Signal inference in Galactic astrophysics

Niels Oppermann

München 2013

Signal inference in Galactic astrophysics

Niels Oppermann

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

vorgelegt von
Niels Oppermann
aus Göttingen

München, den 8. Januar 2013

Erstgutachter: Prof. Dr. Simon D. M. White
Zweitgutachter: Prof. Dr. Gerhard Boerner
Tag der mündlichen Prüfung: 26. Februar 2013

If you shut the door to all errors, truth will be shut out.

Rabrindranath Tagore

Contents

List of Figures

List of Tables

Zusammenfassung

Diese Arbeit enthält eine Kombination aus methodologischer Arbeit mit sehr weitem Fokus und einigen konkreten astrophysikalischen Anwendungen. Der Wissensstand über das galaktische interstellare Medium wird durch Studien neuer Methoden für die Entschlüsselung ausgewählter Eigenschaften im Zusammenhang mit seinem Magnetfeld erweitert. Dabei werden die notwendigen statistischen Werkzeuge auf rigorose Weise aus wahrscheinlichkeitstheoretischen Überlegungen hergeleitet.

Eine Größe, welche die statistischen Eigenschaften von Magnetfeldern beschreibt, ist deren Helizität. Eine jüngst entwickelte Methode zum Nachweis magnetischer Helizität unter Benutzung astronomischer Beobachtungen wird hier auf Daten der Milchstraße angewandt. Dies liefert zwar kein positives Ergebnis, könnte aber durch die komplexe Messsituation bedingt sein. Um dies zu klären wird eine Reihe von galaktischen Simulationen vorgenommen. So kann gezeigt werden, dass die Helizitätsdetektion mit dieser Methode in Fällen fehlschlägt, in denen die zugrundeliegende Elektronendichte zu stark variiert. Folglich kann geschlossen werden, dass dies in der Milchstraße entweder der Fall ist oder das galaktische Magnetfeld nicht helisch ist.

Desweiteren wird eine Methode zur Rekonstruktion kontinuierlicher Signale aus verrauschten Daten entwickelt, für die das Rauschniveau unbekannt ist. Dies ist unter anderem nötig um eine korrekte Anwendung des Helizitätstests zu gewährleisten. Eine solche Rekonstruktion wird durch die Verwendung der räumlichen Korrelationsstruktur des Signals möglich, die aus dem selben Datensatz auf selbstkonsistente Weise rekonstruiert wird. Fluktuationen in den Daten, die inkonsistent mit dieser Korrelationsstruktur sind, werden dann dem Rauschen in den Daten zugeschrieben.

Die Entwicklung dieser Methode ist zum Teil durch das Ziel motiviert, eine Himmelskarte des galaktischen Beitrags zum astronomischen Faradayrotationseffekt zu erstellen. Diese liefert Informationen sowohl über das galaktische Magnetfeld als auch über die Dichte der freien thermischen Elektronen. Da die beobachtete Größe die Faradayrotation einer Radioquelle ist, die durch sämtliche Magnetfelder zwischen der Quelle und dem Beobachter beeinflusst wird, müssen extragalaktische Beitrage herausgefiltert werden. Die Methode zur Rekonstruktion bei unsicherem Rauschniveau wird benutzt um die extragalaktischen Beiträge – sowie einige andere Beobachtungseffekte – dem Rauschen in den Daten zuzuschreiben und so den galaktischen Beitrag zu separieren. Die resultierende Karte ist die detailgenaueste und präziseste ihrer Art und die einzige in der die extragalaktischen Beiträge herausgefiltert wurden.

Schlussendlich wird eine Methode zu Rekonstruktion von log-normal-verteilten Signal-feldern entwickelt, also von strikt positiven Größen, bei denen die Stärke der Fluktuationen über mehrere Größenordnungen schwankt. Dies geschieht mit Blick auf die Rekonstruktion von Emissionskarten für verschiedene Prozesse im galaktischen interstellaren Medium.

Summary

In this thesis we present a combination of methodological work with very wide focus and some specific astrophysical applications. We advance the knowledge on the Galactic interstellar medium by studying new ways of inferring select properties related to its magnetic field. We derive the statistical tools needed in a rigorous way from probabilistic considerations.

One quantity describing the statistical properties of magnetic fields is their helicity. We apply a recently developed technique to detect magnetic helicity from astronomical observations to data from the Milky Way. No indications of helicity are found. Using a series of simulations in the Galactic setting, we are able to show that the technique fails to detect helicity in cases in which the underlying electron density varies too strongly. Thus, we are able to conclude that either this is the case in the Milky Way or the Galactic magnetic field is non-helical.

We further develop a technique, needed among other things to enable the correct application of the helicity test, to reconstruct continuous signals from noisy data for which the noise level is unknown. To do this, we make use of the statistical correlation structure of the signal which we reconstruct from the same data set in a self-consistent way. Fluctuations in the data that are inconsistent with this correlation structure are then assigned to the data's error budget.

The development of this technique is partly motivated by the goal of creating an all-sky map of the Galactic contribution to the astronomical Faraday rotation effect, which probes both the Galactic magnetic field and the density of free thermal electrons. Since the quantity that is observed is the Faraday rotation of a radio source, influenced by all magnetic fields between the source and the observer, extragalactic contributions need to be filtered out. We use the technique for reconstructions in the presence of an uncertain degree of noisiness to assign the extragalactic contributions – as well as some other observational effects – to the error budget of the data and thus single out the Galactic contribution. The resulting map is the most detailed and precise map of its kind and the only one in which the extragalactic contributions have been filtered out.

Finally, we develop a method to reconstruct log-normal signal fields, i.e. strictly positive signal fields for which the strengths of the fluctuations vary over several orders of magnitude. This is done with a view to reconstructions of emission maps due to different processes in the Galactic interstellar medium.

Chapter 1

Introduction

In this thesis we study the Milky Way galaxy, especially aspects related to its magnetic field. In our attempts to learn about the Galactic magnetic field, we also explore the general question how to deduce knowledge from observations. This problem is tackled using the statistical framework of signal inference. The learning techniques that we develop within this framework are then applied in the Galactic setting.

In this introductory chapter, we start by setting the stage for studies of the Milky Way by summarizing what is known about the most important aspects of our home galaxy and introducing the physical processes that make these aspects tangible for astronomers. We then go on to introduce the statistical framework in which we will develop signal inference techniques in later chapters. We conclude this chapter by giving an outline of the content of the thesis.

1.1 Galactic astrophysics

In this section, we briefly introduce the physical background of the astronomical signals with which we will deal in later chapters. We start by giving a rough overview over the Milky Way galaxy, continue by describing the physical processes that make the interstellar medium observable, and finish by describing some constituents of the Milky Way galaxy, focussing on the components of the interstellar medium that will be important in this work.

1.1.1 The Milky Way

The Milky Way is a spiral galaxy, i.e. most of its stars, gas, and dust are situated within a flat disk. The gas and dust distribution is more concentrated in the vertical direction than the distribution of the stars. The mass density due to the different components of the disk can be approximated as

$$\rho(r, z) = \rho_0 \, e^{-r/r_0} \, e^{-|z|/z_0},\tag{1.1}$$

where ρ_0 is the density in the Galactic center, r is the radial distance from the center, and z is the vertical distance from the mid-plane of the disk. For the stellar disk, the scale

Figure 1.1: The near-infrared sky as seen by the *Diffuse Infrared Background Experiment* aboard the *Cosmic Background Explorer* satellite. Measurements at wavelengths of 1.25, 2.2, and 3.5 μm are represented by blue, green, and red colors, respectively. The main part of the radiation is starlight. The red areas show directions in which the shorter wavelengths have been absorbed and thus indicate the presence of large amounts of dust. The image was provided by NASA's Legacy Archive for Microwave Background Data Analysis (`http://lambda.gsfc.nasa.gov/`).

lengths are $r_0^{(\text{stellar})} \approx 3.5\,\text{kpc}$ and $z_0^{(\text{stellar})} \approx 330\,\text{pc}$. The gas and dust components have a similar radial extent but are somewhat flatter with $z_0^{(\text{gas, dust})} \approx 160\,\text{pc}$ (Maoz 2011).

Within the disk, the stellar, gas, and dust densities exhibit a spiral structure that might be caused by gravitational interaction and the resulting density waves (Lin & Shu 1964; Francis & Anderson 2009). The higher gas density in the spiral arms leads to enhanced star formation. Thus, the stellar population within the arms tends to be younger. The Galactic center is surrounded by a roughly spherical stellar bulge of about $1\,\text{kpc}$ in radius (Maoz 2011), which is made up of older stars. At the very center of the Milky Way resides a supermassive Black Hole of about $4 \times 10^6\,\text{M}_\odot$.

The solar system is located slightly above the midplane of the disk at a height of $z_\odot \approx 30\,\text{pc}$ and a distance of $r_\odot \approx 8\,\text{kpc}$ from the Galactic center (Maoz 2011). Unfortunately, experimental studies of the Milky Way and its components are only possible within the solar system. Thus, all observable quantities are always projections along the line of sight. Overcoming these projection effects is one of the great challenges of signal inference in the astrophysical context.

Figure 1.1 shows an example for such a projected observation of the Milky Way, a Mollweide projection of the near-infrared sky as measured by the *Diffuse Infrared Background Experiment* aboard the *Cosmic Background Explorer* satellite (e.g. Boggess et al. 1992). This image shall also serve to explain the Mollweide projection and the Galactic coordinate system that we will use throughout this work whenever we present full-sky maps. The Galactic plane, seen as a stripe on the sky from our vantage point, extends horizontally along the middle of the image and defines the great circle of zero Galactic latitude ($b = 0°$). The northern part ($0° < b < 90°$) is plotted in the top half, and the southern part ($0° > b > -90°$) in the bottom half. The Galactic center lies in the direction that is plotted in the center of the image at zero longitude and zero latitude ($l = 0°, b = 0°$). In the image, the stellar bulge is visible around the Galactic center. The eastern hemisphere ($0° < l < 180°$) is plotted to the left from the center, while the western hemisphere ($-180° < l < 0°$ or $180° < l < 360°$) is plotted to the right. The oval circumference corresponds to the half great circle of $l = \pm180°$.

The whole Galaxy is embedded in a halo of Dark Matter, whose existence and shape can be inferred from observations of stellar velocities (Battaglia et al. 2005, 2006). This halo is roughly spherical in shape and its radial mass density profile can be fitted by an NFW profile (Navarro et al. 1997) given by

$$\rho(r) = \frac{\rho_0}{\frac{r}{r_0}\left(1 + \frac{r}{r_0}\right)^2},$$ (1.2)

where ρ_0 is the central density and r_0 is a radial scale length. Such a fit leads to a total mass for the Dark Matter halo of about $10^{12}\,\mathrm{M_\odot}$ (Battaglia et al. 2005, 2006), roughly 100 times the mass of the visible matter in the Milky Way.

The interstellar medium, i.e. the ordinary matter that is not bound in compact objects such as stars and the electromagnetic fields that permeate this matter, will be the main focus of this thesis. Neutral and ionized gas, dust grains, and highly energetic charged particles confined by the interstellar magnetic field are the main actors in interstellar space. These interact with each other and with the stars through a multitude of different channels. Stars form heavy elements which are then distributed into interstellar space at the end of their lifetime in supernova explosions, thus enriching the gaseous phase of the Galaxy; gas atoms and molecules condense to form larger dust grains, large gas clouds collapse to form stars; magnetic fields exert pressure, changing or preventing such collapses; dust grains destroy each other in collisions or grow through adhesion; starlight and cosmic rays ionize the gas and break up molecules and dust grains; dust absorbs starlight and reemits the energy in the form of thermal infrared radiation; magnetic fields trap the ionized cosmic ray particles and guide the plasma in its movement, the plasma drags the magnetic field lines along as it moves; etc. (see Spitzer 2007, for a broader review).

Supernova explosions and inhomogeneous radiation pressure can cause turbulence in this complex multi-component fluid. The evolution of this turbulence will be influenced by the presence of a magnetic field. Vice versa, turbulent motions of ionized gas can effectively twist up magnetic field lines and thus enhance the field's strength. The origin

of the Galactic magnetic field, its large-scale and small-scale structure, and the details of its interaction with the other components of the interstellar medium are all far from being completely understood. Fortunately, some of the involved components and processes leave imprints on the electromagnetic radiation that passes through, or is emitted from, interstellar space and thus open up opportunities to observationally constrain the happenings there. Chapters 2 and 4 are examples for attempts to shed some light on the interstellar medium by making efficient use of these imprints.

1.1.2   Physical processes

The most direct way of imaging the gaseous phase of the interstellar medium is the observation of electromagnetic radiation that is emitted by the atoms or molecules. By identifying atomic or molecular transition lines in emission or absorption of background light, it is possible to study the chemical composition of the gas. Hotter gas exhibits wider lines than cold gas and therefore the temperature can also be determined spectroscopically.

By far the most common element in the interstellar medium is hydrogen. Thus, transitions in atomic hydrogen are often used to study the gas in the Galaxy. One of the most widely used transitions is the H_α line, i.e. the transition from the $n = 3$ level to the $n = 2$ level of hydrogen, where n is the principal quantum number. The intensity of this line can, after correction for dust absorption, be used as a proxy for the line of sight integrated squared density of free electrons,

$$E_\mathrm{m} = \int_0^\infty \mathrm{d}l \; n_\mathrm{e}^2, \tag{1.3}$$

known as emission measure (e.g. Reynolds et al. 1974; Finkbeiner 2003). Another prominent line is the hyperfine transition of the ground level of atomic hydrogen resulting in emission at a wavelength of $\lambda \approx 21\,\mathrm{cm}$, which is often used in radio-astronomical studies (e.g. McClure-Griffiths et al. 2009). By studying the exact frequency of the observed lines and associating the shifts to velocity differences via the Doppler effect, one can on the one hand constrain the velocities of gas clouds and on the other hand, constrain the gas distribution in the Milky Way by mapping velocities to distances via a model for the differential Galactic rotation (e.g. Kerr 1969; Lockman 2002; Nakanishi & Sofue 2003; Jones & Dickey 2012). Molecules with useful emission and absorption lines for these types of studies include OH (e.g. Dickey et al. 2012), H_2O (e.g. Batchelor et al. 1980; Walsh et al. 2011), and CO (e.g Nakanishi & Sofue 2006).

The Zeeman splitting of atomic and molecular energy levels due to the presence of a magnetic field can cause new line transitions and thus lead to observational signatures of the Milky Way's magnetic field. The polarization properties of these lines depend on the orientation of the magnetic field with respect to the line of sight and thus both the strength and orientation of the magnetic field can be constrained from their observation (e.g. Heiles & Troland 2005; Heiles & Robishaw 2009).

The free thermal electrons (as opposed to the high energy cosmic ray electrons) in the interstellar medium scatter off each other and off ions. When accelerated in such a way,

they emit bremsstrahlung, in this context known as free-free emission. The intensity of this radiation is sensitive to the emission measure, Eq. (1.3), and can in principle be used to constrain the distribution of these electrons. However, its direct observation is complicated by the fact that the total radiation is dominated by other Galactic emission mechanisms at all frequencies except for a narrow range at microwave frequencies (see e.g. Smoot 1998).

A more useful effect for constraining the distribution of thermal electrons is the dispersion that they cause on electromagnetic waves. The oscillating electric field of radiation passing through the interstellar plasma induces oscillations in the free electrons. The eigenfrequency of the electronic oscillations is given by the plasma frequency

$$\nu_{\mathrm{p}} = \sqrt{\frac{n_{\mathrm{e}} e^2}{\pi m_{\mathrm{e}}}}, \tag{1.4}$$

where e is the elementary charge and m_{e} is the electron mass (see e.g. Rybicki & Lightman 2008). This leads to a dielectric constant and therefore a refractive index which are frequency dependent. Therefore, radiation with a higher frequency travels faster through the plasma. This effect is detectable in the observations of pulsars.

Pulsars are rotating neutron stars that emit periodic pulses of radiation with a continuous frequency spectrum. Due to the dispersion within the interstellar plasma, the shorter wavelengths arrive at Earth earlier than the longer wavelengths. In the limit of high frequencies compared to the plasma frequency, $\nu \gg \nu_p$, which is satisfied for observable radiation in interstellar space, this time delay per frequency becomes

$$\frac{\mathrm{d}t}{\mathrm{d}\nu} = -\frac{e^2}{\pi c m_{\mathrm{e}} \nu^3} D_{\mathrm{m}}, \tag{1.5}$$

where

$$D_{\mathrm{m}} = \int_0^L \mathrm{d}l \; n_{\mathrm{e}} \tag{1.6}$$

is the dispersion measure (see e.g. Rybicki & Lightman 2008). Thus, if the electron density n_{e} is assumed to be known, the dispersion of the pulses can be used to determine the distance L of the pulsar (e.g. Lorimer 2008). Vice versa, if the distance to the pulsar is known, the dispersion measure constrains the distribution of free thermal electrons along the line of sight (e.g. Cordes & Lazio (2002), see also Sec. 1.1.3).

Another effect that can be used to investigate the properties of the interstellar plasma is the scattering of radiation on density fluctuations, known as interstellar scintillation. Armstrong et al. (1995) have used this effect to constrain the statistical properties of the fluctuations in the thermal electron density (see also Johnston et al. 1998; Bhat & Gupta 2002).

Finally, the dusty component of the interstellar medium leaves imprints on the observed radiation as well. Due to the multitude of different molecules and the wealth of their excitation modes, the physics involving dust grains is inarguably more complicated than what we have discussed so far. Therefore, we only attempt to give a rough qualitative overview over the involved processes.
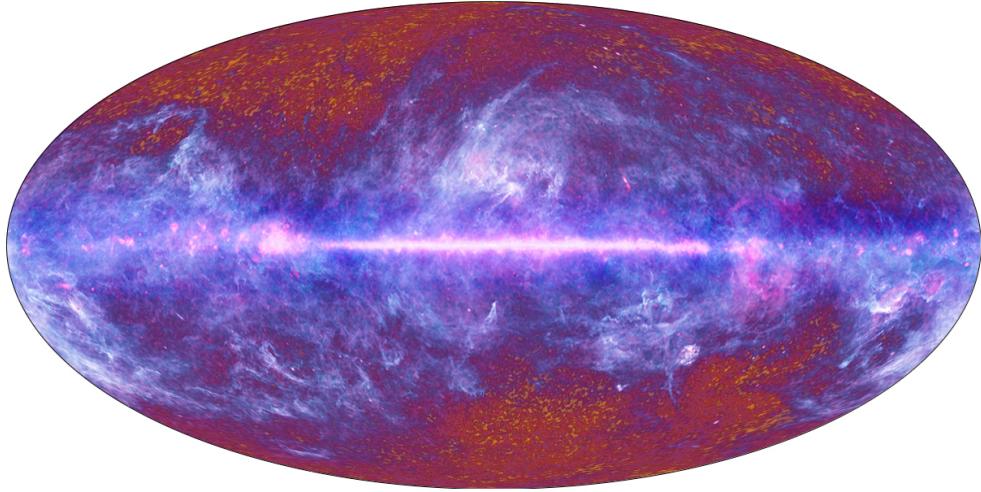
Figure 1.2: Composite image of the radiation detected by the *Planck* satellite. Observations in nine frequency channels from 30 GHz to 857 GHz are combined. The blue and pink colors correspond to thermal emission from dust grains. Image credit ESA / LFI & HFI Consortia.

Optical and near-optical light suffers from interstellar extinction, i.e. absorption and scattering through dust grains. Since this extinction is most efficient if the wavelength of the radiation is near the grain size and typical interstellar dust grains have diameters of at most a few hundred nm (Weingartner & Draine 2001), blue light is absorbed more efficiently. Thus, the presence of dust along a line of sight leads to the effect of interstellar reddening (Draine 2003), visible in Fig. 1.1. The extinction can be used to constrain the composition and three-dimensional distribution of dust in the Galaxy (e.g. Marshall et al. 2006).

The scattered radiation is reemitted at lower frequencies and can be directly observed at infrared wavelengths. Thus, sky maps of thermal dust radiation have been created e.g. by Schlegel et al. (1998). These maps are commonly used to predict the thermal dust emission at microwave frequencies (e.g. Finkbeiner et al. 1999). Figure 1.2 shows such a direct observation by the *Planck* satellite.

One important aspect of non-spherical dust-grains is their tendency to be aligned with the Galactic magnetic field, with their longest axis perpendicular to the field lines. While the detailed processes leading to this alignment are not known with certainty, it seems that particle collisions and radiative torques could be the main contributors (Greaves et al. 1999; Lazarian & Hoang 2009). As a consequence of the grain alignment, interstellar dust extinction is more efficient for radiation polarized orthogonal to the magnetic field, leading to an effective polarization of starlight parallel to the intervening magnetic field (e.g. Heiles 2000). On the other hand, the far-infrared radiation emitted by the dust grains has a net polarization parallel to the long axis of the grains and therefore orthogonal to the magnetic field (e.g. Ponthieu et al. 2005; Draine & Fraisse 2009, 2012).

Another observable consequence of non-spherical dust grains is the recently discovered so-called anomalous microwave emission (Kogut 1999), an emission component in the microwave band that is spatially correlated with thermal dust emission at far-infrared frequencies (Kogut et al. 1996). This component has since been identified to result from dust grains that are spinning or precessing around magnetic field lines (Draine & Lazarian 1998a,b; Hoang & Lazarian 2012).

The constituents of the interstellar medium that are most important for the present study are free thermal electrons, magnetic fields, and cosmic rays. In the following sections, we will briefly describe the processes leading to the observational consequences that we will use in later chapters, namely Faraday rotation and synchrotron emission, and summarize the most important aspects of the current understanding of these components.

Faraday rotation

In Chapters 2 and 4, we use the effect of Faraday rotation to study the magnetic field and density fluctuations in the interstellar medium. We dedicate this section to the description of the underlying physical processes, following Rybicki & Lightman (2008).

The relevant setting is that of an interstellar plasma, containing mainly protons and thermal electrons with density n_e, permeated by a magnetic field described by a field vector $\vec{B}$. The presence of the magnetic field breaks the isotropy of the plasma and introduces the cyclotron frequency,

$$\nu_c = \frac{eB}{2\pi m_e c},\tag{1.7}$$

as the frequency with which the electrons gyrate about the magnetic field lines. Here, c is the speed of light in vacuum.

We will study the propagation of electromagnetic radiation along the magnetic field line under the assumption that the strength of the oscillating magnetic field due to the propagating wave – and the resulting Lorentz force on free electrons – can be neglected with respect to the strength of the fixed magnetic field $\vec{B}$. Consider first a circularly polarized electromagnetic wave, propagating along the z-axis defined by the magnetic field direction, whose electric field vector is described by

$$\vec{E}_{r/l}(t) = E e^{2\pi i \nu t} \left(\hat{e}_x \mp i \hat{e}_y \right),\tag{1.8}$$

where the indices r and l stand for right and left-handed polarization and correspond to the minus and plus sign on the right hand side, respectively. Here, $\hat{e}_x$, $\hat{e}_y$, and $\hat{e}_z$ are the basis vectors of a cartesian coordinate system. Qualitatively speaking, right- and left-handed circularly polarized radiation will induce gyration movements in the electron field. Due to the presence of the magnetic field, the favored gyration frequency is shifted from zero in one direction, depending on the direction of the magnetic field, thus leading to different refractive indices for right- and left-handed polarization.

More quantitatively, the electric field will induce motions in the free electrons (the

protons can be neglected due to their high mass) governed by

$$m_{\mathrm{e}} \frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = -e\vec{E} - \frac{e}{c}\vec{v} \times B\hat{e}_z, \tag{1.9}$$

where $\vec{v}$ is the velocity of an electron. Substituting the expression for the electric field into this equation gives the velocity field

$$\vec{v}_{\mathrm{r/l}} = -\frac{ie}{2\pi m_{\mathrm{e}}\left(\nu \pm \nu_{\mathrm{c}}\right)}\vec{E}. \tag{1.10}$$

Using the electric current density, $j = -n_{\mathrm{e}}ev$, one finds for the electric conductivity

$$\sigma_{\mathrm{r/l}} = \frac{j}{E} = \frac{ie^2 n_{\mathrm{e}}}{2\pi m_{\mathrm{e}}\left(\nu \pm \nu_{\mathrm{c}}\right)}. \tag{1.11}$$

This expression can be used to calculate the dielectric constant

$$\epsilon_{\mathrm{r/l}} = 1 - \frac{2\sigma_{\mathrm{r/l}}}{i\nu} = 1 - \frac{\nu_{\mathrm{p}}^2}{\nu\left(\nu \pm \nu_{\mathrm{c}}\right)}, \tag{1.12}$$

which is now dependent on polarization. Therefore, right- and left-handed circularly polarized waves travel at different speeds along the magnetic field in the interstellar plasma.

Now consider a linearly polarized wave, which can be regarded as a superposition of right- and left-handed circularly polarized waves. After having traveled a distance L through the magnetized plasma, the angle by which the electric field vector of each of these components has rotated is

$$\varphi_{\mathrm{r/l}} = \int_0^L \mathrm{d}l\ k_{\mathrm{r/l}} = \int_0^L \mathrm{d}l\ \frac{2\pi\nu}{c}\sqrt{\epsilon_{\mathrm{r/l}}}, \tag{1.13}$$

where we have made use of the wave number $k_{\mathrm{r/l}}$ of the two polarization states. Superposing these two waves after a distance L therefore leads again to a linearly polarized wave, but with a polarization plane that has been rotated with respect to the original one by an angle $\Delta_\varphi = \left(\varphi_{\mathrm{r}} - \varphi_{\mathrm{l}}\right)/2$. This rotation of the polarization plane is the effect known as Faraday rotation.

In the limit of $\nu \gg \nu_{\mathrm{p}}$ and $\nu \gg \nu_{\mathrm{c}}$, the integrand in the last equation can be expanded, yielding for the rotation of the polarization plane

$$\Delta_\varphi = \frac{\pi}{c}\int_0^L \frac{\nu_{\mathrm{c}}\nu_{\mathrm{p}}^2}{\nu^2} = \frac{e^3\lambda^2}{2\pi m_{\mathrm{e}}^2 c^4}\int_0^L n_{\mathrm{e}}\vec{B}\cdot\mathrm{d}\vec{l}. \tag{1.14}$$

In the last step, we have generalized the expression from a magnetic field oriented along the line of sight to a general magnetic field, where the scalar product ensures that only the line-of-sight component contributes. The orthogonal component of the magnetic field does in general influence the rotation angle as well, however, this contribution can be neglected in the limit that we consider here (see e.g. Schlickeiser 2002).

Taking out the wavelength-dependence, we define the Faraday depth corresponding to a physical distance L as

$$\phi = \frac{e^3}{2\pi m_e^2 c^4} \int_0^L n_e \vec{B} \cdot \mathrm{d}\vec{l}. \tag{1.15}$$

Note that there is no one-to-one correspondence between physical distances and Faraday depths. While each physical distance can be assigned an unambiguous Faraday depth via the last equation, the converse is not true. Due to reversals of the magnetic field, one Faraday depth will in general correspond to several physical distances.

The Faraday depth as defined in Eq. (1.15) is sometimes also referred to as rotation measure (RM) because of the most common means of its estimation. In this procedure, the position angle of the polarization plane is measured at several frequencies and the Faraday depth is estimated as the proportionality constant between the position angle and the squared wavelength. For clarity, we will reserve the term rotation measure for observational estimates of Faraday depths done in this way. Chapter 4 contains a more extensive discussion of this distinction.

The estimation of the Faraday depth of a source via a rotation measure fit has a couple of drawbacks. One problem is the inherent ambiguity of position angle measurements with respect to integer multiples of π. Especially if only a few wavelengths are used for the observation, there is a risk of mistaking a rotation by a small angle for a rotation by a little more than 180° and vice versa. Another problem is the difficulty with which multiple emission components along a line of sight are identified. If in reality instead of one source, several sources are observed, each located at a different Faraday depth, this will in general not be obvious. At best, the fit of ϕ versus λ^2 will appear very noisy, however, if one source is dominant, the fitting procedure can lead to a distorted estimate of this source's Faraday depth without this being apparent.

In the derivation of Eq. (1.15) we have assumed that the observed polarized radiation is generated at a distance L from the observer. However, in general there will be some emission all along the line of sight. In this case the polarization plane of the radiation is rotated by a different amount, depending on where it was emitted, and the simple relationship between rotation angle and wavelength breaks down and is replaced by a Fourier relationship,

$$P(\lambda^2) = \int_{-\infty}^{+\infty} \mathrm{d}\phi \, F(\phi) \mathrm{e}^{2i\phi\lambda^2}, \tag{1.16}$$

between the Faraday depth, ϕ, and the squared wavelength, λ^2 (Brentjens & de Bruyn 2005). Here, $F(\phi)$ is the complex polarized emissivity per Faraday depth and $P(\lambda^2)$ is the complex polarized intensity observed at a wavelength λ. The technique of RM synthesis (Burn 1966; Brentjens & de Bruyn 2005) builds on this Fourier relationship. By taking observations at many different frequencies and approximately inverting the above equation, several sources or even continuous emission along a line of sight can be dealt with. This technique is also less prone to the $n\pi$ ambiguity.

It is evident from Eq. (1.15), that observations of Faraday depths can be used to study the density of thermal electrons, n_e, or the Galactic magnetic field, $\vec{B}$. However, neither

of these quantities is known, therefore, directly inferring one of them from observations of Faraday depths is problematic. Sources of linearly polarized radiation that are used for these kinds of studies are mainly pulsars and radio galaxies. In both cases, the integral projection in Eq. (1.15) is a problem. Not only are the observations only sensitive to one component of the magnetic field, they also only yield a projection along the line of sight. The usage of pulsars has the benefit that by using different pulsars that are embedded in the Milky Way, one can in principle probe different distances. However, only very few pulsars have accurate distance information (e.g. Schnitzeler 2012). Extragalactic sources, on the other hand, always probe the full line of sight through the Galaxy. They have, however, the disadvantage of probing extragalactic environments as well. In Chapter 4, we present an attempt at statistically separating the extragalactic contributions from the Galactic ones.

Synchrotron emission

Synchrotron emission is another probe of the Galactic magnetic field, which we will use in Chapter 2. Here, we briefly discuss the underlying physical mechanism and observational consequences.

The ingredients needed for synchrotron radiation to occur are highly relativistic electrons and a magnetic field. Any electrons that have a velocity component orthogonal to the magnetic field will be accelerated by the Lorentz force, gyrate about the magnetic field lines, and emit radiation due to their accelerated motion.

Due to the high energy of the involved electrons, relativistic effects become important. The relativistic gyration frequency becomes

$$\nu_{\mathrm{s}} = \frac{eB}{2\pi\gamma m_{\mathrm{e}}c},$$
(1.17)

where $\gamma = \left(1 - \frac{v^2}{c^2}\right)^{-1/2}$ is the Lorentz factor for an electron moving with velocity v. The acceleration of the electron, perpendicular to its velocity, can be written as

$$a_{\perp} = 2\pi\nu_{\mathrm{s}}v_{\perp},$$
(1.18)

where $v_{\perp}$ is the component of the electron's velocity that is oriented perpendicular to the magnetic field, and the power emitted in form of radiation becomes

$$P = \frac{2e^2}{3c^3}\gamma^4 a_{\perp}^2 = \frac{2}{3}\frac{\gamma^2 e^4}{m_{\mathrm{e}}^2 c^5}B^2 v_{\perp}^2$$
(1.19)

(Rybicki & Lightman 2008). Note the dependence on the squared magnetic field strength. To get the average power from an electron moving with velocity v, one needs to average over all possible orientations of the velocity vector. Assuming an isotropic distribution for the velocities, this gives another factor 2/3, yielding as the radiated power

$$P = \left(\frac{2}{3}\right)^2 \frac{\gamma^2 e^4}{m_{\mathrm{e}}^2 c^5}B^2 v^2.$$
(1.20)

For non-relativistic electrons, the frequency of the radiation would simply equal the gyration frequency, however relativistic beaming effects lead to the radiation emitted during each rotation appearing as a pulse of much shorter duration than the gyration period. The superposition of these pulses leads to an effective spectrum that is much wider than for non-relativistic electrons. In addition, the gyrating electrons will not all have the same velocity v or energy E. The energy or velocity distribution of particles will influence the observed spectrum of the synchrotron radiation.

If the energy distribution of cosmic ray electrons can be approximated by a power law with index p, i.e.

$$\mathrm{d}n_{\mathrm{CRE}} \propto E^{-p}\,\mathrm{d}E, \tag{1.21}$$

the observed power of synchrotron radiation can be approximated as a power law as well,

$$P(\nu) \propto \nu^{-s}, \tag{1.22}$$

where the synchrotron spectral index is given by $s = (p-1)/2$, as shown e.g. by Rybicki & Lightman (2008).

Calculating the polarization properties of this radiation, one arrives at a net linear polarization perpendicular to the magnetic field projected onto the plane of the sky. The maximum obtainable degree of polarization is 75% (Ginzburg & Syrovatskii 1965). However, in observations of Galactic synchrotron emission, the polarization fractions are typically much lower (e.g. Burigana 2006), due to several depolarization effects. The change of magnetic field orientation within the instrumental beam leads to a partial compensation of different polarization directions, known as beam depolarization. The presence of magnetic fields along the line of sight leads to Faraday rotation of the emitted synchrotron radiation. Radiation at different wavelengths gets rotated by different amounts and therefore the total polarization within a finite instrumental frequency window gets diminished. This effect is called Faraday depolarization and is typically only important at frequencies $\nu \lesssim 1\,\mathrm{GHz}$ (Stroman & Pohl 2009). Finally, any observed synchrotron radiation will be a superposition of radiation emitted at different distances along the line of sight. Since the orientation of the magnetic field can change with distance, so can the polarization direction. The superposition therefore has a reduced polarization fraction.

