
Statistical Inference In Radio Astronomy

Henrik Junklewitz



München 2014

Statistical Inference In Radio Astronomy

Henrik Junklewitz

Dissertation
an der Fakultät für Physik
der Ludwig–Maximilians–Universität
München

vorgelegt von
Henrik Junklewitz
aus Wilhelmshaven

München, den 14. Februar 2014

Erstgutachter: Prof. Dr. Simon D. M. White

Zweitgutachter: PD. Dr. Torsten A. Enßlin

Tag der mündlichen Prüfung: 10. April 2014

Contents

1	Introduction	1
1.1	Statistical inference	1
1.1.1	Bayesian inference	2
1.2	Radio astronomy	3
1.2.1	Instruments	3
1.2.2	Radio astrophysics	5
1.2.3	Radio synchrotron observations	6
1.2.4	Turbulence, dynamo theory and Gaussian magnetic fields	8
1.3	Outline	10
2	RESOLVE: A new algorithm for aperture synthesis imaging of extended emission in radio astronomy	13
2.1	Introduction	13
2.2	The algorithm	15
2.2.1	Aperture Synthesis	15
2.2.2	Signal Inference in Radio Astronomy	16
2.2.3	RESOLVE: Radio Extended Sources Lognormal Deconvolution Estimator	20
2.3	Mathematical derivation of RESOLVE	25
2.3.1	Reconstruction of the signal field s	26
2.3.2	Uncertainty of the signal reconstruction	27
2.3.3	Reconstruction of the power spectrum parameters p	28
2.4	Implementation of RESOLVE	31
2.4.1	General implementation	31
2.4.2	Analysis of algorithmic efficiency	32
2.5	A signal inference view on visibility weighting	33
2.6	Test Simulations	35
2.6.1	Main Test Results	37
2.6.2	Comparison to standard imaging methods	37
2.6.3	Comparison with a real signal	44
2.6.4	Signal Uncertainty	44
2.6.5	Power Spectrum Reconstructions	48
2.7	Conclusions	49

3	A first application of RESOLVE: The galaxy cluster Abell 2256	53
3.1	Introduction: Galaxy clusters as an optimal field of application for RESOLVE	53
3.1.1	The galaxy cluster Abell 2256	54
3.2	Results	54
4	A new approach to multi-frequency synthesis in radio interferometry	57
4.1	Introduction	57
4.2	Theory & Algorithm	59
4.2.1	Multi-frequency aperture synthesis and spectral index reconstruction	59
4.2.2	A multi-frequency extension to the RESOLVE algorithm	61
4.3	Details of the algorithm	66
4.3.1	Reconstruction of the sky brightness signal field s	66
4.3.2	Reconstruction of the spectral index α	67
4.3.3	Combined algorithm	67
4.4	Tests	69
4.4.1	Main test results	70
4.4.2	Comparison to standard methods	72
4.4.3	Power spectrum reconstructions	73
4.5	Conclusion	73
5	Magnetic field statistics	79
5.1	Faraday Caustics	79
5.1.1	Distribution of Faraday caustics in a Gaussian random magnetic field	81
5.2	The LITMUS test	89
5.2.1	Mathematical derivation	90
5.3	Conclusions	92
6	Conclusions	93
A	The maximum entropy principle	97
	Bibliography	99

List of Figures

2.1	Flow chart for RESOLVE.	25
2.2	VLA uv -coverage and psf.	36
2.3	RESOLVE reconstruction for simulated data.	38
2.4	Comparison of RESOLVE with MS-CLEAN I	41
2.5	Comparison of RESOLVE with MS-CLEAN II	42
2.6	Comparison of RESOLVE with MEM	45
2.7	RESOLVE uncertainty maps.	46
2.8	RESOLVE reconstruction with more realistic simulations.	47
2.9	RESOLVE power spectrum reconstruction.	49
3.1	CLEAN images of Abell 2256	55
3.2	First results of RESOLVE on real data.	56
4.1	VLA single- and multi-frequency uv -coverages	69
4.2	RESOLVE spectral index power spectrum reconstruction.	74
4.3	Multi-frequency RESOLVE reconstruction with simulated data.	75
4.4	Comparison of different methods for spectral index reconstruction	76
4.5	Comparison of multi-frequency RESOLVE and MS-MF-CLEAN	77
5.1	Magnetic field correlation.	88
5.2	Number distribution of caustic spikes per logarithmic interval of \mathcal{F}	88

Zusammenfassung

Diese Arbeit vereinigt eine Reihe wissenschaftlicher Studien, die sich alle Problemen aus der statistischen Datenanalyse in Radioastronomie und Astrophysik widmen.

Die Radioastronomie, wie auch die Astronomie als Ganzes, hat durch Einführung neuer Instrumente und Technologien in den letzten zwanzig Jahren eine bemerkenswerte Entwicklung erlebt. Neue Teleskope, wie das erweiterte VLA, LOFAR, oder das SKA und seine Pathfinder-Missionen, versprechen bisher unerreichte Messgenauigkeit, die Erkundung bis heute kaum untersuchter Frequenzbereiche und neuartige Fähigkeiten zur Himmelsdurchmusterung (Garrett 2012). Diese Teleskope haben das Potential die Astrophysik und Kosmologie bedeutend voran zu bringen, von Sonnen - und Sternenphysik, der Astrophysik der Milchstraße und kosmischen Magnetfeldern, zur Astrophysik der Galaxienhaufen oder Signalen aus der Epoche der kosmischen Reionisation (das deutsche SKA-Weißbuch bietet hier einen hervorragenden wissenschaftlichen Überblick, Aharonian et al. 2013).

Parallel dazu machen auch Radiodatenanalyse und Bildgebungsverfahren eine ähnliche Phase neuer Entwicklungen durch, in der die Grenzen des Machbaren vorangetrieben werden um das Feld an die neuen Instrumente und an das wissenschaftlich Mögliche anzupassen. Die vorliegende Arbeit trägt zu diesen Entwicklungen in zwei spezifischen Themengebieten bei: Der Bildgebung in der Radiointerferometrie und der statistischen Analyse kosmischer Magnetfelder.

Die Methoden zur Datenanalyse, die dabei im Laufe der Arbeit verwendet werden, sind zwar durchaus unterschiedlich, können aber allgemein dem Anwendungsbereich der *Inferenzstatistik* zugeordnet werden. Unter diesem Begriff findet sich eine Vielfalt an statistischen Methoden, die sich allgemein mit dem Problem befassen, wie Deduktionen auf der Grundlage unvollständiger Informationen, Messungen, oder Daten vorgenommen werden können. Diese Arbeit bedient sich dabei speziell der Bayesianischen Statistik, die sich durch eine subjektive Definition von Wahrscheinlichkeiten auszeichnet, mit der es auch möglich ist statistische Informationen und Wahrscheinlichkeiten *vor* einer eigentlichen Messung festzulegen.

Die Arbeit enthält Anwendungen solcher Techniken aus zwei unterschiedlichen Bereichen. Erstens, Situationen in denen ein kompliziertes, nicht exakt lösbares Messproblem durch ein Verfahren bearbeitet wird, in welchem die gesuchte Messgröße näherungsweise bestimmt wird, indem für diese eine Priorwahrscheinlichkeit angenommen wird. Probleme dieser Art finden sich oft in Situationen, in welchen ein Messgerät weniger Daten aufnimmt als nötig wären um alle Freiheitsgrade des Problems abzudecken. In dieser Arbeit

ist dies das Messproblem eines Radiointerferometers, welches ein unvollständig gerastertes Bild der Fourier-transformierten Radiointensität am Himmel aufnimmt, so dass eine eindeutige Bildrekonstruktion unmöglich wird. Der Bildgebungsalgorithmus RESOLVE wird vorgestellt, der eine optimale Lösung dieses Messproblems für ausgedehnte Radioquellen erzeugt. Des Weiteren wird das Verhalten des Algorithmus in einem ersten Versuch an echten Daten getestet. Außerdem wird ein neuer Bayesianischer Ansatz zur Multifrequenz-Bildgebung mit einem Radiointerferometer eingeführt und in den RESOLVE Algorithmus integriert.

Der zweite Anwendungsbereich kommt aus Problemen der Astrophysik, in denen die inhärente stochastische Natur eines physikalischen Prozesses notwendigerweise eine Beschreibung benötigt, in der physikalische Größen nur statistisch gemessen werden können. Ein Beispiel sind astrophysikalische Plasmen, die sich oft in einem turbulenten Zustand befinden und deswegen durch statistische Gesetzmäßigkeiten der Hydrodynamik bestimmt sind. Aus diesem Bereich werden zwei Studien präsentiert die beispielhaft zeigen, wie die Eigenschaften turbulenter Plasmamagnetfelder aus Radiobeobachtungen statistisch ermittelt werden können.

Summary

This thesis unifies several studies, which all are dedicated to the subject of statistical data analysis in radio astronomy and radio astrophysics.

Radio astronomy, like astronomy as a whole, has undergone a remarkable development in the past twenty years in introducing new instruments and technologies. New telescopes like the upgraded VLA, LOFAR, or the SKA and its pathfinder missions offer unprecedented sensitivities, previously uncharted frequency domains and unmatched survey capabilities (see Garrett 2012). Many of these have the potential to significantly advance the science of radio astrophysics and cosmology on all scales, from solar and stellar physics, Galactic astrophysics and cosmic magnetic fields, to Galaxy cluster astrophysics and signals from the epoch of reionization (see the German SKA white paper for a broad scientific overview, Aharonian et al. 2013).

Since then, radio data analysis, calibration and imaging techniques have entered a similar phase of new development to push the boundaries and adapt the field to the new instruments and scientific opportunities (e.g. Smirnov 2011a). This thesis contributes to these greater developments in two specific subjects, radio interferometric imaging and cosmic magnetic field statistics.

Throughout this study, different data analysis techniques are presented and employed in various settings, but all can be summarized under the broad term of *statistical inference*. This subject encompasses a huge variety of statistical techniques, developed to solve problems in which deductions have to be made from incomplete knowledge, data or measurements (Mood et al. 1974; Jaynes 2003). This study focuses especially on Bayesian inference methods that make use of a subjective definition of probabilities, allowing for the expression of probabilities and statistical knowledge *prior* to an actual measurement (for an introduction see Jaynes 2003, and Ch. 2).

The thesis contains two different sets of application for such techniques. First, situations where a complicated, and generally ill-posed measurement problem can be approached by assuming a statistical signal model prior to infer the desired measured variable. Such a problem very often is met should the measurement device take less data than needed to constrain all degrees of freedom of the problem. The principal case investigated in this thesis is the measurement problem of a radio interferometer, which takes incomplete samples of the Fourier transformed intensity of the radio emission in the sky, such that it is impossible to exactly recover the signal. The new imaging algorithm RESOLVE is presented, optimal for extended radio sources. A first showcase demonstrates the performance of the

new technique on real data. Further, a new Bayesian approach to multi-frequency radio interferometric imaging is presented and integrated into RESOLVE.

The second field of application are astrophysical problems, in which the inherent stochastic nature of a physical process demands a description, where properties of physical quantities can only be statistically estimated. Astrophysical plasmas for instance are very often in a turbulent state, and thus governed by statistical hydrodynamical laws (see e.g. Ostriker 2006). Two studies are presented that show how properties of turbulent plasma magnetic fields can be inferred from radio observations.

Chapter 1

Introduction

This introductory chapter covers the two main scientific themes of this thesis, *statistical inference* and *radio astronomy*. Both are introduced in a brief manner since significant derivations and remarks additionally appear in subsequent chapters of the thesis, as they are part of published articles.

1.1 Statistical inference

Science relies on experiment, where a measurement device takes data, from which conclusions may be drawn on underlying processes. In many cases, this measurement process cannot be perfect. The measurement device itself might be flawed or imperfect and thus can introduce errors or limitations, the process could be inherently subject to random noise or even of random nature itself, or unwanted events from other sources than the scientific target might be accidentally recorded.

In general, if these measurement imperfections are unknown to the experimentalist (or the observer, in astronomy) or introduce inevitable loss of information, the conclusions that can be drawn from such data *cannot be exact*. Instead, drawing conclusions from incomplete knowledge can often only be of *statistical nature*. In statistics, the technique of searching a probabilistic solution to an incomplete measurement problem is summarized in the subject of *statistical inference* (Mood et al. 1974; Jaynes 2003; Upton et al. 2008)

In astronomy, the measurement process usually includes a telescope observing a celestial object, which very often poses a challenging measurement problem. Telescopes are complicated devices, for not all of their technical characteristics are generally known exactly¹ and their electronics always introduces random measurement noise. Furthermore, a specific property of astronomical observations is that they usually cannot be repeated under the exact same conditions. Thus, in astronomy robust data analysis is of special importance to guarantee the scientific value of conclusions drawn from observations.

¹Examples include unknown point spread functions and sky sensitivity patterns, or non-trivial elevation dependence of the performance of moving dishes.

In this work, an instrumental measurement problem can be sufficiently described by the equation

$$d = R(s) + n \quad (1.1)$$

where d are the measured data of the original physical signal s , which have been accumulated by a process that incorporates an instrumental response R and a random noise contribution n . An application of this type of measurement problem to an observation with a radio interferometer, for which the form of R inherently leads to incomplete data, can be found in Chs. 2, 3 and 4. Conversely, Ch. 5 introduces a type of inference problem, where the instrumental measurement problem is considered to be solved and the results are used to estimate further physical quantities, in this case to constrain magnetic field properties from radio polarization images.

1.1.1 Bayesian inference

In statistics, two different approaches to inference problems exist. The *frequentist* school only allows for probabilities strictly defined as measured frequencies and thus draws conclusions in considering the statistical results of the repeated outcomes to measurements (see e.g. Hogg and Tanis 2006). In contrast, this thesis utilizes a *Bayesian* approach to inference, where probabilities are rigorously defined as subjective degrees of belief, which allows the expression of probabilistic knowledge *prior* to an actual measurement, and the meaningful description of statistics from not-repeated measurements (for an exhaustive introduction see Cox 1990; Jaynes 2003). From now on, statistical derivations in this thesis will always use elements from Bayesian reasoning, even when not explicitly stated.

The cornerstone of Bayesian statistics is Bayes' theorem

$$\mathcal{P}(s|d) = \frac{\mathcal{P}(d|s)\mathcal{P}(s)}{\mathcal{P}(d)} \quad (1.2)$$

that connects the *likelihood* distribution $\mathcal{P}(d|s)$, the *prior* distribution $\mathcal{P}(s)$ and the *evidence* $\mathcal{P}(d)$ distribution to the *posterior* distribution $\mathcal{P}(s|d)$. In case of a measurement problem like (1.1), where d and s are defined as data and the physical signal, the posterior summarizes all important knowledge on the problem. The likelihood describes the measurement process, i.e. the probability that a specific data realization has been taken given a signal s , while the prior encapsulates all pre-measurement knowledge on the signal². For a more thorough introduction to these terms, see Ch. 2 and references therein.

Inferences on the signal can be drawn by calculating an estimator from the posterior. Typical approaches evaluate a statistic on the posterior, for instance its mean, or its mode (called *Maximum a Posteriori* which is used frequently and introduced further in Ch. 2).

Very often, the posterior cannot be calculated analytically. This is especially true, should the measurement problem (1.1) involve a non-linear response R , or more generally,

²The evidence does not depend on the signal and is thus unimportant for inferences thereof. For this thesis, it is handled as a normalization factor.

should the posterior include higher order statistics than a Gaussian distribution. This is the case for the radio imaging applications in Chs. 2, 3 and 4), where numerical methods are used to calculate a solution. A similar problem occurs in Ch. 5, where the assumption of Gaussian magnetic field statistics is only sufficient to solve the problem for restricted cases. In principle, these problems can be solved using perturbative methods similar to statistical or quantum field theory as shown by Enßlin et al. (2009), who have developed *information field theory* (IFT). Since these approaches are computationally rather complex, this thesis does not make explicit use of them. However, the mathematical framework used in the following chapters relies in many cases on the terminology and concepts introduced by Enßlin et al. (2009).

For many realistic applications, the prior knowledge on the signal is only minimal. In such a case, the derivations in this thesis rely on an approach of least information and invoke the *maximum entropy principle* of statistics to specify a prior distribution (see Caticha (2008) and App. A). Very often, this justifies the usage of a Gaussian prior with unknown covariance. This leads to a class of inference problems, where the unknown signal covariance needs to be inferred together with the signal itself. Such problems have been solved already for various problem settings, in particular using the framework of IFT (Enßlin and Frommert 2011; Enßlin and Weig 2010; Oppermann et al. 2011b, 2013). A similar solution is derived in this thesis for the radio measurement problem in Ch. 2.

1.2 Radio astronomy

Radio astronomy is the part of astronomy concerned with observing objects at wavelengths ranging from sub-millimeters to meters. At these wavelengths, observations from the surface of the earth are possible. At longer wavelengths, radio waves are scattered and absorbed by the ionosphere, at smaller wavelengths, water vapor from the troposphere imposes limits on observability. This makes the radio band the only wavelength regime besides the optical in which earthbound telescopes are feasible for electromagnetic radiation³.

1.2.1 Instruments

Radio telescopes are simply receiving instruments for radio waves, either in the form of antennas or dishes depending on the wavelength regime that the instrument is constructed for (Thompson et al. 1986; Wilson et al. 2012). Antennas are much cheaper, but can be used efficiently only at longest wavelengths, so most radio telescopes are found to be dishes.

The diameter of a dish determines the maximum collecting area of a radio telescope and therefore its sensitivity and signal-to-noise characteristics. But it also limits the maximum resolution at which objects, observed with only a single dish, can still be resolved. This has led to a second type of radio telescope, interferometers, where an array of single

³Earthbound cosmic ray experiments and neutrino telescopes are possible, which both are observing other particles than photons.

telescopes is connected together to significantly improve the maximum resolution (Ryle and Hewish 1960). A radio interferometer typically has inferior brightness sensitivity than a single dish, mostly because the individual dishes in an array are usually smaller than for a single dish telescope. It is also less sensitive to the sky brightness structures on largest angular scales. So, in reality, there is always a trade-off between the different instruments, and for many scientific applications data from both single dishes and interferometers are effectively combined. For an exhaustive treatment and technical details of both classes, refer to Thompson et al. (1986), Taylor et al. (1999) and Wilson et al. (2012).

At each frequency, the outcome of an observation with a radio telescope is given in units of flux density S_ν , which are $[Wm^{-2}Hz^{-1}]$, where S_ν is defined as the integral of the surface (sky) brightness or intensity I_ν , in units of $[Wm^{-2}Hz^{-1}Sr^{-1}]$, of the observed source over a solid angle Ω on the sky

$$S_\nu = \int I_\nu d\Omega. \quad (1.3)$$

A single dish telescope directly observes S_ν , since the received power is related to S_ν via antenna gain terms and a telescope beam convolution that both can be determined in principal (Wilson et al. 2012).

In contrast, for a radio interferometer it can be shown that the telescope array effectively samples an incomplete portion of the Fourier transformed sky brightness I_ν (Thompson et al. 1986). This means that the data cannot be used to draw absolutely exact conclusions on the sky brightness, and some sort of imaging algorithm is needed. In the Chs. 2, 3 and 4, a new approach to interferometric imaging of extended radio sources is developed using techniques of statistical inference.

Important single dish instruments are the Effelsberg telescope in Germany, the Green Bank Telescope (GBT) and the Arecibo telescope in the US, and recently the Sardinia Radio Telescope (SRT) in Italy. Known operational radio interferometer facilities include the Very Large Array (VLA) and the Very Long Baseline Array (VLBA) in the US, the Giant Metrewave Radio Telescope (GMRT) in India, the Westerbork Synthesis Radio Telescope (WSRT) in the Netherlands, the Australian Telescope Compact Array (ATCA) in Australia, the Atacama Large Millimeter Array (ALMA) in Chile, and since recently the Low Frequency Array (LOFAR) in Europe. The latter is the first large-scale radio interferometer that operates at very long wavelengths, and it was thus sufficient to solely construct the telescope with simple antenna fields. Ch. 3 covers an application of the new imaging algorithm RESOLVE (see Ch. 2) to VLA data.

Recent years have seen an important development, since a number of new era radio telescopes become operational right now with advancing technology. Of them, LOFAR (examples for recent results are Wise and The LOFAR Surveys KSP Cluster Working Group 2014; Heald and LOFAR collaboration 2014), ALMA (ALMA press group 2012) and an upgraded version of the VLA (Perley et al. 2011) are the first being fully commissioned already. Right now in the planning phase is the Square Kilometer Array (SKA), an envisaged huge radio interferometer to be build in South Africa, Australia and New Zealand (e.g SKA press group 2012). Almost online are two precursor instruments that

serve as pathfinder missions for the SKA, namely the Australian Square Kilometre Array Pathfinder (ASKAP) in Australia (Schinckel et al. 2012) and MeerKAT in South Africa (Booth and Jonas 2012). For a deeper overview of future instruments in radio astronomy, see Garrett (2012).

These instruments pose a challenge for current data analysis methods, first because of data quantities, which are orders of magnitude larger than previously known, and second because of data quality, for which current algorithms possibly cannot match the full potential since currently acceptable models and approximations might break down with the new fidelity.

This development is an important incentive for the studies of this thesis, both because of the promises of future observation capabilities (see Ch. 5, where radio polarization studies are introduced that in part only become possible with these new developments) and because of its challenges (see Chs. 2, 3 and 4 which cover an extensive study on advancing radio imaging algorithms).

1.2.2 Radio astrophysics

With radio astronomical observations, many different celestial objects are investigated, connected to a diversity of astrophysical topics. Some of the most important findings in the history of astronomy and cosmology were made in radio astronomy, most notably the discoveries of quasars, binary pulsars and of the cosmic microwave background (CMB) (Wilson et al. 2012).

Typical scientific targets range through all scales encountered in the cosmos, including planetary and solar science with the Sun and Jupiter as notable radio sources (e.g. Burke 2006; Cook 2011), exoplanet search (e.g. Lecavelier Des Etangs et al. 2011), the interstellar medium of the Milky Way as a foreground for extragalactic or CMB observations or as its own target (e.g. Haverkorn and Spangler 2013), supernova remnants (SNR) (e.g. Green 2001) and pulsars (e.g. D’Amico 2005), radio galaxies of various types (e.g. Miley and De Breuck 2008), SagA*, the source believed to be the central supermassive black hole of the Milky Way (e.g. Lu et al. 2014), other active galactic nuclei (AGN) and Quasars (e.g. Hirabayashi 2004), galaxy clusters (e.g. Feretti et al. 2012), signals from the epoch of re-ionization (e.g. de Bruyn and LOFAR EoR Key Science Project Team 2012) and the CMB (e.g. Planck Collaboration et al. 2013).

A branch of radio astronomy searches for spectral lines, e.g. of planet atmospheres, stars, or the 21 cm line from HI regions or the epoch of re-ionization. Another analyzes continuous emission of thermal or non-thermal origin, like the CMB or synchrotron emission from environments such as SNR, radio galaxies and jets, AGN or clusters.

A wide range of scientific questions are addressed in these studies, from mapping the interstellar medium of the Milky Way and tracing the structure of spiral arms of other galaxies using HI studies, over understanding the complicated physical processes in various high energetic events like jets, accretion discs of AGN, cluster mergers, resolving distant and compact objects, to cosmological studies like the search of signals from the epoch of re-ionization or measuring basic cosmological parameters using the CMB.

Most important for this work are observations of radio synchrotron sources, which are generally very well suited for cosmic magnetic field studies⁴. Synchrotron emissivity directly depends on the magnetic field strength, and synchrotron polarization on the magnetic field orientation (see next section). This makes radio astronomy the prime method to study magnetic fields for many objects in the sky. Continuous synchrotron emission and (turbulent) magnetic fields are central topics of this thesis and will be covered in the next two sections.

1.2.3 Radio synchrotron observations

The radio astronomical data in the focus of this study all result from radio continuum observations of synchrotron sources. This section briefly provides the necessary background.

Synchrotron radiation is emitted by charged relativistic particles gyrating around magnetic field lines, mainly by electrons. Among other quantities, the totally emitted power of an accelerated charge depends on its Lorentz factor, and on the magnetic field energy density:

$$P \propto \gamma^2 \frac{B^2}{8\pi}. \quad (1.4)$$

Synchrotron radiation has a characteristic polarization with a high percentage of linear polarization, and a negligible contribution from circular one. Furthermore, due to relativistic beaming effects, most of the energy is radiated within a narrow cone around the direction of the moving charge.

Synchrotron radiation is most conveniently described using the Stokes parameters I , Q , U and V . I corresponds to the total intensity of radiation. In a polarized state, the field vectors of the electromagnetic wave trace in general an ellipse, and Q and U measure the orientation of this ellipse relative to an axis defined by one of the vectors. V is a circularity parameter that measures the ratio of the principal axes of the ellipse and the handedness of the polarization (V negative usually corresponds to right-handed polarization, and vice-versa). $Q = U = 0$ corresponds to a perfectly circular polarization state, while $V = 0$ is the condition for linear polarization.

A number of important relations can be derived⁵ for the Stokes parameters (see Rybicki and Lightman 1985). For exactly monochromatic radiation, the basic relation is

$$I^2 = Q^2 + U^2 + V^2, \quad (1.5)$$

which becomes an inequality in the case of non-monochromatic radiation

$$I^2 \geq Q^2 + U^2 + V^2. \quad (1.6)$$

⁴Other methods include the use of Zeeman-splitting of spectral lines or the polarization of starlight.

⁵This can be straight-forwardly done using basic relationships of the Stokes parameters with the electric (or magnetic) field vector components.

Since synchrotron emission is mostly linearly polarized ($V = 0$), this work is not further concerned with circular polarization, and V is therefore assumed to be zero for the remainder of the thesis. Then, the complex polarized intensity can be defined by

$$P = |P| \exp[2i\chi] = Q + iU \quad (1.7)$$

where χ is the polarization angle. In this way, a rotation of the polarization plane $\chi \rightarrow \chi + \chi'$ can be easily included by an extra exponential factor

$$P_{\text{rot}} = P \exp[2i\chi'] = |P| \exp[2i(\chi + \chi')]. \quad (1.8)$$

The magnitude $|P|$ is identified as

$$|P| = \sqrt{Q^2 + U^2}. \quad (1.9)$$

For synchrotron radiation, under a number of standard assumptions, the Stokes parameters can be directly related to the involved magnetic field in a simple way. The main assumption is an isotropic distribution and power-law spectrum of the relativistic electrons with a specific power index of 3. This value is in general not exactly met in nature, but known to be a good approximation. Typical values for the spectral index of relativistic electrons in our Galaxy, measured directly by cosmic rays reaching Earth or indirectly via their induced synchrotron emission are around $p \approx 2.7$ (Amsler et al. 2008).

The involved, but standard calculation for the synchrotron emissivity can be found e.g. in Rybicki and Lightman (1985) and leads to the compact notation:

$$I = \int dz (B_x^2 + B_y^2), \quad (1.10)$$

$$P = \int dz (B_x^2 - B_y^2 + 2iB_x B_y), \quad (1.11)$$

where z denotes the line-of-sight (LOS) of the observation, and complicated fore-factors, which mainly depend on physical constants, but are unnecessary for this work are suppressed for convenience. It should be noted that these observables are only sensitive to the magnetic field components *perpendicular* to the LOS.

A powerful tool for analyzing polarization for the field component *parallel* to the LOS is Faraday rotation. The effect stands for the rotation of the polarization plane of a linearly polarized wave in a medium with a non-scalar dielectric constant due to a magnetic field component along the direction of the traveling wave. In such an environment, the dielectric constant ϵ differs for left and right circular polarization (Rybicki and Lightman 1985)

$$\epsilon = 1 - \frac{\omega_p^2}{\omega(\omega \pm \omega_p)}, \quad (1.12)$$

where ω is the frequency of the transversing wave, and $\omega_p = eB/mc$ is the cyclotron frequency. The \pm signs discriminate between right and left circular polarization. If a linearly

polarized wave is described as a superposition of a wave with right circular polarization and a wave with left circular polarization, the polarization plane will not remain constant during the traverse of such a medium.

Including Faraday rotation, the total polarization angle χ at the location of the observer is given by

$$\chi(\mathbf{x}) = \chi_0(\mathbf{x}) + \lambda^2 \phi(\mathbf{x}). \quad (1.13)$$

Here χ_0 denotes the polarization angle at the origin of emission, and the Faraday depth $\phi(\mathbf{x})$ is defined as

$$\phi(\mathbf{x}) = a_0 \int dz n_e B_z, \quad (1.14)$$

where the fore-factor $a_0 = e^3/2\pi m_e^2 c^4$ will be suppressed for the remainder of this thesis, as well as the thermal electron density n_e along the LOS, which is assumed constant⁶ if not stated otherwise.

The Faraday depth $\phi(\mathbf{x})$ describes the phase angle through which the electric vector rotates due to Faraday rotation. If the rotating medium is *not* emitting itself, (1.13) can be used to directly fit for $\phi(\mathbf{x})$ (and the magnetic field) by using observations at different wavelengths. In this context, the inferred Faraday depth is referred to as *rotation measure* (RM)⁷. In a more complex situation, (1.13) becomes a superposition of many differently rotating processes and a more elaborate Fourier-technique called *Rotation Measure Synthesis* (RM-Synthesis) is needed which will be the subject of Sec. 5.1.

1.2.4 Turbulence, dynamo theory and Gaussian magnetic fields

The second part of this thesis (see Ch. 5) presents two inference problems to deduce properties of turbulent magnetic fields from radio synchrotron polarization measurements. This section introduces the necessary physical background to turbulent magnetic field statistics.

Cosmic magnetic fields are often embedded into physically active environments and dragged along with turbulent plasma flows. In an ideal plasma, because of the large number of free charge carriers, no electric field of significance can develop and the electric resistivity is very low. Conversely, this implies that magnetic fields must be abundant. Such a physical situation is well described by magneto-hydrodynamics (MHD), which describes the kinematics of a magnetized plasma flow through the mutual interaction of the magnetic fields and the plasma. Ideal MHD includes the effect of flux-freezing, where the magnetic field is totally confined within a turbulent plasma flux tube (see e.g. Ostriker 2006). This leads to a description of the magnetic field as a turbulent quantity on its own. Turbulent MHD usually directly transfers concepts from classical hydrodynamics to magnetic fields,

⁶This can be interpreted as using magnetic field strength units, which have absorbed these quantities.

⁷Correctly this term should exclusively describe the experimental factor between the rotation angle and λ^2 , obtained through an observational fit, *not* the physical quantity Faraday depth ϕ .

such as magnetic Reynolds numbers (Ostriker 2006) or Kolmogorov-turbulence spectra (Kolmogorov 1941). In single cases, the latter have indeed been observed for magnetic fields in turbulent environments like the inter-stellar medium (Armstrong et al. 1981) or in galaxy clusters (Kuchar and Enßlin 2011). Despite successes this is still an active and somewhat experimental field of research, for a review of the current state see Schekochihin and Cowley (2007).

Of particular importance for contemporary cosmic magnetic field research are helical magnetic fields. Magnetic helicity is defined as the integral over a volume of space of the scalar product of the magnetic field vector with the magnetic vector potential

$$H = \int_V \mathbf{A} \cdot \mathbf{B} \, dx^3. \quad (1.15)$$

It can be interpreted as a topological measure of how much the magnetic field spirals within some given region of space. In ideal MHD, magnetic helicity is a conserved quantity (Ostriker 2006; Schnack 2009). This property earns helicity a key role in modern magnetic field astrophysics.

There, a major topic is to understand the origin of magnetic fields in the universe. Two scenarios are discussed: either the fields are of primordial origin, or have evolved through some physical mechanism (Widrow 2002). Currently, it is assumed that dynamo processes could be responsible for the second scenario. In an astrophysical setting, a dynamo is a physical mechanism that transforms kinetic energy into magnetic energy. For magnetic fields on small scales and of random orientation, a scenario seems feasible in which the dynamo is fueled by the kinetic energy of a turbulent plasma (Brandenburg and Subramanian 2005). But, to generate ordered, large scale magnetic fields different processes are needed, and many dynamo scenarios are proposed, most of which belong to the framework of magnetic mean field theory (Brandenburg and Subramanian 2005). Theory and simulations indicate that such dynamos need a net transfer of magnetic helicity between small and large magnetic field scales in order to amplify large scale magnetic fields while conserving magnetic helicity. To prove and discern these theories, it is thus very important to test available data for helical magnetic field structures in turbulent environments as unambiguously as possible. The second part of Ch. 5 is dedicated to a study of how to use radio polarization data for a statistical measurement of magnetic helicity.

In general, the turbulent nature of a significant portion of cosmic magnetic field structures makes a statistical description inevitable. The exact statistics of turbulent magnetic fields can be very complicated and are, in fact, not very well known in general. But since all but the smallest scale magnetic fields are coherent over some spatial region, they can be expected to be spatially correlated. In this case, without any further detailed information on higher order statistics, the principle of maximum entropy (see App. A) enforces a choice of Gaussian statistics as the simplest statistical description of spatially correlated fields. Furthermore, the Gaussian statistical model is the mathematical basis for implementing more complicated statistics in form of higher order corrections, and thus embodies a natural starting point for any statistical estimator.

Thus, for this work, the magnetic field statistics is generally assumed Gaussian

$$\mathcal{G}(B, M) = \frac{1}{\sqrt{2\pi \det(M)}} \exp\left[-\frac{1}{2} B^+ M^{-1} B\right] \quad (1.16)$$

where B is the three-dimensional magnetic field vector, and M its covariance.

Moreover, the assumption of statistical homogeneity and isotropy, also used later in Ch. 2 for the development of RESOLVE, is very common in MHD magnetic field turbulence. With this assumption, the covariance $M_{ij}(\mathbf{x}, \mathbf{x}') = \langle B_i(\mathbf{x}) B_j(\mathbf{x}') \rangle$ of the Gaussian magnetic field can be expressed as the magnetic autocorrelation tensor (see e.g. Subramanian 1999)

$$M_{ij}(\mathbf{r}) = M_N(r) \delta_{ij} + (M_L(r) - M_N(r)) \frac{r_i r_j}{r^2} + M_H(r) \epsilon_{ijm} r_m, \quad (1.17)$$

or in Fourier space using the solenoidal condition $\nabla \cdot \mathbf{B} = 0$

$$\hat{M}_{ij}(\mathbf{k}) = \hat{M}_N(k) \left(\delta_{ij} - \frac{k_i k_j}{k^2} \right) - i \epsilon_{ijm} \hat{H}(k) \frac{k_m}{k}. \quad (1.18)$$

Here, r and k are defined as the magnitude of the full three dimensional vectors \mathbf{r} and \mathbf{k} .

In (1.18), the degrees of freedom of the the magnetic correlation has been reduced to a normal and a helical part of the spectrum, $\hat{M}_N(k)$ and $\hat{H}(k)$. Drawing inferences on the properties of the turbulent magnetic field often involves the statistical estimation of properties of these functions. Both methods presented later in this chapter use different observational relations to constrain properties of (1.18). For instance, it can be shown that $\hat{M}_N(k)$ and $\hat{H}(k)$ are closely related to the normal and helical 1D-energy densities $\epsilon_B(k)$ and $\epsilon_H(k)$ of the magnetic field

$$\epsilon_B(k) = \frac{k^2 \hat{M}_N(k)}{8\pi^3}, \quad \epsilon_H(k) = -\frac{k^3 \hat{H}(k)}{\pi^2}. \quad (1.19)$$

For this derivation and all further details on the magnetic autocorrelation tensor, especially in the context of the work presented in Ch. 5, the reader is referred to Junklewitz and Enßlin (2011)⁸

1.3 Outline

The first part of the thesis ranges from Ch. 2 to Ch. 4 and presents a study of the problem of imaging extended radio emission with a radio interferometer. In Ch. 2 the new imaging algorithm RESOLVE is presented and shown to outperform standard algorithms on simulated data. In addition, the chapter contains further introductions to Bayesian inference

⁸Only parts of this paper that were not reproduced in the author's diploma thesis are reported in this PhD thesis.

and interferometry in the context of radio astronomy. This chapter has been submitted as an independent publication to the journal *Astronomy & Astrophysics* (for the arXiv print see Junklewitz et al. 2013). The following chapter 3 presents first results from an application of RESOLVE to real data in form of an observation of the galaxy cluster Abell 2256 with the VLA in its D-configuration at 1369 GHz. Ch. 4 introduces a new Bayesian approach to multi-frequency radio interferometric imaging and also offers a brief introduction to classical multi-frequency imaging techniques, where not only the total intensity, but also the spectral index is imaged. Furthermore, an extended, multi-frequency version of RESOLVE is presented together with a set of simulated tests showing that the new approach is superior to classical multi-frequency imaging techniques within the testing parameters. This chapter has been submitted as an independent publication to the journal *Astronomy & Astrophysics* (for the arXiv print see Junklewitz et al. 2014).

