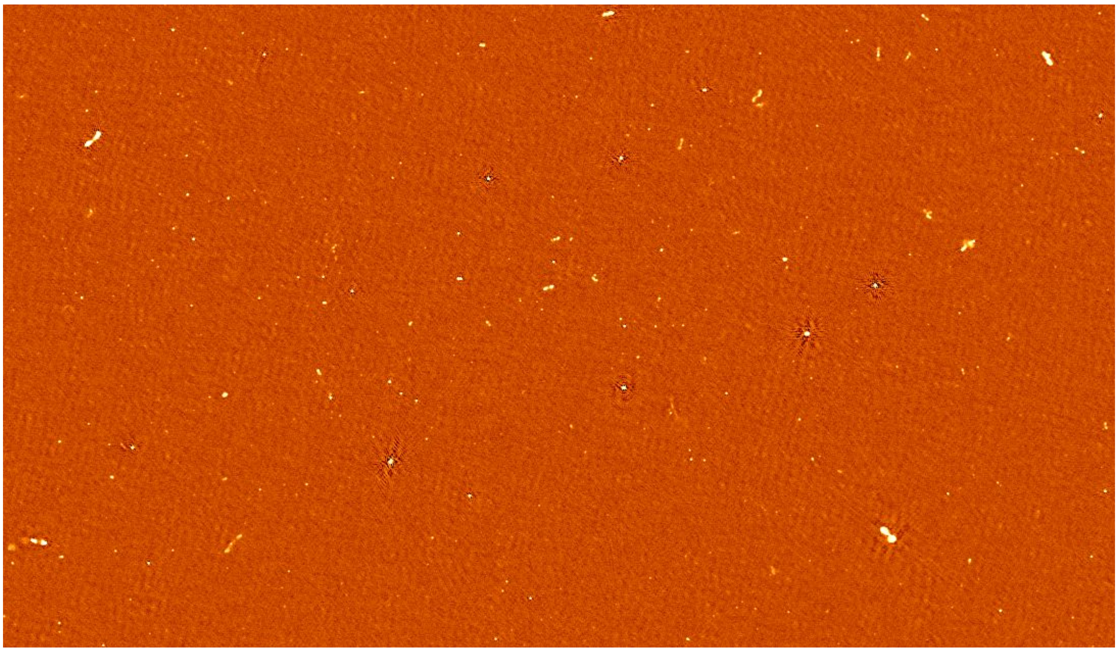


Figure 2.13. *Voronoi facets, as implemented in DDFacet. Image is made before application of direction-dependent correction. This is an image of the Bootes field, made by Cyril Tasse.*

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. 2.13. The facet distribution shown in Fig. 2.13 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. 2.14, 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 2.14. *Difference between imaging made with (Fig. 2.14b)and without (Fig. 2.14a) third-generation calibration.*

2.6 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 thermal radiation detected at low frequencies (associated with dust & protoplanetary disks, and therefore a tracer of star formation) along with 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 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. 2.15[15].

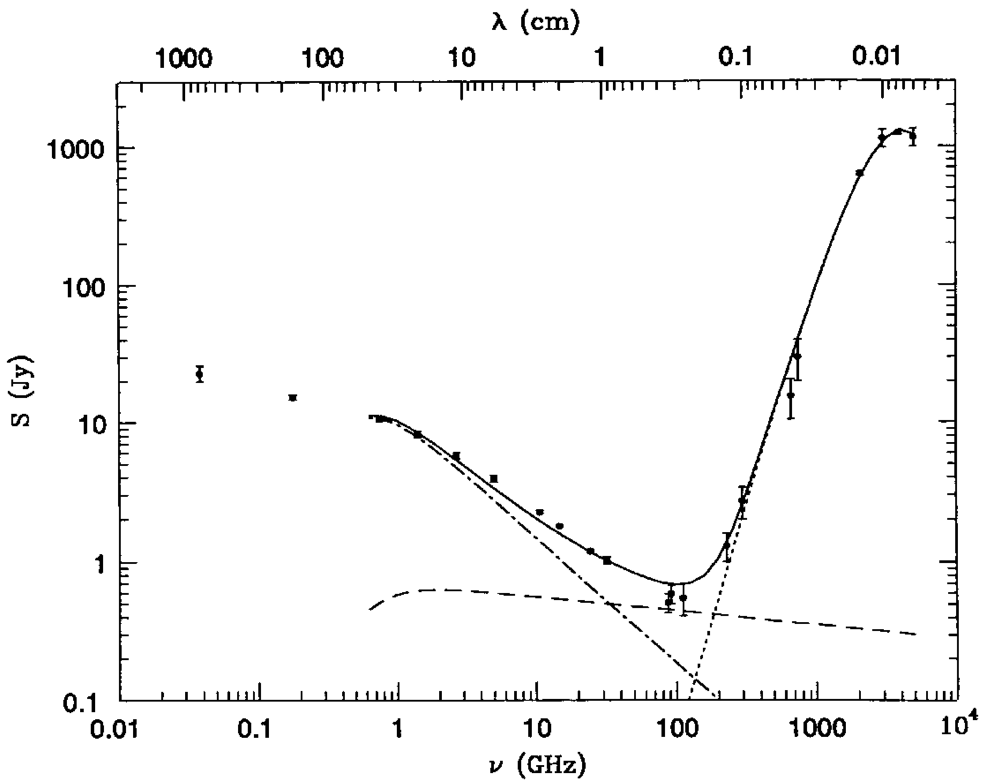


Figure 2.15. *Radio and far-infrared spectrum for galaxy M82, as estimated by the NRAO online course. The flat curve corresponds to free-free emission, while synchrotron radiation and thermal dust emission dominate at low and high frequencies respectively.*

[15]Figure and work taken from the NRAO website. For further information, see here

2.6.1 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 that 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 lectures, as well as Alan Loh's doctoral thesis and the NRAO online course.

Synchrotron radiation (or "magnetobremsstrahlung") occurs when the trajectory of a high-energy charged particle is deflected by a magnetic field. As the German name suggests, it is the magnetic equivalent of free-free radiation. It is the tracer of extremely violent processes, such as AGN jets. 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 angle[16] (Longair 1994). The typical radiation spectrum for a single electron is shown in Fig. 2.16 as a function of the critical frequency, which is itself a function of the emission angle and the synchrotron frequency[17]: $\nu_c = 3/2\nu_{\text{sync}} \sin(\theta)$.

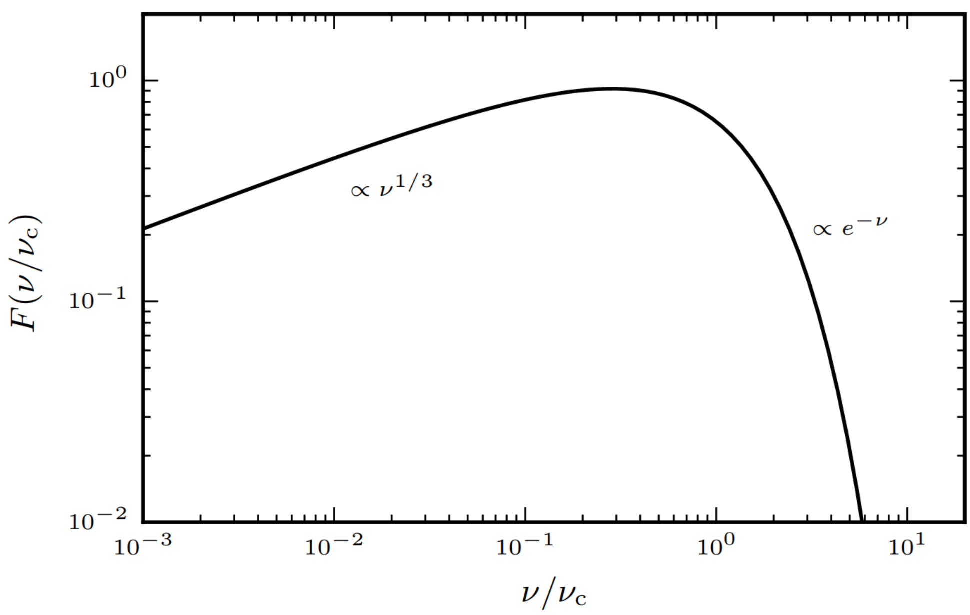


Figure 2.16. *Characteristic spectrum of synchrotron emission for a population of single-energy electrons. Arbitrary unit scale. Source: NRAO online course.*

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

[16] The emission angle is the angle between the electron's velocity and its radial velocity.

[17] Which is itself a function of the Lorentz factor and the magnetic field strength - $\nu_{\text{sync}} = \frac{\gamma^2 eB}{2\pi m_e c}$.

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: $S_\nu \propto \nu^\alpha$, where α is the object's spectral index. 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 synchrotron 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) the underlying electron energy power-law distribution can be seen. An example is shown in Fig. 2.17, taken from Tasse et al. (2007); **?**.

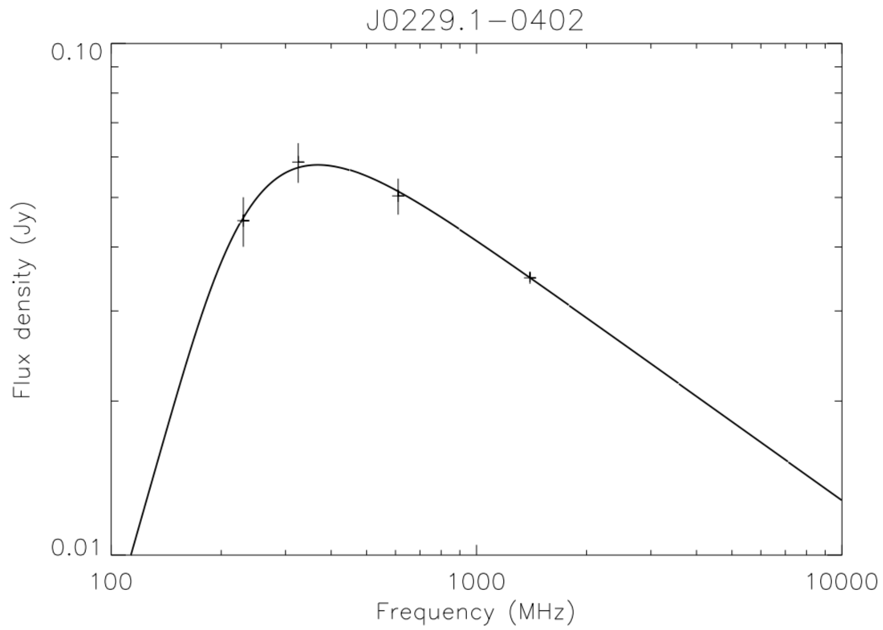


Figure 2.17. *Power spectrum for J0229.1-0402.*

In the optically-thick regime, we know that the most effective 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 approxmation (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 ($S_\nu \propto e^{-\nu}$) and black-body radiation (which has a frequency-dependent α). Empirically, we generally observe $\alpha \sim 0.7$.

2.6.2 Thermal Radiation

Also known as black-body radiation, its spectral intensity is given by Planck's law, given in Eq. 2.6.

$$B_\nu(\nu, T) = \frac{2h\nu^3}{c^2}\left(e^{\frac{h\nu}{k_B T}} - 1\right)^{-1} \tag{2.6}$$

where B_λ is the flux density at frequency ν for a source with temperature T, and $k_B = 1.381\,[J/K]$ is the Boltzmann constant. Near protostellar disks, synchrotron emission is absorbed by the ambient interstellar medium, heating it up to an average of $\sim 10^4$ (see Anderson et al. (2009), Spitzer (1978) and references therein). This gives the following spectral curve:

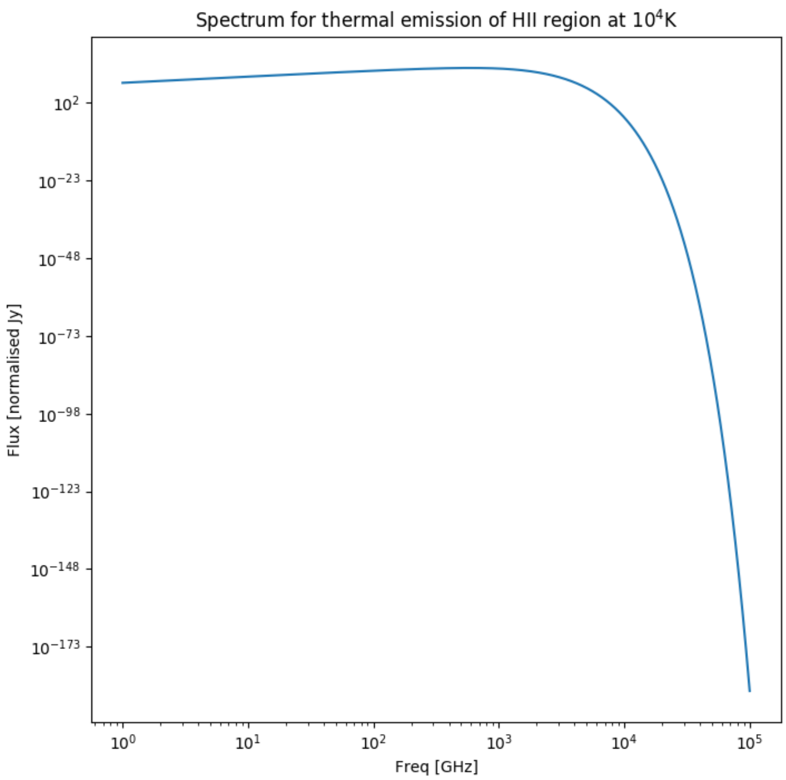


Figure 2.18. *Log-Log plot of blackbody radiation emitted by a particle at 10^4 Kelvin. Note that the y-axis is arbitrary, since it will in practice be modulated by the number of particles in a region, radiation efficiency, resolution etc.*

This is the dominant mode of emission for distant galaxies in the infrared band. In the LOFAR regime, it is not expected to dominate for extragalactic sources, but ought to be detected for resolved galaxies.

2.6.3 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 without being captured.

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. 2.19.

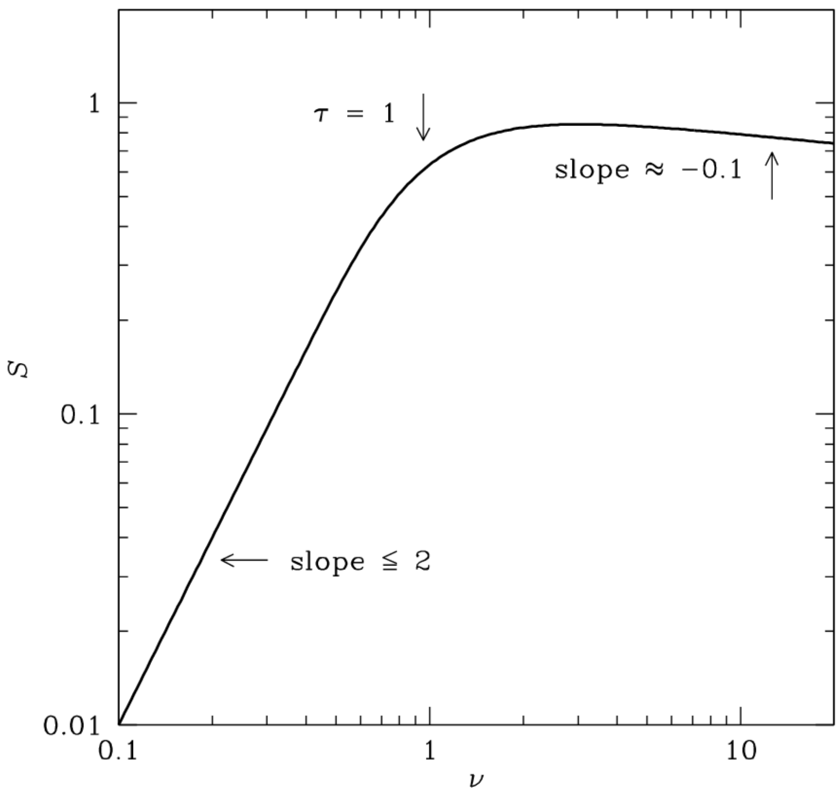


Figure 2.19. *Characteristic spectrum of free-free radiation. Arbitrary unit scale. Source: NRAO online course.*

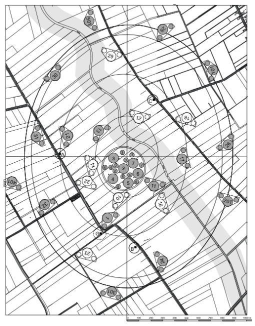
As we can see, this spectrum falls off sharply with frequency. As such, while it is expected to dominate over thermal emission in the absence of synchrotron radiation, it is synchrotron emission that is expected to dominate overall - if present - in the LOFAR band. Free-free emission acts as a tracer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

2.7 LOFAR: The LOw-Frequency Array

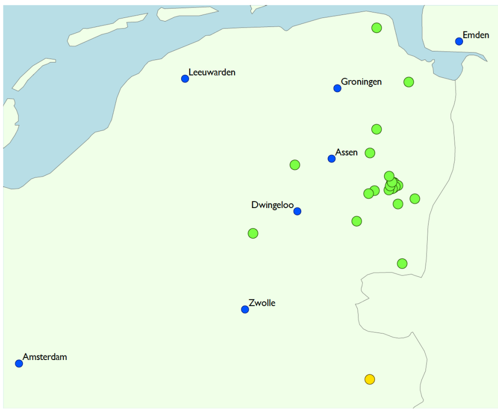
In this section, we will describe the LOw Frequency Array LOFAR (van Haarlem et al. 2013), its technical properties and its current state of the art. In particular, the distinction between "Dutch" LOFAR and "international" LOFAR - and the technical problems associated with each - will be made explicit 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 phase array. It thus consists of antennas which are combined to form stations, which are themselves distributed throughout the Netherlands and Europe.

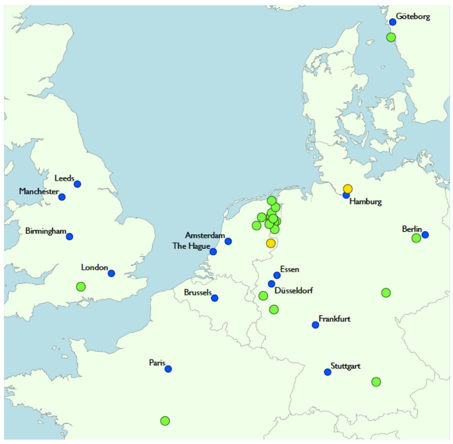
(a) *LOFAR core, known as the Superterp.*



(b) *Distribution of the so-called LOFAR "core" stations, which include the Superterp.*



(c) *LOFAR with both "core" and "remote" stations.*



(d) *International LOFAR. Newer stations (one in Ireland, three in Poland) are not shown here.*

Figure 2.20. *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*

There are two bands to LOFAR, which are known as LOFAR-HBA (High-Band Antennas) and LOFAR LBA (Low-Band Antennas). Figs. 2.20a and 2.21 show the layout of the 6 innermost core stations and the French international LOFAR station, respectively.

Figure 2.21. *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.*

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 48 such tiles, distributed in a "filled-disk" pattern, where the disk fill ratio is determined by the size of the tile. This pattern is shown in Fig. 2.22. Dutch "remote" stations have 96 such tiles spread in 2 disks (HBA_INNER and HBA_OUTER), while "core" station HBA tiles are split into two 24-tile crosses. 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 beam[18]. This station beam is then pointed digitally.

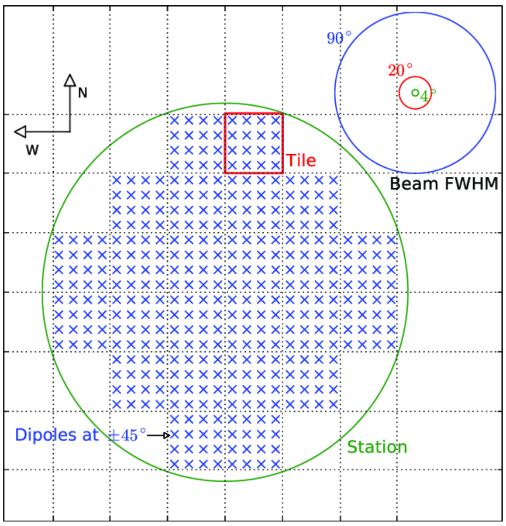


Figure 2.22. *Layout of a single HBA tile. Source: Sarod Yatawatta researchgate profile*

[18]Referece: ASTRON technical description

LOFAR-HBA is sensitive to higher frequencies, from 120 MHz to 240 MHz[19]. It has a smaller field of view than LOFAR-LBA, but a better angular resolution.

LBA antennas, meanwhile, follow a very simple - and cost-effective - design. A typical LBA antenna is shown in Fig. 2.23. 96 such antennas are spread in a semi-random pattern in each LOFAR station[20]. In Dutch stations, 48 of these are used at all times. They are combined as a single phased array.

Figure 2.23. *Picture of a single LBA antenna. Source: LOFAR technology website.*

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 observed phase before averaging the data. In this sense, the pointing is achieved in a similar way as individual HBA tile pointing, albeit in analog. These 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.

[19]Reference: LOFAR.org website
[20]Reference: LOFAR.org website.

Chapter 3

Mathematical Framework: the RIME Formalism

Contents

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 papers[1]. 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.

[1]Smirnov (2011b), Smirnov (2011c) and Smirnov (2011d)

3.1 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.

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 vector $\boldsymbol{e}$. We can then represent physical processes which affect this signal's propagation using *Jones matrices*. In other words:

$$\boldsymbol{v} = \boldsymbol{J}\boldsymbol{e} \tag{3.1}$$

where $\boldsymbol{v}$ is the voltage measured by our antenna, and $\boldsymbol{J}$ the Jones matrix describing the *net* propagation effects - atmospheric, astrophysical, and instrumental - which affect our source's signal. $\boldsymbol{J}$ 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 $\boldsymbol{J}$ is often called the *total Jones matrix*.

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

$$\mathbf{V}_{pq} = 2\boldsymbol{v}_p(\boldsymbol{v}_q)^H \tag{3.2}$$
$$\mathbf{V}_{pq} = 2\boldsymbol{J}_p\boldsymbol{e}(\boldsymbol{J}_q\boldsymbol{e})^H \tag{3.3}$$
$$\mathbf{V}_{pq} = 2\boldsymbol{J}_p\boldsymbol{e}\boldsymbol{e}^H\boldsymbol{J}_q^H \tag{3.4}$$

where $\boldsymbol{J}_p$ corresponds to the signal propagation chain of antenna p, and $\boldsymbol{J}_q$ 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 $\boldsymbol{e}$ outer product.