In conclusion, although synchrotron radiation can in principle be used to probe both the strength of the Galactic magnetic field (see Eq. (1.19)) and its orientation, it suffers from similar complications as Faraday rotation observations. While the latter are weighted with the thermal electron density, the former is sensitive to the density of cosmic ray electrons and even to their energetic composition, and both effects suffer from a line of sight projection. Nevertheless, both effects are being used to constrain the configuration of the Galactic magnetic field (see Sec. 1.1.3).

1.1.3 Components related to the Galactic magnetic field

In this section, we describe the current understanding of the components of the interstellar medium that will play a role in the following chapters. These are the components that

are observationally most entangled with the Galactic magnetic field, namely the free thermal electrons influencing Faraday rotation observations, the cosmic rays influencing the emission of synchrotron radiation, and the magnetic field itself.

Thermal electrons

The term 'thermal electrons' is used here to describe the free electrons that are part of the thermally ionized interstellar plasma, as opposed to the relativistic cosmic ray electrons. Sometimes the term warm ionized medium is also used. In the context of Faraday rotation, the electrons of this phase act as the resonators in the interstellar plasma and are therefore important for magnetic field studies.

The spatial distribution of thermal electrons is governed by the large-scale disk and spiral structure of the Galaxy, compression in shock fronts e.g. from supernova explosions, and turbulent motions. On all levels, the motion of the thermal electrons is also influenced by the presence of magnetic fields and vice versa.

The first non-trivial model for the three-dimensional distribution of the thermal electrons in the Milky Way was deduced by Taylor & Cordes (1993) who for the first time included information on the location of the spiral arms deduced from radio observations of ionized hydrogen. They constrained their model using pulsar dispersion measure data (see Eq. (1.6)) and observations of interstellar scintillation.

The most commonly used model for the electron distribution to date was devised by Cordes & Lazio (2002) and can be seen as an extended and updated version of the model by Taylor & Cordes (1993). The electron distribution within the Galactic plane in this model, known as NE2001, is shown in Fig. 1.3. This model benefitted from a larger number of available observations of D_{m}. Furthermore, some knowledge about the structure of the local interstellar medium, i.e. the solar neighborhood, was explicitly included in the model. Cordes & Lazio (2002) also included pulsar scintillation observations as well as accurate measurements of H_α radiation as a proxy for emission measure, to determine the large-scale variations in the fluctuation strength of the free electron density. Altogether, the NE2001 model is made up of a thin Galactic disk, a thick Galactic disk, which is less concentrated both in the radial and in the vertical direction, the spiral arm structure, an extra component around the Galactic center, and the local structure in the solar neighborhood.

Gaensler et al. (2008) argued using observations of pulsar dispersion measures and diffuse H_α radiation that the scale height of the thick disk component is in reality about twice as large as advocated by Cordes & Lazio (2002). However, Schnitzeler (2012) recently argued that this extended thick disk can be ruled out using pulsar dispersion measure observations.

One important point to note is that most of the models rely heavily on pulsar dispersion measurements, which need to be complemented with a distance estimate for the pulsars to yield information on the three-dimensional electron distribution. However, many of the estimated pulsar distances used e.g. by Cordes & Lazio (2002) are themselves derived from dispersion measure observations under the assumption of a certain electron distribution. This circular argumentation is not always solved self-consistently.
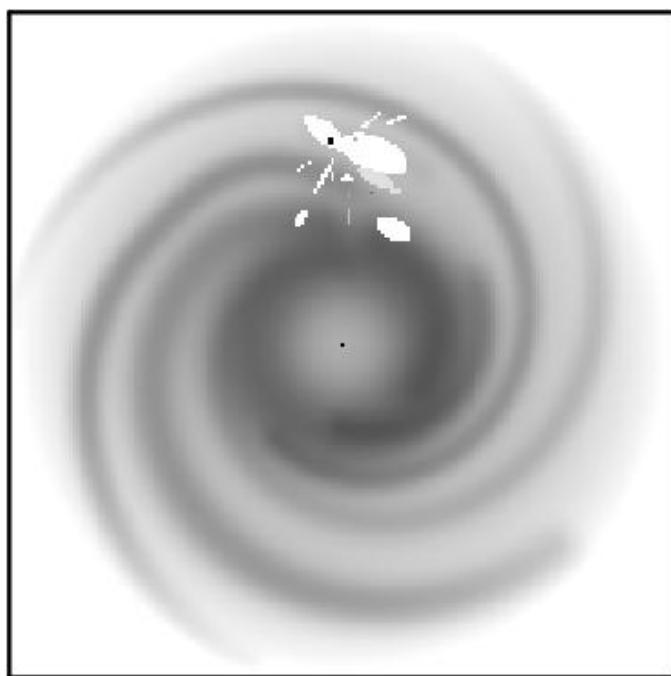
Figure 1.3: Cut through the Galactic plane of the NE2001 model for the thermal electron density. Dark colors correspond to high densities, light colors to low ones. The size of the depicted area is $30 \, \text{kpc} \times 30 \, \text{kpc}$. The Galactic center is in the middle of the plot, the sun is located in the horizontal center next to the small dark spot (the gum nebula) in the upper half. Figure from Cordes & Lazio (2002).

Schnitzeler (2012) confronts several models of differing complexity for the Galactic electron distribution with measurements of dispersion measures for pulsars for which independent and accurate distance estimates exist. He shows that more complex models including spiral arms are not necessarily better supported by the currently available data than very simplistic axisymmetric models. In conclusion, while a rough consensus exists in the literature regarding the broad geometry of the thermal electron distribution in the Milky Way, observational data are at present not suited to constrain any detailed features.

Cosmic rays

The interstellar medium is permeated by a fluid of highly energetic, charged particles, mainly protons, around ten percent α-particles, even fewer more massive nuclei, and relativistic electrons, collectively known as cosmic rays (e.g. Schlickeiser 2002). The main importance of this component of the interstellar medium will be the synchrotron radiation originating from the cosmic ray electrons gyrating around magnetic field lines. We will use the polarization properties of this radiation to probe the Galactic magnetic field in Chapter 2.

The origin of these particles and the means of their acceleration are a topic of active research. One likely source of particles are supernova explosions in which these particles can be created and ejected and undergo ensuing shock acceleration. A similar acceleration mechanism can happen in stellar winds. Additionally, the most energetic particles are likely to be of extragalactic origin and might be accelerated in active galactic nuclei (see Schlickeiser 2002, for details). For observations on Earth, solar wind particles are a significant contaminant. However, these are unimportant when studying effects on Galactic scales, such as synchrotron emission.

The evolution of the cosmic ray fluid in the interstellar medium is a complicated process. Effects like diffusion, convection, turbulence, confinement of the charged particles by the magnetic field, reacceleration, interactions with other cosmic ray particles and background photons all play a role (e.g. Strong & Moskalenko 2001; Strong et al. 2007; Yan & Lazarian 2008). Due to scattering of the particles by plasma waves in the interstellar medium, their resulting velocity distribution is isotropic (Schlickeiser 2002).

Phenomenological models of the distribution of the cosmic ray electrons have been created e.g. by Page et al. (2007) and Sun et al. (2008) using observations of synchrotron radiation. However, both these studies simultaneously fit a model for the Galactic magnetic field and therefore restrict themselves to rather simplistic forms of the cosmic ray electron distribution, consisting essentially of a disk with two scale heights for the radial and vertical directions.

A recent direct measurement of electron energies using the *Fermi* satellite by Ackermann et al. (2010) showed that their energy spectrum is consistent with a power law, as given by Eq. (1.21), with a spectral index of $p = 3.08 \pm 0.05$ for energies between $7\,\mathrm{GeV}$ and $1\,\mathrm{TeV}$. This, however, of course only represents a local measurement and does not reveal anything about spatial variations of the energy distribution. Spatial variations in the spectral index of the observed synchrotron radiation have indeed been detected e.g. by Bennett et al.

(2003), indicating that the electron distribution is likely to depend on the environment.

So altogether, relatively little is known from observations about the spatial variations of the cosmic ray electron distribution both in energy and in number density.

Magnetic fields

The Galactic magnetic field is traceable with a variety of observable effects, among them emission from Zeeman lines, synchrotron emission, Faraday rotation, polarization of thermal dust emission and partially absorbed starlight. From observations of Zeeman splitting, Heiles (1990) deduced a typical magnetic field strength in the disk of the Milky Way of a few µG.

A rough picture of the large-scale coherent structure of the Galactic magnetic field can be obtained from observations of other spiral galaxies. It is in many cases easier to infer the magnetic field structure of other galaxies from an outside perspective than it is to disentangle the observations of the Milky Way made from within. Typically, spiral galaxies exhibit a magnetic field that is aligned with the disk and roughly follows the spiral pattern (e.g. Beck & Hoernes 1996; Fletcher et al. 2011). When seen edge-on, they typically show an X-shaped magnetic field (Krause 2008). The natural assumption, which is also supported by observational data, is that the Milky Way's magnetic field has a similar structure.

The existence of coherent magnetic fields raises the question of their origin. The most promising concept that has been developed to answer this question is dynamo theory, in which mechanical energy is extracted from turbulent gas motions and an initially weak magnetic seed field is amplified and smoothed (e.g. Brandenburg & Subramanian 2005). Magnetic helicity, which we study in Chapter 2 is a quantity that is conserved to high precision during these processes. However, helicity of opposite sign is generically produced on large and small scales (e.g. Subramanian 2002). Observationally detecting this helicity would be a strong point in favor of turbulent dynamo theory as an explanantion for galactic magnetic fields. As an explanation for the existence of the seed fields, several ideas are being discussed, e.g. the Biermann battery mechanism (e.g. Widrow 2002) and the existence of primordial magnetic fields (e.g. Gasperini et al. 1995).

The Galactic magnetic field can be regarded as the sum of three qualitatively different components (Jaffe et al. 2010). The first is a large-scale component, which is coherent over kpc length-scales. This component is likely to be the product of a dynamo acting on Galactic scales. The second component is an isotropic random field, sometimes also referred to as turbulent field. Strength, orientation, and sign of the magnetic field of this component change randomly from position to position. This field is thought to result directly from turbulent motions in the magnetized interstellar plasma, as well as from supernova outflows. The third component is referred to as ordered field, anisotropic random field, or striated random field. Its field vectors share the same orientation, but their sign and direction varies randomly. This last field configuration arises for example when a medium containing an isotropic random field is strongly compressed along one or two axes. Since the same plasma deformations also shape the coherent component, the orientations

of the two field components are commonly assumed to be the same (e.g. Jansson & Farrar 2012a).

These three components can be distinguished by their impact on observations of Faraday rotation and synchrotron radiation. On average, non-zero Faraday rotation is only produced by the coherent component when observed along the orientation of the magnetic field. Spatially strongly varying Faraday rotation is also produced by the isotropic random component and by the anisotropic random component when observed along the direction of this field. The coherent and anisotropic random components both produce polarized synchrotron radiation, but only in directions orthogonal to the field lines. The isotropic random component, on the other hand, produces synchrotrom emission in all directions, however, its polarization cancels on average.

Simple parametric models of the coherent magnetic field exhibiting a spiral structure in the Galactic plane were devised e.g. by Page et al. (2007) and Sun et al. (2008). While Page et al. (2007) used only synchrotron data to constrain the parameters of their model, Sun et al. (2008) added observations of Faraday rotation. Jansson et al. (2009) used updated synchrotron and Faraday rotation data to optimize several parametric models of the coherent field component and found that most best-fit models had changed significantly due to the new data. Detailed studies of the coherent magnetic field within the Galactic disk were performed e.g. by Brown et al. (2003b); Han et al. (2006); Van Eck et al. (2011), especially with a view on the number and locations of possible field reversals. The conclusions range from a total absence of field reversals to a field reversal between each spiral arm and inter-arm region. Haverkorn et al. (2006a) studied the fluctuations of the isotropic random component in arm and inter-arm regions using Faraday rotation observations. They found that the fluctuations happen on smaller spatial scales inside the spiral arms. Combined models of the coherent and isotropic random components of the Galactic magnetic field were created by Fauvet et al. (2011, 2012), who used observations of synchrotron and thermal dust emission to constrain their model parameters. In these models, however, the only fit parameter describing the isotropic random components is the relative amplitude of this component with respect to the coherent one, assumed to be the same throughout the Galaxy. Jaffe et al. (2010) for the first time modeled all three components to fit synchrotron and Faraday rotation observations, albeit only within the Galactic plane.

Recently, Jansson & Farrar (2012a,b) finally created a full three-dimensional model of the Galactic magnetic field allowing for all three components, based on up-to-date synchrotron and Faraday rotation data. Their coherent field model itself consists of three components. The first is a disk field, which is purely azimuthal in the inner Galaxy and is made up of eight magnetic spiral arms in the outer Galaxy. The other two coherent components are a purely toroidal and a purely poloidal component that together make up the magnetic field in the Galactic halo. Additionally, an anisotropic random field component is assumed, which is aligned with the coherent field and whose relative amplitude is the same everywhere. Taking into account all data, they fit a total of 35 parameters describing the random and coherent fields and find that the anisotropic random component is about 35% stronger than the coherent component. Fig. 1.4 shows a comparison of the strengths
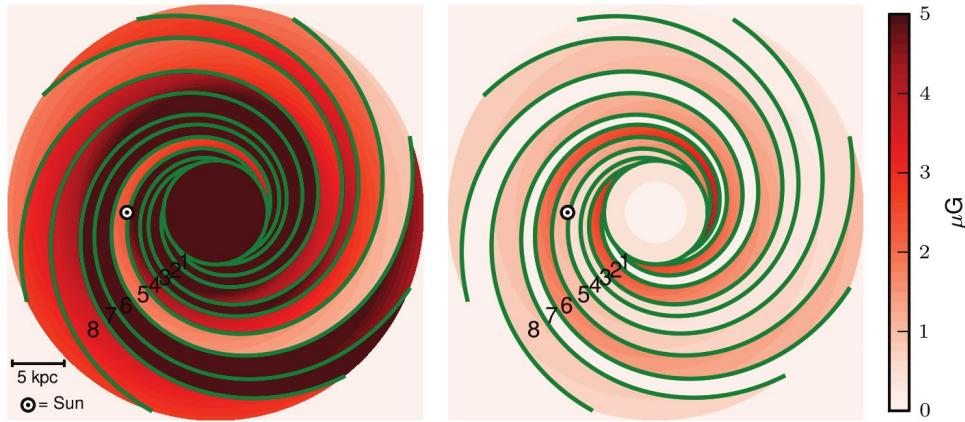
Figure 1.4: Comparison of coherent and random magnetic field strengths in the Galactic plane according to Jansson & Farrar (2012b). The left panel shows the strength of the random field, the right panel the strength of the disk component of the coherent field. The direction of the coherent field is clockwise (in this top-down view) in rings 3-6 and counterclockwise in the others. Figure from Jansson & Farrar (2012b)

of the coherent and random fields in the disk. The coherent magnetic field model also has the desirable property of being divergence-free and reproducing the characteristic X-shape when seen edge-on.

All these modeling attempts, however, suffer from the complexity of having to deal with line of sight projections as observables and the degeneracies between magnetic features and features in the electron densities (cosmic rays or thermal electrons) that this causes. The evidence for certain features in the coherent component of the Galactic magnetic field changes almost on a yearly basis as new data become available and modeling techniques are improved. The random components can only be described statistically, so a detailed prediction of the magnetic field strength and direction for each location in the Milky Way is at the moment out of the question. More detailed three-dimensional reconstruction attempts taking into account all available observables and optimizing the thermal and cosmic ray electron densities at the same time need to be undertaken in the future.

1.2   Signal inference

In the previous section we have already seen that deducing information about physical quantities from observations is not always a trivial task. In this section, we will elaborate a bit more on the challenges and the methodology that can be used to overcome these challenges. This will serve as background for the statistical techniques developed in Chapters 3 and 5. Throughout this work, we generally follow the notational conventions

of *information field theory* (Enßlin et al. 2009).

1.2.1 General framework

In physics and astronomy, we do experiments because we would like to learn something about a physical quantity. It is, however, not always possible to measure the properties of this quantity directly, as we have seen for example for the distribution of the different components of the interstellar medium. The outcome of the measurement is therefore related in a not necessarily trivial way to the quantity of interest. We call this quantity the signal s and the observation results the data d and write

$$d = g(s, \dots), \tag{1.23}$$

where the function g encodes the way in which our observation is sensitive to the underlying signal.

This function can in general depend on other unknown quantities that we might not be interested in. One such quantity that appears in any experimental setting is measurement noise. If we assume that the noise n is additive in our measurement, we can separate it from the function g and write

$$d = f(s) + n. \tag{1.24}$$

Assuming that the fuction f is fully known, we are left with two problems. First, we have to get rid of the noise contribution to the data, which is difficult because it is unknown to us. Second, we have to invert the function f to get an estimate for the quantity s that we are interested in. However, this may be impossible since the value of $f(s)$ may well be the same for different values of s, i.e. the observation may be degenerate with respect to different signal configurations. These are the two main challenges that need to be overcome when doing inference on the signal.

So far, we have not specified what kinds of mathematical objects s, n, and d are. They each could be a single number or a vector of numbers. In the cases that we are interested in here, the signal s will even be a field, i.e. a continuous function on some manifold $\mathcal{M}$,

$$s: \ \mathcal{M} \longrightarrow \mathbb{C}. \tag{1.25}$$

To deal with the statistical inference of continuous fields with infinitely many degrees of freedom, the framework of *information field theory* was developed by Enßlin et al. (2009) (see also Lemm 2003). Much of Chapters 3 and 5 builds on this framework.

In our applications, the manifold $\mathcal{M}$ will either be an interval, a flat two-dimensional rectangle, or the two-sphere $\mathcal{S}^2$. The last case has special relevance in astronomy since the sky can be regarded as a sphere, so that one naturally uses $\mathcal{S}^2$ as signal domain if one attempts to reconstruct the dependence of some quantity on the direction in the sky, but not the dependence on the distance from Earth.

The data, on the other hand, are always finite, since the outcome of any measurement can be summarized in finitely many numbers. Therefore, the function f in Eq. (1.24) can never be invertible when the signal is a continuous field.

A special case of Eq. (1.24) that we will make heavy use of is the case of a linear mapping f, which can be written as an operator applied to the signal,

$$d = Rs + n. \tag{1.26}$$

The operator R is known as response operator. For signals with a finite number of degrees of freedom, it becomes a matrix, which is not necessarily square. For continuous signals, we can write the expression for the i-th data point as an integral over the manifold,

$$d_i = \int_{\mathcal{M}} \mathrm{d}x \ R_i(x) \, s(x) + n_i. \tag{1.27}$$

This specialization is still sufficiently general to accomodate a large variety of observational scenarios. The response operator can for example incorporate line of sight integrations as often needed in Galactic astrophysics, convolutions with a beam function, projections onto specific regions or positions in which the signal was observed, Fourier transformations that are part of interferometric observations, etc.

When trying to answer questions about the signal, we need to know which signal configurations are likely and which are not. In other words, we need to assign probabilities. In Bayesian probability theory one distinguishes between the prior probability distribution, $\mathcal{P}(s)$, which describes the probability of different signal configurations based on any theoretical or empirical knowledge that one might have before conducting the measurement, and the posterior probability distribution, $\mathcal{P}(s|d)$, which is the same, only after the data have been observed. The prior and posterior distribution describe our knowledge on the signal before and after the experiment. These two are connected by the data likelihood $\mathcal{P}(d|s)$, which encodes how likely the observed data are, given a particular signal configuration, via Bayes' theorem,

$$\mathcal{P}(s|d) = \frac{\mathcal{P}(d|s) \, \mathcal{P}(s)}{\mathcal{P}(d)}. \tag{1.28}$$

The denominator in this equation is the so-called evidence. Since it is not signal dependent, it is unimportant when making inferences about s, but can in principle be calculated via

$$\mathcal{P}(d) = \int \mathcal{D}s \ \mathcal{P}(d|s) \, \mathcal{P}(s). \tag{1.29}$$

Here and in the following, $\int \mathcal{D}s$ denotes an integral over all possible signal configurations, i.e. all possible signal values for each location on $\mathcal{M}$. The shape of the likelihood function $\mathcal{P}(d|s)$ follows from Eq. (1.24) or (1.26) in combination with the noise statistics. A first example for this will be given in the next section.

In information field theory, the right hand side of Bayes' theorem is rewritten as

$$\frac{\mathcal{P}(d, s)}{\mathcal{P}(d)} = \frac{1}{Z_d} \mathrm{e}^{-H(d,s)}, \tag{1.30}$$

to highlight the identity of statistics and thermodynamics. The partition function $Z_d = \mathcal{P}(d)$ again only depends on the data and the Hamiltonian

$$H(d, s) = -\ln\left(\mathcal{P}(d, s)\right) \tag{1.31}$$

captures all the important information and is often easier to handle mathematically.

Once the posterior probability distribution is derived, the question arises which signal configuration is the right answer to the inference problem. In principle, there is no single right answer; the posterior distribution is all that we can know about the signal. However, it is nevertheless useful to condense the information into a single result, ideally with an uncertainty estimate. For this, different choices are possible.

One possibility is to look for the signal configuration with the highest probability by simply maximizing the posterior distribution – or equivalently minimizing the Hamiltonian – with respect to s,

$$\left.\frac{\delta H(d, s)}{\delta s}\right|_{s = m^{\text{(MAP)}}} = 0. \tag{1.32}$$

Here, δ denotes a variational derivative. This gives the maximum a posteriori solution $m^{\text{(MAP)}}$. An uncertainty estimate for this solution can be calculated by expanding the Hamiltonian up to second order around its minimum and identifying the second order term with the inverse uncertainty covariance,

$$\left(D^{\text{(MAP)}}\right)^{-1} := \left.\frac{\delta^2 H(d, s)}{\delta s^2}\right|_{s = m^{\text{(MAP)}}}. \tag{1.33}$$

This corresponds to replacing the posterior distribution with a Gaussian distribution that passes through its maximum and has $D^{\text{(MAP)}}$ as covariance matrix. The uncertainty covariance includes information on the typical size of the uncertainty at each location of $\mathcal{M}$ and the correlation of variations at different locations. The benefit of the maximum a posteriori solution is that it is usually comparatively easy to calculate. The drawback is that it takes into account only local information at the maximum of the posterior distribution and the presence of any heavy tails or further local maxima goes unnoticed.

A choice that is influenced by the whole posterior distribution is the posterior mean,

$$m := \langle s \rangle_{\mathcal{P}(s|d)} = \int \mathcal{D}s \, s \, \mathcal{P}(s|d), \tag{1.34}$$

which is an average over all possible signal configurations, each weighted with its posterior probability. Here, we have also introduced an angular bracket notation for expectation values with respect to probability distributions, which we will use throughout this thesis. An uncertainty estimate for the posterior mean is straightforwardly defined as the second connected moment of the posterior distribution,

$$D := \left\langle \left(s - \langle s \rangle_{\mathcal{P}(s|d)}\right)\left(s - \langle s \rangle_{\mathcal{P}(s|d)}\right)^\dagger \right\rangle_{\mathcal{P}(s|d)}. \tag{1.35}$$

The $\dagger$ symbol denotes a transposed and complex conjugated quantity[1]. While the posterior mean of course also does not capture any distinct features of the posterior distribution, it is a more comprehensive summary than the maximum a posteriori. It can be shown (Weig & Enßlin 2010) that the posterior mean minimizes the expectation for the L^2-norm of the error,

$$m = \underset{\tilde{m}}{\mathrm{argmin}}\Big(\langle ||s - \tilde{m}||_{L^2}\rangle_{\mathcal{P}(s|d)}\Big). \tag{1.36}$$

We will use the posterior mean whenever feasible.

1.2.2 The Gaussian case

In this section, we demonstrate how to calculate the posterior distribution for a Gaussian signal under the assumption of Gaussian noise and discuss the result.

Gaussianity for the signal means that we assume a Gaussian prior probability distribution, given by

$$\mathcal{P}(s) = \mathcal{G}(s, S) := \frac{1}{|2\pi S|^{1/2}}\mathrm{e}^{-\frac{1}{2}s^\dagger S^{-1}s}. \tag{1.37}$$

Here,

$$S = \left\langle ss^\dagger\right\rangle_{\mathcal{P}(s)} \tag{1.38}$$

is the signal covariance describing the size of the typical fluctuations in the signal and the correlation structure and the vertical lines denote a determinant. If the signal is just a single number, the covariance becomes the variance. The term in the exponential of the Gaussian can also be written explicitly as

$$s^\dagger S^{-1} s = \int_{\mathcal{M}} \mathrm{d}x \int_{\mathcal{M}} \mathrm{d}y \; \bar{s}(x)\, S^{-1}(x,y)\, s(y). \tag{1.39}$$

In Eq. (1.37) we have also assumed that the prior is centered on zero, i.e. $\langle s\rangle_{\mathcal{P}(s)} = 0$. Generalizing the calculation to a Gaussian prior with a different mean is straightforward. We also assume signal-independent Gaussian noise with zero mean, i.e.

$$\mathcal{P}(n) = \mathcal{G}(n, N), \tag{1.40}$$

where the exponentiated term now reads

$$n^\dagger N^{-1} n = \sum_i \sum_j \bar{n}_i \left(N^{-1}\right)_{ij} n_j, \tag{1.41}$$

since the noise, like the data, is always finite-dimensional.

---

[1] If one thinks of s as a finite-dimensional vector, it becomes clear that $ss^\dagger$ is a matrix and $s^\dagger s$ is a scalar.

The first step is to calculate the likelihood function, which under the assumption of Eq. (1.26) as data model, can be done easily by marginalizing over the noise,

$$\mathcal{P}(d|s) = \int \mathcal{D}n \, \mathcal{P}(d, n|s) = \int \mathcal{D}n \, \mathcal{P}(d|n, s) \, \mathcal{G}(n, N)$$

$$= \int \mathcal{D}n \, \delta(d - Rs - n) \, \mathcal{G}(n, N) = \mathcal{G}(d - Rs, N). \tag{1.42}$$

Here, the δ-distribution arises since the data are fully determined once the signal and noise are given. Next, we simply multiply the result with the signal prior to obtain the posterior

$$\mathcal{P}(s|d) \propto \mathcal{P}(d|s) \, \mathcal{P}(s) \propto \mathcal{G}(d - Rs, N) \, \mathcal{G}(s, S) \propto \mathcal{G}(s - m, D). \tag{1.43}$$

The last step involves a little algebra, after which one finds

$$D = \left(S^{-1} + R^\dagger N^{-1} R\right)^{-1} \tag{1.44}$$

for the posterior covariance and

$$m = DR^\dagger N^{-1} d \tag{1.45}$$

for the posterior mean. This last equation is known as the Wiener filter (e.g. Enßlin et al. 2009). Since the posterior is again Gaussian in this case, the posterior mean and maximum a posteriori solution agree and are both given by the Wiener filter.

The terms in the Wiener filter equation deserve a short discussion. Starting from the right side, the field $R^\dagger N^{-1} d = j$ is known as information source in the context of *information field theory*. It is a response-over-noise weighted representation of the data. The weighting scheme downweights data points that have been poorly observed and the ones for which the noise contribution is expected to be large. This field drives the reconstructed map m away from the zero mean of the prior.

Next is the application of the posterior covariance operator, D, called information propagator. The reconstructed map at one position, $m(x)$, is made up of the values of the information source at all positions, each weighted with the appropriate value of the information propagator,

$$m(x) = \int_{\mathcal{M}} \mathrm{d}y \, D(x, y) \, j(y). \tag{1.46}$$

Hence, the information is propagated from each position y to its final destination x.

The information propagator, given in Eq. (1.44), consists of two parts. One part is a projection of the inverse noise covariance into the space of possible signals by means of the response operator, $R^\dagger N^{-1} R$. Qualitatively, this operator distinguishes between regions of $\mathcal{M}$ that are well observed and those which are not. The other part of D is the inverse signal covariance, S^{-1}, which separates regions of high expected signal fluctuations from those with low variations. In the limiting case of a region with high signal fluctuations that are well observed, $S^{-1} \approx 0$ and $R^\dagger N^{-1} R \neq 0$, the information propagator becomes $D \approx \left(R^\dagger N^{-1} R\right)^{-1}$, thus effectively undoing the response-over-noise weighting of the data and applying something close to an inverse response operator to compensate for the response

operator in Eq. (1.26), without extrapolations. In the other extreme case, in which the observations are very poor, $R^\dagger N^{-1} R \approx 0$, the information propagator is the same as the signal covariance and extrapolates widely. In general, there will be poorly observed regions, for which information has to be propagated across longer distances, and well observed regions, for which little extrapolation is necessary. The information propagator automatically balances the information from all locations in the appropriate manner for each point of the reconstruction. Note, however, that the two parts of the propagator will in general have different eigenbases, so one always has to consider the whole of $\mathcal{M}$ and cannot consider points or regions individually. The Wiener filter theory forms the basis for the signal inference techniques that we develop in Chapters 3 and 5.

It should be clear by now, that the signal and noise covariances play an important role in the reconstruction of Gaussian signals. Their precise knowledge is important for an accurate result. However, in general, at least the signal covariance S is not known. To overcome this problem, several strategies have been developed (Wandelt et al. 2004; Jasche et al. 2010b; Enßlin & Weig 2010; Enßlin & Frommert 2011). We will briefly outline the derivation as done by Enßlin & Weig (2010), since Chapters 3 and 5 build on it directly.

The basic idea is to augment the Gaussian signal prior with a well-chosen prior for the signal covariance, $\mathcal{P}(S)$, and derive the posterior via marginalization over the covariance,

$$\mathcal{P}(s|d) \propto \mathcal{P}(d|s)\, \mathcal{P}(s) = \mathcal{P}(d|s) \int \mathcal{D}S\, \mathcal{P}(s|S)\, \mathcal{P}(S). \tag{1.47}$$

The resulting posterior, however, is highly non-Gaussian and calculating its mean is therefore not possible analytically. Enßlin & Weig (2010) therefore suggest to find an optimal Gaussian approximation to this posterior, described by its mean m and its covariance D. Fitting for m and D then gives the *critical filter* estimate for the signal and its uncertainty. The detailed assumptions regarding the structure of the signal covariance, the choice of priors, and the approximations made can be found e.g. in Chapter 3.

1.2.3 Statistical homogeneity and isotropy

One assumption that we will make in all later chapters is the assumption of homogeneous and isotropic prior statistics. Statistical homogeneity means that the prior is invariant under translations in $\mathcal{M}$, statistical isotropy means that it is invariant under rotations. In the case of $\mathcal{M} = \mathcal{S}^2$, which will be important in later chapters, translational invariance is not defined.

In the case of a zero-mean Gaussian prior, the statistics are fully described by the covariance. Consider therefore a signal on the two-sphere with a covariance given by

$$S(\hat{n}, \hat{n}') = \langle s(\hat{n})\bar{s}(\hat{n}')\rangle_{\mathcal{P}(s)}. \tag{1.48}$$

Here, $\hat{n}$ and $\hat{n}'$ are two unit vectors, denoting two directions or two positions on $\mathcal{S}^2$. Assuming statistical isotropy now means that the covariance cannot depend on the two directions individually, but can only depend on their angular separation, given by $\arccos(\hat{n} \cdot \hat{n}')$.