The second part of the thesis comprises Ch. 5. It offers two applications of magnetic field statistics to different settings of radio polarization observations, and shows how statistical properties of the magnetic fields can be inferred from them. It also briefly introduces the astrophysics of turbulent magnetic fields and radio polarization observations. Sections of the chapter have been published in *Astronomy & Astrophysics* as part of accepted papers (Junklewitz and Enßlin 2011; Bell et al. 2011).

Chapter 2

RESOLVE: A new algorithm for aperture synthesis imaging of extended emission in radio astronomy¹

2.1 Introduction

Aperture synthesis techniques using large interferometers have a long and successful history in radio astronomy (Ryle and Hewish 1960; Thompson et al. 1986; Finley and Goss 2000). While enabling observers to achieve very high resolutions, data processing is considerably more complicated than with a single dish instrument. A radio interferometer effectively measures the Fourier transformation of the sky brightness (see e.g. Thompson et al. 1986). Unfortunately, inverting this relationship to achieve an estimate of the desired source brightness is a non-trivial task since an interferometer only samples a fraction of the Fourier plane, effectively convolving the true image brightness with an observation-dependent point-spread function. A crucial part in data reduction is therefore the *imaging*, i.e. estimating the sky brightness distribution from the observed data.

To date the most successful and widely used imaging algorithm in radio astronomy is CLEAN (Högbom 1974). It assumes the image to be comprised of uncorrelated point sources and iteratively approximates the true image with a large set of delta functions. CLEAN has been demonstrated to be very accurate for observations of point source dominated fields (Thompson et al. 1986; Taylor et al. 1999; Sault and Oosterloo 2007) and over time many variants and more elaborate extensions have been developed to improve various aspects of its performance (Clark 1980; Schwab 1984; Cornwell 2008; Sault and Wieringa 1994; Rau and Cornwell 2011).

However, there are drawbacks with the CLEAN algorithm. Since it effectively assumes the image to be a large superposition of point sources, its performance is naturally non-

¹Note: This chapter has been submitted as a paper for publishing in *Astronomy & Astrophysics*.

optimal for highly resolved, extended and diffuse sources (Sault and Oosterloo 2007). Some of the newest enhancements of CLEAN try to address this problem using a multiscale approach, assuming differently scaled kernel functions like Gaussians instead of sharp delta peaks (Cornwell 2008; Rau and Cornwell 2011), but it is still not clear how to properly choose the scales. Another important drawback of CLEAN is that it is not known how to appropriately propagate measurement uncertainty (e.g. Thompson et al. 1986; Taylor et al. 1999) and thus, no reliable uncertainty estimates are available.

There are other approaches than CLEAN that try to address the problem of imaging extended sources. Among them are the Maximum Entropy Method (MEM) (Cornwell and Evans 1985), the non-negative-least-squares (NNLS) approach, which has been shown to improve over CLEAN on mildly extended sources (Briggs 1995a; Sault and Oosterloo 2007), and, approaches using wavelets within the framework of *Compressed Sensing* (Wiaux et al. 2009; Carrillo et al. 2012, 2013). We will come back to these in Sec. 2.6.2.

In this paper, we introduce RESOLVE (**R**adio **E**xtended **S**ources **L**ognormal deconvolution **E**stimator), a novel algorithm for the imaging of diffuse and extended radio sources in total intensity. A new approach to the problem is taken, using Bayesian statistics in the framework of *Information Field Theory* (Enßlin et al. 2009) and based on clearly formulated mathematical principles. RESOLVE is designed to fulfill two main requirements:

1. It should be optimal for extended and diffuse radio sources.
2. It should include reliable uncertainty propagation and provide an error estimate together with an image reconstruction.

An important incentive for the development of RESOLVE are recent advances in radio astronomical instrumentation. The new generation of radio telescopes, such as the upgraded VLA, LOFAR, the SKA pathfinder missions or ultimately the SKA itself, are opening new horizons in radio astronomy (see e.g. Garrett 2012). Their unprecedented capabilities of simultaneous, broadband frequency coverage including previously unexplored wavelength regimes, sensitivity, and wide fields of view, while still being sensitive to a large range of spatial frequencies, will almost certainly advance astrophysical and cosmological sciences (see e.g. the German SKA white paper, Aharonian et al. 2013). At the same time, new developments in signal processing and data analysis will be required to exploit these new capabilities. For instance, as yet unreached levels of sensitivity allow in principle for more detailed detection of structures in diffuse emission. RESOLVE takes advantage of this and uses the rich correlation structure prominently present in such high sensitivity data to guide itself toward an optimal reconstruction of extended sources.

The main astrophysical focus of RESOLVE is by definition on extended and diffuse radio sources. Among those are galaxy clusters with their weak diffuse halos and strong extended relic structures, lobes of radio galaxies, giant radio galaxies, supernova remnants, galactic radio halos, and the radio emission from the Milky Way.

Ultimately, with this paper, we do not only aim to present a new algorithm but we also propose and discuss a statistical framework (see Sec. 2.2) that, we believe, will be

advantageous to formulate and solve upcoming and more complex imaging problems in radio data analysis. Among these could be for instance multi-frequency techniques for GHz - broadband data, direction-dependent calibration problems, unknown beam reconstructions, polarization imaging, and many more. We will come back to an outlook in Sec. 2.7.

2.2 The algorithm

2.2.1 Aperture Synthesis

In aperture synthesis, we try to connect an array of telescopes in such a way that we can effectively synthesize a combined instrument with a much larger aperture and therefore resolution. Using the van Zittert-Cernike theorem from the theory of optical coherence (Born and Wolf 1999), it can be shown that such a radio interferometer takes incomplete samples of the Fourier transformed brightness distribution in the sky (Thompson et al. 1986). We would like to measure a signal, the sky brightness distribution I , which is a real, continuous function of position in the sky. In the most basic model, taking an observation of the signal I translates into

$$V(u, v, w) = W(u, v, w) \int dl dm \frac{I(l, m)}{\sqrt{1 - l^2 - m^2}} e^{-2\pi i(ul + vm + w\sqrt{1 - l^2 - m^2})}. \quad (2.1)$$

The quantity $V(u, v, w)$ is the *visibility* function following classical terminology of optical interferometry. The coordinates u , v , and w are vector components describing the distance between a pair of antennas in an interferometric array, where this distance is usually referred to as a *baseline*. They are given in numbers of wavelengths, with u and v usually parallel to geographic east-west and north-south, respectively, and w pointing in the direction of the center of the image plane (i.e. the phase center). The coordinates l and m are a measure of the angular distance from the phase center along axes parallel to u and v , respectively. $W(u, v, w)$ is a sampling function defined by the layout of the interferometric array. It is zero throughout most of the u, v, w -space, apart from where measurements have been made where it is taken to be unity.

For simplicity, we now restrict ourselves to the common approximation of measuring the sky as flat in a plane tangent to the phase center of the observation, such that $w\sqrt{1 - l^2 - m^2} \approx 0$. Nevertheless, we note that this is not a necessary requirement of our formalism (see Sec. 2.2.2).

With this assumption, (2.1) simplifies approximatively to a two-dimensional Fourier transformation

$$V(u, v) \approx W(u, v) \int dl dm I(l, m) e^{-2\pi i(ul + vm)}. \quad (2.2)$$

The visibility function is what our instrument measures, but we are actually interested in the brightness distribution of the source in the sky. This means that we ideally want to invert the relationship (2.2). Unfortunately, this is not possible, since we have lost all information on the Fourier modes that have not been measured due to the incomplete sampling of the Fourier plane. Thus, an inversion of (2.2) gives us not the true brightness distribution, but its convolution with the inverse Fourier transform of the sampling function, better known as the *point spread function* (psf) or, in common radio astronomical terminology, the *dirty beam* $I_{db} = \mathcal{F}^{-1}W$:

$$I_D = \mathcal{F}^{-1}V = \mathcal{F}^{-1}W\mathcal{F}I = I_{db} * I. \quad (2.3)$$

Here, we have introduced a symbolic Fourier operator \mathcal{F} to be strictly defined later, the common notation I_D , *dirty image*, for the simple Fourier inversion of the visibilities, and the symbol $*$ to denote a convolution operation.

Reconstructing the real brightness distribution is therefore an ill-posed inverse problem. In principle, infinitely many signal realizations could have led to the measured visibility function and we have no way to *exactly* discriminate between them. However, we can find a statistical description that may produce the *most probable* signal given the measured visibility function.

2.2.2 Signal Inference in Radio Astronomy

In the following, we develop a statistical solution to the inverse problem (2.2) using Bayesian inference techniques. Later, under the condition of a spatially extended source brightness distribution, this will lead us to the formulation of RESOLVE. Our derivation relies on notation and methods developed within the framework of *information field theory* (Enßlin et al. 2009; Enßlin 2013).

To start, we condense our mathematical notation considerably by rewriting all equations using indexed quantities. If we properly define the Fourier operator in (2.3) as $\mathcal{F}_{kx} = \exp(-i(ul + vm))$ with $x = (l, m)$ and $k = (u, v)$, (2.2) becomes

$$\begin{aligned} V_k &= W_k \int dx \mathcal{F}_{kx} I_x \\ &= W\mathcal{F}I. \end{aligned} \quad (2.4)$$

We have now translated our functions and operations on them into a notation that allows us to interpret them as vectors and operators defined on an arbitrary-dimensional functional vector space. For the sake of brevity, we will often even drop the indices and use a notation as in the second line of (2.4). We can do that if we define the inner product between vectors and operators appropriately for discrete and continuous spaces:

$$\begin{aligned}
\text{discrete space :} & & a^\dagger b & := \sum_x \mathcal{V}_x \overline{a_x} b_x \\
\text{continuous space :} & & a^\dagger b & := \int dx \overline{a(x)} b(x) dx
\end{aligned} \tag{2.5}$$

where the \dagger symbol stands for a transposing operation (and a possible complex conjugation in case of a complex field). In contrast, where needed explicitly, the \cdot symbol will denote component-wise multiplication, so that $(a \cdot b)_x = a(x) b(x)$. The symbol \mathcal{V}_x indicates the possible need for a volume factor in the sum, if the inner product actually is just a discretized version of a continuous one. In practice, this is unavoidable, since all quantities effectively become discrete when finally calculated on a computer (for details see Selig et al. 2013).

That way, we now can effortlessly combine discrete and continuous quantities in our notation. This is important, since, in real observations, the visibility V_k is always a function defined over a discrete, complex Fourier space, spanned by n_d measurements, whereas the sky brightness I_x is in principle a continuous function, defined over an infinitely large, real space.

Following the notation of Enßlin et al. (2009), we define two fundamental quantities, the signal s and the data d . The signal is the ideal, true physical quantity we would like to investigate with our observation. The data is what our measurement device has delivered us. In this radio astronomical application, the signal is the true brightness distribution in the sky $s := I(l, m)$ and the data is our visibility function $d := V(u, v)$ including measurement noise. From now on, we will use this definition, but will occasionally translate equations into traditional radio astronomical notation for a more transparent presentation.

If we know how to translate the actions of our measurement device into mathematical operations, we can write down a fundamental data model, connecting signal s and data d with a response operator R

$$d = Rs. \tag{2.6}$$

ignoring measurement noise for a moment.

This is basically equation (2.4), if we identify the response operator with

$$R = W\mathcal{F}, \tag{2.7}$$

We can add more terms to this response operator, slowly introducing more complexity. An inevitable addition is to consider a gridding and degriding operation within the sampling $W' = WG$. This is not a feature of the instrument itself, but is needed in its computational representation for purely numerical reasons to put the visibilities onto a regularly spaced grid, in order to apply the Fast Fourier Transform algorithm (Cooley and Tukey 1965; Bracewell 1965), improving computational speed enormously:

$$R = W'\mathcal{F} \tag{2.8}$$

Henceforth, if not explicitly shown, we drop the prime and consider G to be contained in the sampling operator W .

An important extension might be to introduce a mathematical representation of the antenna sensitivity pattern on the sky, usually called primary beam A :

$$R = WFA. \quad (2.9)$$

Even more sophisticated instrumental effects like beam smearing or directional dependent sampling could as well be included here. Also an extension of the response to non-coplanar baselines, and thus allowing for a non-negligible w - term in Eq. (2.1), could be directly incorporated without fundamental complication, e.g. in similar form to the w - projection algorithm (Cornwell et al. 2008).

Another relevant extension is to include multi-frequency synthesis by adding a new dimension to signal and data using e.g. a common spectral model $I(x, \nu) = I(x, \nu_0) \left(\frac{\nu}{\nu_0}\right)^{-\alpha(x)}$:

$$\begin{aligned} V_{k'} &= \int dx R_{kx} I_{x\nu} \\ &= W_k \int dx \mathcal{F}_{kx} A_x I_{x\nu_0} \left(\frac{\nu}{\nu_0}\right)^{-\alpha_x} \end{aligned} \quad (2.10)$$

with $k' = k\nu$.

Going a step further, a full approach using all four Stokes polarizations is conceivable. In that case, the response representation can in principle be expanded into a full RIME (radio interferometer measurement equation) description, as presented e.g. by Smirnov (2011a,b).

However, both, multi-frequency and polarization imaging, are outside the scope of the present work.

In a real observation, our data is always corrupted by measurement noise. This means we have to add such a noise contribution n to our data model:

$$d = Rs + n. \quad (2.11)$$

As already noted, even without noise, we cannot exactly invert this relationship. We thus instead seek for the optimal statistical solution for the signal s given our data d . To find the optimal reconstruction, we regard the signal as a random field following certain statistics and being constrained by the data. In probabilistic terms, we look for an expression of the *posterior distribution* $\mathcal{P}(s|d)$ of the signal s given the data d . It expresses how the data constrain the space of possible signal realizations by quantifying probabilities for each of them. It comprises all the information we might have obtained through a measurement.

With the posterior probability, we can in principle estimate the real signal by calculating for instance its posterior mean $\langle s \rangle_{\mathcal{P}(s|d)}$, equivalent to minimizing the posterior-averaged \mathcal{L}_2 - norm of the quadratic reconstruction error $\operatorname{argmin}_m \langle \| (s - m) \|_{\mathcal{L}_2} \rangle_{\mathcal{P}(s|d)}$ (see e.g. Enßlin

et al. 2009). This is exactly the type of solution to the ill-posed inverse problem (2.11) that we want.

Probability theory shows that we can calculate $\mathcal{P}(s|d)$ if we have expressions for the *likelihood distribution* $\mathcal{P}(d|s)$, describing our model of the measurement process and the noise statistics, and for the statistics of the signal alone, the *prior distribution* $\mathcal{P}(s)$. The renowned Bayes' theorem states this as

$$\mathcal{P}(s|d) = \frac{\mathcal{P}(d|s)\mathcal{P}(s)}{\mathcal{P}(d)} \quad (2.12)$$

where $\mathcal{P}(d)$ is called the *evidence distribution*. It effectively acts as a normalization factor since it does not depend on s and thus is unimportant for statistical inferences on the signal.

To specify the likelihood for a radio interferometer observation, we only need a good model for the measurement process. With (2.11), we see that this involves detailed knowledge of the instrument response R and the statistical properties of the measurement noise n .

Throughout this work we will assume the response representation (2.9) to be exact, or expressed differently, the data to be fully calibrated. On the perspective of combining calibration and imaging into one inference step see Sec. 2.7.

As for the thermal noise of a radio interferometer, it is fair to assume Gaussian statistics, mainly induced by the antenna electronics and independent between measurements at different time steps of the observation (Thompson et al. 1986). Henceforth, the noise field n will be assumed to be drawn from a multivariate, zero mean Gaussian distribution of dimension n_d :

$$\begin{aligned} \mathcal{P}(n) &= \mathcal{G}(n, N) \\ &:= \frac{1}{\det(2\pi N)^{1/2}} \exp\left(-\frac{1}{2}n^\dagger N^{-1}n\right). \end{aligned} \quad (2.13)$$

The assumption of uncorrelated Gaussian noise leads to a diagonal covariance matrix $N_{kk'} = \delta_{kk'}\sigma_k^2$. For this work, we will assume the noise variance σ_k^2 to be known.

We can now derive an expression for the likelihood by marginalizing over the noise field:

$$\begin{aligned} \mathcal{P}(d|s) &= \int \mathcal{D}n \mathcal{P}(d|s, n) \mathcal{P}(n) \\ &= \int \mathcal{D}n \mathcal{P}(d|s, n) \mathcal{G}(n, N) \\ &= \int \mathcal{D}n \delta(n - (d - Rs)) \mathcal{G}(n, N) \end{aligned} \quad (2.14)$$

$$= \mathcal{G}(d - Rs, N), \quad (2.15)$$

where the integral is meant to be taken over the infinite space of all possible noise realizations. By inserting the delta function in (2.14) we have stated the implicit assumption that our response (2.9) is exact.

We are left with the crucial question of how to statistically represent our signal. Until now, the derivation was kept general and we effectively formulated an inference framework for aperture synthesis imaging. Now, we need to specify a prior $\mathcal{P}(s)$, depending on the type of signal field to which the statistical estimation should be optimal.

In the next section, we present a solution to the inference problem with a signal prior chosen to represent the properties of extended and diffuse emission.

2.2.3 RESOLVE: Radio Extended Sources Lognormal Deconvolution Estimator

To specify the prior distribution, we choose to follow an approach of least information. The question is: What is the most fundamental, minimal state of knowledge we have about the signal, prior to the measurement and without introducing any specific biases?

In this work, we want to focus on diffuse and extended sources in total intensity. Stating this alone enables us to give a few central assumptions we want to be reflected in the prior distribution:

1. An extended source exhibits a certain, *a priori* translationally and rotationally invariant (but usually unknown) spatial correlation structure.
2. The signal field must be strictly positive, since it should represent a physical intensity.
3. Typically, signal fields in radio astronomy show high variation in structures across the observed field of view, with a few strong components surrounded by weak extended structure, going over to large regions basically dominated by noise, usually spanning many orders of magnitude in intensity.

Apart from these statements, we assume that we know nothing more specific about our signal, and the prior should be chosen accordingly. For instance, we do not want to include specific source shapes or intensity profiles.

The assumption of translational and rotational invariance is very common and useful in signal inference, where it translates into homogeneity and isotropy of the prior statistics. Given our just stated, restricted prior assumptions, there is no reason, in general, to assume *a priori* that the correlation of the signal should change under spatial translation or rotation². We thus keep this assumption as valid throughout this paper.

The first constraint (1.) urges us to consider how to include the fact that the signal exhibits a spatial correlation of unknown structure. First we might argue just to use an

²It should be emphasized that this *a priori* assumption is not in contradiction with an *a posteriori* solution not exhibiting homogeneity and isotropy. Ultimately, if the combination of data and measurement noise allow for a specific source shape, the likelihood will dominate the prior and drive the reconstruction in this direction.

uninformative prior, not favoring any particular configuration. But, in fact, we do know something, namely that there *is* a spatial correlation, although its exact structure is obscure to us. Thus, we search for the statistics of a random field about whose correlation we know the least possible, i. e. only the two-point correlation function, equivalent to the second moment of the statistics. Now, the maximum entropy principle of statistics (see App. A) states that if we search for such a probability distribution, it must be Gaussian. Of course, a priori, we might even have no information about the two-point correlation. Nevertheless, the data itself yields such information, which we can extract during the inference procedure.

For the problem of reconstructing a Gaussian signal field with unknown covariance, an optimal solution to the inference problem (2.11) can actually be found analytically or at least approximatively in calculating the posterior mean $\langle s \rangle_{\mathcal{P}(s|d)}$ of the signal. A number of methods have been derived to do this, e.g. the critical filter and variants thereof (Enßlin and Weig 2010; Enßlin and Frommert 2011; Oppermann et al. 2011b, 2013) or approaches using the method of Gibbs sampling (Jasche et al. 2010; Sutter et al. 2012; Karakci et al. 2013).

Unfortunately, if we consider the second (2.) and third (3.) constraints from above more closely, we must come to the conclusion that Gaussian signal fields are inappropriate for our problem since they are neither positive definite nor strongly fluctuative over orders of magnitude in strength.

We consider instead that the logarithm of our signal field is Gaussian. If s is a Gaussian field, $I = e^s$ exhibits all the desired properties (1-3). It is known as a log-normal field. If we adapt the data model (2.11)

$$d = RI + n = RI_0 e^s + n \quad (2.16)$$

we are now faced with a considerably more complicated, non-linear problem. The factor I_0 can be set to account for the right units, w.l.o.g., we set it to one for the rest of this work.

The likelihood $\mathcal{P}(d|s)$ and the signal prior $\mathcal{P}(s)$ take the following form

$$\begin{aligned} \mathcal{P}(d|s) &= \mathcal{G}(d - Re^s, N) \\ &= \frac{1}{\det(2\pi N)^{1/2}} e^{-\frac{1}{2} (d - Re^s)^\dagger N^{-1} (d - Re^s)}, \end{aligned} \quad (2.17)$$

$$\begin{aligned} \mathcal{P}(s) &= \mathcal{G}(s, S) \\ &= \frac{1}{\det(2\pi S)^{1/2}} e^{-\frac{1}{2} s^\dagger S^{-1} s}. \end{aligned} \quad (2.18)$$

Then, the posterior of s

$$\mathcal{P}(s|d) \propto \mathcal{G}(d - Re^s, N) \mathcal{G}(s, S) \quad (2.19)$$

possibly becomes highly non-Gaussian due to the non-linearity introduced by (2.16).

2. RESOLVE: A new algorithm for aperture synthesis imaging of extended emission in radio astronomy

22

Indeed, the resulting problem cannot be solved analytically. A possible approach would be to separate the quadratic and higher terms in (2.19)

$$\mathcal{P}(s|d) \propto e^{-1/2 s^\dagger (S^{-1}+M)s + s^\dagger j + \sum_{n=3}^{\infty} \Lambda_{x_1 \dots x_n}^n s_{x_1} \dots s_{x_n}} \quad (2.20)$$

where Λ^n is a rank $-n$ tensor, and

$$j = R^\dagger N^{-1} d \quad (2.21)$$

$$M = R^\dagger N^{-1} R. \quad (2.22)$$

The higher order terms could be handled either by invoking perturbative methods as known in statistical or quantum field theory (Huang 1963; Peskin and Schroeder 1995), and already further developed for statistical inference (e.g. Enßlin et al. 2009), or by using a Monte Carlo Gibbs sampling method (Hastings 1970; Geman and Geman 1984; Neal 1993). Since these methods are computationally very expensive for this log-normal ansatz and the high dimensionality of the problem, we do not follow them any further in this work.

Instead, we seek an approximate solution in the signal field that maximizes the posterior:

$$\langle s \rangle_{\mathcal{P}(s|d)} \approx \operatorname{argmax}_s \mathcal{P}(s|d). \quad (2.23)$$

This method is known as *Maximum a posteriori* (MAP) in statistical inference³. For the present problem it leads to a non-linear optimization problem of a gradient equation for the posterior. With this approach, it is further possible to calculate a consistent uncertainty estimate. In principle, the uncertainty of a signal reconstruction can be estimated by the width of the posterior. In this case, we use the inverse curvature of the posterior at its maximum to approximate the relative uncertainty D (see App. 2.3 for details).

In this context, we still need to specify how to deal with the unknown correlation structure, i.e. the Gaussian signal covariance $S = \langle s s^\dagger \rangle$. As mentioned earlier, the problem of reconstructing a Gaussian random field with unknown covariance has been solved already (Jasche et al. 2010; Enßlin and Weig 2010; Enßlin and Frommert 2011; Oppermann et al. 2011b; Sutter et al. 2012), and even the respective problem for a log-normal random field has been partly solved before (Oppermann et al. 2013). Unfortunately, none of these methods can be readily applied to the inference problem at hand, since they require the signal response to have a diagonal representation in signal space. This is not necessarily fulfilled for the Fourier-response (2.9). We therefore develop a different approach, which nevertheless closely follows the previously mentioned works.

³The maximum a posteriori approach can also be interpreted as an approximation to the posterior mean $\langle s \rangle_{\mathcal{P}(s|d)}$, but is not guaranteed to yield a close result, especially not for highly non-Gaussian posterior shapes. Alternatively, it can be derived by minimizing an \mathcal{L}_∞ -norm error measure instead of the \mathcal{L}_2 minimization underlying the posterior mean approach.

Crucially, as explained above, our prior knowledge signal statistics is homogeneous and isotropic. This implies that the unknown signal covariance becomes diagonal in its conjugate Fourier space and can be expressed by its power spectrum $P_s(|k|)$ (see the Wiener-Kinchin theorem in Bracewell 1965)

$$S(k, k') = \langle s(k)s(k')^\dagger \rangle = (2\pi)^{n_s} \delta(k - k') P_s(|k|) \quad (2.24)$$

where $P_s(|k|)$ is just the Fourier transformation of the homogeneous and isotropic autocorrelation function $C(r) = S(|x - y|)$

$$P_s(|k|) = \int dr C(r) \exp(ikr). \quad (2.25)$$

We note that due to the assumption of isotropy, the power spectrum only depends on the length $|k|$ of the Fourier vector k . It is therefore sensitive to scales but not to full modes in Fourier space. Where the distinction is needed, we will make it explicit using the notation $|k|$.

We now parameterize the unknown covariance S as a decomposition into spectral parameters p_i and positive, disjoint projection operators $S^{(i)}$ onto a number of spectral bands such that the bands fill the complete Fourier domain

$$S = \sum_i p_i S^{(i)}. \quad (2.26)$$

These parameters can be introduced into the inference problem as a second set of fields to infer.

We therefore add a second MAP algorithm to the signal MAP, solving for these unknown parameters p_i . We then iterate between both solvers until convergence is achieved. The algorithm produces a signal estimate m , an approximation to the reconstruction uncertainty D , and a power spectrum estimate parameter set p_i . The equations to be solved iteratively are

$$S_p^{-1} m + e^m \cdot M e^m - j \cdot e^m = 0 \quad (2.27)$$

$$(D)_{xy} = S_p^{-1}{}_{xy} + e^{m_x} M_{xy} e^{m_y} + e^{m_y} \int dz M(x, z) e^{m(z)} - j_x \cdot e^{s_x} \delta_{xy} \quad (2.28)$$

$$p_i = \frac{q_i + \frac{1}{2} \text{tr} [(mm^\dagger + D)S^{(i)}]}{\alpha_i - 1 + \frac{\rho_i}{2} + (Tp)_i}. \quad (2.29)$$

A detailed derivation can be found in App. 2.3. The two quantities j and M are defined as above, q and α are parameters of a power spectrum parameter prior, ρ is a measure for the

number of degrees of freedom of each Fourier band, and T is an operator, which enforces a smooth solution of the power spectrum p_i . A thorough explanation of all these terms can be found in App. 2.3. Eq. (4.18) is the fix point equation that needs to be solved numerically to find a Maximum a Posteriori signal estimate m for the current iteration. The second equation (4.19) results from calculating the second derivative of the posterior for the signal estimate m , its inverse serves as an approximation to the signal uncertainty D at each iteration step. The last equation (4.20) represents an estimate for the signal power spectrum (and therefore its autocorrelation function), using the signal uncertainty D to correct for missing signal power in the current estimate m . The iteration is stopped after a suitable convergence criterion is met (see App. 2.4). The whole algorithm is visualized in a flow chart in Fig. 2.1.

It should be noted that solving these equations can be relatively time-consuming compared to e.g. MS-CLEAN, depending on the complexity of the problem at hand, since it involves a non-linear optimization scheme (4.18) and the numerical inversion and random probing of an implicitly defined matrix (4.19)⁴ (for details, see App. 2.4).

We call the combined algorithm RESOLVE (**R**adio **E**xtended **S**Ources **L**ognormal **d**econvolution **E**stimator).

Since the most severe problem in radio imaging is effectively how to extrapolate into unmeasured regions in uv -space to deconvolve the dirty beam from the dirty image (see Eq. 2.3), an explanation is in order of how RESOLVE achieves this deconvolution.

In the fix-point equation (4.18) to calculate the signal estimate m , the multiplicative term e^m acts as an effective convolution beam in Fourier space. Regions, where it has a significant value require only little modifications through the iterations in order to explain features in the data entering the equation via j . In contrast, regions where m is very negative require a drastic modification to capture data features. Therefore, sidelobe structures of the dirty image are more comfortably accounted for by restructuring the existing stronger emission regions than by enforcing weaker sidelobes structures in the final signal estimate. By concentrating the resolved structures into the strong emission regions, the lognormal model extrapolates information in uv -space. The multiplicative e^m term acts as a convolution kernel in Fourier space, enforcing some amount of smoothness in the visibility structures. This smoothness is exploited by RESOLVE for extrapolating the measured visibilities into the regions of uv -space without direct measurements. In this way, RESOLVE is also capable of achieving some degree of superresolution by extrapolating beyond the largest visibilities.

⁴The overall computational costs go roughly with $N_{\text{global}}N_{\text{pr}}O(\sqrt{n_s n_d})$ in the limit of a large number of visibility measurements n_d . The n_s are the number of pixels in image space, N_{pr} is the number of used random probing vectors to estimate matrix traces, and N_{global} is the global number of iterations RESOLVE needs to converge (see App. 2.4)

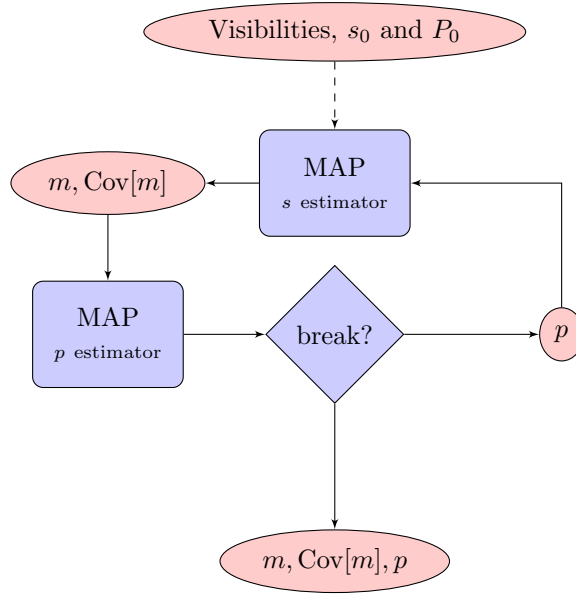


Figure 2.1: Flow chart, illustrating the basic workflow of the RESOLVE algorithm

2.3 Mathematical derivation of RESOLVE⁵

For a complete derivation of RESOLVE, we first give some general remarks, and then divide the rest into two parts, where we derive a Maximum a Posteriori solution for the signal field and for its power spectrum, respectively.

From Sec. 2.2, we recall the basic premises of the inference problem to be solved. We want to find the statistically optimal reconstruction of the total intensity signal I given a data model

$$d = RI + n = Re^s + n, \quad (2.30)$$

under the assumptions that

- I follows log-normal statistics, such that $s = \log I$ follows Gaussian statistics,
- the noise n follows Gaussian statistics as well,
- and R models the linear response of a radio interferometer (see Eq. (2.9) in Sec. 2.2).

Under these assumptions the likelihood $\mathcal{P}(d|s)$ and the signal prior $\mathcal{P}(s)$ take the following form as was shown in (2.15)

⁵Note: This section was originally part of the appendix of Junklewitz et al. (2013) and has been shifted into the main body of the thesis for clarity.

$$\begin{aligned}\mathcal{P}(d|s) &= \mathcal{G}(d - Re^s, N) \\ &= \frac{1}{\det(2\pi N)^{1/2}} e^{-1/2 \left((d - Re^s)^\dagger N^{-1} (d - Re^s) \right)},\end{aligned}\tag{2.31}$$

$$\begin{aligned}\mathcal{P}(s) &= \mathcal{G}(s, S) \\ &= \frac{1}{\det(2\pi S)^{1/2}} e^{-1/2 \left(s^\dagger S^{-1} s \right)}.\end{aligned}\tag{2.32}$$

Then, the posterior of s

$$\mathcal{P}(s|d) \propto \mathcal{G}(d - Re^s, N) \mathcal{G}(s, S)\tag{2.33}$$

can become highly non-Gaussian due to the non-linearity introduced by (2.30).

As a further complication, we have to assume *a priori* that the signal covariance $S = \langle ss^\dagger \rangle$ is unknown. Assuming statistical homogeneity and isotropy for the signal statistics, we parameterize its power spectrum $P(k)$ as a decomposition into spectral parameters p_i and positive projection operators $S^{(i)}$ onto a number of spectral bands such that the bands fill the complete Fourier domain

$$S = \sum_i p_i S^{(i)}.\tag{2.34}$$

RESOLVE consists of two inference steps to solve the main problem (2.16) iteratively for s and all p_i . We fully describe both steps individually in the following subsections.

2.3.1 Reconstruction of the signal field s

For the reconstruction of the signal field s , we assume the power spectrum parameters p_i to be known from a previous inference step. This can formally be expressed by marginalizing over them while assuming a delta distribution for the known parameters p^*

$$\begin{aligned}\mathcal{P}(s|d, p^*) &= \int \mathcal{D}p \mathcal{P}(s|d, p) \mathcal{P}(p|p^*) \\ &= \int \mathcal{D}p \mathcal{P}(s|d, p) \delta(p - p^*).\end{aligned}\tag{2.35}$$

For convenience, we rewrite our notation to work with the Hamiltonian $H(s, d)$ instead of the posterior $\mathcal{P}(s|d)$

$$\mathcal{P}(s|d) := \frac{e^{-H(d,s)}}{Z}\tag{2.36}$$

with $Z := \mathcal{P}(d)$. This effectively expresses our problem in more familiar terms of statistical physics, while the Hamiltonian $H(s, d) = -\log(P(d|s)P(s))$ still comprises all important signal-dependent terms and is usually easier to handle than the posterior.

The Hamiltonian of problem (2.33) reads

$$\begin{aligned} H(s, d) &= -\log(\mathcal{G}(d - Re^s, N) \mathcal{G}(s, S)) \\ &= \frac{1}{2} s^\dagger S_{p^*}^{-1} s + \frac{1}{2} (e^s)^\dagger M e^s - j^\dagger e^s + H_0 \end{aligned} \quad (2.37)$$

where $j = R^\dagger N^{-1} d$, $M = R^\dagger N^{-1} R$ and H_0 summarizes all terms which are not dependent on the signal s .

Using the Gibbs free energy ansatz of Enßlin and Weig (2010), Oppermann et al. (2013) have shown that it is possible to re-derive the critical filter for this Hamiltonian. However, in practice, it is only solvable under the assumption of a diagonal M in signal space. Otherwise we would be forced to explicitly compute arbitrary components of the very large matrix of size n_s^2 , representing the operator M , which is computational infeasible. Unfortunately, for the response under consideration here (2.9), with non-complete sampling of the Fourier plane in data space, M will not be diagonal in general.

Thus, we instead use the MAP principle to solve the inference problem for s . Maximizing the posterior readily translates to minimizing the Hamiltonian (2.36). If we take the derivative of the Hamiltonian (2.37) with respect to the signal field s and set it to zero, we get

$$\frac{\delta H(s)}{\delta s} = S_{p^*}^{-1} s + e^s \cdot M e^s - j \cdot e^s = 0. \quad (2.38)$$

This is a high dimensional, non-linear equation, which can be solved numerically using an iterative optimization algorithm, in our case a steepest descent method. We call the solution of this equation $m = \operatorname{argmax}_s \mathcal{P}(s|d)$.

The solution m is an estimate for the Gaussian field s . To calculate a signal estimate \hat{I} for the original log-normal signal $I = e^s$, we just take the exponential of m

$$\hat{I} = e^m. \quad (2.39)$$

2.3.2 Uncertainty of the signal reconstruction

A full statistical analysis involves accounting for the uncertainty of the signal estimate. For this, we use the information encoded in the second posterior moment (or covariance) $D = \langle (s - m)(s - m)^\dagger \rangle$ as a measure of the expected uncertainty of the signal reconstruction. Within the MAP approach, we approximate the inverse posterior covariance D^{-1} with the second derivative of the Hamiltonian

$$D^{-1} \approx -\left. \frac{\delta^2 H(s)}{\delta s_x \delta s_y} \right|_{s=m} = S_{p^*}^{-1}{}_{xy} + e^{s_x} M_{xy} e^{s_y} + e^{s_y} \int dz M_{xz} e^{s_z} - j_x \cdot e^{s_x} \delta_{xy}, \quad (2.40)$$

which needs to be inverted numerically in practice. In this way, we effectively assume that the real signal posterior is approximated with a Gaussian $\mathcal{G}(m, D)$. Unfortunately, D only

approximates the posterior covariance of the Gaussian field m . We need to translate this into a posterior covariance for the full estimate $\hat{I} = e^m$.