If we choose to use xyz as our coordinates, with z the direction of propagation of our signal, then Hamaker et al. (1996) show that the following relation holds:

$$2\boldsymbol{e}\boldsymbol{e}^H = 2\begin{bmatrix} < e_x e_x^\star > & < e_x e_y^\star > \\ < e_y e_x^\star > & < e_y e_y^\star > \end{bmatrix} = \begin{bmatrix} I + Q & U + iV \\ U - iV & I - Q \end{bmatrix} = \mathbf{B} \tag{3.5}$$

We can thus write:

$$\mathbf{V}_{pq} = \boldsymbol{J}_p\mathbf{B}\boldsymbol{J}_q^H \tag{3.6}$$

This gives us an elegant formulation of the relationship between the signal emitted by an astrophysical source and its corresponding measured visibility. The Stokes parameters of our source can be directly calculated from our measured visibility, provided $\boldsymbol{J}_p$ and $\boldsymbol{J}_q$ 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.

3.2 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. $\boldsymbol{J}_p$ 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:

$$\boldsymbol{J}_{sp}^H = \prod_1^n \boldsymbol{J}_{nsp}^H = \boldsymbol{J}_{1sp}^H \boldsymbol{J}_{2sp}^H ... \boldsymbol{J}_{nsp}^H \tag{3.7}$$

$$\boldsymbol{J}_{sp} = \boldsymbol{J}_{nsp} ... \boldsymbol{J}_{2sp} \boldsymbol{J}_{1sp} \tag{3.8}$$

Since the Jones terms are applied sequentially to the brightness matrix, starting with $\boldsymbol{J}_{1sp}$, the terms to the right of our decomposition of $\boldsymbol{J}_{sp}$ 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.

3.2.1 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* between our measured voltages $\boldsymbol{v}_p$ and $\boldsymbol{v}_q$. 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*.

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 $\boldsymbol{u}_p = (u_p, v_p, w_p)$. The phase difference between p and phase centre (where, by definition, $\boldsymbol{u} = 0$) for a signal arriving from direction $\boldsymbol{\sigma}$ is then:

$$k_p = 2\pi i((u_p l + v_p m + w_p(n-1))) \tag{3.9}$$

where $n = \sqrt{1 - l^2 - m^2}$. Here, l, m, n are the direction cosines[2] of $\boldsymbol{\sigma}$, and $\boldsymbol{u}$ is defined in units of wavelength[3].

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 be written as:

$$\boldsymbol{K}_p = e^{-ik_p} = e^{-2\pi i(u_p l + v_p m + w_p(n-1))} \tag{3.10}$$

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

[3] This means that, for a given baseline pq, $\boldsymbol{u}$ will vary as a function of frequency.

3.2.2 The Coherency Matrix

The commutation property of $\boldsymbol{K}_p$ 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:

$$\boldsymbol{J}_p = \boldsymbol{G}_p \boldsymbol{K}_p \tag{3.11}$$
$$\tag{3.12}$$

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

$$\mathbf{X}_{pq} = \boldsymbol{K}_p \mathbf{B}_{pq} \boldsymbol{K}_p^H \tag{3.13}$$
$$\mathbf{V}_{pq} = \boldsymbol{G}_p \mathbf{X}_{pq} \boldsymbol{G}_q^H \tag{3.14}$$

This is useful on a conceptual level: the process of *calibration* is now to find a best estimate for $\boldsymbol{G}_p$, 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.

3.3 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. If our sky contains s point sources, then our RIME becomes:

$$\mathbf{V}_{pq} = \sum_s \boldsymbol{J}_{sp} \mathbf{B}_s \boldsymbol{J}_{sq}^H \tag{3.15}$$

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:

$$\boldsymbol{J}_{sp} = G_p E_{sp} K_{sp} \tag{3.16}$$

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:

$$\mathbf{V}_{pq} = \boldsymbol{G}_p \left(\sum_s \boldsymbol{E}_{sp} \boldsymbol{K}_{sp} \mathbf{B}_s \boldsymbol{K}_{sq}^H \boldsymbol{E}_{sq}^H \right) \boldsymbol{G}_q^H \tag{3.17}$$

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 $\boldsymbol{K}$, the direction-dependent effects modeled into $\boldsymbol{E}$, and the direction-independent effects (antenna-based effects) modeled into $\boldsymbol{G}$.

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.

3.4 Diffuse Emission: the Full-sky RIME

3.4.1 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 $\mathbf{B}(\boldsymbol{\sigma})$ over all possible directions. This gives us the following expression:

$$\mathbf{V}_{pq} = \int_{4\pi} \boldsymbol{J}_p(\boldsymbol{\sigma})\mathbf{B}(\boldsymbol{\sigma})\boldsymbol{J}_q^H(\boldsymbol{\sigma})d\Omega \tag{3.18}$$

Here, effects such as our radio 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:

$$\mathbf{V}_{pq} = \int_{l,m} \boldsymbol{J}_p(\boldsymbol{\sigma})\mathbf{B}(\boldsymbol{\sigma})\boldsymbol{J}_q^H(\boldsymbol{\sigma})\frac{dldm}{n} \tag{3.19}$$

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

$$\mathbf{V}_{pq} = \boldsymbol{G}_p\left(\int_{\boldsymbol{l}} \frac{1}{n}\boldsymbol{E}_p(\boldsymbol{l})\mathbf{B}\boldsymbol{E}_q(\boldsymbol{l})\boldsymbol{K}_p(\boldsymbol{l})\boldsymbol{K}_q^H(\boldsymbol{l})dldm\right)\boldsymbol{G}_q^H \tag{3.20}$$

$$= \boldsymbol{G}_p\left(\int_{\boldsymbol{l}} \frac{1}{n}\boldsymbol{E}_p(\boldsymbol{l})\mathbf{B}\boldsymbol{E}_q(\boldsymbol{l})e^{-2\pi i\left(u_{pq}l+v_{pq}m+w_{pq}(n-1)\right)}dldm\right)\boldsymbol{G}_q^H \tag{3.21}$$

where

$$u_{pq} = u_p - u_q \qquad v_{pq} = v_p - v_q \qquad w_{pq} = w_p - w_q \tag{3.22}$$

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:

$$\boldsymbol{W}_p = \frac{1}{\sqrt{n}}e^{-2\pi i w_p(n-1)} \tag{3.23}$$

$$B_{pq} = \boldsymbol{E}_p \boldsymbol{W}_p \mathbf{B} \boldsymbol{W}_q^H \boldsymbol{E}_q^h \tag{3.24}$$

and write our RIME in its final form:

$$\mathbf{V}_{pq} = \boldsymbol{G}_p \Bigg(\int_{\boldsymbol{l}} \mathbf{B}_{pq}(\boldsymbol{l}) e^{-2\pi i \left(u_{pq} l + v_{pq} m + w_{pq}(n-1) \right)} dl dm \Bigg) \boldsymbol{G}_q^H \tag{3.25}$$

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* - there is no *a priori* reason this should be the same for all baselines.

3.4.2 Recovering the van Cittert-Zernike Theorem

Earlier calibration models and algorithm 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 2.5.1. 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 baselines[4], as it allows us to write:

$$\boldsymbol{E}_p(\boldsymbol{l}) \boldsymbol{W}_p(\boldsymbol{l}) = \boldsymbol{E}(\boldsymbol{l}) \boldsymbol{W}(\boldsymbol{l}) \tag{3.26}$$

$$\mathbf{B}_{pq}(\boldsymbol{l}) = \mathbf{B}_{app}(\boldsymbol{l}) = \boldsymbol{E}(\boldsymbol{l}) \boldsymbol{W}(\boldsymbol{l}) \mathbf{B})(\boldsymbol{l}) \boldsymbol{W}^H(\boldsymbol{l}) \boldsymbol{E}^H(\boldsymbol{l}) \tag{3.27}$$

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

$$\mathbf{V}_{pq} = \boldsymbol{G}_p \mathbf{X}_{pq} \boldsymbol{G}_q^H \tag{3.28}$$

where $\mathbf{X}_{pq} = \mathbf{X}(u_{pq}, v_{pq}) = \mathbf{X}(\boldsymbol{u})$. Comparing Eqs. (3.13) and (3.25) directly shows that the coherency matrix is simply the two-dimensional Fourier transform of $\mathbf{B}_{app}(\boldsymbol{l})$. We can thus call $\mathbf{X}(\boldsymbol{u})$ 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 right[5]. A significant consequence of this is that DDEs can be incorporated within the RIME formalism.

3.5 Time-dependent Jones matrices, Smearing, and Decoherence

3.5.1 Time-dependence of Jones matrices

Until this point, we have not considered any kind of sky variability in time. In effect, Eq. (3.28) 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

[4] $\mathbf{B}_{pq}$ would then correspond to the 'true' sky attenuated by the power beam, in the traditional view

[5] Noordm, J.E. 1996, AIPS++, Note 185: The Measurement Equation of a Generic Radio Telescope

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)[6]) For Eq. (3.28) 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

$$\boldsymbol{E}_p(t,\boldsymbol{l}) = \boldsymbol{E}_p(\boldsymbol{l}) = \boldsymbol{E}(\boldsymbol{l}) \quad \forall (t,p) \tag{3.29}$$

The above means that direction-dependent effects must be time-independent for Eq. (3.28) to hold. We can thus split direction-dependent effects into two categories: *trivial* DDEs, which satisfy this condition (and can thus be solved for within the RIME formalism), and *non-trivial* DDEs, which are - as the name indicates - non-trivial to account for within the RIME formalism. An example of a trivial DDE is the primary gain; and example of a non-trivial (i.e. time-variable) DDE is ionospheric turbulence.

3.5.2 Smearing and Decoherence

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 $\boldsymbol{u}$ becomes a *time-dependent* variable.

Furthermore, if we choose to define $\boldsymbol{u}$ in terms of wavelength, $\boldsymbol{u}_p$ 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:

$$\langle \mathbf{V}_{pw} \rangle = \frac{1}{\Delta t \Delta \nu} \int_{t_0}^{t_1} \int_{\nu_0}^{\nu_1} \mathbf{V}_{pq}(t,\nu) d\nu dt \tag{3.30}$$

$$= \frac{1}{\Delta t \Delta \nu} \int_{t_0}^{t_1} \int_{\nu_0}^{\nu_1} \boldsymbol{J}_p(t,\nu) \mathbf{B} \boldsymbol{J}_q^H(t,\nu) \tag{3.31}$$

$\boldsymbol{J}(t,\nu)$ 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 *decoherence*, and the specific case of decoherence caused by the $\boldsymbol{K}$ term is referred to as *smearing.*

[6](link to paper)

Smearing increases with baseline length ($\boldsymbol{u}_{pq}$) 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 Chapter 8).

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:

$$\langle \mathbf{V}_{pw} \rangle \simeq \operatorname{sinc}\frac{\Delta\boldsymbol{\Psi}}{2} \ \operatorname{sinc}\frac{\Delta\boldsymbol{\Phi}}{2} \ \mathbf{V}_{pq}(t_{mid}, \nu_{mid}) \tag{3.32}$$

where

$$t_{mid} = (t_0 + t_1)/2 \tag{3.33}$$

$$\nu_{mid} = (\nu_0 + \nu_1)/2 \tag{3.34}$$

$$\Delta\boldsymbol{\Psi} = \arg \mathbf{V}_{pq}(t_1, \nu_{mid}) - \arg \mathbf{V}_{pq}(t_0, \nu_{mid}) \tag{3.35}$$

$$\Delta\boldsymbol{\Phi} = \arg \mathbf{V}_{pq}(t_{mid}, \nu_1) - \arg \mathbf{V}_{pq}(t_{mid}, \nu_0) \tag{3.36}$$

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. (3.17), then it can be calculated for each source individually.

3.6 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 noise[7], which means that a RIME of the same form as Eq. (3.17) would have to be formuled at follows:

$$\mathbf{V}_{pq} = \boldsymbol{G}_p \left(\sum_s \boldsymbol{E}_{sp} \boldsymbol{K}_{sp} \mathbf{B}_s \boldsymbol{K}_{sq}^H \boldsymbol{E}_{sq}^H \right) \boldsymbol{G}_q + \sigma_{pq} \tag{3.37}$$

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.

[7] Detailed treatment given in (Thompson et al. 2001), section 6.2

Chapter 4

Analysing the Variance of Gain Solutions

Contents

4.1 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 problems for astronomers to invert: calibration and imaging. Both of these problems are ill-conditioned.

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 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 gain variation is generally slower than some given scale: 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 pixels[1]. 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 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, 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

[1]The 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".

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. 4.4).

The remainder of this chapter is split into four main sections. In Section 4.2, 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 4.3. We follow in Section 4.4 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.

4.2 Building the Noise Map

In this section, we derive our first fundamental result: the Cov-Cov relationship, Eq. 4.25, 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 paper are given in Table 4.1, along with the scalar numbers used to denote specific dimensions. All other variables are scalars.

4.2.1 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:

$$\boldsymbol{V}_{pq}^{t\nu} = \sum_d \boldsymbol{K}_{p,t\nu}^d \boldsymbol{J}_{p,t\nu}^d \boldsymbol{B}_\nu^d \left(\boldsymbol{J}_{q,t\nu}^d\right)^H \left(\boldsymbol{K}_{q,t\nu}^d\right)^H + \boldsymbol{N} \tag{4.1}$$

All the quantities above are 2×2 matrices. Eq. 4.1, implies a linear relationship between the coherency matrix $\boldsymbol{B}_\nu^d$ and the visibilities recorded by a given baseline ($\boldsymbol{V}_{pq}^{t\nu}$), with the addition of a thermal noise matrix $\boldsymbol{N}$, 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 $\boldsymbol{B}_\nu^d$ is not a function of time. The Jones matrices ($\boldsymbol{J}_{\dots,t\nu}^d$) contain the antenna gain information in matrix form, while $\boldsymbol{K}_{\dots,t\nu}^d$ is the Fourier kernel. Let us limit ourselves to the scalar case, which corresponds to assuming that emission is unpolarised. We assume that $\boldsymbol{B}^d = s_d \mathbf{I}$, where s is the flux of our single point source and $\mathbf{I}$ is the 2×2 identity matrix. We also assume that $\boldsymbol{J}_{p,t\nu} = g_{p,t\nu}\mathbf{I}$, 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 $\boldsymbol{J}_{\dots,t\nu}^d$ becomes $\boldsymbol{J}_{\dots,t\nu}$. Similarly, $\boldsymbol{K}_{p,t\nu}^d = k_{p,t\nu}^d \mathbf{I}$,

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	
$\tilde{\boldsymbol{y}}$	Residual image vector, size n_{pix}
$\boldsymbol{\epsilon}$	Vector of ϵ, size n_{pix}
$\widetilde{\boldsymbol{\gamma}}$	Contains gain products, size n_b. See Eq. 4.12
$\mathbf{1}$	Vector containing 1 in every cell, size n_b
$\boldsymbol{\delta u}_{bb'}$	Vector of coordinates in differential Fourier plane, of length 3.
$\boldsymbol{l}_d$	Vector of sky coordinates, of length 3.
Matrices	
$\boldsymbol{V}_{pq}^{t\nu}$	Visibility seen by a baseline pq at time and frequency t, ν. Size 2×2.
$\boldsymbol{K}_{p,t\nu}^{d}$	Fourier kernel for direction d, antenna p and one (t, ν) pair. Size 2×2.
$\boldsymbol{J}_{p,t\nu}$	Jones matrix for antenna p for one (t, ν) pair. Contains the gains information. Size 2×2
$\boldsymbol{B}$	Sky brightness distribution matrix, of size 2×2.
$\boldsymbol{N}$	Noise matrix, of size 2×2. Contains a single realisation n of the thermal noise in each cell.
$\boldsymbol{\mathcal{F}}$	Fourier transform matrix, of size $n_{\text{pix}} \times n_{\text{pix}}$.
$\boldsymbol{\mathcal{S}}_b$	Baseline selection matrix, which picks out 1 visibility out of the full set. Size $n_{\text{pix}} \times n_b$
$\boldsymbol{\mathcal{C}}_b$	$n_{\text{pix}} \times n_{\text{pix}}$ convolution kernel that defines the PSF.
$\boldsymbol{\mathcal{F}}_{bb'}$	Convolution matrix mapping one $\delta u \delta v$ to $\delta l \delta m$. The set of all $\boldsymbol{\mathcal{F}}_{bb'}$ determines the noise-PSF. Size $n_{\text{pix}} \times n_{\text{pix}}$.

Table 4.1. *Table recapitulating the meaning and dimensions of vectors and matrices used in Sec. 4.2.1. Only scalars which give matrix dimensions or indices are given here.*

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

$$\epsilon \sim \mathcal{N}(0, \sigma) + i\mathcal{N}(0, \sigma) \tag{4.2}$$

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. 4.1:

$$V_{pq}^{\tau} = \left(\sum_d s_d k_{p,\tau}^d \overline{k_{q,\tau}^d}\right) g_p^{\tau} \overline{g_q^{\tau}} + \epsilon \tag{4.3}$$

$$k_{p,\tau}^d = \exp\left(2\pi i \left(u_{p,\tau} l_d + v_{p,\tau} m_d + w_{p,\tau}\left(n_p - 1\right)\right)\right) \tag{4.4}$$

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. $\widehat{g_p^{\tau}}$ 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:

$$\tilde{r}_{pq}^{\tau} = \sum_d s_d \left(k_{p,\tau}^d \overline{k_{q,\tau}^d}\right)\left(g_p^{\tau}\overline{g_q^{\tau}} - \widehat{g_p^{\tau}}\overline{\widehat{g_q^{\tau}}}\right) + \epsilon \tag{4.5}$$

The flux values in the image-plane pixels[2] are the Fourier transform of the visibility values mapped onto each pixel. This can be written as follows:

$$\tilde{\boldsymbol{y}} = \begin{pmatrix} \vdots \\ \sum_{pq} I_{pq,lm}^r \\ \vdots \end{pmatrix} \tag{4.6}$$

$$I_{pq,lm}^r = \sum_{\tau} \omega_{pq,\tau} \tilde{r}_{pq}^{\tau} k_{pq,\tau}^{lm} \tag{4.7}$$

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:

$$\tilde{\boldsymbol{y}}_b = \boldsymbol{\mathcal{F}}^H \boldsymbol{\mathcal{S}}_b \omega_b \left(\kappa_b \tilde{\gamma} + \boldsymbol{\epsilon}\right) \tag{4.8}$$

$$\tilde{\boldsymbol{y}} = \sum_b \tilde{\boldsymbol{y}}_b \tag{4.9}$$

where $\boldsymbol{\epsilon}$ is a vector of the ϵ of Eq. 4.2 and $\tilde{\boldsymbol{y}}$ is a vector of size n_{pix}, with

$$\kappa_b = \sum_d s_d \kappa_b^d \tag{4.10}$$

$$\kappa_b^d = k_{p,\tau}^d \overline{k_{q,\tau}^d} = k_{pq,\tau}^{lm} \tag{4.11}$$

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

$$\widetilde{\gamma}_b = g_p^\tau \overline{g_q^\tau} - \widehat{g_p^\tau}\overline{\widehat{g_p^\tau}} \tag{4.12}$$

$$\tilde{\gamma} = \begin{pmatrix} \vdots \\ \widetilde{\gamma}_b \\ \vdots \end{pmatrix} \tag{4.13}$$

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 $\tilde{\boldsymbol{y}}$ then becomes the average of all $\tilde{\boldsymbol{y}}_b$. $\tilde{\gamma}$ is a vector of all $\widetilde{\gamma}_b$, and thus of size n_b. $\boldsymbol{\mathcal{F}}$ is the Fourier kernel, of size $n_{\text{pix}} \times n_{\text{pix}}$. $\boldsymbol{S}_b$ is a matrix of size $n_{pix} \times n_b$: its purpose is to encode the uv-coverage. Each $\boldsymbol{S}_b$ contains only a single non-zero cell, different for different $\boldsymbol{S}_b$. The height (number of rows) of $\boldsymbol{S}_b$ is determined 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 $(\kappa_b\tilde{\gamma} + \boldsymbol{n})$ is assigned some weight w_b and its uv-coordinates are set by $\boldsymbol{\mathcal{S}}_b$. The inverse Fourier transform $(\boldsymbol{\mathcal{F}}^H)$ 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.