We can therefore write

$$S(\hat{n}, \hat{n}') = S(\hat{n} \cdot \hat{n}'). \tag{1.49}$$

Instead of as a function of positions on $\mathcal{S}^2$, we can also express the covariance in spherical harmonic components as

$$S_{(\ell m),(\ell' m')} = \langle s_{\ell m} \bar{s}_{\ell' m'} \rangle_{\mathcal{P}(s)}, \tag{1.50}$$

where

$$s_{\ell m} = \int_{\mathcal{S}^2} \mathrm{d}\Omega \ \bar{Y}_{\ell m}(\hat{n}) \, s(\hat{n}) \tag{1.51}$$

is a spherical harmonic component of the signal field. Using the orthogonality of the spherical harmonic functions,

$$\int_{\mathcal{S}^2} \mathrm{d}\Omega \ \bar{Y}_{\ell m}(\hat{n}) \, Y_{\ell' m'}(\hat{n}) = \delta_{\ell \ell'} \delta_{mm'}, \tag{1.52}$$

and the expansion of a Legendre polynomial in terms of spherical harmonics,

$$P_\ell(\hat{n} \cdot \hat{n}') = \frac{4\pi}{2\ell + 1} \sum_{m=-\ell}^{\ell} \bar{Y}_{\ell m}(\hat{n}) Y_{\ell m}(\hat{n}') \tag{1.53}$$

(e.g. Abramowitz & Stegun 1964), as well as an expansion of the signal covariance in terms of Legendre polynomials,

$$S(\hat{n} \cdot \hat{n}') = \sum_{\ell=0}^{\infty} S^{(\ell)} P_\ell(\hat{n} \cdot \hat{n}'), \tag{1.54}$$

we can calculate the covariance in spherical harmonic representation as

$$
\begin{aligned}
S_{(\ell m),(\ell' m')} = \langle s_{\ell m} \bar{s}_{\ell' m'} \rangle_{\mathcal{P}(s)} &= \int_{\mathcal{S}^2} \mathrm{d}\Omega \int_{\mathcal{S}^2} \mathrm{d}\Omega' \ \bar{Y}_{\ell m}(\hat{n}) \, Y_{\ell' m'}(\hat{n}') \, \langle s(\hat{n}) \bar{s}(\hat{n}') \rangle_{\mathcal{P}(s)} \\
&= \sum_{\ell''=0}^{\infty} \frac{4\pi}{2\ell + 1} S^{(\ell'')} \sum_{m''=-\ell''}^{\ell''} \int_{\mathcal{S}^2} \mathrm{d}\Omega \int_{\mathcal{S}^2} \mathrm{d}\Omega' \ \bar{Y}_{\ell m}(\hat{n}) Y_{\ell' m'}(\hat{n}') \bar{Y}_{\ell'' m''}(\hat{n}) Y_{\ell'' m''}(\hat{n}') \\
&= \sum_{\ell''=0}^{\infty} \frac{4\pi}{2\ell'' + 1} S^{(\ell'')} \sum_{m''=-\ell''}^{\ell''} \delta_{\ell \ell''} \delta_{mm''} \delta_{\ell' \ell''} \delta_{m'm''} \\
&= \delta_{\ell \ell'} \delta_{mm'} \frac{4\pi}{2\ell + 1} S^{(\ell)}. \tag{1.55}
\end{aligned}
$$

We thus see that as a consequence of the statistical isotropy the signal covariance becomes diagonal in the basis given by the spherical harmonics and the diagonal entries depend only on ℓ. The combination $\frac{4\pi}{2\ell+1} S^{(\ell)}$ is called the angular power spectrum of the signal and usually denoted as C_ℓ. A similar result can be derived for signals in Euclidean space, replacing the spherical harmonics with Fourier modes. In this case, statistical homogeneity leads to a diagonal covariance matrix in the Fourier basis with entries that depend on the
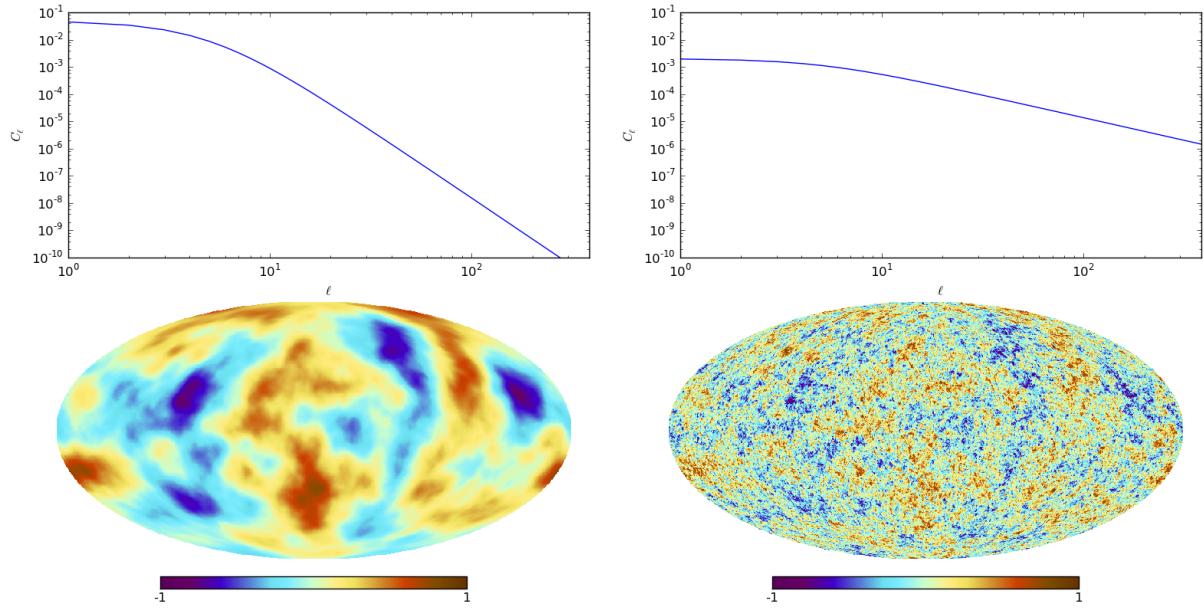
Figure 1.5: Comparison of Gaussian random fields with different power spectra. The top row shows the angular power spectra and the bottom row the corresponding signals.

wavevector $\vec{k}$; statistical isotropy then leads to diagonal entries that only depend on the absolute value of the wavevector.

The information contained in the power spectrum is the average strength of the signal fluctuations on different scales. For signals on the sphere, the scales of the fluctuations are given by ℓ (roughly as $\Delta_\theta = 180°/\ell$) and C_ℓ is the variance of the spherical harmonics components $s_{\ell m}$. To illustrate the importance of the power spectrum, we plot two versions of a randomly generated signal realization in Fig. 1.5. The spherical harmonics components are drawn randomly from Gaussian distributions with variances C_ℓ. For the two versions, the random seeds were chosen equal, only the values of C_ℓ were changed. As can be readily seen in the plots, higher C_ℓ-values for small ℓ correspond to more pronounced large-scale features in the signal and higher C_ℓ-values for large ℓ correspond to more pronounced small-scale features.

The assumption of statistical isotropy simplifies the inference formalism greatly, insofar as the covariance is now fully described by the power spectrum. It does, on the other hand, also restrict the discussion somewhat, since the generality is partly lost. One should be aware, however, that making this assumption for the prior statistics only amounts to assuming that there are no *a priori* special directions. If the observed data distinguish some directions from others, this will be taken into account in the posterior distribution, which will not be isotropic.

1.2.4   Gaussianity assumption

In all of the applications that we discuss in this thesis, we assume Gaussian statistics in one way or another. The reasons for this are manifold and shall be outlined here.

One obvious reason is that Gaussian probability distributions are easy to calculate with. Gaussian expectation values of powers of the signal can for example easily be calculated using Isserlis' theorem (often also called Wick theorem, Isserlis (1918)). The other, far better reason is that Gaussianity is often a good approximation. If the signal of interest arises for example as the sum of many independent components that can be assumed to approximatively follow the same statistics, the central limit theorem (e.g. Caticha 2008) tells us that the signal will be roughly Gaussian distributed. This applies for instance to many astrophysical signals defined on the sky that arise as projections along the line of sight through many independent regions of the universe or the Milky Way.

Another reason why Gaussianity is a justified assumption stems directly from probability theory. Assuming that a priori we know nothing about the signal of interest, a natural choice for the signal prior would be a flat prior, i.e. assuming equal probability for all signal realizations. If we now assume that we gain knowledge about the signal's covariance, the question arises how we should update our flat prior to incorporate this knowledge.

The most widely accepted recipe for updating probability distributions when knowledge about a certain expectation value is gained is the maximum entropy principle (e.g. Caticha & Preuss 2004). This principle is based on three axioms, namely that the update should be conducted in such a way that new information on a subset of all possible signal configurations only affects this subset, that it is independent of any coordinates, and that it does not matter whether two independent systems are updated together or individually. From these three axioms, a rule can be derived how to update probability distributions in such a way that they are only changed in ways that are absolutely necessary to accomodate the newly available information.

Applying this rule to the case where we start with a flat probability distribution and gain knowledge about the signal's covariance leads directly to a Gaussian distribution, as shown e.g. by Caticha & Preuss (2004). Thus one can say that the zero-mean Gaussian prior is the only prior that contains information about the signal's covariance and nothing else. Therefore, the Gaussian assumption is perfectly justified in cases in which only the covariance is a priori known.

More commonly, we do not know the covariance a priori but we do know that it is an important part of the problem. This is what Caticha (2012) calls "the important case". He argues that it should not make a difference for the probability distribution that best describes the system whether we happen to know certain things or not. He therefore concludes that whenever we know that a certain expectation value contains important information on the system under consideration, we should apply the maximum entropy principle as if this expectation value was known to us. Thus, since our discussion in the last two sections has shown that the signal covariance is indeed an important factor in our considerations, following this line of argumentation, we arrive again at the conclusion that we should choose a Gaussian signal prior.

One caveat, however is included. While the updating scheme is by design coordinate independent, our choice of a flat prior to start with is not. Since we are dealing with continuous signal values, we are always talking about probability densities. Therefore, it makes a difference whether we choose to pick a uniform prior for a certain signal or e.g. for its logarithm. We will explore this a bit more in the next section.

1.2.5 Beyond Gaussianity

There are certainly situations in which a Gaussian signal prior is a poor choice. Such a situation arises if we know for physical reasons that our signal has certain properties that are not well described by a Gaussian distribution. A common case for example is a signal that can only be positive. The tails of the Gaussian that extend infinitely far in both directions are certainly nonsensical in this case. Another case that arises often in astrophysics is a signal that has lots of small features and a few that are orders of magnitude stronger. Here, also, the Gaussian distribution reaches its limit.

In cases in which both positivity and the vast range of fluctuation strengths apply, a good solution is to simply consider the logarithm of the field of interest as the signal. If s is a Gaussian signal, e^s will be a field that exhibits these properties, known as a log-normal field. The application of the exponential function thereby has to be interpreted pointwise, i.e. individually for the values of s at each location on $\mathcal{M}$. If we then adapt Eq. (1.26) to read

$$d = Re^s + n, \tag{1.56}$$

the likelihood will change and the posterior $\mathcal{P}(s|d)$ will become highly non-Gaussian. We discuss this case in Chapter 5.

Thus, by choosing an appropriate transformation for the signal field, one can in some circumstances avoid the use of a non-Gaussian prior, albeit at the expense of having to deal with a complicated likelihood.

Whenever one uses non-Gaussian probability density distributions, be it due to the choice of a non-Gaussian prior or due to the presence of non-linearities in the system, such as in Eq. (1.56), one quickly gets to a point where arriving at exact results analytically becomes impossible. One possible solution to this conundrum is to try to solve the exact problem numerically. For this, a whole variety of sampling techniques has been developed (see e.g. Neal 1993; Robert & Casella 2004, for reviews of the field). The basic idea is to draw a large number of samples from a given probability distribution and to calculate expectation values by averaging over the samples. A different approach is to use approximations to reformulate the problem in a way that it becomes solveable analytically. Examples are the use of renormalization techniques (Enßlin & Frommert 2011) or the use of perturbative expansions as developed by Enßlin et al. (2009). Another example is the Gaussian approximation to the non-Gaussian posterior in the fashion introduced by Enßlin & Weig (2010) that we mentioned at the end of Sec. 1.2.2. This last technique is the method of choice for treating non-Gaussian probability distributions for the remainder of this work.

1.3   Outline

In the remaining chapters, we tackle a few of the open inference problems in the context of Galactic astrophysics. We focus on studies of the Galactic magnetic field and the related components of the interstellar medium.

In Chapter 2, we study the applicability of a technique developed recently by Junklewitz & Enßlin (2011) to detect magnetic helicity from observations of Faraday rotation and synchrotron emission. Using several sets of simulated observations in the Galactic setting, we test under which conditions the technique can be expected to yield reliable results. Of special importance are the assumptions that one makes about the distribution of free thermal electrons and cosmic ray electrons throughout the Milky Way. As we pointed out in Sec. 1.1.3, little is known about these distributions and studying the impact of false assumptions is therefore important. Thus, only conclusions regarding a combination of the electron densities and the magnetic field can be drawn.

To apply the technique to observational data, a continuous map of the Galactic Faraday depth is needed. As a first attempt, we use observational data from extragalactic radio sources along with the *critical filter* formalism of Enßlin & Weig (2010) to construct such a map. However, due to the use of extragalactic sources, the map created in this way contains not only the Galactic contribution to Faraday rotation but also unknown extragalactic contaminations. These should ideally have been assigned to the error budget of the data.

To overcome the problem of an unknown error budget, we develop the *extended critical filter* formalism in Chapter 3. The idea is to make use of the correlation structure of the signal, itself reconstructed from the entirety of the available data under the assumption of statistical isotropy, to separate fluctuations in the data that are consistent with this correlation structure from other ones that can then be assigned to the error budget. Thus, we are able to make use of data with an uncertain amount of uncertainty. We study the performance of this methodology in a variety of test cases with simulated signals and controlled noise characteristics.

In Chapter 4, we employ this newly developed algorithm to create an improved map of the Galactic Faraday depth. For this, we want to assign any extragalactic contributions to the error budget of the data points. However, the size of the extragalactic contributions is not known. At best, an expectation for the average extragalactic contribution can be deduced, but it is expected to vary vastly from one data point to the other. Additionally, there are several observational reasons why a few of the data points are likely to be faulty. For these as well, we should drastically increase the error budget, but we have no way of telling in advance which of the data points they are. Therefore, we are in a situation that is ideally suited for the *extended critical filter* formalism. Before using it to reconstruct a map, we assemble an extensive data catalog of Faraday rotation observations of extragalactic sources, consisting of newly provided data as well as data from the literature. We finally present our statistical results and discuss their possible astrophysical implications.

After this, we train our view on the reconstruction of Galactic emission maps. The emission intensity due to most physical processes discussed in Sec. 1.1.2 can be expected to be orders of magnitude higher within the Galactic disk than within the halo. Also, it

is always positive. Thus, we find ourselves in exactly the situation described in Sec. 1.2.5, in which a log-normal description is appropriate. Therefore we develop in Chapter 5 a technique to reconstruct log-normal signal fields from observational data in cases in which the signal covariance is unknown.

Another aspect introduced in that chapter is the notion of spectral smoothness. For statistically isotropic signals, the covariance is determined by the power spectrum. In many cases, this power spectrum is expected to be a smooth function of the scale-length. We study in Chapter 5 how to incorporate this expectation in a fully Bayesian way with the usage of an appropriate prior, thus alleviating the need for an ad-hoc treatment of spectral smoothness for the *critical* and *extended critical filter*. We again study the performance of the resulting algorithm in a series of simulated scenarios.

Finally, we summarize our conclusions in Chapter 6 and give a brief outlook on future work.

Chapter 2

Probing magnetic helicity with synchrotron radiation and Faraday rotation

Note: This chapter has been published in Astronomy & Astrophysics *(Oppermann et al. 2011a).*

2.1 Introduction

Helicity is of utmost interest in the study of astrophysical magnetism. Mean field theories for turbulent dynamos operating in the Galactic interstellar medium have been successful in explaining how the observed magnetic field strengths are maintained (e.g. Subramanian 2002). These theories predict that helicity is present on small scales in interstellar magnetic fields. Observationally detecting or excluding helicity in these fields would therefore either strongly suggest that these theories are valid or indicate that there are some flaws in them.

However, since helicity is a quantity that describes the three-dimensional structure of a magnetic field and most observation techniques produce at best two-dimensional images leading to an informational deficit, it has thus far largely eluded observers. Previous work on the detection of magnetic helicity in astrophysical contexts has focused mainly on either magnetic fields of specific objects, such as the Sun (see e.g. Zhang 2010, and references therein) or astrophysical jets (cf. e.g. Enßlin 2003; Gabuzda et al. 2004), or cosmological primordial magnetic fields (e.g. Kahniashvili & Ratra 2005; Kahniashvili et al. 2005). Two exceptions are the work by Volegova & Stepanov (2010), in which the use of Faraday rotation and synchrotron radiation for detecting magnetic helicity was suggested for the first time, and the work of Kahniashvili & Vachaspati (2006), in which the use of charged ultra high energy cosmic rays of known sources is suggested for probing the three-dimensional structure of magnetic fields through which they pass. However, the sources of ultra high energy cosmic rays are not known yet and the applicability of this test is therefore limited.

The *LITMUS* (**L**ocal **I**nference **T**est for **M**agnetic fields which **U**ncovers helice**S**) procedure for the detection of magnetic helicity suggested by Junklewitz & Enßlin (2011) probes the local current helicity density $\vec{B} \cdot \vec{j}$, which for an ideally conducting plasma becomes

$$\vec{B} \cdot \vec{j} \propto \vec{B} \cdot \left(\vec{\nabla} \times \vec{B} \right). \tag{2.1}$$

Here, the magnetic field is denoted by $\vec{B}$ and the electric current density by $\vec{j}$. The test uses measurements of the Faraday depth and of the polarization direction of synchrotron radiation to probe the magnetic field components along the line of sight and perpendicular to it, respectively. Its simple geometrical motivation should make it applicable in a general setting, provided these quantities can be measured. The results depend only on the properties of the magnetic field along a line of sight and are therefore purely local in the two-dimensional sky projection. Our aim is to test this idea on observational as well as on simulated data, thereby determining the conditions under which the test will yield useful results.

This paper is organized as follows. In Sect. 2.2, the basic equations used in the *LITMUS* test are reviewed. They are applied to observational data describing the Galactic magnetic field in Sect. 2.3, with special emphasis on a sophisticated reconstruction of the Faraday depth, described in Sect. 2.3.2. Section 2.4 is devoted to a thorough general assessment of the test's reliability. To this end it is applied to simulated observations of increasing complexity. Section 2.4.1 describes the application in a flat sky approximation, whereas Sect. 2.4.2 examines all-sky simulations, finally arriving at complete simulations of the Galactic setting in Sect. 2.4.2, where realistic electron distributions are added. We discuss our results and conclude in Sect. 2.5.

2.2 The helicity test

For a thorough introduction into the ideas behind the *LITMUS* test, the reader is referred to Junklewitz & Enßlin (2011). Here, we only summarize the resulting equations.

On the one hand side, synchrotron emission produced by cosmic ray electrons is used to probe the magnetic field component perpendicular to the line of sight. Its polarization is described by the complex field

$$P = Q + iU = |P| \, e^{2i\chi}, \tag{2.2}$$

where Q and U are the usual Stokes parameters quantifying the linearly polarized components of the radiation with respect to some orthogonal coordinate system and χ is the polarization angle with respect to the first coordinate direction. On the other hand, the Faraday depth

$$\phi \propto \int_{\text{LOS}} n_{\text{e}} \vec{B} \cdot \mathrm{d}\tilde{\mathrm{l}} \tag{2.3}$$

is used to probe the magnetic field component parallel to the line of sight (LOS).

A helical magnetic field will lead to a gradient of the Faraday depth that is parallel to the polarization direction of the synchrotron emission, as was argued in Junklewitz & Enßlin (2011). In order to compare the directions of the two quantities, this gradient is also formulated as a complex field

$$G = \left(\left(\vec{\nabla}\phi \right)_x + i \left(\vec{\nabla}\phi \right)_y \right)^2 = |G| \, e^{2i\alpha}, \tag{2.4}$$

with

$$\alpha = \arctan \left(\frac{\left(\vec{\nabla}\phi \right)_y}{\left(\vec{\nabla}\phi \right)_x} \right), \tag{2.5}$$

where the indices x and y denote its components with respect to the coordinates used. The helicity test that is performed in this work consists simply of multiplying G with the complex conjugate of the polarization P^*. If the two angles χ and α differ by a multiple of π (i.e. the gradient and the polarization direction are parallel), the product will be real and positive. If they differ by an odd multiple of $\pi/2$ (i.e. the two directions are perpendicular), it will be real and negative. Any orientation in between will produce varying real and imaginary parts in the product. Thus, observational directions along which a magnetic field is helical are indicated by a positive real part and a vanishing imaginary part of the product. Averaging over all observational directions will give an indication of the global helicity of the field.

It was furthermore shown by Junklewitz & Enßlin (2011) that the ensemble average of this product over all magnetic field realizations given a magnetic correlation tensor (and therefore a helicity power spectrum) is a measure for the squared integrated spectral current helicity density

$$\langle GP^* \rangle_{\vec{B}} \propto \left(\int_0^\infty \mathrm{d}k \frac{\epsilon_H(k)}{k} \right)^2, \tag{2.6}$$

with large scales weighted more strongly than small scales.

2.3 Application to Galactic observations

In this section, we try to answer the question whether the magnetic field of the Milky Way is helical by applying the *LITMUS* test to the available observational data. Since the magnetic field is localized in a region that surrounds the observer, all relevant quantities will be given as fields on the sphere $\mathcal{S}^2$, i.e. as functions of the observational direction, specified by two angles ϑ and φ, which are taken to represent the standard spherical polar coordinates in a Galactic coordinate system.

2.3.1   Observational data

For the synchrotron emission, we use the data gathered by the *WMAP* satellite after seven years of observations[1], described in Page et al. (2007). Since the foreground synchrotron emission is most intense at low frequencies, we use the measurement in the K-Band, which is centered at a frequency of $\nu = 23$ GHz. Furthermore, we assume that the detected polarized intensity is solely due to Galactic synchrotron emission. Thus, the Stokes Q and U parameter maps (defined with respect to the spherical polar coordinate directions $\hat{\vec{e}}_\vartheta$ and $\hat{\vec{e}}_\varphi$ in the Galactic coordinate system) can be simply combined according to Eq. (2.2) to give the complex quantity P whose argument is twice the rotation angle of the plane of polarization with respect to the $\hat{\vec{e}}_\vartheta$-direction

$$\chi(\vartheta, \varphi) = \frac{1}{2} \arctan\left(\frac{\text{Im}(P(\vartheta, \varphi))}{\text{Re}(P(\vartheta, \varphi))} \right) \tag{2.7}$$

(cf. Junklewitz & Enßlin 2011).

There are several depolarizing effects that have to be considered when dealing with polarization data. Faraday depolarization, which is important only at low frequencies due to the proportionality of the Faraday rotation angle to the square of the wavelength, can be safely neglected in the K-Band. Depolarization effects due to different magnetic field orientations along the line of sight are certainly present. However, they are present as well in the numerical test cases presented in Sect. 2.4, which yield good results. Additionally, depolarization due to the finite beam-size of the *WMAP* satellite and the finite pixel size of the polarization maps used in this study can play a role. The only way to limit this effect is to use higher resolution maps, ultimately necessitating the use of *Planck* data in the future.

In order to construct a map of the Faraday depth, we use the catalog of rotation measurements provided by Taylor et al. (2009)[2]. These provide an observational estimate of the Faraday depth for certain directions in the sky where polarized radio point-sources could be observed. Since the catalog encompasses a large number (37 543) of point-sources, it paints a rather clear picture of the structure of the Faraday depth. However, Earth's shadow prevents observations in a considerably large region within the southern hemisphere.

2.3.2   Reconstructing the Faraday depth map

The reconstruction is conducted according to the *critical filter* method first presented in Enßlin & Frommert (2011). A more elegant derivation of the same filter can be found in Enßlin & Weig (2010). Since this formalism takes into account available information on the statistical properties of the signal in the form of the power spectrum[3], it is able to interpolate into regions where no direct information on the signal is provided by the data,

---

[1]The data are available from NASA's Legacy Archive for Microwave Background Data Analysis at `http://lambda.gsfc.nasa.gov`.

[2]The catalog is available at `http://www.ucalgary.ca/ras/rmcatatlogue`.

[3]See App. A for notes on the definition and usage of power spectra for real and complex fields.
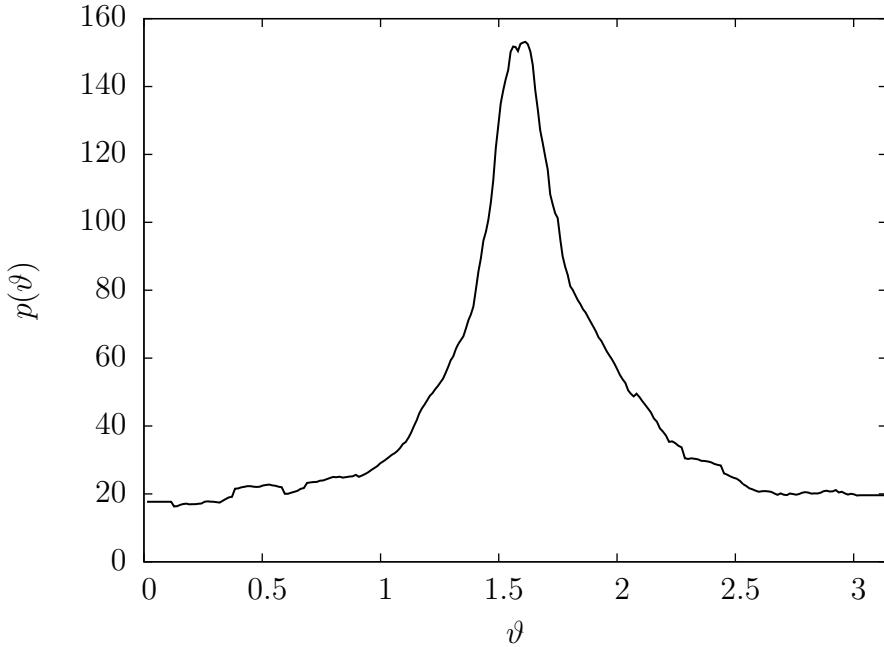
Figure 2.1: Vertical Galactic profile $p(\vartheta)$ of the Faraday depth.

such as the shadow of Earth in this case. Furthermore, it takes into account the available information on the uncertainty of the measurements. All in all it is expected to lead to a reconstructed map of the Faraday depth that is much closer to reality than e.g. a map in which the data were simply smoothed to cover the sphere. Small-scale features that are lost in such a smoothing process are for example reproduced by the *critical filter* algorithm.

Data model

The field that is to be reconstructed here is the sky-map of the Faraday depth. In order to apply the *critical filter* formula, the signal should be an isotropic Gaussian field. Since the Faraday depth clearly is larger along directions passing through the Galactic plane, the condition of isotropy is not satisfied. Therefore, a vertical profile is calculated by binning the observations into intervals $[\vartheta_i, \vartheta_i + \Delta\vartheta)$, calculating the root mean square rotation measure value for each bin and smoothing the resulting values to obtain a smooth function $p(\vartheta)$. The result is shown in Fig. 2.1. This profile is used to approximatively correct the anisotropies induced by the Galactic structure and the resulting signal field

$$s(\vartheta, \varphi) = \frac{\phi(\vartheta, \varphi)}{p(\vartheta)} \tag{2.8}$$

is assumed to be isotropic and Gaussian with a covariance matrix S. The Gaussian co-variance matrix is determined solely by the angular power spectrum coefficients C_l, the reconstruction of which is part of the problem at hand.

The data d, i.e. the rotation measure values in the catalog, are taken to arise from the signal s by multiplication with a response matrix R, which consists of a part encoding the specific directions in which the signal field is probed in order to produce the measurements and another part that is a simple multiplication with the vertical profile $p(\vartheta)$. Additionally, a Gaussian noise component n is assumed with a covariance matrix $N = \text{diag}(\sigma_1^2, \sigma_2^2, \dots)$, where σ_i is the one sigma error bar for the ith measurement in the catalog. Thus, the data are given by[4]

$$d = Rs + n = \tilde{R}ps + n. \tag{2.9}$$

Recent discussions in the literature (see e.g. Stil et al. 2011) have shown, however, that the error estimates as quoted in the catalog are probably too low. In addition, any contribution to the measured data from intrinsic Faraday rotation within the sources will further increase the error budget since the signal field in this context is only the contribution of the Milky Way. We therefore adapt the error bars of Taylor et al. (2009) according to the formula

$$\sigma^{(\text{corrected})} = \sqrt{(f_\sigma \sigma)^2 + (\sigma^{(\text{int})})^2}, \tag{2.10}$$

respectively. Here, the factor f_σ accounts for the general underestimation of the errors in the catalog of Taylor et al. (2009), whereas the additive constant $\sigma^{(\text{int})}$ represents the average contribution of the sources' intrinsic Faraday rotation. As numerical values, we use $f_\sigma = 1.22$, which was found by Stil et al. (2011) by comparing the data of Taylor et al. (2009) and Mao et al. (2010) on sources that are contained in both catalogs, and $\sigma^{(\text{int})} = 6.6\,\text{m}^{-2}$, which corresponds to the upper end of the numbers found by Schnitzeler (2010). Since the error contribution from the internal Faraday rotation is not correlated with the measurement error, the two contributions add up quadratically in the noise covariance matrix.

Reconstruction method

For signal and noise fields with a Gaussian prior probability distribution, the posterior probability distribution of the signal field is again a Gaussian, i.e. it is of the form $\mathcal{P}(s|d, C_l) = \mathcal{G}(s - m, D)$, where a multivariate Gaussian probability distribution with covariance matrix X is denoted by

$$\mathcal{G}(x, X) = \frac{1}{|2\pi X|^{1/2}} \exp\left(-\frac{1}{2}x^\dagger X^{-1} x\right), \tag{2.11}$$

where the $\dagger$ symbol denotes a transposed and complex conjugated quantity. In order to reconstruct the mean signal field $m = \langle s \rangle$, where the brackets denote the posterior mean, it is also necessary to reconstruct the angular power spectrum C_l. To do so, the *critical filter* formulas

$$m = Dj \tag{2.12}$$

---

[4]The discretized version used in the implementations is $d_i = \sum_j R_{ij} s_j + n_i = \sum_j \tilde{R}_{ij} p_j s_j + n_i$, where the index j determines a pixel on the sphere, so that $s_j = s(\vartheta_j, \varphi_j)$ and $p_j = p(\vartheta_j)$.
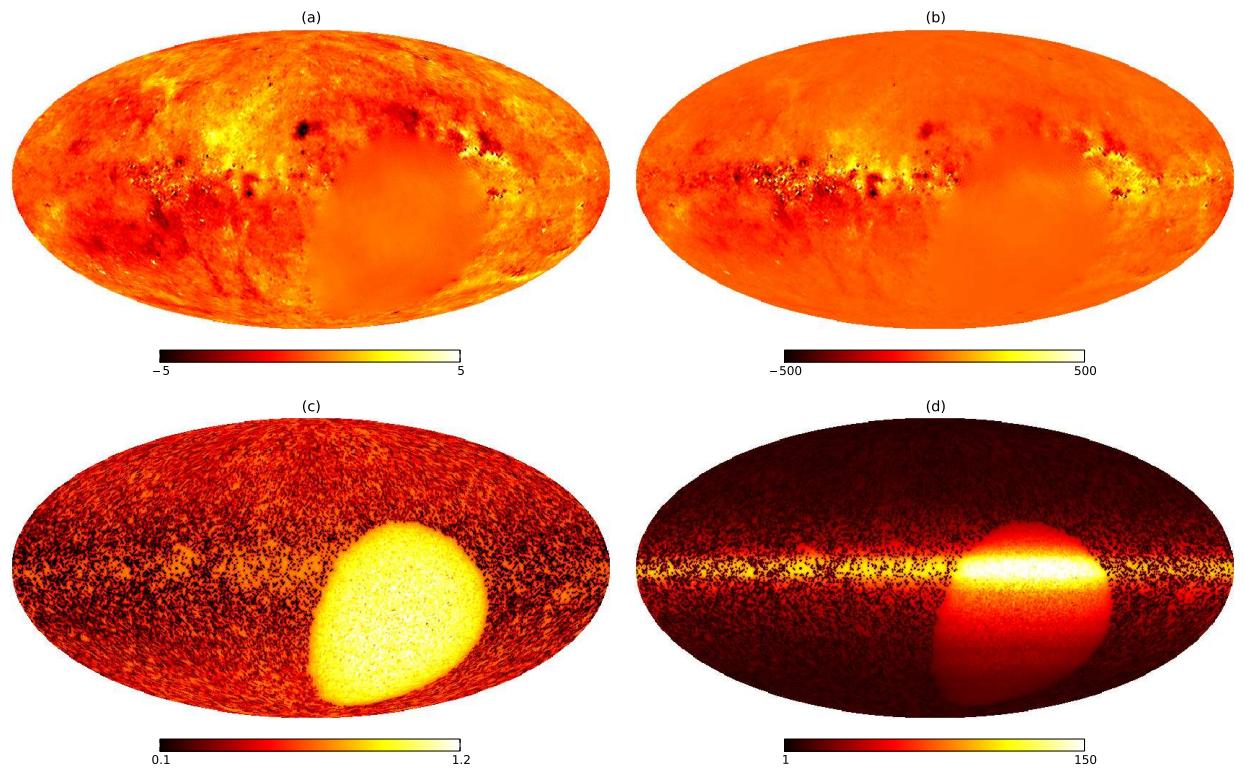
Figure 2.2: Results of the reconstruction of the signal field and Faraday depth. The left column shows the posterior mean of the signal field m (panel (a)) and its one-sigma uncertainty $\sqrt{\hat{D}}$ (panel (c)). The right column shows the resulting map of the Faraday depth pm (panel(b)) and the corresponding one-sigma uncertainty $\sqrt{p^2\hat{D}}$ (panel (d)) in m^{-2}.

and

$$C_l = \frac{1}{2l+1} \mathrm{tr}\left(\left(mm^\dagger + D\right) S_l\right) \tag{2.13}$$

are iterated, starting with some initial guess for the power spectrum. Here, the signal covariance matrix is expanded as $S = \sum_l C_l S_l$, where S_l is the projection onto the spherical harmonic components with index l. Furthermore, D is the posterior covariance matrix,

$$D = \left(S^{-1} + R^\dagger N^{-1} R\right)^{-1}, \tag{2.14}$$

and j is the information source term,

$$j = R^\dagger N^{-1} d. \tag{2.15}$$

Since the *critical filter* is on the brink of exhibiting a perception threshold (cf. Enßlin & Frommert 2011) and it is generally more desirable to overestimate a power spectrum entering a filter than to underestimate it, the coefficients C_l are subjected to a procedure in which the value of C_l is replaced by $\max\{C_{l-1}, C_l, C_{l+1}\}$ after each iteration step. The advantage of overestimating the power spectrum can be seen by considering the limit of $C_l \to \infty$ in Eq. (2.12) and (2.14). For high values of C_l, the first term in Eq. (2.14) can be neglected and Eq. (2.12) becomes

$$m_{C_l \to \infty} = s + R^{-1} n. \tag{2.16}$$

Thus, by overestimating the power spectrum the importance of its exact shape is diminished and the reconstruction will instead follow the information given directly by the data more closely. Considering an extreme underestimation of the power spectrum, $C_l \to 0$, on the other hand, would lead to

$$m_{C_l \to 0} = 0, \tag{2.17}$$

suppressing the information given by the data.