If the signal posterior *were* exactly Gaussian, we could just assume our posterior estimate to be of exact log-normal statistics, solve for the mean and variance analytically and thus write

$$\langle e^{s_x} \rangle_{\mathcal{G}(m,D)} = e^{m_x + \frac{1}{2}D_{xx}} \quad (2.41)$$

$$\langle (e^{s_x})^2 \rangle_{\mathcal{G}(m,D)} - \langle e^{s_x} \rangle_{\mathcal{G}(m,D)}^2 = e^{2m_x + D_{xx}} [e^{D_{xx}} - 1] \quad (2.42)$$

using the definitions for the mean and variance of a log-normal distribution (see e.g. Mood et al. 1974). But since the posterior is *not* Gaussian in general, we cannot solve Eqs. (2.41, 2.42) analytically. This was, in the first place, the reason why RESOLVE uses the MAP approach (see Sec. 2.2.3). Nevertheless, since we effectively approximate the full posterior with a Gaussian $\mathcal{G}(m, D)$ when using Eq. (2.40) as the posterior covariance, one might be tempted to just use Eqs.(2.41, 2.42) anyhow.

However, in practice, it turns out that within the MAP approach this procedure is prone to overestimating signal estimate and its uncertainty. This is because usually the maximum of a log-normal distribution lies *above* its mean (for details see Greiner 2013). We thus drop the extra terms of D in the argument of the exponentials in Eqs. (2.41, 2.42), keep (2.39), and write

$$\hat{I}_x = e^{m_x} \pm \sqrt{e^{2m_x} [e^{D_{xx}} - 1]} \quad (2.43)$$

if we want to account for the uncertainty in the reconstruction.

2.3.3 Reconstruction of the power spectrum parameters p

In the second step of RESOLVE, we assume to have a solution for m and D from the last iteration and estimate the unknown spectral parameters p from the signal-marginalized probability of data and power spectrum $\mathcal{P}(p, d)$:

$$\begin{aligned} \mathcal{P}(p, d) &= \int \mathcal{D}s \mathcal{P}(s, d|p) \mathcal{P}(p) \\ &= \int \mathcal{D}s \mathcal{G}(d - Re^s, N) \mathcal{G}(s, S_p) \mathcal{P}(p) \end{aligned} \quad (2.44)$$

This approach was first derived in Oppermann et al. (2013) for Gaussian signal fields. We closely follow their argument and show its approximate validity also for log-normal fields.

In order to do this, we first need to define a prior for the power spectrum parameters p . In this, we follow Enßlin and Frommert (2011), Enßlin and Weig (2010) and Oppermann et al. (2013), and choose independent inverse-gamma distributions for each spectral parameter p_i

$$\mathcal{P}(p) = \prod_i \mathcal{P}_{\text{IG}}(p_i) = \prod_i \frac{1}{q_i \Gamma(\alpha_i - 1)} \left(\frac{p_i}{q_i}\right)^{-\alpha_i} \exp\left(-\frac{q_i}{p_i}\right). \quad (2.45)$$

where $\Gamma(\cdot)$ denotes the gamma function, q_i defines an exponential cutoff in the prior for low values of p_i , and α_i is the slope of the power-law decay for large values of p_i . In principle, by tuning these parameters, the prior can be adapted according to the a priori knowledge about the power spectrum. Usually, we use the limits of $q_k \rightarrow 0$ and $\alpha_k \rightarrow 1$ for all k . This turns the inverse-gamma prior into Jeffreys prior (Jaynes 2003), which is flat on a logarithmic scale. In some tests though, we have allowed for non-unity α_k parameters to suppress unmeasured Fourier modes.

During the reconstruction of the power spectrum, we additionally introduce a smoothness prior as developed by Oppermann et al. (2013) to punish most probably unphysical and numerically unwanted random fluctuations in the power spectrum. In that prescription, the inverse-gamma prior (2.45) is augmented with a probability distribution that enforces smoothness of the power spectrum

$$\mathcal{P}(p) = \mathcal{P}_{\text{sm}}(p) \prod_k \mathcal{P}_{\text{IG}}(p_k). \quad (2.46)$$

The spectral smoothness prior can be written as a Gaussian distribution in $\tau = \log p$:

$$\begin{aligned} \mathcal{P}_{\text{sm}}(p) &\propto \exp\left(-\frac{1}{2\sigma_p^2} \int d(\log k) \left(\frac{\partial^2 \log p_k}{\partial (\log k)^2}\right)^2\right) \\ &\propto \exp\left(-\frac{1}{2} \tau^\dagger T \tau\right), \end{aligned} \quad (2.47)$$

where the differential operator T includes the second derivative of $\tau = \log p$ and a scaling constant σ_p^2 that determines how strict the smoothness should be enforced. This particular form of the prior favors smooth power-law spectra. For all details we refer to (Oppermann et al. 2013).

As was shown there, the corresponding inverse-gamma prior for the τ parameters can easily be derived from the conservation of probability under transformations

$$\mathcal{P}(\tau) = \mathcal{P}(p) \left| \frac{dp}{d\tau} \right| = \prod_i \frac{q_i^{\alpha_i - 1}}{\Gamma(\alpha_i - 1)} e^{-[(\alpha_i - 1)\tau_i + q_i e^{-\tau_i}]}. \quad (2.48)$$

With this prior, we can calculate the signal-marginalized joint probability (2.44) if we apply one crucial approximation. Since $\mathcal{P}(s, d|\tau)$ in (2.44) is non-Gaussian due to the high non-linearity of the $e^{(d - \text{Re}^s)}$ - terms, we cannot just move on analytically. We instead use a saddle point method and approximate the argument of the exponential occurring

2. RESOLVE: A new algorithm for aperture synthesis imaging of extended emission in radio astronomy

30

in $\mathcal{P}(s, d|\tau)$, which can be written as $e^{-H(s, d)}$ using (2.36). To perform the saddle point approximation, we replace $H(s, d)$ with its Taylor expansion up to second order around the maximum of the Posterior m , derived in the previous iteration of the signal reconstruction:

$$\begin{aligned} e^{-H(s, d)} &\propto e^{(-\frac{1}{2}(d - Re^s)^\dagger N^{-1}(d - Re^s) - \frac{1}{2}s^\dagger S_\tau^{-1}s)} \\ &\approx e^{(H(m) + \frac{1}{2}(s - m)^\dagger D(m)^{-1}(s - m))} \end{aligned} \quad (2.49)$$

This effectively approximates the non-Gaussian signal posterior $\mathcal{P}(s, d|\tau)$ with a Gaussian with mean m and covariance D . We note that this procedure is similar to a mean field approximation in statistical physics (Huang 1963).

With this approximation, we can solve the marginalization integral in (2.44) and calculate $\mathcal{P}(\tau, d)$, or alternatively the Hamiltonian

$$\begin{aligned} H(d, \tau) &= -\log \mathcal{P}(d, \tau) \\ &= -\log \int \mathcal{D}s \mathcal{G}(d - Re^s, N) \mathcal{G}(s, S) \mathcal{P}(p) \\ &\approx \frac{1}{2} \text{tr}(\log S_\tau) - \frac{1}{2} \text{tr}(\log D_\tau) + H(m, \tau) \\ &\quad + \sum_i ((\alpha_i - 1) \tau_i + q_i e^{-\tau_i}) + \frac{1}{2} \tau^\dagger T \tau \\ &\quad + H_0 \end{aligned} \quad (2.50)$$

where we have used the matrix theorem $\log |S| = \text{tr}(\log S)$, and have collected all terms not depending on τ into a constant H_0 .

Taking the derivative of (2.50) with respect to one parameter τ_i and replacing $p_i = e^{\tau_i}$, we find

$$p_i = \frac{q_i + \frac{1}{2} \text{tr}((mm^\dagger + D)S^{(i)})}{\alpha_i - 1 + \frac{q_i}{2} + (T \log p)_i}. \quad (2.51)$$

With this equation we can update the power spectrum parameters for each iteration using the current m and D .

This is in perfect accordance with previous findings (Enßlin and Frommert 2011; Enßlin and Weig 2010; Oppermann et al. 2013) and shows effectively that we can re-discover the critical filter for a pure MAP approach if we accept the approximation (2.49) as valid.

2.4 Implementation of RESOLVE⁶

2.4.1 General implementation

We have implemented RESOLVE in PYTHON, where crucial parts have been translated into more efficient C code using CYTHON⁷. The actual implementation of the algorithm makes heavy use of the versatile inference library NIFTY (Selig et al. 2013).

To perform the gridding and degridding operations needed in radio astronomical applications, we use the generalized Fast Fourier transformations package GFFT⁸. The grid convolution is performed using a Kaiser-Bessel kernel following Beatty et al. (2005).

For numerical optimization, we use a self-written steepest descent solver and in some cases the conjugate gradient routine provided by the SCIPY package.

The algorithm is controlled by a number of numerical procedures and parameters, governing the grade of convergence and the degree of accuracy. Apart from standard parameters, such as the maximum number of iterations or the accuracy of the steepest descent, the most important are:

- Different starting guesses for s and p might have strong impact on the performance or the solution of RESOLVE. In non-linear optimization, there is, for instance, always the danger to only converge to a local minimum. Experience showed that in most cases, it is optimal to use constant fields and simple generic power spectra as starting guesses to prevent any biases. But other options are available, e.g. a CLEAN or a dirty map, and/or their respective empirical power spectra, in some cases allowing for an improvement in computation time.
- To calculate D for (2.51), we have to numerically invert D^{-1} and statistically probe the needed matrix entries (Selig et al. 2012) using an implicit representation of the operator as a coded function. For this, we employ a conjugate gradient routine whose convergence and accuracy parameters must be set. This numerical inversion is usually the most serious bottleneck in computation time (see Sec. 2.4.2). Especially calculating D for an estimate of the signal uncertainty can be a time consuming task, depending on the accuracy needed.
- For observations with rather poor uv -coverage, problems might occur with the inversion of the operator D , which sometimes tends to be numerically non-positive definite during early iterations. In that case, we have implemented a solution where a diagonal matrix with a user-defined positive constant M_0 gets added to D^{-1} to ensure positive-definiteness. While the solution is slowly converging over the global iterations, M_0 is constantly decreased. This is a standard approach in numerical optimization, see for instance Transtrum and Sethna (2012).

⁶Note: This section was originally part of the appendix of Junklewitz et al. (2013) and has been shifted into the main body of the thesis for clarity.

⁷See <http://docs.cython.org/>.

⁸See <https://github.com/mrbell/gfft>

- For large data sets, it is sometimes of high advantage to bin the power spectrum instead of mapping it over all possibly allowed modes set by the user defined image size. Otherwise, the calculation might take prohibitively long.

2.4.2 Analysis of algorithmic efficiency

As visualized in Fig. 2.1, RESOLVE mainly consists of two parts, a signal estimator, and a power spectrum estimator. They are iterated N_{global} times, until convergence is achieved, while both the maximum number of iterations and the exact convergence criteria can be set by the user. The signal estimator utilizes a steepest descent algorithm to solve Eq. (4.18), which needs N_{sd} internal iterations. The power spectrum is estimated with Eq. (4.20), where the trace of the inverse operator given by Eq. (4.19) needs to be calculated. Since the operator is only given implicitly, its diagonal entries need to be probed N_{pr} - times using random vectors (Selig et al. 2012), where, for each probe, the operator equation (4.19) has to be inverted using a conjugate gradient algorithm.

The steepest descent iterations are dominated by the operations needed to calculate M (see Eq. (4.18)), which involves the response operator R with a FFT and a subsequent Gridding operation. Therefore, its computational cost goes roughly with $N_{\text{sd}}(O(n_d) + O(n_s \log(n_s)))$, where n_d is the total number of visibilities, and n_s the number of pixels in image space.

The conjugate gradient is dominated by the need to compute the same operation, only, at least some fraction of n_s times, and for each probe individually. Usually a maximum of $\sqrt{n_s}$ iterations of the conjugate gradient are performed. This leads to a total computational cost of roughly $N_{\text{pr}}(O(\sqrt{n_s}n_d) + O(\sqrt{n_s}n_s \log(n_s)))$.

A realistic assessment of the asymptotic overall algorithmic efficiency is complicated, because all of the iteration numbers, N_{global} , N_{sd} , and N_{pr} can in principle vary strongly from case to case. Although N_{sd} usually will be larger than N_{pr} ⁹, the conjugate gradient term will likely dominate the algorithmic costs. In realistic applications, n_d will usually be larger than n_s , because, for modern instrument data sets, the number of visibilities can reach the millions. In that case, the algorithmic efficiency probably tends to $N_{\text{global}}N_{\text{pr}}O(\sqrt{n_s}n_d)$.

In addition, this analysis shows that calculating an estimate for the uncertainty of the signal reconstruction is very costly. To accurately compute the diagonal of D , a large number of probes is needed so that N_{pr} can easily exceed the thousands.

On our development machine, with up to 8 used CPUs and a maximum of 64GB working memory, the non-optimized code produced the results presented in Sec. 2.6.1 in roughly a couple of hours for the low noise case, and a couple of days for the high noise case. For the relatively small size of the simulated VLA snapshot data sets, we never used more than a few percent of the memory but this would most likely change for larger data sets.

⁹At least empirically taken from the simulations, the number of probes can be kept well below a couple of hundreds.

2.5 A signal inference view on visibility weighting¹⁰

In radio astronomy, the imaging step of aperture synthesis is usually combined with a weighting scheme that is included in the Fourier inversion of the visibilities. Essentially, the term W in (2.3), defining the dirty image I^D , can be expanded to hold more factors than the mere sampling function

$$I^D = \mathcal{F}^{-1}(T \cdot B \cdot w \cdot S \cdot \mathcal{F}I) \quad (2.52)$$

with $W = T \cdot B \cdot w \cdot S$, where T is a possible tapering of outer visibilities, B is a user-defined baseline weighting, w are the statistical noise weights obtained from an analysis of the thermal noise, and S is the sampling function.

Historically, mainly two weighting schemes have been employed. *Natural* weighting just multiplies every visibility point with the inverse thermal noise variance for the particular baseline and is therefore a simple, noise-dependent down-weighting mechanism. *Uniform* weighting ensures that the weight per gridded visibility cell is constant and, hence, effectively gives higher weight to outer baselines, where usually less visibility points are found in a grid cell.

In a seminal work (Briggs 1995b), Briggs has shown that natural weighting can be obtained under the constraint that the sample variance of the image should be minimized. In contrast, uniform weighting can be shown to reduce sidelobe levels, but actually down-grades sensitivity at the same time.

In the same work, a new weighting scheme was devised that interpolates between these two extremes, called *robust* weighting. The robust weights are determined as

$$W(k) \propto \frac{1}{1 + \sigma^2(k)/s_p^2(k)} \quad (2.53)$$

where σ^2 is the thermal noise variance, and s_p^2 is some parameter that originally was derived having in mind some measure of the source power at the given visibility (Briggs 1995b). In practice, s_p^2 is usually adjusted by hand to meet the needs of the astronomer for having a trade-off between sensitivity and resolution.

This form of weighting can be explained within the presented Bayesian framework, and, furthermore, we will show that an algorithm like RESOLVE *automatically* chooses the optimal weighting parameters according to the ratio of estimated noise and signal power.

For this, we consider the negative logarithm of the posterior (2.19), i.e. the Hamiltonian of our inference problem (see Eq. (2.33) in App. 2.3 for details)

$$H(s, d) = \frac{1}{2} s^\dagger S^{-1} s + \frac{1}{2} (e^s)^\dagger M e^s - j^\dagger e^s + H_0. \quad (2.54)$$

¹⁰Note: This section was originally part of the appendix of Junklewitz et al. (2013) and has been shifted into the main body of the thesis for clarity.

We can expand the exponents in a Taylor series and separate the quadratic from the higher orders in s as we have done in (2.20):

$$H(s, d) = \frac{1}{2} s^\dagger (S^{-1} + M) s - s^\dagger j + H_0 + \sum_{k=3}^{\infty} \frac{1}{k!} \Lambda(M, j)_{x_1 \dots x_k}^k s_{x_1} \dots s_{x_k}. \quad (2.55)$$

If we now apply the MAP principle and set the derivative with respect to s to zero, we find

$$(S^{-1} + M) s - j + \Delta(M, j, s) = 0 \quad (2.56)$$

where we have defined $\Delta(M, j, s) = \frac{\delta}{\delta s} \sum_{k=3}^{\infty} \frac{1}{k!} \Lambda(M, j)_{x_1 \dots x_k}^k s_{x_1} \dots s_{x_k}$. We can partly solve this equation for s :

$$s = (S^{-1} + M)^{-1} j - (S^{-1} + M)^{-1} \Delta(M, j, s). \quad (2.57)$$

The first term is the analytic solution to the quadratic part of the full log-normal Hamiltonian. It was shown to be equivalent to a Wiener Filter applied to the data d (Enßlin et al. 2009), which would be the optimal solution for a purely Gaussian signal field.

Using (2.24) for the covariance matrices S and N and $j = R^\dagger N^{-1} d$, we can write the Wiener Filter operator in (2.57), $F = (S^{-1} + M)^{-1} R^\dagger N^{-1}$, in Fourier space:

$$F(k) = \frac{1}{1 + P_n^g(k)/P_s(k)} \quad (2.58)$$

where $P_n^g = G_{ku} P_n(u)$ is the noise power spectrum on the regular grid, defined by the gridding operator G from (2.8).

This has the exact same form as the definition of the robust weights (2.53), and even the original premise is fulfilled that the factor s_p^2 in (2.53) should be connected to the source power. The great difference is that the Wiener Filter automatically weights each mode in Fourier space differently, given that the signal power spectrum $P_s(k)$ is known.

We conclude that the classical robust weighting can be theoretically understood as the optimal solution to a signal reconstruction problem of a Gaussian signal field, equivalent to a Wiener Filter operation. In fact, this similarity between the robust weights and Wiener Filtering was already mentioned by Briggs himself (Briggs 1995b), although in that work, no clear explanation of the connection was given.

In common practice, of course, the weights are set manually, as only the knowledge of the signal power spectrum would allow for an automatic assignment. Since RESOLVE

reconstructs this power spectrum, it does implicitly assign these weights. Of course, RESOLVE solves (2.57) iteratively, and only the converged solution will give optimal weights for the log-normal inference problem. No simple and direct equivalence can be given between these effective weights and robust weighting. It is even unclear how to write them down explicitly since the sum in $\Delta(M, j, s)$ in principle extends infinitely.

2.6 Test Simulations

In what follows, we present a range of tests of RESOLVE using simulated data. We have implemented the algorithm¹¹ in PYTHON using the versatile signal inference library NIFTY (Selig et al. 2013). For all details of the implementation, we refer the reader to Sec. 2.2 and App. 2.4. We also show comparisons to CLEAN and MEM to benchmark the performance and fidelity of our algorithm.

For all tests, we constructed simulated observations with the tool MAKEMS¹² using a realistic uv -coverage from a VLA observation in its A-Configuration. The VLA samples the uv -plane non-uniformly at irregular intervals, and the response includes thereby a convolutional gridding and degridding operator using a Kaiser-Bessel kernel (for details see Eq. 2.8 and App. 2.4). We simulated observations at a single frequency, approximatively 20 minutes snapshot observation with a total of 42 120 visibility measurements at a central frequency of 1 GHz (see Fig. 4.1). This setting leads to an especially sparse sampling of the uv -plane. For ease of code development and testing, we have not used longer observations. On the other hand, if we can solve the more demanding cases of sparse uv -coverage, we certainly can handle better suited data.

Through all simulations, we varied thermal visibility noise levels and input signals.

For the next two sections (2.6.1 and 2.6.2), the signals were drawn from a log-normal distribution, exactly meeting our prior assumptions. In Sec. 2.6.3, we go beyond that and illustrate the validity of our statistical model by using a signal derived from a CLEAN image of a real source.

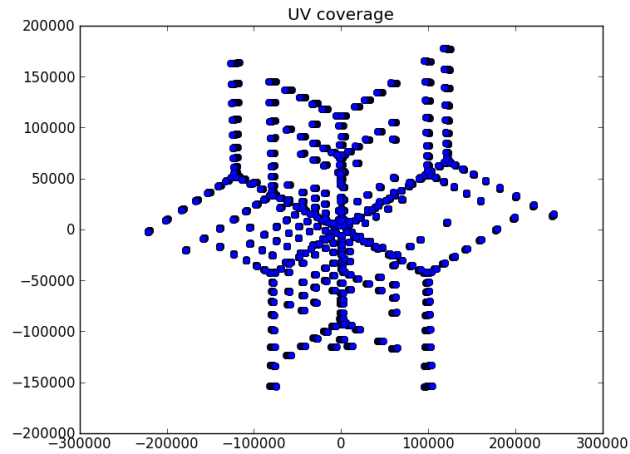
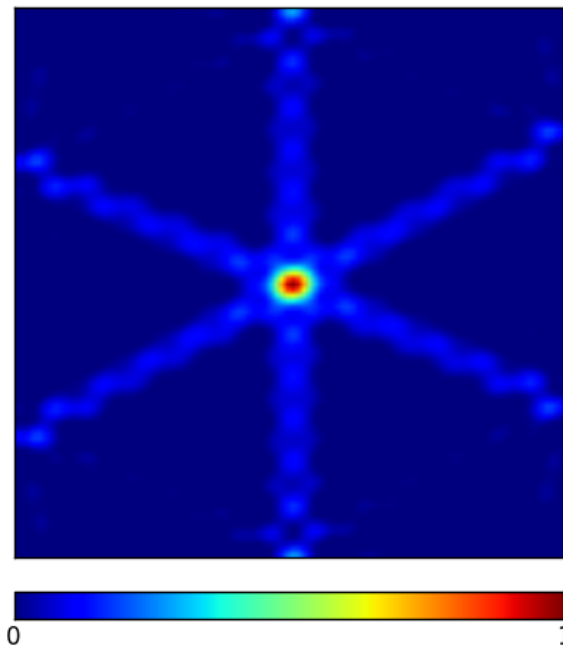
The complex, Gaussian input noise in uv -space is defined by a user-defined variance, equal for all visibilities. The code does not require equal noise variances and can in principle handle varying variances as well. Always in the following, *low noise* refers to $\sigma_{\text{ln}}^2 = 10^{-3} \text{Jy}^2$, whereas *high noise* denotes $\sigma_{\text{hn}}^2 = 10^5 \text{Jy}^2$ ¹³. These numbers are of course somewhat arbitrary, only chosen for demonstrational reasons. They are not intended to necessarily reflect realistic visibility noise values in every possible aspect, but to serve as examples for particularly low or high noise cases.

To give a quantitative account of the accuracy of the reconstructions, we use a relative

¹¹To get access to the code prior to its envisaged public release, please contact henrikju@mpa-garching.mpg.de or ensslin@mpa-garching.mpg.de.

¹²See <http://www.lofar.org/wiki/lib/exe/fetch.php?media=software:makems.pdf>.

¹³The unit Jy was used here for convenience. Effectively, it stands for whatever units the simulated signal is interpreted to be given in.

(a) uv -coverage in units of # of wavelengths.

(b) Point spread function.

Figure 2.2: uv -coverage and point spread function for the simulated 20 minutes snapshot observation in VLA-a configuration. The image of the point spread function is 100^2 pixels large, the pixel size corresponds to roughly 0.2 arcsec.

\mathcal{L}_2 - norm measure of the difference of signal to map:

$$\delta = \sqrt{\left(\frac{\sum (e^s - e^m)^2}{\sum (e^s)^2}\right)} \quad (2.59)$$

where the sums are taken over all pixels of the reconstruction. This choice is motivated by the fact that the inference approach underlying RESOLVE approximates a reconstruction that is optimal in the sense of minimizing this error measure (see Sec. 2.2 and Eq. 4.16 therein).

In Secs. 2.6.1 - 2.6.3, we focus exclusively on the reconstruction of the signal, i.e. the sky brightness distribution. The reconstruction of the power spectrum is discussed separately in Sec. 2.6.5.

2.6.1 Main Test Results

Here, we describe the main test results for the reconstruction of a simulated signal using RESOLVE.

In Fig. 2.3, an artificial log-normal signal is shown alongside with the results from RESOLVE for observations with low and high noise. The error measures are $\delta_{\text{ln}} = 0.12$ and $\delta_{\text{hn}} = 0.3$ for the low and high noise case respectively.

We see that we can recover all the structures of the original surface brightness, down to even very small features in the low noise case and at least all main features in the high noise case. All strong effects of the point spread function have been successfully removed, thus showing that RESOLVE is effective in deconvolving the dirty image.

In fact, the reconstruction is expected to be smoothed out on the smaller scales and lose overall power due to noise in the observation. This is simply because all information in the power spectrum gets lost for powers comparable to the noise variance (see Sec. 2.6.5 for details). Effectively, RESOLVE performs an automatic visibility weighting by comparing noise and signal power for all Fourier modes, much in the way robust weighting originally was conceived (Briggs 1995b,a). For a detailed discussion of this topic see App. 2.5.

2.6.2 Comparison to standard imaging methods

In this section, we give a short introduction to common imaging algorithms in radio interferometry and show comparisons to RESOLVE. We focus on two of them, MS-CLEAN and MEM, which are probably the most widespread methods to date.

In addition, we should mention recent developments in the application of Compressed Sensing (CS) (Candes et al. 2006; Donoho 2006) to radio imaging, most notably the development of the imaging algorithm SARA (Carrillo et al. 2012). Another approach was taken recently to apply Gibbs sampling methods to imaging in radio interferometry (Sutter et al. 2013), also within the framework of Bayesian inference, but restricted to pure Gaussian priors. A direct comparison of RESOLVE to either SARA or Gibbs Sampling methods is

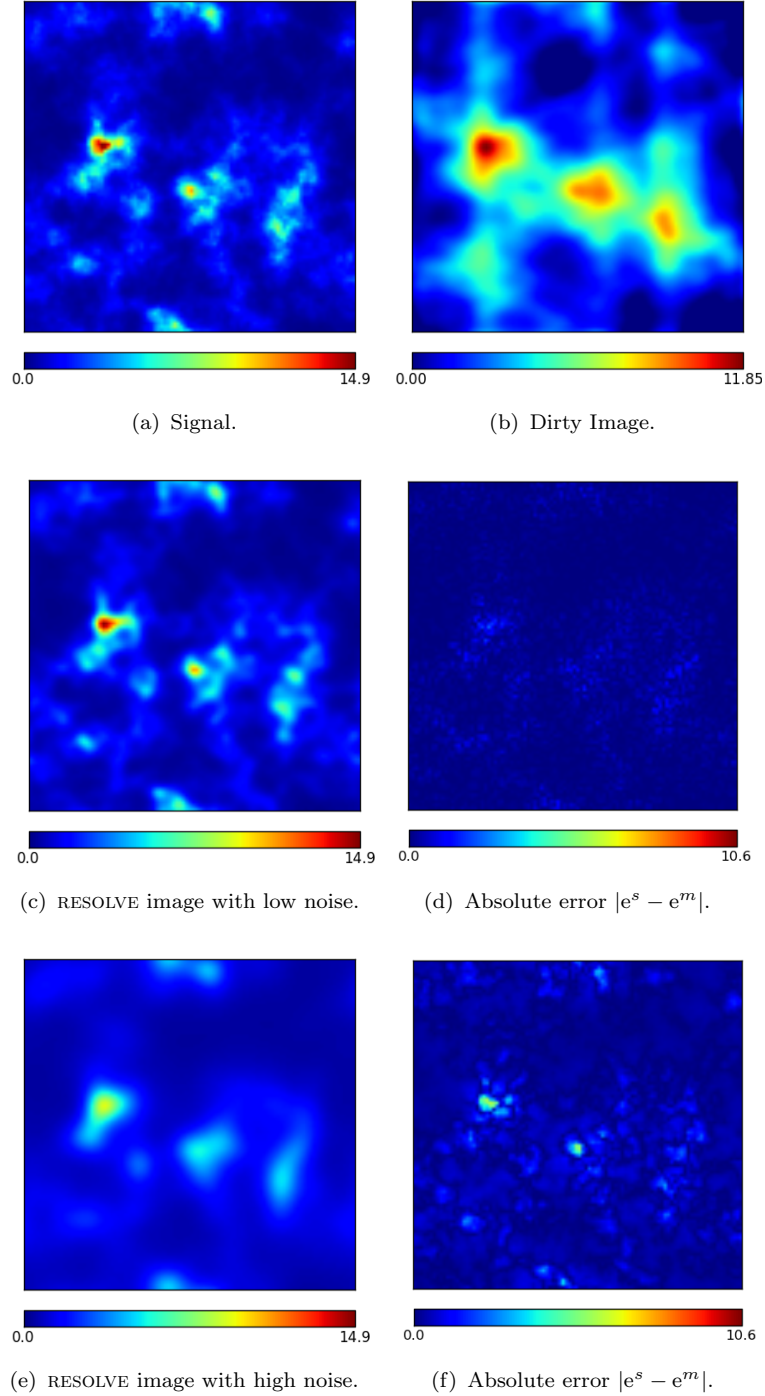


Figure 2.3: Reconstruction of a log-normal signal field, observed with a sparse uv -coverage from a VLA-A-configuration and different noise levels. The images are 100^2 pixels large, the pixel size corresponds to roughly 0.2 arcsec. The brightness units are in Jy/px. The ridge-like structures in the difference maps simply stem from taking the absolute value and mark zero-crossings between positive and negative errors. *First row left*: Signal field. *First row right*: Dirty map. *Second row left* RESOLVE reconstruction with low noise. *Second row right*: Absolute per-pixel difference between the signal and the RESOLVE reconstruction with low noise. *Third row left*: RESOLVE reconstruction with high noise. *Third row right*: Absolute per-pixel difference between the signal and the RESOLVE reconstruction with high noise.

out of the scope of this work, but we will discuss possible ways to include the CS approach into our Bayesian framework in the conclusions (see Sec. 2.7).

Standard imaging methods almost always use an additional visibility weighting that is manually set by the observer. We do not include such weights explicitly in RESOLVE (i.e. we set them to one). However, as already mentioned, an algorithm like RESOLVE *automatically* chooses optimal weights according to the ratio of reconstructed source power to noise power. A detailed derivation can be found in App. 2.5. For CLEAN, we compare to different weighting schemes in order to be as unbiased as possible.

For both CLEAN and MEM we used the implementation in the radio astronomical software package CASA (Reid and CASA Team 2010).

Comparison to CLEAN

CLEAN was first presented by (Högbom 1974) and is surely the most widely used deconvolution algorithm in radio astronomy. It works around the major assumption that the image is comprised of point sources. In its simplest variant, it iteratively finds the highest peak in the dirty map, subtracts a psf-convolved fraction of a delta function fitted to the peak, and saves the delta components in a separate image. After some noise threshold is reached, the algorithm stops and re-convolves the components with a so called clean beam, usually the main lobe of the point spread function or a broader version of it to downgrade resolution.

Over time, many variants of CLEAN have been developed, most notably Clark CLEAN (Clark 1980), Cotton-Schwab CLEAN (Schwab 1984), multifrequency CLEAN (Sault and Wieringa 1994) and multiscale CLEAN (MS-CLEAN) (Cornwell 2008). The latter was constructed to better reflect extended emission by subtracting Gaussians of various shapes instead of pure point sources. We will thus compare the results of RESOLVE to MS-CLEAN.

It has been proved in the framework of compressed sensing that CLEAN is in fact a variant of a matching pursuit algorithm (Lannes et al. 1997). This class of algorithms can be shown to be optimal for signals that are sparse in some basis (like a point source signal on the sky). It can be understood as minimizing the \mathcal{L}_1 -norm of the signal field and can therefore be cast into the assumption of a Laplacian prior distribution, which would allow in principle a representation of CLEAN in a Bayesian inference framework (see Wiaux et al. 2009, and references therein).

In Figs. 2.4 and 2.5, a comparison is shown between the results of RESOLVE and MS-CLEAN as implemented in the radio astronomical software package CASA. For this test, the same simulated low noise data were used as in Sec. 2.6.1. We compare to three different CLEAN reconstructions with natural, uniform and robust weighting (robust parameter $r = 0$, which gives an intermediate result between the other two schemes). We used a very small noise threshold and a standard gain factor of 0.1. In total, we choose to run the algorithm interactively for around 1000 iterations. We used around ten different scales for the multi-scale settings, ranging from a few pixels to enough to roughly match the scales found in the signal. Together with the reconstructions, we show maps of the squared error $(e^s - m)^2$ for each of them. The \mathcal{L}_2 - error measures are shown in Table 2.1.

Algorithm	δ
RESOLVE	0.12
MS-CLEAN, natural	1.46
MS-CLEAN, uniform	0.67
MS-CLEAN, robust	0.69
MEM	1.07

Table 2.1: \mathcal{L}_2 error measures for RESOLVE, MS-CLEAN and MEM for the low-noise simulation and the reconstruction shown in Figs. 2.4, 2.5 and 2.6.

Both quantitative analysis and visual comparison show that RESOLVE clearly outperforms MS-CLEAN in this case. Its result is closer to the signal in the \mathcal{L}_2 error measure sense and it is clearly superior in reconstructing the detailed extended structure of the surface brightness signal. Especially the very weak emission around all the brighter sources is much better resolved and denoised than in the MS-CLEAN images. The reconstruction with natural weighting is overestimating the flux scales considerably, while uniform and robust weighting roughly find the same correct solution as RESOLVE. However, it should be noted that, at least for natural weighting, this is a somewhat biased comparison, since the natural weighting scheme is by construction enhancing point-source sensitivity while preserving larger side-lobe structures (Briggs 1995b) and thus not the optimal choice for resolving extended emission.

Comparison to the Maximum Entropy Method (MEM)

The maximum entropy method (MEM) is an imaging algorithm introduced into radio astronomy by (Cornwell and Evans 1985). It actually goes back to earlier developments in statistical inference, connected to the broad field of entropic priors (Gull and Daniell 1979; Skilling et al. 1979). It should not be confused with the maximum entropy principle of statistics (see App. A that describes how to update probability distributions when new information has to be included).

MEM aims to maximize a quantity called image entropy S_{im} , which is defined for strictly positive signal images s as

$$S_{\text{im}} = - \int dx s(x) \log(s(x)/m(x)) \tag{2.60}$$

where $m(x)$ is a model image of the observed signal, thus allowing to introduce some kind of prior information into the problem. The data enter this formalism as a constraint for the maximization problem. Usually, one adds a term to (2.60) that measures the closeness of the entropic signal reconstruction to the data in the form of a $\chi^2(d, Rs)$ distribution, which is nothing else but the log-likelihood of (2.15):

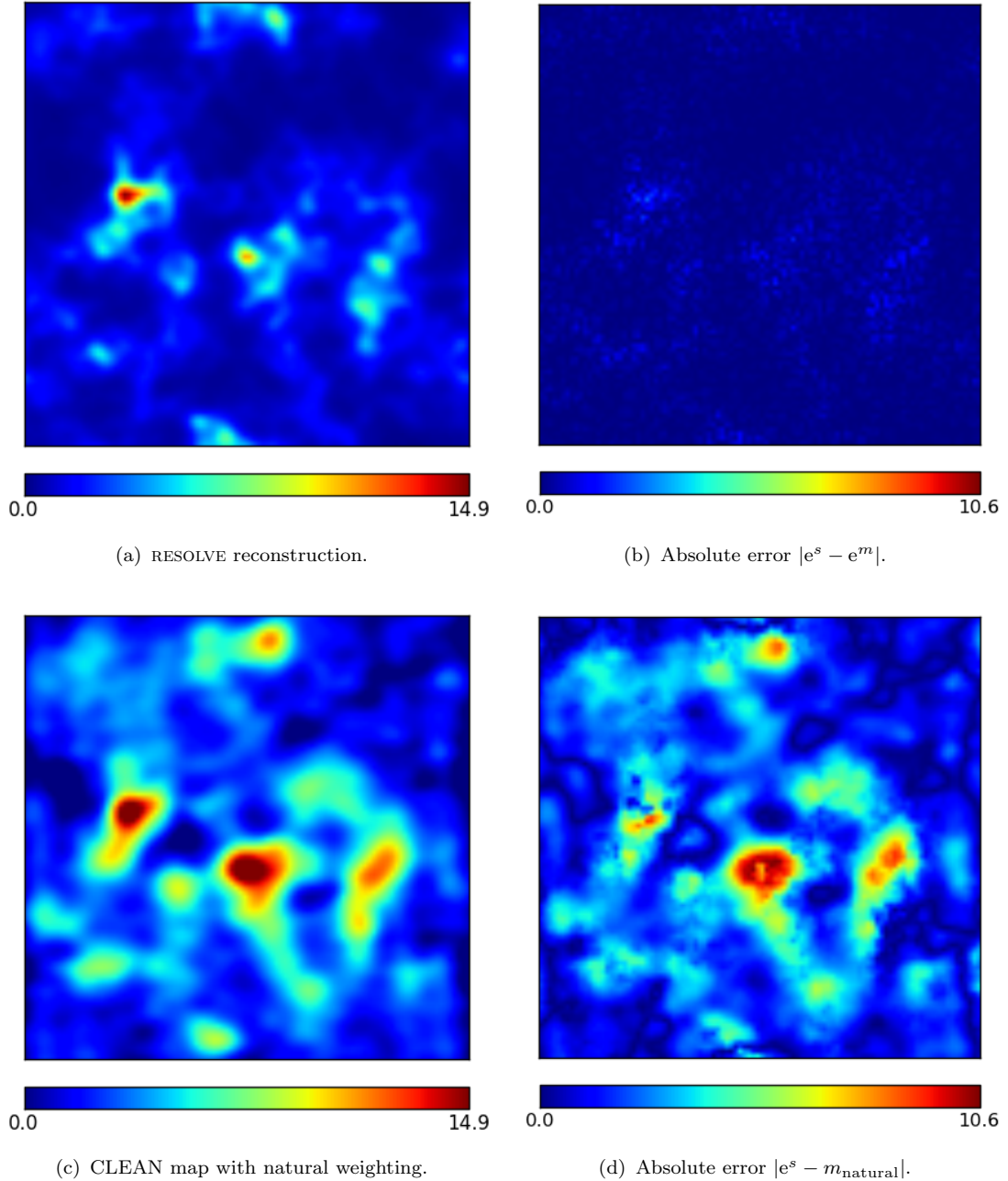


Figure 2.4: Comparison of RESOLVE with MS-CLEAN for the simulated low noise observation of Sec. 2.6.1. The images are 100^2 pixels large, the pixel size corresponds to roughly 0.2 arcsec. The brightness units are in Jy/px. The ridge-like structures simply stem from taking the absolute value and mark zero-crossings between positive and negative errors. *First row left*: RESOLVE reconstruction. *First row right*: Absolute per-pixel difference between the signal and the RESOLVE reconstruction with. *Second row left*: MS-CLEAN reconstruction with natural weighting using the radio astronomical software package CASA. *Second row right*: Absolute per-pixel difference between the signal and the MS-CLEAN reconstruction with natural weighting.