4.2.2 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, $\text{Cov}\{\tilde{\boldsymbol{y}}\}$. The diagonal of this matrix gives the variance for each pixel, while the wings give the covariance between pixels. Using the property that $\text{Cov}\{\boldsymbol{A}\boldsymbol{x}\} = \boldsymbol{A}\text{Cov}\{\boldsymbol{x}\}\boldsymbol{A}^H$, we can apply the Cov{} operator to Eq. 4.9 to write:

$$\text{Cov}\{\tilde{\boldsymbol{y}}\} = \sum_{bb'} \boldsymbol{\mathcal{F}}^H \underbrace{w_b w_{b'} \kappa_b \overline{\kappa_{b'}}}_{\stackrel{\text{def}}{=} \phi_{bb'}} \boldsymbol{S}_b \text{Cov}\{\tilde{\gamma}\} \boldsymbol{S}_{b'}^T \boldsymbol{\mathcal{F}} \tag{4.14}$$

$$+ \sum_b w_b^2 \sigma^2 \underbrace{\boldsymbol{\mathcal{F}}^H \boldsymbol{S}_b \mathbf{I} \boldsymbol{S}_b^T \boldsymbol{\mathcal{F}}}_{\stackrel{\text{def}}{=} \boldsymbol{\mathcal{C}}_b} \tag{4.15}$$

$$= \sum_{bb'} \phi_{bb'} \boldsymbol{\mathcal{F}}^H \boldsymbol{S}_b \text{Cov}\{\tilde{\gamma}\} \boldsymbol{S}_{b'}^T \boldsymbol{\mathcal{F}} + \sum_b w_b^2 \sigma^2 \boldsymbol{\mathcal{C}}_b \tag{4.16}$$

So far, we have only applied definitions. The net effect of $\boldsymbol{S}_b\text{Cov}\{\tilde{\gamma}\}\boldsymbol{S}_{b'}^T$ (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. $\boldsymbol{\mathcal{S}}_b$ 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 $\boldsymbol{S}_b\text{Cov}\{\tilde{\gamma}\}\boldsymbol{S}_{b'}^T = [\text{Cov}\{\tilde{\gamma}\}]_{bb'}\boldsymbol{S}_b\mathbf{1}\mathbf{1}^T\boldsymbol{S}_{b'}^T$, 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:

$$\text{Cov}\{\tilde{\boldsymbol{y}}\} = \sum_{bb'} \phi_{bb'} [\text{Cov}\{\tilde{\gamma}\}]_{bb'} \boldsymbol{\mathcal{F}}_{bb'} + \sum_b w_b^2 \sigma^2 \boldsymbol{\mathcal{C}}_b \tag{4.17}$$

$$\text{with} \qquad \boldsymbol{\mathcal{F}}_{bb'} = (\boldsymbol{\mathcal{F}}_b)^H \underbrace{\boldsymbol{\mathcal{F}}_{b'}}_{\stackrel{\text{def}}{=} \mathbf{1}^T \boldsymbol{S}_{b'}^T \boldsymbol{\mathcal{F}}} \tag{4.18}$$

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

$$\boldsymbol{\mathcal{F}}_{bb'}[d, d'] = e^{2i\pi(u_b l_d - u_{b'} l_{d'} + v_b m_d - v_{b'} m_{d'} + (n_d - 1)w_b - (n_{d'} - 1)w_{b'})} \tag{4.19}$$

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:

$$\begin{aligned} \text{Cov}\{\tilde{\boldsymbol{y}}\} = &\sum_b \left(\phi_{bb}[\text{Cov}\{\tilde{\gamma}\}]_{bb} + w_b^2\sigma^2\right)\boldsymbol{\mathcal{C}}_b \\ &+ \sum_{b, b' \neq b} \phi_{bb'}[\text{Cov}\{\tilde{\gamma}\}]_{bb'}\boldsymbol{\mathcal{F}}_{bb'} \end{aligned} \tag{4.20}$$

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 $\kappa_b^d \kappa_{b'}^{d'} = (\kappa_b^d)^2 \delta_{bb'}$)[3] we can write:

$$\phi_{bb'} = w_b w_{b'} \kappa_b \overline{\kappa_{b'}} \tag{4.21}$$

$$= w_b w_{b'} \left(\sum_d s_d \kappa_b^d\right)\left(\sum_{d'} s_{d'} \overline{\kappa_{b'}^{d'}}\right) \tag{4.22}$$

$$\approx \sum_d w_b w_{b'} s_d^2 \kappa_b^d \overline{\kappa_{b'}^d} \tag{4.23}$$

$$\phi_{bb'} \approx \sum_d \phi_{bb'}^d \tag{4.24}$$

Note that $\boldsymbol{\mathcal{F}}_{bb} = \boldsymbol{\mathcal{C}}_b$, since those are the coordinates along the diagonal: for these values, the matrix-of-ones at the centre of $\boldsymbol{\mathcal{F}}_{bb'}$ 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. 4.20 as:

$$\begin{aligned} \text{Cov}\{\tilde{\boldsymbol{y}}\} = \sum_d \Bigg(&\sum_b \phi_{bb}^d \left([\text{Cov}\{\tilde{\gamma}\}]_{bb} + \frac{w_b^2\sigma^2}{\phi_{bb}}\right)\boldsymbol{\mathcal{C}}_b \\ &+ \sum_{b, b' \neq b} \phi_{bb'}^d [\text{Cov}\{\tilde{\gamma}\}]_{bb'}\boldsymbol{\mathcal{F}}_{bb'}\Bigg) \end{aligned} \tag{4.25}$$

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

[3] This hypothesis is equivalent to assuming that the sky is dominated by distant point sources, where "distant" means that the sources are multiple PSF Full-Width Half-Maximum apart

This is our fundamental result: assuming unpolarised emission coming from distant 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 Diag {} operator (which returns the diagonal of an input matrix as a vector) to both sides of Eq.4.25, we can find an expression for the variance map in the image-plane:

$$\mathrm{Var}\{\tilde{\boldsymbol{y}}\} = \mathrm{Diag}\left\{\mathrm{Cov}\{\tilde{\boldsymbol{y}}\}\right\} \tag{4.26}$$

$$= \sum_d \left(\sum_b \left(\phi_{bb}^d [\mathrm{Cov}\{\tilde{\gamma}\}]_{bb} + w_b^2\sigma^2\right) \underbrace{\mathrm{Diag}\left\{\boldsymbol{\mathcal{C}}_b\right\}}_{=\mathbf{1}} + \sum_{b,b'\neq b} \phi_{bb'}^d [\mathrm{Cov}\{\tilde{\gamma}\}]_{bb'} \mathrm{Diag}\left\{\boldsymbol{\mathcal{F}}_{bb'}\right\} \right) \tag{4.27}$$

$$\text{where} \quad \boldsymbol{l}_d = (l_d, m_d, (n_d - 1)) \tag{4.28}$$

$$\boldsymbol{\delta u}_{bb'} = (\delta u_{bb'}, \delta v_{bb'}, \delta w_{bb'}) \tag{4.29}$$

$$= (u_b - u_{b'}, v_b - v_{b'}, w_b - w_{b'}) \tag{4.30}$$

In Eq 4.27, we have:

$$\mathrm{Diag}\left\{\boldsymbol{\mathcal{F}}_{bb'}\right\}[d] = e^{2i\pi \boldsymbol{l}_d \cdot \boldsymbol{\delta u}_{bb'}} \tag{4.31}$$

For $b \neq b'$, the diagonals of $\boldsymbol{\mathcal{F}}_{bb'}$ are the Fourier kernels mapping $\delta u \delta v$ space to $\delta l \delta m$. $\boldsymbol{\mathcal{F}}_{bb'}$ 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. 4.1 shows, for a given uv-track (Fig. 4.1b), both the corresponding $\delta u \delta v$ domain (Fig. 4.1c) and point-spread function (Fig. 4.1a). 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 $\boldsymbol{\mathcal{F}}_{bb'}$, 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 to every source in the field.

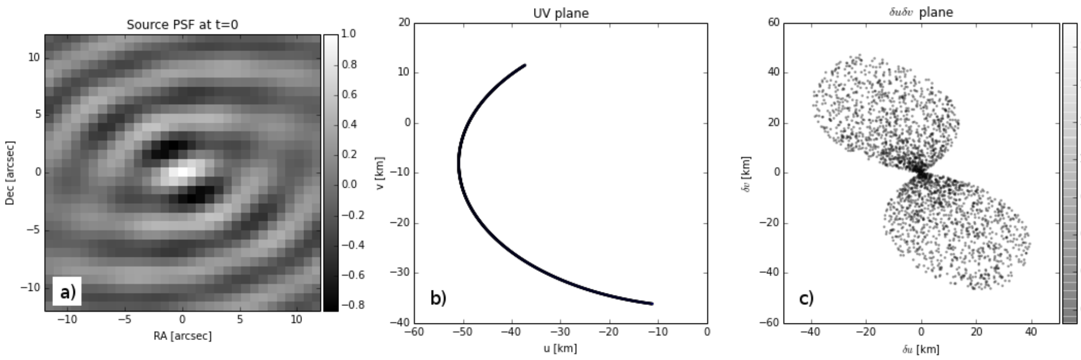


Figure 4.1. *Fig. 4.1a shows the PSF image of a simulated 1Jy source at phase centre. Colourbar units are in Jansky. Fig. 4.1b shows the associated UV track, and Fig. 4.1c the corresponding $(\delta u, \delta v)$ tracks. Note that the $\delta u \delta v$ plane does not have a homogeneous point density, 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, $[\mathrm{Cov}\{\tilde{\gamma}\}]_{b\neq b'}$ are all zero and $\mathrm{Cov}\{\tilde{\gamma}\}$ 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.

4.2.3 Noise Map Simulations

We have shown in Eq. 4.25 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 of correlated random numbers with zero mean and a distribution following a specified covariance matrix $\boldsymbol{C}$. 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

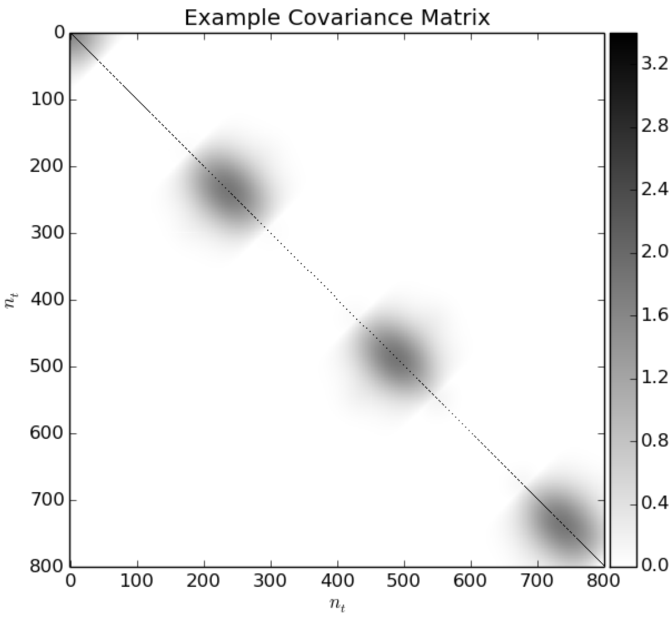


Figure 4.2. *Example of a non-stationary covariance matrix, which can be used to simulate Cov*$\{\tilde{\gamma}\}$*.The colourbar units are Jy*^{2}*. The correlation scale* σ_τ *is 40 cells, and the variability period is 500 cells. The matrix is made positive semi-definite (and therefore a covariance matrix) through SVD decomposition. The maximum size of the "bubbles" is determined by* σ_τ*.*

others. An example of this covariance matrix for arbitrary parameter values is shown in Fig. 4.2. 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: $\boldsymbol{C}$ is built following user specifications as described above. We then find its matrix square root $\boldsymbol{C}_0$ so as to apply it to random numbers generated from a normal distribution. We generate 2000 realisations i of our random variables $\boldsymbol{r}_i$:

$$\tilde{\boldsymbol{y}}_i = \sum_b \boldsymbol{\mathcal{F}}_b \boldsymbol{r}_i \quad \text{with} \quad \boldsymbol{r}_i = \boldsymbol{C}_0 \boldsymbol{x} \tag{4.32}$$

$$\text{and} \quad \boldsymbol{x} \leftarrow \mathcal{N}(0,1) \tag{4.33}$$

$$\tilde{\boldsymbol{y}} = \left(\cdots \tilde{\boldsymbol{y}}_i \cdots\right) \tag{4.34}$$

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

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 Bootes 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 $\boldsymbol{V}_y^m$, built by measuring the variance across realisations at each pixel in the image-plane, with the *predicted* variance map $\text{Var}\{\tilde{\boldsymbol{y}}\}$, built using the Cov-Cov relationship (Eq. 4.25). Since we are only interested in the variance map, rather than the covariance between pixels, we compute only the diagonal terms.

$$\boldsymbol{V}_{\tilde{y}}^{m} = \mathrm{Diag}\left\{\tilde{\boldsymbol{y}}\tilde{\boldsymbol{y}}^{H}\right\} \tag{4.35}$$
$$\boldsymbol{V}_{\tilde{y}}^{pr} = \sum_{b}[\boldsymbol{C}]_{bb}\mathbf{I} + \sum_{b,b'\neq b} \mathrm{Diag}\left\{\boldsymbol{\mathcal{F}}_{bb'}\right\}[\boldsymbol{C}]_{bb'} \tag{4.36}$$

where the thermal noise is already incorporated into the diagonal of $\boldsymbol{C}$ and Eq. 4.36 is merely the diagonal operator applied to the Cov-Cov relationship.

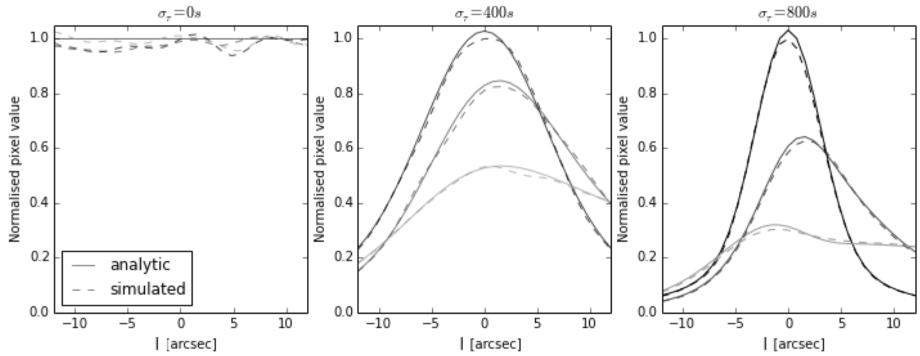


Figure 4.3. *The three lines in each figure correspond to three horizontal cross-sections from images in Fig. 4.4. The units on the y-axis are dimensionless [Jy^2/Jy^2]. σ_τ is the maximum characteristic error correlation length. In decreasing intensity, they correspond to $m = 0''$, $m = 4''$, and $m = 8''$. 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. 4.1. 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. 4.35, we can estimate $\mathrm{Var}\{\tilde{\boldsymbol{y}}\}$. The *predicted* noise-map, meanwhile, was found by assigning each cell of $\boldsymbol{C}$ to the appropriate point in the $(\delta u, \delta v)$ plane and Fourier transforming from this plane into the image-plane, as per Eq. 4.36. 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. 4.4: the predicted and simulated noise-PSFs match. The peak-normalised predicted noise-PSF is less noisy, as shown in Fig. 4.3 for different correlation scales. This is expected, since it is calculated directly from the underlying distribution, rather than an estimate thereof. As

$n_{\mathrm{realis}} \to \infty$, we expect the two methods to fully converge. As the maximum characteristic correlation length σ_τ increases, the variance becomes ever more sharply peaked.

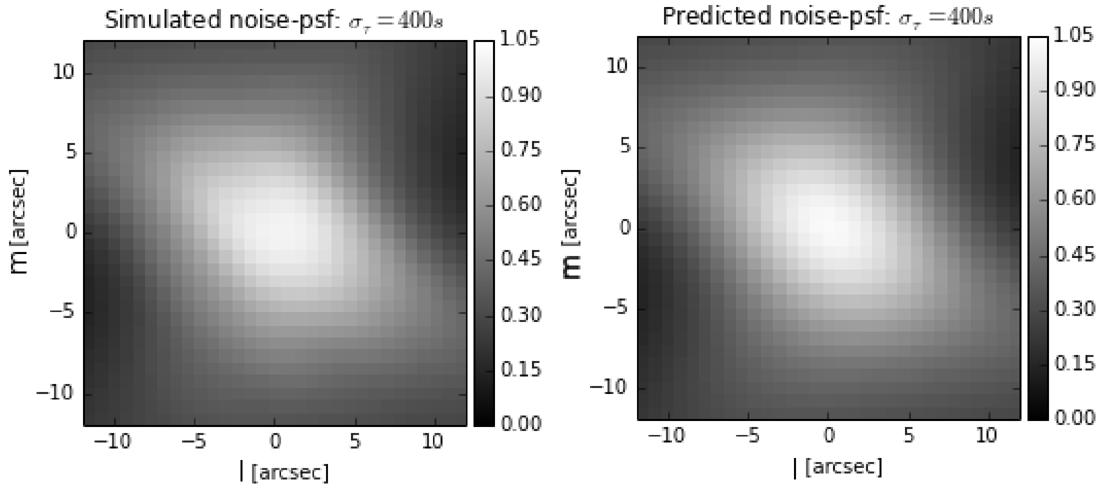


Figure 4.4. *Simulated noise-maps, compared with theoretical prediction. The pixel values are normalised by the average value of the covariance matrix: the units of the colourbar are thus dimensionless (Jy^2/Jy^2). These are on the same angular scale as the source shown in Fig. 4.1.*

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. 4.5d.

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. 4.5a. Secondly, we perform a DFT from the differential Fourier plane to the (l, m) plane as before, assigning one cell of $\mathrm{Cov}\{\tilde{\gamma}\}$ to each point of the differential Fourier plane. This time, however, $\phi^d_{bb'}$ 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. 4.5b. Finally, Fig. 4.5c 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 imags clearly overlap. We are thus not quite in the regime where emission is fully spatially incoherent. Nevertheless, our predictions hold to better than 5% accuracy.

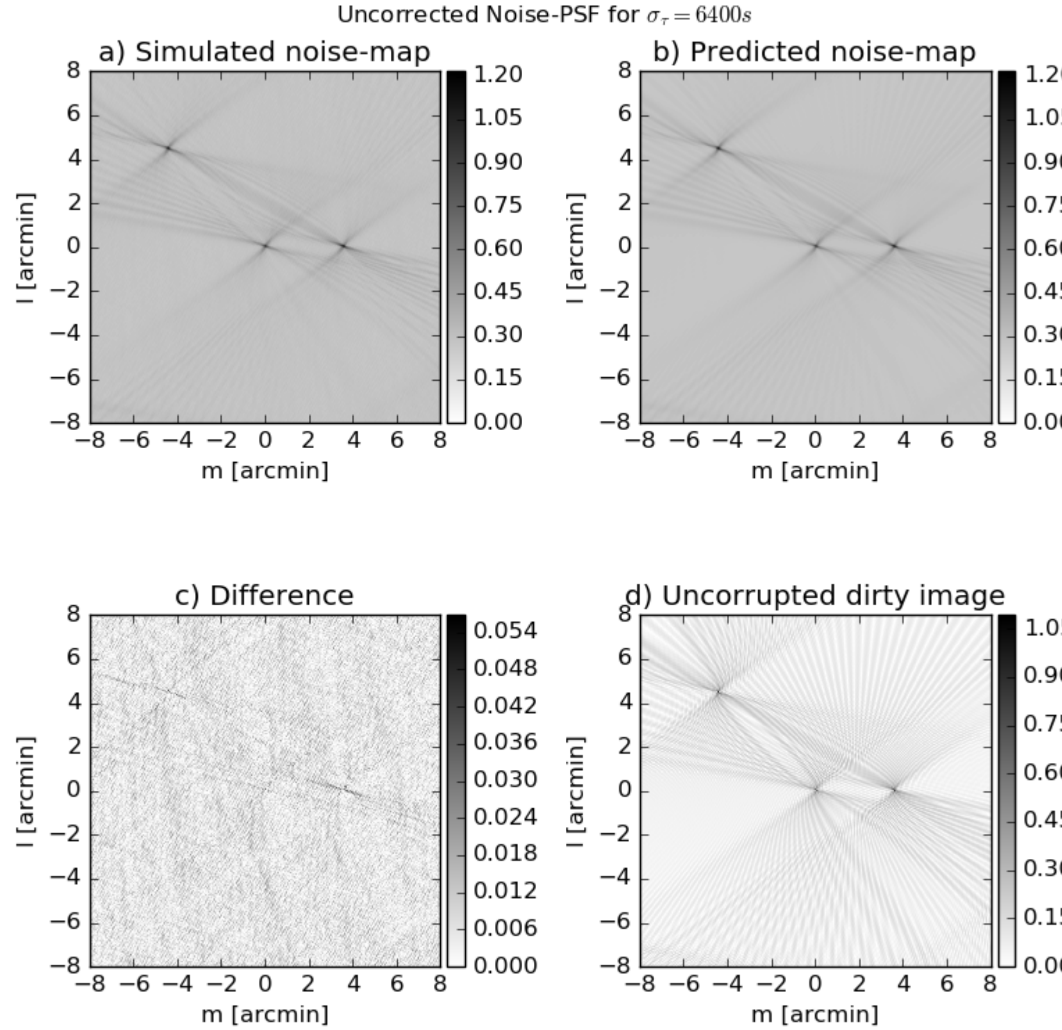


Figure 4.5. *Noise-map of sky with correlated gain errors and three point sources. The colourbars of (a), (b) and (c) have dimensionless units, while that of (d) is in Jansky. Note the presence of structure in the residuals (c): these show the limits of our hypothesis that sources are spatially incoherent.*

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.*

4.3 Adaptive Quality-based Weighting Schemes

As discussed in Section 4.1, 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. 4.27. 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. 4.27). 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.

4.3.1 Optimising sensitivity

The Cov-Cov relationship (Eq. 4.25) tells us that, far from any sources, the variance map (Eq. 4.27) 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 $\mathrm{Diag}\left\{\mathrm{Cov}\{\tilde{\gamma}\}\right\}$. This is equivalent to assigning visibilities weights inversely proportional to their variance:

$$w_b = \frac{1}{\mathrm{Var}\{\tilde{\gamma}_b\}} \tag{4.37}$$

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.

4.3.2 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. 4.25), 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:

$$V\left(\boldsymbol{w}\right) = \frac{\boldsymbol{w}^T \mathrm{Cov}\{\tilde{\gamma}\}\boldsymbol{w}}{\boldsymbol{w}^T \mathbf{1}\mathbf{1}^T \boldsymbol{w}} \tag{4.38}$$

Our optimality condition is then, after some algebra:

$$0 = \frac{\partial}{\partial \boldsymbol{w}}\left(V\right) \tag{4.39}$$

$$\leftrightarrow \mathrm{Cov}\{\tilde{\gamma}\}\boldsymbol{w} = \mathbf{1}\mathbf{1}^T\boldsymbol{w}\left(\boldsymbol{w}^T\mathbf{1}\mathbf{1}^T\boldsymbol{w}\right)^{-1}\boldsymbol{w}^T\mathrm{Cov}\{\tilde{\gamma}\}\boldsymbol{w} \tag{4.40}$$

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

$$\boldsymbol{w} = \mathrm{Cov}\{\tilde{\gamma}\}^{-1}\mathbf{1} \tag{4.41}$$

where $\mathbf{1}$ is a vector of ones. Provided that the hypotheses made so far hold[4], 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.

4.4 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.

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.

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

4.4.1 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 $\widehat{\text{Cov}\{\tilde{\boldsymbol{\gamma}}\}}$of the covariance matrix $\text{Cov}\{\tilde{\boldsymbol{\gamma}}\}$ is built by applying the covariance operator:

$$\widehat{\text{Cov}\{\tilde{\boldsymbol{\gamma}}\}} \overset{\text{def}}{=} \widehat{\boldsymbol{C}}_{\tilde{\gamma}} = \frac{1}{n_{\text{int}}} \sum_{i \in n_{\text{int}}} \left(\tilde{\gamma}_i - \langle\tilde{\gamma}\rangle\right)\left(\tilde{\gamma}_i - \langle\tilde{\gamma}\rangle\right)^H \tag{4.42}$$

where the $\langle\cdots\rangle$ operator denotes taking the average over the full vector. If the calibration solver's gain estimates are unbiased (i.e. $\text{E}\{\hat{\mathbf{g}}\} = \mathbf{g}$) *and* the model of the sky is sufficiently complete, this quantity should be zero. Having created $\widehat{\boldsymbol{C}}_{\tilde{\gamma}}$, 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 4.2. 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:

$$r_b = w_b \kappa_b \tilde{\gamma}_b + \epsilon \tag{4.43}$$

$$\rho_b = \frac{r_b}{k_b} \tag{4.44}$$

We then organise the residuals in cells:

$$\boldsymbol{r}_{\mathcal{C}} = \begin{pmatrix} \vdots \\ \rho_{b \in \mathcal{C}} \\ \vdots \end{pmatrix} \tag{4.45}$$

$$\boldsymbol{R} = \begin{pmatrix} \cdots & \boldsymbol{r}_{\mathcal{C}} & \cdots \end{pmatrix} \tag{4.46}$$

$\boldsymbol{R}$ 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 $\boldsymbol{R}$ with its Hermitian conjugate:

$$\widehat{\boldsymbol{C}}_{\tilde{\gamma}}[b \in \mathcal{C}, b' \in \mathcal{C}'] = \left(\boldsymbol{R}^H \boldsymbol{R}\right)[\mathcal{C}, \mathcal{C}'] \tag{4.47}$$

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. $\widehat{\boldsymbol{C}}_{\tilde{\gamma}}$ is then an estimate of the residual visibility covariance matrix.