Calculating $\langle G \rangle_{\mathcal{G}(s-m,D)}$

Once the posterior mean of the signal is reconstructed, the corresponding field G can be calculated according to Eq. (2.4), assuming $\phi = m$. However, since Eq. (2.4) is a nonlinear transformation of the Faraday depth, that is not the same as the posterior mean of G. This can be estimated by averaging over a large number of samples, according to

$$\langle G(s) \rangle_{\mathcal{G}(s-m,D)} \approx \frac{1}{N_{\mathrm{samples}}} \sum_{k=1}^{N_{\mathrm{samples}}} G(s_k). \tag{2.18}$$

Here, the samples s_k are drawn from the posterior probability distribution $\mathcal{G}(s-m, D)$. In our implementation we use the method of Jasche et al. (2010b) to generate these samples, i.e. we generate corrections y_k to the posterior mean m by drawing signal and noise realizations from the respective prior probability distributions,

$$\tilde{s}_k \hookleftarrow \mathcal{P}(s|S), \tag{2.19}$$
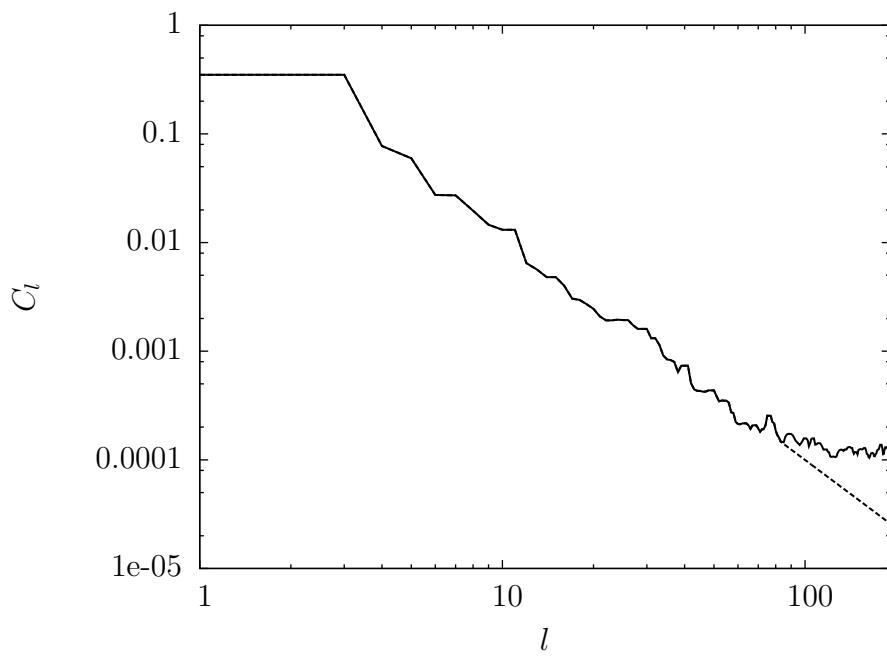
Figure 2.3: Angular power spectrum of the signal field s. The solid curve shows the resulting power spectrum as calculated with the *critical filter* formalism. The dashed line depicts the signal power spectrum used to generate large-scale uncertainty corrections for the calculation of G according to Sect. 2.3.2.
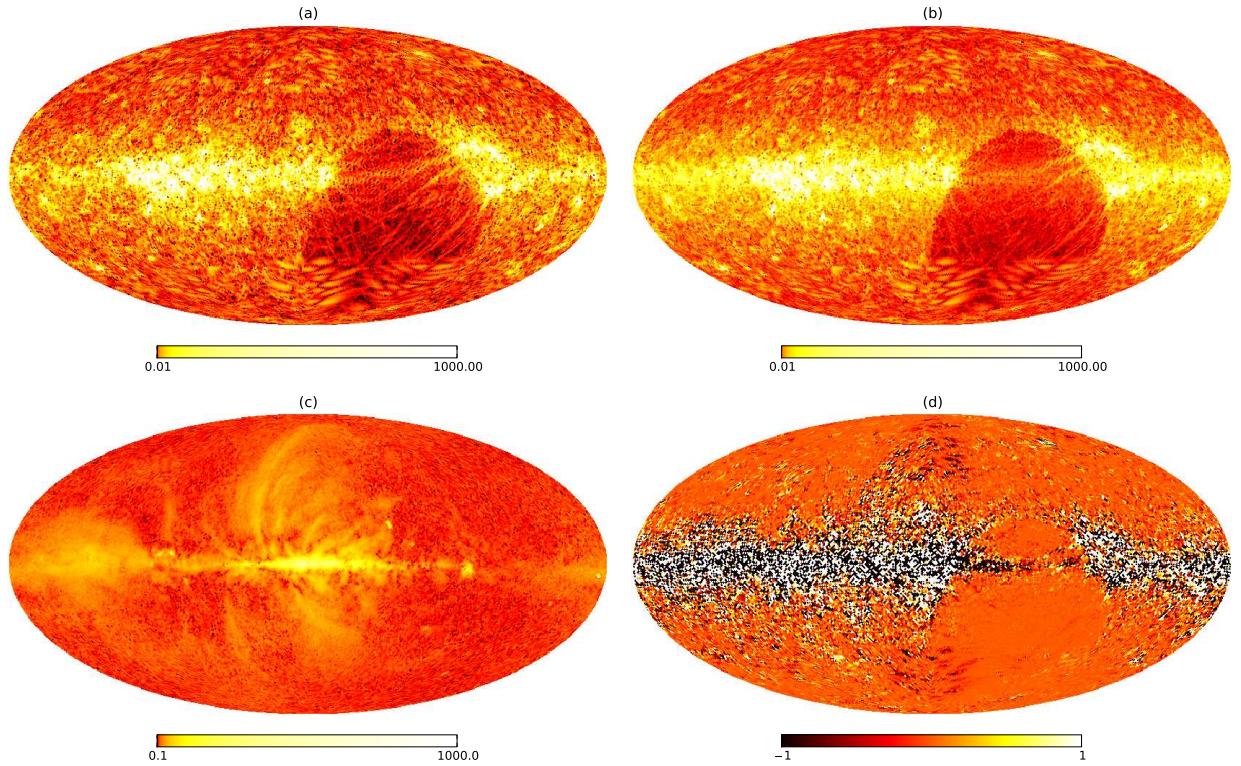
Figure 2.4: Results of the calculation of the gradient field G and the application of the *LITMUS* test. The first row shows estimates for the absolute values of the field G. Panel (a) shows G as calculated directly from the posterior mean m of the signal field, whereas panel (b) takes into account large-scale uncertainty corrections as described in Sect. 2.3.2. Panel (c) shows the absolute value of the polarized intensity P as measured by the *WMAP* satellite and panel (d) shows the result of the *LITMUS* test, namely $\mathrm{Re}\,(GP^*)$, in arbitrary units, using G and P as shown in panels (b) and (c), respectively. Note the logarithmic color code of panels (a)–(c).

$$n_k \hookleftarrow \mathcal{P}(n|N), \tag{2.20}$$

combining them into a data realization according to

$$d_k = R\tilde{s}_k + n_k, \tag{2.21}$$

and calculating the difference between the signal realization and the Wiener filtered data realization

$$y_k = \tilde{s}_k - DR^\dagger N^{-1} d_k. \tag{2.22}$$

The signal realizations

$$s_k = m + y_k \tag{2.23}$$

are shown by Jasche et al. (2010b) to follow the posterior probability distribution[5]. For reasons of numerical feasibility, we consider only large scale corrections to the posterior mean in our calculation, i.e. we introduce a small-scale power cut-off for the signal power spectrum that describes the prior probability distribution used in Eq. (2.19).

Thus, given m and D, the posterior mean for the Faraday depth is given by $\langle \phi \rangle = pm$, its one sigma error bars by $\pm p\sqrt{\hat{D}}$, and the posterior mean for the field G by the above sampling approximation, where the signal samples s_k are multiplied with the Galactic profile function $p(\theta)$ to give realizations of the Farady depth. Here, $\hat{D}$ is the vector that contains the diagonal elements of the matrix D.

2.3.3 Results

Figures 2.2 – 2.4 summarize the results of the Faraday depth reconstruction and the application of the *LITMUS* test to these data. All calculations are conducted at a HEALPix[6] resolution $N_{\text{side}} = 64$. The left column of Fig. 2.2 shows the reconstructed signal field as well as its uncertainty as they are calculated using the *critical filter* method. Evidently, the reconstruction method is able to extrapolate from the available information into regions where no data are taken, i.e. the Earth's shadow in the lower right of the projection. However, only structures on scales comparable to the extent of the region without information are reconstructed within it and the reconstruction's uncertainty becomes large in this region, as well as near the Galactic poles, where the signal-response to noise ratio of the data is low.

The outcome for the angular power spectrum is shown in Fig. 2.3. Clearly, the *critical filter* predicts a large amount of power on small scales. Therefore the rather high values seen in Fig. 2.2.c are mainly due to uncertainty on small scales. This may in part be due to an underestimation of the error bars of extreme data points which are not accounted for by our rather crude correction of the error bars as described by Eq. (2.10).[7]

---

[5]A cautionary note regarding a bias for the posterior variance calculated from these samples is given in App. B.

[6]The HEALPix package is available from `http://healpix.jpl.nasa.gov`.

[7]A thorough study of the problem of reconstructing a signal field with unknown power spectrum from data with unknown error bars will be presented in a future paper.

The right column of Fig. 2.2 shows the reconstructed Faraday depth and its uncertainty, obtained from the left column of Fig. 2.2 by a simple multiplication with the Galactic profile function $p(\vartheta)$. Clearly, the field $\langle \phi \rangle$ takes on only low values within the information-less region, whereas, again, its uncertainty is especially large there.

Figure 2.4 shows the two fields used in the *LITMUS* test as well as its result. The field G calculated according to the sampling prescription, Eq. (2.18), is shown in panel (b). In the sampling procedure, large-scale uncertainty corrections that are described by the signal power spectrum shown in Fig. 2.3 are taken into account. For comparison, the field G as calculated from the posterior mean of the signal field without any uncertainty corrections is shown in panel (a). The main effect of the uncertainty corrections is to increase the values within the region where no data points were measured and in the polar regions. The spatial structure of G, however, remains largely unchanged.

Finally, the real part of the product GP^* is shown in the last panel of Fig. 2.4. By visual inspection, it is not immediately clear whether positive values dominate this image. Taking the spatial average $\mathrm{Re}\, \langle GP^* \rangle_{\mathcal{S}^2}$ does in fact yield a negative value. Nevertheless, we performed another test for helicity. In order to assess helicity on small scales, the two fields G and P are rotated with respect to one another about an angle β around the Galactic axis and their product is again averaged over the whole sky. For a helical magnetic field, this procedure should result in a curve $\mathrm{Re}\, \langle GP^* \rangle_{\mathcal{S}^2} (\beta)$ that exhibits a maximum at $\beta = 0$, since the correlations between the fields G and P are expected to be local. The resulting curve for this test is shown in Fig. 2.5. Not only does it not take on its global maximum at $\beta = 0$, but the point $\beta = 0$ does not seem to be special in any way. Therefore, no indication of helicity is found.

We also performed the analysis with a different function $p(\vartheta)$, created by doubling the smoothing length in its calculation. The results of the reconstructed map of the Faraday depth changed only slightly, i.e. the variance of the difference between the reconstructions with the two different profiles is 0.65% of the variance of the original $\langle \phi \rangle$-map shown in Fig. 2.2.b.

2.4   Application to simulated data

In order to check whether the nondetection of helicity in the previous section allows the conclusion that the Galactic magnetic field is in fact nonhelical, we now apply the same helicity test to a number of artificially generated magnetic fields with known helicity.

2.4.1   Planar implementation

The most simple setting that can be considered is the observation of a well localized magnetic field structure. In the limit of great distances between the magnetic field under consideration and the observer, the lines of sight penetrating the field become parallel. We assume the field to be contained in a cubic box which is oriented along the lines of sight.
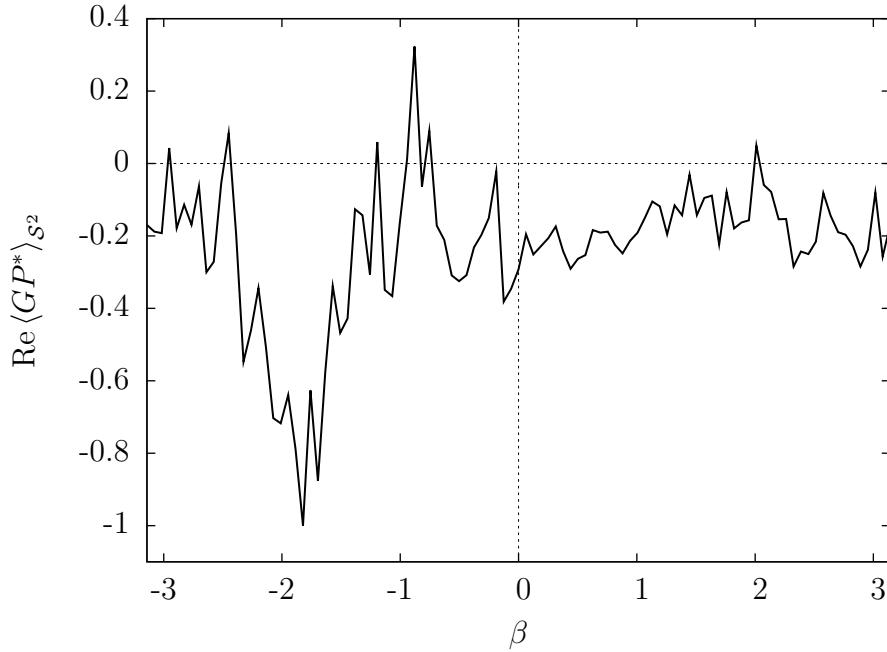
Figure 2.5: Rotational curve for $\mathrm{Re}\,\langle GP^*\rangle_{\mathcal{S}^2}$ in arbitrary units.

The field in the box is generated by the GARFIELDS code (first applied in Kitaura & Enßlin 2008). This code draws the three cartesian components of the magnetic field in Fourier space independently of a common power spectrum, assumed here to be a Kolmogorov-type spectrum of the form $P_B(k) \propto k^{-5/3-2}$, according to Gaussian statistics. In order to produce a magnetic field without divergence, its frequency components parallel to the respective k-vector are then subtracted

$$\vec{B}_{\mathrm{div\text{-}free}}(\vec{k}) \propto \vec{B}(\vec{k}) - \vec{k}\frac{\vec{k}\cdot\vec{B}(\vec{k})}{k^2}. \tag{2.24}$$

A degree of helicity is then imprinted onto the field by applying the formula

$$\vec{B}_{\mathrm{div\text{-}free,hel}}(\vec{k}) \propto \vec{B}_{\mathrm{div\text{-}free}}(\vec{k}) + \eta\frac{i\vec{k}\times\vec{B}_{\mathrm{div\text{-}free}}(\vec{k})}{k}, \tag{2.25}$$

where $\eta = 0$ leaves the field unaffected and $\eta = \pm 1$ produces the highest degree of helicity.

Finally, we assume the thermal and cosmic ray electron densities to be constant throughout the box. Thus, the observables Q, U, and ϕ can be obtained by simply integrating the appropriate magnetic field components along the box direction associated with the line of sight. Then the complex quantities G and P are easily calculated and multiplied, yielding two-dimensional images of GP^*. This procedure is conducted for various realizations of random magnetic fields both without helicity ($\eta = 0$) and with maximal helicity ($\eta = 1$).
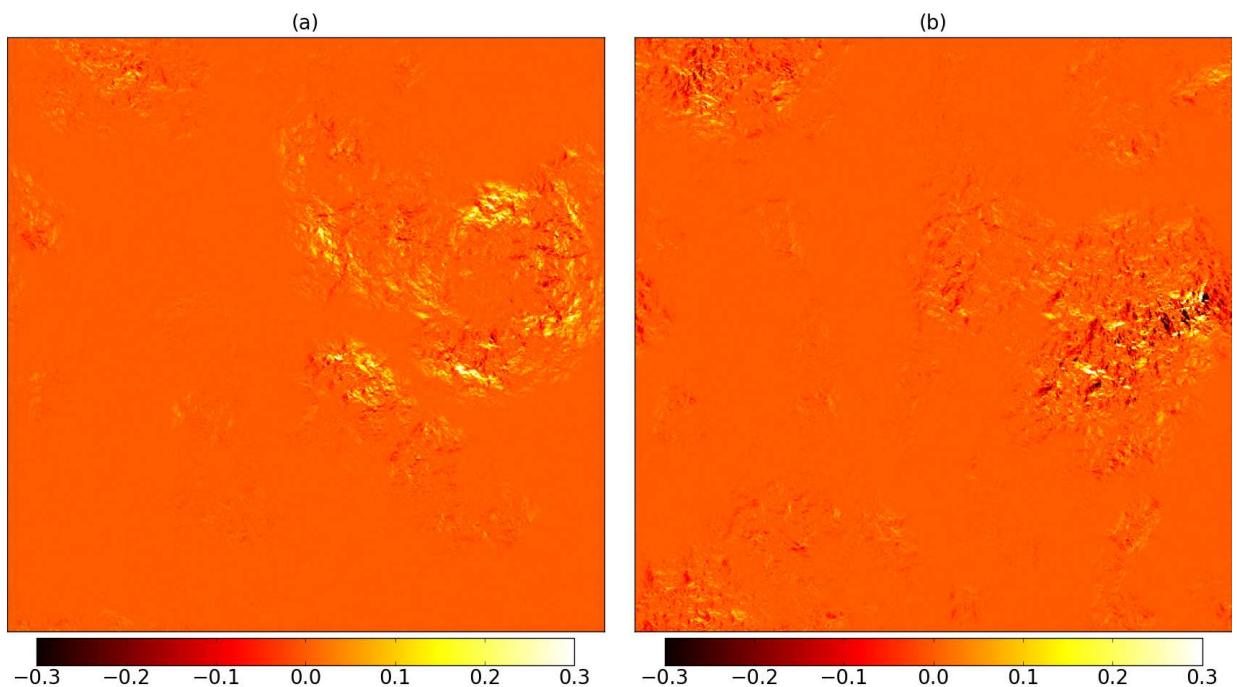
Figure 2.6: Maps of Re (GP^*) for a particular magnetic field realization in arbitrary units. Panel (a) shows the case with maximal helicity $(\eta = 1)$, panel (b) the one without helicity $(\eta = 0)$.

Results

The resulting images for one random magnetic field realization are shown in Fig. 2.6. The cube was discretized for the calculation into 512^3 pixels. It can already be seen by eye that positive values of $\mathrm{Re}\,(GP^*)$ dominate in the case with maximal helicity (panel (a)), whereas in the case without helicity (panel (b)), positive and negative values seem to be roughly equally represented. Calculating the spatial averages over the whole square yields $\mathrm{Re}\,\langle GP^*\rangle_\square = 2.3\cdot 10^{-3}$ and $\mathrm{Re}\,\langle GP^*\rangle_\square = -1.5\cdot 10^{-4}$ for the case with and without helicity, respectively.

We calculated this spatial average for the results of the *LITMUS* test applied to 100 different random magnetic field realizations, both with and without helicity. Averaging these values for the helical fields and for the nonhelical fields separately yields a positive value in the helical case. Normalizing all values such that this average becomes equal to one yields

$$\langle \mathrm{Re}\,\langle GP^*\rangle_\square\rangle_{\mathrm{samples}} = 1.0, \quad \sigma_{\mathrm{Re}\langle GP^*\rangle_\square} = 0.63$$

in the case with helicity and

$$\langle \mathrm{Re}\,\langle GP^*\rangle_\square\rangle_{\mathrm{samples}} = -0.025, \quad \sigma_{\mathrm{Re}\langle GP^*\rangle_\square} = 0.34$$

in the case without. Clearly, the *LITMUS* test yields positive results if applied to helical fields, whereas its results fluctuate around zero if applied to nonhelical fields. This is exactly the behavior that should be expected and the basic functioning of the *LITMUS* test is thereby demonstrated in this setting.

2.4.2 Spherical implementation

As a next step, the applicability of the *LITMUS* test is checked for magnetic fields surrounding the observer, as in the case of the Galactic field. Again, several sets of mock observations are produced. These simulations are conducted using the HAMMURABI code (see Waelkens et al. 2009) in connection with the GARFIELDS code[8]. The HAMMURABI code allows for a large scale analytic field model and an additional Gaussian random field component, which can be generated by the GARFIELDS code with a preset degree of helicity as described in Sect. 2.4.1. It then integrates the different field components, weighted with the appropriate electron density, along radial lines of sight and produces sky maps of simulated observations of the Stokes parameters Q and U and Faraday depth ϕ (among others), thus providing all necessary ingredients to perform the *LITMUS* procedure.

Constant electron densities

For simplicity, we start again by setting the densities of the thermal electrons and the cosmic ray electrons to constant values throughout the simulated galaxy.

---

[8]Both codes are available from `http://www.mpa-garching.mpg.de/hammurabi/hammurabi11`.
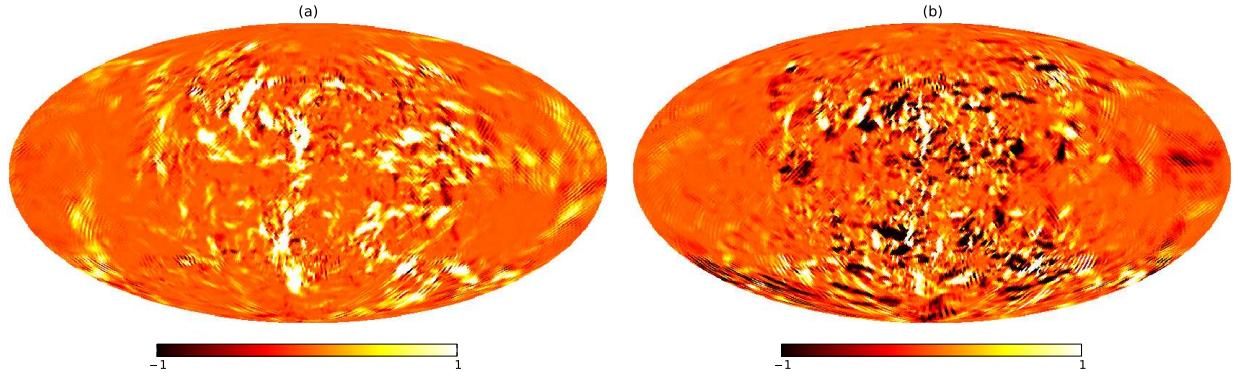
Figure 2.7: Maps of $\mathrm{Re}\,(GP^*)$ for a particular magnetic field realization in arbitrary units in a spherical setting. Panel (a) shows the case with maximal helicity ($\eta = 1$), panel (b) the one without helicity ($\eta = 0$).

Gaussian random field.   First, in order to apply the test to a field with a well-defined degree of helicity, the field strength of the large scale analytic component is set to zero, such that the simulated Galactic field is a purely random one with a chosen degree of helicity. As in the planar case, we choose either no helicity ($\eta = 0$) or maximal helicity ($\eta = 1$).

Figure 2.7 shows the maps of $\mathrm{Re}\,(GP^*)$ for one particular Gaussian random magnetic field with a power law index of $-5/3$, with helicity and without helicity, respectively. As for the planar implementation, one can immediately see that positive values dominate in the helical case and positive and negative values are roughly equally represented in the nonhelical case.

Figure 2.8 shows the results of the rotational test described in Sect. 2.3.3 for this particular magnetic field realization. Clearly, the spatial average $\mathrm{Re}\,\langle GP^*\rangle_{\mathcal{S}^2}$ takes on a sharp maximum at $\beta = 0$ and is positive in the case with helicity, while it does not have a maximum there and in fact happens to be negative in the case without helicity. This is exactly the result expected from the *LITMUS* test.

The sharpness of the peak in Fig. 2.8 indicates that the helicity is to be found in small-scale features. As a test of this assertion, we calculated the spherical multipole components G_{lm} and P_{lm} of the complex gradient and polarization fields. Note that the spatial average over the product of the fields is proportional to the sum of the products of the multipole components, i.e.

$$\mathrm{Re}\,\langle GP^*\rangle_{\mathcal{S}^2} = \frac{1}{4\pi}\mathrm{Re}\left(\sum_{l=0}^{l_{\mathrm{max}}}\sum_{m=-l}^{l} G_{lm}P_{lm}^*\right), \tag{2.26}$$

where l_{max} is determined by the finite resolution of the map. If we now neglect the first terms in the sum, i.e. the small-l contributions, we arrive at a spatial average over the product in which all large-scale features were neglected. The resulting rotational curves for the same magnetic fields used for Fig. 2.8 are shown in Fig. 2.9, where only multipole

moments with $l \geq 25$ were considered. It can be seen that this procedure further sharpens the peak at $\beta = 0$ and strengthens it relative to other local maxima in the curves, thus facilitating the detection of small-scale helicity. The same result for the observational data studied in Sect. 2.3 is shown in Fig. 2.10. Clearly, there is still no sign of helicity in this case.

Furthermore, we created a set of 100 different Gaussian random magnetic fields, performed the *LITMUS* test, and calculated the spatial average Re $\langle GP^* \rangle_{\mathcal{S}^2}$ for all of them. Each field realization was considered in a version without helicity and a version with maximal helicity, just as in the case of the planar implementation. Averaging over the 100 samples yields again a positive value in the helical case. Normalizing all values such that this average is equal to one yields

$$\langle \mathrm{Re}\, \langle GP^* \rangle_{\mathcal{S}^2} \rangle_{\mathrm{samples}} = 1.0, \quad \sigma_{\mathrm{Re}\langle GP^* \rangle_{\mathcal{S}^2}} = 0.25$$

in the helical case and

$$\langle \mathrm{Re}\, \langle GP^* \rangle_{\mathcal{S}^2} \rangle_{\mathrm{samples}} = -0.27, \quad \sigma_{\mathrm{Re}\langle GP^* \rangle_{\mathcal{S}^2}} = 0.23$$

in the nonhelical case. This clearly underlines the success of the *LITMUS* test in the spherical setting.

Since the magnetic field of the Milky Way is known to be much stronger in the vicinity of the Galactic plane than in the halo region, we repeated the analysis with random realizations of Gaussian magnetic fields that are confined to the region within a vertical distance of 8 kpc of the Galactic plane. The results of the *LITMUS* procedure for these fields still turn out to be a reliable indicator for magnetic helicity.

Large scale field models. As a next step, the (helical or nonhelical) random magnetic field component is switched off completely and replaced by an analytic large scale magnetic field model. Several sets of simulations are performed using different models for the Galactic large scale field. For these analytic models, the current helicity can be calculated directly, giving an expectation as to whether the helicity test should produce positive results or not.

The results of the rotational helicity test are shown in Fig. 2.11 for the large scale magnetic field model described in Page et al. (2007), i.e.

$$\begin{aligned}
\vec{B}(r, \varphi, z) = B_0 \big[&\cos\left(\psi(r)\right) \cos\left(\chi(z)\right) \hat{\vec{e}}_r \\
&+ \sin\left(\psi(r)\right) \cos\left(\chi(z)\right) \hat{\vec{e}}_\varphi \\
&+ \sin\left(\chi(z)\right) \hat{\vec{e}}_z \big]
\end{aligned} \tag{2.27}$$

in Galactic cylindrical coordinates, where

$$\psi(r) = \psi_0 + \psi_1 \ln\left(\frac{r}{8\mathrm{kpc}}\right) \tag{2.28}$$
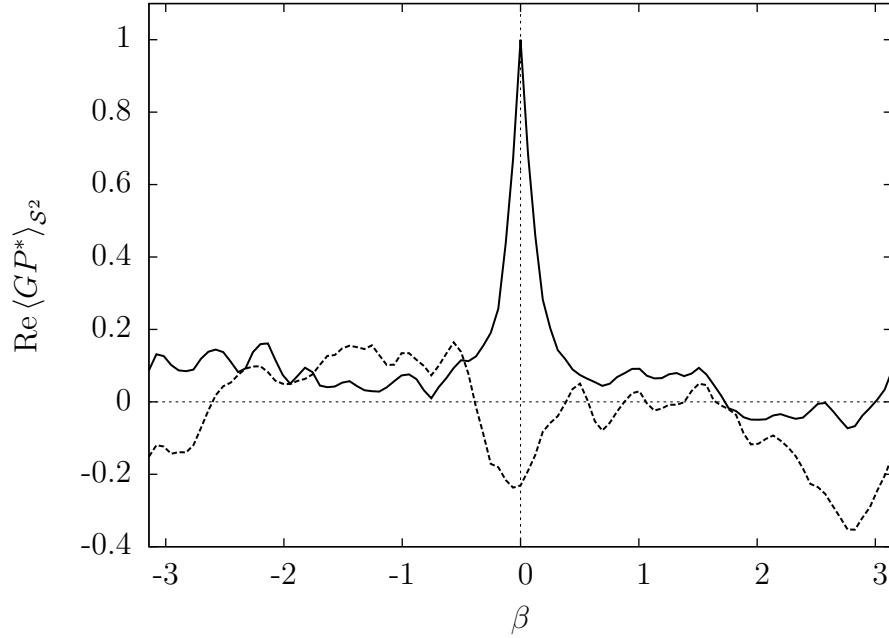
Figure 2.8: Rotational curve for $\mathrm{Re}\,\langle GP^*\rangle_{\mathcal{S}^2}$ in arbitrary units. The solid and dashed lines depict the results for the same magnetic field with helicity parameter $\eta = 1$ and $\eta = 0$ respectively. Constant electron densities are assumed.
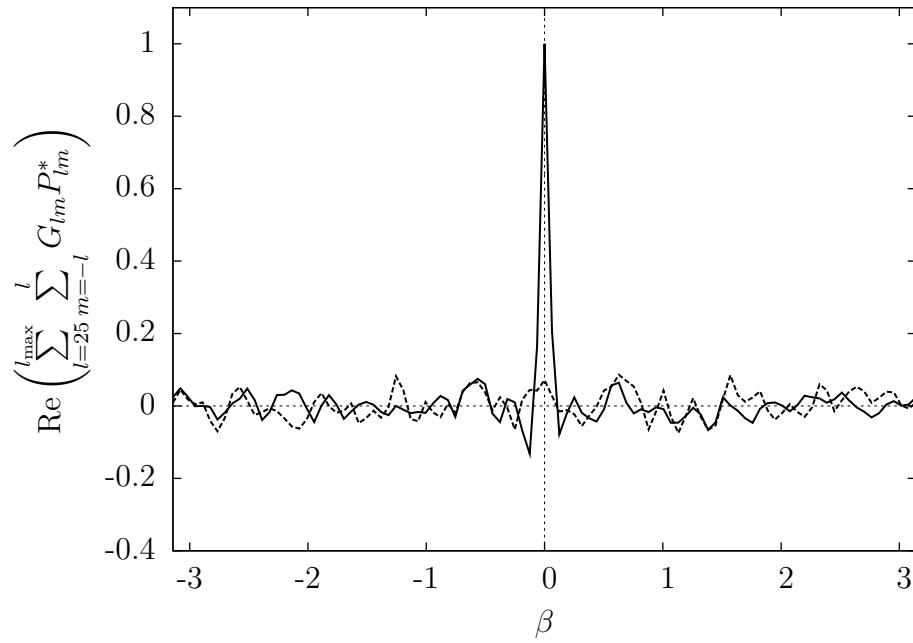
Figure 2.9: Same as Fig. 2.8, only with large-scale contributions neglected.
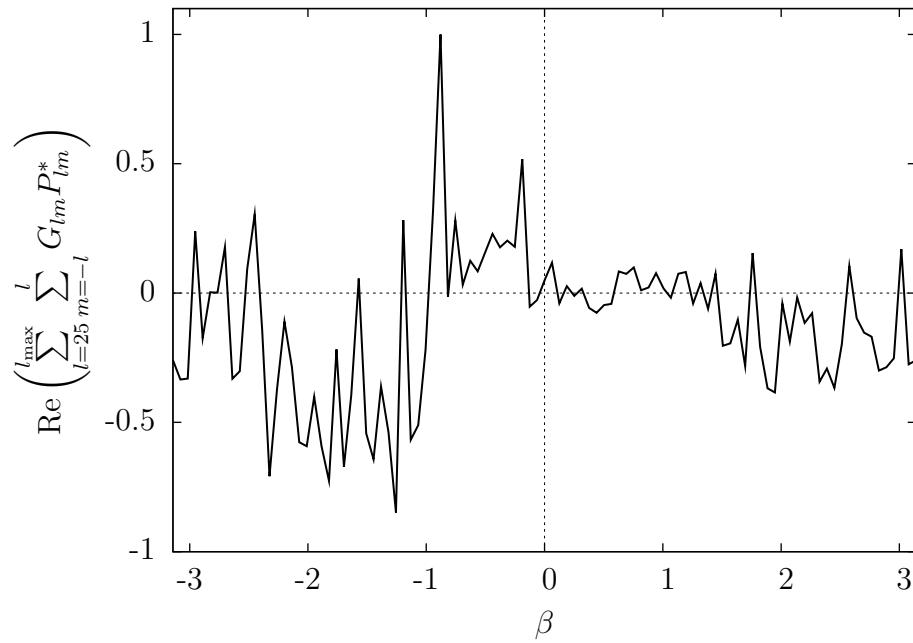
Figure 2.10: Same as Fig. 2.5, only with large-scale contributions neglected.
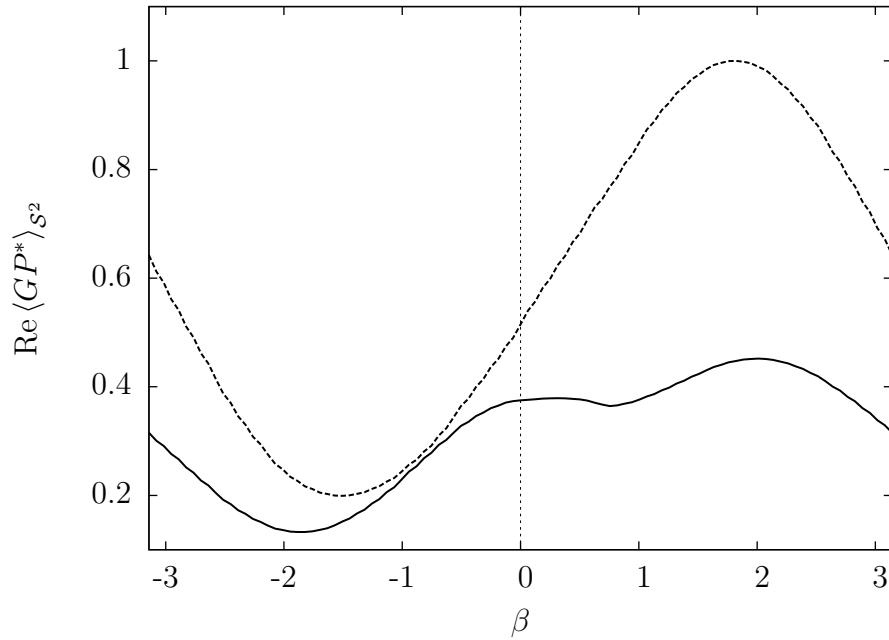
Figure 2.11: Rotational curve for $\mathrm{Re}\,\langle GP^* \rangle_{\mathcal{S}^2}$ in arbitrary units. The solid line corresponds to the best-fit magnetic field model of Page et al. (2007), the dashed line to a model where the parameter χ_0 is set to zero. Constant electron densities are assumed.

and

$$\chi(z) = \chi_0 \tanh\left(\frac{z}{1\text{kpc}}\right). \tag{2.29}$$

The solid curve corresponds to the parameters favored by Page et al. (2007), namely $\chi_0 = 25°$, $\psi_0 = 27°$, and $\psi_1 = 0.9°$. This corresponds to a simple flat spiral in the Galactic plane which becomes more and more screw-like with vertical distance z from the Galactic plane, so that a slight degree of helicity is inherent in the field geometry. This can be verified by direct calculation according to Eq. (2.1), yielding

$$\vec{j} \cdot \vec{B} = B_0^2 \frac{\sin(\chi(z))\cos(\chi(z))}{r} \left(\sin(\psi(r)) + \psi_1 \cos(\psi(r))\right), \tag{2.30}$$

which is nonzero for any generic point away from the Galactic plane. The resulting line in Fig. 2.11 is *not* a clear indication for this helicity. However, the curve is nevertheless sensitive to the angle χ_0, which produces the helicity. Lowering its value, i.e. making the spirals more and more parallel to the Galactic plane, changes the results of the *LITMUS* test. The extreme case of $\chi_0 = 0$, i.e. $B_z = 0$ everywhere, for which the value of Eq. (2.30) becomes zero everywhere, is also shown in Fig. 2.11. This curve's value at $\beta = 0$ is even more distinct from its maximum than in the case of the solid line. This example shows that while the results react in a systematic way on changes in the parameters, the test is not suited to detect helicity on the largest scales.