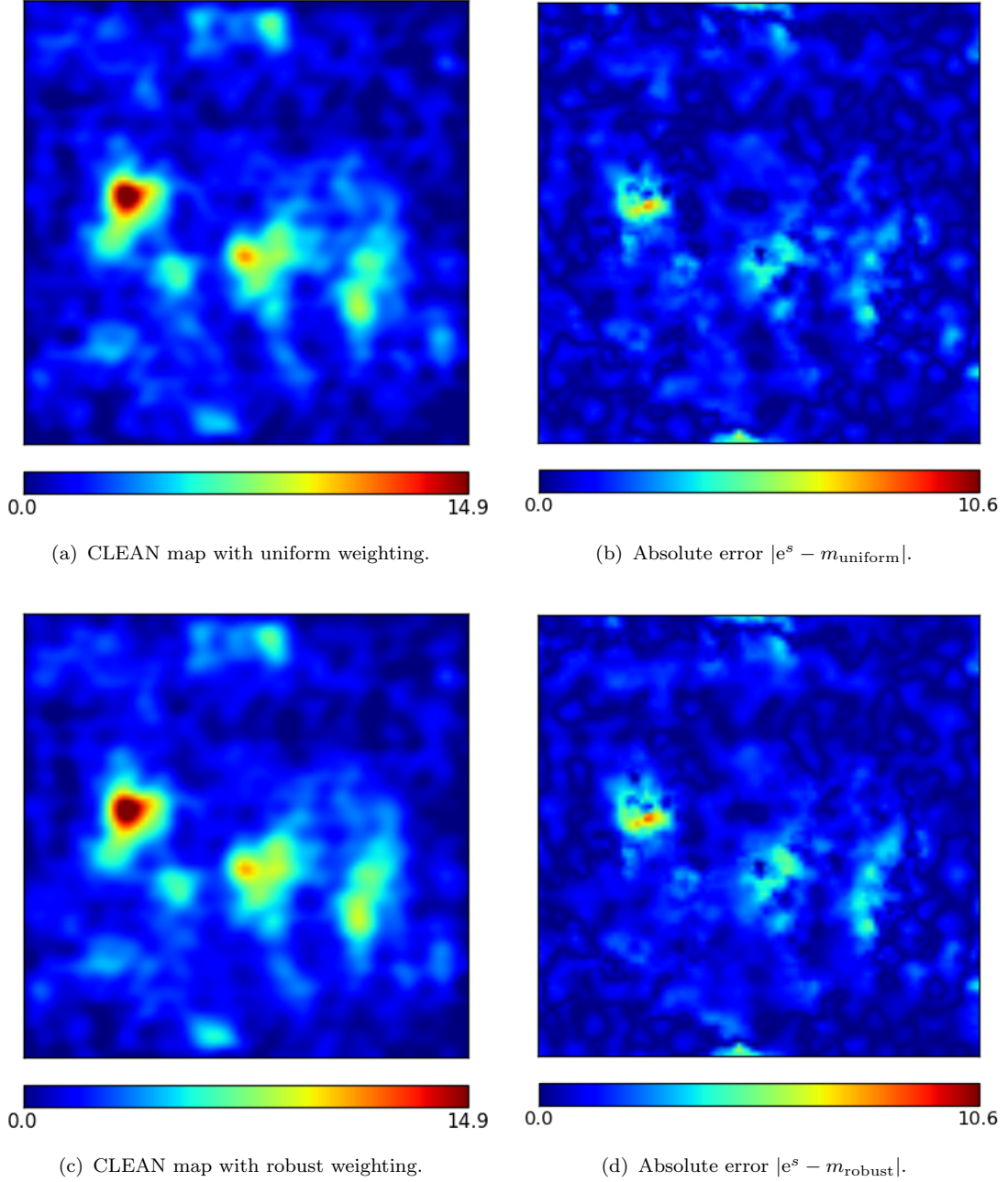


Figure 2.5: Comparison of RESOLVE with MS-CLEAN for the simulated low noise observation of Sec. 2.6.1. The images are 100^2 pixels large, the pixel size corresponds to roughly 0.2 arcsec. The brightness units are in Jy/px. The ridge-like structures simply stem from taking the absolute value and mark zero-crossings between positive and negative errors. *First row left*: MS-CLEAN reconstruction with uniform weighting using the radio astronomical software package CASA. *First row right*: Absolute per-pixel difference between the signal and the MS-CLEAN reconstruction with uniform weighting. *Second row left*: MS-CLEAN reconstruction with robust ($r = 0$) weighting using the radio astronomical software package CASA. *Second row right*: Absolute per-pixel difference between the signal and the MS-CLEAN reconstruction with uniform weighting.

$$\begin{aligned}\frac{1}{2}\chi^2(d, Rs) &= \frac{1}{2}(d - Rs)^\dagger N^{-1}(d - Rs) \\ &= -\log(P(d|s)) + \text{const.}\end{aligned}\tag{2.61}$$

With (2.60) and (2.61), MEM achieves a solution by extremizing

$$J(d, s) = -\log P(d|s) - \mu S_{\text{im}}\tag{2.62}$$

for s . The multiplier μ is usually adjusted during the extremization to meet numerical constraints (see Cornwell and Evans 1985, for details).

We now repeat a short section from (Enßlin and Weig 2010), analyzing the assumptions of this approach from the viewpoint of Bayesian signal inference.

As we have identified (2.15) as the log-likelihood, it is also possible to re-identify the prior distribution. If we interpret $J(d, s)$ as a Hamiltonian $H(d, s)$, than the entropy term can be understood as a log-prior

$$\mu S_{\text{im}}(s) = \log \mathcal{P}(s).\tag{2.63}$$

With this, we can read off the underlying prior distribution implicitly assumed in MEM

$$\begin{aligned}\mathcal{P}(s) &= \exp \left[-\mu \int dx s(x) \log \left(\frac{s(x)}{m(x)} \right) \right] \\ &= \prod_x \left(\frac{s(x)}{m(x)} \right)^{-\mu s(x)}.\end{aligned}\tag{2.64}$$

This prior is very specific. It extremely suppresses strong pixel values and thereby favors to smooth out emission over all pixels in the image while sharp peaks are heavily down-weighted. It implicitly assumes no correlation between pixels, and a more than exponentially falling brightness distribution. In the case of the model $m(x)$ being a close approximation to real signal, the prior becomes effectively flat and MEM turns basically into a Maximum Likelihood fit.

In Fig. 2.6, a comparison is shown between the results of RESOLVE and MEM as implemented in the radio astronomical software package CASA. Again, the same simulated low noise data were used as in Sec. 2.6.1. As a model image, we used an MS-CLEAN reconstruction with uniform weighting. We again show maps of the squared error $(e^s - m)^2$ for the reconstruction with RESOLVE and MEM respectively. The \mathcal{L}_2 error measures are shown in Table 2.1.

It can be clearly seen that RESOLVE also outperforms MEM. There is much more false structure in the MEM reconstruction, as reflected by the ℓ_2 - norm analysis. Partly, this might be due to the specific MEM prior that enforces to smooth out the signal over

all pixels, partly, it seems to be due to badly deconvolved remnants of the point spread function. However, we note that the MEM implementation in CASA still is considered to be somewhat experimental, and that a more stable code or a longer time of parameter adjustment and fine-tuning might improve these results.

2.6.3 Comparison with a real signal

So far we have only shown reconstructions of signals that were drawn from log-normal statistics, using the exact assumptions that we use to specify the prior distribution. To some degree, it is expected that RESOLVE should be optimal for these simulated signals.

To further demonstrate the validity of our assumptions, we have conducted a test where we did not use a signal drawn from log-normal statistics. Instead, we took an MS-CLEAN image, obtained from real data of the galaxy cluster Abell 2256 (Clarke and Enßlin 2006), and reused this as a signal for the simulated observation using the same VLA configuration as before. The original data were taken with the VLA at 1.369 GHz in D-configuration. The surface brightness values are not in the original range but chosen arbitrarily in our simulation, effectively given in Jy/px. The signal (i.e. the adapted CLEAN image of Abell 2256) and the reconstruction from RESOLVE are shown in Fig. 2.8.

Although this time we have at no point introduced log-normal statistics into the simulation process, the prior assumption still seems to be valid and leads to results comparable in exactness to the tests using explicit log-normal signals.

2.6.4 Signal Uncertainty

As already stated in Sec. 2.2.3, RESOLVE provides also an estimate of the uncertainty of the signal reconstruction. The algorithm uses the inverse second derivative D of the posterior, evaluated at the specific signal estimate m , to approximate the posterior covariance. In App. 2.3.2, it is shown that a full signal estimate taking into account approximative uncertainty leads to

$$I \approx e^{m_x} \pm \sqrt{e^{2m_x} [e^{D_{xx}} - 1]}. \quad (2.65)$$

In Fig. 2.7, we present an example of the approximated relative uncertainty

$$\sqrt{\frac{\langle (e^{s_x})^2 \rangle_{\mathcal{G}(m,D)} - \langle e^{s_x} \rangle_{\mathcal{G}(m,D)}^2}{\langle e^{s_x} \rangle_{\mathcal{G}(m,D)}^2}} = \sqrt{[e^{D_{xx}} - 1]} \quad (2.66)$$

for the low noise reconstruction of Sec. 2.6.1, together with the signal estimate, and absolute and relative difference map between signal and estimate. The subscripts indicate that our approach effectively involves to approximate the full posterior with a Gaussian $\mathcal{G}(m, D)$ centered on the signal estimate and with a covariance of D (see App. 2.3.2).

Fig. 2.7 shows that the uncertainty follows the structure of the reconstruction. Where the signal is strong, the relative uncertainty is much lower than in regions that are mainly

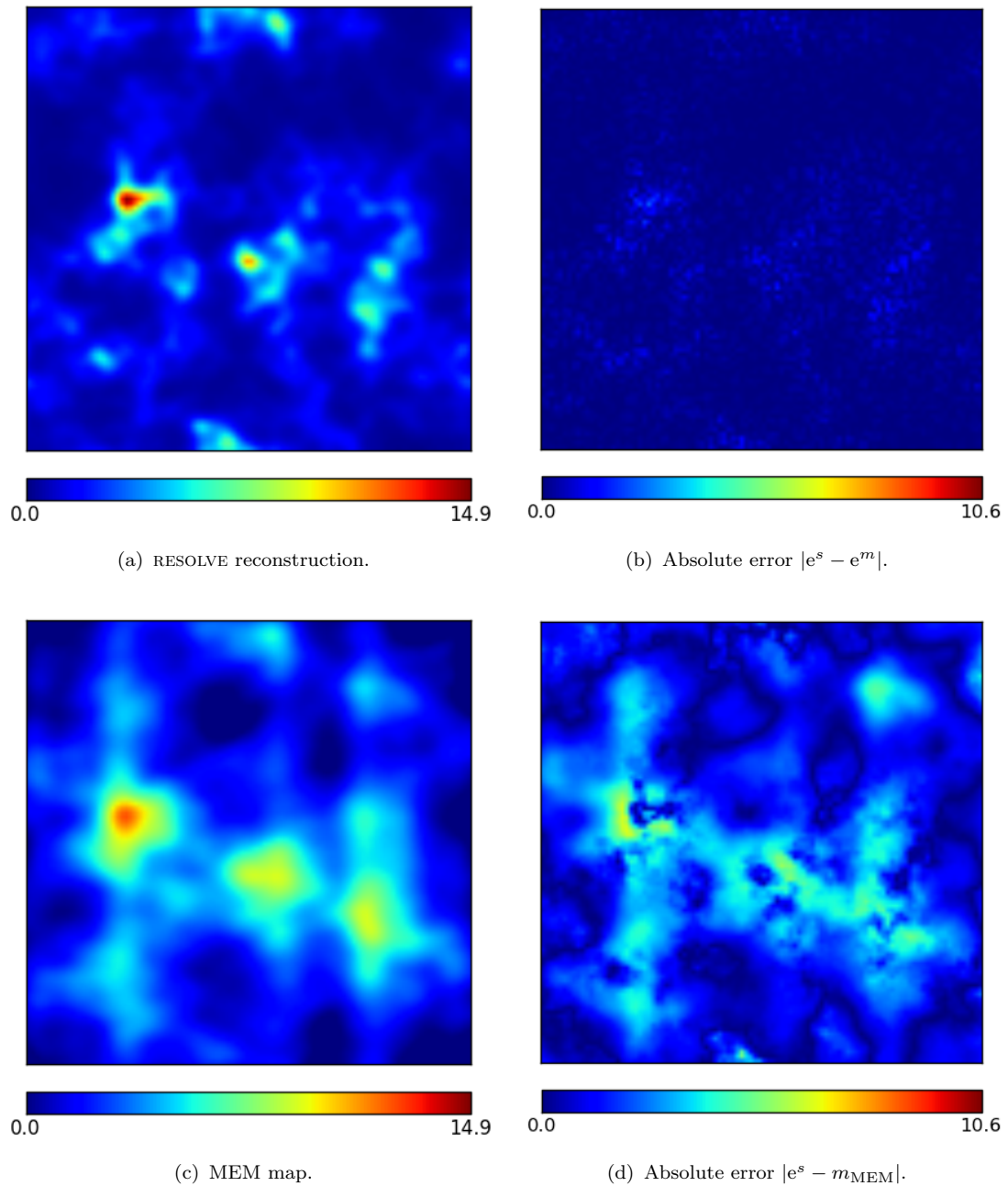


Figure 2.6: Comparison of RESOLVE with MEM for the simulated low noise observation of Sec. 2.6.1. The images are 100^2 pixels large, the pixel size corresponds to roughly 0.2 arcsec. The brightness units are in Jy/px. The ridge-like structures simply stem from taking the absolute value and mark zero-crossings between positive and negative errors. *First row left*: RESOLVE reconstruction. *First row right*: Absolute per-pixel difference between the signal and the RESOLVE reconstruction. *Second row left*: MEM reconstruction using the radio astronomical software package CASA. *Second row right*: Absolute per-pixel difference between the signal and the MEM reconstruction.

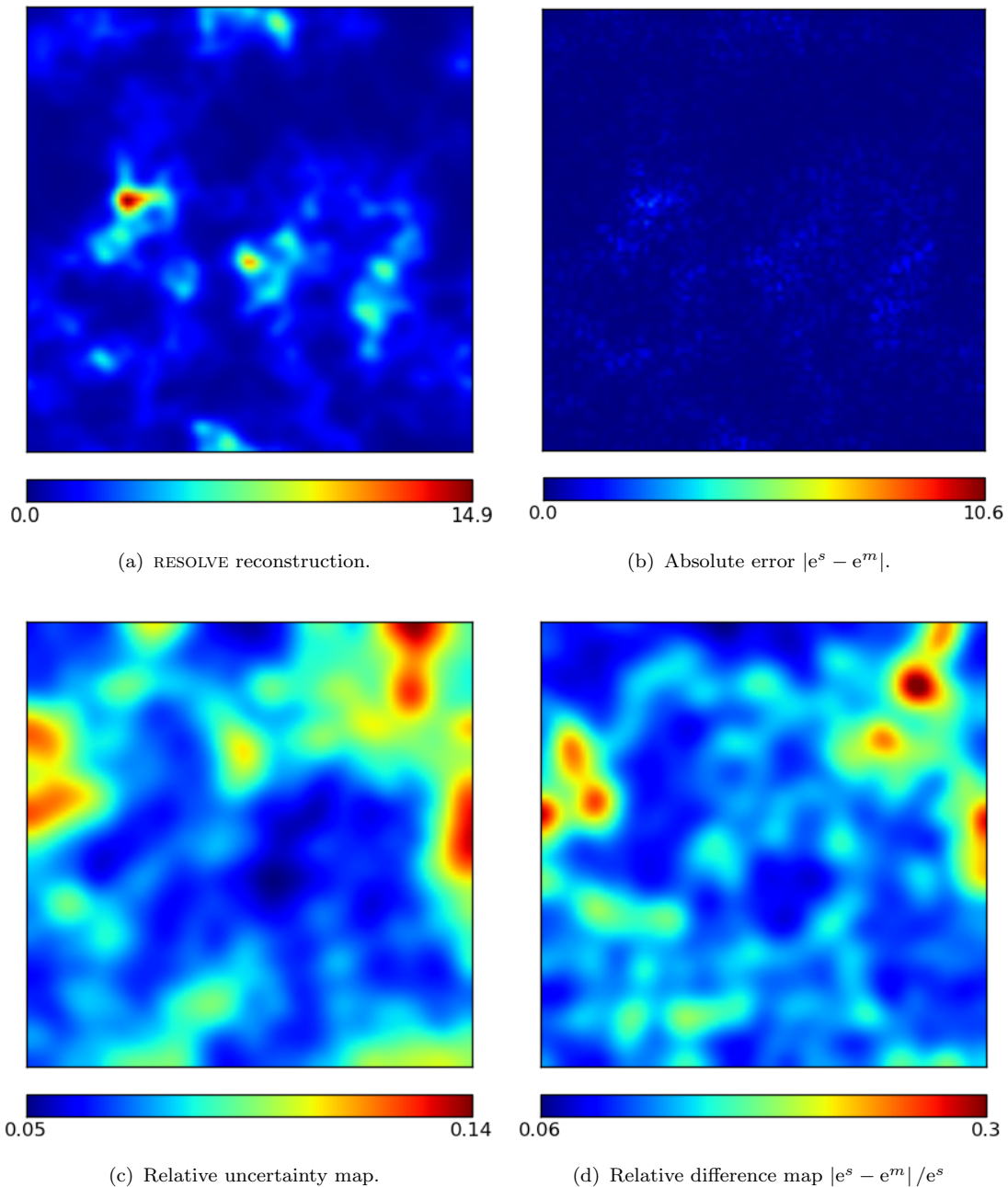


Figure 2.7: *First row left*: RESOLVE reconstruction for the low noise reconstruction of Sec. 2.6.1. *First row right*: Absolute per-pixel difference between the signal and the RESOLVE reconstruction. The ridge-like structures simply stem from taking the absolute value and mark zero-crossings between positive and negative errors. *Second row left*: Relative Uncertainty map derived from the RESOLVE reconstruction. *Second row right*: Relative difference map between signal and RESOLVE reconstruction.

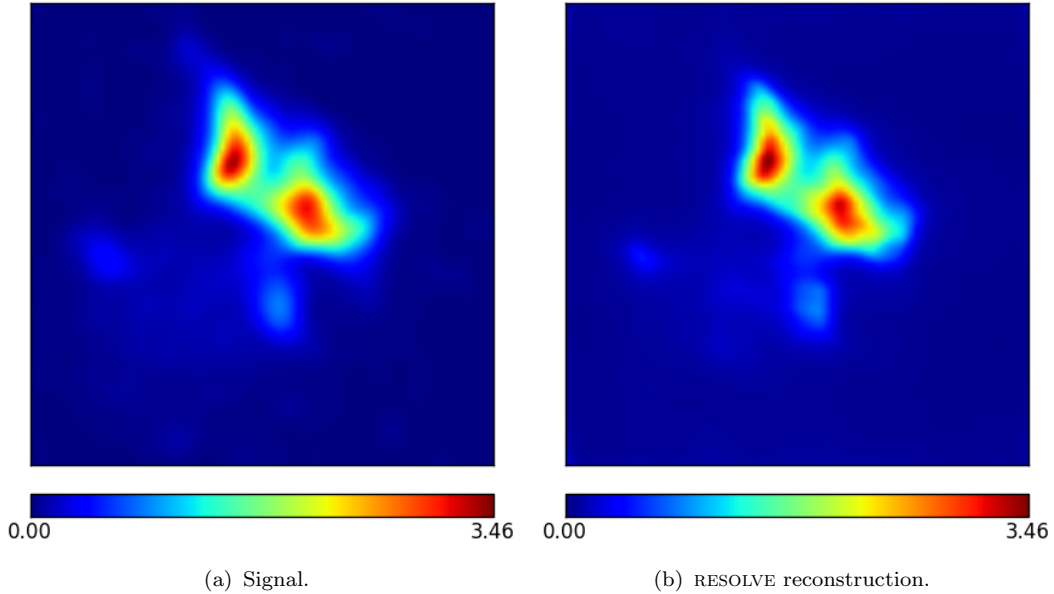


Figure 2.8: Reconstruction of a signal field that was obtained from a CLEAN image of the real extended emission of Galaxy cluster Abell 2256. For the simulation, the same setup with low noise was used as in Sec. 2.6.1.

dominated by noise. A comparison between the estimated relative uncertainty and the real relative difference map shows the approximative nature of the theoretical estimate. While both maps agree nicely in structure, they do not fully match in terms of values. Overall, the theoretical uncertainty underestimates the real relative difference. However, it should be noted that the deviations between both maps are much stronger in the outer regions, where the signal is only weak. In the center of the map, where the source mainly is located, both agree relatively well.

If we further use (2.65) to calculate the absolute uncertainty for the low noise reconstruction of Sec. 2.6.1, we find that roughly 40 % of the original signal values lie within a $1 - \sigma$ region, and roughly 70 % within a $2 - \sigma$ region. Although this result deviates from pure Gaussian expectations, this is a reasonable outcome. Since the posterior is in general non-Gaussian, the assumption of posterior Gaussianity needed to exactly define (2.65) can only result in an approximation.

Calculating the uncertainty to a very high precision is computationally expensive¹⁴. It involves the probing of an implicitly defined matrix and a numerical algorithm to invert this matrix (see App. 2.4). In this case, we have stopped the stochastic probing of D at some point for computational reasons and smoothed the outcome a bit to obtain Fig. 2.7. This might add to the deviations from pure Gaussian expectations on the absolute

¹⁴The estimation of the uncertainty goes roughly with $N_{\text{pr}} (O(\sqrt{n_s n_d}) + O(\sqrt{n_s n_s} \log(n_s)))$, where N_{pr} is the number of probes, n_d the number of visibility measurements, and n_s the number of pixels in image space (see App. 2.4)

uncertainty, mentioned earlier. Nevertheless, since the theoretical matrix representation of D must be smooth, this procedure should be acceptable as long as this example simply serves as a showcase to fundamentally demonstrate how to obtain an uncertainty estimate with RESOLVE.

2.6.5 Power Spectrum Reconstructions

Until now, we have focused entirely on the reconstruction of signal maps. Now we discuss the reconstruction of the signal power spectrum that RESOLVE achieves automatically in order to infer the best signal solution. The signal power spectrum is defined as the Fourier transformation of the autocorrelation function of the signal, assuming translationally and rotationally invariant statistics:

$$P(|k|) = \int dr C(r) \exp(ikr). \quad (2.67)$$

(for more details, see Sec. 2.2.3).

Qualitatively, it can be understood as decomposing the signal autocorrelation into its different contributions from various scales. High power on low Fourier modes means strong correlations on larger scales and high power on high Fourier modes means strong correlations on smaller scales.

In the first row of Fig. 2.9, we show the reconstruction of power spectra for the low and high noise reconstructions of Sec. 2.6.1. The figure shows the original power spectrum, which defines the correlation structure of the signal field, and the final results of RESOLVE after 6 iterations in the low, and 80 iterations in the high noise case. It can be seen that, with more noise, the reconstruction loses sensitivity for the smaller scales. This is reflected in the high noise map reconstruction in Fig. 2.3, where the smallest scales are smoothed out by the algorithm.

The second row of Fig. 2.9 serves as an example for the actual reconstruction process, where all of the 80 iterations for the high noise power spectrum are shown, together with the starting guess, which was a simple and generic power law $P_{\text{sg}} \propto k^{-2}$. The power spectrum first dropped, to slowly rise up again. This is a consequence of a numerical procedure to ensure the convergence of the underlying non-linear optimization routines, where a constant diagonal is first added to the uncertainty estimate D^{-1} used in the power spectrum reconstruction, and then suppressed again with converging iterations (see App. 2.4).

We emphasize that an accurate power spectrum reconstruction can be a scientific result on its own and should not only be regarded as a mere by-product. Since this is a rather unusual topic for observations of radio total intensity, it might be in place to explain a little further its meaning and to outline possible scientific merits.

The most typical physical source of extended emission in radio astronomy is synchrotron radiation. By spelling the power spectrum of the total intensity from some astronomical synchrotron source we effectively measure its correlation structure. Since synchrotron intensity is in part determined by the magnetic field strength (Rybicki and Lightman 1985)

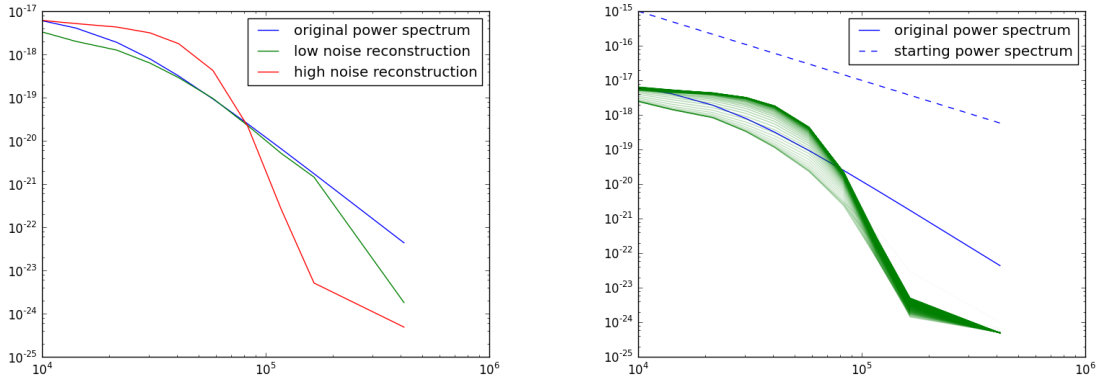


Figure 2.9: *First row*: Power spectrum reconstruction for the simulated low noise and high noise observations of Sec. 2.6.1. *Second Row*: Evolution of the high noise power spectrum reconstruction over 80 iterations. The iteration process is indicated from transparent to full green.

in the source, we automatically gather valuable scientific information on the magnetic field statistics as well:

$$C_I(r) = \langle I(x)I(x+r) \rangle \propto \langle B(x)^2 B(x+r)^2 \rangle. \quad (2.68)$$

Detailed derivations of this and related statistical quantities, together with many discussions on its scientific use, mostly in the context of analyzing turbulent magnetic fields, can be found in a series of astrophysical papers (e.g. Spangler 1982, 1983; Eilek 1989; Waelkens et al. 2009; Junklewitz and Enßlin 2011; Oppermann et al. 2011a; Lazarian and Pogosyan 2012)

For future observations, it might be especially interesting to use these results from RESOLVE to compare data of specific astrophysical synchrotron sources, e.g. supernova remnants or radio halos of galaxies and clusters, to simulations thereof. In simulations, the inputs are under control, and (2.68) can actually be calculated and compared with real data.

2.7 Conclusions

In this paper we presented a new approach to signal inference and imaging in radio astronomy and especially radio interferometry. The inference algorithm RESOLVE is targeted to be optimal for the imaging of extended and diffuse radio sources in total intensity. In simulations, RESOLVE demonstrated to produce high fidelity reconstructions of such extended signals, drawn from pure log-normal statistics or from real data. Comparisons showed that RESOLVE can outperform current imaging algorithms in these tasks.

Furthermore, RESOLVE is capable of producing an approximative uncertainty estimate for the inferred image through consistent propagation of measurement uncertainty. This is not possible with current imaging algorithms.

In addition to the inferred signal reconstruction, RESOLVE also estimates the power spectrum of the signal, i.e. its two-point correlation structure. The power spectrum is used for the signal reconstruction, but can be regarded as a new scientific outcome by itself. For instance, it opens opportunities to study the statistical properties of magnetic fields that lead to observed synchrotron emission. At the same time it offers a unique tool to compare simulations of turbulent, magneto-ionic media in extended radio sources to observations.

It was shown that instead of using classical visibility weights directly, RESOLVE chooses these internally, according to the ratio of reconstructed signal power to noise power. This is much in the spirit in which the robust weighting approach was originally conceived by Briggs (Briggs 1995b,a).

It should be noted, however, that obtaining all results with extremely high accuracy, especially to produce the uncertainty map, can be more time consuming than traditional imaging methods because of the complicated numerical procedures necessarily involved to solve Eqs. (4.18,4.19).

In this paper, only simulated data was analyzed and the fundamental principles underlying RESOLVE were reviewed. To simplify the analysis, some typical complexities of radio interferometers have been omitted. However, the response operator R (see Eq. 2.9), describing the act of observation, can easily be expanded to cover more effects, thereby adapting to the needs of the actual observational situation.

It is most straight-forward to include the effects of a primary beam, as long as it is known accurately for the instrument in question. Also a direction- or time-dependent point spread function can be included without any further fundamental complication, although computational complexity would be considerably higher.

Furthermore, it should be highlighted that the inclusion of single dish data is almost readily possible. A radio interferometer is not sensitive to the largest scales of the sky brightness because it cannot measure at arbitrarily small uv -points, leaving a gap in the center of the uv -plane. This problem can in principle be overcome by combining the radio interferometric data with single dish observations on the same source. Always, when using CLEAN-derived imaging algorithms, there is a problem with the choice of the correct restoring beam, since it is not possible to just trivially use the point spread function of the radio interferometer for the combined data. There is no such problem with the imaging approach presented in this work.

The extension to multi-frequency synthesis (see Eq. 2.10) and polarization imaging is already being worked on and will be the subject of upcoming publications.

Another future topic is the possible inclusion of calibration into the framework. A first step could be to include the calibrational errors into the error budget and use an approach similar to the extended critical filter (Oppermann et al. 2011b), where the noise covariance is subject to the inference itself. In principle, calibration itself can be understood as a reconstruction problem for which the presented methods could be useful. In the long run,

the distinction between *calibration* and *imaging* is somewhat artificial and should ideally be merged into one step of complete reconstruction (see also Smirnov 2011a,b).

Finally, a future goal should be to extend the imaging algorithm RESOLVE to a broader approach that can handle diffuse emission and point sources simultaneously (see e.g. Selig and Enßlin 2013, for an example from photon count imaging). It could be worthwhile to think about merging the approaches of compressed sensing, where optimal imaging strategies for sparse signals are already known, with the presented Bayesian approach into which they could be included in form of a Laplacian prior.

Chapter 3

A first application of RESOLVE: The galaxy cluster Abell 2256

3.1 Introduction: Galaxy clusters as an optimal field of application for RESOLVE

RESOLVE is optimized for the imaging of extended radio sources. This chapter focuses on an application on radio observations of galaxy clusters, in particular on diffuse non-thermal radio emission found in the inter-cluster medium (ICM). First results of RESOLVE on the galaxy cluster Abell 2256 are presented in Sec. 3.2.

Typical radio astronomical applications include such different objects as diffuse emission from the Milky Way, supernova remnants, galactic radio halos, lobes of radio galaxies and galaxy clusters.

Some galaxy clusters are known to host very distinct extended non-thermal radio structures that cannot unambiguously be ascribed to single member galaxies and are thus identified with the ICM. This extended emission is mainly divided into two different types¹, called *radio halos* and *radio relics*. The former are roughly spherical regions of weak diffuse synchrotron emission, usually centered on the gravitational center of clusters, whereas the latter are distinct, often elongated sources of stronger synchrotron emission found at the edges of clusters. The exact physical origins of these sources are still debated, especially for the halos. The theoretical paradigm explains them to be the results of cluster merger activities, but with different origins. Relics are widely accepted to trace shock fronts of former cluster merger collisions. Halos are less well understood, they must be the result of some kind of reheating or injection mechanism of relativistic electrons in the ICM, but this is still subject to ongoing research and scientific debate (for a review see Feretti et al. 2012). By far not all observed galaxy clusters are also known to host halos and relics. But since in particular the halo emission usually is of low surface brightness compared to other bright compact or point sources in cluster fields, very sensitive observations with

¹Further subclassifications exists, which are of minor importance for this work.

high resolutions and exact data analysis are needed to detect and accurately image diffuse cluster emission. In many cases, high fidelity observations on clusters have only become accessible recently, mainly through the new, upgraded VLA, which is considered to be the best instrument for this kind of science in the moment (see Feretti et al. 2012). It is thus far from clear, whether halos and relics really only occur in particular environments, or are common to many more clusters but are below our to-date detection capabilities.

This is an optimal field of application for RESOLVE², since weak diffuse structure is known to be a problem for the classical imaging method CLEAN, and RESOLVE is constructed to be optimal for this type of sources.

3.1.1 The galaxy cluster Abell 2256

In this chapter, a first application of RESOLVE on real data is presented. The chosen target is the galaxy cluster Abell 2256. The data is from a VLA observation in its D-configuration at 1.369 GHz, taken in 2006 (Clarke and Enßlin 2006). The courtesy of all data goes to T.A. Clarke to whom the author expresses his thanks. Two CLEAN images of the cluster using the same data set are shown in Fig. 3.1. The first shows the full cluster field, the second only the diffuse emission after careful subtraction of point sources and cluster galaxies. In these images, the cluster hosts a prominent relic structure to the north-west of the image center, and a relatively small and weak halo roughly centered a bit to the south-east of the cluster center (better visible in the second image without the other stronger point-like structures in the field). The images cover a region of roughly 2.6×2.5 Mpc. The pixel size is 15 arcseconds in both images, but the resolution is slightly degraded for the diffuse emission image.

3.2 Results

Fig. 3.2 shows an image of Abell 2256 obtained with RESOLVE in its first iteration cycle, using the data from Clarke and Enßlin (2006). The reconstruction was conducted only using the visibilities corresponding to the diffuse emission image shown in Fig. 3.1. Since RESOLVE is only optimal for extended emission, strong point sources in the field can significantly degrade the quality of the reconstruction with RESOLVE. For comparison, a MS-CLEAN image is shown as well, obtained for the exact same settings. For the CLEAN image, the full set of visibilities was used, since experiences showed that this leads to a reconstruction with less subtraction artifacts. The stronger point-like emission has been clipped away so that only the emission on comparable flux scales is shown. The background of the CLEAN image is reduced to a minimal, very small threshold to suppress typical reconstruction artifacts, like regions of negative flux. No such procedure needed to be applied to the RESOLVE image. The pixel size in both images is roughly 7.2 arcseconds.

Currently, these are only preliminary results. In practice, the proper estimation of the signal power spectrum proves to be numerically more complicated than in the simulated

²Regardless of whether new high-quality data is at hand or older data sets are re-analyzed.