4.4.2 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:

$$\hat{\mathbf{g}}_c = \begin{pmatrix} \vdots \\ \hat{g}_p^{\tau \in c} \\ \vdots \end{pmatrix} \tag{4.48}$$

$$\hat{\mathbf{G}} = \begin{pmatrix} \cdots & \hat{\mathbf{g}}_c & \cdots \end{pmatrix} \tag{4.49}$$

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

$$= \mathrm{E}\{\hat{\mathbf{g}}_c \hat{\mathbf{g}}_c^H\} - \underbrace{\mathrm{E}\{\hat{\mathbf{g}}_c\}\mathrm{E}\{\hat{\mathbf{g}}_c\}^H}_{=\mathbf{g}_c \mathbf{g}_c^H} \tag{4.50}$$

As we can see, Eq. 4.50 is simply a vector form of Eq. 4.12. The residual gains of Eq. 4.12 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*:

$$\widehat{\boldsymbol{V}_{\mathcal{C}}} = \widehat{\mathrm{Var}\{\mathbf{g}_c\}} \tag{4.51}$$

$$= \sum_{\tau \in \mathcal{C}} \left(\hat{\mathbf{g}}_\tau \hat{\mathbf{g}}_\tau^H - \mathbf{g}_\tau \mathbf{g}_\tau^H \right) \tag{4.52}$$

We define the residual matrix as:

$$\boldsymbol{r}_\tau = \sum_d s_d \boldsymbol{K}_{d,\tau} \left(\hat{\mathbf{g}}_\tau \hat{\mathbf{g}}_\tau^H - \mathbf{g}_\tau \mathbf{g}_\tau^H \right) \boldsymbol{K}_{d,\tau}^H + \epsilon \tag{4.53}$$

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. 4.53 lies $\widehat{\boldsymbol{V}}_{\hat{\mathbf{g}}_\tau}$, where $\sum_{\tau \in \mathcal{C}} \widehat{\boldsymbol{V}}_{\hat{\mathbf{g}}_\tau} = \widehat{\boldsymbol{V}_{\mathcal{C}}}$. The K-matrix is defined as follows:

$$\boldsymbol{K}_{d,\tau} = \begin{pmatrix} k_{p,\tau}^d & & 0 \\ & k_{q,\tau}^d & \\ 0 & & \ddots \end{pmatrix} \tag{4.54}$$

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

$$\boldsymbol{r}_{\mathcal{C}} = \begin{pmatrix} \vdots \\ \boldsymbol{r}_{\tau \in \mathcal{C}} \\ \vdots \end{pmatrix} \tag{4.55}$$

$$\boldsymbol{R} = \begin{pmatrix} \cdots & \boldsymbol{r}_c & \cdots \end{pmatrix} \tag{4.56}$$

$\boldsymbol{r}_{\mathcal{C}}$ 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:

$$\mathbb{V} = \boldsymbol{R}^H \boldsymbol{R} \tag{4.57}$$

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:

$$\mathbb{V} = \sum_{\tau \in \mathcal{C}} \left(\sum_{d,d'} s_d \boldsymbol{K}_{d,\tau} \left(\widehat{\boldsymbol{V}_{\mathcal{C}}}\right)^H \boldsymbol{K}_{d,\tau}^H s_{d'} \boldsymbol{K}_{d',\tau} \left(\widehat{\boldsymbol{V}_{\mathcal{C}}}\right) \boldsymbol{K}_{d',\tau}^H + \mathbf{I}\sigma^2 \right) \tag{4.58}$$

To reach this point, in Eq. 4.23, we made the hypothesis that the sky brightness distribution is dominated by spatially incoherent emission. 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:

$$\mathbb{V} \approx \sum_{\tau} \left(\sum_{d} s_d^2 \boldsymbol{K}_{d,\tau} \left(\widehat{\boldsymbol{V}_{\mathcal{C}}}\right)^H \left(\widehat{\boldsymbol{V}_{\mathcal{C}}}\right) \boldsymbol{K}_{d,\tau}^H + \mathbf{I}\sigma^2 \right) \tag{4.59}$$

$$= \left(\widehat{\boldsymbol{V}_{\mathcal{C}}}\right)^2 \circ \left(\underbrace{\sum_{\tau} \sum_{d} s_d^2 \boldsymbol{k}_{d,\tau} \boldsymbol{k}_{d,\tau}^H}_{\overset{\text{def}}{=} \mathcal{S}} \right) + n_{\mathcal{C}} \mathbf{I}\sigma^2 \tag{4.60}$$

$$\widehat{\boldsymbol{V}_{\mathcal{C}}} = \sqrt{\mathcal{S}^{\circ -1} \left(\mathbb{V} - n_{\mathcal{C}} \mathbf{I}\sigma^2\right)} \tag{4.61}$$

where $\circ$ denotes the Hadamard or entrywise product and $\boldsymbol{k} = \mathrm{Diag}\left\{\boldsymbol{K}\right\}$. 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}.

4.5 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. 4.6. The two remain in such agreement throughout the cross-section as to be nearly indistinguishable.

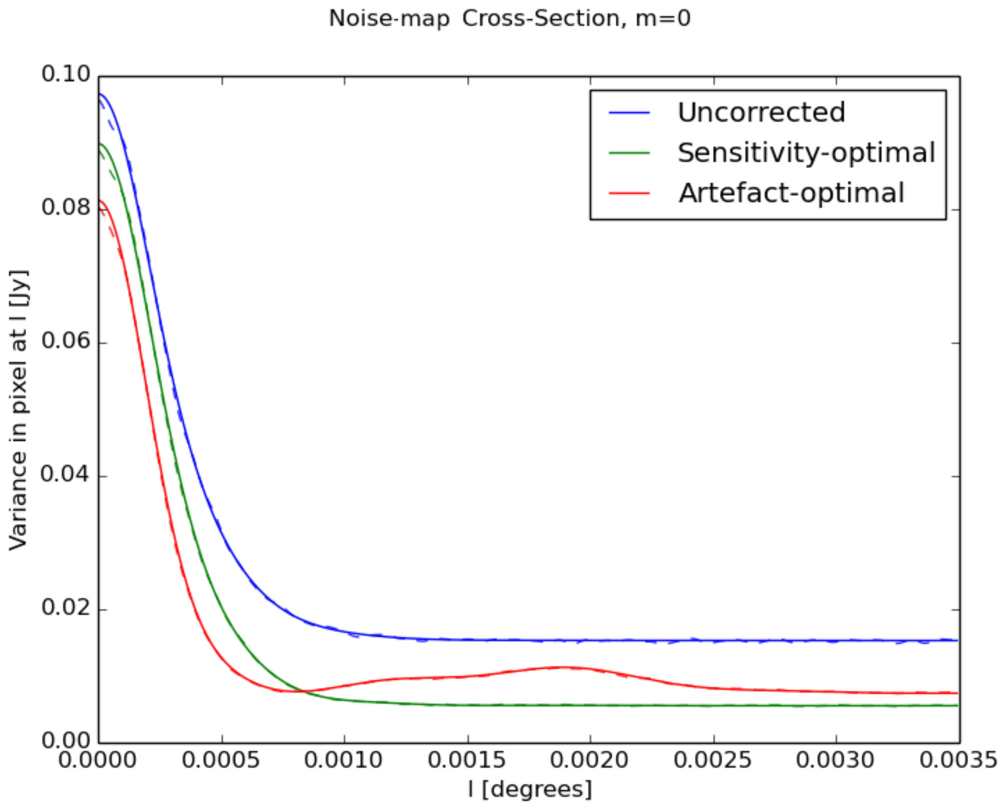


Figure 4.6. *The sensitivity-optimal (green) and artefact-optimal (red) weights both give improvements over the unweighted noise-map (blue).*

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.

4.6 Applying the Correction to Real Data

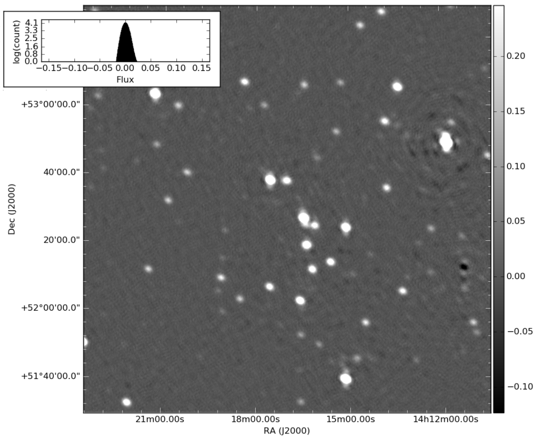
In this section, we show the effect of adaptive quality-based weighting on real data. The dataset used in this section is a single sub-band from an 8-hour LOFAR observation centred on the Extended Groth Strip (α=14:19:17.84,δ=52:49:26.49). The observation was performed on August 28th, 2014. The subband 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. 4.7a) 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., in press).

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. 4.7b. 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.

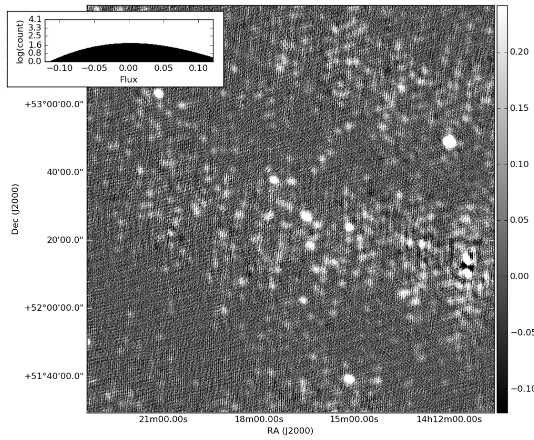
Weights were then calculated based on the badly-calibrated residual visibilities. Fig. 4.7c was made using the same visibilities as Fig. 4.7b and applying baseline-based, sensitivity-optimal weight. Similarly, Fig. 4.7d used the poorly-calibrated residual visibilities with the application of baseline-based, 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.

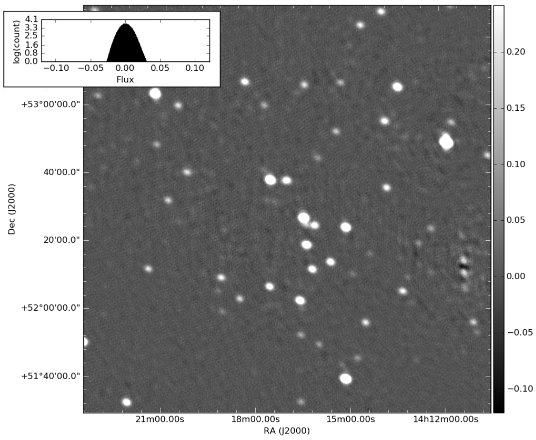
The pixel histograms show us that the weights do not completely mitigate the poor calibration interval choice, but certainly give a dramatic improvement over the unweighted, poorly-calibrated residuals. This is compatible with our statement that the weights give similar residuals in the image with a dramatic improvement in time at some cost in sensitivity. It is interesting to note that while Fig. 4.7d looks noisier than Fig 4.7c, its residual flux histogram is actually closer to that of Fig. 4.7a.



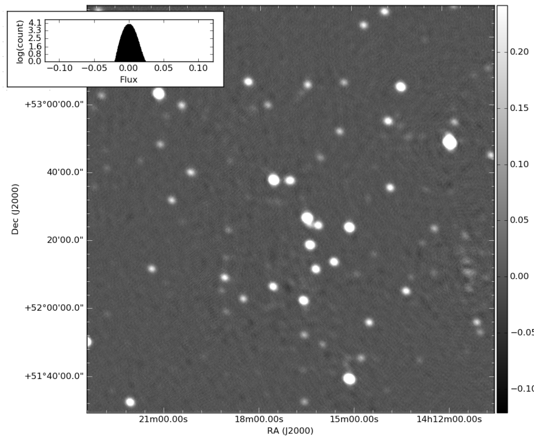
(a) *Well-calibrated, unweighted restored image of the sky near the centre of the Extended Groth Strip. Used for comparison with the other images. Units of colourbar are Janskys. This image was made with data calibrated following best practice (solution intervals of 8 seconds, half the bandwidth). rms=5.87mJy/beam*



(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. 4.7a, but averaged in time and calibrated using larger gain solution intervals: 2 minutes and half the bandwidth. rms=86.4mJy/beam*



(c) *Image made using the same imaging parameters and corrected visibilities as Fig. 4.7b, using sensitivity-optimal weighting. Units of colourbar are Janskys. rms=9.69mJy/beam*

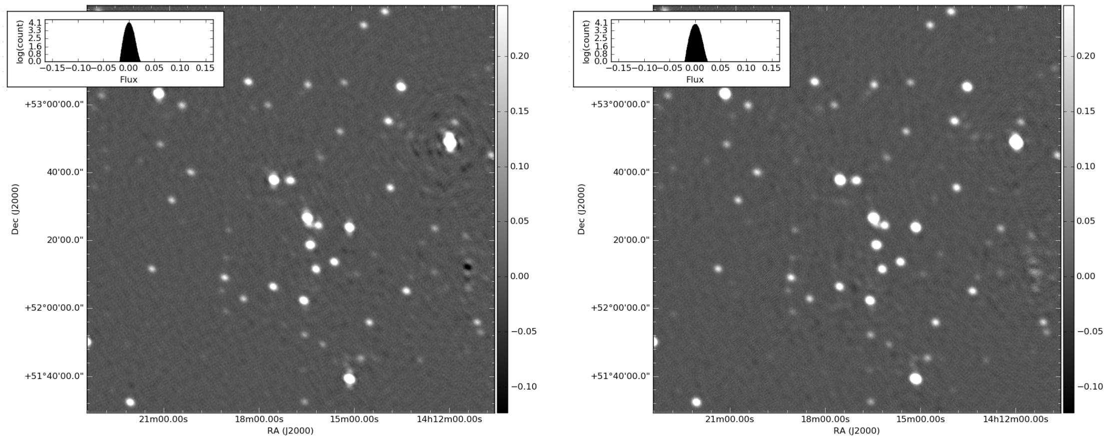


(d) *Image made using the same imaging parameters and corrected visibilities as Fig. 4.7b, using artefact-optimal weighting. Units of colourbar are Janskys. rms=15.8mJy/beam*

Figure 4.7. *Restored images of the centre of the Extended Groth Strip, as seen with an 8-hour observation using the full LOFAR array. Fig. 4.7a shows an image of the field made with good calibration intervals. Fig. 4.7b shows an image of the field made with poor calibration intervals. Fig. 4.7c shows image made with the same visibilities and imaging parameters, but with the application of the sensitivity-optimal weighting scheme. Fig. 4.7d, similarly, differs from Fig 4.7c 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.*

As for performance, the weights used for Fig. 4.7d took 8 hours of computing time on a single core[5], working on a 29 GB dataset, which is not particularly large for LOFAR

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



(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=5.87mJy/beam*

(b) *Image made using the same imaging parameters and corrected visibilities as Fig. 4.7b, with the application of antenna-based, sensitivity-optimal weighting. Solution interval of 2 minutes, half the bandwidth. rms=6.69mJy/beam*

Figure 4.8. *Comparison between the well-calibrated image (i.e. the same image as Fig 4.7a) 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.*

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. 4.8b, they are computed in 1min 6s on the same single core.

4.7 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 $\mathrm{Cov}\{\tilde{\boldsymbol{y}}\}$) 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 $\mathrm{Cov}\{\tilde{\boldsymbol{y}}\}$. 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 4.7b to Fig. 4.7c. 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. 4.7b to 4.7d 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.

Chapter 5

Scientific Overview: Goals & Strategy

Contents

5.1 Science Goal

So far, we have shown the technical work done on interferometric techniques as part of this PhD. In this section, we outline the application of this work, along with other modern tools, to the creation of large, deep and high-resolution surveys of the 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.

Why, then, use this approach? One reason is the challenge itself. A high-resolution (matching HST resolution), wide-area (multiple degrees wide), high-quality[1] radio survey has *never* been achieved at the time of writing. Many surveys of the radio sky exist, of course, but they tend to combine only two of those three qualities. FIRST (see Helfand et al. 2015, and references therein), for example, has achieved tremendous results with a $4''$ resolution, relatively low sensitivity, and no short baselines (thereby missing diffuse emission and extended structure) at ~1.4 GHz. The VLASS (see Hales 2013, and references therein) will aim to go to a $2.5''$ resolution, medium sensitivity, and few short baselines - at frequencies ranging from 2 to 4 GHz. There exist many LOFAR HBA (~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 ($\sim 5''$).

[1] Reliably artefact-free, with known & limited decorrelation, etc.

In this context, a sky survey made using international LOFAR would be competitive until the SKA-MID survey (and considering it maps the Northern sky rather than the Southern sky, would arguably remain competitive even then). It would also allow LOFAR to fulfill the role of pathfinder in a very literal way, by providing datasets on which algorithms and imaging strategies could be tested before handling SKA data volumes in real-time.

Making a high-resolution full-sky survey using international LOFAR is thus arguably useful in and of itself in terms of interferometric techniques & instrumentation. What then of its science value? The two most obvious advantages of a good large-sky survey are that they are complete (i.e. they are an accurate sample of the underlying source distribution, up to the survey's limiting flux) while still giving information on rare outlier objects which may happen to lie in the field. This gives us access to a wealth of statistics from which to derive more robust estimations of population distributions.

Of course, LOFAR is already providing such catalogues - LOTSS (see Shimwell et al. 2017, and references therein) is only the first of many to come. But these catalogues do not make use of the LOFAR international stations, and thus do not make full use of LOFAR's resolution ($5''$ resolution vs. $0.5''$). This is so due to the technical challenges introduced by the use of international stations, which are due to decorrelation and an already limited signal-to-noise on those very short spatial frequencies.

What would they add to a LOFAR catalogue? A high-resolution sky survey has all the advantages listed above with more besides. For AGN science, for example, the higher resolution gives information on smaller scales, which is very relevant to study the turbulent processes in the lobes and the nearby intergalactic medium. We also know that there exist small AGNs: size therefore matters, and resolving these small AGNs (which tend to lie at larger redshifts, hence their smaller angular size) can give extremely salient information on AGN population distributions as a function of the age of the universe. Indeed, for AGN science, the LOFAR international baselines can be the difference between having absolutely no information on size distribution (no resolved AGN) and knowing everything about the AGN size distribution within a given extragalactic field. Finally, resolving sources is very relevant to spectral studies of AGN populations: we know very little about the physics of AGNs at low frequencies. Information on large populations of resolved sources (allowing us to separate compact structure from extended emission) is a prerequisite to begin the statistical work needed to understand these low-frequency AGN physics. As such, the international baselines can bridge a critical gap in the data available to peers studying AGN science.

AGN science is far from the only field which could benefit from the sub-arcsecond resolution that international LOFAR could provide. High-resolution imaging can be of tremendous interest to the study of star-forming galaxies: it could give insight into cosmic ray structure and access to spectral information. This is relevant because the entire star-forming history of galaxies is encoded in radio emission, but can only be accessed with sufficient spectral information and spatial resolution. A high-resolution survey, in particular, allows comparisons between optical and radio emission for known star-forming galaxies on a truly industrial scale, which would allow for the mapping of free-free absorption, supernova remnants, HII regions, etc...onto optical maps, and this for sources going up to high redshifts.

Of course, this does not run the full range of science cases which would benefit from high-resolution maps of large parts of the radio sky. Colleagues studying gravitational lensing would benefit greatly from large samples of lensed galaxies. Resolved extragalactic recombination lines (for bright objects) would doubtlessly interest peers studying galactic evolution. We expect these samples to be a natural byproduct of a large, high-resolution map of the radio sky: if they do not appear, that could indeed be a very interesting scientific result in and of itself.

More generally, high-resolution surveys provide high-resolution images of everyone's favourite objects and sources. It allows for better optical matching and identification, which is particularly relevant e.g. when needing to associate either a low-z or high-z sources to optical counterpart sources (in the case of the Extended Groth Strip, for example, using the international stations means matching the Hubble Space Telescope's resolution). They are thus a strong contribution to multi-wavelength datasets. It also improves the image's sensitivity to compact sources embedded in diffuse emission: if the source is better-resolved, then the emission associated with compact sources is more easily separated from emission emitted by its surroundings. This can be of tremendous help in image interpretation. Finally, actually resolving objects allows us to have better morphological classification for compact objects, which can be extremely useful e.g. when interested in AGN populations vs star-forming galaxy populations.

5.2 3C295 and the Extended Groth Strip

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 5.1. 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 $\alpha = 14^h 17^m, \delta = +52\,\mathrm{deg}\,30'$, 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 5.1. *Table recapitulating the multi-wavelength coverage of the Extended Groth Strip, with observations performed as part of the AEGIS (later, CANDELS) collaborations.*

3C295 is a 3C[2] source, and thus 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.

The first step towards acquiring a deep, high-resolution image of the Extended Groth Strip therefore consists of creating a good, high-resolution model of 3C295 at LOFAR frequencies. 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 6.

5.3 Imaging the Full Primary Beam

After creating a high-resolution model of 3C295 to calibrate the international stations, we must create a low-resolution model of all the sources in the primary beam to both improve final calibration quality and subtract sources outside of the Extended Groth Strip when creating the final high-resolution images.

Initial direction-independent calibration is done using the same pipeline as for the Lofar Two-metre Sky Survey (LOTSS, Shimwell et al. 2017). The killMS-DDF facet-based pipeline (cf Smirnov & Tasse 2015; van Weeren et al. 2016; Tasse et al. 2017) is then used to perform direction-dependent calibration for the entire LOFAR primary beam. This allows us to reach a noise threshold of ~233μJy.