We used the model of Page et al. (2007) in the demonstration of this effect mainly because of its mathematical simplicity. More sophisticated models can be found e.g. in Sun et al. (2008), Jansson et al. (2009), Jaffe et al. (2010), and references therein.

The role of the electron densies

In order to get closer to a realistic model of the Milky Way, as a next step we replace the constant electron densities with realistic models. The HAMMURABI code allows the use of the NE2001 model for the thermal electron density (cf. Cordes & Lazio 2002, 2003) to compute the Faraday depth and several analytic models for the cosmic ray electron density to compute the synchrotron emissivity (see also Waelkens et al. 2009). In the calculations performed to obtain the results presented here, the cosmic ray electron distribution model of Page et al. (2007) was used.

Gaussian random field. The resulting rotational curves, as calculated with the realistic electron distributions, for the case of the Gaussian random field are shown in Fig. 2.12. Neither in the case with helicity ($\eta = 1$), nor in the case without helicity ($\eta = 0$) does the curve take on its maximum at $\beta = 0$. This is the expected result in the latter case but contradicts the expectation in the former one. Therefore, the helicity imprinted onto the small scale magnetic field clearly fails to be detected by the test applied.
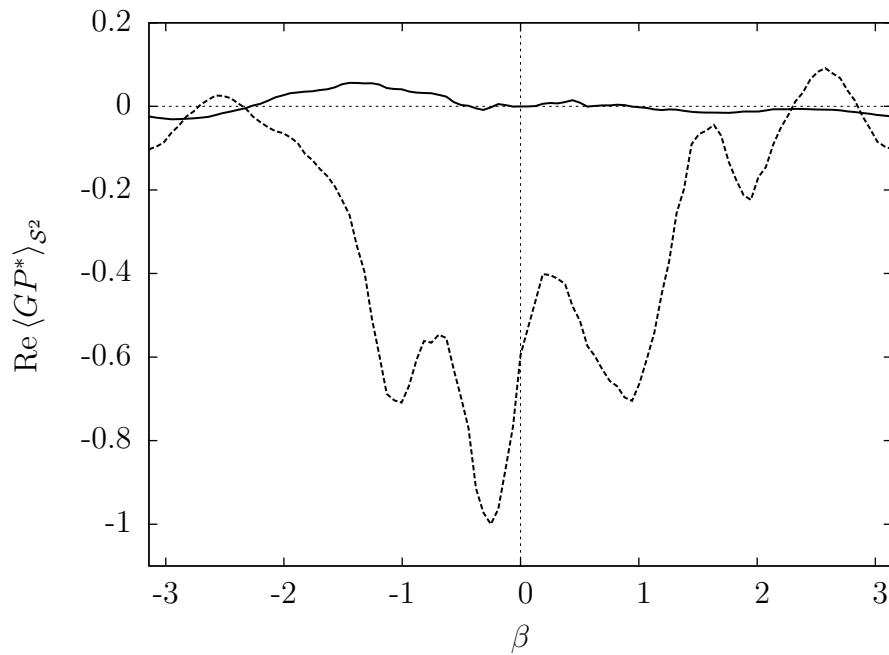
Figure 2.12: Same as Fig. 2.8, only with realistic electron densities used in the calculation.
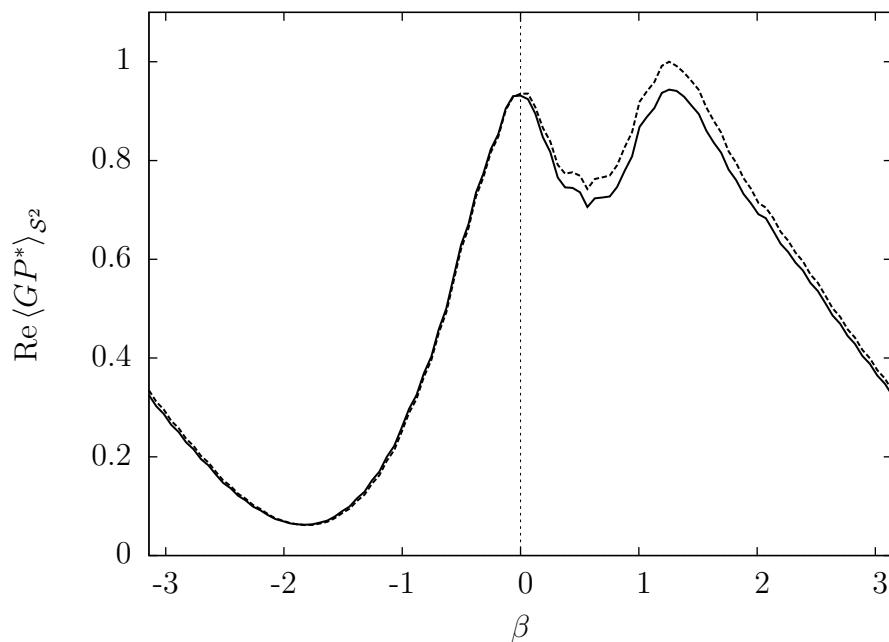
Figure 2.13: Same as Fig. 2.11, only with realistic electron densities used in the calculation.

Large scale field models. The result for the large scale magnetic fields is shown in Fig. 2.13. There actually seems to be a maximum in the vicinity of $\beta = 0$ now. This is true, however, for the case of the planar spiral model ($\chi_0 = 0$) as well as for the model with helicity ($\chi_0 = 25°$). Therefore, the proposed helicity test might under certain circumstances even indicate helicity on large scales where there is none, if the observer is surrounded by the field. Other magnetic field models with planar spirals, such as the bisymmetric (i.e. $\vec{B}(r, \varphi, z) = -\vec{B}(r, \varphi + \pi, z)$) spiral model of Stanev (1997) lead to similar results.

2.5 Discussion and conclusion

The present work presents the first application of the *LITMUS* test for magnetic helicity proposed by Junklewitz & Enßlin (2011) to actual data. The application of the test involves thoroughly reconstructing a map of the Faraday depth distribution, calculating its transformed gradient field G, creating a map of GP^*, averaging over this map, shifting the two fields with respect to each other to see whether any signal vanishes, and filtering out large-scale contributions for a better detection of small-scale helicity. This procedure, applied to observations of the Faraday depth and polarization properties of the synchrotron radiation within our own galaxy in Sect. 2.3, does not show any signs of helicity in the Milky Way's magnetic field.

In order to assess the significance of this, the applicability of the test was probed in different artificial settings. The complexity of these settings was increased bit by bit to find out under what circumstances exactly the *LITMUS* test yields reliable results. It was found that meaningful results can be achieved if the electron densities do not vary on the scales of the magnetic field, both in the regime of magnetic field structures whose distance from the observer is much greater than their extension, as shown in Sect. 2.4.1, and in the regime of magnetic fields surrounding the observer, as shown in Sect. 2.4.2. We showed that the performance of the *LITMUS* test with regard to small-scale helicity is further improved by dropping the first few terms in Eq. (2.26). However, indications of helicity on large scales are unreliable, as shown in Sect. 2.4.2. Furthermore, it was demonstrated in Sect. 2.4.2 that any non-trivial electron density may distort the outcome of the test to a point where even small-scale helical structures fail to be detected. This is not too surprising since e.g. a variation in the thermal electron density will introduce a gradient in the Faraday depth that is not caused by the magnetic field structure. Therefore the nondetection of helicity for the Galactic magnetic field does not necessarily mean that the field is nonhelical on small scales. It may be the case that small-scale fluctuations of the electron density introduce effects in the observational data that prevent the detection of helicity.

So, as a natural next step, the hunt for helicity in astrophysical magnetic fields should focus on a region that is small and/or homogeneous enough for the assumption of constant electron densities to hold at least approximatively. Although our work has shown that the helicity test that we studied is not suitable for all astrophysical settings, we are confident that it may nevertheless yield useful results if applied in a setting with constant electron

densities.

As a side effect of this paper, it was demonstrated in Sect. 2.3.2 that the method proposed by Enßlin & Frommert (2011) to reconstruct a Gaussian signal with unknown power spectrum is very well suited for practical application.

Chapter 3

Reconstructing signals from noisy data with unknown signal and noise covariance

Note: This chapter, as well as Appendix C, has been published in Physical Review E *(Oppermann et al. 2011b).*

3.1   Introduction

The problem of signal inference consists of reconstructing a set of parameters or even a continuous field s from some data set d, which is influenced in some way by the signal,

$$d = f(s) + n. \tag{3.1}$$

Two problems will arise. First, the function f may not be invertible and, second, the noise term n will not be known. In the Bayesian framework, one uses prior information on the signal and the noise term to calculate a best estimate for the true signal realization or, ideally, the whole probability distribution for the signal given the prior information and the information contained in the data.

Symmetry considerations and knowledge about the underlying physics of the signal and the measurement process may restrict the class of priors that one has to consider. They might, however, still contain some free parameters that then become part of the inference problem. The case in which the signal covariance contains uncertain parameters was tackled by Enßlin & Frommert (2011), producing a whole class of filters for this problem. The filter that we extend in this work was reproduced by Enßlin & Weig (2010), where the principle of minimum Gibbs free energy was introduced (cf. also Sect. 3.3), and successfully applied in an astrophysical setting by Oppermann et al. (2011a).

Here, we focus on the case where we can assume zero-mean Gaussian priors both for the signal and for the noise. The priors are therefore completely characterized by the respective covariance matrices. Our goal is to extend the study of Enßlin & Weig (2010) to the case

in which both the signal covariance and the noise covariance contain parameters that are not known a priori. This is motivated mainly by applications from the field of astrophysics. The theory and resulting filter formulas, however, are of general applicability. Gaussian noise, e.g., is omnipresent in nearly every area of the natural sciences and the situation in which its variance is not precisely known should be a rather common one.

Previous work dealing with the problem of unknown noise variance has mainly dealt with specific applications. One of these applications is the field of image reconstruction. Here, it is usually assumed that the measured picture is the sum of the underlying signal and a white Gaussian noise term. Often, it is further assumed that the noise level, i.e. its variance, is the same in every image pixel. A comparison of different algorithms for noise estimation under these assumptions was conducted e.g. by Olsen (1993). An example for an algorithm allowing for inhomogeneous noise was presented by Starck & Murtagh (1998), where a wavelet transform of the image is applied and the lack of correlated noise is exploited. Most of these algorithms, however, are not derived by rigorous statistical calculations but rather by a combination of intuition and experience.

From a mathematical viewpoint, the problem of an unknown noise prior has received some attention in the theory of density deconvolution, which deals with the inference of the probability density for a signal from measurements with additive noise. Here, the signal is usually assumed to consist of independent identically distributed variables. The case of Gaussian noise with unknown variance has been considered e.g. by Koltchinskii (2000) and Schwarz & van Bellegem (2010).

In this work, we create a general setting with well defined assumptions and a traceable derivation of a general filter formula within a Bayesian framework, not loosing sight of its applicability. Our result can accomodate a host of different assumptions and models, such as correlated or uncorrelated noise. It allows for a distinction between the data space and the signal space, with possibly different numbers of degrees of freedom.

The remainder of the paper is organized as follows. In Sect. 3.2 we introduce our model for the measurement process and the notation that is required. The derivation of the filter formulas follows in Sect. 3.3. We then demonstrate the usefulness of our filter by applying it in a set of mock observational situations in Sect. 3.4 and discuss the implications in Sect. 3.5.

3.2 Signal model and notation

We assume a linear measurement process where the data are a superposition of a linear signal response and a noise term,

$$d = Rs + n. \tag{3.2}$$

Here, both the data and noise and the signal can be finite-dimensional vectors or continuous fields defined on some manifold. The response matrix R maps a field in the signal space to a field in the data space. In the continuous limit the matrix vector product becomes

$$(Rs)_i = \int \mathrm{d}x \ R_{ix} s_x, \tag{3.3}$$

where the index denotes the value of a field at this position. In physical applications the data vector will always be discrete, since only a finite number of measurements can be taken, while the signal space might well be continuous. The result of any numerical signal reconstruction, however, will at best be a discretized version of the continuous field.

Further, we assume Gaussian prior statistics both for the signal and for the noise contribution, i.e. $s \hookleftarrow \mathcal{G}(s, S)$, $n \hookleftarrow \mathcal{G}(n, N)$, where

$$\mathcal{G}(a, A) = \frac{1}{\sqrt{|2\pi A|}} \exp\left(-\frac{1}{2} a^\dagger A^{-1} a\right) \tag{3.4}$$

denotes a multivariate Gaussian distribution in a with covariance matrix A. We use the dagger symbol to indicate a scalar product,

$$a^\dagger b = \int \mathrm{d}x \; a_x^* b_x, \tag{3.5}$$

and the asterisk to denote complex conjugation. This corresponds to the notation introduced by Enßlin et al. (2009).

The problem of signal reconstruction is to find an optimal estimate m for the signal realization that the measured data arose from. Optimality in an L^2-norm sense leads to

$$m = \langle s \rangle_{\mathcal{P}(s|d)} := \int \mathcal{D}s \; s\mathcal{P}(s|d), \tag{3.6}$$

i.e. the posterior mean. The integration is performed over all possible signal configurations. In the discrete case this becomes a product of one-dimensional integrals,

$$\int \mathcal{D}s = \int_{-\infty}^{+\infty} \mathrm{d}s_1 \int_{-\infty}^{\infty} \mathrm{d}s_2 \cdots, \tag{3.7}$$

where $s = (s_1, s_2, \dots)$ is the vector of signal values at locations $1, 2, \dots$. Ideally, we would also like to obtain some information on the posterior distribution $\mathcal{P}(s|d)$ other than its mean. If the signal and noise covariances are known, the posterior is a Gaussian $\mathcal{G}(s-m, D)$ with mean

$$m = Dj, \tag{3.8}$$

and covariance D, where $j = R^\dagger N^{-1} d$ is called the *information source* and $D = (S^{-1} + R^\dagger N^{-1} R)^{-1}$ the *information propagator* (cf. Enßlin et al. 2009) and the dagger attached to a matrix denotes its hermitian conjugate.

In this paper we are concerned with the case in which neither the signal covariance matrix S nor the noise covariance matrix N are known. We parameterize these matrices as sums of their eigenvalues $\tilde{p}_k$ and $\tilde{\eta}_j$ multiplied with the projectors onto the respective eigenspaces $\tilde{S}_k$ and $\tilde{N}_j$. The parameters can be rescaled by including some numerical values

$\tilde{s}_k$ and $\tilde{n}_j$ in the projection-like matrices, making the rescaled version of the parameterization

$$S = \sum_k p_k S_k, \tag{3.9}$$

$$N = \sum_j \eta_j N_j, \tag{3.10}$$

where

$$p_k = \frac{\tilde{p}_k}{\tilde{s}_k}, \ \eta_j = \frac{\tilde{\eta}_j}{\tilde{n}_j} \tag{3.11}$$

and

$$S_k = \tilde{s}_k \tilde{S}_k, \ N_j = \tilde{n}_j \tilde{N}_j. \tag{3.12}$$

Furthermore, we define the pseudo-inverse matrices $S_k^{-1} = \tilde{s}_k^{-1} \tilde{S}_k$ and $N_j^{-1} = \tilde{n}_j^{-1} \tilde{N}_j$, so that $S_k^{-1} S_k$ and $N_j^{-1} N_j$ are identity operators on the respective eigenspaces.

We assume here that the eigenspaces corresponding to the different eigenvalues are known a priori, e.g. from symmetry considerations. However, the formalism allows for eigenvalues of different eigenspaces becoming equal a posteriori.

Finally, we also need to define some priors for the parameters p_k and η_j. As was done by Enßlin & Frommert (2011) and Enßlin & Weig (2010), we assume each parameter to be a priori independent from all the others and use inverse Gamma distributions, i.e. power laws with exponential cutoff, as priors for the individual parameters,

$$\mathcal{P}(p, \eta) = \mathcal{P}(p)\mathcal{P}(\eta), \tag{3.13}$$

$$\mathcal{P}(p) = \prod_k \frac{1}{q_k \Gamma(\alpha_k - 1)} \left(\frac{p_k}{q_k}\right)^{-\alpha_k} \exp\left(-\frac{q_k}{p_k}\right), \tag{3.14}$$

$$\mathcal{P}(\eta) = \prod_j \frac{1}{r_j \Gamma(\beta_j - 1)} \left(\frac{\eta_j}{r_j}\right)^{-\beta_j} \exp\left(-\frac{r_j}{\eta_j}\right). \tag{3.15}$$

The parameters α_k and β_j determine the steepness of the power law and the parameters q_k and r_j give the position of the cutoff. In the limit $(\alpha_k, \beta_j) \to (1, 1)$ and $(q_k, r_j) \to (0, 0)$, this turns into the so-called Jeffreys prior, which is flat on a logarithmic scale and can therefore be characterized as non-informative.

3.3 Derivation of the filter formulas

With the priors for s, n, p, and η, we can calculate the joint probability of the signal and the data by marginalizing over the parameters p and η,

$$
\begin{aligned}
\mathcal{P}(s, d) &= \int \mathcal{D}p \int \mathcal{D}\eta \; \mathcal{P}(s, d|p, \eta)\mathcal{P}(p, \eta) \\
&= \int \mathcal{D}p \int \mathcal{D}\eta \; \mathcal{P}(d|s, p, \eta)\mathcal{P}(s|p)\mathcal{P}(p, \eta) \\
&= \int \mathcal{D}p \int \mathcal{D}\eta \; \mathcal{G}(d - Rs, N)\mathcal{G}(s, S)\mathcal{P}(p, \eta).
\end{aligned}
\tag{3.16}
$$

Solving the integrals yields

$$
\begin{aligned}
\mathcal{P}(s, d) = \; \prod_k \; & \frac{\Gamma\left(\gamma_k\right) q_k^{\alpha_k - 1}}{\Gamma\left(\alpha_k - 1\right)(2\pi)^{\rho_k/2}} \\
& \left(q_k + \frac{1}{2}s^\dagger S_k^{-1} s\right)^{-\gamma_k} \\
\prod_j \; & \frac{\Gamma\left(\delta_j\right) r_j^{\beta_j - 1}}{\Gamma\left(\beta_j - 1\right)(2\pi)^{\mu_j/2}} \\
& \left(r_j + \frac{1}{2}(d - Rs)^\dagger N_j^{-1}(d - Rs)\right)^{-\delta_j},
\end{aligned}
\tag{3.17}
$$

where $\rho_k = \operatorname{tr}\left(S_k^{-1} S_k\right)$, $\mu_j = \operatorname{tr}\left(N_j^{-1} N_j\right)$, $\gamma_k = \rho_k/2 + \alpha_k - 1$, and $\delta_j = \mu_j/2 + \beta_j - 1$. Note that the posterior is proportional to this joint likelihood for any given dataset.

One could construct the *maximum a posteriori* estimator, however, this was shown by Enßlin & Frommert (2011) to perform poorly due to a perception threshold, i.e. modes with too little power in the data are completely filtered out. A better estimate for the posterior mean of the signal can be constructed using the formalism of minimum Gibbs free energy, derived by Enßlin & Weig (2010), who introduce thermodynamic quantities by identifying the posterior probability density with a canonical density function according to

$$
\mathcal{P}(s|d) = \frac{\mathcal{P}(s, d)}{\mathcal{P}(d)} = \frac{\mathrm{e}^{-T^{-1}H(s,d)}}{Z(d)}.
\tag{3.18}
$$

The Gibbs energy is then

$$
G = U - TS_B,
\tag{3.19}
$$

where $U = \langle H \rangle_{\mathcal{P}(s|d)}$ is the internal energy, $S_B = \langle -\log \mathcal{P}(s|d) \rangle_{\mathcal{P}(s|d)}$ the Boltzmann entropy, and $H = -\log \mathcal{P}(s, d)$ is called the information Hamiltonian. The temperature T serves as a tuning parameter, shifting the weight beteween the internal energy term and the entropy term in the Gibbs free energy.

Approximating the posterior with a Gaussian with mean m and covariance D,

$$
\mathcal{P}(s|d) \approx \mathcal{G}(s - m, D),
\tag{3.20}
$$

gives an approximate internal energy $\tilde{U}$, an approximate entropy $\tilde{S}_B$, and therefore an approximate Gibbs energy

$$
\begin{aligned}
\tilde{G}(m, D) &= \tilde{U}(m, D) - T\tilde{S}_B \\
&= \langle H(s, d) \rangle_{\mathcal{G}(s-m, D)} - \frac{T}{2} \operatorname{tr}\left(1 + \log\left(2\pi D\right)\right).
\end{aligned}
\tag{3.21}
$$

For $T = 1$, this approximate energy is, apart from an additive constant, identical to the non-symmetric Kullback-Leibler distance (Kullback & Leibler 1951) between the full posterior and the Gaussian approximation,

$$
\begin{aligned}
\tilde{G}(m, D) &= \langle H(s, d) + \log\left(\mathcal{G}(s - m, D)\right) \rangle_{\mathcal{G}(s-m, D)} \\
&= \int \mathcal{D}s \, \mathcal{G}(s - m, D) \log\left(\frac{\mathcal{G}(s - m, D)}{\mathcal{P}(s, d)}\right) \\
&= \int \mathcal{D}s \, \mathcal{G}(s - m, D) \log\left(\frac{\mathcal{G}(s - m, D)}{\mathcal{P}(s|d)}\right) \\
&\quad + \log\left(\mathcal{P}(d)\right) \\
&= d_{\mathrm{KL}}\left[\mathcal{G}(s - m, D), \mathcal{P}(s|d)\right] + \log\left(Z(d)\right),
\end{aligned}
\tag{3.22}
$$

as was shown already by Enßlin & Weig (2010).

The approximate internal energy in our case, calculated from the joint probability of Eq. (3.17), is

$$
\begin{aligned}
\tilde{U}(m, D) \cong \sum_k \gamma_k \underbrace{\left\langle \log\left(q_k + \frac{1}{2} s^\dagger S_k^{-1} s\right) \right\rangle_{\mathcal{G}(s-m, D)}}_{=:\mathcal{A}_k} \\
+ \sum_j \delta_j \underbrace{\left\langle \log\left(r_j + \frac{1}{2}\left(d - Rs\right)^\dagger N_j^{-1}\left(d - Rs\right)\right) \right\rangle_{\mathcal{G}(s-m, D)}}_{=:\mathcal{B}_j},
\end{aligned}
\tag{3.23}
$$

where we have dropped terms that are independent of m and D. The logarithms can be expanded in an asymptotic power series, giving

$$
\begin{aligned}
\mathcal{A}_k &= \log\left(\tilde{q}_k\right) \\
&\quad - \sum_{i=1}^{\infty} \frac{(-1)^i}{i\tilde{q}_k^i} \underbrace{\left\langle \left(q_k + \frac{1}{2} s^\dagger S_k^{-1} s - \tilde{q}_k\right)^i \right\rangle_{\mathcal{G}(s-m, D)}}_{=:\tilde{\mathcal{A}}_k i}
\end{aligned}
\tag{3.24}
$$

and

$$\mathcal{B}_j = \log\left(\tilde{r}_j\right) - \sum_{i=1}^{\infty} \frac{(-1)^i}{i\tilde{r}_j^i}$$
$$\underbrace{\left\langle \left(r_j + \frac{1}{2}\left(d - Rs\right)^\dagger N_j^{-1}\left(d - Rs\right) - \tilde{r}_j\right)^i \right\rangle_{\mathcal{G}(s-m,D)}}_{=:\tilde{\mathcal{B}}_j i}, \tag{3.25}$$

where we have chosen the linear dependencies to be captured by

$$\tilde{q}_k = \left\langle q_k + \frac{1}{2}s^\dagger S_k^{-1} s \right\rangle_{\mathcal{G}(s-m,D)}$$
$$= q_k + \frac{1}{2}\text{tr}\left(\left(mm^\dagger + D\right)S_k^{-1}\right) \tag{3.26}$$

and

$$\tilde{r}_j = \left\langle r_j + \frac{1}{2}\left(d - Rs\right)^\dagger N_j^{-1}\left(d - Rs\right) \right\rangle_{\mathcal{G}(s-m,D)}$$
$$= r_j + \frac{1}{2}\text{tr}\left(\left(\left(d - Rm\right)\left(d - Rm\right)^\dagger + RDR^\dagger\right)N_j^{-1}\right), \tag{3.27}$$

respectively.

Here, we restrict ourselves to the zeroth order solution, i.e. we neglect all contributions from $\tilde{\mathcal{A}}$ and $\tilde{\mathcal{B}}$. Furthermore we set $T = 1$. The case with $T \neq 1$ is discussed up to second order in Appendix C.

Now we search for the optimal Gaussian approximation to the posterior by minimizing the approximate Gibbs energy, which is equivalent to minimizing the Kullback-Leibler distance between the two probability densities, according to Eq. (3.22). Taking the functional derivatives of Eq. (3.21) with respect to m and D and equating them to zero yields the equations

$$m = Dj, \tag{3.28}$$

$$j = \sum_j \frac{\delta_j}{\tilde{r}_j} R^\dagger N_j^{-1} d, \tag{3.29}$$

$$D = \left(\sum_k \frac{\gamma_k}{\tilde{q}_k} S_k^{-1} + \sum_j \frac{\delta_j}{\tilde{r}_j} R^\dagger N_j^{-1} R\right)^{-1}. \tag{3.30}$$

By comparing these expressions to the Wiener filter formula, Eq. (3.8), we can read off the parameters $p_k = \frac{\tilde{q}_k}{\gamma_k}$ and $\eta_j = \frac{\tilde{r}_j}{\delta_j}$ for the signal and noise covariance matrix, respectively.

So altogether the equations that need to be solved simultaneously are

$$m = Dj \tag{3.31a}$$

$$p_k = \frac{q_k + \frac{1}{2}\mathrm{tr}\left(\left(mm^\dagger + D\right)S_k^{-1}\right)}{\frac{\rho_k}{2} + \alpha_k - 1} \tag{3.31b}$$

$$\eta_j = \frac{r_j + \frac{1}{2}\mathrm{tr}\left(\left(\left(d - Rm\right)\left(d - Rm\right)^\dagger + RDR^\dagger\right)N_j^{-1}\right)}{\frac{\mu_j}{2} + \beta_j - 1}. \tag{3.31c}$$

Thus, we find both the posterior mean and the posterior covariance for the signal. Note that the first two of these three equations were already found by Enßlin & Frommert (2011) and Enßlin & Weig (2010), who discussed the reconstruction of signals with unknown power spectra. The term *critical filter* was coined by Enßlin & Frommert (2011) to refer to this filter since it belongs to a family of filters lying on a line in the parameter plane used by Enßlin & Frommert (2011) that separates the filters with a perception threshold from those without. The additional uncertainty in the noise covariance that we introduce here simply adds one more equation, leading to an *extended critical filter*.

3.4 Application to simulated signals

Here we demonstrate the performance of our signal reconstruction algorithm under different circumstances.

3.4.1 Setup

We consider two different scenarios. First, we consider a simple one-dimensional test case, where the signal is supposed to be a real field defined over some interval with periodic boundary conditions. We discretize the interval into 2048 pixels. For simplicity, we set the response matrix R to be the identity operator, so that our data set consists of 2048 individual points as well. We further assume statistical homogeneity for the signal field, leading to a covariance matrix that is diagonal in Fourier representation,

$$S_{kk'} = \langle s_k s_{k'}^* \rangle_{\mathcal{P}(s)} = \delta_{kk'} P_s(k), \tag{3.32}$$

with the power spectrum $P_s(k)$ on its diagonal. For this power spectrum we choose a simple power law

$$P_s(k) \propto (1 + k)^{-2} \tag{3.33}$$

and draw a random realization of the signal from it.