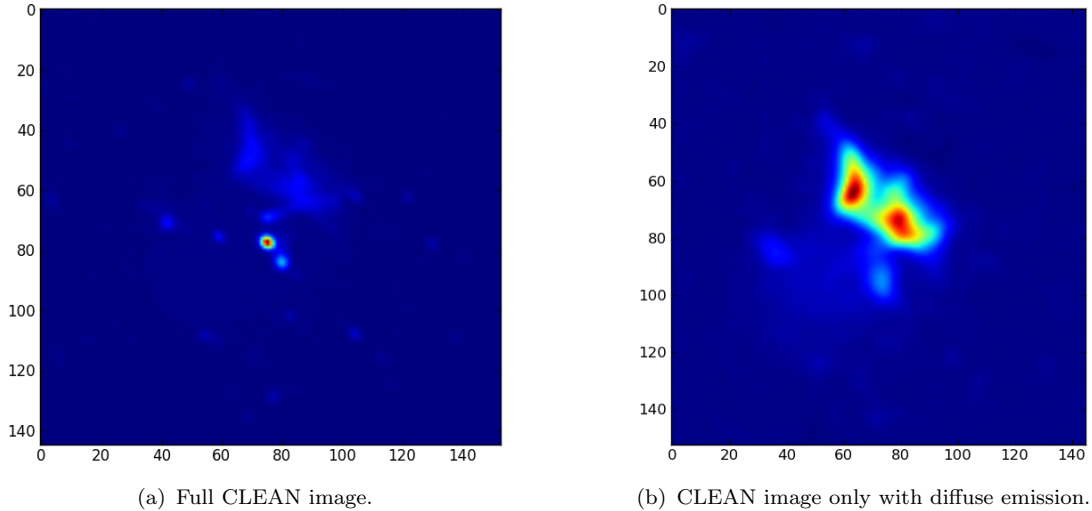
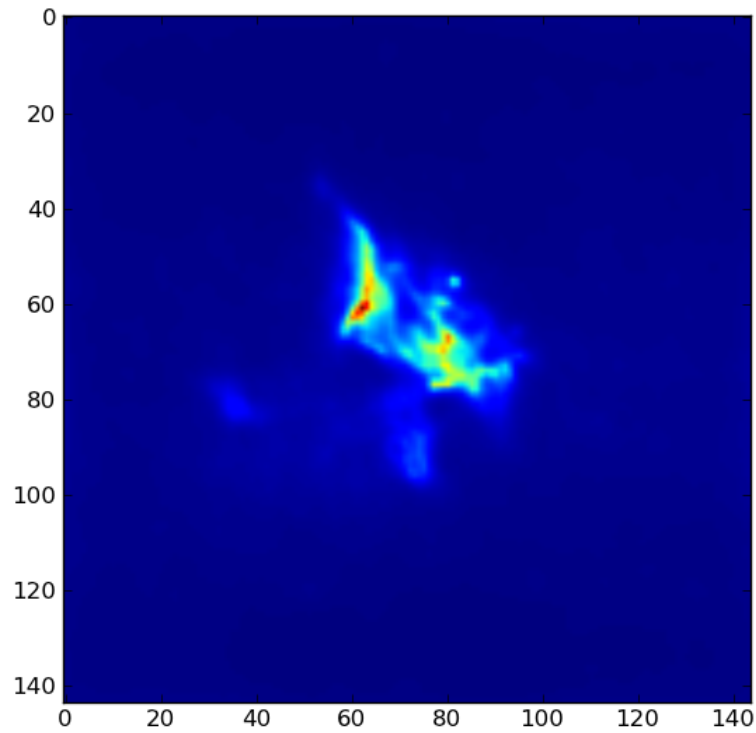


Figure 3.1: CLEAN images of Abell 2256 from an observation taken with the VLA in D-configuration at 1369 GHz. The pixel size is 15 arcseconds. The first image shows the full cluster field, the second only the diffuse emission after subtraction of point and foreground sources. Image courtesy goes to T. A Clarke.

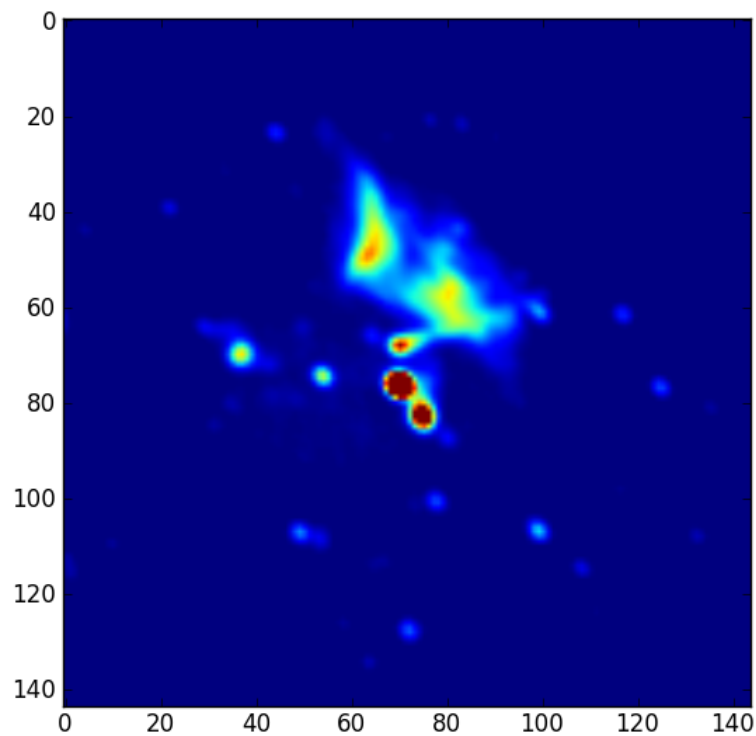
test cases shown in Ch. 2. To estimate the power spectrum, an implicit inversion of the operator $D^{-1} \approx \frac{\delta^2 H(s)}{\delta s_x \delta s_y} \Big|_{s=m}$ needs to be calculated with a conjugate gradient algorithm (see Ch. 2 for details). This is not guaranteed to work always stably in a non-linear regime. Fig. 3.2 thus only shows the first iteration cycle before the first intermediate power spectrum reconstruction,

Nevertheless, this comparison clearly indicates that RESOLVE can outperform MS-CLEAN also in a real data application. In the presented case, RESOLVE is capable to reveal structures on smaller scales than the CLEAN image. In a first simple estimation, it also detects more overall total flux in the halo. Within roughly the same circular halo region as defined by Clarke and Enßlin (2006) while replacing the stronger relic emission with the mean value from the rest of the halo region, RESOLVE finds total flux values of the order of 160 mJy in comparison to 103.4 ± 1.1 mJy, as found by Clarke and Enßlin (2006) for the same data. This hints toward a possibly important result, since it could mean that the weak halo of Abell 2256 is actually much stronger (and possibly larger) at 1.369 GHz than previously thought³. This would also change the halo spectrum, with possible implications for different halo formation scenarios (Ferretti et al. 2012). However, to confirm this, of course, more solid results using RESOLVE are needed.

³Which could be true also for other clusters.



(a) RESOLVE image.



(b) CLEAN image.

Figure 3.2: Comparison between the first iteration cycle reconstruction of Abell 2256 from RESOLVE (*top*) and MS-CLEAN (*bottom*) for the data presented in Sec 3.1.1. The pixel size is roughly 7.2 arcseconds. Note that RESOLVE was run on point-source subtracted data.

Chapter 4

A new approach to multi-frequency synthesis in radio interferometry¹

4.1 Introduction

In radio astronomy, multi-frequency observations are widely used for many different purposes. Examples include the investigation of spectral line emission, analyzing continuous synchrotron spectra of radio sources, the determination of Faraday rotation in polarization imaging, correlating brightness structures at largely different wavelengths or improving the sensitivity and coverage of an interferometer without introducing more antennas into the array (for a review see e.g. Taylor et al. 1999). Most of these only become possible with observations that span over many frequencies and data analysis methods that can handle this kind of observations.

For the remainder of this work, we focus entirely on multi-wavelength radio continuum studies in total intensity. Neither spectral line observations, polarization imaging, nor wavelengths studies beyond pure radio observations will be a direct topic. In Sec. 4.5, we will comment on possible extensions of the presented work into other domains of multi-frequency radio astronomy.

Historically in radio interferometry, the term *multi-frequency synthesis* is mainly used to denote techniques that focus on the direct combination of single-instrument observations at different frequencies to improve the resolution and sensitivity of a radio interferometer (see Conway et al. (1990) and references therein). For this purpose, a number of methods have been devised, most notably double deconvolution (Conway et al. 1990) and multi-frequency CLEAN (Sault and Wieringa 1994). These methods usually work by Taylor-expanding a spectral model function around a reference frequency. We will comment on these methods in more detail in Sec. 4.2.1.

Conversely, investigating the spectral behavior of a radio synchrotron source on its own

¹Note: This chapter has been submitted as a paper to *Astronomy and Astrophysics*

usually is achieved through totally independent standard imaging of the surface brightness in the sky at a number of frequencies. Resolution is kept uniform over all frequencies to correctly recombine the images at individual frequencies. Subsequently, the spectral parameters are determined by fitting a function through all the single frequency images, usually assuming a power-law shaped spectral evolution.

Only recently, implementations of MF-CLEAN also make it possible to constrain the spectral properties (see Reid and CASA Team 2010), effectively starting to merge both multi-frequency high resolution imaging and spectral analysis into one algorithm. It is this combined notion, how we like to understand and use the term multi-frequency synthesis.

Further development of multi-frequency imaging techniques is paramount for being able to fully exploit the data from the new generation of radio telescopes, such as the upgraded VLA, LOFAR, the SKA pathfinder missions or ultimately the SKA itself (see e.g. Garrett 2012). Their unprecedented, broadband frequency coverages of many GHz of possible bandwidths and previously unknown frequency regimes offer many advances in astrophysical and cosmological sciences. But at the same time, they are a challenge for current data analysis methods (see Sec. 4.2.1). One part of the challenge rises from the fact that current algorithms might not meet the expected sensitivity or fidelity because previously acceptable models and approximations break down with the quality and quantity of data now available. The other side is that the huge quantities in which new data sets come, represent a huge computational challenge on their own, regardless of the algorithms used. It should be emphasized that this study focuses entirely on the first part of the challenge (for this topic, see also Sec. 4.5).

In this paper, we present a new approach that uses Bayesian statistical inference techniques to combine a *simultaneous* statistical estimation of the sky brightness and the spectral behavior with the benefits in resolution and sensitivity of classical multi-frequency synthesis. Conceptually, this new approach has a number of advantages over standard methods. In principle, the full bandwidth of the observation is used for maximum theoretical sensitivity. No subsequent imaging at all single frequencies is employed and, thereby, less reconstruction artifacts are introduced and no artificial downgrading of resolution is necessary. Our model for the spectral behavior is not approximated through Taylor-expansions around a frequency, and in the moment only is limited by possible higher order effects like spectral curvature. In this way, our approach is conceptually very similar to the new method of *Faraday Synthesis* (Bell and Enßlin 2012) in polarization imaging (see Sec. 4.2.2 for more details). Furthermore, the spatial correlation of the parameters describing the spectral behavior is used to constrain and improve the estimation of spectral properties, and can further be viewed as a new scientific result on its own. Finally, it is possible to approximate the statistical uncertainty of the spectral index estimate in addition to that of the total intensity.

To demonstrate the viability of our new approach, we present a multi-frequency extension to the radio extended emission imager RESOLVE (see Ch. 2 and Junklewitz et al. (2013)) as an alternative to multi-scale-multi-frequency CLEAN (Rau and Cornwell 2011), the standard method for wide-band multi-frequency synthesis of extended radio sources.

For our derivations, we will often refer to the work presented in Ch. 2 and Junklewitz

et al. (2013).

4.2 Theory & Algorithm

In this section, we first briefly review the current status of multi-frequency synthesis techniques (Sec. 4.2.1), then develop the new approach (Sec. 4.2.2) based on *information field theory* (Enßlin et al. 2009) and the inference framework presented in Ch. 2. We further present a multi-frequency-extension to the imaging algorithm RESOLVE (see Ch. 2) using the developed framework (Sec. 4.2.2).

4.2.1 Multi-frequency aperture synthesis and spectral index reconstruction

Aperture synthesis is the technique of connecting an array of telescopes in such a way that we can effectively synthesize a combined instrument with a much larger aperture and therefore resolution (Ryle and Hewish 1960; Thompson et al. 1986). It can be shown that for observations at a single frequency ν_0 , and under the assumption of measuring the sky as flat in a plane tangent to the phase center of the observation, such a radio interferometer approximatively takes incomplete samples of the Fourier transformed brightness distribution in the sky (Thompson et al. 1986):

$$V(u, v, \nu_0) \approx W(u, v, \nu_0) \int dl dm I_0(l, m, \nu_0) e^{-2\pi i(ul+vm)}. \quad (4.1)$$

For our purposes, working under this assumption suffices for the analysis undertaken in this paper. For a detailed derivation of (4.1) consider (Thompson et al. 1986). The quantity $V(u, v)$ is called the *visibility* function. The coordinates u and v are vector components describing the distance between a pair of antennas in an interferometric array, where this distance is usually referred to as a *baseline*. They are given in numbers of wavelengths, with u and v usually parallel to geographic east-west and north-south, respectively. The coordinates l and m are a measure of the angular distance from the phase center along axes parallel to u and v , respectively. $W(u, v)$ is a sampling function defined by the layout of the interferometric array. It is zero throughout most of the u, v -space, apart from where measurements have been made where it is taken to be unity.

The visibility function is what our instrument measures, but we are actually interested in the brightness distribution of the source in the sky. Unfortunately, an inversion of (4.1) gives us not the true brightness distribution, but its convolution with the inverse Fourier transform of the sampling function, better known as the *dirty beam* $I_{db} = \mathcal{F}^{-1}W$:

$$I_D = \mathcal{F}^{-1}V = \mathcal{F}^{-1}WFI = I_{db} * I. \quad (4.2)$$

Here, we have introduced a symbolic Fourier operator $\mathcal{F}_{kx} = \exp(-i(ul + vm))$ with $x = (l, m)$ and $k = (u, v)$, the common notation I_D , *dirty image*, for the simple Fourier inversion of the visibilities, and the symbol $*$ to denote a convolution operation.

For imaging at a single frequency, one usually proceeds using a deconvolution algorithm aiming to solve (4.2) approximatively. The most common algorithm is CLEAN (Högbom 1974), or one of its many variants (Clark 1980; Schwab 1984; Sault and Wieringa 1994; Cornwell 2008; Rau and Cornwell 2011), which basically model the sky brightness to be a collection of delta peak point sources that are iteratively assigned by searching for the peak values in the dirty image. But alternatives exist, especially for imaging extended emission, like the Maximum Entropy Method (Cornwell and Evans 1985), Adaptive Scale Pixel decomposition (Bhatnagar and Cornwell 2004) or the recently published Bayesian extended emission imager RESOLVE (see Ch. 2), which will be used in this paper in an expanded multi-frequency version (see Secs. 4.2.2 and 4.4).

Originally, radio telescopes observed with relatively narrow bandwidths and at only a few different frequencies. The standard approach for imaging the spectral properties of a source since then is to take single frequency observations, or observations averaged over close channels for higher sensitivity, image them separately, and then fit a spectral model to the observations. Since continuous synchrotron emission is known to show a power-law spectrum (Rybicki and Lightman 1985), for many astrophysical purposes such a model is sufficient:

$$I(l, m, \nu) = I_0(l, m, \nu_0) \left(\frac{\nu}{\nu_0} \right)^{-\alpha}. \quad (4.3)$$

Sometimes, higher order spectral deviations are modeled with a second term in the exponential of (4.3):

$$I(l, m, \nu) = I_0(l, m, \nu_0) \left(\frac{\nu}{\nu_0} \right)^{(-\alpha + \log \nu / \nu_0 \beta)}, \quad (4.4)$$

effectively enhancing the linear function in $\log I$ vs. $\log \nu$ space (4.3) to a quadratic polynomial. This model is referred to as spectral curvature.

Simultaneous multi-frequency imaging at different frequencies has been introduced to radio astronomy upon the realization that observations at many frequencies, if stacked together appropriately, can improve the sampling in the uv -plane (see Conway et al. (1990) and references therein). That way, the sensitivity of a radio interferometric observation can be enhanced considerably. Because the uv -coordinates are measured in numbers of wavelengths, the same interferometer samples different parts of the Fourier space at different frequencies. An interferometer includes $N(N - 1)/2$ baselines and therefore uv points, where N is the number of antennas. In theory, observing at a number of frequencies N_f enhances the measured baselines to $N_f N(N - 1)/2$, which is the equivalent effect of introducing roughly $\sqrt{N_f}$ extra antennas (Conway et al. 1990).

This approach was first developed by Conway et al. (1990) for mid-sized fractional bandwidths of around $\pm 10\%$, together with the method of double deconvolution to mitigate

spectral errors when using CLEAN for a spectrally combined data set. Later, double deconvolution was developed into the more advanced multi-frequency CLEAN (MF-CLEAN) (Sault and Wieringa 1994) and multi-frequency multi-scale CLEAN (MF-MS-CLEAN) (Rau and Cornwell 2011).

All these methods assume the spectral dependence to be a power-law, like Eqs. (4.3) or (4.4) and propose to approximate it during the CLEAN-like-deconvolution process with a Taylor-expansion. To our knowledge, current implementations use a few terms of a Taylor-expansion in ν or $\log \nu$ around a reference frequency ν_0 . In Rau and Cornwell (2011) it is discussed that a direct decomposition into one or two terms of a polynomial in $\log I$ vs. $\log \nu$ would actually be the most accurate representation², but for current implementations this is discarded by the authors because of numerical instabilities.

During the CLEAN deconvolution process, the iteratively updated sky model is used to also update the coefficients of the spectral Taylor expansion. In this way, spectral index or even spectral curvature can be constrained by these coefficients. Care must be taken, since this expansion is usually stopped after a few terms and might not be valid over large bandwidths. Because of this, the only implementation of MF-MS CLEAN known to the authors to date, within the radio astronomical software package CASA (Reid and CASA Team 2010), is considered to be experimental and on a shared-risk basis with regard to the spectral reconstructions with a higher number of terms.

For modern wide-band data sets from the new generation of instruments, spanning several GHz of bandwidths, this means that probably even more emphasis must be put onto the handling of the spectral effects, either by invoking higher order terms in the expansions (see Sault and Wieringa 1994; Rau and Cornwell 2011) or by shifting to a different approach, where the sky brightness and its spectral properties are fully considered simultaneously, and estimated to fit the entire data using some global minimization function. This last approach was actually mentioned by Conway et al. (1990), but found to be unnecessarily complicated and computationally expensive for the typically modest fractional bandwidths at the time.

We will now consider the latter approach and present a statistical solution.

4.2.2 A multi-frequency extension to the RESOLVE algorithm

In the course of this section, we often refer to the detailed derivations layed out in Sec. 2 of Ch. 2 that form the basis from which we derive our multi-frequency algorithm.

We start by summarizing (4.1) and (4.3) into a multi-frequency measurement model:

$$d(k, \nu) = W(k, \nu) \int dx \mathcal{F}(k, x) I(x, \nu) + n(k, \nu) \quad (4.5)$$

where again $x = (l, m)$ and $k = (u, v)$, the term $n(k, \nu)$ introduces measurement noise, the data $d(k, \nu)$ has been introduced, which is basically the visibility function with mea-

²Actually, for the spectral model (4.4), this would be no approximation at all since it simply is a quadratic polynomial in $\log I$ vs. $\log \nu$ space.

surement noise, and the spectral dependence of the source $I(x, \nu)$ is kept general for the moment.

In order to simplify notation and to identify (4.5) as a general inverse inference problem as analyzed in Ch. 2, we henceforth drop all explicit dependence on k and x and combine all known instrumental effects into a response function $R_\nu := W(k, \nu)\mathcal{F}$, leading to

$$d_\nu = R_\nu I_\nu + n_\nu. \quad (4.6)$$

There are many instrumental effects beyond the sampling in the uv -plane given by $W(k, \nu)$ like an antenna sensitivity pattern or an direction dependent, variable sampling. For this work, we stay with the basic definition $R_\nu := W(k, \nu)\mathcal{F}$ and refer the reader to Ch. 2 concerning the possibility of including other effects within this framework. We just emphasize that, without loss of generality, most of these effects can be in principle included³. We also assume the instrument to be fully calibrated, and thus the response R_ν to be known.

In accordance with standard radio interferometric literature (Thompson et al. 1986), we assume Gaussian noise statistics, mainly induced by the antenna electronics and independent between measurements at different frequencies and time steps of the observation (Thompson et al. 1986). Henceforth, the noise n_ν will be assumed to be drawn from a multivariate, zero mean Gaussian distribution of dimension n_d :

$$\begin{aligned} \mathcal{P}(n) &= \mathcal{G}(n, N) \\ &:= \frac{1}{\det(2\pi N_\nu)^{1/2}} \exp\left(-\frac{1}{2}n_\nu^\dagger N_\nu^{-1}n_\nu\right). \end{aligned} \quad (4.7)$$

Solving (4.6) exactly for I_ν is not possible, since all information is lost on the Fourier modes not sampled by R_ν . This is just a different way of stating that a direct Fourier inversion of (4.1) yields the dirty image and not an exact representation of the sky brightness.

Instead, the aim is to find a statistical estimate for the *most probable* sky brightness signal, given all observational and noise constraints. In Ch. 2, it was shown in detail how this can be done in a Bayesian statistical framework for a radio astronomical data model like (4.1). We briefly repeat the main points, and otherwise refer the reader to Ch. 2.

To find an optimal statistical estimate for the sky brightness signal I_ν , we regard it as a random field with certain *a priori* statistical properties expressed in the prior distribution $P(I)$, but fully constrained by the data through the statistics of the *likelihood* distribution $P(d|I)$. The likelihood distribution summarizes how the data are obtained with a measurement of the true sky brightness signal, and for our problem can be expressed as

$$\mathcal{P}(d|I) = \mathcal{G}(d_\nu - R_\nu I_\nu, N_\nu), \quad (4.8)$$

³In principle this approach can be extended into a full RIME (Radio Interferometer Measurement Equation), considering all Stokes parameters and instrumental gains (Smirnov 2011a,b).

which is a Gaussian over $d_\nu - R_\nu I_\nu$ with the covariance structure of the uncorrelated noise $N_{kk'\nu} = \delta_{kk'}\sigma_{k\nu}^2$.

Prior and likelihood statistics can be combined into the *posterior* distribution $P(I|d) \propto P(d|I)P(I)$ that holds the important information of how much the sky brightness signal is statistically constrained by the data. From there, an estimate for the signal can be obtained by calculating a suitable statistic of the posterior, most prominently its mean or its mode, corresponding to the minimization of different error norm measures between the signal and its estimate (see Ch. 2; or Jaynes (2003), Caticha (2008) and Lemm (1999), Enßlin et al. (2009) for a comprehensive review of Bayesian statistics or inference on fields respectively).

The exact choice of an appropriate inference algorithm at this stage largely depends on the complexity of the problem (i.e. the posterior). For the problem at hand, since the likelihood is already known (4.8), this comes down to the question of the prior statistics of I_ν .

The most general way conceivable would be *not* to explicitly model the spectral dependence of $I(x, \nu)$ at all, as for instance in (4.3) or (4.4). Instead, $I(x, \nu)$ can be interpreted as a three dimensional continuous field and should be inferred as a whole from the entire data. In general, let us assume not only point-like but also extended emission in the sky, and some kind of extended structure in spectral space as well. In such a setting, $I(x, \nu)$ could be set *a priori* as a statistical field with an unknown and probably non-isotropic cross-correlation structure in the combined sky and spectral space. Such a complex, unknown cross-correlation structure at the one hand complicates the problem enormously, but could also be used to guide the reconstruction, if correctly estimated with the field itself.

A number of statistical methods have already been developed to handle simpler problems of signal reconstructions with unknown but isotropic correlation structure, many of them solving the problem for Gaussian fields using information field theory (Enßlin and Frommert 2011; Enßlin and Weig 2010) or Gibbs-Sampling Monte-Carlo methods (Sutter et al. 2012; Karakci et al. 2013), or recently also for log-normal fields (Enßlin and Frommert 2011; Enßlin and Weig 2010; Oppermann et al. 2013; Greiner 2013; Selig and Enßlin 2013). Most notably this method was also used to create RESOLVE (see Ch. 2). A full combined spatial and spectral reconstruction as outlined above, would require substantial further development and is outside of the scope of this work.

We can choose a more direct strategy, still residing within our approach of statistical inference, but fix a spectral model and infer instead the spectral index (or curvature) as a field on its own. For the rest of the paper, we will choose the model (4.3) for the simplicity of the approach, and for a functional similarity with the algorithm RESOLVE that makes it very natural to include into a combined method.

RESOLVE works under the assumption that the extended surface brightness at a single frequency is *a priori* assumed as a random field drawn from log-normal statistics (see Ch. 2). For our multi-frequency problem, this basically turns (4.6) into

$$\begin{aligned}
 d_\nu &= R_\nu I_\nu + n_\nu \\
 &= R_\nu \left[I_0(l, m, \nu_0) \left(\frac{\nu}{\nu_0} \right)^{-\alpha} \right] + n_\nu \\
 &= R_\nu \left[\rho_0 e^{s(l, m)} \left(\frac{\nu}{\nu_0} \right)^{-\alpha} \right] + n_\nu,
 \end{aligned} \tag{4.9}$$

where s is a Gaussian random field (such that the logarithm of e^s is a Gaussian random field again), and ρ_0 is a constant to e.g. normalize the system to the right units. Although, the frequency dependence of s was not explicitly shown in Ch. 2, the derivation of RESOLVE implicitly assumed the algorithm to work for a single frequency in the way presented here. RESOLVE assumes the spatial correlation of an extended source in the sky to be reflected by the covariance of the Gaussian random field s , which is unknown *a priori*, and thus estimated from the data itself together with the sky brightness. The covariance S of a Gaussian field is equivalent to its two-point correlation function $S(l, m) = \langle s(l)s(m)^\dagger \rangle$ and is handled as such by RESOLVE in form of the power spectrum $S(k, k') = \langle s(k)s(k')^\dagger \rangle$, which is the Fourier transformation of the correlation function⁴. A deeper analysis of this can be found in Ch. 2.

We now make the central assumption that the spectral index α can be modeled *a priori* as a Gaussian random field with its own spatial correlation structure in the sky. At least for an extended source, we have every reason to assume that the spectral index should be a field with spatial extension itself. Observational constraints strongly imply that typical extended radio structures show as well extended and smooth spectral index structures, for instance radio halos and relics of galaxy clusters (Feretti et al. 2012), radio galaxy lobes (Kassim et al. 2005), or supernova remnants (Green 2001).

In Ch. 2, we argued extensively that a Gaussian random field would be the ideal choice for a signal prior of an extended field with a priori unknown correlation structure, as long as the field is not assumed to vary strongly on orders of magnitude and not necessarily needs to be positive definite. Both constraints apply very well to known spectral index maps, where variations usually do not reach even one order of magnitude, and nothing prevents the spectral index in principle to change the sign. For details of these arguments see Ch. 2, we now proceed under this assumption.

If we rewrite (4.3) only slightly,

$$\begin{aligned}
 I(l, m, \nu) &= I_0(l, m, \nu_0) \left(\frac{\nu}{\nu_0} \right)^{-\alpha} \\
 &= I_0(l, m, \nu_0) e^{-\ln(\nu/\nu_0)\alpha},
 \end{aligned} \tag{4.10}$$

⁴We actually assume the spatial correlation to be *a priori* rotationally and translationally invariant, and thus the power spectrum to be diagonal $S(k, k') = \langle s(k)s(k')^\dagger \rangle = (2\pi)^{n_s} \delta(k - k') P_s(|k|)$. However, this does not imply that the correlation structure must be invariant under any transformation *a posteriori* as well. A more detailed discussion can be found in Ch. 2.

it reveals that, if α has a Gaussian prior, (4.10) naturally turns into a model for a log-normal prior, only different in shape (and more complicated) from (4.9) because of the term $-\ln(\nu/\nu_0)$.

It is important to note that we have not specified $I_0(l, m, \nu_0)$ yet. Thus, at this point, our inference approach to multi-frequency synthesis is in principle compatible with any method that reconstructs and deconvolves the surface brightness $I_0(l, m, \nu_0)$ at a single reference frequency ν . At least as long as it seems consistent with the source of interest to assume that the spectral index is an extended and spatially correlated field. In an extreme case, like single, unresolved point sources, this method probably will not yield optimal results.

For this paper, we take the choice to combine (4.9) and (4.10) into one single method, where we assume our double log-normal measurement model to be

$$d_\nu = R_\nu [\rho_0 e^{s(l,m) - \ln(\nu/\nu_0)\alpha(l,m)}] + n_\nu, \quad (4.11)$$

with ρ_0 again a constant, from now on set to one w.l.o.g., and the signal fields s and α having Gaussian prior distributions $P(s)$ and $P(\alpha)$:

$$\begin{aligned} \mathcal{P}(s) &= \mathcal{G}(s, S) \\ &= \frac{1}{\det(2\pi S)^{1/2}} e^{-\frac{1}{2} s^\dagger S^{-1} s}, \end{aligned} \quad (4.12)$$

$$\begin{aligned} \mathcal{P}(\alpha) &= \mathcal{G}(\alpha, A) \\ &= \frac{1}{\det(2\pi A)^{1/2}} e^{-\frac{1}{2} \alpha^\dagger A^{-1} \alpha}. \end{aligned} \quad (4.13)$$

We now write down a posterior distribution for each signal field, while the other field (and its covariance) is assumed to be known, held constant and regarded as part of two distinct versions of the response operator R_ν , henceforth called $R_{(s)} = W(k, \nu)\mathcal{F}(e^{-\ln(\nu/\nu_0)\alpha} \circ)$ and $R_{(\alpha)} = W(k, \nu)\mathcal{F}(e^{s(l,m)} \circ)$ (where the symbol \circ denotes where the field needs to be inserted that the operator acts on):

$$\mathcal{P}(s|d) \propto \mathcal{G}(d - R_{(s)}e^s, N) \mathcal{G}(s, S), \quad (4.14)$$

$$\mathcal{P}(\alpha|d) \propto \mathcal{G}(d - R_{(\alpha)}e^{-\ln(\nu/\nu_0)\alpha}, N) \mathcal{G}(\alpha, A). \quad (4.15)$$

As in Ch. 2, it is not possible to calculate the posterior mean for either of the two signal fields s or α without invoking complicated or expensive perturbation or sampling methods (see Ch. 2). We therefore continue to use the procedure already presented there, and calculate the posterior maximum to estimate both signals

$$\begin{aligned}
 m_s &= \operatorname{argmax}_s \mathcal{P}(s|d), \\
 m_\alpha &= \operatorname{argmax}_\alpha \mathcal{P}(\alpha|d).
 \end{aligned}
 \tag{4.16}$$

In signal inference, this procedure is called Maximum A Posteriori (MAP) (Jaynes 2003). The resulting fix-point equations from Eqs. (4.16) need to be solved numerically using a non-linear optimization scheme. For this, we resort to the same implementations as in Ch. 2.

With this choice, we basically extend RESOLVE to a multi-frequency algorithm by integrating a second complete RESOLVE step for the spectral index into the method, and iterating between the statistical estimation of s with its covariance S , and α with its covariance A . As outlined above, we always hold one of the fields (and their respective covariances) as constant and regard them as part of the response during this process.

It should be emphasized again that the RESOLVE step for the spectral index could in principle be combined with any other method to reconstruct $I_0(l, m, \nu_0)$ at a single frequency ν_0 .

The exact equations that need to be solved to calculate (4.16) for either of the two fields and estimate their power spectra (i.e. their correlation structure in form of their Gaussian covariances, see above) are laid out in App. 4.3 and derived rigorously in Ch. 2.

4.3 Details of the algorithm

In this section, we repeat some of the basic derivations from Ch. 2, and derive the details of how reconstructing the spectral index using RESOLVE differs from the standard procedure presented there. In principle, many of the original derivations are still valid for multi-frequency RESOLVE, estimating s and α . The derivations for power spectrum reconstructions and uncertainty calculations stay especially close to the equations already presented in Ch. 2, and we do not repeat them explicitly here.

4.3.1 Reconstruction of the sky brightness signal field s

The estimation of the sky brightness at a reference frequency $I_0(l, m, \nu_0) = \rho_0 e^{s(l, m)}$ can mostly be conducted with the standard single-frequency RESOLVE.

The only difference is the more complex response operator $R_{(s)} = W(k, \nu) \mathcal{F} (e^{-\ln(\nu/\nu_0)\alpha} \circ)$. Dropping the explicit signal index for a moment and writing the specific measurement model (4.11) for the signal s with explicit operator and field indices

$$d_{k\nu} = R_{k\nu x} (\rho_0 e^s)_x + n_{k\nu},
 \tag{4.17}$$

reveals that the operator $R_{k\nu x}$ actually spans the single-frequency Fourier transform over all observed frequencies into a large, many-frequency data space⁵. Conversely, the adjoint response R^\dagger now includes a sum over all frequencies to collapse everything back into the two dimensional sky at reference frequency. Since both operations are used in RESOLVE, it is by this procedure that the multi-frequency RESOLVE effectively uses uv -information at all frequencies to constrain the estimation of s . Of course, the quality of this reconstruction depends on the accuracy of the current estimation for α , used in the reconstruction step for s to define $R_{(s)}$.

4.3.2 Reconstruction of the spectral index α

The log-normal model for the spectral index, $e^{-\ln(\nu/\nu_0)\alpha}$, only differs in the term $b = -\ln(\nu/\nu_0)$ from the model used for the sky brightness signal. This slightly complicates the calculations for the signal estimation with MAP and the power spectrum reconstruction and eventual uncertainty calculation. For both, derivatives of the posterior (4.15) with respect to α are needed (see Ch. 2), and this will add extra b -terms into the equations. The results of this are summarized in Sec. 4.3.3.

As for the sky brightness signal s (see App. 4.3.1), the response operator for the α reconstructions $R_{(\alpha)} = W(k, \nu) \mathcal{F}(e^{s(l,m)} \circ)$ is more complex. Effectively, the b -term in the basic log-normal model $e^{-\ln(\nu/\nu_0)\alpha}$ would prevent us from just defining our signal space to be the two-dimensional sky since $R_{(\alpha)}$ acts on $e^{-\ln(\nu/\nu_0)\alpha}$. In the actual implementation of the multi-frequency RESOLVE algorithm, we circumvent this problem by assuming that $R_{(\alpha)}$ acts on α and all other terms, including the exponential operation of the log-normal model, are part of the operator⁶. In this way, $R_{(\alpha)}$ and $R_{(\alpha)}^\dagger$ can be understood to act in the same way as their respective counterparts for s do (i.e. they also span up and collapse into the full frequency data space).

4.3.3 Combined algorithm

Using our findings in Ch. 2 and in the previous subsections, multi-frequency RESOLVE comes down to solving iteratively two only slightly different, subsequent sets of equations:

Estimation of s

$$S_p^{-1} m_s + e^{m_s} \cdot M_s e^{m_s} - j_s \cdot e^{m_s} = 0 \quad (4.18)$$

$$(D_s)_{xy} = S_p^{-1} + e^{(m_s)_x} (M_s)_{xy} e^{(m_s)_y} + e^{(m_s)_y} \int dz M_s(x, z) e^{(m_s)_z} - (j_s)_x \cdot e^{(m_s)_x} \delta_{xy} \quad (4.19)$$

⁵As a reminder: In (4.17) we use an implicit convention to sum or taking the integral over repeated discrete or continuous indices, see Ch. 2 for details.

⁶This actually renders $R_{(\alpha)}$ a non-linear operator.

$$p_i^s = \frac{(q_i^s + \frac{1}{2}\text{tr}((m_s m_s^\dagger + D_s)S^{(i)}))}{\left(\alpha_i^{(s-\text{pr})} - 1 + \frac{\varrho_i}{2} + (Tp)_i\right)} \quad (4.20)$$

Estimation of α

$$A_p^{-1} m_\alpha + b e^{bm_\alpha} \cdot M_\alpha e^{bm_\alpha} - j_\alpha \cdot b e^{bm_\alpha} = 0 \quad (4.21)$$

$$(D_\alpha)_{xy} = A_p^{-1}{}_{xy} + b^2 e^{b(m_\alpha)_x} (M_\alpha)_{xy} e^{b(m_\alpha)_y} + b^2 e^{b(m_\alpha)_x} \int dz M_\alpha(x, z) e^{b(m_\alpha)_z} - (j_\alpha)_x \cdot e^{(m_\alpha)_x} \delta_{xy} \quad (4.22)$$

$$p_i^\alpha = \frac{(q_i^\alpha + \frac{1}{2}\text{tr}((m_\alpha m_\alpha^\dagger + D_\alpha)A^{(i)}))}{\left(\alpha_i^{(\alpha-\text{pr})} - 1 + \frac{\varrho_i}{2} + (Tp)_i\right)} \quad (4.23)$$

A rigorous derivation for all equations can be found in Ch. 2. The two sets of equations only differ in form by the $b = -\ln(\nu/\nu_0)$ - terms that show up in the spectral index reconstruction because of the derivatives used in order to calculate the MAP estimate. The quantities j_s , M_s and j_α , M_α are defined as

$$j_s = R_s^\dagger N^{-1} d, \quad (4.24)$$

$$M_s = R_s^\dagger N^{-1} R_s, \quad (4.25)$$

$$j_\alpha = R_\alpha^\dagger N^{-1} d, \quad (4.26)$$

$$M_\alpha = R_\alpha^\dagger N^{-1} R_\alpha. \quad (4.27)$$

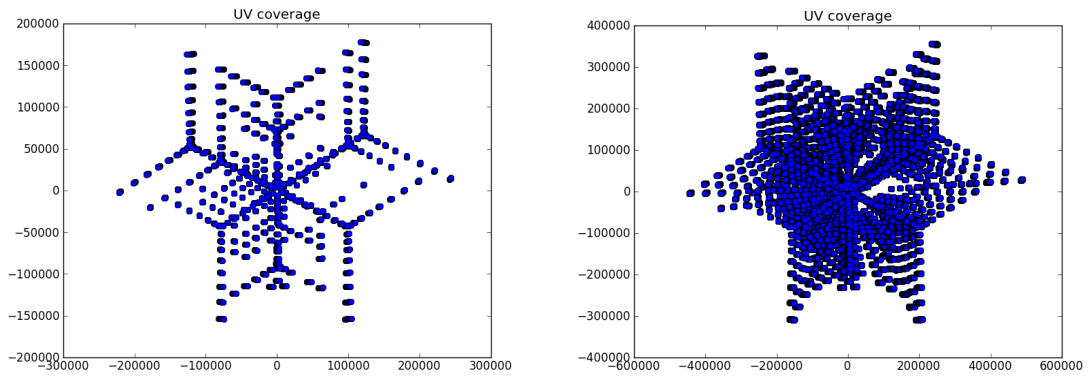
$S^{(i)}$, or $A^{(i)}$ are projection operators onto a band of Fourier modes denoted by the index i , while p_i are parameters to model the unknown power spectrum into a number of such bands $S = \sum_i p_i S^{(i)}$, or $A = \sum_i p_i A^{(i)}$ (see Ch. 2 for details). The quantities q , $\alpha^{(\text{pr})}$ and ϱ are parameters of a power spectrum prior for the signal or the spectral index, and T is an operator, which enforces a smooth solution of the power spectrum p_i .