This work, and its results, are shown in Chapter 7. Note that the AEGIS observations summarised in Table 5.1 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 producing 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 Groth Strip, and are therefore only imaged at low resolution, but follow-up images and observations could be made should the need arise.

[2]Third Cambridge Catalogue of Radio Sources, a 1959 survey of of the Northern radio sky

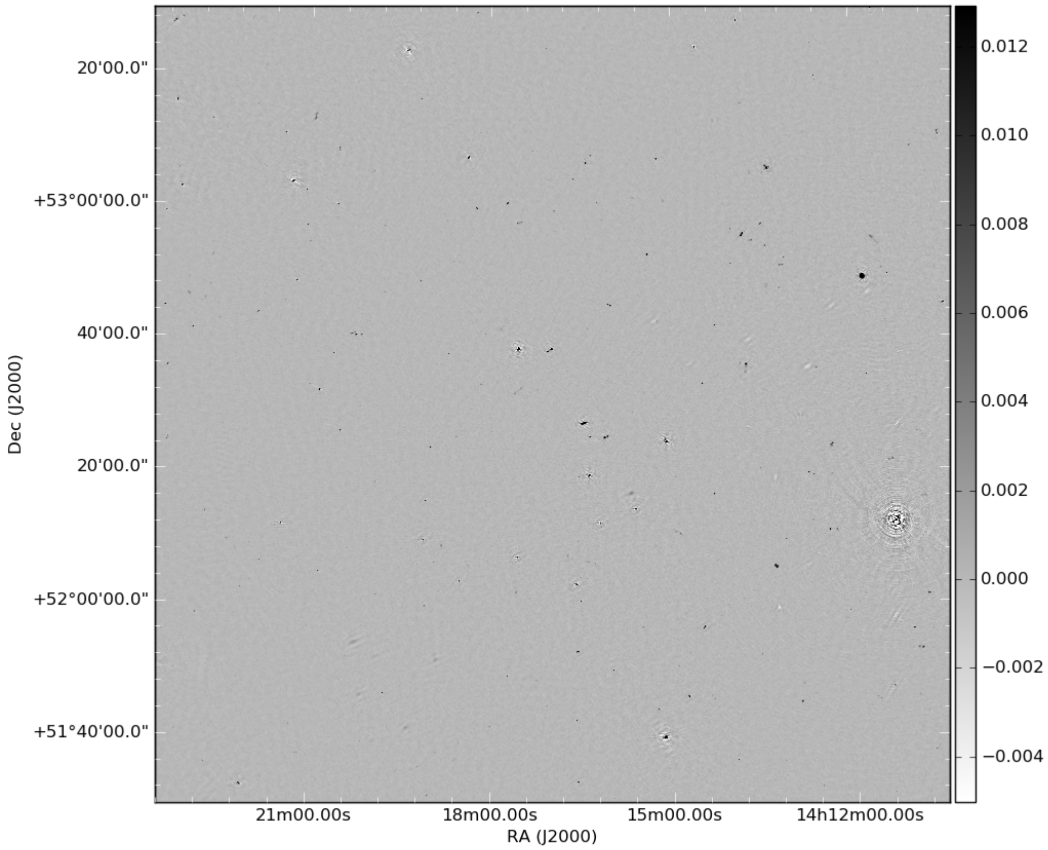


Figure 5.1. *Full direction-dependent calibrated image of the LOFAR primary beam, centred on the EGS. Image made without international stations.*

5.4 Test Decorrelation

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 imager used, whereas the second is a function of the gain variability as a function of distance from the calibrator. Seven calibrator sources, lying at various distances and directions from the calibrator source, are examined. They are summarised in Table 8.1, which is reproduced below.

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 5.2. *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.*

5.5 Imaging the EGS with LOFAR international stations

5.5.1 Science Goals

if decorrelation is merciful, proceed to patchwise imaging of EGS by using the results from the sections above (DI calibration using 3c295 model, followed by subtraction of all sources seen at low-res except within the patch we want to image; image, change patch; repeat until all EGS imaged)

Chapter 6

Imaging 3C295 with International LOFAR

Contents

6.1 Aims & Methodology

Our aim in this section is to create a high-resolution model of 3C295, something that - as of yet - does not exist. More specifically, we seek to create a model that will allow us to find good phase-calibration solutions for LOFAR international baselines. While we also solve for amplitude gains, we know that we will need to correct the total source flux in each frequency subband as our initial spectral model is not necessarily correct. We select 6 sub-bands out of the total LOFAR HBA bandwidth, evenly spread throughout the bandwidth as shown in Fig. 6.1.

This approach should allow us to appropriately constrain the final spectral behaviour of our high-resolution model, once the amplitude correction is applied to ensure our model is compliant with pre-existing flux measurements for this source (Scaife & Heald 2012). [put image of source flux as function of frequency for 3c295, citing source and ideally showing position of our subbands]

Our procedure is as follows: we self-calibrate individual subbands, starting from a model extracted from a high-resolution VLA observation of 3C295 at 8.7 GHz (cf. Perley & Taylor 1991). Once this is done, we extract and apply a scalar flux correction factor for each subband so that the integrated flux of our 3C295 image is compatible with the existing literature at all frequencies. The resulting model is then reliable enough to calibrate our Groth Strip data using the full LOFAR array and the entire HBA bandwidth.

6.2 Data Reduction

6.2.1 Data & Observation Properties

The dataset used for this PhD project is part of the LOFAR Surveys KSP Tier-1 survey, which consists of a number of 8-hour pointings covering as much of the sky visibile to LOFAR as possible. We analyse one of these pointings, an observation performed on the 28th of August 2014 and centred on the Extended Groth Strip. As part of the Tier-1 survey, it is an 8-hour-long observation. We limit ourselves to analysing only the HBA observation, meaning that we use 365 sub-bands which sample a total bandwidth of [give bandwidth]. We use all Core and Remote LOFAR stations, as well as some of the International stations which were online at the time (specifically the German stations 1-5 and 7, along with the Swedish and the British station).

The data was acquired through the LOFAR Long-Term Archive, and is thus pre-processed and flagged for RFI using the standard tools ([cite aoflagger, ndppp?]).

6.2.2 Calibrating the Data

Since we plan on applying an amplitude correction based on the works of Scaife & Heald (see Scaife & Heald 2012), our overriding concern is to find good phase and amplitude gain without concern for the overall scaling factor. As such, our calibration strategy consists of calibrating 6 subbands, chosen across the LOFAR bandwidth, simultaneously. The chosen subbands, and their position in the total bandwidth, are shown here:

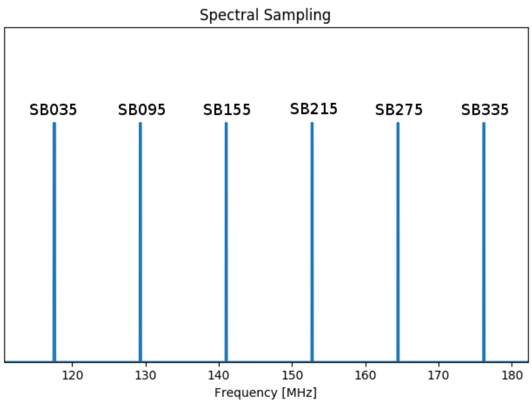


Figure 6.1. *Position of subbands chosen across full LOFAR bandwidth*

Subband	$\nu_{\rm min} - \nu_{\rm max}$
SB035	117.45-117.69
SB095	129.21-129.41
SB155	140.93-141.12
SB215	152.65-152.84
SB275	164.37-164.56
SB335	176.08-176.28

Table 6.1. *Frequency bounds for the subbands chosen*

Because no high-resolution models exist for 3C295 at our observing frequencies, which span quite a large bandwidth, care must be taken not to bias our model in an unphysical direction. We acquire the initial calibration model by extracting features from a NASA/IPAC Extragalactic Database[1] image of 3C295 (see Perley & Taylor 1991), shown in Fig. 6.2. This feature extraction is carried out using PyBDSM (Mohan & Rafferty 2015), changing all extracted Gaussians into points[2].

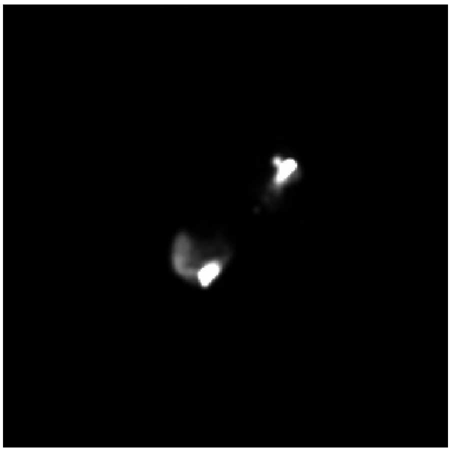


Figure 6.2. *VLA observation of 3C295 at 8.7 GHz. Pixel size is* $0.2''$.

For SB035, the gain curves for all antenna when calibrating with the initial model are as shown in Fig 6.3. Note that, especially for the international baselines, the gain amplitudes show a lot of structure: this is unphysical, and represents the unmodelled flux being absorbed into the gain solutions.

A few features are immediately visible in Fig. 6.3. Firstly, we see in the gain amplitude curves of the core and remote stations that the gains for the last 2 hours of observation are noisier than the rest. This behaviour can be seen for most core and remote stations. Thankfully, the quality-based weighting schemes ought to optimise the contribution from these noisy visibilities, and so this is not too concerning. We see that the amplitude curves for the core and remote stations are nice and flat, with little structure to be seen: this is very encouraging, as this is what we expect "physical" gain curves to look like. In other words, we see little pollution from unmodelled sources in the calibration model for core and remote stations. The phase seems reasonable, with some structure but not dominated by noise. Note that the phase is expected to wrap from $-\pi$ to π and vice-versa, which is what the fast oscillations in the last half of the phase gains of RS407HBA correspond to.

As for the international station gains, we see the presence of structure in the amplitude gain curve. This is indicative of the presence of sky model errors, which is expected: this is, after all, the gain curves after the first pass of calibration (i.e. without self-calibration). Very little information can be extracted from the phase curves, which wrap very fast: this is expected behaviour, as longer baselines rotate faster than shorter ones.

[1] https://ned.ipac.caltech.edu/

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

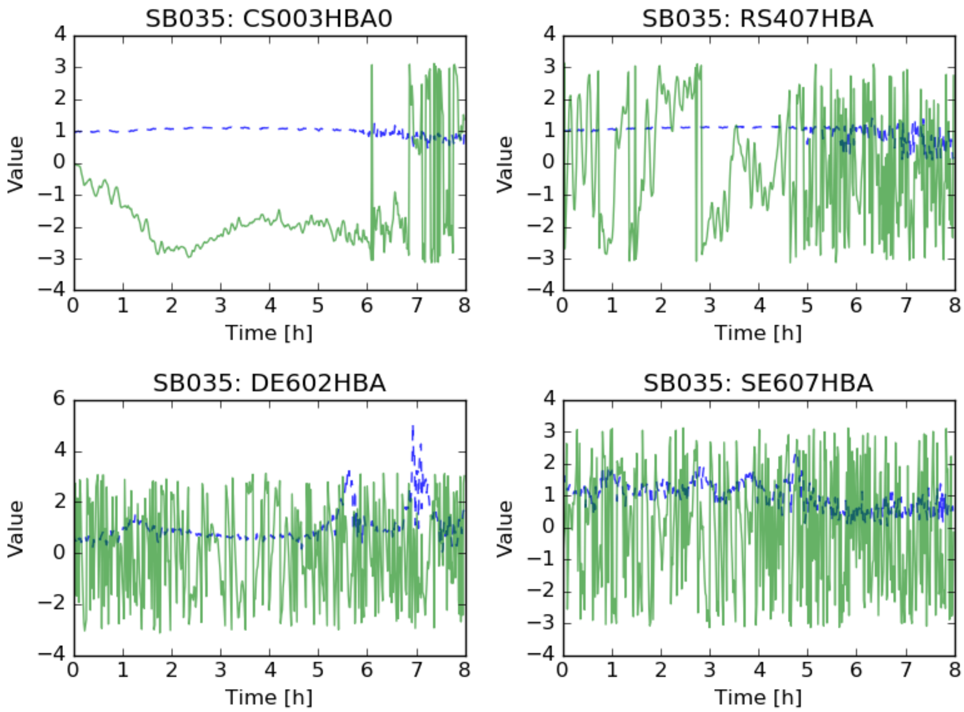


Figure 6.3. *Gain curves for four LOFAR stations in SB035. Calibration was done using the VLA model. The blue curve shows the gain amplitude, while the green curve shows the phase. Here, only the values for the XX correlation are shown; they are indicative of the other Jones term properties.*

As calibration improves, however, we expect to get more of the true sky in our model. This, in turn, ought to decrease the structure in the gain curves. Fig 6.4 shows the gain curves for the same antennas after 3 passes of self-calibration.

As we can see, the last two hours remain noisy compared to the first two. What is interesting here is the evolution of the amplitude gain curves for DE602HBA and SE607HBA; while there is clearly still structure in the amplitude curves, we see that some of the structure is no longer present in the curves. What this indicates is that the sky model has improved compared to what it had been, but there is still room for improvement: indeed, this is as expected, as there are many more sources than 3C295 within the primary beam.

This self-calibration was performed over all 6 subbands simultaneously; this allowed us to constrain the flux distribution in the sky as a function of frequency, even though the absolute scale in each image is known to be wrong. The absolute scale distribution being wrong is no issue, however; indeed, for each given subband, the gains are all wrong by the same multiplicative factor. This factor changes between subbands, however, and so the corrected visibilities must be corrected by the appropriate multiplicative factor for the true position-dependent spectral indices to be found. Mathematically, we have:

$$\mathbf{V}_{pq}^{\text{corr.}}(t,\nu) = (a_{\text{false}}(\nu)\tilde{\mathbf{G}}_{p,t}^{\nu})^{-1}\mathbf{V}_{pq}^{\text{meas.}}(t,\nu)(a_{\text{false}}(\nu)\tilde{\mathbf{G}}^{\nu})^{-1} \tag{6.1}$$

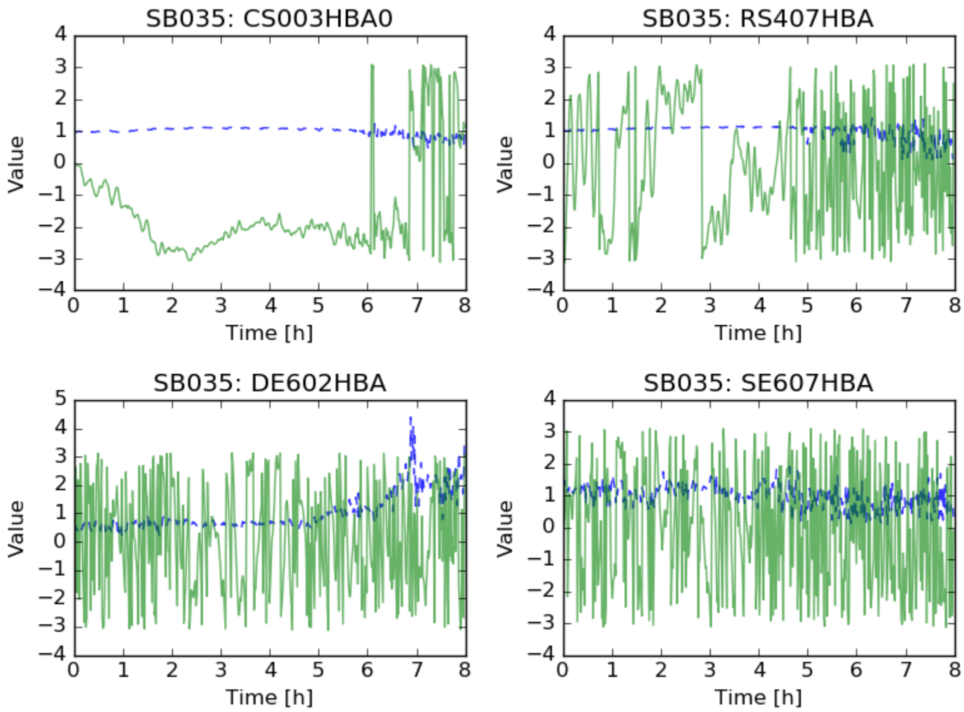


Figure 6.4. *Gain curves for four LOFAR stations in SB035 after 3 passes of self-calibration. The blue curve shows the gain amplitude, while the green curve shows the phase. Here, only the values for the XX correlation are shown; they are indicative of the other Jones term properties.*

$$\mathbf{V}^{\text{meas.}}_{pq}(t,\nu) = (a_{\text{true}}(\nu)\mathbf{G}^{\nu}_{p,t})\mathbf{X}_{pq}(t,\nu)(a_{\text{true}}(\nu)\mathbf{G}^{\nu}_{p,t}) \tag{6.2}$$

where uppercase, boldface letters denote 2×2 Jones matrices, and lowercase letters are scalars. Here, $a_{\text{false}}(\nu)$ is the incorrect scaling factor applied at frequency ν by calibrating without the proper spectral indices for all the components of 3C295, while $a_{\text{true}}(\nu)$ is the correct scaling factor. The true scaling factor is known from low-resolution calibrator analysis, while the false scaling factor is an arbitrary function of frequency. Assuming that our estimate for the Jones matrices is accurate (modulo the scaling factor), i.e. that $\tilde{\mathbf{G}}^{\nu}_{q,t} \approx \mathbf{G}^{\nu}_{q,t}$, then

$$(a_{\text{false}}(\nu)\tilde{\mathbf{G}}^{\nu}_{p,t})^{-1}(a_{\text{true}}(\nu)\mathbf{G}^{\nu}_{p,t}) = \frac{a_{\text{false}}(\nu)}{a_{\text{true}}(\nu)}(\tilde{\mathbf{G}}^{\nu}_{p,t})^{-1}\mathbf{G}^{\nu}_{p,t} \tag{6.3}$$

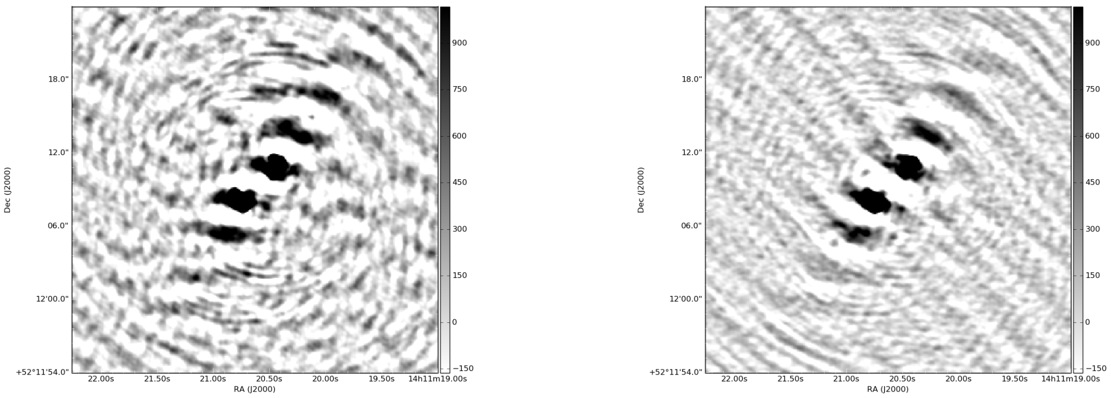
$$\approx \frac{a_{\text{false}}(\nu)}{a_{\text{true}}(\nu)}\mathbf{I} \tag{6.4}$$

and so

$$\mathbf{V}^{\text{corr.}}_{pq}(t,\nu) \approx \left(\frac{a_{\text{false}}(\nu)}{a_{\text{true}}(\nu)}\right)^2 \mathbf{X}_{pq}(t,\nu) \tag{6.5}$$

we can thus recover the properly-scaled visibilities in each subband by normalising the visibilities (i.e. divide by whatever a_{false} happens to be with a given calibration at a given frequency) and subsequently scaling them by a_{true} so that the average value in the corrected visibilities is equal to what's expected at this frequency.

Having calibrated our 6 subbands with the VLA model, we deconvolve them simultaneously. This allows us to improve the conditioning of the imaging inverse problem. After three passes of self-calibration, we have a spectral cube with 6 frequency slices centred on the various subbands. The stacked residual images are in Fig 6.5, before and after self-calibration. We see that noise falls by a factor of 1.17. Note that this measure is not as significant as might be expected: indeed, since the scaling factors for each subband is arbitrary, the overall decrease in noise could be attributed to a single subband, should its incorrect scaling factor dominate. The overall decrease is nevertheless encouraging. The frequency-dependent model is shown in Fig. 6.6. Note that while



(a) *Calibrated with the VLA model: there is quite a bit of noise.*

(b) *Same data after 3 passes of self-calibration: the noise is decreased.*

Figure 6.5. *Images made with 6 subbands spread across the LOFAR bandwidth.*

there is some frequency-dependent flux distribution changes, it is very small.

!!! SHOW GAIN CURVES, LOTS OF PLOTS ETC

!!! ANALYSE CONVERGENCE OF CALIBRATION ACROSS SUBBANDS

apply beam to account for different sensitivities of different baselines

use full-Jones solver (i.e. solve for a 2x2 complex matrix) to account for effects such as Faraday rotation etc

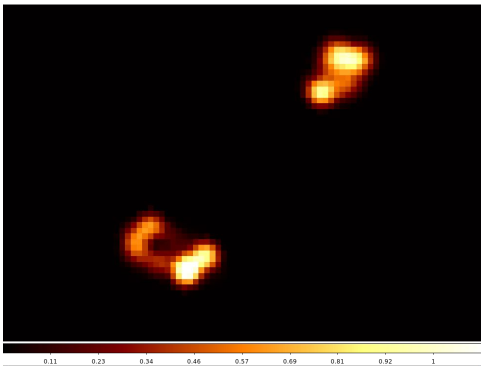
!!! SHOW JONES CHAIN WE SOLVE FOR AND EXPLAIN IT, MATHS ETC: ***BJ*** for Beam*Jones

!!! SHOW DIRECTION-DEPENDENT PSF EFFECT WHICH WE MODEL

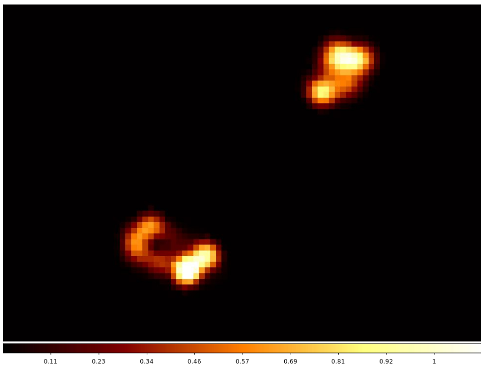
6.3 Overlays & Spectral Analaysis

6.3.1 Overlays of 3C295 at Multiple Frequencies

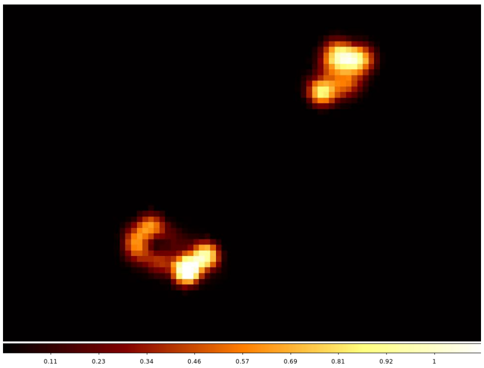
Show overlays of our low-freq model onto images of 3c295 at various other freqs - VLA, optical, IR, X-ray, etc.