Motivated by astrophysical applications, we also consider a real signal field on the sphere,

$$s : \mathcal{S}^2 \longrightarrow \mathbb{R}. \tag{3.34}$$
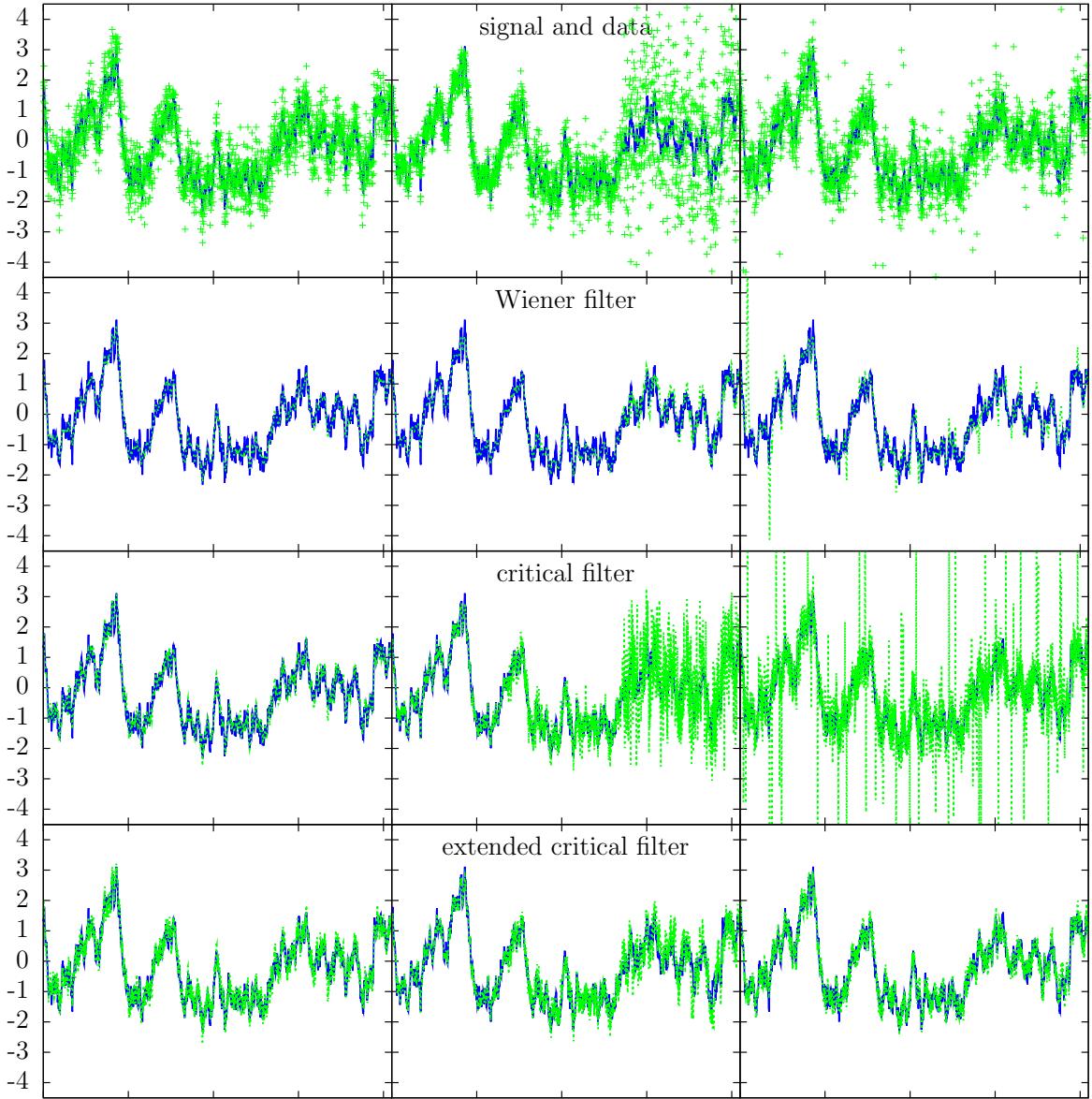
Figure 3.1: Comparison of different filter algorithms in the one-dimensional test case. Each column corresponds to a different setting. The signal, drawn from a power law power spectrum, is the same in each case and depicted in each panel with a solid line. The left column contains homogeneous noise, while in the middle column, the noise is suppressed in the left third of the interval and enhanced in the right third, and in the right column the noise is enhanced in some individual pixels. The first row shows the signal realization along with the data. The second row shows the reconstruction using the Wiener filter formula, assuming the correct power spectrum and under the assumption of homogeneous noise; the third row shows the *critical filter* reconstruction, assuming the power spectrum to be unknown, but still assuming homogeneous noise. The last row, finally, shows the *extended critical filter* reconstruction in which both the signal power spectrum as well as the noise variance are assumed to be unknown. The respective reconstructions are depicted by a dashed line which lies on top of the solid one in many cases. In the two right panels of the first row, some of the data points lie outside the area that is shown.
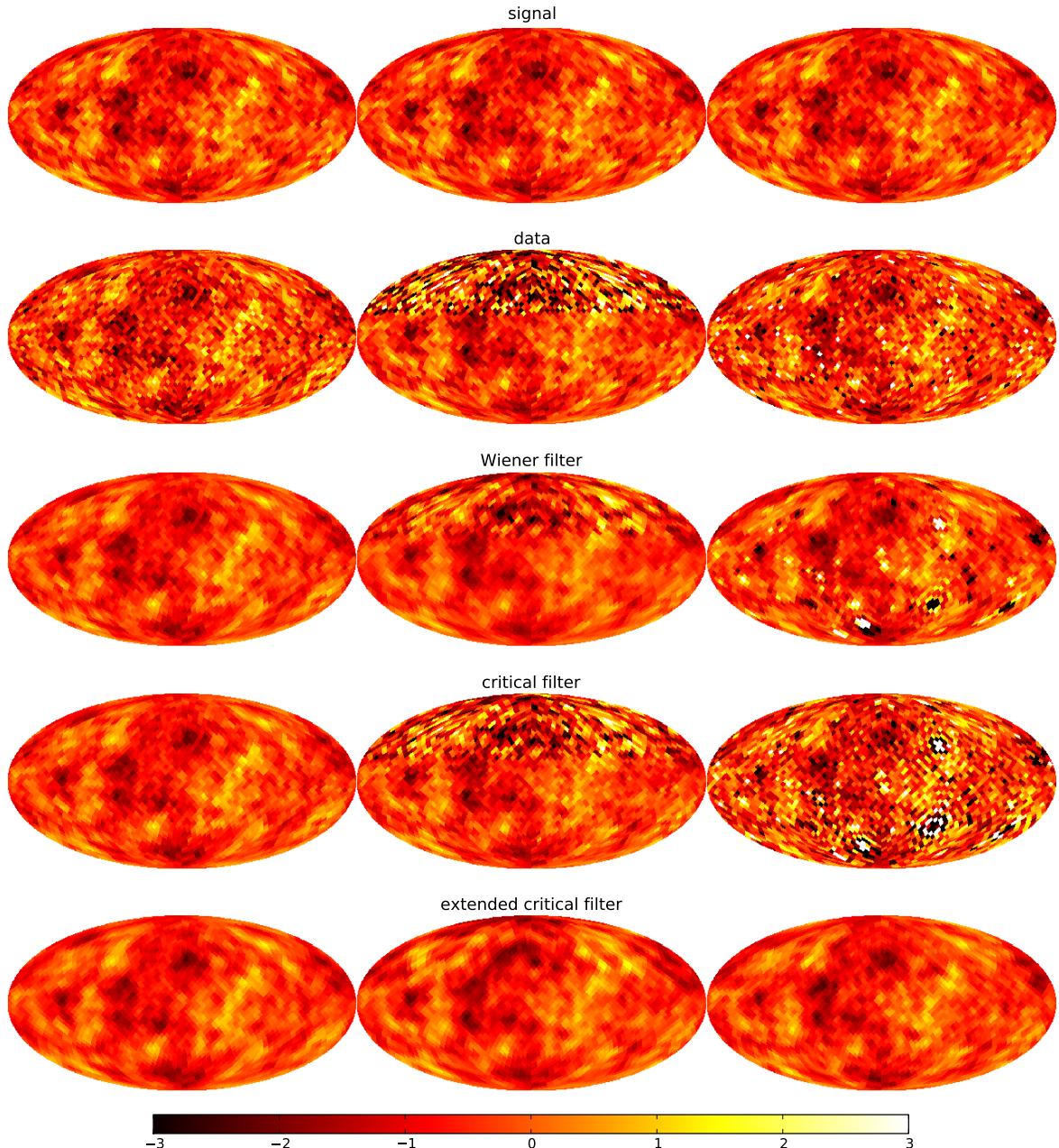
Figure 3.2: Comparison of different filter algorithms in the spherical case. Each column corresponds to a different setting. The signal, drawn from a power law power spectrum, is the same in each case. The left column contains homogeneous noise, while in the middle column, the noise is suppressed in the southern third of the sphere and enhanced in the northern third, and in the right column the noise is enhanced in some individual pixels. The first row shows the signal realization and the second row the data. The third row shows the reconstruction using the Wiener filter formula, assuming the correct power spectrum and under the assumption of homogeneous noise; the fourth row shows the *critical filter* reconstruction, assuming the power spectrum to be unknown, but still assuming homogeneous noise. The last row, finally, shows the *extended critical filter* reconstruction in which both the signal power spectrum as well as the noise variance are assumed to be unknown.

Using again $R = \mathbb{1}$, the data and noise are also fields on the sphere,

$$d, n : \ \mathcal{S}^2 \longrightarrow \mathbb{R}. \tag{3.35}$$

In the numerical implementation, we use the HEALPix[1] discretization scheme at a resolution of $N_{\mathrm{side}} = 16$, which leads to $3\,072$ pixels. Under the assumption of statistical homogeneity and isotropy, the signal covariance matrix in this case becomes diagonal in the basis given by the spherical harmonics components,

$$S_{(\ell m)(\ell' m')} = \langle s_{\ell m} s^*_{\ell' m'} \rangle_{\mathcal{P}(s)} = \delta_{\ell \ell'} \delta_{m m'} C_\ell, \tag{3.36}$$

where C_ℓ are the angular power spectrum components. We draw our signal realization again from a power law spectrum,

$$C_\ell = (1 + \ell)^{-2}. \tag{3.37}$$

We assume the noise to be uncorrelated in the position basis, making the noise covariance matrix diagonal in this basis,

$$N_{\hat{n}\hat{n}'} = \langle n_{\hat{n}} n_{\hat{n}'} \rangle_{\mathcal{P}(n)} = \delta_{\hat{n}\hat{n}'} \sigma^2_{\hat{n}}, \tag{3.38}$$

where $\hat{n}$ and $\hat{n}'$ denote positions on the sphere or on the interval, respectively. Within this framework, we consider three cases for the noise statistics. In the first one, we use homogeneous noise with variance $\sigma^2_{\hat{n}} = 1/4$ independent of $\hat{n}$. For the second case we divide the data space into three zones. In the left/southern third, we suppress the noise variance by a factor of nine and in the right/northern third we enhance it by a factor of nine, while we leave it unchanged in the middle. Finally, in the third case we again assume homogeneous noise with variance $1/4$, but we enhance the variance in five percent of the pixels, randomly selected, by a factor of 100. Both the signal and the three resulting data realizations are shown in Fig. 3.1 for the one-dimensional case and in Fig. 3.2 for the spherical case, along with the results of different reconstructions that we discuss next.

3.4.2 Reconstructions

We first apply the standard Wiener filter formula, the results of which are shown in the second row of Fig. 3.1 for the one-dimensional case and in the middle row of Fig. 3.2 for the spherical case. For this we assume the correct power spectrum to be known, but we assume homogeneous noise with variance $1/4$ in all three cases. In the case where this assumption is correct, the reconstruction is known to be optimal and this is confirmed by visual inspection of the outcome. In the cases with inhomogeneous noise, the Wiener filter fails to completely filter out the noise structures in the data in the regions where the noise is underestimated and therefore reproduces some of them in the reconstruction, as one would expect. This is true for the right (northern) third in the middle column of Fig. 3.1

---

[1]The HEALPix package is available from `http://healpix.jpl.nasa.gov`.

(Fig. 3.2), as well as the noisy pixels in the right column. The opposite should happen in the left (southern) third in the middle column of Fig. 3.1 (Fig. 3.2), where the noise is overestimated. One would expect that structures in the data that are actually due to the signal get filtered out. This is actually happening, although it is barely visible in the resulting plots.

Next we assume that the power spectrum is not known a priori, i.e. we apply the *critical filter*. The resulting plots are shown in the third row of Fig. 3.1 for the one-dimensional case and the fourth row of Fig. 3.2 for the two-dimensional case. In the one-dimensional case we define the S_k operators of Eq. (3.9) to be projections onto bins of width $\Delta_k = 2$ in Fourier space, effectively assuming that the two scales that enter the bin have the same power. This binned power is then represented by the paramer p_k in Eq. (3.9). This binning is necessary in one dimension since each individual Fourier component contains only two degrees of freedom.[2] In the spherical case, we can directly use the angular power spectrum components C_ℓ as parameters and the projection-like operators S_k become actual projections onto the ℓ-th angular scale, which contains $2\ell + 1$ degrees of freedom. In both cases we assume Jeffreys prior for the unknown parameters. Then we simply iterate the first two lines of Eq. (3.31), while keeping the assumption of homogeneous noise with variance $1/4$.

In the cases where our assumptions about the noise are true, the resulting map is very close to the Wiener filter reconstruction, confirming the assessment by Enßlin & Frommert (2011); Enßlin & Weig (2010); Oppermann et al. (2011a) that the *critical filter* can yield a very accurate reconstruction, even if the power spectrum is completely unknown. In the cases where we have made false assumptions about the noise, however, we see the same problems that the Wiener filter reconstruction has, only much stronger pronounced. This is because the reconstructed power spectrum now actually accounts for the features in the data that are due to noise where this is underestimated. With this power spectrum, the map reconstruction tends to favor these features even more than when the correct power spectrum is used. This amplifying effect is again much more prominent where the noise was underestimated than where it was overestimated.

Finally, we account for the possibility that we might have misestimated the noise statistics by applying the full *extended critical filter*, derived in Sect. 3.3. As projection-like matrices N_j we choose projections onto the j-th pixel of the interval and sphere, respectively, multiplied with our original guess for the noise variance in that pixel, $\sigma_j^2 = 1/4$. In this way, the parameters η_j become correction factors for the noise variance of each data point. For the prior parameters we choose $\beta_j = 2$ and we adapt r_j such that $\langle \log (\eta_j) \rangle_{\mathcal{P}(\eta)} = 0$. After iterating the full set of equations (3.31), we obtain the results shown in the bottom rows of Fig. 3.1 and Fig. 3.2. In the case with homogeneous noise, we still get a result that is similar to the Wiener filter one. This shows that we do not lose much by allowing for some uncertainty in the noise covariance. In the cases in which our original noise estimate was wrong, however, we obtain reconstructions of a much higher quality than from the

---

[2]Note that since our signal is real, the Fourier components associated with negative k-values do not contain additional degrees of freedom.

critical filter. Obviously, our algorithm succeeds in uncovering the false error bars in our dataset and correcting them. This works especially well in the case where only individual pixels have underestimated noise variance. This setting makes it especially easy for the algorithm to infer the signal statistics from all the other pixels and find the pixels in which the data points and the signal are inconsistent with one another. However, even in the case where one third of the space is covered with underestimated noise, our algorithm still does a good job in reconstructing the original signal. In the spherical case, the *extended critical filter* performs even better than the Wiener filter. This is also true for the one-dimensional case in the scenario where the noise in individual pixels is enhanced. In the scenario with enhanced and suppressed noise in one third of the one-dimensional interval, the Wiener filter performs better than the *extended critical filter*. It should be noted, however, that using the Wiener filter is not an option if the power spectrum of the signal is not known a priori.

Some further insight can be gained by looking at the reconstructed angular power spectra for the spherical case. These are shown in Fig. 3.3. In the case with homogeneous noise, the *critical filter* recovers the true power spectrum almost perfectly, while the *extended critical filter* misses some power on the smallest scales, i.e. some of the small-scale power in the data is falsely attributed to noise and therefore not represented in the signal power spectrum. This effect is small, however, and does not greatly influence the resulting map.

In the case in which the noise is highly inhomogeneous, being higher and lower in one third of the data space each, the *extended critical filter* misses quite a lot of power on small scales. This results in the slightly oversmoothed map seen in Fig. 3.2. The *critical filter*, however, that operates under wrong assumptions for the noise statistics, overestimates the power on small scales significantly. This is in agreement with the very noisy reconstructed map.

It is in the third case, in which the noise is greatly enhanced in individual pixels, that the *extended critical filter* shows its full strength. While the *critical filter* attributes the power in the faulty pixels to the signal and therefore overestimates the signal power by orders of magnitude, the extended version accounts for the misestimated error bars and does not account for these pixels in the signal power spectrum. While the result is a power spectrum that is slightly underestimated on the smallest scales, this is again only a comparatively small error, still leading to a good reconstruction.

These findings are confirmed by Fig. 3.4, which shows the differences of the nine reconstructions and the signal realization. Our extension of the *critical filter* clearly brings the strongest improvement in the case where the noise is enhanced in individual pixels, while also lowering the error in the case with an extended region of underestimated noise. The same can directly be seen for the one-dimensional case in Fig. 3.1.

Finally, we plot the standard deviation per pixel of the Gaussian approximation (3.20) to the posterior probability distribution, i.e. the square root of the diagonal of the covariance matrix in the pixel basis, $\sigma = \sqrt{\mathrm{diag}(D)}$, in Fig. 3.5 for the *extended critical filter*. This can be interpreted as an estimate for the 1σ-error bar of the reconstructed maps. The region with enhanced noise in the second scenario is clearly marked out by a
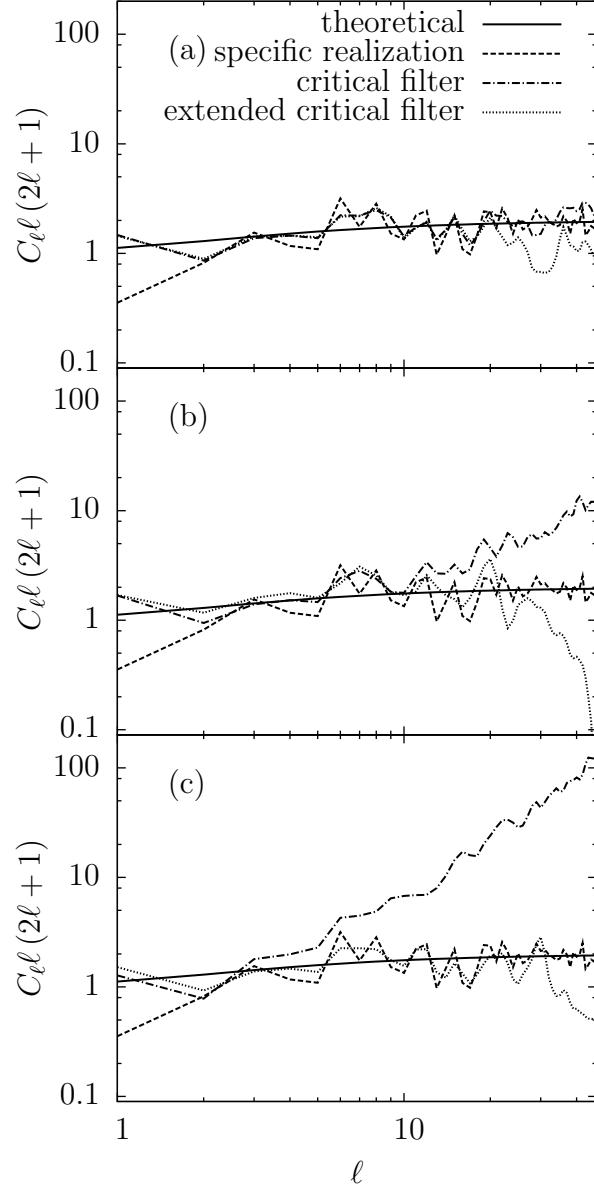
Figure 3.3: Comparison of the different reconstructed angular power spectra for the spherical scenario. The solid line depicts the theoretical power spectrum which is also used in the Wiener filter reconstructions. The dashed line corresponds to the power of the specific signal realization and the dash-dotted and dotted lines to the power spectrum reconstructed with the *critical filter* and the *extended critical filter*, respectively. Panel (a) shows the case with homogeneous noise, panel (b) the one in which the noise is enhanced and suppressed in one third of the sphere each, and panel (c) the one in which the noise is enhanced in individual pixels.
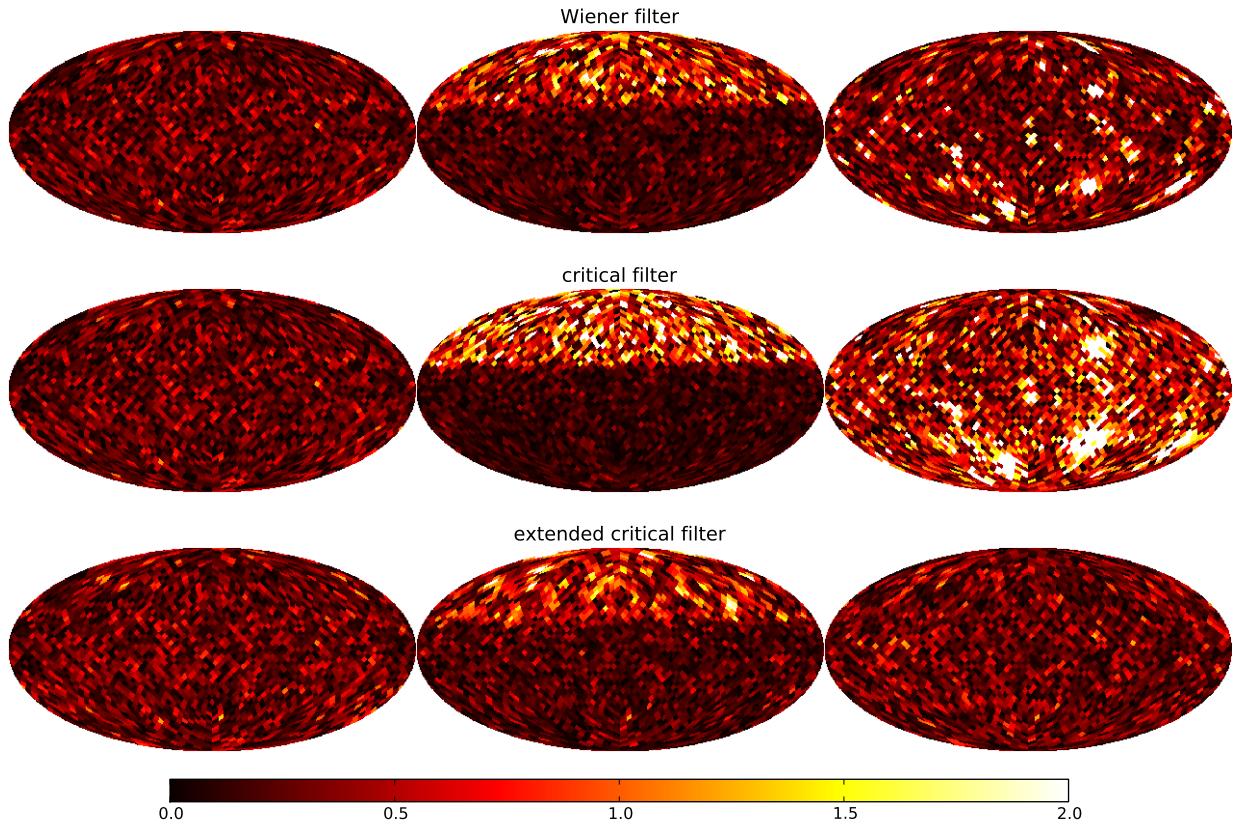
Figure 3.4: Absolute value of the pixelwise difference between the reconstructed maps and the signal realization for the spherical scenario. Each row shows the results for a different filter algorithm. As in Fig. 3.2, the left column shows the case with homogeneous noise, the middle column the one with enhanced noise in the northern third of the sphere and suppressed noise in the southern third, and the right column shows the case where the noise is enhanced in individual pixels. Note that the color bar differs from the one used in Fig. 3.2.
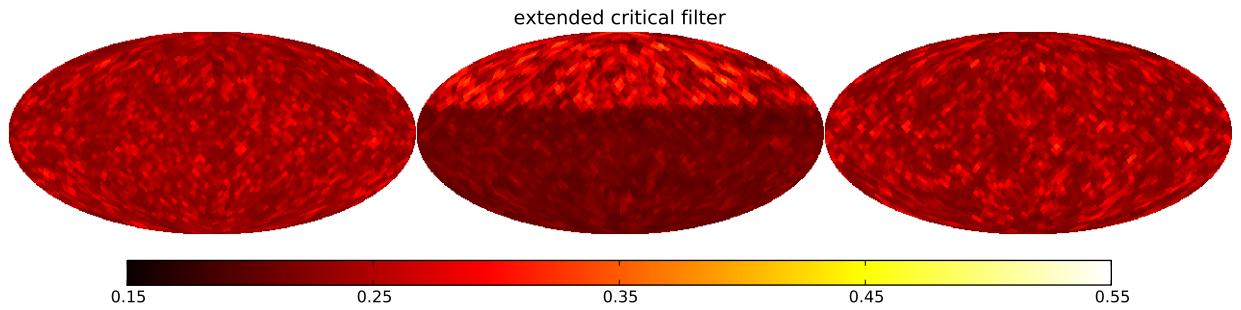
Figure 3.5: Pixelwise uncertainty of the *extended critical filter* reconstructions in the spherical scenario. The left panel shows the case with homogeneous noise, the middle panel the case with enhanced noise in the northern third and suppressed noise in the southern third, and the right panel the case with enhanced noise in individual pixels.
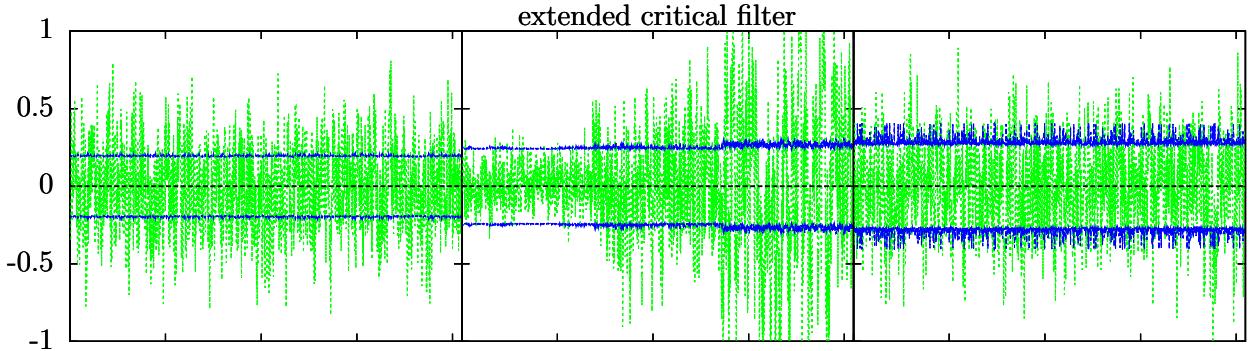
Figure 3.6: Pixelwise uncertainty of the *extended critical filter* reconstructions in the one-dimensional scenario. The left panel shows the case with homogeneous noise, the middle panel the case with enhanced noise in the right third and suppressed noise in the left third, and the right panel the case with enhanced noise in individual pixels. The dark curves represent $\pm\sigma$ and the light curve the difference between the reconstruction and the correct signal.

higher uncertainty of the reconstruction due to the corrected entries of the noise covariance matrix entering the information propagator D. Note, however, that the full posterior is non-Gaussian and the 1σ range can therefore not necessarily be interpreted as a 68% confidence interval, especially since we are using a zeroth order approximation to calculate the Gaussian approximation. In fact, in our spherical example only about 50% of the pixels of the reconstructions lie within 1σ of the correct signal in all three noise scenarios.

Fig. 3.6 extends this study to the one-dimensional case. We plot the difference of the signal and the reconstruction result of the *extended critical filter*, along with lines depicting $\pm\sigma$, for the three different noise settings. In the case where the noise variance is constant within each third of the interval, the one-sigma curve exhibits a step-like behavior at the boundaries of the thirds, although this effect is relatively small. In the case of homogeneous noise, the one-sigma curve is roughly constant while the individual noisy pixels in the last scenario are reflected in the one-sigma curve by its large variations from pixel to pixel. The fraction of the pixels for which the reconstruction lies within 1σ of the correct signal in the one-dimensional case is 50% in the case with homogeneous noise, 63% in the case with enhanced and suppressed noise in a third of the interval each, and 69% in the case with enhanced noise in individual pixels.

3.5 Discussion

Using the formalism of minimum Gibbs free energy we have extended the *critical filter* algorithm, developed by Enßlin & Frommert (2011) and Enßlin & Weig (2010), to an algorithm that allows for uncertainties both in the signal covariance and in the noise covariance. We have demonstrated the performance of our algorithm, Eqs. (3.31), by applying it to a set of mock observations on the sphere, as well as in a simple one-dimensional test case.

These applications have shown that the *extended critical filter* performs outstandingly if only a few individual data points have a misestimated error bar. However, even in a case where large portions of the data are affected, the algorithm was shown to perform inarguably better than the *critical filter*, using a fixed – and faulty – assumption about the noise statistics. We have also compared the results to those obtained from a Wiener filter reconstruction, using the correct power spectrum, which is known to be optimal if the assumptions about the noise statistics are correct. This filter was demonstrated, however, to lead to reconstructed maps that are much further from the true signal than the results of the *extended critical filter* in some cases where the assumptions are not correct.

The choice of the two-sphere as the space on which our signal is defined was motivated by astrophysical applications, where we could think of the signal as an all-sky field or some quantity defined on the surface of a star or a planet. Applications in other fields of physics are abundant. However, it should be noted that there is nothing special about the sphere. We could equally well have chosen a more-dimensional euclidean space, using the power spectrum defined in Fourier space instead of the angular power spectrum, as we have done in the one-dimensional scenario.

Furthermore, our choice of the identity operator as response matrix was made only on account of simplicity. It allowed us to represent the data in the same fashion as the signal. It should be clear, however, that the derived filter formulas, Eqs. (3.31), are valid for any response matrix, even a singular one. Applications of the *critical filter* with non-trivial response matrices were presented by Enßlin & Frommert (2011) and Oppermann et al. (2011a) and such a response would not pose a problem for the extended version of the filter.

The problem of signal reconstruction with some uncertainty in the noise variance is certainly one of general interest. There are several ways in which uncertainty in the noise variance might arise. It may be due to questionable assumptions that enter in the calculation of the error bars of the data. Another possibility is that it arises from the definition of the signal itself. The quantity of interest may only be part of what has been measured in the first place in which case the rest of the data would be noise with essentially unknown variance. All these factors come together in the reconstruction problem considered by Oppermann et al. (2011a). An extension of that reconstruction, using additional sets of data and the improved algorithm presented here, is planned (Oppermann et al. 2012a).

It should be noted, however, that even with the *extended critical filter*, some knowledge about either the parameters in the signal covariance or the ones in the noise covariance is needed to arrive at a sensible reconstruction. Leaving them both completely free would lead to a degeneracy between signal and noise that cannot be resolved. Only by assigning an informative prior to at least one of the two sets of parameters is this degeneracy broken. Furthermore, the functional bases in which the signal and noise covariances are diagonal, i.e. their eingenspaces, need to differ to allow for a separation of noise from signal.

Chapter 4

An improved map of the Galactic Faraday sky

Note: This chapter has been published in Astronomy & Astrophysics *(Oppermann et al. 2012a). Parts of Sect. 4.4.1 have been written by Jeroen M. Stil.*

4.1 Introduction

Magnetic fields are ubiquitous in the interstellar medium. They are likely to play a major dynamical role in the evolution of galaxies. It is by comparing theoretical predictions and simulations to observations of galactic magnetic fields that the generation and dynamical role of these magnetic fields can be understood (see e.g. Beck 2011, and references therein). It is natural to look first and foremost at our own galaxy, the Milky Way, and try to study its magnetic field. However, its observation is complicated by several effects. The magnetic field is a three-dimensional vector field that varies on multiple scales throughout the Galaxy. Thus, very many measurements of the field would be needed to determine even its large-scale properties. Furthermore, virtually any observation suffers from a projection effect because local effects add up along the line of sight. And finally the magnetic field cannot be measured directly, so that related observables have to be used. These observables, however, are not only sensitive to the magnetic field itself but also to other quantities that are not necessarily better understood, introducing ambiguities when inferring properties of the magnetic field. The intensity of synchrotron radiation is sensitive to the strength of the magnetic field component orthogonal to the line of sight, but it is modulated by the density of cosmic ray electrons (e.g. Ginzburg & Syrovatskii 1965). The direction of this magnetic field component can be studied via the polarization direction of synchrotron radiation and thermal dust emission (e.g. Gardner & Whiteoak 1966; Lazarian 2003). A magnetic field component along the line of sight, on the other hand, gives rise to the effect of Faraday rotation (e.g. Nicholson 1983; Gardner & Whiteoak 1966; Burn 1966). The strength of this effect is influenced not only by the magnetic field but also by the density of thermal electrons. Furthermore, when observing this effect for extragalactic sources, it

contains contributions not only from the Galaxy, but from every position along the line of sight to the source with a non-vanishing magnetic field and thermal electron density.

To find an unambiguous terminology that can capture these subtleties, we introduce the concept of Faraday depth, which depends on position and is independent of any astrophysical source. The Faraday depth corresponds to a position at a distance r_0 from an observer and is given by a line-of-sight integral,

$$\phi(r_0) = \frac{e^3}{2\pi m_e^2 c^4} \int_{r_0}^{0} \mathrm{d}r\ n_{\mathrm{e}}(r)B_r(r), \tag{4.1}$$

over the thermal electron density n_{e} and the line-of-sight component of the magnetic field B_r. Here, e and m_e are the electron charge and mass and c is the speed of light. The Galactic Faraday depth is therefore exactly this integral, where the lower boundary is the outer edge of the Milky Way. It is this integral that contains the information on the Galactic magnetic field.

The observational consequence of Faraday rotation on a single linearly polarized source is a rotation of its plane of polarization about an angle that is proportional to the square of the wavelength. The proportionality constant is equal to the source's Faraday depth, i.e the above integral expression, where the lower boundary is now the source's position. Often, the assumption that the observed polarized radiation stems from a single source is made implicitly and a linear fit to the position angle of the plane of polarization as a function of the squared wavelength is made. We refer to the slope of such a λ^2-fit as rotation measure (RM). For a single source this is the same as the source's Faraday depth. However, the polarized radiation will in general be emitted over a range of physical distances and also over a range of Faraday depths, and the position angle will no longer vary linearly with λ^2. This emission spectrum in Faraday space can be recovered using the technique of RM synthesis (Burn 1966; Brentjens & de Bruyn 2005). In this work we create a map of the Galactic Faraday depth using both data that are based on RM synthesis and data that are based on linear λ^2-fits. Neither measures the Galactic Faraday depth exclusively and we use the term Faraday rotation data when referring to data values without specifying whether they are rotation measures or the result of a synthesis study.

A review of early work on the inference of the regular component features of the Galactic magnetic field from RM measurements is included in the work of Frick et al. (2001). Several studies were conducted by Morris & Berge (1964), Gardner et al. (1969), Vallée & Kronberg (1973), Ruzmaikin & Sokolov (1977), Ruzmaikin et al. (1978), Simard-Normandin & Kronberg (1979), Andreasian (1980), Andreasian (1982), Inoue & Tabara (1981), Sofue & Fujimoto (1983), Vallée (1983), Agafonov et al. (1988), Clegg et al. (1992), Han & Qiao (1994), Han et al. (1997), as well as Rand & Kulkarni (1989), Rand & Lyne (1994), who used RM data of pulsars, and Seymour (1966, 1984), who used spherical harmonics to obtain an all-sky RM map. Among the more recent studies aiming to constrain the Galactic magnetic field with rotation measures of extragalactic radio sources are those by Brown & Taylor (2001), Mao et al. (2010), Kronberg & Newton-McGee (2011), Pshirkov et al. (2011), and Brown et al. (2003b, 2007), Nota & Katgert (2010), and Van Eck et al.

(2011), who supplemented extragalactic RMs with pulsar rotation measures. Weisberg et al. (2004), Vallée (2005), Vallée (2008), Han et al. (2006), and Men et al. (2008) relied entirely on pulsar rotation measures for estimating the Galactic magnetic field, while Sun et al. (2008), Jansson et al. (2009), and Jaffe et al. (2010) used rotation measures of extragalactic sources in combination with synchrotron polarization and intensity data.

Recent attempts to create an all-sky map of Faraday rotation measure were made by Frick et al. (2001), Johnston-Hollitt et al. (2004), Dineen & Coles (2005), and Xu et al. (2006). However, owing to the limited number of data points available at the time, their reconstructions are limited to the largest-scale features. A fairly sophisticated attempt was made by Short et al. (2007), who used Monte Carlo Markov chain methods and accounted for uncertainty in the noise covariance while avoiding the direct involvement of covariance matrices. Realistic attempts to create all-sky maps including smaller-scale features have been possible only since Taylor et al. (2009) published the NRAO VLA Sky Survey (NVSS) (Condon et al. 1998) rotation measure catalog that contains data on sources distributed roughly equally over the sky at declinations higher than $-40°$. One such attempt was made in the same publication where the data were simply smoothed to cover the celestial sphere in regions where the data were taken. Another attempt has been made by Oppermann et al. (2011a), using a more sophisticated signal reconstruction algorithm that takes into account spatial correlations without oversmoothing any maxima or minima.