Eqs. (4.18) and (4.21) are the fix point equations that need to be solved numerically to find a MAP signal estimate m_s or m_α for the current iteration. The second equations (4.19) and (4.22) result from calculating the second derivative of the respective posteriors for the signal estimates m_s or m_α , their inverses serve as an approximation to the signal uncertainty $D_s = \langle (s - m_s)(s - m_s)^\dagger \rangle$ or $D_\alpha = \langle (\alpha - m_\alpha)(\alpha - m_\alpha)^\dagger \rangle$ at each iteration step. The last equations (4.20) and (4.23) represent an estimate for the signal power spectra (and therefore their autocorrelation functions), using the signal uncertainties D_s or D_α to correct for missing signal power in the current estimates m_s or m_α . The iteration is stopped after a suitable convergence criterion is met (see Ch. 2).

4.4 Tests

We have integrated multi-frequency capability into the existing implementation of RESOLVE and tested the algorithm using simulated data⁷. The code is written in PYTHON using the signal inference library NIFTY (Selig et al. 2013), for details of the implementation we refer the reader to Ch. 2.

As in Ch. 2, we constructed simulated observations with the tool MAKEMS⁸ using a realistic uv -coverage from a VLA observation in its A-Configuration. The VLA samples the uv -plane non-uniformly at irregular intervals, and the response includes thereby a convolutional gridding and de-gridding operator using a Kaiser-Bessel kernel (for details see Ch. 2). We simulated observations over a range of 2 GHz, with 20 separate frequency channels. The observations are short snapshots of approximately 20 minutes per frequency, with a total of 42 120 visibility measurements at each frequency channel (see Fig. 4.1). This setting leads to an especially sparse sampling of the uv -plane at a single frequency, but to a much better coverage for the combined multi-frequency data (see Fig. 4.1).



(a) Single frequency uv -coverage in units of # of lengths. (b) Full multi-frequency uv -coverage in units of # of wavelengths.

Figure 4.1: Single-frequency and full-frequency uv -coverage from all frequencies of a simulated 20 minutes snapshot observation in VLA-a configuration.

For all tests, the signal s is drawn from a Gaussian distribution, finally entering the formalism as an exact log-normal field⁹

In order to use a spectral index signal with some correlation to the brightness signal, we model the spatial source dependence of the spectral index in an ad hoc fashion.

⁷To get access to the preliminary code prior to its envisaged public release, please contact henrikju@mpa-garching.mpg.de or enssln@mpa-garching.mpg.de.

⁸See <http://www.lofar.org/wiki/lib/exe/fetch.php?media=software:makems.pdf>.

⁹In Ch. 2 it was demonstrated that RESOLVE also works beyond that on realistic signals drawn from real CLEAN maps.

For the spectral index signal α , we use a sum of two Gaussian fields (see Sec. 4.4.1), one being independently drawn, while the second is the surface brightness signal s itself, down-weighted with a suitable factor. This is done in order to introduce a spectral index signal with some correlation to the surface brightness signal, at least in an ad-hoc fashion. Typically, extended sources in radio astrophysics show a spatial cross-correlation between both, which is a result of different physical processes underlying the structure and formation of these sources. A notable example are radio halos and relics in galaxy clusters (Feretti et al. 2012). Of course, our model is only an ad-hoc approximation for the proof of concept undertaken in this work. We also emphasize that such a cross-correlation between α and s is *not* exploited by RESOLVE (see Sec. 4.2.2 for an outlook).

The complex, Gaussian input noise variance in uv -space is equal for all visibilities and frequencies. As with single-frequency RESOLVE, the algorithm does not require equal noise variances and can in principle handle varying variances. The noise variance was set to a low¹⁰ value of $\sigma^2 = 10^{-3}\text{Jy}^2$.

In continuation with Ch. 2, we use a relative \mathcal{L}_2 - norm measure of the difference in the signal to the reconstruction to measure the accuracy of the estimate of both brightness and spectral reconstructions:

$$\delta_s = \sqrt{\left(\frac{\sum (e^s - e^m)^2}{\sum (e^s)^2}\right)}, \quad (4.28)$$

$$\delta_\alpha = \sqrt{\left(\frac{\sum (\alpha - m_\alpha)^2}{\sum (\alpha)^2}\right)}, \quad (4.29)$$

where the sums are taken over all pixels of the reconstruction. For the motivation behind this choice see Sec. 3 in Ch. 2.

In Sec. 4.4.1 we focus exclusively on the reconstruction of the two signals s and α . Then, in Sec. 4.4.2, we compare the results from multi-frequency RESOLVE with standard imaging procedures. The reconstruction of the signal power spectra is discussed separately in Sec. 4.4.3.

4.4.1 Main test results

We start by showing the reconstruction of the presented simulated observations using multi-frequency RESOLVE. In Fig. 4.3, a surface brightness and a spectral index signal are shown, together with the respective reconstructions obtained with RESOLVE and absolute difference maps of both signals to their reconstructions. We show the spectral index maps in full, and over-layed with a mask that focuses on the part of the observed field that

¹⁰In comparison to the signal strength in Fourier space, the chosen value ensures a high signal to noise ratio. The unit Jy was used here for convenience. Effectively, it stands for whatever units the simulated signal is interpreted to be given in.

contains the brightest part of the surface brightness signal. Later, in Fig. 4.4, we show the reconstruction for different masks (along with a comparison to other methods, see Sec. 4.4.2). We choose the three different masks mainly qualitatively by visual comparison to show only the parts of the sky with a surface brightness signal above a certain threshold. The actual threshold values for the surface brightness were 2 Jy/px, 4 Jy/px, and 6 Jy/px. In Fig. 4.3, the second conservative mask is used. The error measures are $\delta_I = 0.13$ for the surface brightness, and $\delta_\alpha = 0.35$ for the spectral index.

It can be seen that RESOLVE recovers very accurately the original surface brightness. For this particular choice of relative low noise, the algorithm succeeds in reconstructing even small scale features of the signal. The effects of the instrumental point spread function are successfully deconvolved. These findings are in perfect agreement with the results presented on single-frequency RESOLVE in Ch. 2.

The general structure of the spectral index is well reconstructed, even for outer regions where the brightness signal is very weak. Overall, small scale features are much better recovered in the inner regions, where the main brightness sources are located, as can be seen by comparison between the full and the masked images in Fig. 4.3. It is expected that the quality of the spectral index reconstruction depends on the strength of the observed surface brightness, as is illustrated by the poor performance of standard methods in recovering any structures outside the strong source regions, presented in the next section (see Fig. 4.4). In most real applications, the outer parts of the observed fields should therefore usually be not a focus of the investigation. We thus choose the relatively conservative mask in Fig. 4.3 to highlight this important part of the reconstruction. Nevertheless, it should be noted that – at least for this low noise example with relatively high sensitivity due to the broad bandwidth – RESOLVE is able to reliably extrapolate its estimation also into the weaker regions around the main sources.

In addition, with RESOLVE, an uncertainty of the spectral index reconstruction can be estimated. As for the single frequency reconstructions presented in Ch. 2, the second derivative of the posterior is used to approximate its covariance D_α (for details, see App. 4.3 and Ch. 2). This way, the full estimate for the spectral index signal becomes

$$\alpha_x \approx (m_\alpha)_x \pm \left(\sqrt{D_\alpha} \right)_{xx}. \quad (4.30)$$

For the given set of simulated data, roughly 60 % of the original signal values within the unmasked regions in Fig. 4.3 lie within a $1 - \sigma$ interval, but only roughly 20 % lie within a $1 - \sigma$ interval for the full spectral index image. Due to the non-linear nature of the inference problem (4.16), it is expected that the uncertainty estimate is not exactly what would be expected from a pure Gaussian covariance (i.e. 68 % in the $1 - \sigma$ interval), the MAP estimate is not guaranteed to lie very close to the real posterior covariance (see Ch. 2 for this problem). It is no surprise that the estimate worsens for all the outer regions with only very weak brightness structures present, which explains the poor performance on the whole spectral index map. Furthermore, as discussed in Ch. 2, the calculation of an uncertainty estimate is computationally costly. Due to a lack of accessible computer power only for testing purposes, we stopped the calculations at some point and smoothed

Algorithm \ Mask	liberal mask	medium mask	conservative mask
RESOLVE	0.34	0.33	0.32
power-law fit	132.79	0.94	0.48
MS-MF CLEAN	3.07	1.89	1.84

Table 4.1: \mathcal{L}_2 error measures for RESOLVE, the power law fit, and MS-MF CLEAN for the simulation and the reconstruction observation of Sec. 4.4.1 and the three different masks defined in Sec. 4.4.1.

the outcome. This should not be a great problem, since the uncertainty is expected to be smooth and we are mainly interested in a proof of concept at this point.

4.4.2 Comparison to standard methods

In this section, we compare the results of multi-frequency RESOLVE with two standard methods: A straight forward power-law fit for (4.3) using single-frequency CLEAN images, and a MF-MS-CLEAN reconstruction for the surface brightness and the spectral index. The reconstructions were performed on the same simulated observation as in Sec. 4.4.1. All results were obtained using the radio-astronomical software package CASA (Reid and CASA Team 2010). The MF-MS-CLEAN reconstructions were obtained using 1500 iterations, a small gain factor of 0.1, uniform weighting and ten different multi-scales ranging from a single pixel to moderately large structures. We further used two terms for the Taylor expansion in ν (see Sec. 4.2.1) and only half the available bandwidth, since a larger frequency range seems to lie outside the convergence radius of the Taylor-expansion (see Sec. 4.2.1 used in the implementation of CASA).

In Fig. 4.4, a comparison is shown between RESOLVE, a power-law fit, and MF-CLEAN results for the spectral index and differently strong masks. The full field is not shown, because neither the power-law fit nor the CLEAN reconstruction could recover any structures in the remaining regions. The error measures are listed in Tab. 4.1. It can be seen that for this example RESOLVE overall outperforms the other methods. The advantage of RESOLVE is more pronounced in the outer regions, where the surface brightness signal is weak, but RESOLVE still gives the best result even for the most central parts of the reconstruction (see also the full reconstruction in Fig. 4.3).

Fig. 4.5 shows a comparison of multi-frequency surface brightness reconstructions with RESOLVE and MS-MF-CLEAN for the simulated observation of Sec. 4.4.1. For this simulation, RESOLVE is more successful in reconstructing the overall structure, and especially recovers more of the small scales. It should be noted that this particular CLEAN image was achieved using the full bandwidth and coverage (other than for the spectral index images, as stated at the beginning of the section).

4.4.3 Power spectrum reconstructions

The power spectrum of the spectral index is reconstructed together with the field itself. It represents the spatial correlation of the spectral index over the observed sky. As explained in Sec. 4.2.2, we expect the typical spectral index structure of an extended (i.e. spatially correlated) radio source to be spatially correlated to itself. For all the details on power spectrum reconstructions we refer the reader to Ch. 2, since almost all derivations for standard single-frequency RESOLVE power spectrum reconstructions are valid as well for the spectral index.

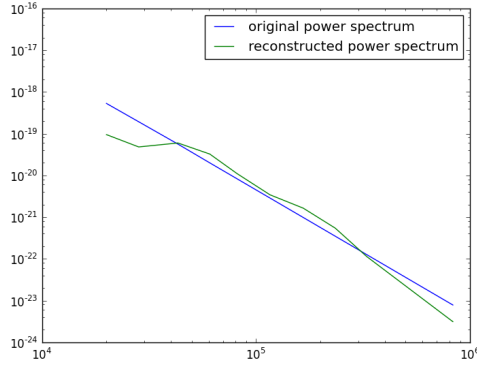
In this section we only discuss the spectral index power spectrum, since the reconstructions for the brightness power spectra are identical to the ones already presented in Ch. 2. We simply note that, in principle, a multi-frequency reconstruction recovers structure on smaller scales, and thus, it is expected that multi-frequency RESOLVE should be able to map spatial power spectra up to higher Fourier modes (i.e. smaller correlation structures).

A typical result from multi-frequency RESOLVE is illustrated in Fig. 4.2. It shows the original spectral index power spectrum of the Gaussian signal field, used in the simulated observations of Sec. 4.4.1, and its reconstruction that belongs to the same iteration step as the presented signal reconstructions earlier. The power spectrum is reconstructed relatively well, no power is lost on high modes (i.e. small scales), which is just a consequence of the fact that the simulated observation was conducted with low noise.

As for the brightness reconstructions in Ch. 2, we emphasize that the power spectrum should not only be viewed as a by-product of the algorithm in order to accurately estimate the spectral index signal. For instance, some astronomical objects show very distinct spectral index structures and can even be classified after this criterion. A prominent example might be radio halos and relics of galaxy clusters, both of which typically show steep spectral indices that evolve spatially roughly like the source itself (Feretti et al. 2012). Measuring the spatial power spectra of these objects might lead to a more exact and quantitative classification scheme. Another application might lie in the investigation of different physical processes within a single source that lead to spectrally very different regions. Estimating the spectral correlation structure over such regions offers a new way of quantitative analysis of the interplay of these processes.

4.5 Conclusion

We presented a multi-frequency extension to the imaging algorithm RESOLVE (Junklewitz et al. 2013). The combined algorithm is optimal for multi-frequency imaging of extended radio sources. It simultaneously estimates the surface brightness at a reference frequency, and the spectral index of the source. Within the assumption of a spectral index model, no further expansions or parameter-dependent modeling is used in the reconstruction. Multi-frequency RESOLVE is thus capable of exploiting the full bandwidth of a modern radio observation for maximum sensitivity and resolution, only limited by higher order spectral effects like spectral curvature.



(a) Spectral power spectrum reconstruction

Figure 4.2: Power spectrum reconstruction of the spectral index α for the reconstruction shown in Sec. 4.4.1 using multi-frequency RESOLVE.

Multi-frequency RESOLVE has been tested successfully using simulated observations of the VLA in its A-configuration. For the presented tests, the algorithm can outperform standard imaging methods in both surface brightness and spectral index estimations.

The algorithm uses a Gaussian prior and an effective log-normal model for the spectral index. This approach is not necessarily restricted to a combination with single-frequency RESOLVE, and can in principle be combined with any other imaging or deconvolution method for the surface brightness reconstruction.

For the sake of feasibility, many details of radio interferometric observations have been left out of the analysis. The response might realistically contain a number of additional effects, for instance an instrumental primary beam or wide-field and direction dependent effects. We refer the reader to Junklewitz et al. (2013), where many of these problems already have been discussed in the outlook.

In its current form, the presented algorithm has relatively high computational costs and numerical demands (for an analysis of algorithmic efficiency, see Junklewitz et al. 2013). In general, this is true for most Bayesian statistical inference algorithms. This could pose an obstacle for day-to-day applicability, since especially modern broad band data sets tend to be very large, and therefore are already a numerical challenge on their own. Since this study was centered on fundamental algorithmic development, numerical efficiency was not a focus. Future work might be needed to obtain a more efficient implementation of the algorithm.

For the future, it seems to be desirable to refrain completely from an explicit spectral model, and try to infer a full, non-parametric, three dimensional spectral intensity $I(l, m, \nu)$. Possibly, such a development could benefit greatly from reconstructing a full cross-correlation structure between the sky and spectral space. We leave this more complete approach for a future publication.

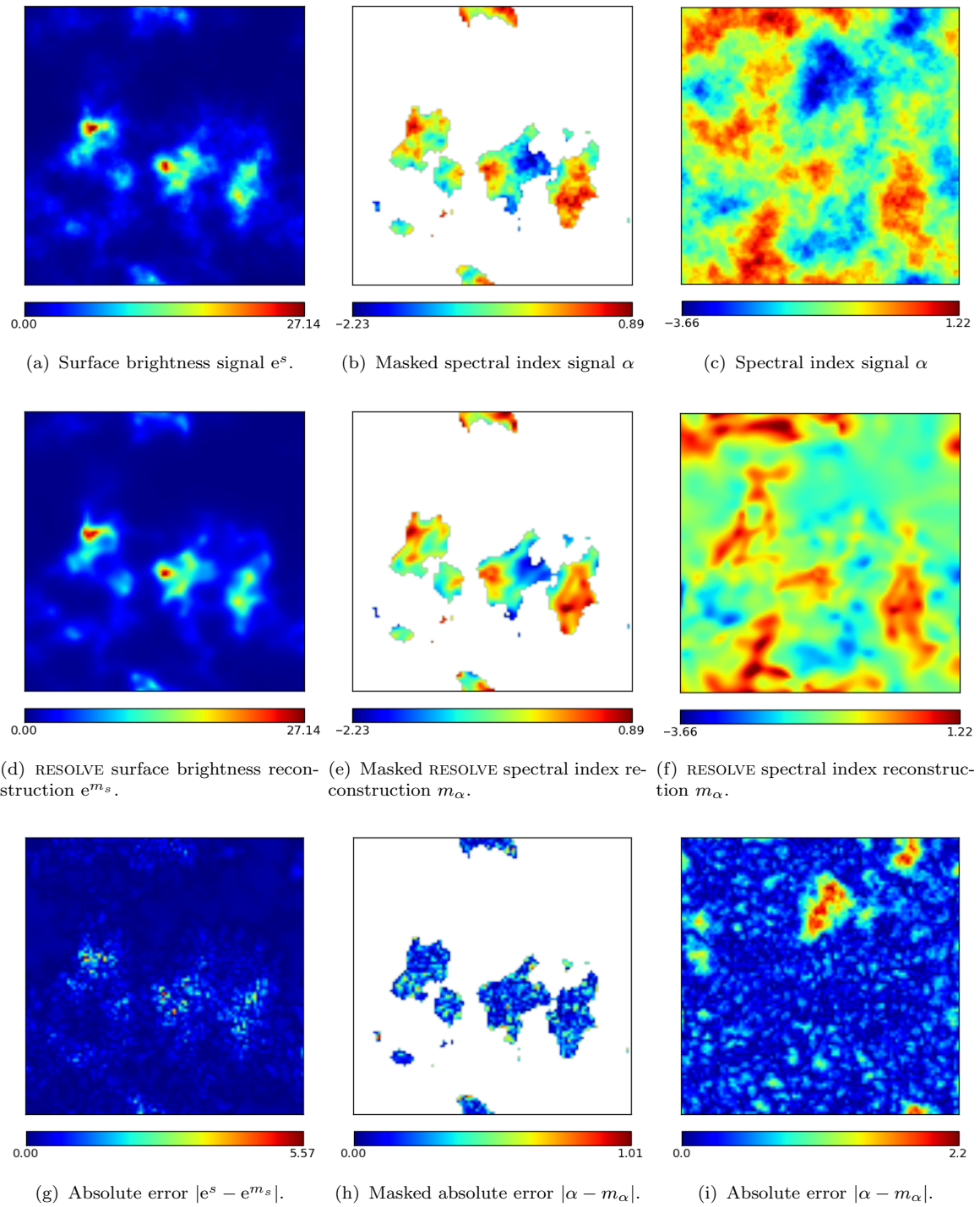


Figure 4.3: Multi-frequency reconstruction of the two signal fields e^s and α , observed with a sparse uv -coverage from a VLA-A-configuration (see Fig. 4.1). The images are 100^2 pixels large, the pixel size corresponds to roughly 0.1 arcsec. The brightness units are in Jy/px. The ridge-like structures in the difference maps simply stem from taking the absolute value and mark zero-crossings between positive and negative errors.

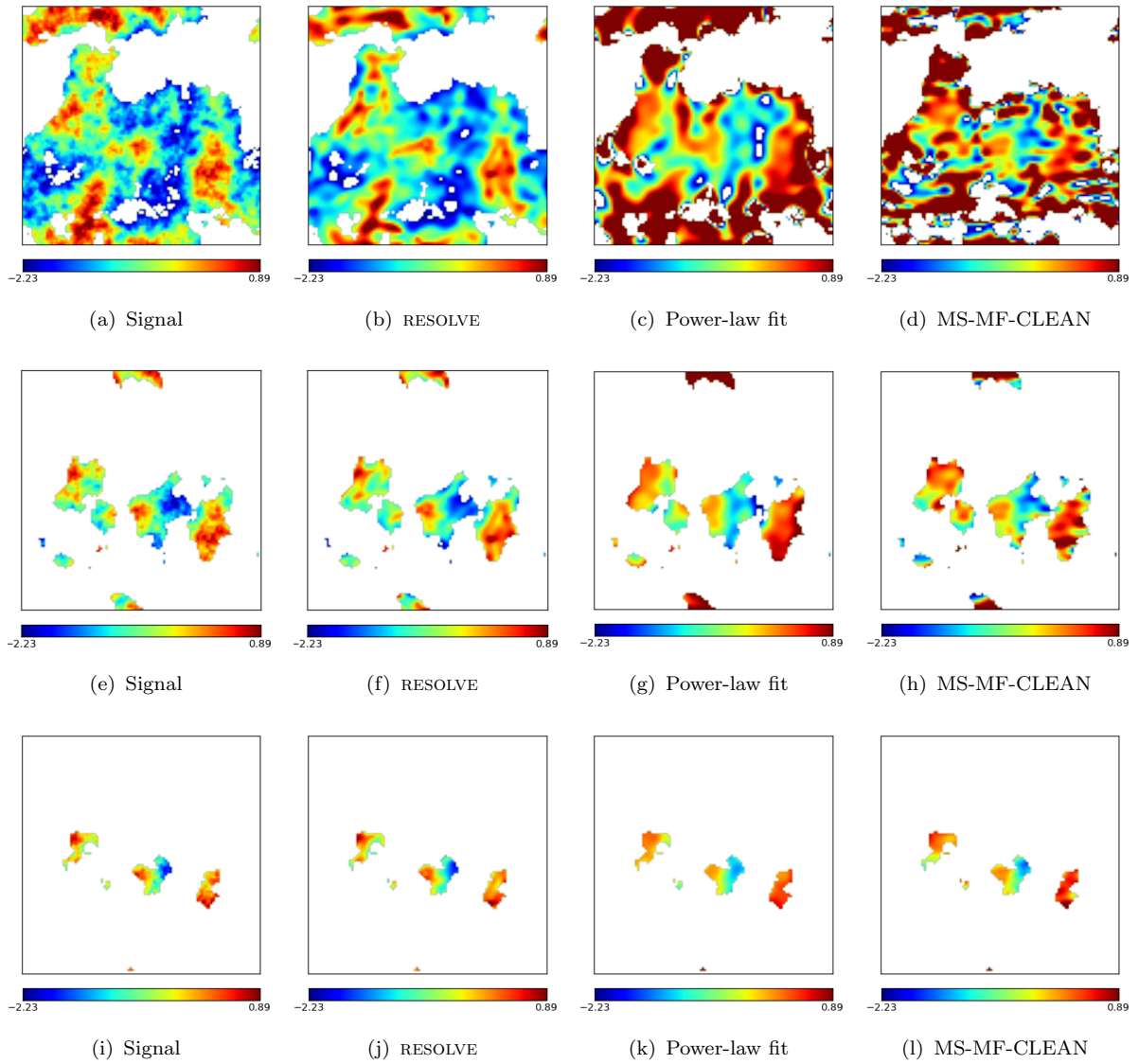


Figure 4.4: Comparison of different methods for spectral index reconstruction with differently strong masks. The images are 100^2 pixels large, the pixel size corresponds to roughly 0.1 arcsec. *First column*: Signal. *Second column*: RESOLVE reconstruction. *Third column*: Power-law fit. *Fourth column*: MS-MF-CLEAN.

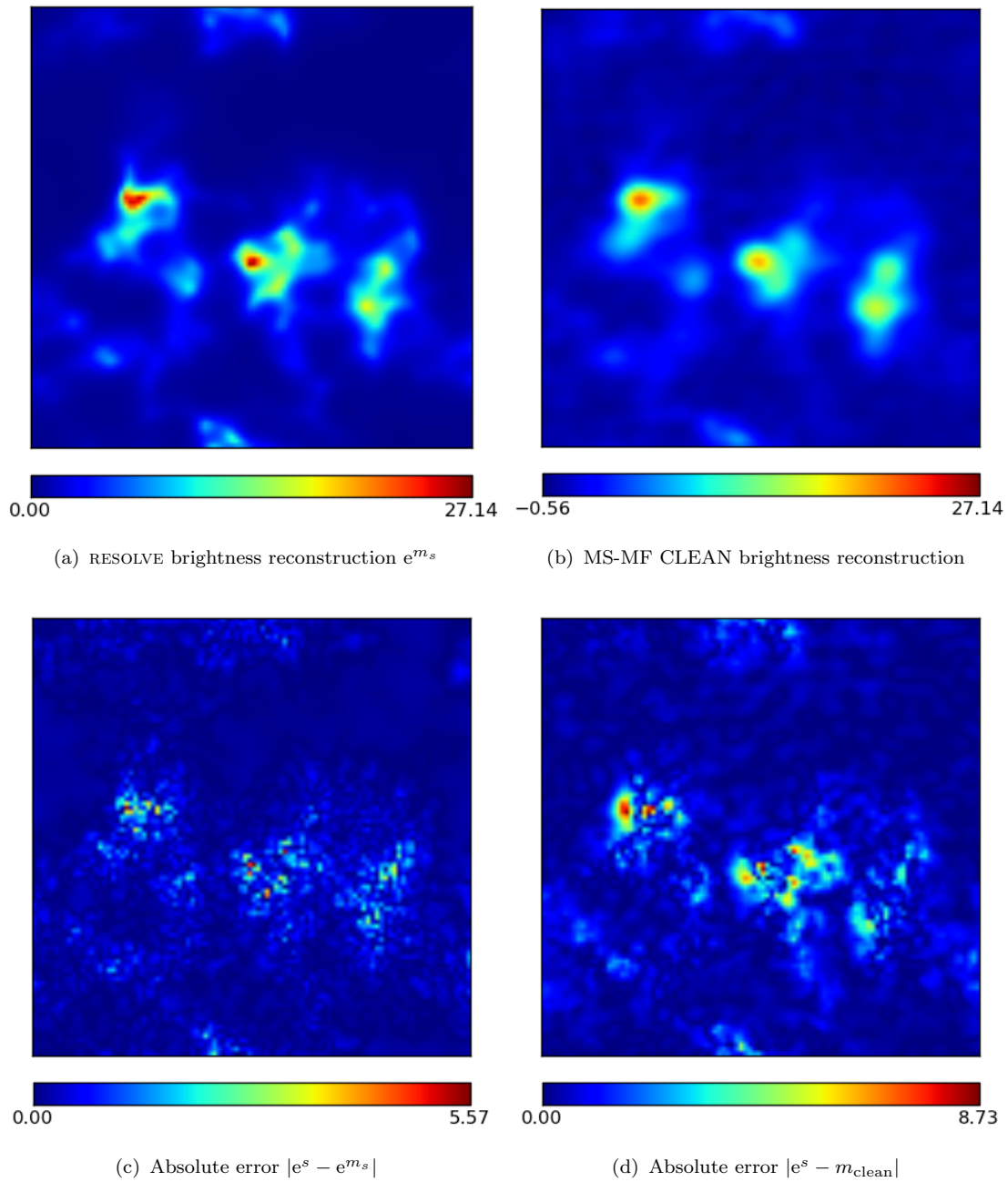


Figure 4.5: Comparison of multi-frequency RESOLVE and MS-MF-CLEAN surface brightness reconstructions. The images are 100^2 pixels large, the pixel size corresponds to roughly 0.2 arcsec. The brightness units are in Jy/px. The ridge-like structures in the difference maps simply stem from taking the absolute value and mark zero-crossings between positive and negative errors. *First row left:* RESOLVE reconstruction. *First row right:* MS-MF CLEAN reconstruction. *Second row left:* Absolute per-pixel difference between the signal and the RESOLVE reconstruction. *Second row right:* Absolute per-pixel difference between the signal and the MS-MF CLEAN reconstruction.

Chapter 5

Magnetic field statistics

A strong incentive for the developments in the previous chapters has been to meet some of the challenges of modern radio data analysis. Ultimately, it is scientific interest that drives these efforts, since advances in radio astrophysics rely on new developments in observing and imaging techniques. This chapter presents two works from magnetic field astrophysics, which both crucially depend on a careful, high fidelity imaging analysis. Both engage in studies of turbulent magnetic fields, whose stochastic nature prevents a direct analysis. In two different settings, it is presented how magnetic field properties still can be measured using statistical analysis tools. In this, the chapter continues to make use of inference techniques very similar to the ones previously presented in the thesis.

Both studies were originally presented as parts of recently published works. The first is a statistical assessment of the occurrence of specific features that are expected to be found in rotation measure synthesis analysis (see Sec. 1.2.3), called *Faraday Caustics* and was presented as part of Bell et al. (2011). The second presents the mathematical derivation of a previously ad-hoc devised magnetic helicity estimator using the imprint of magnetic field statistics on radio polarimetry, called *LITMUS* and was presented as part of Junklewitz and Enßlin (2011).

5.1 Faraday Caustics

The first method presented in this chapter involves the analysis of *Faraday Caustics*. These are characteristic, asymmetrical structures that indicate magnetic field reversals. They are predicted to be present in images of polarized radio emission, generated using the technique of RM synthesis (Bell et al. 2011).

This technique rests on the work of Burn (1966). He showed that the complex polarized intensity P , introduced in Sec. 1.2.3, as a function of wavelength-squared λ^2 is related to the intrinsic complex polarized intensity F as a function of Faraday depth ϕ^1 , called the

¹In this work, any explicit spectral dependence of $F(\phi)$ on λ^2 itself is not considered (see Bell et al. (2011) for details).

Faraday spectrum², by a Fourier transformation

$$P(\lambda^2) = \int_{-\infty}^{\infty} F(\phi) e^{2i\phi\lambda^2} d\phi. \quad (5.1)$$

The Faraday spectrum can in principle be reconstructed by inverting Eq. (5.1). The technique is used when the Faraday rotating and synchrotron radiating media along the LOS are intermixed to discern the complex structure of polarization plane rotations into $F(\phi)$ (see also Sec. 1.2.3).

This is in general not possible due to our inability to completely sample P at all values of λ^2 , since no observation can cover a continuous band of frequencies and furthermore, negative frequencies cannot be measured by principle. This is in complete analogy to the imaging problem with radio-interferometers (see Ch. 2 or Thompson et al. (1986) for comparison). However, good imaging results of Eq. (5.1) can be achieved using the RM synthesis technique, developed by Brentjens and de Bruyn (2005), in which the inversion is treated much the same way as in aperture synthesis imaging in radio astronomy (Thompson et al. 1986). Even further, both techniques, aperture synthesis and RM synthesis can be combined into a single approach, Faraday synthesis (Bell and Enßlin 2012).

Bell et al. (2011) predicted asymmetric spike-like features in the Faraday spectrum, the Faraday caustics. They arise when the LOS magnetic field changes sign (i.e. direction), causing a pile-up of emission at a single Faraday depth. This can be understood directly in a qualitative way. A volume filled with magnetic fields and a plasma of thermal and relativistic electrons in general emits synchrotron radiation and rotates the plane of polarization of traversing radiation. Along a LOS of an observation, the polarized emission from each point maps in general to a different location in ϕ - space. However, when the magnetic field along the LOS becomes very small and approaches zero, according to Eq. (1.14) the emission from a whole region gets mapped to a single ϕ value and will create a spiky feature in the Faraday spectrum. This means that Faraday caustics indicate magnetic field reversals along the LOS, which renders them a possible tool to study the magnetic field, especially in a turbulent environment, where many such reversals are expected to be present.

In Bell et al. (2011) it was shown that the Faraday spectrum for such a field reversal follows

$$F(\phi) = \frac{D \Theta[(\phi - \phi_0)B'_3(0)]}{\sqrt{\phi - \phi_0}}, \quad (5.2)$$

where $D = \frac{2a_s |B_\perp(0)|^2}{\sqrt{a_1 B'_3(0)}}$, Θ is the Heaviside function, B_3 the magnetic field component along the LOS parallel to the z -axis, $B'_3 = \partial B_3 / \partial z$, and the position $z = 0$ is identified with the location of the caustic.

It can be further shown that the integrated intensity over a range $\delta\phi$, roughly identical to a resolution element of the observation, and centered on a Faraday caustic structure,

²Usually it is also referred to as the Faraday dispersion function (see Burn 1966).

gives

$$\mathcal{F} = \left| 2D\sqrt{\delta\phi} \right| = 4a_s |B_{\perp}(0)|^2 \sqrt{\frac{\delta\phi}{a_1 B'_3(0)}}, \quad (5.3)$$

where a_s is a parameter that summarizes all important constants from synchrotron physics, unimportant for this study and exactly defined in Bell et al. (2011).

This way the more realistic setting of a finite instrument resolution is accounted for, where probably many caustics will be smeared out in the actual observed Faraday spectrum. It can be seen that small values of $B'_3(0)$, corresponding to locations where the magnetic field is small and does not change rapidly along the LOS, result in strong caustics. This also shows that measuring the flux of a Faraday caustic in principle provides a way to investigate the LOS magnetic field.

5.1.1 Distribution of Faraday caustics in a Gaussian random magnetic field³

We have just shown how Faraday caustics can provide information about the structure of the LOS magnetic field. It is variation and turbulence in the magnetic field that leads to caustics; after all, a purely uniform magnetic field would not produce such signatures. It may be possible to recover information about the statistical properties of the underlying magnetic field in a source by studying the distribution of caustics within its Faraday spectrum. We therefore investigate Faraday caustics generated by a Gaussian random magnetic field and calculate their probability distribution as a function of strength, i.e. their luminosity function.

The probability distribution for caustics

We derive the probability of measuring a spike with an integrated intensity \mathcal{F} given that the spike is due to a caustic (i.e. $B_3 = 0$) and the magnetic field has a covariance M . We limit our discussion to the regime where Eq. (5.3) adequately describes the integrated flux of the caustic.

In Bell et al. (2011), we found that the integrated polarized intensity of a caustic is $\mathcal{F} = 4a_s \sqrt{\delta\phi} |B_{\perp}(0)|^2 / \sqrt{B'_3(0)a_1}$. We assume a Gaussian magnetic field distribution, $\mathcal{G}(B, M)$. The probability distribution of integrated intensities \mathcal{F} , given that $B_3 = 0$ and the magnetic correlation tensor is M is

$$P(\mathcal{F}|\text{caustic}, M) = \frac{P(\mathcal{F}, \text{caustic}|M)}{P(\text{caustic}|M)}. \quad (5.4)$$

³Note: This section was published in *Astronomy and Astrophysics* as part of Bell et al. (2011) as Section 4 and Appendix A therein. Section and Appendix have been merged and slightly rewritten for better clarity in the context of this thesis. These sections include the main contributions of this thesis' author to the paper.