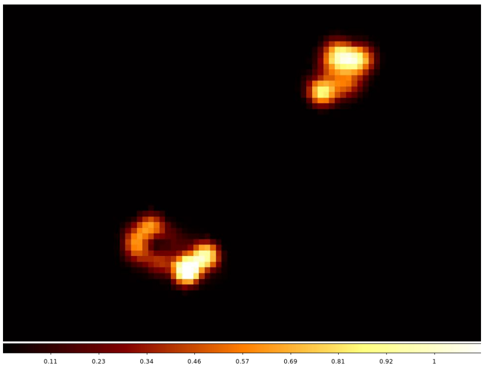
(a) *SB035*



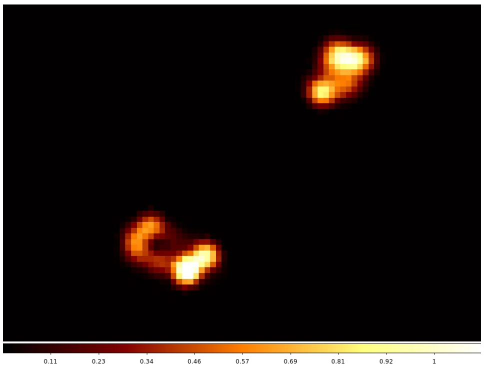
(b) *SB095*



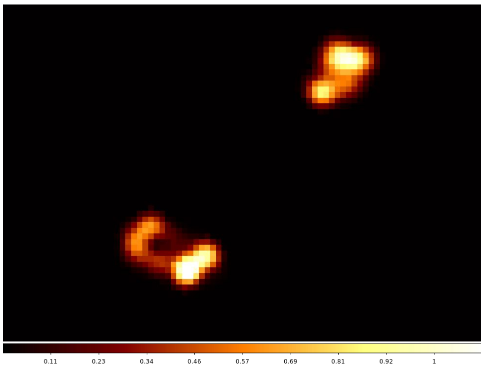
(c) *SB155*



(d) *SB215*



(e) *SB275*



(f) *SB335*

Figure 6.6. *Source model for each subband after 4 passes of self-calibration.*

Comment on behaviour of various components, astrometric accuracy, etc

6.3.2 Spectral Analaysis

Work with Astron guy to extract spectral behaviour of various components of 3C295 over multiple wavelengths, since he's got a purpose-built package

Chapter 7

The Groth Strip with Dutch LOFAR

Contents

7.1 Aims & Methodology

In this section, we describe the work done to make a wide-field image of the entire LOFAR primary beam, centred on 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, the international stations are not used for this part of the project. 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 Extended Groth Strip 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 use our high-resolution model of 3C295.

We aim to achieve a sensitivity of a few hundreds of μJy using visibilities recorded between 110 – 180 MHz. This corresponds to a sensitivity of ~ 20 μJy 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 goal 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 7.3.

7.2 Data Reduction

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 3.5.2) means that data can only be averaged with care. As visibilities are averaged in time and frequency[1], the relative contribution of signal picked up on longer baselines 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, antenna gains 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 kHz[2]. 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:

$$\mathbf{V}_{pq} = \boldsymbol{G}_p \left(\sum_s \boldsymbol{E}_{sp} \boldsymbol{K}_{sp} \mathbf{B}_s \boldsymbol{K}_{sq}^H \boldsymbol{E}_{sq}^H \right) \boldsymbol{G}_q^H \tag{7.1}$$

which is the same as Eq. (3.17). To account for the faceting explicitly, we can write:

$$\mathbf{V}_{pq} = \boldsymbol{G}_p \sum_\Omega \left(\sum_{s\in\Omega} \boldsymbol{E}_{sp} \boldsymbol{K}_{sp} \mathbf{B}_s \boldsymbol{K}_{sq}^H \boldsymbol{E}_{sq}^H \right) \boldsymbol{G}_q^H \tag{7.2}$$

where Ω indices denote different facets, and s indices denote individual sources. 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:

$$\mathbf{V}_{pq} = \boldsymbol{G}_p \sum_\Omega \left(\sum_{s\in\Omega} \boldsymbol{E}_{\Omega p} \boldsymbol{K}_{sp} \mathbf{B}_s \boldsymbol{K}_{sq}^H \boldsymbol{E}_{\Omega q}^H \right) \boldsymbol{G}_q^H \tag{7.3}$$

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

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

We begin with direction-independent calibration to find an estimate of $\boldsymbol{G}_p \forall p$. 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:

$$\mathbf{V}_{pq}^{\mathrm{corr}} = \hat{\boldsymbol{G}}_p^{-1} \mathbf{V}_{pq} \left(\hat{\boldsymbol{G}}_q^H \right)^{-1} \tag{7.4}$$

$$= \hat{\boldsymbol{G}}_p^{-1} \boldsymbol{G}_p \sum_{\Omega} \left(\sum_{s \in \Omega} \boldsymbol{E}_{\Omega p} \boldsymbol{K}_{sp} \mathbf{B}_s \boldsymbol{K}_{\Omega q}^H \boldsymbol{E}_{sq}^H \right) \boldsymbol{G}_q^H \left(\hat{\boldsymbol{G}}_q^H \right)^{-1} \tag{7.5}$$

In so doing, we assume that $(\boldsymbol{E}_{sp} \forall s, p) = 1$ when solving for $\hat{\boldsymbol{G}}$. In general, this is of course invalid. For the calibrator source, however, we have $Ejones_{sp} = 1 \forall p$ 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 $\boldsymbol{G}$. We then have $\hat{\boldsymbol{G}}_p^{-1} \boldsymbol{G}_p \sim 1$, and

$$\mathbf{V}_{pq}^{\mathrm{corr}} \sim \sum_{\Omega} \left(\sum_{s \in \Omega} \boldsymbol{E}_{\Omega p} \boldsymbol{K}_{sp} \mathbf{B}_s \boldsymbol{K}_{sq}^H \boldsymbol{E}_{\Omega q}^H \right) \tag{7.6}$$

and direction-dependent calibration then consists of solving for $\hat{\boldsymbol{E}}_{\Omega p} = \hat{\boldsymbol{G}}_p^{-1} \boldsymbol{G}_p \boldsymbol{E}_{\Omega p} \sim \boldsymbol{E}_{\Omega p}$ for all facets.

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 4 was used to improve self-calibration convergence and the final image.

Direction-dependent calibration was 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. 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 stations[3] introduces a phase error proportional to ν, while differences

[3]Each 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.

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.

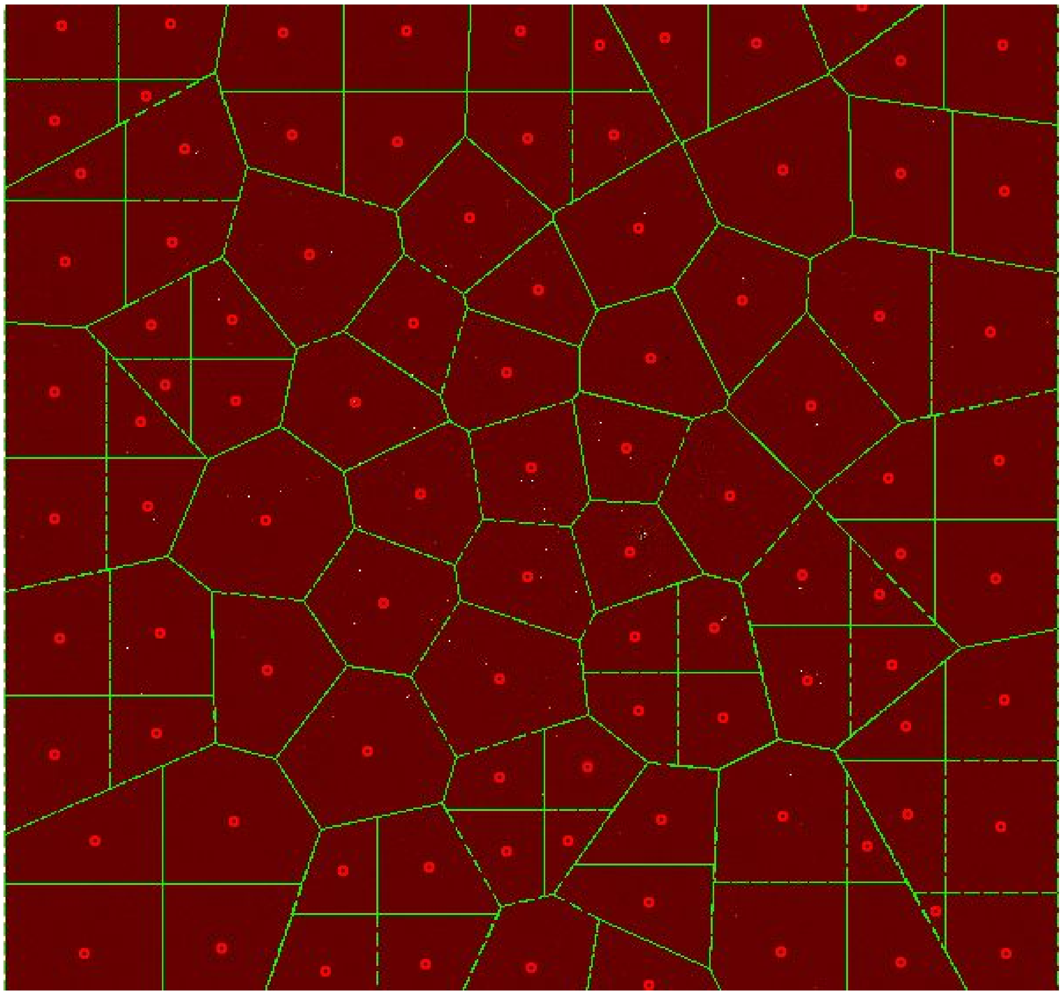
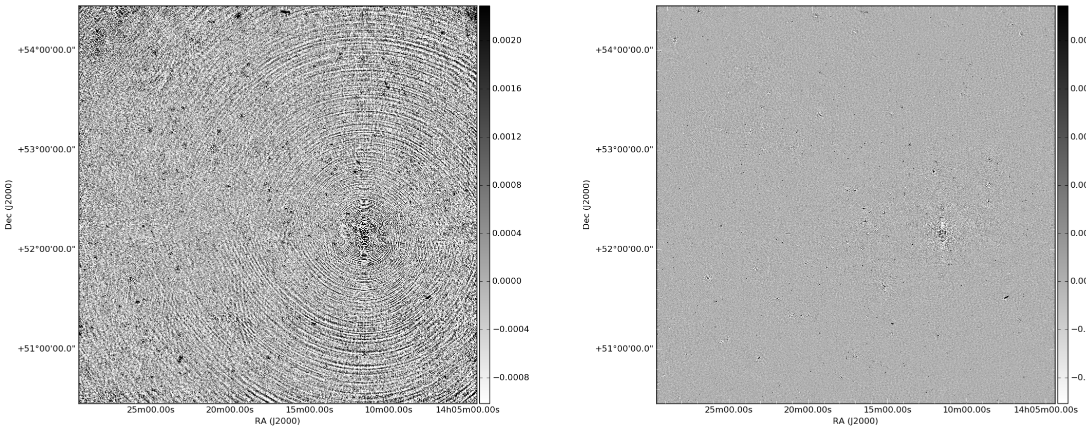


Figure 7.1. *Facet distribution for the full direction-dependent calibrated image of the LOFAR primary beam. Image made without international stations.*

The direction-dependent calibration was done using 121 facets, spread as shown in Fig. 7.1. 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. Calibration is done solving for full, complex Jones matrices, rather than a physics-based or diagonal-and-rotation matrix. 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 AWImager (Tasse et al. 2013), so the image gives integrated flux values rather than apparent flux values. Deconvolution was performed on the full bandwidth[4] using SSD-clean with an external mask.



(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. 7.2a with direction-dependent Jones terms.*

Figure 7.2. *Wide-field image cutouts of the full primary beam, centred on the EGS, before and after selfcal.*

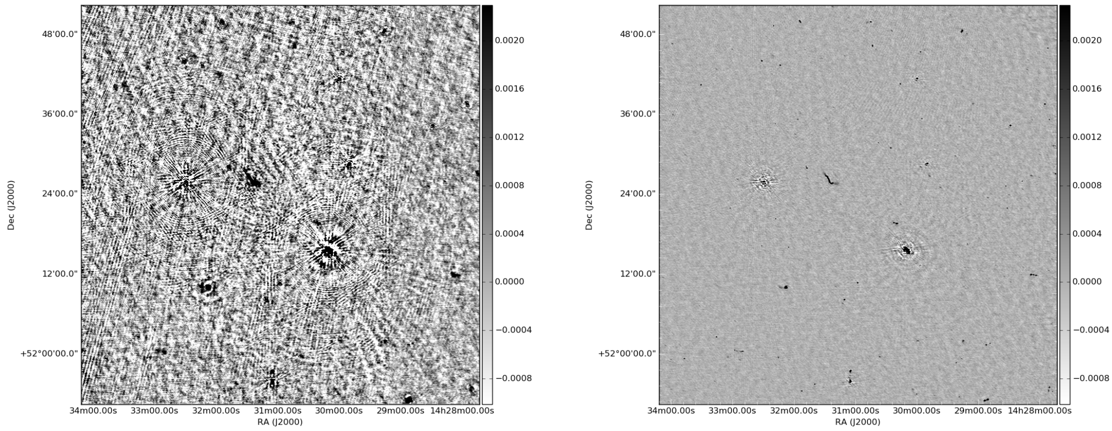
Direction-dependent self-calibration dramatically improves the noise level in the final image, as shown in Fig. 7.2. Note that not only do artefacts centred on 3C295 (which dominate the noise level in Fig. 7.2a) decrease dramatically, but many of the "diffuse" sources in the field also seem to disappear: 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. 7.3. 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.

7.3 Overlays & Images

We select 12 sources in the wide-field image. 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 source positions are given in Table 7.1.

[4]This 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 hightest-frequency band.



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

(b) *LOFAR wide-field image cutout, self-calibrated from Fig. 7.2a with direction-dependent Jones terms.*

Figure 7.3. *Cutouts of smaller regions of Fig. 7.2.*

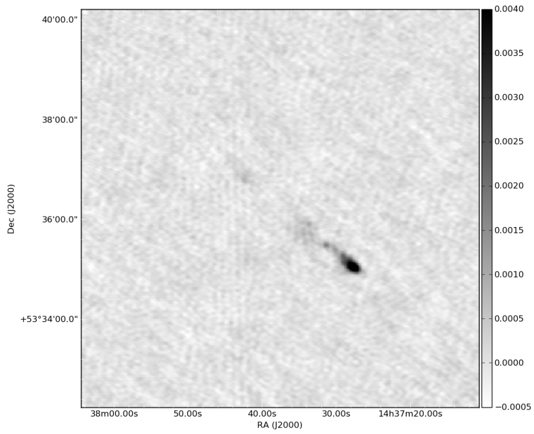
Name	RA [hms]	Dec [dms]	Likely Association	LOFAR thumbnail
EGS-1	14:37:39.53	53:36:31.24	NVSS Radio Galaxy	Fig. 7.4a
EGS-2	14:35:27.84	55:07:56.32	Radio Galaxy in Cluster	Fig. 7.5a
EGS-3	14:29:34.13	54:43:46.93	Radio Galaxy in Cluster	Fig. 7.6a
EGS-4	14:31:36.62	52:27:33.75	Radio galaxy, z = .292	Fig. 7.7a
EGS-5	14:29:48.89	51:10:30.52	Lobes matched individually	Fig. 7.8a
EGS-6	14:26:04.09	51:29:35.07	One lobe matched individually	Fig. 7.9a
EGS-7	14:17:55.58	50:08:01.75	Radio galaxy, z = .186	Fig. 7.10a
EGS-8	14:14:40.42	51:17:41.00	Radio galaxy, z unknown	Fig. 7.11a
EGS-9	14:11:36.43	52:54:25.53	Galaxy cluster, z = .525	Fig. 7.12a
EGS-10	14:07:09.90	55:04:22.23	Galaxy cluster, z = .250	Fig. 7.13a
EGS-11	14:03:16.00	51:43:35.23	Lobes matched individually	Fig. 7.14a
EGS-12	14:02:43.49	51:03:14.69	Radio Galaxy in Cluster	Fig. 7.15a

Table 7.1. *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.*

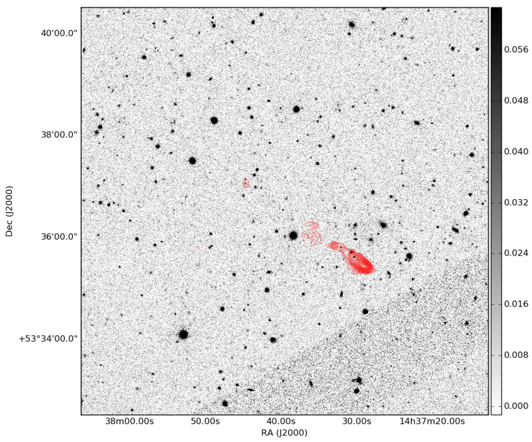
Table 7.1 also gives the likely associations for each chosen EGS source, giving redshift information when available. The analysis here is done by hand using NED (NASA/IPAC Extragalactic Database (NED)), Simbad (operated at CDS, Strasbourg) and the associated catalogues and papers, often accessible through 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, this was not done here, but would be a promising avenue of future work.

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.

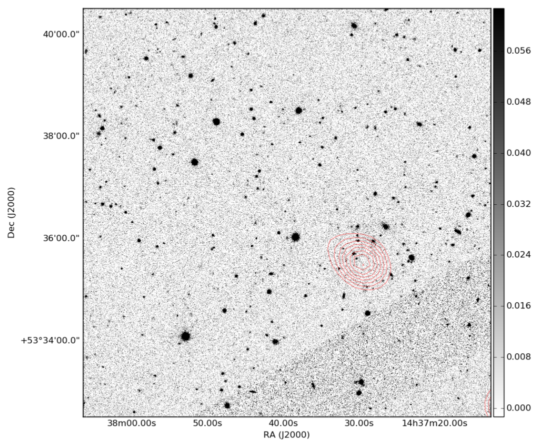
EGS-1



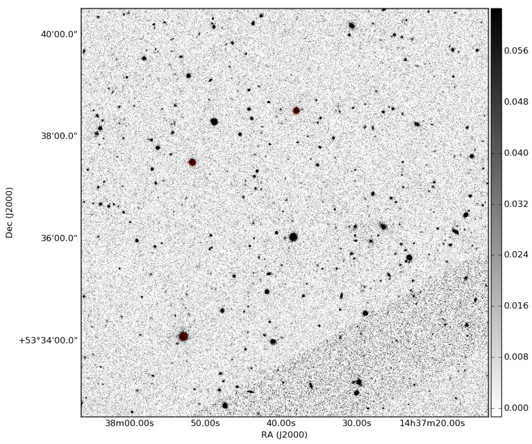
(a) *Cutout of our LOFAR wide-field image centred on EGS-1.*



(b) *SDSS image of EGS-1, with Fig. 7.4a as an overlay.*



(c) *Same image as Fig. 7.4b, with NVSS overlay.*

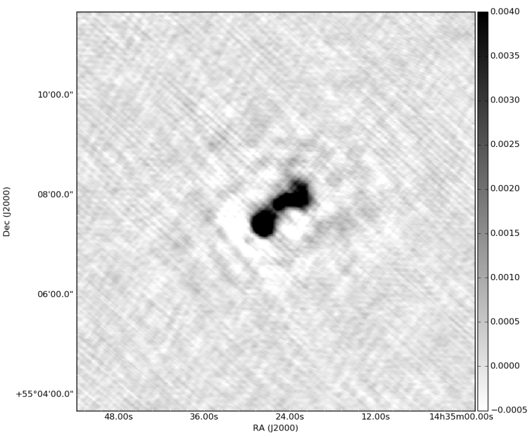


(d) *Same image as Fig. 7.4b, with WISE overlay.*

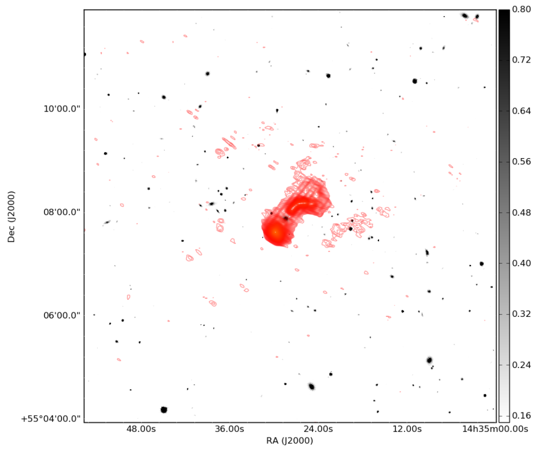
Figure 7.4. *Images & Overlays for EGS-1.*

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. 7.4a. This is congruent with a radio galaxy with one Doppler-boosted jet. The absence of an optical counterpart could be due to dust extinction or to insufficient sensitivity.