The NVSS rotation measure catalog is, however, suboptimal in two respects. It lacks data in a large region in the southern sky below the declination of $-40°$ because of the position of the observing telescope (VLA) and its rotation measure values were deduced using only two nearby frequency channels (see Table 4.1). This increases the risk of introducing offsets of integer multiples of π in the rotation angle, as discussed by Sunstrum et al. (2010), and makes it impossible to detect any deviations from a proportionality to λ^2 in the polarization angle. Consequently, sources with a non-trivial Faraday spectrum could not be identified and were assigned a possibly misleading RM value.

In this work we aim to create a map of the Galactic Faraday depth that summarizes the current state of knowledge. To this end we combine the NVSS rotation measure catalog of Taylor et al. (2009) with several other catalogs of Faraday rotation data of polarized extragalactic radio sources, increasing the spatial coverage and additionally constraining the signal also in regions where several data sets overlap. We improve on the map of Oppermann et al. (2011a) by using this more extensive data set and by using an extended version of the reconstruction algorithm, which takes into account uncertainties in the noise covariance, presented by Oppermann et al. (2011b). The resulting all-sky map of the Galactic Faraday depth will be useful in many respects. On the one hand, all-sky information can help in bringing forth global features of the underlying physics, such as the Galactic magnetic field or the electron distribution. On the other hand, an all-sky map can also be useful when studying local or extragalactic features. It could, for example, serve as a look-up table for Galactic contributions to the Faraday depth when studying extragalactic objects.

The remainder of this paper is organized as follows: In Sect. 4.2 we briefly review the main features of the *extended critical filter* algorithm that we used in our map-making

procedure and discuss how it is applied to the situation at hand. The data sets entering the reconstruction are listed in Sect. 4.3 and the results are presented in Sect. 4.4. In the results section, we also include a brief discussion of the reconstructed angular power spectrum. We summarize our findings in Sect. 4.5.

4.2 Reconstruction algorithm

To reconstruct the Galactic Faraday depth from the point source measurements, we used the *extended critical filter* formalism that was presented by Oppermann et al. (2011b). This filter is based on the *critical filter* that was used for the reconstruction by Oppermann et al. (2011a) and derived by Enßlin & Frommert (2011) and Enßlin & Weig (2010) within the framework of the *information field theory* developed by Enßlin et al. (2009).

4.2.1 Signal model

The signal model we used is the same as the one used by Oppermann et al. (2011a). We briefly review the essentials.

In the inference formalism we employ, it is assumed that a linear relationship, subject to additive noise, exists between the observed data d and the signal field s that we try to reconstruct, i.e.

$$d = Rs + n. \tag{4.2}$$

Here, the response operator R describes the linear dependence of the data on the signal. Formally, the signal could be a continuous field, e.g. some field like the Galactic Faraday depth on the celestial sphere. In practice, however, the best we can hope for is to reconstruct a discretized version of such a field, i.e. a pixelized sky-map. In this case, one can think of the signal field s on the sphere as a vector of dimension N_{pixels}, each component of which corresponds to one pixel, and the whole set of data points d as another vector of dimension N_{data}. The response operator then becomes a matrix of dimension $N_{\text{data}} \times N_{\text{pixels}}$ and n is another vector of dimension N_{data} that contains the noise contributions to each data point. Next, we specify the definitions of the signal field and the response matrix for our specific application.

The *critical filter* algorithm, as well as the *extended critical filter*, is intended to statistically reconstruct isotropic and homogeneous random signal fields. We briefly recapture the meaning of this.

It is assumed in the derivation of the filter formulas (see Oppermann et al. 2011b, for details) that the signal field that describes nature is one realization of infinitely many possible ones. Furthermore, it is assumed that some of these possibilities are a priori more likely to be realized in nature than others, i.e. a prior probability distribution function on the space of all possible signal realizations is defined. We assume this probability distribution to ba a multivariate Gaussian with an autocorrelation function $S(\hat{n}, \hat{n}')$. Here, $\hat{n}$ and $\hat{n}'$ denote two positions on the celestial sphere. Now, assuming statistical homogeneity and isotropy means assuming that $S(\hat{n}, \hat{n}')$ depends only on the angle between the two positions $\hat{n}$ and

$\hat{n}'$. This means that the correlation of the value of the signal field at one position with another one at a certain distance depends only on this distance, not on the position on the sphere (homogeneity) and not on the direction of their separation (isotropy). Note, however, that we are making this assumption only for the prior probability distribution, i.e. the inherent probability for signal realizations. The data can (and do) break this symmetry, making the posterior probability distribution, i.e. the probability for a signal realization given the measured data, anisotropic. Furthermore, any single realization of a signal with isotropic statistics can appear arbitrarily anisotropic. Extremely anisotropic realizations will, however, be a priori more unlikely than others.

For this reason we divide out the most obvious largest-scale anisotropy introduced by the presence of the Galactic disk. We do this by defining our signal as

$$s(l, b) = \frac{\phi(l, b)}{p(b)}, \tag{4.3}$$

i.e. the dimensionless ratio of the Galactic Faraday depth ϕ and a variance profile p that is a function of Galactic latitude only. We use this simplistic ansatz for the Galactic variance profile to account for the largest-scale anisotropies without using any specific Galactic model in the analysis.

The profile function is calculated in a multi-step procedure. In the first step, we sort the data points into bins of Galactic latitude and calculate the root-mean-square value for the Faraday rotation data of each bin, disregarding any information on Galactic longitude of the data points. We then smooth these values with a kernel with $10°$ FWHM[1] to form an initial profile function $\tilde{p}$. In the second step, we reconstruct the signal field, resulting in a map $\tilde{m}$ and the corresponding 1σ uncertainty map $\hat{\tilde{D}}^{1/2}$. We use these to calculate the corresponding posterior mean of the squared Faraday depth according to

$$\left\langle \phi^2 \right\rangle_{\mathcal{P}(s|d)} = \tilde{p}^2 \tilde{m}^2 + \tilde{p}^2 \hat{\tilde{D}}. \tag{4.4}$$

The posterior mean is the ensemble average over all possible signal configurations weighted with their posterior probability distribution $\mathcal{P}(s|d)$, i.e. their probability given the measured data, and is denoted by $\langle \cdot \rangle_{\mathcal{P}(s|d)}$. From this expected map of the squared Faraday depth, we then calculate a new variance profile p, now using the pixel values of the map instead of the data points. A few data points were added before repeating this final step yet another time. The final reconstruction is then conducted with the resulting profile function. The initial variance profile and the one used in the final reconstruction are shown in Fig. 4.1. The drop-off toward the Galactic poles of the first-guess profile function is less pronounced since the relatively high noise component of the Faraday rotation data in these regions enters in the root-mean-square that is calculated from the data points. The variance profile as calculated from the final results is also shown in Fig. 4.1.

---

[1]Oppermann et al. (2011a) experimented with different smoothing lengths and found that a factor two difference does not matter for the end result. We chose $10°$ by visual inspection of the smoothness of the resulting profile.
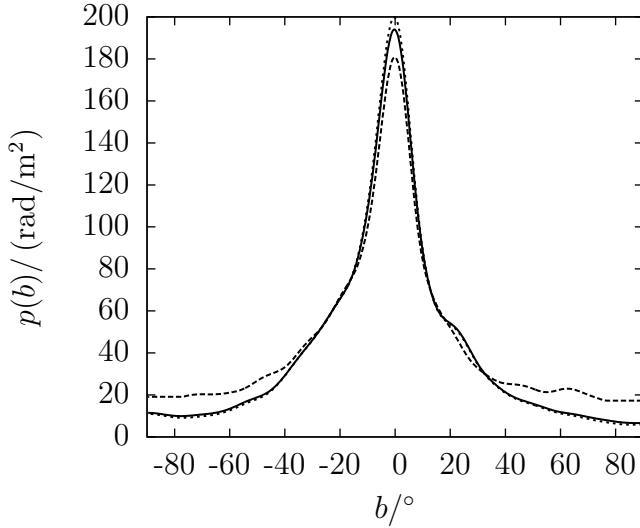
Figure 4.1: Root-mean-square Galactic profile that is used in the definition of the signal field and is part of the response matrix as a function of Galactic latitude. The dashed curve represents the initial profile function $\tilde{p}$ and the solid curve the one used in the final reconstruction, p. The dotted curve shows the profile as calculated from the final results.

After introducing the Galactic variance profile, we can now specify the response operator. In our application, the response matrix R needs to contain both the multiplication of the signal field with this profile function and the probing of the resulting Faraday depth in the directions of the point sources. It is a matrix of dimension $N_{\mathrm{data}} \times N_{\mathrm{pixels}}$. Each row corresponds to one data point and each column to one pixel of the sky map. Here, the row corresponding to the i-th data point contains a non-zero entry only in the column corresponding to the pixel in which the i-th observed extragalactic source lies, modeling the probing of the Faraday depth in the observed directions. This entry is the value of the Galactic variance profile p at the latitude of the pixel, effectively rescaling the local signal field value into a Faraday depth.

Furthermore, we assume Gaussian priors both for the signal and for the noise with covariance matrices S and N, respectively. Since our signal field is assumed to be statistically homogeneous and isotropic, its covariance matrix S is completely determined by its angular power spectrum[2] $(C_\ell)_\ell$, $\ell = 0, 1, \ldots, \ell_{\mathrm{max}}$. The minimum length scale ℓ_{max} is determined by the finite resolution of the discretization. Assuming uncorrelated noise for all data points, the noise covariance N becomes diagonal. The diagonal entries are given by the variance calculated from the error bars given in the data catalogs, modified to account

---

[2]The angular power spectrum is defined by $C_\ell = \langle s_{\ell m} s_{\ell m}^* \rangle_{\mathcal{P}(s)}$, where $s_{\ell m}$ denotes the signal's spherical harmonic component of a certain azimuthal quantum number ℓ and an arbitrary magnetic quantum number m, the asterisk denotes complex conjugation, and the angular brackets denote an ensemble average weighted with the prior probability distribution.

for the expected average extragalactic contribution,

$$\sigma^2 = \sigma^2_{\text{(measurement)}} + \sigma^2_{\text{(extragalactic)}}. \tag{4.5}$$

We include a multiplicative correction factor η that will be determined during the reconstruction, making the diagonal entry of N corresponding to the i-th data point

$$N_{ii} = \eta_i \sigma_i^2. \tag{4.6}$$

As the extragalactic contribution, we use the value $\sigma_{\text{(extragalactic)}} = 6.6 \text{ rad/m}^2$, motivated by the study of Schnitzeler (2010).

Reasons for a deviation of η from unity could be a general underestimation of the measurement error, as was discussed for the NVSS catalog by Stil et al. (2011), a misestimation of the extragalactic contribution, a multi-component Faraday depth spectrum, but also an offset of an integer multiple of π in the rotation angle.

4.2.2 The *extended critical filter*

The *extended critical filter* (see Oppermann et al. 2011b) is a method to simultaneously reconstruct the signal, its covariance, given here by its angular power spectrum $(C_\ell)_\ell$, and the noise covariance, given here by the correction factors $(\eta_i)_i$. To this end, inverse Gamma distributions are assumed as priors for the parameters of the covariances, i.e.

$$\mathcal{P}(C_\ell) = \frac{1}{q_\ell \Gamma(\alpha_\ell - 1)} \left(\frac{C_\ell}{q_\ell} \right)^{-\alpha_\ell} \exp \left(-\frac{q_\ell}{C_\ell} \right) \tag{4.7}$$

and

$$\mathcal{P}(\eta_i) = \frac{1}{r_i \Gamma(\beta_i - 1)} \left(\frac{\eta_i}{r_i} \right)^{-\beta_i} \exp \left(-\frac{r_i}{\eta_i} \right), \tag{4.8}$$

and all these parameters are assumed to be independent. We choose $\alpha_\ell = 1$ for the parameter describing the slope of the power law and $q_\ell = 0$ for the parameter giving the location of the exponential low-amplitude cutoff, turning the prior for each C_ℓ into Jeffreys prior, which is flat on a logarithmic scale, enforcing the fact that we have no a priori information on the power spectrum. For the prior of the correction factors we choose the parameter $\beta_i = 2$, since we already have information on the expected noise covariance from the data catalogs. We adapt the value of r_i such that the a priori expectation value of $\log \eta$ becomes 0, thereby conforming with the catalogs.

With these values, the actual filtering process consists of iterating the three equations[3]

$$m = DR^\dagger N^{-1} d, \tag{4.9}$$

$$C_\ell = \frac{1}{2\ell + 1} \text{tr} \left(\left(mm^\dagger + D \right) S_\ell^{-1} \right), \tag{4.10}$$

---

[3]This is the first-order version of the *extended critical filter*. See Oppermann et al. (2011b) for details.

and

$$\eta_i = \frac{1}{2\beta_i - 1} \left[2r_i + \frac{1}{\sigma_i^2} \left((d - Rm)_i^2 + \left(RDR^\dagger \right)_{ii} \right) \right] \tag{4.11}$$

until convergence is reached. Here, m is the reconstructed signal map, the †-symbol denotes a transposed quantity, and $D = \left(S^{-1} + R^\dagger N^{-1} R \right)^{-1}$ is the so-called information propagator (Enßlin et al. 2009). The matrix S_ℓ^{-1} projects a signal vector onto the ℓ-th length-scale by keeping only the degrees of freedom represented by spherical harmonics components with the appropriate azimuthal quantum number. Although we have chosen $\beta_i = 2$ for our reconstruction, we leave the parameter unspecified in these equations, since we later compare our results to those obtained with $\beta \neq 2$ (see Sect. 4.4.2).

The three equations can be qualitatively explained. Eq. (4.9) links the reconstructed map to the data. It consists of a response over noise weighting of the data and an application of the information propagator to the result. The information propagator combines knowledge about the observational procedure encoded in the response matrix R and the noise covariance matrix N with information on the signal's correlation structure contained in the signal covariance matrix S. It is used in Eq. (4.9) to reconstruct the map at a given location by weighting the contributions of all data points using this information. The information propagator is also (approximately) the covariance matrix of the posterior probability distribution. Therefore, it can be used to obtain a measure for the uncertainty of the map estimate. The 1σ uncertainty of the map estimate in the j-th pixel is given by $\hat{D}_j^{1/2} = D_{jj}^{1/2}$. Eq. (4.10) estimates the angular power spectrum from two contributions. The first term in the trace gives the power contained within a reconstructed map, while the second term compensates for the power lost in the filtering procedure generating this map. This second contribution is not contained in the map calculated via Eq. (4.9) since the data are not informative enough to determine the locations of all features. In a very similar fashion, Eq. (4.11) estimates the correction factors for the error bars also from two main contributions. The first contribution simply uses the difference between the observed data and the data expected from the reconstructed map and the second contribution compensates partly for the attraction the data exhibit onto the map in the reconstruction step, which lets some fraction of the noise imprint itself onto the map. Both contributions are rescaled by the inverse noise variance to turn this estimate of the noise variance into a correction factor. There is a third term in Eq. (4.11) that is solely due to the prior we chose for η. It prevents the error bars from vanishing if a data point by chance perfectly agrees with the map. For a detailed derivation of these formulas, the reader is referred to Oppermann et al. (2011b).

We include a smoothing step for the angular power spectrum in each step of the iteration, where we smooth with a kernel with $\Delta_\ell = 8$ FWHM, lowering Δ_ℓ for the lowest ℓ-modes. This is done to avoid a possible perception threshold on scales with little power in the data (see Enßlin & Frommert 2011). The smoothing step is also justified because none of the underlying physical fields, i.e. the thermal electron density and the line of sight component of the magnetic field, are expected to have vastly different power on neighboring scales.
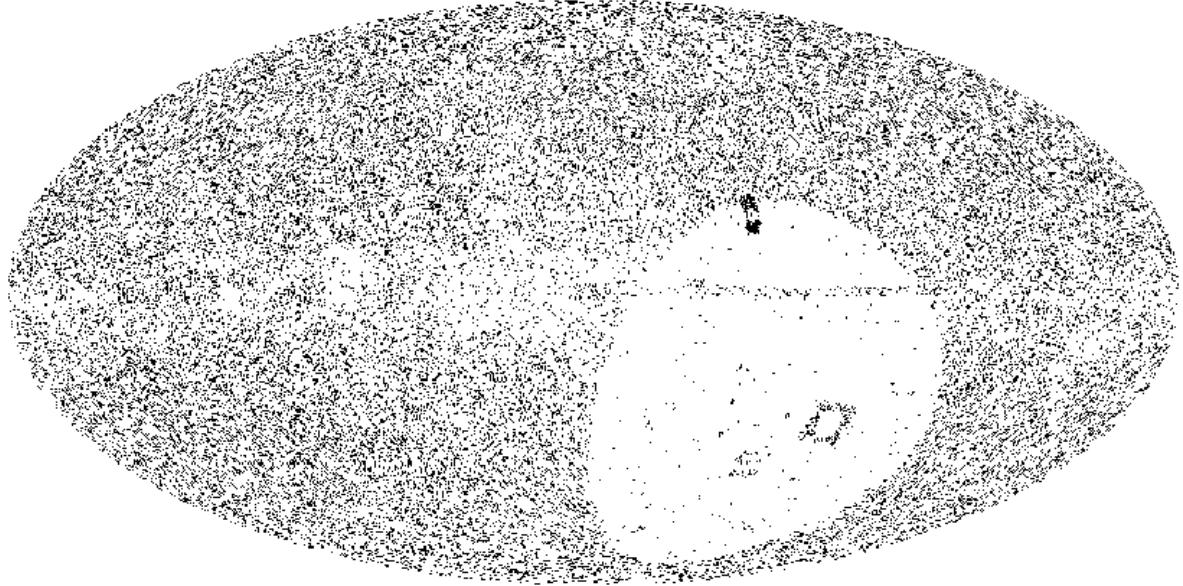
Figure 4.2: Distribution of the data points on the sky. Shown is a HEALPix map at a resolution of $N_{\text{side}} = 128$, using Galactic coordinates. The map is centered on the Galactic center, latitudes increase upward, and longitudes increase to the left. Each black pixel contains at least one data point.

4.3   Data sets

Tables 4.1 and 4.2 summarizes the data catalogs that we used for the reconstruction. Altogether, the catalogs contain 41 330 measurements of the Faraday rotation of extragalactic point sources. Fig. 4.2 shows their distribution on the sky. The coverage is clearly far from complete, especially at declinations below $-40°$ where the Taylor-catalog does not provide any data. However, 24% of the data points from the other catalogs lie within this region, so some toeholds are present even there. The densely sampled region that stands out at the top of the empty patch in Fig. 4.2 is Centaurus A, studied in the Feain-catalog. The relative scarcity of data points near the Galactic plane is due to numerous depolarization effects caused by nearby structures in the magneto-ionic medium, as explained by Stil & Taylor (2007). We used only extragalactic sources, and not pulsar rotation measures, because this ensures that each measurement contains the full Galactic Faraday depth.

Since the regions of coverage of the different catalogs overlap, some of the data points have the same underlying radio source. While this does not constitute a problem for the reconstruction algorithm, it does in principle lead to noise correlations because the intrinsic

Faraday rotation of this source, which is part of the noise in our formalism, enters each of these data points in the same way. We ignore this effect in favor of a greatly simplified analysis. The combination of the response matrix and the inverse noise covariance matrix in Eq. (4.9) corresponds to an inverse noise weighted averaging of all data points that fall within one sky pixel. If the error bars were only caused by the intrinsic Faraday rotation of the sources, this would amount to an underestimation of the error bar by a factor $1/\sqrt{k}$ for a source that appears in k different catalogs. In reality, the intrinsic Faraday rotation constitutes only a fraction of the total error budget. The effect is therefore weaker.

Some of the catalogs listed in Tables 4.1 and 4.2 are themselves compilations of earlier measurements. As a consequence, some individual observations are contained in several of the catalogs. We have removed data points where we suspect such duplications so that each observation is used only once. Note that this does not apply to different observations of the same source, as discussed above. The number of data points given in Table 4.2 is the effective number of data points that we used in our analysis from the respective catalog.

Any variation of the Galactic Faraday depth within one pixel of our map can naturally not be reconstructed. Such variations on very small scales have been detected by Braun et al. (2010) for a region around $(l, b) \approx (94°, -21°)$. Should several sources fall within a pixel in such a region, our algorithm will yield an appropriate average value for the pixel and increase the error bars of the data points until they are consistent with this average value.

The sources studied in the Bonafede-catalog and some of the sources in the Clarke-catalog lie within or behind galaxy clusters. They are therefore expected to have an increased extragalactic contribution to their measured Faraday rotation. To take the cluster contribution into account, we corrected the error bars of these points accoring to

$$\sigma^2_{(\text{corrected})} = \sigma^2 + \sigma^2_{(\text{cluster})}. \tag{4.12}$$

To estimate the cluster contribution $\sigma_{(\text{cluster})}$, Bonafede et al. (2010) studied resolved background sources for which several independent RM measurements are possible. $\sigma_{(\text{cluster})}$ was then identified with the empirical value of the standard deviation of these measurements. Clarke et al. (2001) estimated the cluster contribution by comparing the RM values of sources within the cluster to those of sources behind the cluster. The Johnston-Hollit-B-catalog also contains sources associated with galaxy clusters. However, owing to the low density of sources, an estimation of the cluster contribution is not possible in this case. We expect a fraction of the other sources to be affected by clusters as well. However, since information on which sources exactly are affected is missing in general, we leave it to our algorithm to increase the error bars of the appropriate data points. The same problem exists in principle for satellite galaxies of the Milky Way, such as the Large and Small Magellanic Clouds. We do not attempt to separate their contribution to the Faraday depth from the one of the Milky Way, so that the map we reconstruct is strictly speaking not a pure map of the Galactic Faraday depth, but rather a map of the Faraday depth of the Milky Way and its surroundings. Because we use spatial correlations in the reconstruction algorithm, the Faraday depth contribution intrinsic to the sources will, however, be largely removed.
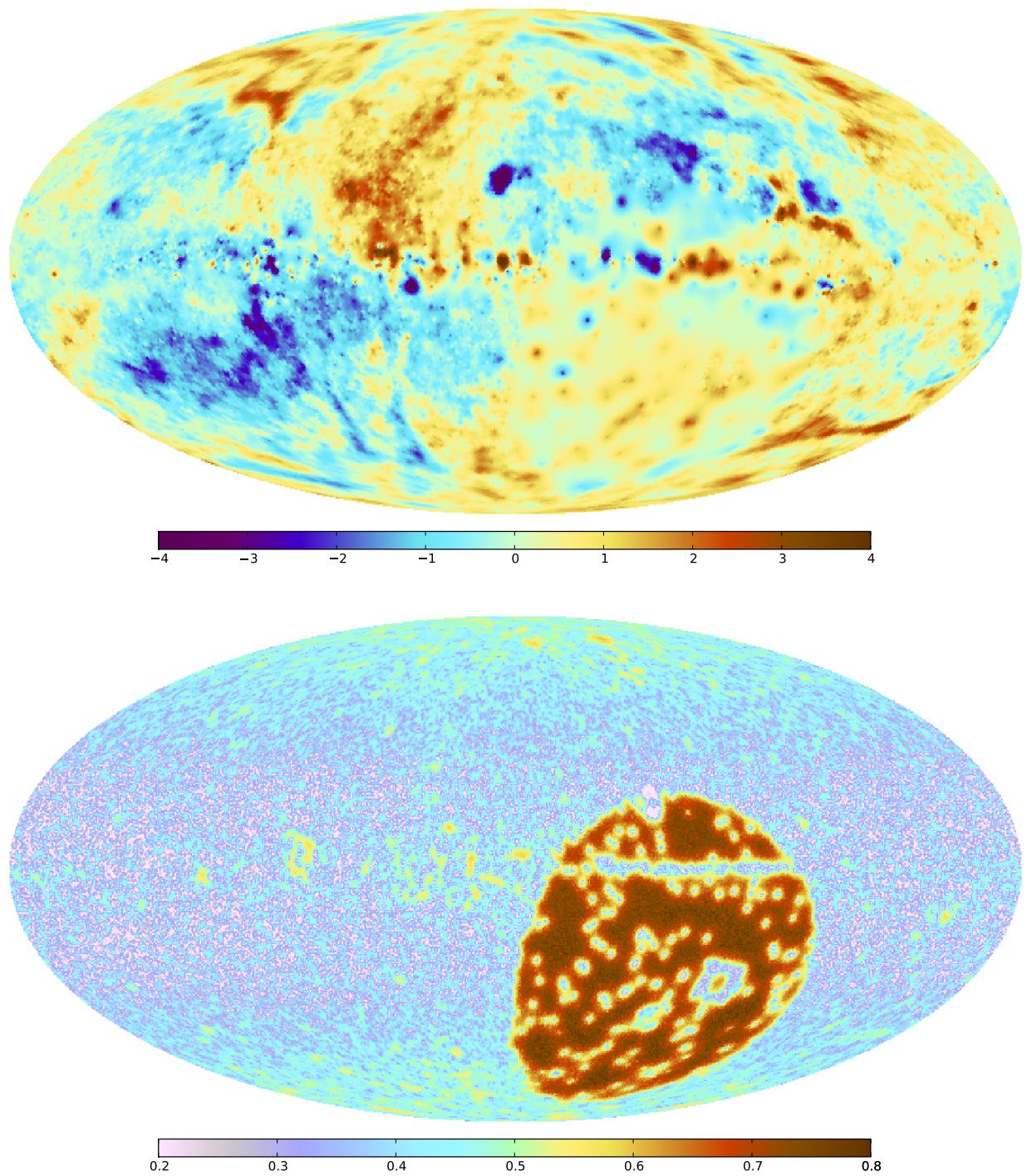
Figure 4.3: Reconstructed dimensionless signal map m (top) and its uncertainty $\hat{D}^{1/2}$ (bottom). Note the different color codes.
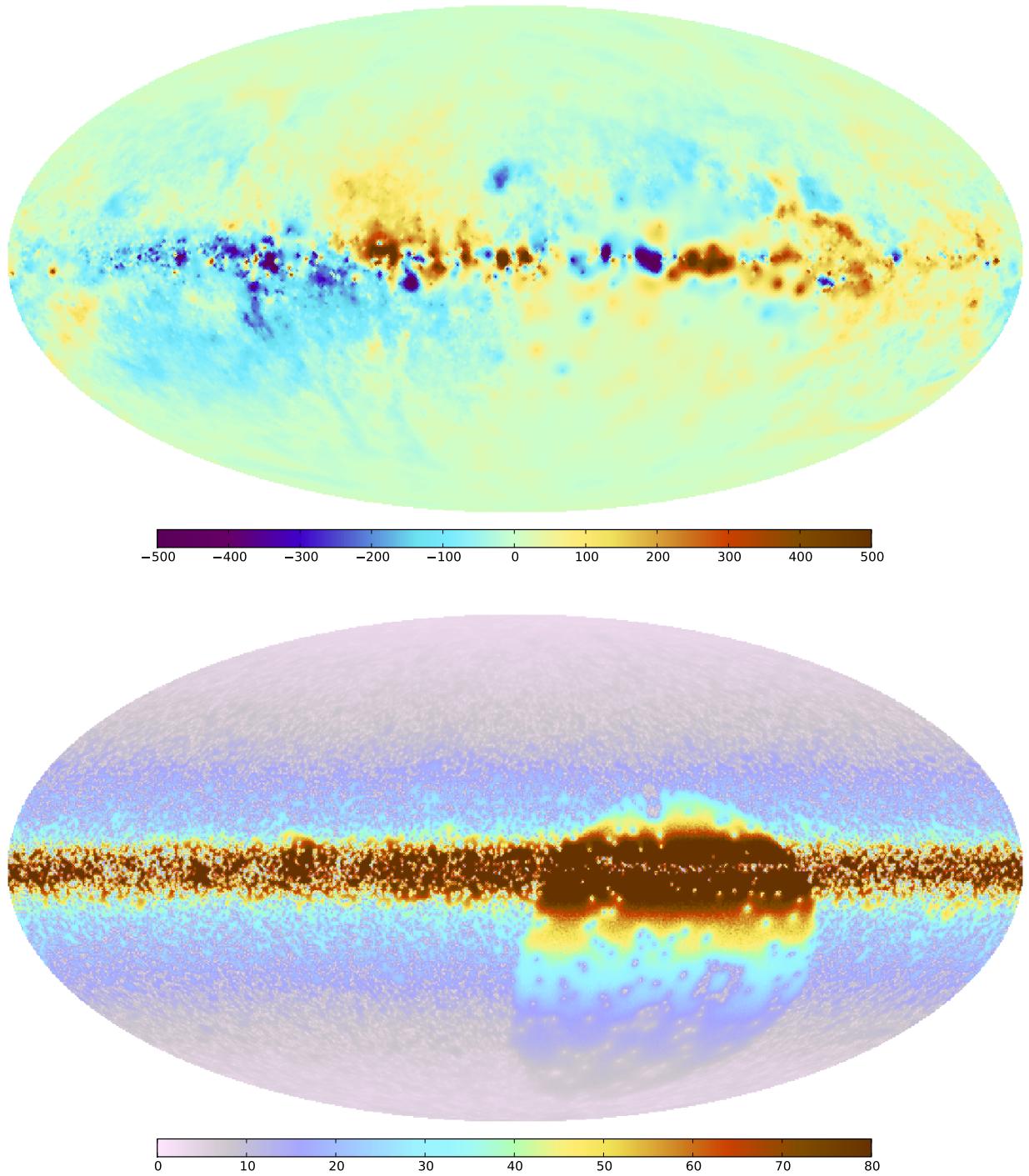
Figure 4.4: Reconstruction of the Galactic Faraday depth pm (top) and its uncertainty $p\hat{D}^{1/2}$ (bottom) in rad/m^2. Note the different color codes.

Furthermore, some of the sources will have a non-trivial Faraday spectrum, i.e. they exhibit polarized emission at more than one Faraday depth. While the technique of RM synthesis (Burn 1966; Brentjens & de Bruyn 2005) is able to distinguish these sources, such features are not described by a λ^2-fit, which may therefore lead to an erroneous rotation measure value. This problem becomes more severe if only a few frequencies are used in the fit. In the limit of two frequencies, multi-component Faraday spectra necessarily go unnoticed. We used the data points obtained by λ^2-fits of only a few frequencies nevertheless, and left it to the reconstruction algorithm to increase the error bars of those with an underlying multi-component spectrum accordingly.

4.4   Results

All results shown here are calculated at a HEALPix[4] resolution of $N_{\text{side}} = 128$, i.e. the maps contain 196 608 pixels. The minimum angular scale that we consider is $\ell_{\text{max}} = 383$, corresponding roughly to half a degree. These results are publicly available and can be downloaded from `http://www.mpa-garching.mpg.de/ift/faraday/`. The maps that we show are all centered on the Galactic center with positive Galactic latitudes at the top and positive Galactic longitudes plotted to the left.

4.4.1   Map

Figure 4.3 shows the reconstructed dimensionless signal map m and an estimate for its uncertainty, given by $\hat{D}^{1/2}$. The same for the physical Galactic Faraday depth pm, i.e. the signal multiplied by the Galactic variance profile, is shown in Fig. 4.4. As expected, the signal reconstruction is more uncertain in regions that lack data. Furthermore, the uncertainty in Fig. 4.3 tends to be smaller in the Galactic plane. This is because of the higher signal response caused by the Galactic variance profile in this area. When considering the uncertainty of the final map of the Faraday depth, i.e. the bottom panel of Fig. 4.4, this feature is turned around. The values within the Galactic plane now tend to be more uncertain than those near the poles. Note, however, that this is the absolute uncertainty. Because the Galactic Faraday depths are greater for lines of sight through the Galactic disk as well, the relative uncertainty is smaller there. This corresponds roughly to the uncertainty shown in the bottom panel of Fig. 4.3, which can be interpreted as the uncertainty of the Galactic Faraday depth relative to the value of the Galactic variance profile at the specific latitude. Also, the uncertainty is only high in the Galactic plane in pixels that do not contain any data. In the pixels that contain measurements, the uncertainty is comparable to the error bars of the data. It should be noted, however, that due to the approximations made in the derivation of the filter formulas (for details, see Oppermann et al. 2011b), the presented 1σ intervals cannot be interpreted as containing 68% of the correct pixel values of the signal. Oppermann et al. (2011b) found in their mock tests that about 50% of the correct pixel values lie within this range.

---

[4]The HEALPix package is available from `http://healpix.jpl.nasa.gov`.
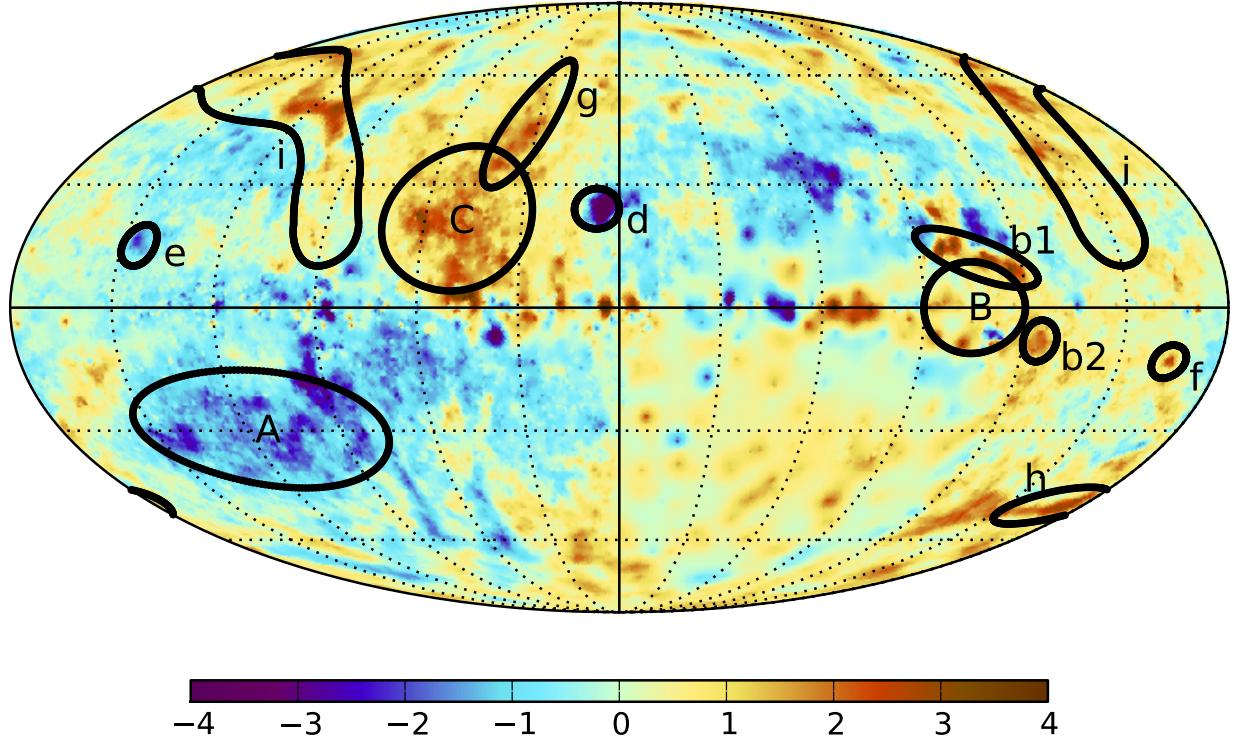
Figure 4.5: Same as the top panel of Fig. 4.3, with markings around the regions discussed in the text. The letters labeling the regions are used for reference in the main text. Dashed lines denote lines of constant Galactic longitude or latitude. Their angular separation is 30°.