The denominator, which is the probability that a caustic occurs in a Gaussian random magnetic field, is relatively straightforward to compute. We impose the condition that $B_3(0) = 0$ using a delta function in order to compute the probability

$$P(\text{caustic}|M) = \int \mathcal{D}B \delta(B_3(0)) \mathcal{G}(B, M), \quad (5.5)$$

where the integral is a path-integral over all possible realizations of the magnetic field. We then replace the delta function using the Fourier transform

$$P(\text{caustic}|M) = \frac{1}{\sqrt{|2\pi M|}} \int \mathcal{D}B \int d\eta e^{2\pi i \eta B_3} e^{-\frac{1}{2} B^\dagger M^{-1} B}. \quad (5.6)$$

We now replace ηB_3 with $\eta \delta_{i3} \vec{B}$ and after completing the square we have

$$P(\text{caustic}|M) = \frac{1}{\sqrt{|2\pi M|}} \int \mathcal{D}B \int d\eta \exp \left[-\frac{1}{2} (B - J^\dagger M)^\dagger M^{-1} (B - J^\dagger M) + \frac{1}{2} J^\dagger M J \right], \quad (5.7)$$

where $J^\dagger = 2\pi i \eta \delta_{i3} \delta(z)$. The \vec{B} integral is the integral of a Gaussian function, which gives a factor of $\sqrt{|2\pi M|}$. The η integral is also a Gaussian integral. After integration, the result is

$$P(\text{caustic}|M) = \sqrt{\frac{2\pi}{M_{33}(0)}}. \quad (5.8)$$

The calculation of the numerator of Eq. (5.4) is more complicated. We again start by imposing our conditions using delta functions

$$P(\mathcal{F}, \text{caustic}|M) = \int \frac{\mathcal{D}B}{\sqrt{|2\pi M|}} \delta(B_3(0)) \delta\left(\mathcal{F} - \frac{a_s \sqrt{\delta\phi} B_\perp^2(0)}{\sqrt{B'_3(0) a_1}}\right) \exp \left[-\frac{1}{2} B^\dagger M^{-1} B \right]. \quad (5.9)$$

We rewrite the second delta function in terms of B_3

$$\delta\left(\mathcal{F} - \frac{4a_s B_\perp^2(0) \sqrt{\delta\phi}}{\sqrt{B'_3(0) a_1}}\right) = 4K \frac{B_\perp^4}{\mathcal{F}^3} \delta\left(B'_3(0) - K \frac{B_\perp^4(0)}{\mathcal{F}^2}\right), \quad (5.10)$$

where $K = 16a_s^2 \delta\phi / a_1$.

Inserting this into Eq. (5.9), we now proceed with the calculation using the Fourier representation of the delta function

$$\begin{aligned} & \int \frac{\mathcal{D}B}{\sqrt{|2\pi M|}} \delta(B_3(0)) 2K \frac{\delta\phi^2 B_\perp^4}{a_1^2 \mathcal{F}^3} \delta\left(B'_3(0) - K \frac{B_\perp^4(0)}{\mathcal{F}^2}\right) \exp \left[-\frac{1}{2} B^\dagger M^{-1} B \right] \\ &= \frac{2K \delta\phi^2}{a_1^2 \mathcal{F}^3} \int \frac{\mathcal{D}B}{\sqrt{|2\pi M|}} \int \frac{d\mu}{2\pi} \int \frac{d\eta}{2\pi} B_\perp^4(0) \exp \left[-\frac{1}{2} B^\dagger M^{-1} B \right] \exp[i\eta B_3(0)] \\ & \quad \exp \left[i\mu \left(B'_3(0) - \frac{K B_\perp^4(0)}{\mathcal{F}^2} \right) \right]. \end{aligned} \quad (5.11)$$

We introduce two generalized fields $\mu' = \delta_{i3} \delta(x - x_0) \delta(y - y_0) \mu$ and $\eta' = \delta_{i3} \delta(x - x_0) \delta(y - y_0) \eta$ that permit us to work with the total magnetic field $B(z)$ rather than the LOS-component $B_3(z)$ in Eq. (5.11):

$$\begin{aligned} & \frac{4K}{\mathcal{F}^3} \int \frac{\mathcal{D}B}{\sqrt{|2\pi M|}} \int \frac{d\mu}{2\pi} \int \frac{d\eta}{2\pi} B_{\perp}^4(0) \exp \left[-i\mu \left(\frac{KB_{\perp}^4(0)}{\mathcal{F}^2} \right) \right] \\ & \exp \left[-\frac{1}{2} B^{\dagger} M^{-1} B + i(\eta' + \mu' \partial_3)^{\dagger} B \right]. \end{aligned} \quad (5.12)$$

The functional integral over $\mathcal{D}B$ cannot be solved analytically because of the $B_{\perp}^4(z)$ -term in the exponential. In general, we need to employ diagrammatic perturbation theory in order to proceed.

If we restrict our attention to strong caustics, we can proceed analytically. This is a reasonable restriction because it will be the strong caustics that are the most likely features to be observed. We note that a strong caustic can occur when either B_{\perp} is large or B_3' is small. Strong caustics caused by an exceptionally large sky plane magnetic field are unlikely because of our assumption of Gaussian statistics, and therefore we concentrate on caustics with large \mathcal{F} owing to a slowly changing LOS magnetic field. For these caustics, we can neglect the $B_{\perp}^4(z)/\mathcal{F}^4$ -term and solve the remaining problem by introducing a generating functional J and repeatedly applying the Gaussian integrations

$$\begin{aligned} & \frac{4K}{\mathcal{F}^3} \int \frac{\mathcal{D}B}{\sqrt{|2\pi M|}} \int \frac{d\mu}{2\pi} \int \frac{d\eta}{2\pi} (B_1^2(0) + B_2^2(0))^2 \exp \left[-\frac{1}{2} B^{\dagger} M^{-1} B + i(\eta' + \mu' \partial_3)^{\dagger} B \right] \\ & = \frac{4K}{\mathcal{F}^3} \int \frac{\mathcal{D}B}{\sqrt{|2\pi M|}} \int \frac{d\mu}{2\pi} \int \frac{d\eta}{2\pi} \left[\left(\frac{\delta}{\delta J_1(z)} \right)^2 + \left(\frac{\delta}{\delta J_2(z)} \right)^2 \right]^2 \\ & \quad \exp \left[-\frac{1}{2} B^{\dagger} M^{-1} B + (i\eta' + i\mu' \partial_3 + J)^{\dagger} B \right] \Bigg|_{J=0} \\ & = \frac{4K}{\mathcal{F}^3} \left[\left(\frac{\delta}{\delta J_1(z)} \right)^2 + \left(\frac{\delta}{\delta J_2(z)} \right)^2 \right]^2 \int \frac{d\mu}{2\pi} \int \frac{d\eta}{2\pi} \\ & \quad \exp \left[\frac{1}{2} (i\eta' + i\mu' \partial_3 + J)^{\dagger} M (i\eta' + i\mu' \partial_3 + J) \right] \Bigg|_{J=0} \\ & = \frac{4K}{\mathcal{F}^3} \left[\left(\frac{\delta}{\delta J_1(z)} \right)^2 + \left(\frac{\delta}{\delta J_2(z)} \right)^2 \right]^2 \int \frac{d\mu}{2\pi} \int \frac{d\eta}{2\pi} \\ & \quad \exp \left[-\frac{1}{2} \eta'^{\dagger} M \eta' - \frac{1}{2} (\mu' \partial_3)^{\dagger} M (\mu' \partial_3) + \frac{1}{2} J^{\dagger} M J \right. \\ & \quad \left. + \underbrace{\left(-\frac{1}{2} (\eta' + J)^{\dagger} M (\mu' \partial_3) - \frac{1}{2} (\mu' \partial_3)^{\dagger} M (\eta' + J) \right)}_{=0, \text{ because } M_{33}(z, z') \text{ is at its maximum for } z, z'=0} + i\eta'^{\dagger} M J \right] \Bigg|_{J=0} \end{aligned}$$

$$\begin{aligned}
&= \frac{4K}{\mathcal{F}^3} \frac{1}{\sqrt{2\pi|M''_{33}(0)|}} \left[\left(\frac{\delta}{\delta J_1(z)} \right)^2 + \left(\frac{\delta}{\delta J_2(z)} \right)^2 \right]^2 \int \frac{d\eta}{2\pi} \\
&\quad \exp \left[-\frac{1}{2}\eta^2 M_{33}(0) + i\eta \underbrace{\int dz' J_i M_{i3}(z', 0)}_{=I} + \frac{1}{2} J^\dagger M J \right] \Big|_{J=0} \\
&= \frac{4K}{\mathcal{F}^3} \frac{1}{\sqrt{2\pi|M_{33}(0)|}} \frac{1}{\sqrt{2\pi|M''_{33}(0)|}} \left[\left(\frac{\delta}{\delta J_1(z)} \right)^2 + \left(\frac{\delta}{\delta J_2(z)} \right)^2 \right]^2 \\
&\quad \exp \left[\frac{1}{2} J^\dagger M J - \frac{1}{2} \frac{I^2}{M_{33}(0)} \right] \Big|_{J=0} \\
&= \frac{4K}{\mathcal{F}^3} \frac{1}{\sqrt{2\pi|M_{33}(0)|}} \frac{1}{\sqrt{2\pi|M''_{33}(0)|}} \left[3(M_{11}^2(0) + M_{22}^2(0)) + M_{22}(0)M_{11}(0) \right. \\
&\quad \left. + \frac{1}{M_{33}^2(0)} \underbrace{(M_{13}^4(0) + M_{23}^4(0) + 2M_{23}^2(0)M_{13}^2(0))}_{=0} \right] \\
&= \frac{28K}{\mathcal{F}^3} \frac{1}{\sqrt{2\pi|M_{33}(0)|}} \frac{1}{\sqrt{2\pi|M''_{33}(0)|}} M_N^2 \\
&= \frac{28K}{\mathcal{F}^3} \frac{1}{\sqrt{2\pi|M_{33}(0)|}} \frac{1}{\sqrt{2\pi|M''_{33}(0)|}} \left[\int \frac{dk^3}{(2\pi)^3} \hat{M}_N(k) \right]^2 \\
&= \frac{28K}{\mathcal{F}^3} \frac{1}{\sqrt{2\pi|M_{33}(0)|}} \frac{1}{\sqrt{2\pi|M''_{33}(0)|}} \left[\int dk^3 \frac{\epsilon_B(k)}{k^2} \right]^2. \tag{5.13}
\end{aligned}$$

In the last step, we identified the covariance matrix M with the magnetic correlation tensor given by Eqs. 1.18. We used the property $M_{ij}(0) = M_N(0) \delta_{ij}$. Finally, from this result and Eq. (5.8), we conclude that in the bright caustic limit

$$\begin{aligned}
P(\mathcal{F}|\text{caustic}, M) &= \frac{P(\mathcal{F}, \text{caustic}|M)}{P(\text{caustic}|M)} = \frac{448a_s^2 \delta \phi^2}{\mathcal{F}^3 a_1} \frac{M_N^2(0)}{\sqrt{2\pi|M''_{33}(0)|}} \\
&= \frac{448a_s^2 \delta \phi^3}{\mathcal{F}^3 a_1^2 \sqrt{2\pi|M''_{33}(0)|}} \left[\int dk^3 \frac{\epsilon_B(k)}{k^2} \right]^2 = \frac{448a_s^2 \delta \phi^3}{\mathcal{F}^3 a_1^2 \sqrt{2\pi}} \frac{\left[\int dk^3 \frac{\epsilon_B(k)}{k^2} \right]^2}{\left[\int dk^3 \epsilon_B(k) \right]^{1/2}}. \tag{5.14}
\end{aligned}$$

Physical consequences

We have just found that, over a limited range in \mathcal{F} , the distribution of the integrated flux of caustics will be proportional to \mathcal{F}^{-3} :

$$P(\mathcal{F}|\text{caustic}, M) = \frac{448a_s^2 \delta \phi^3}{a_1^2} \frac{M_N^2(0)}{\sqrt{2\pi|M''_{33}(0)|}} \frac{1}{\mathcal{F}^3}. \tag{5.15}$$

We note that this indicates that the number of caustics per logarithmic interval of \mathcal{F} is $\propto \mathcal{F}^{-2}$ and thus that the Faraday caustic flux per logarithmic interval of \mathcal{F} goes as \mathcal{F}^{-1} , which is finite as $\mathcal{F} \rightarrow \infty$.

The fore-factors in Eq. (5.15) depend on the properties of the magnetic field and the turbulent flow of the fluid in which they are situated. The term $M_N(0)$ is simply the total magnetic energy density. We also introduce the *Taylor micro-scale*, l_T , which is defined by

$$\left. \frac{d^2 C(z)}{dz^2} \right|_{z=0} = \frac{2}{l_T^2} \quad (5.16)$$

(see e.g. Tennekes and Lumley 1972, sec. 6.4), where $C(z)$ is the covariance in z normalized to unity at $z = 0$. This length scale represents the largest scale on which dissipation is important in a turbulent flow. The second derivative of the LOS component of the magnetic correlation tensor, $M''_{33}(0)$, is proportional to ϵ_B/l_T^2 . We can therefore rewrite Eq. (5.15) as

$$P(\mathcal{F}|\text{caustic}, M) = \frac{448 a_s^2 \delta \phi^3 \epsilon_B^{3/2} l_T}{a_1^2 \sqrt{4\pi} \mathcal{F}^3}, \quad (5.17)$$

where ϵ_B is the total magnetic energy density.

By measuring the luminosity function of Faraday caustics in the \mathcal{F}^{-3} regime, we can see that one is in principle able to measure the Taylor micro-scale in a turbulent medium. Fletcher and Shukurov (2007) show that this length scale is also measurable by observing the mean separation between so-called “depolarization canals,” which are lines of zero polarized intensity in a diffuse polarized field. These canals have been observed in maps of diffuse polarized Galactic emission by, e.g. Uyaniker et al. (1998), Haverkorn et al. (2000), Gaensler et al. (2001), and Reich et al. (2004). Observations of depolarization canals and Faraday caustics in tandem may be a powerful tool for studying the turbulent properties in, e.g. the ISM.

A rather important caveat, applicable to both Faraday caustics and depolarization canals, is that without sufficiently high resolution one will not be able to measure this scale. In the case of caustics, as discussed in the previous section, the effect of small-scale fluctuations is to split a single Faraday caustic into a tight bundle of spikes. Without enough resolution, an observer may simply count such a bundle as one large spike. This has the effect of altering the normalization of the number distribution of caustics in \mathcal{F} , which would lead to an incorrect measurement of l_T . In principle, the same effect applies to depolarization canals, where what appears to be a single canal of net zero polarized intensity may be a fine network of such canals. As described previously, the effect of finite resolution is to smooth the quantity being measured, e.g. the LOS magnetic field distribution in the case of Faraday caustics, thereby throwing away information on any smaller scales. If the Taylor micro-scale is smaller than the smoothing scale set by our experimental resolution, then it will not be measurable. We suspect that this may be the reason why the Taylor scale of the Milky Way is measured to be much larger than expected in the example presented by Fletcher and Shukurov (2007).

The result given in Eq. (5.17) is valid for strong spikes, i.e. when $B'_3(0)$ is small, but we expect that below some value of $\mathcal{F} = \mathcal{F}_{low}$ the distribution will flatten. This is required to ensure that the integral of $P(\mathcal{F}|\text{caustic}, M)$ is finite for $\mathcal{F} \rightarrow 0$, but we can also see that this will be the case by considering the approximation applied during our derivation.

To permit us to compute the probability analytically, we neglect an exponential term in Eq. (5.12). This term is $\exp(-\mu K B_\perp^4/\mathcal{F})$, where $K = 16a_s^2\delta\phi/a_1$. This is, of course, only valid if the argument of the exponent is much less than one. The approximation breaks down when this term approaches unity, which leads to the condition for \mathcal{F}_{low}

$$\mathcal{F}_{low} \approx \langle B_\perp^2 \rangle \sqrt{\frac{K}{\langle B'_3 \rangle}} \quad (5.18)$$

$$\approx K^{\frac{1}{2}} \epsilon_B^{\frac{3}{4}} \lambda_B^{-\frac{1}{2}}, \quad (5.19)$$

where

$$\epsilon_B = \int_0^\infty dk \epsilon_B(k) \quad (5.20)$$

is the average magnetic energy density, $\epsilon_B(k)$ is the 1D magnetic power spectrum, and

$$\lambda_B = \frac{\pi \int_0^\infty dk \epsilon_B(k)/k}{\int_0^\infty dk \epsilon_B(k)} \quad (5.21)$$

is the magnetic correlation length.

If we wish to evaluate Eq. (5.12) in all regimes, we can no longer neglect this term. While this full calculation is beyond the scope of this study, we can consider what effects the inclusion of this term might have on the result. To proceed without approximations, we would use a perturbative expansion in Feynman diagrams as it is used in quantum field theory. The first term in such an expansion would be negative and $\propto \mathcal{F}^{-2}$, the second positive and $\propto \mathcal{F}^{-1}$, and so on. The summation of these terms results in a turnover of the spectrum near \mathcal{F}_{low} .

We also expect the distribution to steepen at larger values of \mathcal{F} because of depolarization. As discussed above, large values of \mathcal{F} imply that the LOS magnetic field changes over long distances. As this happens, it is increasingly more likely that the position angle of the polarized emission at the values of z that contribute to the polarized intensity at a single value of ϕ will be uncorrelated resulting in a reduction in the integrated intensity because of depolarization. This reduces the number of caustic features at high values of \mathcal{F} .

To investigate the conditions for a steepening of the probability distribution of caustics, we consider the correlation of the sky plane magnetic field as a function of LOS distance

$$\mathcal{R} = \frac{\langle B_\perp(\bar{z}) B_\perp(-\bar{z}) \rangle}{\langle B_\perp^2(0) \rangle}, \quad (5.22)$$

where the average is over all possible field configurations. The procedure for computing these expectation values is similar to that used in Sec. 5.1.1 to compute $P(\mathcal{F}|caustic, M)$ and therefore not included here. The result is

$$\mathcal{R} = \frac{3M_N(2\bar{z})}{10M_N(0)} + \frac{7M_N^2(\bar{z})}{10M_N^2(0)}. \quad (5.23)$$

A small value of \mathcal{R} indicates that, on average, the sky plane magnetic fields are uncorrelated. When this is true, the luminosity function is steeper than \mathcal{F}^{-3} .

As an example of estimating \mathcal{F}_{low} and \mathcal{R} , we model the magnetic power spectrum as a broken power law

$$\epsilon_B(k) = \epsilon_0 \left(\frac{k}{k_0}\right)^b \left[1 + \left(\frac{k}{k_0}\right)^2\right]^{-\left(\frac{a+b}{2}\right)}. \quad (5.24)$$

For such a power spectrum, the luminosity function turns over at

$$\mathcal{F}_{low} = \pi^2 \sqrt{K} \left(\frac{\epsilon_0 k_0^{5/3}}{2}\right)^{\frac{3}{4}} \mathcal{B}\left(\frac{a-1}{2}, \frac{b+1}{2}\right)^{\frac{5}{4}} \mathcal{B}\left(\frac{a}{2}, \frac{b}{2}\right)^{-\frac{1}{2}}, \quad (5.25)$$

where $\mathcal{B}(x)$ is the Beta function.

We can also compute \mathcal{R} because

$$M_N(z) = \int dk \hat{M}_N(k) e^{-ikz} \quad (5.26)$$

and

$$\hat{M}_N(k) = \frac{8\pi^3 \epsilon_B(k)}{k^2}. \quad (5.27)$$

The resulting form of $\mathcal{R}(k_0 z)$ is shown for a few representative values of a in Fig. 5.1. For a Kolmogorov spectrum, $a = 5/3$. In each case, $b = 2$. We note that \mathcal{R} is independent of the parameter ϵ_0 .

A calculation of the precise shape of the distribution for low and high values of \mathcal{F} is beyond the scope of this paper. A sketch summarizing the expected shape of the probability distribution, represented as the number of caustics per logarithmic scale of \mathcal{F} as a function of $\log \mathcal{F}$, is shown in Fig. 5.2. The \mathcal{F}^{-2} regime is valid given the assumptions outlined above for sufficiently strong caustics and until depolarization becomes important, as described above. We expect that the exact shape of the distribution at and below \mathcal{F}_{low} will be more strongly dependent on the magnetic field statistics because weak caustic features depend more strongly on $|B_{\perp}(0)|$.

In summary, we have calculated the probability density function of caustics with a particular strength in a Gaussian random magnetic field. This result illustrates the type of analysis that is possible and represents the simplest reference case for future observations of the luminosity function. We caution that the result has a limited scope for several

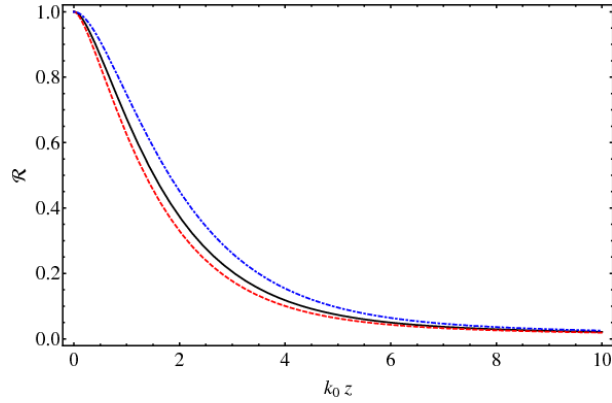


Figure 5.1: The normalized correlation of the sky plane magnetic field as a function of $k_0 z$ for a few values of the power law index a . Dashed red $a = 4/3$, solid black $a = 5/3$, dot-dashed blue $a = 7/3$.

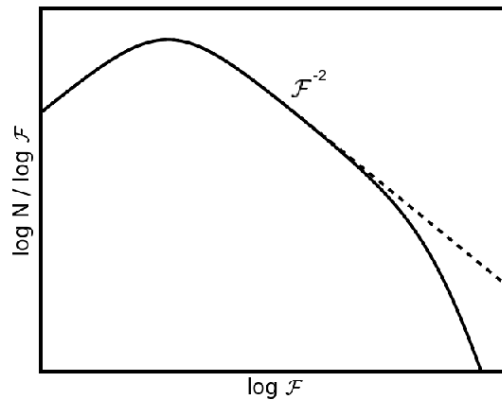


Figure 5.2: A sketch of the number distribution of caustic spikes per logarithmic interval of \mathcal{F} showing the limits to the \mathcal{F}^{-2} regime and the rough shape of the distribution beyond these limits. At high values of \mathcal{F} , the distribution steepens owing to depolarization effects, while at lower \mathcal{F} values the distribution turns over because the total number of caustics is finite. We refer to the text for more details.

reasons. For one, we have assumed an unrealistic magnetic field distribution. This was done for simplicity, but also because if one is unaware of any statistical properties of the magnetic field other than the two point correlation, the assumption of a Gaussian field is the most appropriate starting point. Any further knowledge would always be included by expanding around the Gaussian case. While we expect our result in the \mathcal{F}^{-3} regime to be only weakly dependent on the magnetic field statistics, realistic magnetic field distributions are likely to be significantly non-Gaussian. We have not investigated the degree to which this will effect our results. Secondly, the result in Eq. (5.17) is only valid over a limited range of \mathcal{F} values. The precise extent of this range is unclear at the moment, although we have presented some estimates. It is possible that the flattening and steepening regimes that we describe are near one another or even overlapping. A more detailed investigation into the flattening regime, for example, would require much more complicated calculations and, because this regime depends strongly on the magnetic field statistics, a more realistic model.

Lastly, Eq. (5.17) is truly only valid for very high resolution observations or simulated data. As discussed above, the result will change depending on ϕ -space resolution. In our treatment, we simply count the number of zero-point crossings of the LOS magnetic field. In observations, two or more of these events in physical space may occur at nearly the same ϕ location and may therefore be counted as a single caustic. This would have the effect of changing the overall normalization of the distribution (the total number of caustics would be reduced), but the \mathcal{F}^{-3} prediction should still be valid. Some additional analysis of simulations would be helpful in assessing the extent of this effect.

5.2 The LITMUS test⁴

In Junklewitz and Enßlin (2011), a number of correlation functions between the three radio observables I , P , and ϕ (see Sec. 1.2.3) have been derived and revisited in order to find a statistical estimator for magnetic helicity. Under the assumption of Gaussian magnetic fields and isotropic turbulence (see Sec. 1.2.4), it has been found that the correlation function

$$\langle P(\mathbf{k}_\perp)\phi(\mathbf{k}'_\perp)\phi(\mathbf{k}''_\perp) \rangle_{\mathbf{B}} = 2L_z(2\pi)^2\delta^2(\mathbf{k}_\perp + \mathbf{k}'_\perp + \mathbf{k}''_\perp) \hat{H}(u)\hat{H}(v)/uv \left((u_1v_1 - u_2v_2) + i(u_1v_2 + u_2v_1) \right) \quad (5.28)$$

solely depends on the helical part $\hat{H}(u)$ of the magnetic correlation tensor (see Sec. 1.2.4). The vectors \mathbf{u}, \mathbf{v} are defined as $\mathbf{u} = (\mathbf{k}''_\perp, 0)$, $\mathbf{v} = (\mathbf{k}'_\perp, 0)$, the quantities u, v denote their magnitudes.

⁴Note: Parts of this section and the complete following subsection were published in Junklewitz and Enßlin (2011) in *Astronomy & Astrophysics*. Only parts of the publication that were not in the author's diploma thesis are reported here.

Using this result, Junklewitz and Enßlin (2011) devised a simple helicity estimator only by using heuristic arguments, the *LITMUS* test (**L**ocal **I**nference **T**est for **M**agnetic fields which **U**ncovers helice**S**). The test simply operates on images of polarized intensity P , and of the gradient of the Faraday depth $\mathbf{G} = \nabla\phi$. Its rationale is that in the presence of helical fields, the average of the scalar product \mathcal{G}^*P over all pixels of a \mathcal{G}^*P -map should have a real value significantly larger than 0. Whereas in the case of non-helical fields, the alignment of \mathcal{G} and P should be changing randomly from pixel to pixel so that we would expect the average over \mathcal{G}^*P to be zero. The quantity \mathcal{G} is defined as

$$\mathcal{G} = |\mathbf{G}|^2 \exp[2i\alpha] \quad \text{with } \alpha = \arctan \frac{G_y}{G_x}. \quad (5.29)$$

In short mathematical notation this is stated as

$$\left\langle \mathcal{G}^*P \right\rangle_{\text{helicity}} > 0 \quad \text{and real,} \quad (5.30)$$

$$\left\langle \mathcal{G}^*P \right\rangle_{\text{no helicity}} = 0, \quad (5.31)$$

where the ensemble average over helical or non-helical fields is in practice replaced by an average over all pixels of a \mathcal{G}^*P -map.

5.2.1 Mathematical derivation

It can be shown that the heuristics were right and that this test actually depends directly on $\langle P(\mathbf{k})\phi(\mathbf{k}')\phi(\mathbf{k}'') \rangle$ which was the initial starting point. The full condition for the *LITMUS* test reads:

$$\left\langle \left(\mathcal{G}^*(\mathbf{x})P(\mathbf{x}) \right) \right\rangle = \left\langle \left[\left(\frac{\partial\phi(\mathbf{x})}{\partial x} \right)^2 + \left(\frac{\partial\phi(\mathbf{x})}{\partial y} \right)^2 \right] \exp \left[-2i \arctan G_y/G_x \right] P(\mathbf{x}) \right\rangle. \quad (5.32)$$

Using trigonometrical theorems it is easy to show that

$$\left(G_x^2 + G_y^2 \right) \exp \left[-2i \arctan G_y/G_x \right] = \left(G_x - iG_y \right)^2. \quad (5.33)$$

We now apply a Fourier transformation to our observables $P(x)$ and $\phi(x)$ and rewrite the whole expression:

$$\begin{aligned} \left\langle \mathcal{G}^*(\mathbf{x})P(\mathbf{x}) \right\rangle &= \left\langle \left[\frac{\partial\phi(\mathbf{x})}{\partial x_1} - i \frac{\partial\phi(\mathbf{x})}{\partial x_2} \right]^2 P(\mathbf{x}) \right\rangle \\ &= \left\langle \left(\left[\frac{\partial}{\partial x_1} - i \frac{\partial}{\partial x_2} \right] \phi(\mathbf{x}) \right) \left(\left[\frac{\partial}{\partial x_1} - i \frac{\partial}{\partial x_2} \right] \phi(\mathbf{x}) \right) P(\mathbf{x}) \right\rangle \end{aligned}$$

$$\begin{aligned}
&= \left\langle \int \frac{dk^2}{(2\pi)^2} \int \frac{dk^{2'}}{(2\pi)^2} \int \frac{dk^{2''}}{(2\pi)^2} [k'_1 - ik'_2] [k''_1 - ik''_2] \exp [i\mathbf{x}(\mathbf{k} + \mathbf{k}' + \mathbf{k}'')] \right. \\
&\quad \left. P(\mathbf{k})\phi(\mathbf{k}')\phi(\mathbf{k}'') \right\rangle \\
&= \int \frac{dk^2}{(2\pi)^2} \dots \int \frac{dk^{2''}}{(2\pi)^2} [k'_1 - ik'_2] [k''_1 - ik''_2] \exp [i\mathbf{x}(\mathbf{k} + \mathbf{k}' + \mathbf{k}'')] \\
&\quad \left\langle P(\mathbf{k})\phi(\mathbf{k}')\phi(\mathbf{k}'') \right\rangle
\end{aligned} \tag{5.34}$$

We assumed again that the average over all pixels of a \mathcal{G}^*P -map is equivalent to an ensemble average over the magnetic field statistics. Now we insert our result (5.28) for $\langle P(\mathbf{k})\phi(\mathbf{k}')\phi(\mathbf{k}'') \rangle$ and see the dependence of $\langle \mathcal{G}(\mathbf{x})P^*(\mathbf{x}) \rangle$ on the helical spectra $\hat{H}(k')\hat{H}(k'')$:

$$\begin{aligned}
\langle \mathcal{G}(\mathbf{x})P^*(\mathbf{x}) \rangle &= 2L_z(2\pi)^2 \int \frac{dk^2}{(2\pi)^2} \dots \int \frac{dk^{2''}}{(2\pi)^2} \delta^2(\mathbf{k} + \mathbf{k}' + \mathbf{k}'') [k'_1 - ik'_2] [k''_1 - ik''_2] \\
&\quad \exp [i\mathbf{x}(\mathbf{k} + \mathbf{k}' + \mathbf{k}'')] [(k'_1k''_1 - k'_2k''_2) + i(k'_1k''_2 + k'_2k''_1)] \frac{\hat{H}(k')\hat{H}(k'')}{k'k''} \\
&= 2L_z(2\pi)^2 \int \frac{dk^{2'}}{(2\pi)^2} \int \frac{dk^{2''}}{(2\pi)^2} [k'_1 - ik'_2] [k''_1 - ik''_2] \\
&\quad [(k'_1k''_1 - k'_2k''_2) + i(k'_1k''_2 + k'_2k''_1)] \frac{\hat{H}(k')\hat{H}(k'')}{k'k''}
\end{aligned} \tag{5.35}$$

$$\begin{aligned}
&= 2L_z(2\pi)^2 \int \frac{dk^{2'}}{(2\pi)^2} \int \frac{dk^{2''}}{(2\pi)^2} [k_1^{2'} + k_2^{2'}] [k_1^{2''} + k_2^{2''}] \frac{\hat{H}(k')\hat{H}(k'')}{k'k''} \\
&= 2L_z(2\pi)^2 \int \frac{dk^{2'}}{(2\pi)^2} \int \frac{dk^{2''}}{(2\pi)^2} k' k'' \hat{H}(k')\hat{H}(k'') \\
&= 2L_z \left[\int_0^\infty dk k^2 \hat{H}(k) \right]^2 = 2L_z \pi^4 \left[\int_0^\infty dk \frac{\epsilon_H(k)}{k} \right]^2
\end{aligned} \tag{5.36}$$

In (5.35) we assumed w. l. of g. that $\mathbf{x} = 0$ and in the last line, $\hat{H}(k)$ was substituted with the helical energy density ϵ_H , much as was shown for the normal magnetic energy density in Sec. 1.2.4. The details can be found in Junklewitz and Enßlin (2011). It can be seen that $\langle \mathcal{G}(\mathbf{x})P^*(\mathbf{x}) \rangle$ is a direct and clear estimator which measures the square of the k-space integrated helicity spectrum, giving more weight to the large scales. Therefore, it can be concluded that, except maybe from pathological or fine tuned situations in which the k-weighted helicities at different k-scales cancel each other (since $\hat{H}(k)$ might change sign), we can expect the *LITMUS* test to be able to reveal the presence of magnetic helicity. This is demonstrated in numerical tests on simulated data conducted by Oppermann et al. (2011a). The same work also contains an application on real data of our own galaxy and

a thorough analysis thereof.⁵

5.3 Conclusions

The presented magnetic field estimators demonstrate how statistical properties of turbulent cosmic magnetic fields can be inferred from radio polarization data. In Sec. 5.1 it was shown how the distribution of Faraday caustics, as expected to be found in Faraday spectra, could be used to constrain the Taylor microscale of the underlying turbulent magnetic fields. However, that requires a sufficiently high Faraday resolution of the observational setup. Sec. 5.2 presented a rigorous derivation of the magnetic helicity estimator *LITMUS*. Detecting magnetic helicity unambiguously in astrophysical settings other than the sun would be an important step forward in evaluating dynamo theories of magnetic evolution.

Of course, this type of analysis can be taken further in principle. Usage of the *LITMUS* test is currently fundamentally restricted to cases where the thermal electron density is known (set to be constant in this study), which is in fact not the case for many astronomical sources of interest (e.g. the ISM as in Oppermann et al. (2011a)). New and more polarization data of high quality are needed to further advance the field.

For both methods, Faraday caustic statistics and the *LITMUS* test, higher order magnetic field statistics could be explicitly involved using perturbative methods as developed for statistical inference by Enßlin et al. (2009). But as long as the general knowledge on actual magnetic field distributions remains scarce, such an approach would only be useful if these higher statistical moments could be inferred self-consistently from the analyzed data. Either way, this requires high quality radio data.

However, it should be noted that all derivations in this chapter implicitly assume that the used data has been imaged with highest fidelity, or at least that imaging artifacts are negligibly small. Thus, advances are likely to be expected not only from new data⁶ alone, but from a combination with novel developments in imaging methods.

⁵Unfortunately, it seems that a straightforward application of this test can be hampered by a too large spatial variance in the electron density, which we assumed to be constant in this work (see Oppermann et al. 2011a).

⁶As for instance from the upcoming generation of radio telescopes (see Sec. 1.2.1).

Chapter 6

Conclusions

As demonstrated throughout the chapters of this thesis, astrophysical research can benefit greatly from the interaction with statistics. This study has presented successful cases of the application of statistical inference tools with a special focus to radio astronomy and magnetic field astrophysics.

Much of this work represents of course only a part of a broader research program, and an outlook is in order on problems and possible further developments.

In Ch. 2 a new approach to signal inference and imaging in radio astronomy and especially radio interferometry was presented. The Bayesian inference algorithm RESOLVE was developed, targeted to be optimal for the imaging of extended and diffuse radio sources in total intensity. Comparisons in simulations showed that RESOLVE can outperform current imaging algorithms like Multi-scale-CLEAN and the Maximum Entropy Method in these tasks.

RESOLVE estimates the surface brightness in the sky in total intensity together with its spatial correlation structure, which is used to guide the algorithm to an optimal reconstruction of extended and diffuse sources. For a radio interferometer, it succeeds in deconvolving the instrumental point spread function during the process. Furthermore, RESOLVE provides statistical uncertainty propagation in the form of an approximated uncertainty map. This is not possible with other current imaging algorithms. Also, with RESOLVE, a reviewed approach to robust visibility weighting is introduced that allows for an automated assignment of statistically optimal weights in the imaging process, in contrast to the standard practice of setting these weights manually.

A first application of RESOLVE to VLA data of the galaxy cluster Abell 2256 was presented in Ch. 3. As for the current status, results are only preliminary since further numerical complications of processing the real data are not yet fully under control. Nevertheless, the outcomes indicate that RESOLVE can outperform classical imaging algorithms also in a real data application.

An extension of this work to multi-frequency imaging was presented in Ch. 4, where also a multi-frequency version of RESOLVE was developed. The combined algorithm is optimal for broad-band imaging of extended sources. It simultaneously estimates the surface brightness at a reference frequency and the spectral index across the source. Within

the assumption of pure power law spectra, no further expansions or parameter-dependent modeling is used in the reconstruction. Multi-frequency RESOLVE is thus capable of exploiting the full bandwidth from a modern radio observation for maximum sensitivity and resolution, only limited by higher order spectral effects like spectral curvature.

Multi-frequency RESOLVE has been tested successfully using simulated observations. For the presented tests, the algorithm can outperform standard imaging methods for both surface brightness and spectral index.

The combined algorithm uses a Gaussian prior for the spectral index and a log-normal model for the intensity.

To simplify the analysis in Chs. 2 and 4, some typical complexities of radio interferometers have been omitted for all simulations. However, the response operator R (see Eq. 2.9), describing the act of observation, can easily be expanded to cover more effects, thereby adapting to the needs of the actual observational situation. This includes a primary beam correction, a direction-dependent point spread function or the inclusion of single dish data on the same target source.