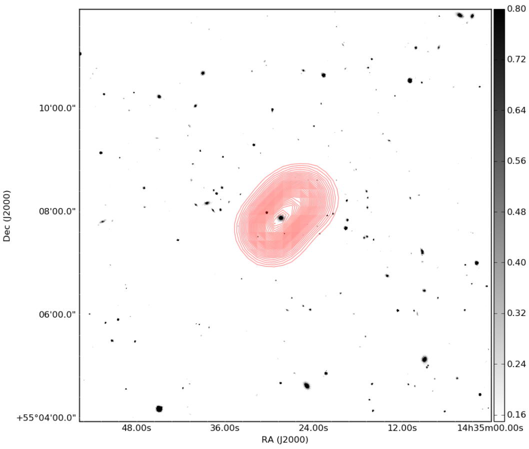
EGS-2



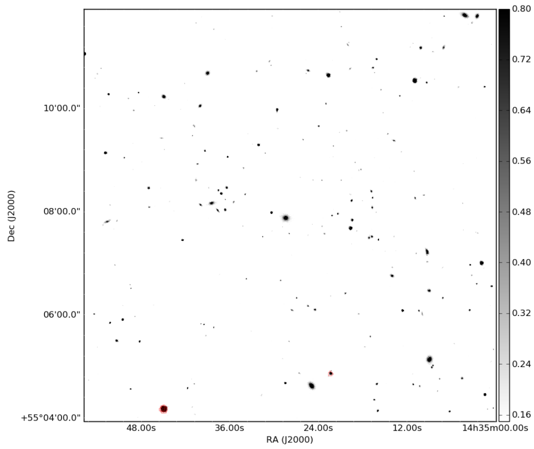
(a) *Cutout of our LOFAR wide-field image centred on EGS-2.*



(b) *SDSS image of EGS-2, with Fig. 7.5a as an overlay.*



(c) *Same image as Fig. 7.5b, with NVSS overlay.*



(d) *Same image as Fig. 7.5b, with WISE overlay.*

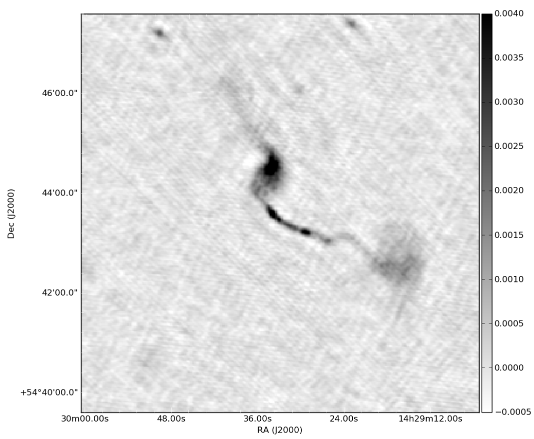
Figure 7.5. *Images & Overlays for EGS-2.*

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+5507519[5] - 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.

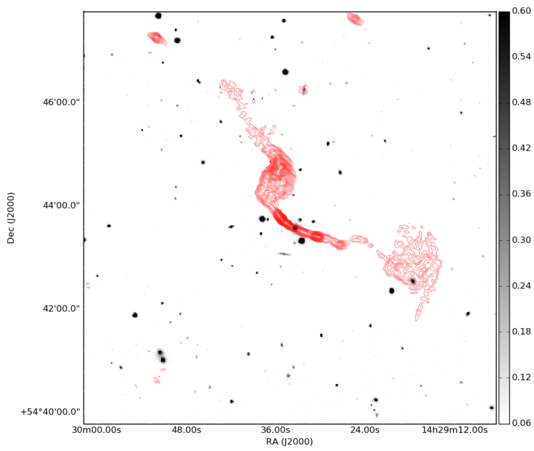
[5] `http://simbad.u-strasbg.fr/simbad/sim-id?Ident=%40499314&Name=2MASX%20J14352846%2b5507519&submit=submit`

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.

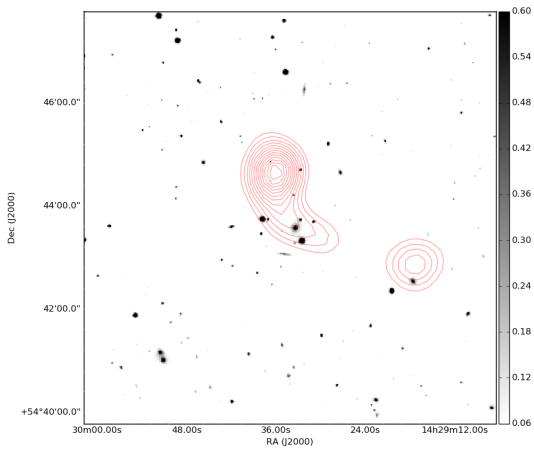
EGS-3



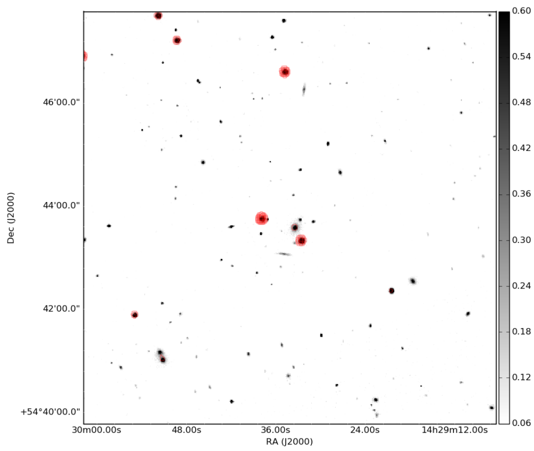
(a) *Cutout of our LOFAR wide-field image centred on EGS-3.*



(b) *SDSS image of EGS-3, with Fig. 7.6a as an overlay.*



(c) *Same image as Fig. 7.6b, with NVSS overlay.*



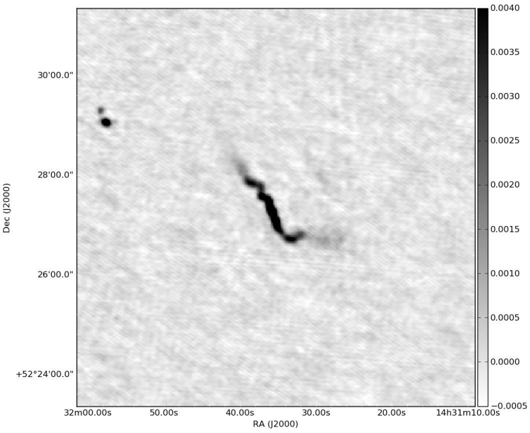
(d) *Same image as Fig. 7.6b, with WISE overlay.*

Figure 7.6. *Images & Overlays for EGS-3.*

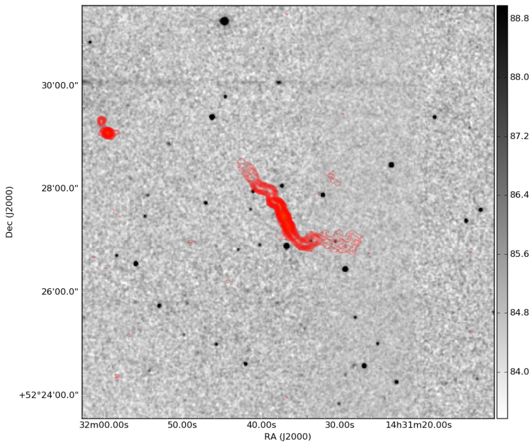
EGS-3 is one of the most impressive sources in the primary beam. It has very complex diffuse structure. It almost certainly consists of two jets, though the interactions of these jets with the neighbouring medium must be complex indeed. From Fig. 7.6c, one can see that the emission can be successfully 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 ambiant 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. 7.6a. EGS-3 can therefore be fairly unambiguously associated to a galaxy cluster.

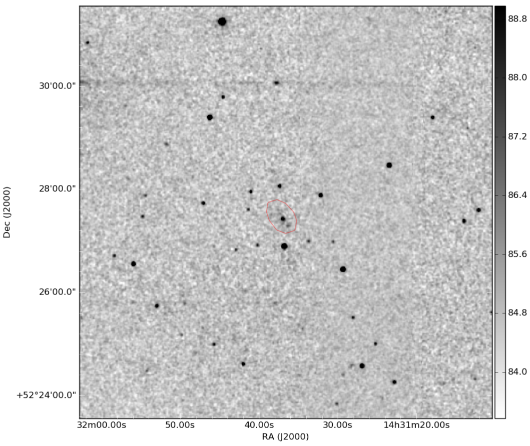
EGS-4



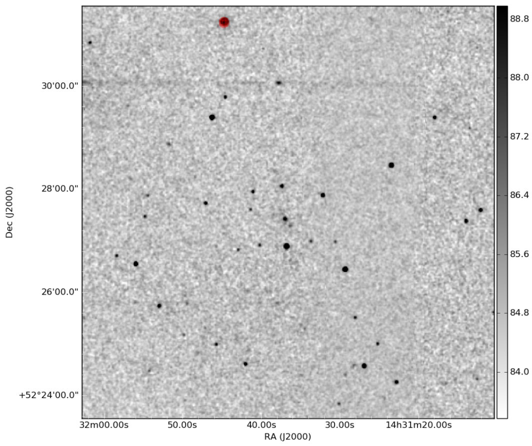
(a) *Cutout of our LOFAR wide-field image centred on EGS-4.*



(b) *SDSS image of EGS-4, with Fig. 7.7a as an overlay.*



(c) *Same image as Fig. 7.7b, with NVSS overlay.*

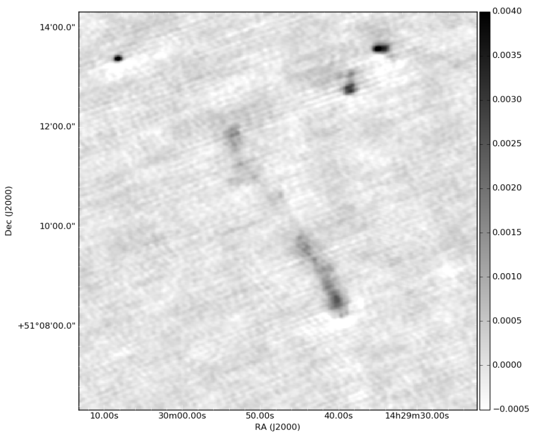


(d) *Same image as Fig. 7.7b, with WISE overlay.*

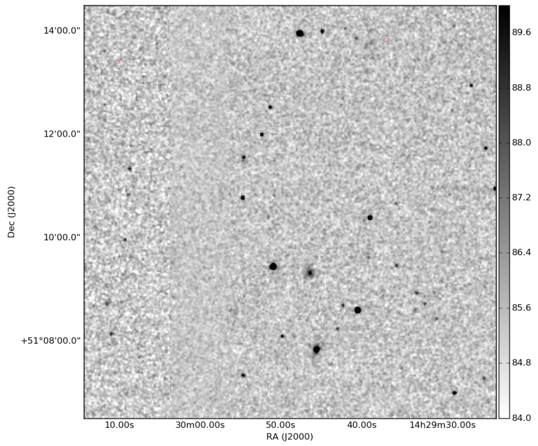
Figure 7.7. *Images & Overlays for EGS-4.*

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. 7.7c, 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.

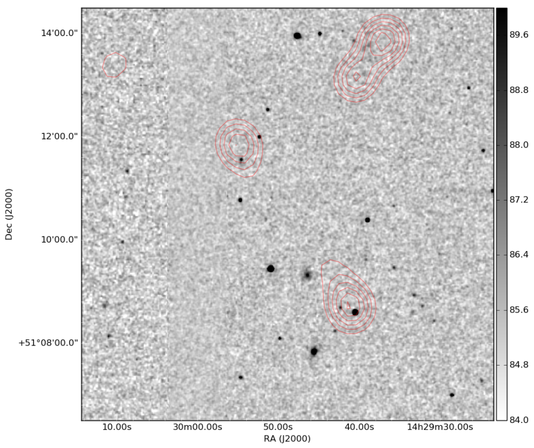
EGS-5



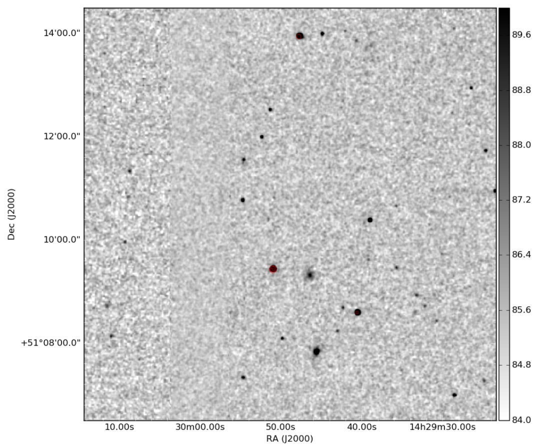
(a) *Cutout of our LOFAR wide-field image centred on EGS-5.*



(b) *SDSS image of EGS-5, with Fig. 7.8a as an overlay.*



(c) *Same image as Fig. 7.8b, with NVSS overlay.*

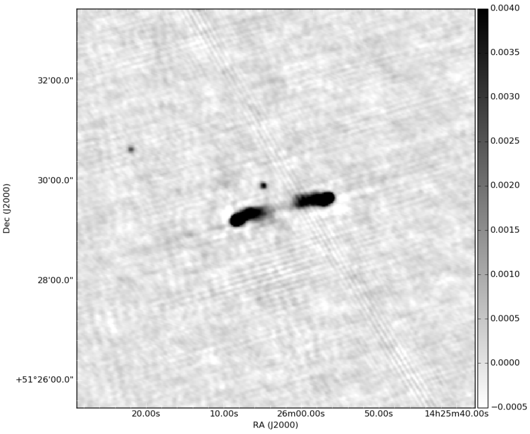


(d) *Same image as Fig. 7.8b, with WISE overlay.*

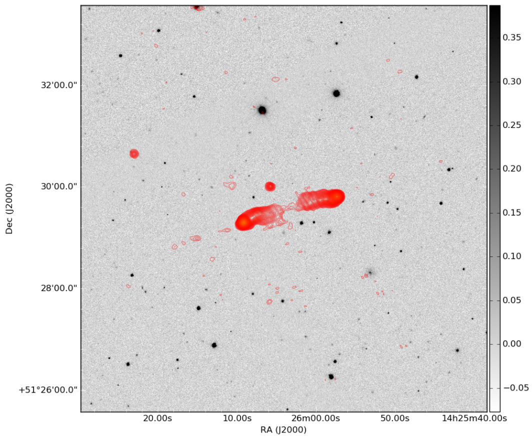
Figure 7.8. *Images & Overlays for EGS-5.*

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.

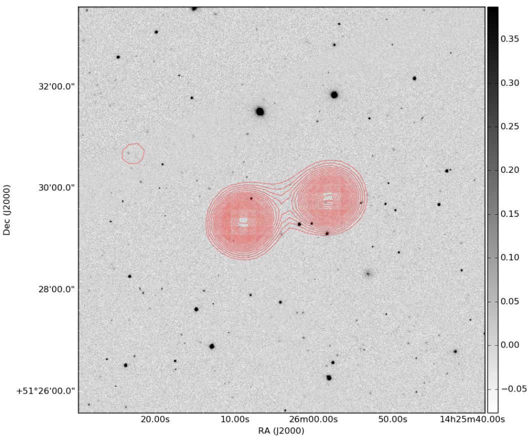
EGS-6



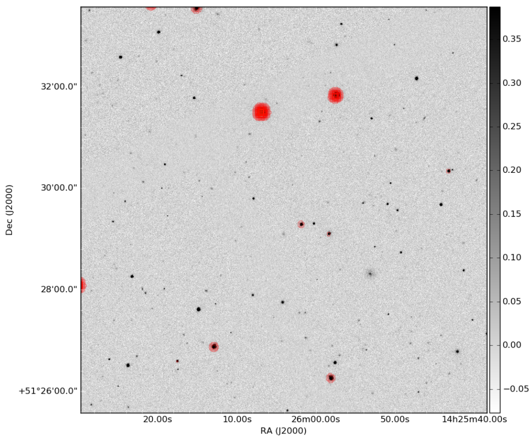
(a) *Cutout of our LOFAR wide-field image centred on EGS-6.*



(b) *SDSS image of EGS-6, with Fig. 7.9a as an overlay.*



(c) *Same image as Fig. 7.9b, with NVSS overlay.*

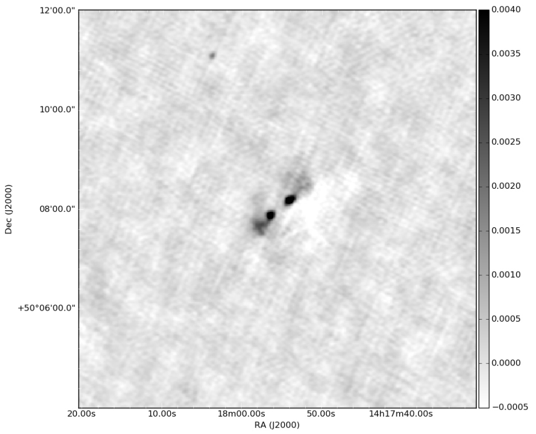


(d) *Same image as Fig. 7.9b, with WISE overlay.*

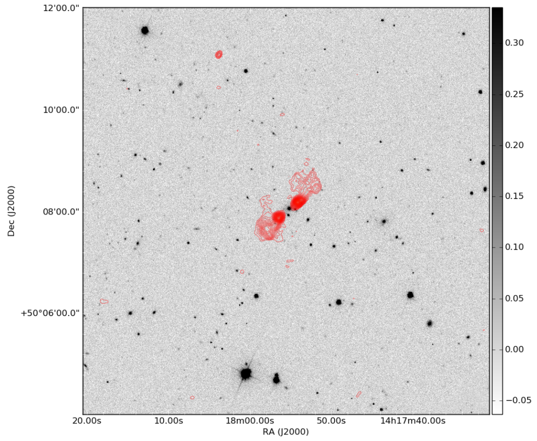
Figure 7.9. *Images & Overlays for EGS-6.*

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.

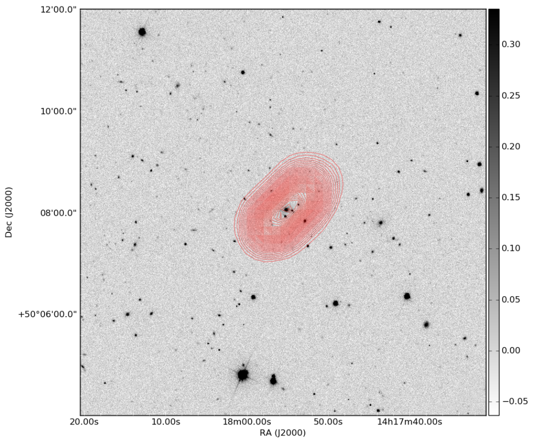
EGS-7



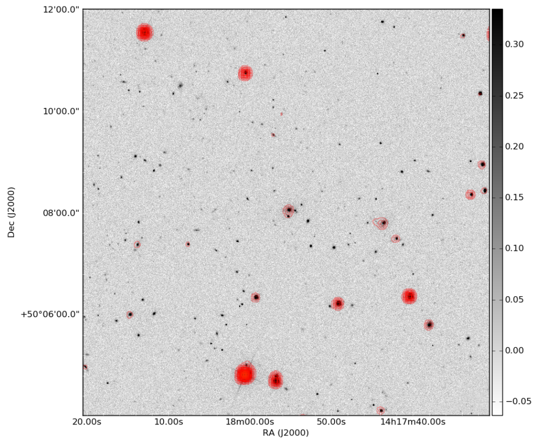
(a) *Cutout of our LOFAR wide-field image centred on EGS-7.*



(b) *SDSS image of EGS-7, with Fig. 7.10a as an overlay.*



(c) *Same image as Fig. 7.10b, with NVSS overlay.*

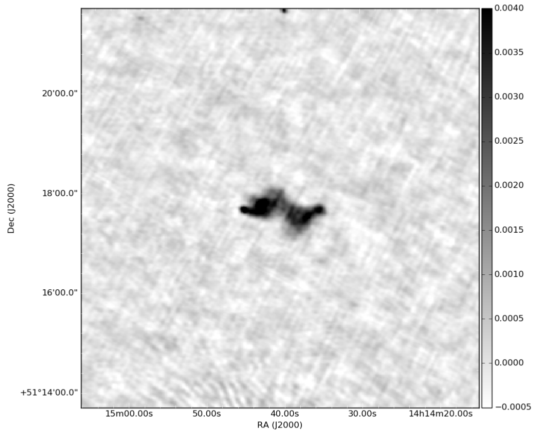


(d) *Same image as Fig. 7.10b, with WISE overlay.*

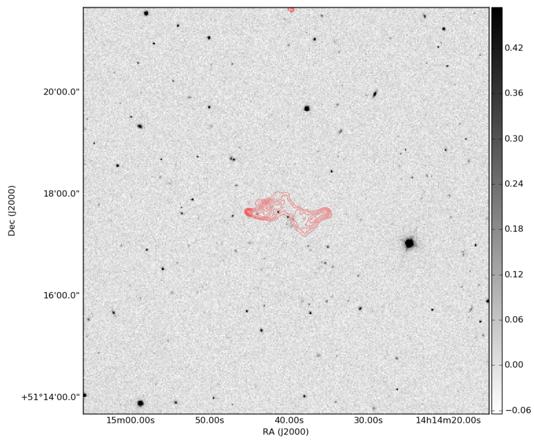
Figure 7.10. *Images & Overlays for EGS-7.*

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.

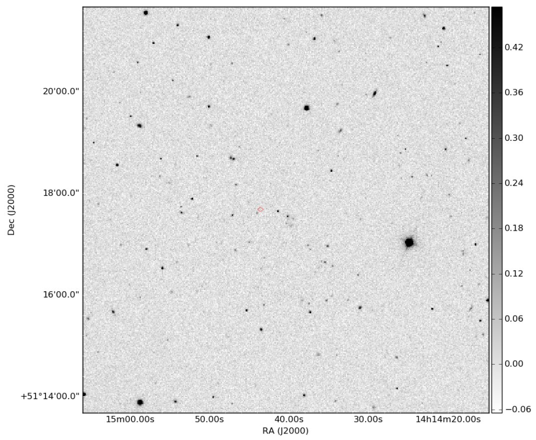
EGS-8



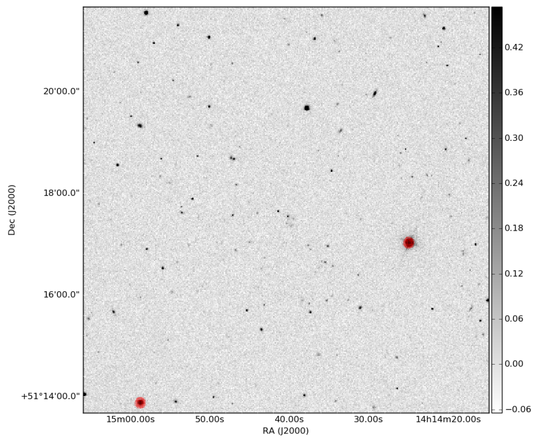
(a) *Cutout of our LOFAR wide-field image centred on EGS-8.*



(b) *SDSS image of EGS-8, with Fig. 7.11a as an overlay.*



(c) *Same image as Fig. 7.11b, with NVSS overlay.*

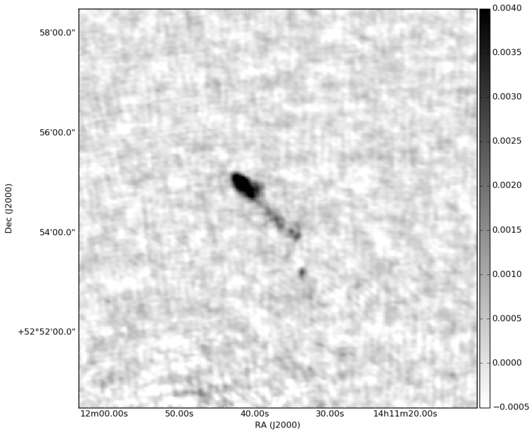


(d) *Same image as Fig. 7.11b, with WISE overlay.*

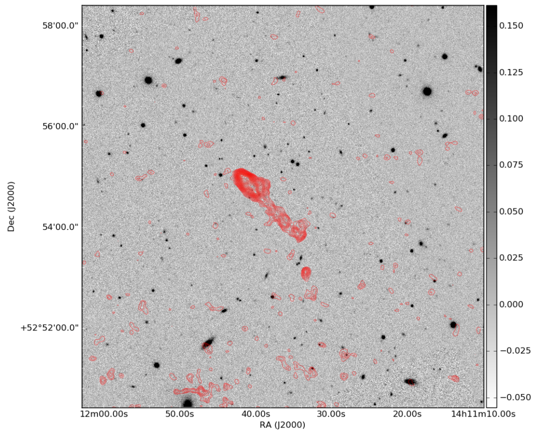
Figure 7.11. *Images & Overlays for EGS-8.*

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.

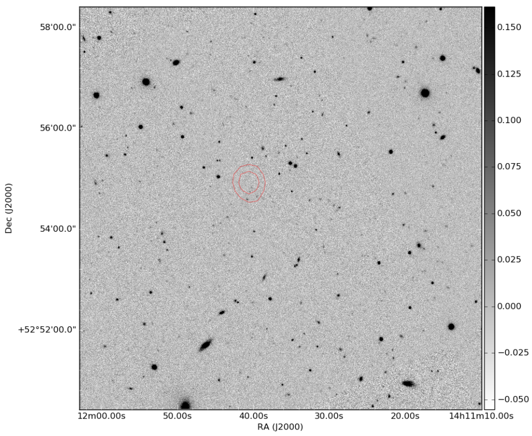
EGS-9



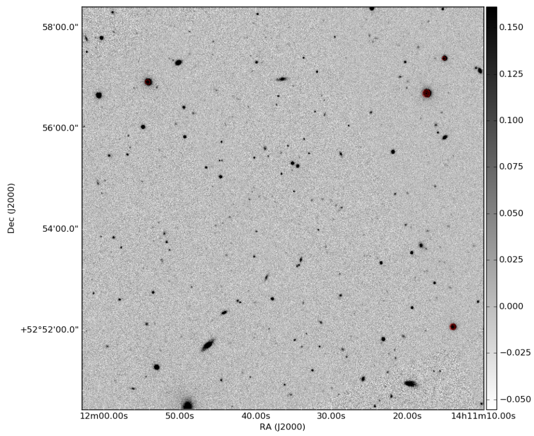
(a) *Cutout of our LOFAR wide-field image centred on EGS-9.*



(b) *SDSS image of EGS-9, with Fig. 7.12a as an overlay.*



(c) *Same image as Fig. 7.12b, with NVSS overlay.*



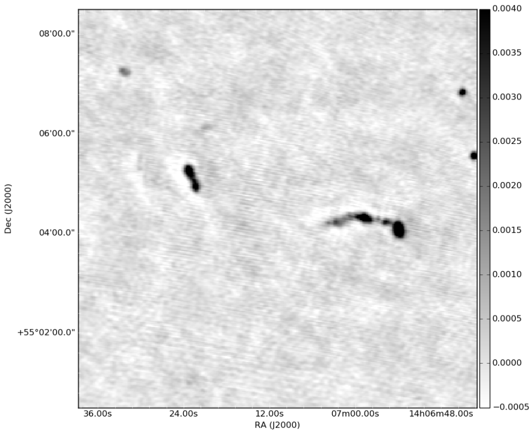
(d) *Same image as Fig. 7.12b, with WISE overlay.*

Figure 7.12. *Images & Overlays for EGS-9.*

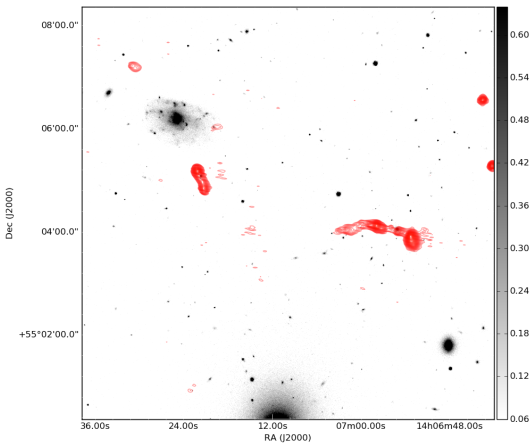
This source was chosen because it had one strong lobe and one weak: it was expected to be a good example of a radio galaxy with a single doppler-boosted lobe. There is a likely association to the northern lobe: the galaxy cluster CFHT-W CL J141141.7+525450 (Wen & Han 2011), which lies at $z = .5252$. 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.

The southern source has an NVSS equivalent, but is still dominated by artefacts in the LOFAR image. We therefore do not seek to extract information from this source at the time of writing.

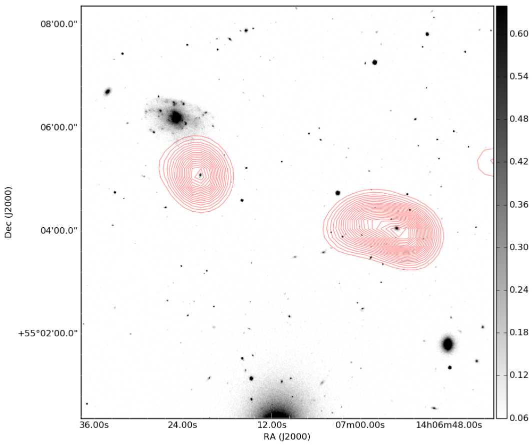
EGS-10



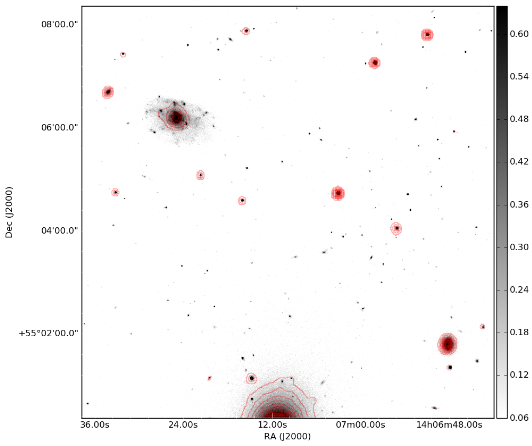
(a) *Cutout of our LOFAR wide-field image centred on EGS-10.*



(b) *SDSS image of EGS-10, with Fig. 7.13a as an overlay.*



(c) *Same image as Fig. 7.13b, with NVSS overlay.*

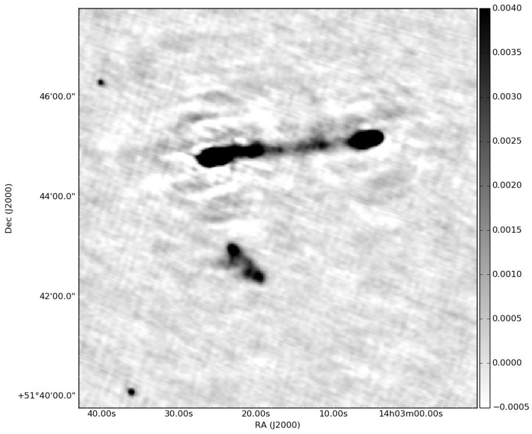


(d) *Same image as Fig. 7.13b, with WISE overlay.*

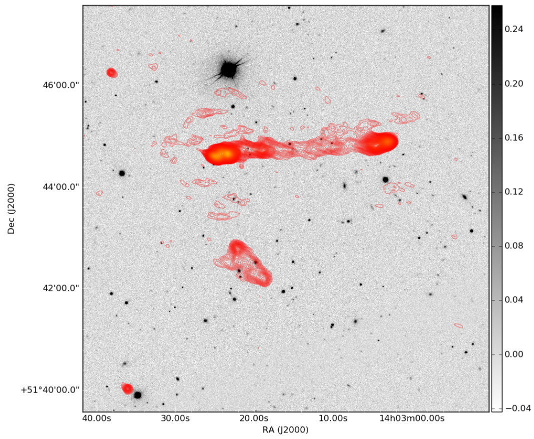
Figure 7.13. *Images & Overlays for EGS-10.*

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.

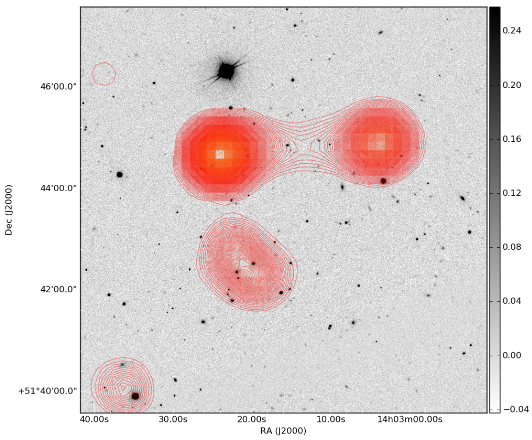
EGS-11



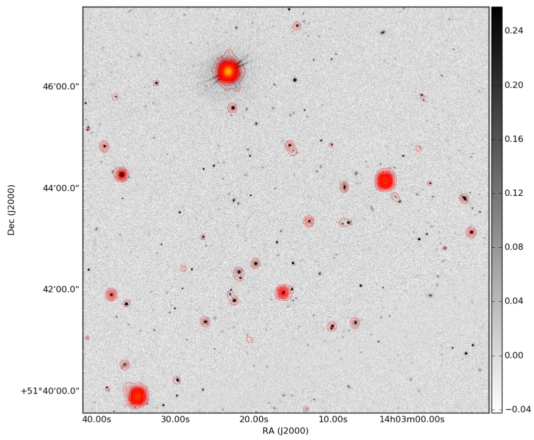
(a) *Cutout of our LOFAR wide-field image centred on EGS-11.*



(b) *SDSS image of EGS-11, with Fig. 7.14a as an overlay.*



(c) *Same image as Fig. 7.14b, with NVSS overlay.*

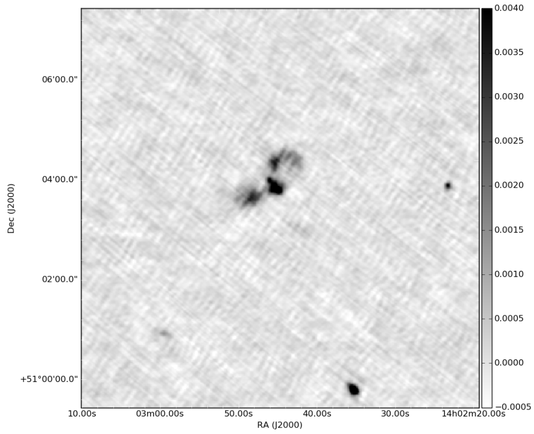


(d) *Same image as Fig. 7.14b, with WISE overlay.*

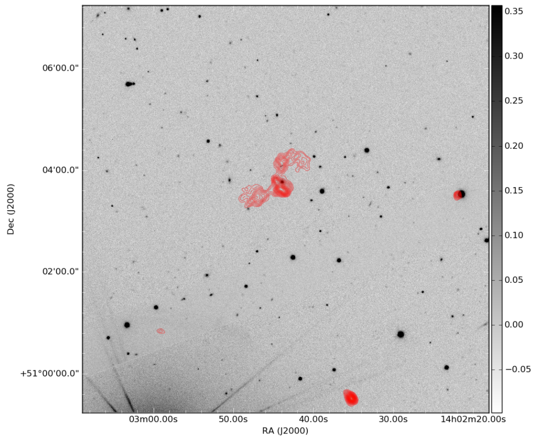
Figure 7.14. *Images & Overlays for EGS-11.*

This source was chosen because of its size and very straight jets/lobes. While Fig. 7.14a 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.

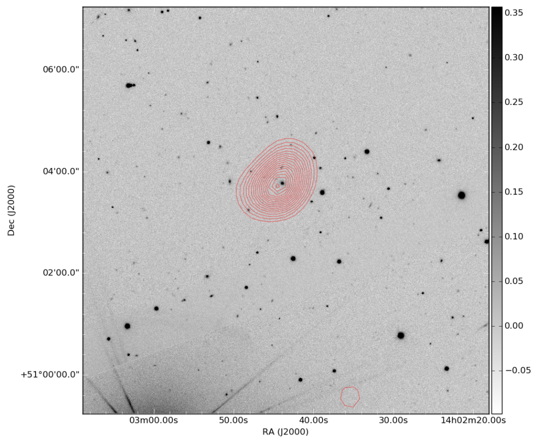
EGS-12



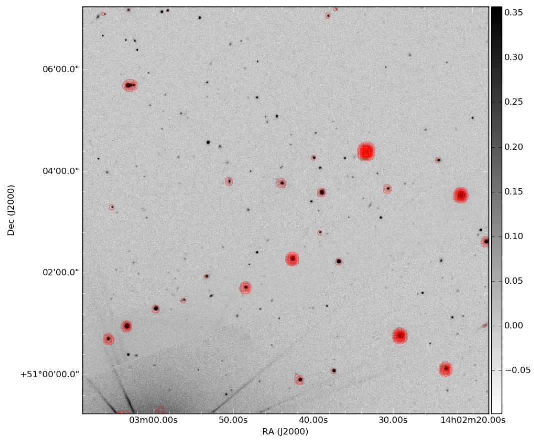
(a) *Cutout of our LOFAR wide-field image centred on EGS-12.*



(b) *SDSS image of EGS-12, with Fig. 7.15a as an overlay.*



(c) *Same image as Fig. 7.15b, with NVSS overlay.*



(d) *Same image as Fig. 7.15b, with WISE overlay.*

Figure 7.15. *Images & Overlays for EGS-12.*

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.

No matching was done for the other two radio sources in the field, as they are unresolved and far from the main source of interest, and thus unlikely to be linked. They do, however, have NVSS counterparts, and are thus likely to be true sources.

7.4 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, low-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 as-of-yet unimplemented in the DDF-kMS pipeline). As such, we look in the large-field image for interesting or unusual sources, and attempt to match them to known sources in other bands in order to learn about them. 12 sources were selected, and the most interesting were consistently matched to brightest galaxies in clusters. This indicates that there are many sources with very interesting diffuse emission in and around the Extended Groth Strip, and some of which merit further observations to better resolve their turbulent structure.

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.

show image of 3C295 calibrated w/ the full source extraction model of the primary beam, compared to result of self-cal?

Chapter 8

Imaging the Extended Groth Strip with the full LOFAR array

Contents

8.1 Aims & Methodology

primary aim: see if direction-independent calibration of EGS good enough, for international baselines, to allow for HR imaging of entire EGS with acceptable upper bounds to decorrelation throughout field.

in other words: can entire EGS be placed on same facet as 3C295 at high resolutions? Link to V.

necessary test: investigate impact of decorrelation as function of distance from 3C295 (calibrator) - do we lose enough signal to lose resolution?

LOBOS survey gives list of VLBI phase calibrators in primary beam: compact sources, should show impact of direction-dependent gain errors directly. These are distributed at various distances from 3C295, conveniently.

Show images (+ overlays?) of direction-dependent PSF (i.e. modelled decorrelation due to time-freq averaging) and images of LOBOS sources

Finally, show plot of ratio of $flux_{peak}$ / $flux_{integrated}$ as function of distance from 3C295: the flatter, the better.

Imaging strategy: image the field patch by patch.

Have source list of entire EGS from section V. Use that to subtract all source save those in the patch from visibilities, then image (self-cal?) over that patch. Rinse and repeat throughout EGS.

8.2 Testing Decorrelation - LOBOS Sources in the Primary Beam

describe LOBOS catalogue & sources chosen - why, how

#	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 8.1. *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.*

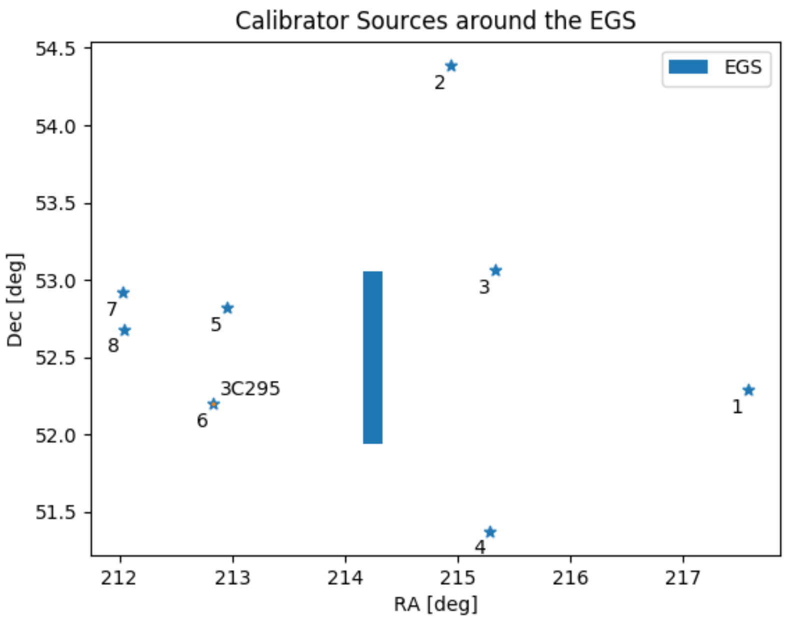


Figure 8.1. *Position of the calibrator sources around the EGS. Note that this is not projected properly onto the celestial sphere.*

show images of decorrelation, residuals etc for all 8 sources

8.3 Patchwise Imaging of the EGS using LOFAR International Stations

Chapter 9

Conclusion & Future Work

Contents

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Morbi vehicula dui vel sem elementum viverra. Nullam molestie egestas finibus. Vestibulum at rhoncus massa, ac mollis nulla. Nam vitae vestibulum ex. Interdum et malesuada fames ac ante ipsum primis in faucibus. Curabitur lobortis quam eget dui pretium pulvinar. Quisque consequat dapibus justo, eu blandit lacus auctor id. Vivamus vestibulum diam id feugiat posuere. Morbi ut vehicula urna, id laoreet turpis. Nam dignissim magna augue, et rhoncus felis lacinia nec. Praesent convallis ultrices hendrerit.

Praesent porta lacus est. Duis tempor augue augue, ac pharetra nisi viverra sed. Integer tristique risus ac metus aliquam, in pretium felis porttitor. Sed eu ipsum vitae tortor faucibus ultrices. Cras sem erat, eleifend eget massa et, dapibus vulputate velit. Duis semper non odio in semper. Mauris quis massa rhoncus ante ullamcorper faucibus. Nam quis sollicitudin diam. Maecenas et posuere augue, ac aliquam sapien. Curabitur ac tristique ante. Sed tempor tellus a magna vehicula fringilla. Suspendisse arcu lacus, bibendum vitae velit vel, posuere commodo ipsum. Suspendisse et congue sem. Maecenas maximus quam sed interdum dignissim. Etiam ac tempor risus. Quisque vel nisi arcu. Nunc consectetur, nisl nec blandit egestas, dolor felis tincidunt odio, et iaculis lacus nunc in sem. Suspendisse potenti. Aenean est ante, egestas ac nulla eu, suscipit maximus orci. Etiam suscipit neque sed diam hendrerit, eu iaculis risus laoreet. Sed venenatis ultricies justo porta placerat. Quisque ac convallis nunc. Suspendisse et risus enim.

Bibliography

Adelman-McCarthy, J. K. & et al. 2009, VizieR Online Data Catalog, 2294

Ahn, C. P. & et al. 2013, VizieR Online Data Catalog, 5139

Anderson, L. D., Bania, T. M., Jackson, J. M., et al. 2009, ApJS, 181, 255

Atemkeng, M. T., Smirnov, O. M., Tasse, C., Foster, G., & Jonas, J. 2016, MNRAS, 462, 2542

Best, P. N. & Heckman, T. M. 2012, MNRAS, 421, 1569

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., Bozyan, 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

Fried, D. L. 1978, Journal of the Optical Society of America (1917-1983), 68, 1651

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

Hales, C. A. 2013, ArXiv e-prints

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

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

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

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

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

Noordam, J. E. & Smirnov, O. M. 2010, A&A, 524, A61

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

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

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. & Tasse, C. 2015, MNRAS, 449, 2668

Spitzer, L. 1978, Physical processes in the interstellar medium

Tasse, C. 2014, ArXiv e-prints

Tasse, C., Hugo, B., Mirmont, M., et al. 2017, ArXiv e-prints

Tasse, C., Röttgering, H. J. A., Best, P. N., et al. 2007, A&A, 471, 1105

Tasse, 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

van Cittert, P. H. 1934, Physica, 1, 201

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

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

Un nouvel algorithme permettant des améliorations dramatiques dans la création d'images interférométriquesà faible coût computationnel est développé dans l'optique d'être utilisée pour LOFAR et autres précurseurs SKA. Cet algorithme est déployé dans le pipeline LOFAR, qui est utilisé pour créer la première image grand-champ et haute-résolution faite avec le LOFAR international. Cette image, du champ extragalactique Extended Groth Strip, possède une résolution comparable à celle du télescope Hubble.

Abstract

A new algorithm was devised to dramatically improve radio interferometric images at near-zero computational cost. This algorithm was deployed in the LOFAR pipeline and used to create a high-resolution, wide-field image of a famous deep extragalactic field, the Extended Groth Strip using international LOFAR. This image will match the Hubble Space Telescope's resolution.

Mots Clés

Caesar licentia post honoratis haec adhibens urbium honoratis nullum Caesar.

Keywords

Delatus delatus nominatus onere aut trahebatur quod tenus et bonorum.