Further developments, beyond the scope of this thesis, include several topics. An extension to polarization imaging is already being worked on and will be the subject of upcoming publications. Another future topic is the possible inclusion of calibration into the framework. In the long run, the distinction between *calibration* and *imaging* is somewhat artificial and should ideally be merged into one step of complete reconstruction (see also Smirnov 2011a,b; Enßlin et al. 2013). Finally, a further goal should be to extend RESOLVE to a broader approach that can handle diffuse emission and point sources simultaneously (see Selig and Enßlin 2013, for an example from photon count imaging). It could be worthwhile to think about merging the approaches of compressed sensing, where optimal imaging strategies for sparse signals are already known, with the presented Bayesian approach into which they could be included in form of a Laplacian prior.

In Ch. 5, two studies on turbulent magnetic field statistics were presented. First, on the statistical analysis of Faraday caustics, spiky features that are predicted to occur in radio polarization imaging. Second, on the estimation of magnetic helicity from radio polarization data. Both works demonstrated how properties of turbulent magnetic fields can be inferred from radio polarization data.

Of course, this type of statistical magnetic field analysis can be taken further in principle. However, as discussed in Ch. 5, further advances crucially depend on high quality radio images, and therefore on developments in telescope technology and data analysis methods.

A joint reflection of the outcome of the different studies of Chs. 2 and 4 and Ch. 5 leads to some more general remarks. A common theme is that using elaborate statistical methods inevitably increases the complexity of the data analysis. This is true both for the mathematical effort and (usually even more important) for the computational costs. Even more so for Bayesian methods, which deal with the complicated posterior statistics and usually have a strong demand on numerical resources and complexity, as evidenced by the findings of Chs. 2, 3 and 4. Going beyond Gaussian statistics in any statistical estimation process often poses an extreme challenge and usually leads to mathematical equations,

which are not analytically solvable any more. This has led to the need for non-linear numerical solvers for RESOLVE, and to a necessary restriction of the findings in Ch. 5 to a certain valid regime for the Faraday caustics statistics. One might agree that there is always a balance to be found between the potential benefits of more exact methods versus problem complexity.

That being said, computational costs or numerical complexity should not hinder new developments in data analysis¹. Much of the work in this study has necessarily the character of a proof of concept, which is the case for new developments in any field. In the first step of fundamental algorithmic development, the focus is naturally not on numerical efficiency. But even more importantly, many of the developments presented in this thesis should be viewed as high fidelity methods, which to use might increase complexity and computing time, but should be justified always when the accuracy is needed for the anticipated scientific results.

¹As a side-remark, the available data in science is not only getting better with technological advances but also growing significantly larger in size, especially in astronomy. Thus, using larger computational resources becomes more and more necessary (and common) anyhow.

Appendix A

The maximum entropy principle

To solve an inference problem, a suitable prior distribution needs to be chosen. This decision process is a fundamental part of all presented studies. Both the development of RESOLVE and the validity of the magnetic field estimators presented in Ch. 5 crucially depend on the choice of prior statistics for the total intensity signal, the spectral index and the magnetic fields.

In inference, probabilities are used to reflect states of incomplete knowledge. In this study, the preferred method to assign these probabilities according to the available information is the *Maximum Entropy Principle* (MEP) (Jaynes 1982, 2003). The basic idea is that a state of knowledge should be reflected by a probability distribution such that the probability expresses all the available information but is maximally ignorant about everything else. Since in statistics, ignorance is measured with an entropy function, the MEP consists of maximizing a suitable entropy, subject to the constraints of all known information¹.

The exact functional form of this entropy is fully determined by three basic requirements (see Jaynes 2003; Caticha 2008), namely that local information only should have local effects, that the entropy is invariant under coordinate system transformations, and that independent systems of knowledge can equally well be treated together or separately.

This analysis results in the functional²

$$\mathcal{S}(\mathcal{P}|\mathcal{Q}) = - \int \mathcal{D}\phi \mathcal{P}(\phi) \log \left(\frac{\mathcal{P}(\phi)}{\mathcal{Q}(\phi)} \right), \quad (\text{A.1})$$

where $\mathcal{P}(\phi)$ is a probability distribution for a field ϕ about which the state of information should be evaluated (for example the signal field s), $\mathcal{Q}(\phi)$ denotes a probability distribution reflecting a possible a priori state of knowledge, and $\int \mathcal{D}\phi$ is a phase space integral over all possible configurations of the field ϕ .

¹This is the same principle that is applied in statistical physics to derive the basic Boltzmann distributions (Huang 1963).

²In information theory, this functional is also known as the *Kullback-Leibler divergence*, used to measure the distance of the probability distributions \mathcal{P} and \mathcal{Q} (see Caticha 2008).

Any available information usually is expressed in form of statistical quantities like the mean or higher moments of $\mathcal{P}(\phi)$, and gets incorporated into the assignment using Lagrange multipliers to constrain the maximization.

Under the MEP, the claim (see Secs. 1.1.1, 1.2.4 and 2.2.3) can be verified that a priori knowledge of only the zeroth moment (normalization), first moment (mean) φ , and second moment $\Phi + \varphi\varphi^\dagger$, with a flat prior \mathcal{Q} chosen to be constant, enforce a choice of a Gaussian probability distribution for $\mathcal{P}(\phi)$. Maximizing (A.1) under the three constraints

$$\begin{aligned}\langle 1 \rangle &= 1, \\ \langle \phi \rangle &= \varphi, \\ \langle \phi\phi^\dagger \rangle &= \Phi + \varphi\varphi^\dagger,\end{aligned}\tag{A.2}$$

yields $\mathcal{P}(\phi) = \mathcal{G}(\phi - \varphi | \Phi)$ (see e.g. Caticha 2008).

Bibliography

Aharonian, F., Arshakian, T. G., Allen, B., Banerjee, R., Beck, R., Becker, W., Bomans, D. J., Breitschwerdt, D., Brüggen, M., Brunthaler, A., Catinella, B., Champion, D., Ciardi, B., Crocker, R., de Avillez, M. A., Dettmar, R. J., Engels, D., Enßlin, T., Enke, H., Fieseler, T., Gizon, L., Hackmann, E., Hartmann, B., Henkel, C., Hoeft, M., Iapichino, L., Innes, D., James, C., Jasche, J., Jones, D., Kagramanova, V., Kauffmann, G., Keane, E., Kerp, J., Klöckner, H.-R., Kokkotas, K., Kramer, M., Krause, M., Krause, M., Krupp, N., Kunz, J., Lämmerzahl, C., Lee, K. J., List, M., Liu, K., Lobanov, A., Mann, G., Merloni, A., Middelberg, E., Niemeyer, J., Noutsos, A., Perlick, V., Reich, W., Richter, P., Roy, A., Saintonge, A., Schäfer, G., Schaffner-Bielich, J., Schinnerer, E., Schleicher, D., Schneider, P., Schwarz, D. J., Sedrakian, A., Sesana, A., Smolčić, V., Solanki, S., Tuffs, R., Vetter, M., Weber, E., Weller, J., Wex, N., Wucknitz, O., and Zwaan, M. (2013). Pathway to the Square Kilometre Array - The German White Paper -. *ArXiv e-prints*.

ALMA press group (2012). Press Review: The world lays its eyes on ALMA/ A hundred media glances about early science. *ALMA Newsletter*, 9:35.

Amsler, C., Doser, M., Antonelli, M., Asner, D., Babu, K., Baer, H., Band, H., Barnett, R., Bergren, E., Beringer, J., Bernardi, G., Bertl, W., Bichsel, H., Biebel, O., Bloch, P., Blucher, E., Blusk, S., Cahn, R., Carena, M., Caso, C., Ceccucci, A., Chakraborty, D., Chen, M.-C., Chivukula, R., Cowan, G., Dahl, O., D'Ambrosio, G., Damour, T., de Gouva, A., DeGrand, T., Dobrescu, B., Drees, M., Edwards, D., Eidelman, S., Elvira, V., Erler, J., Ezhela, V., Feng, J., Fetscher, W., Fields, B., Foster, B., Gaiser, T., Garren, L., Gerber, H.-J., Gerbier, G., Gherghetta, T., Giudice, G., Goodman, M., Grab, C., Gribsan, A., Grivaz, J.-F., Groom, D., Grnewald, M., Gurtu, A., Gutsche, T., Haber, H., Hagiwara, K., Hagemann, C., Hayes, K., Hernandez-Rey, J., Hikasa, K., Hinchliffe, I., Hcker, A., Huston, J., Igo-Kemenes, P., Jackson, J., Johnson, K., Junk, T., Karlen, D., Kayser, B., Kirkby, D., Klein, S., Knowles, I., Kolda, C., Kowalewski, R., Kreitz, P., Krusche, B., Kuyanov, Y., Kwon, Y., Lahav, O., Langacker, P., Liddle, A., Ligeti, Z., Lin, C.-J., Liss, T., Littenberg, L., Liu, J., Lugovsky, K., Lugovsky, S., Mahlke, H., Mangano, M., Mannel, T., Manohar, A., Marciano, W., Martin, A., Masoni, A., Milstead, D., Miquel, R., Mnig, K., Murayama, H., Nakamura, K., Narain, M., Nason, P., Navas, S., Nevski, P., Nir, Y., Olive, K., Pape, L., Patrignani, C., Peacock, J., Piepke, A., Punzi, G., Quadt, A., Raby, S., Raffelt, G., Ratcliff, B., Renk, B., Richardson, P.,

- Roesler, S., Rolli, S., Romaniouk, A., Rosenberg, L., Rosner, J., Sachrajda, C., Sakai, Y., Sarkar, S., Sauli, F., Schneider, O., Scott, D., Seligman, W., Shaevitz, M., Sjstrand, T., Smith, J., Smoot, G., Spanier, S., Spieler, H., Stahl, A., Stanev, T., Stone, S., Sumiyoshi, T., Tanabashi, M., Terning, J., Titov, M., Tkachenko, N., Trnqvist, N., Tovey, D., Trilling, G., Trippe, T., Valencia, G., van Bibber, K., Vinciter, M., Vogel, P., Ward, D., Watari, T., Webber, B., Weiglein, G., Wells, J., Whalley, M., Wheeler, A., Wohl, C., Wolfenstein, L., Womersley, J., Woody, C., Workman, R., Yamamoto, A., Yao, W.-M., Zenin, O., Zhang, J., Zhu, R.-Y., Zyla, P., Harper, G., Lugovsky, V., and Schaffner, P. (2008). Review of particle physics. *Physics Letters B*, 667(15):1 – 6. `jc:title;Review of Particle Physics;ce:title;.`
- Armstrong, J. W., Cordes, J. M., and Rickett, B. J. (1981). Density power spectrum in the local interstellar medium. *Nature*, 291:561–564.
- Beatty, P. J., Nishimura, D. J., and Pauly, J. M. (2005). Rapid gridding reconstruction with a minimal oversampling ration. *IEEE Trans Med Imaging*.
- Bell, M. R. and Enßlin, T. A. (2012). Faraday synthesis. The synergy of aperture and rotation measure synthesis. *A&A*, 540:A80.
- Bell, M. R., Junklewitz, H., and Enßlin, T. A. (2011). Faraday caustics. Singularities in the Faraday spectrum and their utility as probes of magnetic field properties. *A&A*, 535:A85.
- Bhatnagar, S. and Cornwell, T. J. (2004). Scale sensitive deconvolution of interferometric images. I. Adaptive Scale Pixel (Asp) decomposition. *A&A*, 426:747–754.
- Booth, R. S. and Jonas, J. L. (2012). An Overview of the MeerKAT Project. *African Skies*, 16:101.
- Born, M. and Wolf, E. (1999). *Principles of Optics*.
- Bracewell, R. (1965). *The Fourier Transform and its applications*.
- Brandenburg, A. and Subramanian, K. (2005). Astrophysical magnetic fields and nonlinear dynamo theory. *Phys. Rep.*, 417:1–209.
- Brentjens, M. A. and de Bruyn, A. G. (2005). Faraday rotation measure synthesis. *A&A*, 441:1217–1228.
- Briggs, D. S. (1995a). High Fidelity Interferometric Imaging: Robust Weighting and NNLS Deconvolution. In *American Astronomical Society Meeting Abstracts*, volume 27 of *Bulletin of the American Astronomical Society*, page 112.02.
- Briggs, D. S. (1995b). Phd thesis.

- Burke, B. F. (2006). Planetary Radio Astronomy, Fifty Years Ago and Fifty Years Hence. In Rucker, H. O., Kurth, W., and Mann, G., editors, *Planetary Radio Emissions VI*, page 1.
- Burn, B. J. (1966). On the depolarization of discrete radio sources by Faraday dispersion. *MNRAS*, 133:67.
- Candes, E. J., Romberg, J. K., and Tao, T. (2006). Stable signal recovery from incomplete and inaccurate measurements. *Comm. Pure Appl. Math.*, 59:1207–1223.
- Carrillo, R. E., McEwen, J. D., and Wiaux, Y. (2012). Sparsity Averaging Reweighted Analysis (SARA): a novel algorithm for radio-interferometric imaging. *MNRAS*, 426:1223–1234.
- Carrillo, R. E., McEwen, J. D., and Wiaux, Y. (2013). Why CLEAN when you can PURIFY? *ArXiv e-prints*.
- Caticha, A. (2008). Lectures on Probability, Entropy, and Statistical Physics. *ArXiv e-prints*.
- Clark, B. G. (1980). An efficient implementation of the algorithm 'CLEAN'. *A&A*, 89:377.
- Clarke, T. E. and Enßlin, T. A. (2006). Deep 1.4 GHz Very Large Array Observations of the Radio Halo and Relic in Abell 2256. *AJ*, 131:2900–2912.
- Conway, J. E., Cornwell, T. J., and Wilkinson, P. N. (1990). Multi-Frequency Synthesis - a New Technique in Radio Interferometric Imaging. *MNRAS*, 246:490.
- Cook, J. (2011). An introduction to solar radio astronomy. *Journal of the British Astronomical Association*, 121:241–243.
- Cooley, J. W. and Tukey, J. W. (1965). An algorithm for the machine calculation of complex Fourier series. *Math. Comp.*, 19:297 – 301.
- Cornwell, T. J. (2008). Multiscale CLEAN Deconvolution of Radio Synthesis Images. *IEEE Journal of Selected Topics in Signal Processing*, 2:793–801.
- Cornwell, T. J. and Evans, K. F. (1985). A simple maximum entropy deconvolution algorithm. *A&A*, 143:77–83.
- Cornwell, T. J., Golap, K., and Bhatnagar, S. (2008). The Noncoplanar Baselines Effect in Radio Interferometry: The W-Projection Algorithm. *IEEE Journal of Selected Topics in Signal Processing*, 2:647–657.
- Cox, R. T. (1990). Probability, frequency and reasonable expectation. pages 353–365.

- D'Amico, N. (2005). Observations of Radio Pulsars. In Baykal, A., Yerli, S. K., Inam, S. C., and Grebenev, S., editors, *NATO ASIB Proc. 210: The Electromagnetic Spectrum of Neutron Stars*, page 147.
- de Bruyn, A. G. and LOFAR EoR Key Science Project Team (2012). Lofar Deep Imaging And Prospects For Detecting The Eor. In *American Astronomical Society Meeting Abstracts 219*, volume 219 of *American Astronomical Society Meeting Abstracts*, page 214.05.
- Donoho, D. L. (2006). Compressed sensing. *IEEE Transactions of Information Theory*, 52:1289–1306.
- Eilek, J. A. (1989). Magnetic fields in clusters of galaxies. *Bulletin of the American Physical Society*, 34:1286.
- Enßlin, T. (2013). Information field theory. In von Toussaint, U., editor, *American Institute of Physics Conference Series*, volume 1553 of *American Institute of Physics Conference Series*, pages 184–191.
- Enßlin, T. A. and Frommert, M. (2011). Reconstruction of signals with unknown spectra in information field theory with parameter uncertainty. *Phys. Rev. D*, 83(10):105014.
- Enßlin, T. A., Frommert, M., and Kitaura, F. S. (2009). Information field theory for cosmological perturbation reconstruction and nonlinear signal analysis. *Phys. Rev. D*, 80(10):105005.
- Enßlin, T. A., Junklewitz, H., Winderling, L., and Selig, M. (2013). Art and theory of self-calibration. *ArXiv e-prints*.
- Enßlin, T. A. and Weig, C. (2010). Inference with minimal Gibbs free energy in information field theory. *Phys. Rev. E*, 82(5):051112.
- Feretti, L., Giovannini, G., Govoni, F., and Murgia, M. (2012). Clusters of galaxies: observational properties of the diffuse radio emission. *A&A Rev.*, 20:54.
- Finley, D. G. and Goss, W. M., editors (2000). *Radio interferometry : the saga and the science*.
- Fletcher, A. and Shukurov, A. (2007). Depolarization canals and interstellar turbulence. In Miville-Deschênes, M.-A. and Boulanger, F., editors, *EAS Publications Series*, volume 23 of *EAS Publications Series*, pages 109–128.
- Gaensler, B. M., Dickey, J. M., McClure-Griffiths, N. M., Green, A. J., Wieringa, M. H., and Haynes, R. F. (2001). Radio Polarization from the Inner Galaxy at Arcminute Resolution. *ApJ*, 549:959–978.

- Garrett, M. A. (2012). Radio Astronomy Transformed: Aperture Arrays - Past, Present Future. In *From Antikythera to the Square Kilometre Array: Lessons from the Ancients*.
- Geman, S. and Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- Green, D. A. (2001). Galactic Supernova Remnants: An overview of their radio properties. In Aharonian, F. A. and Völk, H. J., editors, *American Institute of Physics Conference Series*, volume 558 of *American Institute of Physics Conference Series*, pages 59–70.
- Greiner, M. (2013). The galactic free electron density: A bayesian reconstruction, master thesis.
- Gull, S. F. and Daniell, G. J. (1979). The Maximum Entropy Method (invited Paper). In van Schooneveld, C., editor, *IAU Colloq. 49: Image Formation from Coherence Functions in Astronomy*, volume 76 of *Astrophysics and Space Science Library*, page 219.
- Hastings, W. K. (1970). Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57:97 – 109.
- Haverkorn, M., Katgert, P., and de Bruyn, A. G. (2000). Structure in the local Galactic ISM on scales down to 1 pc, from multi-band radio polarization observations. *A&A*, 356:L13–L16.
- Haverkorn, M. and Spangler, S. R. (2013). Plasma Diagnostics of the Interstellar Medium with Radio Astronomy. *Space Sci. Rev.*, 178:483–511.
- Heald, G. and LOFAR collaboration (2014). The LOFAR Multifrequency Snapshot Sky Survey (MSSS): Status and Results. In *American Astronomical Society Meeting Abstracts*, volume 223 of *American Astronomical Society Meeting Abstracts*, page 236.07.
- Hirabayashi, H. (2004). High Resolution Radio Observations of AGN. *Progress of Theoretical Physics Supplement*, 155:178–185.
- Högbom, J. A. (1974). Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines. *A&AS*, 15:417.
- Hogg, R. and Tanis, E. (2006). *Probability and Statistical Inference*. Pearson Prentice Hall.
- Huang, K. (1963). *Statistical Mechanics*.
- Jasche, J., Kitaura, F. S., Wandelt, B. D., and Enßlin, T. A. (2010). Bayesian power-spectrum inference for large-scale structure data. *MNRAS*, 406:60–85.
- Jaynes, E. T. (1982). On the Rationale of Maximum Entropy Methods. *IEEE Proceedings*, 70:939–952.

- Jaynes, E. T. (2003). *Probability Theory: The Logic of Science*.
- Junklewitz, H., Bell, M. A., and Enßlin, T. (2014). A new approach to multi-frequency synthesis in radio interferometry. *ArXiv e-prints*.
- Junklewitz, H., Bell, M. R., Selig, M., and Enßlin, T. A. (2013). RESOLVE: A new algorithm for aperture synthesis imaging of extended emission in radio astronomy. *ArXiv e-prints (Res2013)*.
- Junklewitz, H. and Enßlin, T. A. (2011). Imprints of magnetic power and helicity spectra on radio polarimetry statistics. *A&A*, 530:A88.
- Karakci, A., Sutter, P. M., Zhang, L., Bunn, E. F., Korotkov, A., Timbie, P., Tucker, G. S., and Wandelt, B. D. (2013). Bayesian Inference of Polarized Cosmic Microwave Background Power Spectra from Interferometric Data. *ApJS*, 204:10.
- Kassim, N. E., Lazio, T. J. W., Lane, W. M., Cohen, A. S., Clarke, T. E., Brogan, C. L., and Perley, R. A. (2005). Probing Relativistic Plasmas in Supernova Remnants, Radio Galaxies, and Clusters at Long Radio Wavelengths. In Sjouwerman, L. O. and Dyer, K. K., editors, *X-Ray and Radio Connections*.
- Kolmogorov, A. (1941). The Local Structure of Turbulence in Incompressible Viscous Fluid for Very Large Reynolds' Numbers. *Akademiia Nauk SSSR Doklady*, 30:301–305.
- Kuchar, P. and Enßlin, T. A. (2011). Magnetic power spectra from Faraday rotation maps. REALMAF and its use on Hydra A. *A&A*, 529:A13.
- Lannes, A., Anterrieu, E., and Marechal, P. (1997). Clean and Wipe. *A&AS*, 123:183–198.
- Lazarian, A. and Pogosyan, D. (2012). Statistical Description of Synchrotron Intensity Fluctuations: Studies of Astrophysical Magnetic Turbulence. *ApJ*, 747:5.
- Lecavelier Des Etangs, A., Sirothia, S., Gopal-Krishna, and Zarka, P. (2011). Radio emission from exoplanets. In *EPSC-DPS Joint Meeting 2011*, page 992.
- Lemm, J. C. (1999). Bayesian Field Theory: Nonparametric Approaches to Density Estimation, Regression, Classification, and Inverse Quantum Problems. *ArXiv Physics e-prints*.
- Lu, R., Fish, V. L., Doeleman, S., Monnier, J. D., and Baron, F. (2014). Black holes under the microscope: Prospects for imaging with the Event Horizon Telescopes. In *American Astronomical Society Meeting Abstracts*, volume 223 of *American Astronomical Society Meeting Abstracts*, page 443.05.
- Miley, G. and De Breuck, C. (2008). Distant radio galaxies and their environments. *A&A Rev.*, 15:67–144.

- Mood, A. M., Graybill, F. A., and Duane, C. B. (1974). *Introduction to the theory of statistics*.
- Neal, R. M. (1993). Probabilistic inference using markov chain monte carlo methods. *Technical Report CRG-TR-93-1, Dept. of Computer Science, University of Toronto*.
- Oppermann, N., Junklewitz, H., Robbers, G., and Enßlin, T. A. (2011a). Probing magnetic helicity with synchrotron radiation and Faraday rotation. *A&A*, 530:A89.
- Oppermann, N., Robbers, G., and Enßlin, T. A. (2011b). Reconstructing signals from noisy data with unknown signal and noise covariance. *Phys. Rev. E*, 84(4):041118.
- Oppermann, N., Selig, M., Bell, M. R., and Enßlin, T. A. (2013). Reconstruction of Gaussian and log-normal fields with spectral smoothness. *Phys. Rev. E*, 87(3):032136.
- Ostriker, E. C. (2006). Book Review: PLASMA PHYSICS FOR ASTROPHYSICS / Princeton U. Press, 2005. *Physics Today*, 59(1):010000.
- Perley, R. A., Chandler, C. J., Butler, B. J., and Wrobel, J. M. (2011). The Expanded Very Large Array: A New Telescope for New Science. *ApJ*, 739:L1.
- Peskin, M. E. and Schroeder, D. V. (1995). *An Introduction to Quantum Field Theory*. Westview Press.
- Planck Collaboration, Ade, P. A. R., Aghanim, N., Armitage-Caplan, C., Arnaud, M., Ashdown, M., Atrio-Barandela, F., Aumont, J., Baccigalupi, C., Banday, A. J., and et al. (2013). Planck 2013 results. I. Overview of products and scientific results. *ArXiv e-prints*.
- Rau, U. and Cornwell, T. J. (2011). A multi-scale multi-frequency deconvolution algorithm for synthesis imaging in radio interferometry. *A&A*, 532:A71.
- Reich, W., Fürst, E., Reich, P., Uyaniker, B., Wielebinski, R., and Wolleben, M. (2004). The Effelsberg 1.4 GHz Medium Galactic Latitude Survey (EMLS). In Uyaniker, B., Reich, W., and Wielebinski, R., editors, *The Magnetized Interstellar Medium*, pages 45–50.
- Reid, R. I. and CASA Team (2010). CASA: Common Astronomy Software Applications. In *American Astronomical Society Meeting Abstracts 215*, volume 42 of *Bulletin of the American Astronomical Society*, page 479.04.
- Rybicki, G. B. and Lightman, A. P. (1985). *Radiative processes in astrophysics*.
- Ryle, M. and Hewish, A. (1960). The synthesis of large radio telescopes. *MNRAS*, 120:220.
- Sault, R. J. and Oosterloo, T. A. (2007). Imaging algorithms in radio interferometry. *ArXiv Astrophysics e-prints*.

- Sault, R. J. and Wieringa, M. H. (1994). Multi-frequency synthesis techniques in radio interferometric imaging. *A&AS*, 108:585–594.
- Schekochihin, A. A. and Cowley, S. C. (2007). *Turbulence and Magnetic Fields in Astrophysical Plasmas*, page 85. Springer.
- Schinckel, A. E., Bunton, J. D., Cornwell, T. J., Feain, I., and Hay, S. G. (2012). The Australian SKA Pathfinder. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 8444 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*.
- Schnack, D. D. (2009). *Lectures in Magnetohydrodynamics*.
- Schwab, F. R. (1984). Relaxing the isoplanatism assumption in self-calibration; applications to low-frequency radio interferometry. *AJ*, 89:1076–1081.
- Selig, M., Bell, M. R., Junklewitz, H., Oppermann, N., Reinecke, M., Greiner, M., Pachajoa, C., and Enßlin, T. A. (2013). NIFTY - Numerical Information Field Theory. A versatile PYTHON library for signal inference. *A&A*, 554:A26.
- Selig, M. and Enßlin, T. (2013). D3PO - Denoising, Deconvolving, and Decomposing Photon Observations. *ArXiv e-prints*.
- Selig, M., Oppermann, N., and Enßlin, T. A. (2012). Improving stochastic estimates with inference methods: Calculating matrix diagonals. *Phys. Rev. E*, 85(2):021134.
- SKA press group (2012). News Note: News from the SKA. *Monthly Notes of the Astronomical Society of South Africa*, 71:51–54.
- Skilling, J., Strong, A. W., and Bennett, K. (1979). Maximum-entropy image processing in gamma-ray astronomy. *MNRAS*, 187:145–152.
- Smirnov, O. M. (2011a). Revisiting the radio interferometer measurement equation. I. A full-sky Jones formalism. *A&A*, 527:A106.
- Smirnov, O. M. (2011b). Revisiting the radio interferometer measurement equation. II. Calibration and direction-dependent effects. *A&A*, 527:A107.
- Spangler, S. R. (1982). The transport of polarized synchrotron radiation in a turbulent medium. *ApJ*, 261:310–320.
- Spangler, S. R. (1983). Determination of the properties of magnetic turbulence in radio sources. *ApJ*, 271:L49–L53.
- Subramanian, K. (1999). Unified Treatment of Small- and Large-Scale Dynamos in Helical Turbulence. *Physical Review Letters*, 83:2957–2960.

- Sutter, P. M., Wandelt, B. D., and Malu, S. S. (2012). Bayesian Angular Power Spectrum Analysis of Interferometric Data. *ApJS*, 202:9.
- Sutter, P. M., Wandelt, B. D., McEwen, J. D., Bunn, E. F., Karakci, A., Korotkov, A., Timbie, P., Tucker, G. S., and Zhang, L. (2013). Probabilistic image reconstruction for radio interferometers. *ArXiv e-prints*.
- Taylor, G. B., Carilli, C. L., and Perley, R. A., editors (1999). *Synthesis Imaging in Radio Astronomy II*, volume 180 of *Astronomical Society of the Pacific Conference Series*.
- Tennekes, H. and Lumley, J. L. (1972). *First Course in Turbulence*.
- Thompson, A. R., Moran, J. M., and Swenson, G. W. (1986). *Interferometry and synthesis in radio astronomy*.
- Transtrum, M. K. and Sethna, J. P. (2012). Improvements to the Levenberg-Marquardt algorithm for nonlinear least-squares minimization. *ArXiv e-prints*.
- Upton, G., Cook, I., and Cook, I. (2008). *A Dictionary of Statistics*. Opr Series. OUP Oxford.
- Uyaniker, B., Fuerst, E., Reich, W., Reich, P., and Wielebinski, R. (1998). A 1.4 GHz radio continuum and polarization survey at medium Galactic latitudes. I. Observation and reduction technique. *A&AS*, 132:401–411.
- Waelkens, A. H., Schekochihin, A. A., and Enßlin, T. A. (2009). Probing magnetic turbulence by synchrotron polarimetry: statistics and structure of magnetic fields from Stokes correlators. *MNRAS*, 398:1970–1988.
- Wiaux, Y., Jacques, L., Puy, G., Scaife, A. M. M., and Vandergheynst, P. (2009). Compressed sensing imaging techniques for radio interferometry. *MNRAS*, 395:1733–1742.
- Widrow, L. M. (2002). Origin of galactic and extragalactic magnetic fields. *Reviews of Modern Physics*, 74:775–823.
- Wilson, T. L., Rohlfs, K., and Huttemeister, S. (2012). *Tools of Radio Astronomy, 5th edition*.
- Wise, M. W. and The LOFAR Surveys KSP Cluster Working Group (2014). Recent Results on Clusters of Galaxies with LOFAR. In *American Astronomical Society Meeting Abstracts*, volume 223 of *American Astronomical Society Meeting Abstracts*, page 431.05.

Acknowledgments

There are some people who have accompanied me through the last years, and who have contributed to the success of this thesis in many ways.

First and foremost my thesis advisor Torsten. I am very glad for the time that I worked with him, and learned from him, since he is a supervisor as caring and helpful as I could have imagined. There almost wasn't a single day where he wouldn't have listened to problems or suggestions. The dedication to his students' projects, and his sincerity in always promoting and backing their work in public, to me, is just amazing. I enjoy that with his habilitation, he now officially can appear on this thesis.

Next Mike, without whom this work would have been impossible, and from whom I basically learned all I know about good computing.

Further, all the members of Torsten's group at MPA that I had the pleasure to meet and work with during a great time. Niels, who went along with me from the beginning and through countless discussions and many experiences, of scientific and sometimes more important nature. Marco, whose nifty work and ever-lasting patience with volume factors and numerical problems eases the life of every one of us and will continue to do so for all future students in the group. Maksim, with whom I had much to few time of common work at MPA that saw me not in stress, but the best time on a conference trip ever. All the other students, undergrads and internship students that I have seen coming and going over the time: Sebastian, Mahsa, Vanessa, Ashmeet, Lars, Helin, Max & Max, Mona and Jens. And of course the members of the MPA Planck group, Martin, Jörg, Wolfgang, Uwe, Theo and Thomas who helped me out with countless computer problems.

I should also thank all other people and members of the MPA that were around over the years with help, guidance and company: The directors, and especially Simon White for acting as my official advisor, the secretaries and the administration for all the help in the small things that keep everything going, Marat for being a member of my committee, and the many PhD students and postdocs working and living at the institute.

I like to grant special thanks to Rick Perley, Oleg Smirnov, Annalisa Bonafede and Tracy C. Clarke for introducing me into the deeper aspects of observational radio astronomy. Tracy should receive special thanks for her courtesy in offering access to her data on Abell 2256.

Finally, I want to thank all the people outside MPA that have been a part of my life, and who all have contributed to this work in their own ways: Robert, Keno, Christian and Tom for their patience with a stressed old friend; Johannes for the music; Quirin and

Stefanie for their friendship and help; Kathrin for her love; and my family for just always being there.

Curriculum Vitae

Henrik Junklewitz

Friedrichshafenerstr. 39

Tel.: +49(0)171 9103746

81243 Munich, Germany

Email: henrikju@mpa-garching.mpg.de

Finishing PhD student at the Max-Planck-Institute
for Astrophysics in Garching (MPA)

Postdoc at Argelander-Institut für Astronomie in Bonn

Personal Data

Date of birth: September 23rd, 1982

Place of birth: Wilhelmshaven, Germany

Nationality: German

Education and Science

- since 2/2014* **Postdoctoral Researcher**
at the Argelander-Institut für Astronomie in Bonn
- 2010 - present* **PhD studies at the MPA Garching**
Thesis: "Statistical Inference in Radio Astronomy"
Advisors: *Dr. Torsten Enßlin, Dr. Michael R. Bell*
- 2010* **Diploma degree in physics**
Ludwig Maximilians-Universität München (LMU)
Thesis: *Imprints of magnetic power and helicity spectra on radio polarimetry statistics*
Advisor: *Dr. Torsten Enßlin*
- 2006 - 2007* **Studies of physics**
studying abroad with the ERASMUS program at Universite Paris sud in Orsay, France
- 2003 - 2010* **Studies of physics**
Ludwig Maximilians-Universität München (LMU)
- 2002 - 2003* **Civilian service**
social work for school kids with mental and physical disabilities
- 1990 - 2002* **Abitur (high school degree)**
Integrierte Gesamtschule Wilhelmshaven

Teaching

- 2011* Conducting the exercises for the summer term lecture in "Information theory and signal reconstruction" at the LMU.

Programming skills

- | | |
|--|---------------------|
| PYTHON | Extensive knowledge |
| AIPS, CASA
(radio-astronomical software packages) | Good knowledge |
| C/C++ | Basic knowledge |

Publications

“A new approach to multi-frequency synthesis in radio interferometry“ - H. Junklewitz, M.R. Bell, T.A. Enßlin - 2014, *submitted to Astronomy & Astrophysics*

“Art and theory of self-calibration“ - T.A. Enßlin, H. Junklewitz, L. Winderling, M. Selig - 2013, *submitted to Astronomy & Astrophysics*

“RESOLVE: A new algorithm for aperture synthesis imaging of extended emission in radio astronomy“ - H. Junklewitz, M.R. Bell, M. Selig, T.A. Enßlin - 2013, *submitted to Astronomy & Astrophysics*

“NIFTY - Numerical Information Field Theory. A versatile PYTHON library for signal inference“ - M. Selig, M.R. Bell, H. Junklewitz, N. Oppermann, M. Reinecke, M. Greiner, C.Pachajoa, T.A. Enßlin - 2013, *Astronomy & Astrophysics, Vol. 554, A26*

“An improved map of the Galactic Faraday sky“ - N. Oppermann, H. Junklewitz, G. Robbers, M. R. Bell, T. A. Enßlin, et al. - 2012, *Astronomy & Astrophysics, Vol. 542, A93*

”Faraday caustics: Singularities in the Faraday spectrum and their utility as probes of magnetic field properties“ - M. R. Bell, H. Junklewitz and T. A. Enßlin - 2011, *Astronomy & Astrophysics, Vol. 535, A85*

”Probing magnetic helicity with synchrotron radiation and Faraday rotation“ - N. Oppermann, H. Junklewitz, G. Robbers, T. A. Enßlin - 2011, *Astronomy & Astrophysics, Vol. 530, A89*

”Imprints of magnetic power and helicity spectra on radio polarimetry statistics“ - H. Junklewitz, T. A. Enßlin - 2011, *Astronomy & Astrophysics, Vol. 530, A88*

Presentations and conference talks

- 2013 " *Faraday Synthesis - The synergy of aperture and rotation measure synthesis* ", invited talk, given at the meeting on Galactic magnetism in the Era of LOFAR and SKA in Stockholm, Sweden
- 2013 " *Bayesian Approaches on Imaging in Radio Interferometry* ", talk given at the 3GC3 calibration and imaging meeting in Port Alfred, South Africa
- 2012 " *High fidelity imaging and magnetic field inference* ", talk given at the meeting of the German Astronomical Union 2012 in Hamburg, Germany
- 2012 " *A new approach to multi-frequency aperture synthesis* ", talk given at the meeting of the DFG Research Unit RU 1254 "Magnetisation of interstellar and intergalactic media" in Mainz, Germany
- 2012 " *A new approach to multi-frequency aperture synthesis* ", talk given at the meeting of the magnetic key science project of LOFAR in Gothenburg, Sweden
- 2011 " *The magnetic correlation tensor - a short introduction* ", blackboard talk at the DFG RU 1254 summer school at castle Ringberg, Germany
- 2011 " *Magnetic field statistics* ", talk given at the german-russian meeting on magnetic fields in Puschino, Russia
- 2010 " *Imprints of magnetic power and helicity spectra on radio polarimetry statistics* ", talk given at the meeting on Rotation measure analysis of magnetic fields in and around radio galaxies in Riccione, Italy
- 2010 " *Imprints of magnetic power and helicity spectra on radio polarimetry statistics* ", invited talk, held at the ISSI workshop on large-scale magnetic fields in the universe, in Bern, Switzerland

Public outreach

2011 ”Das Radioteleskop LOFAR und was wir damit über das Universum lernen können“ , public science talk in Weilach, Germany, where one german LOFAR station is located

Languages

German: native

English: fluent

Spanish: advanced

French: advanced