In general, Fig. 4.3 is better suited to make out localized features away from the Galactic plane. The most striking of these features is the quadrupole-like structure on large scales that favors positive Faraday depths in the upper left and lower right quadrant and negative Faraday depths in the upper right and lower left quadrant. This has been observed in measurements of Faraday rotation in the past, first by Simard-Normandin & Kronberg (1980), and has often been claimed to be caused by a toroidal component of the large-scale Galactic magnetic field that changes sign over the Galactic plane (see e.g Han et al. 1997). Recent studies by Wolleben et al. (2010) and Mao et al. (2010) have shown, however, that this pattern is probably at least partly caused by local features of the interstellar medium in the solar neighborhood. At Galactic longitudes beyond roughly ±100°, this pattern turns into a dipolar structure, favoring negative values at the left edge of the map and positive ones on the very right, as noted previously by Kronberg & Newton-McGee (2011). This might be a signature of a toroidal magnetic field component that does not change sign over the Galactic plane. But of course this could also be a local effect, independent of the large-scale magnetic field.

Many other features are visible in the top panel of Fig. 4.3. We have marked some of the

features that have already been discussed in the literature in Fig. 4.5 for easier reference.

Simard-Normandin & Kronberg (1980) identified three large regions (A, B, and C in Fig. 4.5) with large angular size that stand out in Galactic Faraday depth amplitude. Stil et al. (2011) narrowed the definitions of the regions A and C down to their more striking parts using the NVSS RM catalog. Region A is a large area of negative Galactic Faraday depth localized roughly at $80° < l < 150°$, $-40° < b < -20°$. This region is seen in the direction of radio Loop II, but there is little evidence that the two are associated. The high-longitude boundary of region A coincides with part of the edge of Loop II. However, pulsar rotation measures suggest that Region A extends more than 3 kpc along the line of sight (Simard-Normandin & Kronberg 1980), which suggests that region A is a much larger structure.

Region B of Simard-Normandin & Kronberg (1980) is associated with the Gum nebula. Vallee & Bignell (1983) and Stil & Taylor (2007) identified a large magnetic shell in the area. The arc of positive Galactic Faraday depth around 250 rad/m^2 at $-120° < l < -90°$, $b \approx 13°$ (region b1 in Fig. 4.5) coincides with the northern Hα arc of the Gum nebula. A small excess in Galactic Faraday depth (region b2 in Fig. 4.5) is associated with the nearby HII region RCW 15 ($l = -125°$, $b = -7°$).

Region C is an area of positive Galactic Faraday depth in the range $33° < l < 68°$, $10° < b < 35°$ near the boundary of Radio Loop I. Wolleben et al. (2010) found diffuse polarized emission at a Faraday depth of 60 rad/m^2 at $l \approx 40°$, $b \approx 30°$ with associated HI structure, and interpreted this structure as part of a separate super shell around a subgroup of the Sco-Cen (Sco OB2_2).

Besides the Gum nebula, some extended HII regions at intermediate Galactic latitude can be identified in the form of a localized excess in Galactic Faraday depth (Stil & Taylor 2007; Harvey-Smith et al. 2011). The HII regions Sh 2-27 around ζ Oph at $l = 8°$, $b = 23.5°$ (region d in Fig. 4.5) and Sivan 3 around α Cam at $l = 144.5°$, $b = 14°$ (region e in Fig. 4.5) stand out as isolated regions of negative Galactic Faraday depth, while Sh 2-264 around λ Ori (region f in Fig. 4.5) is visible as a positive excess at $l = 195°$, $b = -12°$. Stil et al. (2011) presented an image of Hα intensity with rotation measure data overplotted.

Some large shells are also visible in the image of the Galactic Faraday depth. The Galactic anti-center direction is the most favorable direction to see these large structures, because it is less crowded than the inner Galaxy and the line of sight makes a large angle with the large-scale magnetic field. The North Polar Spur (region g in Fig. 4.5) is the notable exception toward the inner Galaxy. The filament of positive Galactic Faraday depth at $180° < l < 200°$, $b \approx -50°$ (region h in Fig. 4.5) is associated with the wall of the Orion-Eridanus superbubble (Heiles 1976; Brown et al. 1995). A large arc of positive Galactic Faraday depth (region i in Fig. 4.5) rises north of the Galactic plane at around $l \approx 95°$ up to $b \approx 65°$ around $l = 180°$ and curves back to the Galactic plane at around $l = 210°$ (Stil et al. 2011). This arc of positive Galactic Faraday depth traces the intermediate-velocity arch of atomic hydrogen gas identified by (Kuntz & Danly 1996).

Xu et al. (2006) reported RM excesses in the direction of the nearby Perseus-Pisces and Hercules super clusters. The higher sampling provided by the new Faraday rotation data catalogs has revealed high-latitude structures in the Galactic Faraday depth that

warrant additional investigation of the Galactic foreground effect. Many more small- and intermediate-scale features are visible in the top panel of Fig. 4.3. A detailed analysis of these features is left for future work.

4.4.2   Reconstruction of the noise covariance

The *extended critical filter* adapts the correction factors $(\eta_i)_i$, introduced in Sect. 4.2, to make the error bars of the data conform with the local map reconstruction. This is influenced by the surrounding data points and the angular power spectrum, which is in turn reconstructed using the entire data. Oppermann et al. (2011b) showed that allowing for this adaptation of the error bars leads to a slight oversmoothing of the reconstructed map since small-scale features that are only supported by individual data points are easily misinterpreted as noise.

In our reconstruction, we find that the median correction factor is $\eta^{(\mathrm{med})} = 0.56$. This indicates that the bulk of the data points are fairly consistent with one another and therefore with the reconstruction as well. As a consequence, their error bars are not enlarged but are instead slightly decreased by the algorithm. Oversmoothing can therefore not be a serious problem for the map as a whole. This is supported by the geometric mean of the correction factors, for which we find $\eta^{(\mathrm{geom})} = 0.75$. This corresponds to the arithmetic mean on a logarithmic scale and its prior expectation value was tuned to be one. Looking at the arithmetic mean on a linear scale, we find $\eta^{(\mathrm{mean})} = 6.40$, indicating that there are at least a few data points whose error bars are significantly corrected upward. Indeed, there are 134 data points with $\eta_i > 400$, meaning that the error bar has been increased by a factor of more than 20. These are isolated outliers in the data that are not consistent with their surroundings.

Figure 4.6 shows the final distribution of η-values. The bulk of these values lies around $\eta = 1$ or even slightly below. Only relatively few data points have highly increased error bars (note the logarithmic scale of the vertical axis in Fig. 4.6). Also plotted in Fig. 4.6 is the distribution of η-values that resulted from a reconstruction in which the slope parameter in the prior for the correction factors was chosen to be $\beta = 3$, as well as the prior probability distributions corresponding to $\beta = 2$ and $\beta = 3$. This shows two things. The resulting distribution does not change much when the value of β is changed and both distributions are better represented by a prior with $\beta = 2$. Our choice for β is thus justified.

The data points with $\eta \gg 1$ do not appear to be spatially clumped, making it improbable that any extended physical features that are present in the data are lost through the increase in the assumed noise covariance. Any real features that might mistakenly be filtered out in this procedure can be expected to be smaller or comparable in size to the distance to the next data point, i.e. one or two pixels or about one degree in most parts of the sky. The data points with strongly corrected error bars are predominantly located near the Galactic plane. This can be clearly seen in Fig. 4.7, where we plot the distribution of the correction factors for three latitude bins separately. While the difference in the distributions for the polar regions and the intermediate latitude bin is not very big, the data points around the Galactic disk clearly are more likely to have correction factors at
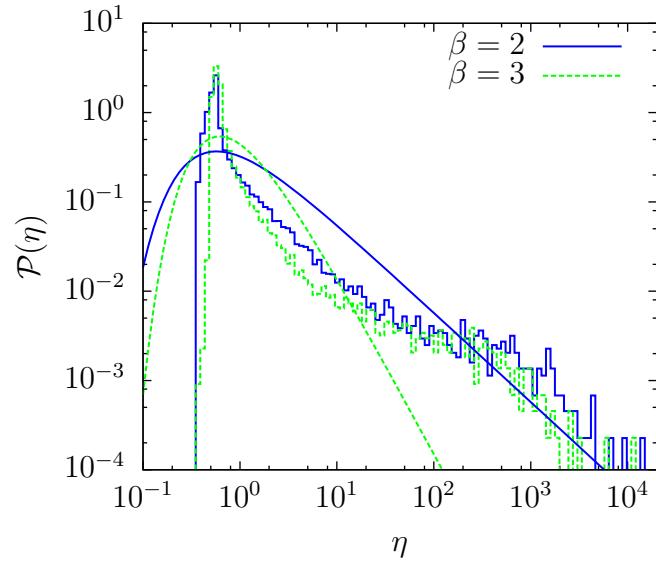
Figure 4.6: Comparison of the reconstructed distribution of the correction factors η that enter the noise covariance matrix and their priors. The dark histogram and line show the normalized distribution and prior for $\beta = 2$, respectively. The light histogram and line show the same for $\beta = 3$.
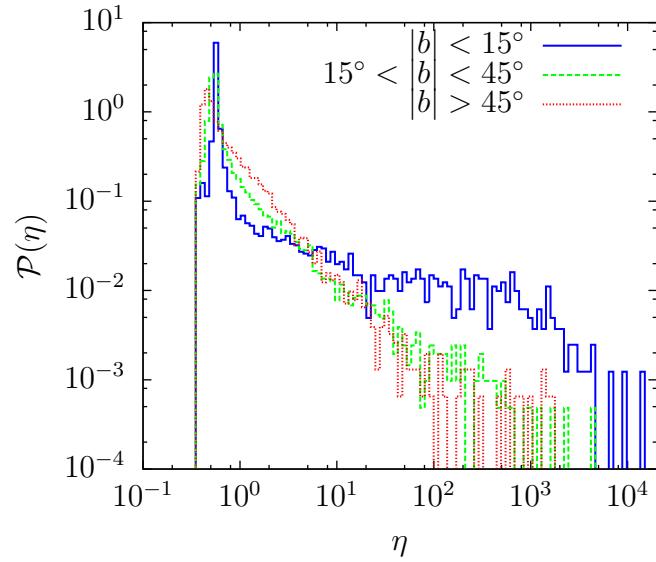
Figure 4.7: Comparison of the reconstructed distributions of the correction factors η for different latitude bins. The dark solid histogram depicts the distribution for data points within the Galactic plane, the light dashed histogram the distribution for data points at intermediate latitudes, and the dotted histogram the one for data points in the polar regions. Only the results obtained with $\beta = 2$ are shown.
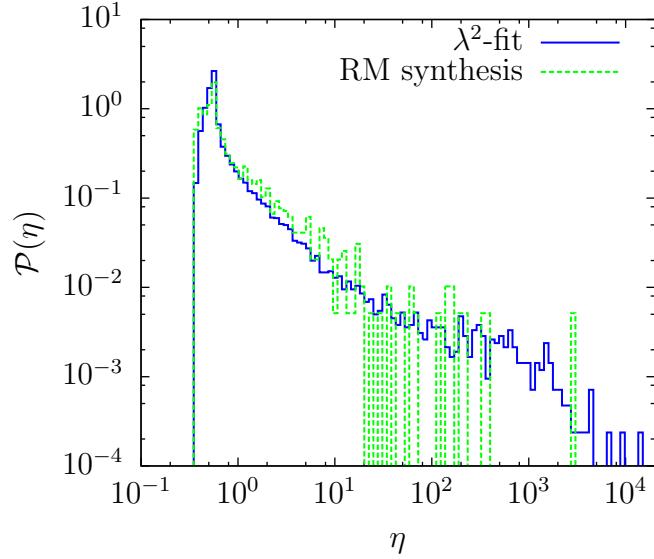
Figure 4.8: Comparison of the reconstructed distributions of the correction factors η for the two data reduction techniques. The dark solid histogram depicts the distribution for data points obtained from a linear λ^2-fit and the light dashed histogram the distribution for data points stemming from RM synthesis studies.

the high end. At least in some cases these high η-values can be interpreted as correcting an offset in the rotation angle of π that has escaped the observational analysis. Others might be due to a high level of polarized emissivity within the Galactic disk that can lead to misleading RM fits. Another reason for high η-values is a higher extragalactic contribution to the measured Faraday rotation, caused e.g. by magnetic fields in galaxy clusters. This last reason, however, would not be expected to show any statistical latitude dependence.

As mentioned above, a non-trivial emission spectrum in Faraday space is hard to identify when using linear λ^2-fits to obtain RM values. We therefore compare the distributions of the correction factors for data points from λ^2-fits and the ones for data points that stem from RM synthesis studies in Fig. 4.8. From the histograms it can indeed be seen that the data from λ^2-fits are more likely to have a high η-value, as expected.

Figure 4.9 shows a comparison of our reconstructed signal map with the reconstruction of Oppermann et al. (2011a), where the *critical filter* formalism was used without accounting for uncertainties in the noise covariance and only data from the Taylor-catalog were used. The differences that can be seen are twofold. On the one hand, our map shows structure due to the additional data points that we used, most prominently at declinations below $-40°$. On the other hand, some of the features present in the older map have vanished because they were supported only by a single data point that was interpreted as being noise-dominated by our algorithm. These features appear prominently both in the old map and in the difference map, where our newly reconstructed map has been subtracted from the old one. They have the same sign in both these maps. Also, our new reconstruction
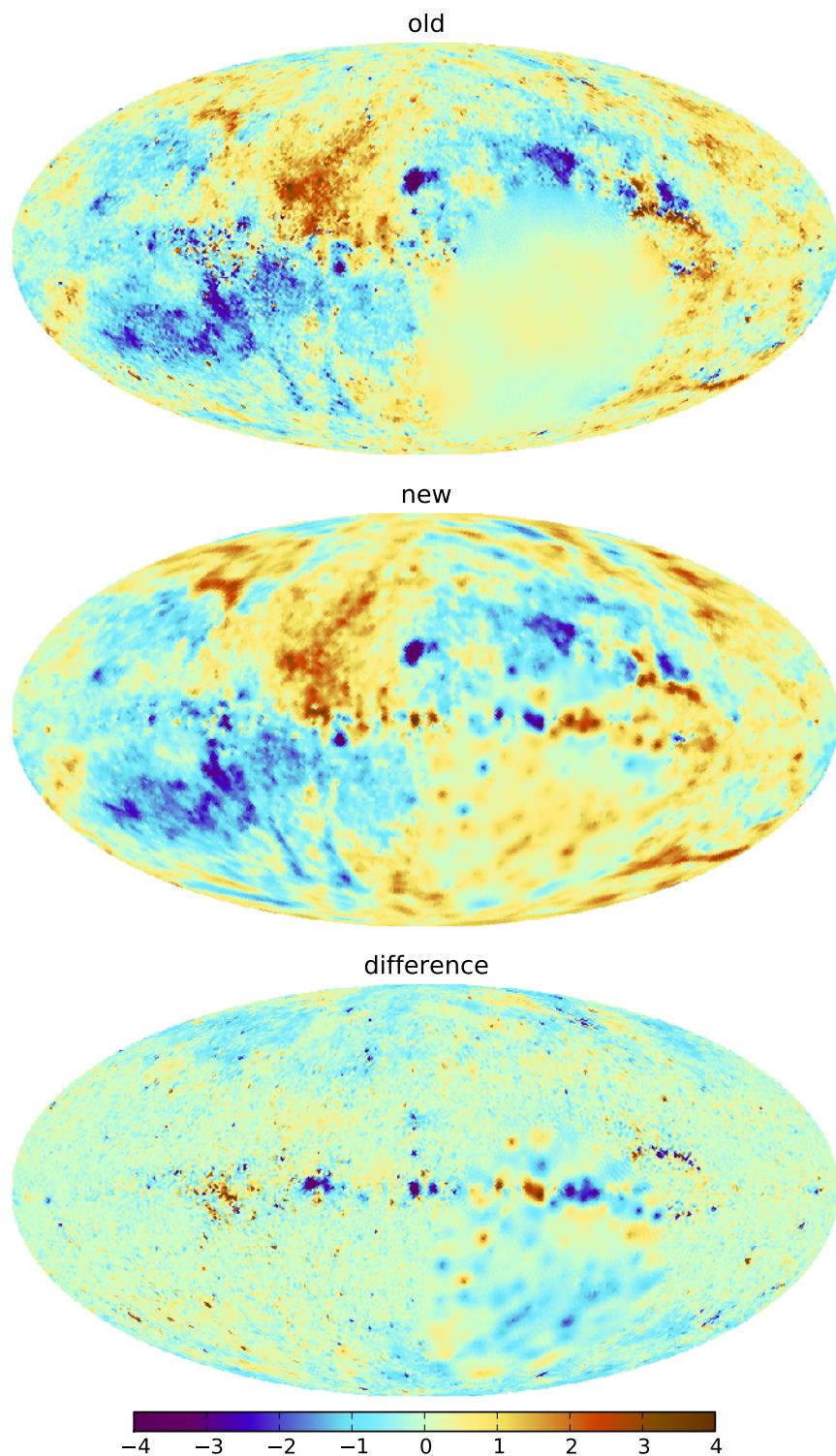
Figure 4.9: Comparison of the reconstruction of the dimensionless signal to earlier results. The top panel shows the reconstructed signal field of Oppermann et al. (2011a), the middle panel shows the same as the top panel of Fig. 4.3, only coarsened to a resolution of $N_{\mathrm{side}} = 64$ to match the resolution of the old reconstruction. The bottom panel shows the difference between the upper panel and the middle panel.
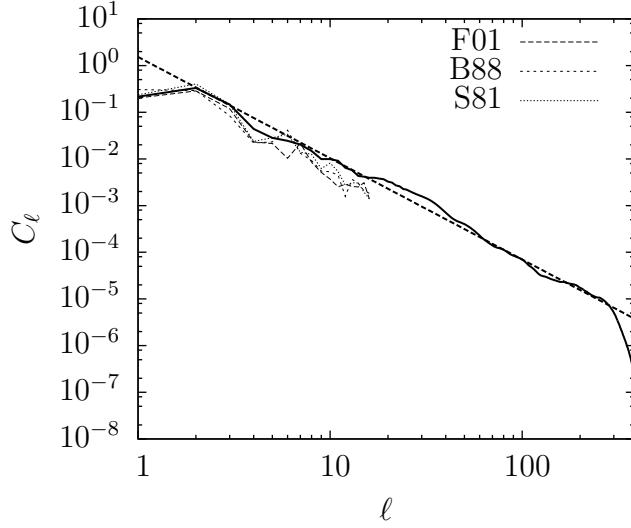
Figure 4.10: Angular power spectrum of the dimensionless signal field (thick solid line), along with a power law fit, $C_\ell \propto \ell^{-2.17}$ (thick dashed line). The thin lines depict the angular power spectra corresponding to the maps reconstructed by Dineen & Coles (2005), corrected for the Galactic variance profile. The three RM catalogs used in their work are from Simard-Normandin et al. (1981) (S81), Broten et al. (1988) (B88), and Frick et al. (2001) (F01).

is less grainy. This is a combined effect of the higher resolution that we used and the adaptation of error bars during our reconstruction.

4.4.3   Power spectrum

The reconstructed angular power spectrum of the dimensionless signal field is shown in Fig. 4.10. It is well-described by a power law. A logarithmic least-squares fit, which is also shown in Fig. 4.10, yields a spectral index of 2.17, i.e.

$$C_\ell \propto \ell^{-2.17}, \tag{4.13}$$

where we have taken scales down to $\ell = 300$ into account. Note that owing to the typical distance of neighboring data points of roughly one degree, structures smaller than this angular size, corresponding to $\ell \gtrsim 180$, will in general not be reconstructed and we might therefore be missing some power on the smallest scales. However, some data points have smaller angular separations and we therefore have some information on the angular power spectrum up to $\ell_{\max} = 383$.

Also shown in Fig. 4.10 is a comparison with the angular power spectra of the maps that Dineen & Coles (2005) reconstructed. They created three separate maps from three differ-
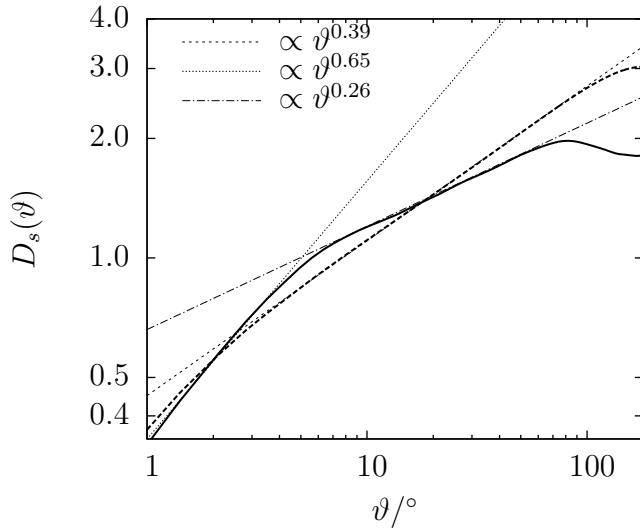
Figure 4.11: Second-order structure function corresponding to the angular power spectrum plotted in Fig. 4.10 (thick solid line) and its power law fit (thick dashed line), along with power law approximations (thin lines).

ent RM catalogs. We used the spherical harmonics components of their maps[5], transformed them into position space, and then divided them by our Galactic variance profile. We plot the angular power spectra of the three resulting dimensionless maps. Evidently, both the slope and the normalization of the spectra agree with our result. Haverkorn et al. (2003) studied the angular power spectrum of rotation measures of diffuse polarized radio emission from the local interstellar medium in two regions of the sky on scales $400 < \ell < 1500$. They fitted power laws with exponents close to -1, i.e. $C_\ell \propto \ell^{-1}$, significantly larger than our result. This is not necessarily a contradiction, however, since a flattening of the angular power spectrum on scales that are too small for our analysis could explain both results. Furthermore, we take into account the full line of sight through the galaxy by using only extragalactic sources, therefore the volume that we probe is significantly larger than that probed by Haverkorn et al. (2003).

To compare our result to other earlier papers, we considered the second-order structure function for the dimensionless signal field,

$$
\begin{aligned}
D_s(\vartheta) &= \left\langle \left(s(\hat{n}) - s(\hat{n}') \right)^2 \right\rangle_{\mathcal{P}(s)} \\
&= 2 \left(S_{\hat{n}\hat{n}} - S_{\hat{n}\hat{n}'} \right),
\end{aligned}
\tag{4.14}
$$

where $\vartheta = \arccos(\hat{n} \cdot \hat{n}')$ and $\hat{n}$ and $\hat{n}'$ are two directions in the sky. Here, S denotes the signal covariance matrix and the angle brackets denote a prior ensemble average. Since we assume statistical homogeneity and isotropy for the signal field, $S_{\hat{n}\hat{n}}$ does not depend on $\hat{n}$,

---

[5]Dineen & Coles (2005) provide their results at `http://astro.ic.ac.uk/~pdineen/rm_maps/`.

$S_{\hat{n}\hat{n}'}$ depends only on ϑ, and both terms are completely determined by the angular power spectrum. This also allows us to exchange the usual spatial average with an ensemble average in Eq. (4.14). The resulting structure function is plotted in Fig. 4.11. Using the final angular power spectrum of our reconstruction (the solid line in Fig. 4.10), we find a broken power law with exponents 0.65 for small angles and 0.26 for large angles with the transition occuring around $\vartheta = 5°$ (the solid line in Fig. 4.11). The power law fit to the angular power spectrum (the dashed line in Fig. 4.10) leads to a structure function that can be approximated by a single power law with exponent 0.39 (the dashed line in Fig. 4.11).

Minter & Spangler (1996) found that the structure function derived from their observations is well described by a power law with exponent 0.64 for angular scales of $\vartheta > 1°$. Sun & Han (2004) studied the structure function in three different regions within the Galactic plane and in the vicinity of the North Galactic pole. An inverse-noise-weighted average of their power law indices yields a value of 0.11. Haverkorn et al. (2006a) and Haverkorn et al. (2008) studied observations through interarm regions in the Galactic plane separately from observations along Galactic arms. These authors found flat structure functions for the observations along Galactic arms. Haverkorn et al. (2006a) found a weighted mean power law index of 0.55 for the structure functions derived from observations through interarm regions, while Haverkorn et al. (2008) found an inverse-noise-weighted mean power law index of 0.40. Haverkorn et al. (2003) found flat structure functions for the two regions they studied. Roy et al. (2008) found a structure function for the region around the Galactic center that is constant on scales above $\vartheta = 0.7°$ and exhibits a power law behavior with an exponent of 0.7 on smaller scales. Stil et al. (2011) fitted broken power laws with the breaking point at $\vartheta = 1°$ to the structure functions they extracted from the NVSS rotation measure catalog (Taylor et al. 2009). They found power law indices that vary spatially. Taking an inverse-noise-weighted average of their power law indices for the regions that they studied in detail yields 0.37 for $\vartheta > 1°$ and 0.59 for $\vartheta < 1°$.

These observational results indicate that the slope of the structure function varies from region to region. Our result is insensitive to these variations since our structure function is just a description of the prior for the dimensionless signal, for which we have assumed statistical isotropy. It can therefore be interpreted as a mean structure function across the whole sky. The observations that yield non-flat structure functions roughly agree with the slopes that we fitted in Fig. 4.11. The dependence of the structure function slope on Galactic latitude (e.g. Simonetti et al. 1984; Sun & Han 2004) is partly removed in our analysis by the division through the Galactic variance profile. Note that Simonetti et al. (1984) and Simonetti & Cordes (1986) already suspected a break in the structure function at roughly five degrees. However, existing studies have not shown convincing evidence for this.

Consequences for the 3D fields

As an illustrative thought experiment, assume that an observer is sitting in the middle of a spherical distribution of magnetoionic medium. Let $\tilde{\varphi}(\vec{x}) \propto n_e(\vec{x})B_r(\vec{x})$ be the product

of the local thermal electron density and the line-of-sight component of the magnetic field as a function of 3D position $\vec{x}$, i.e. the differential contribution to the Faraday depth that this observer is measuring. We model this field as factorizing into two parts,

$$\tilde{\varphi}(\vec{x}) = \bar{\varphi}(r)\varphi(\vec{x}). \tag{4.15}$$

The first part is a spherically symmetric contribution, whose functional dependence on the radial distance from the observer is known, and the second part is assumed to be a realization of a statistically homogeneous and isotropic random field, i.e.

$$\left\langle \varphi(\vec{k})\varphi^*(\vec{k}') \right\rangle = (2\pi)^3 \, \delta^{(3)}(\vec{k} - \vec{k}')P_\varphi(k), \tag{4.16}$$

where the angle-brackets denote an average over all possible field realizations, $P_\varphi(k)$ is the Fourier power spectrum[6] that describes the statistics of φ, and $k = \left|\vec{k}\right|$.

Using the simplest form of $\bar{\varphi}(r)$, namely a constant within some finite radius r_0, i.e.

$$\bar{\varphi}(r) = \left\{ \begin{array}{ll} \varphi_0 & \text{if } r < r_0 \\ 0 & \text{else} \end{array} \right. , \tag{4.17}$$

and a power law for the Fourier power spectrum,

$$P_\varphi(k) \propto k^{-\alpha}, \tag{4.18}$$

we calculated the angular power spectrum of the Faraday depth that the observer would measure and compared the result numerically with Eq. (4.13) (see App. D for the details of the calculation). We find that the two agree well if one chooses α roughly equal to the power law index that was found for the angular power spectrum, i.e. 2.17 in this case.

A similar thought experiment has been conducted by Simonetti et al. (1984). They assumed a Fourier power spectrum $P_\varphi(k) \propto \exp\left(-k^2/k_1^2\right)\left(1 + k^2/k_0^2\right)^{\alpha/2}$, i.e. a power law with a low-wavenumber cutoff at k_0 and a high-wavenumber cutoff at k_1, and calculated the expected structure function. In the power law regime, i.e. $1/k_1 \ll r_0 \sin\vartheta \ll 1/k_0$, they found $D_s(\vartheta) \propto \vartheta^{\alpha-2}$ to lowest order in ϑ. Extending this study to independent variations in the thermal electron density and the magnetic field component along the line of sight, each described by a power law power spectrum with the same index α, Minter & Spangler (1996) found the same dependence on ϑ.[7] Our intermediate fit of $D_s(\vartheta) \propto \vartheta^{0.39}$ (see Fig. 4.11) therefore corresponds to $\alpha = 2.39$, in rough agreement with our numerical finding from the power spectrum analysis.

Armstrong et al. (1995) used observations of effects of interstellar radio scintillation (see also Rickett 1977, 1990), as well as pulsar dispersion measures, to constrain the power spectrum describing the fluctuations of the thermal electron density in the local interstellar

---

[6]Note that the definition of the Fourier power spectrum made in Eq. (4.16) corresponds to what is sometimes referred to as the 3D power spectrum, i.e. the variance of the field φ at each position $\vec{x}$ in real space can be calculated as $\left\langle \varphi^2(\vec{x}) \right\rangle \propto \int_0^\infty \mathrm{d}k \, k^2 P_\varphi(k)$.

[7]Minter & Spangler (1996) assume a rectangular shape for φ_0 instead of a spherical one.

medium. They found a Kolmogorov-type power spectrum, i.e. a power law index of $\alpha = 11/3$ in the present notation. This result was combined by Minter & Spangler (1996) with their own observations of rotation measures of extragalactic sources. Since they did not find the slope expected from the Kolmogorov power law in the structure function of the rotation measure they observed, they concluded that the outer scale of the Kolmogorov-type turbulence is smaller than the smallest scale probed by their RM observations. They fitted model structure functions for the variations of the thermal electron density and the magnetic field to their own observations of RM, as well as observations of Hα intensity and Hα velocity performed by Reynolds (1980), while also taking into account the results of Armstrong et al. (1995) on smaller scales. This procedure led to an estimate of the angular scale corresponding to the outer scale of the turbulence in the region of their observations of $\vartheta^{(\mathrm{out})} \lesssim 0.1°$. Although the outer scale of the turbulence may well vary across the Galaxy, it is probably safe to assume that the scales larger than one degree, which are mainly probed by the observations used in this work, are not dominated by three-dimensional turbulence. Whether or not the simple power law behavior of the angular power spectrum in Eq. (4.13) points to some sort of interaction between the fluctuations on different scales is at the moment an open question.

In any case it is clear that the simplifying assumptions made in the thought experiments presented above are far from the truth in the Galactic setting. A more realistic study will likely have to involve numerical magneto-hydrodynamical simulations of the interstellar medium, which have become more and more sophisticated over the last years (see e.g. de Avillez & Breitschwerdt 2007; Kissmann et al. 2008; Burkhart et al. 2009; Tofflemire et al. 2011). Cross-checking the angular power spectrum of the Faraday depth that is predicted by such a simulation against Eq. (4.13) might be a good indicator of how realistic the simulation actually is. For this, an empiric variance profile would have to be calculated from the simulated observations to create a dimensionless signal field comparable to our reconstruction. Numerical studies will also be able to show whether the simple power law that we find for the angular power spectrum is a functional form that arises generically or an outcome that needs certain ingredients. This may then enable a physical interpretation of the angular power spectrum that we find. On the other hand, if simulations show that different physical processes are needed to create the fluctuation power on different angular scales, our result will directly constrain the relative strength of these processes.

4.5   Conclusions

We have presented a map of the Galactic Faraday depth that summarizes the current state of knowledge, along with its uncertainty. For the map reconstruction we used the *extended critical filter*, a state-of-the-art algorithm, yielding a result that is robust against individual faulty measurements. It is this robustness, along with the usage of the most complete data set on the Faraday rotation of extragalactic sources to date, and the high resolution that we are therefore able to reach, that make our map an improvement over existing studies. Along with the map, the reconstruction algorithm yields the angular power spectrum of

the underlying signal field, $C_\ell \propto \ell^{-2.17}$, which agrees with earlier work. We discussed the implications of this power spectrum for the statistics of the 3D quantities involved in a greatly simplified scenario and suggested future work on simulations with the possibility of checking predicted angular power spectra against our observational result.

All products of this work, i.e. the maps and their uncertainties, as well as the angular power spectrum, are made available to the community[8] for further analysis, interpretation, and for use in other work where the Galactic Faraday depth plays a role.

---

[8]See `http://www.mpa-garching.mpg.de/ift/faraday/` for a fits-file containing all the results and an interactive map to explore the Galactic Faraday sky.

Table 4.1: Details of the surveys used for the map reconstruction (footnotes on page 100).

identifyer	telescope	survey	# observed wavelengths
Bonafede[1]	VLA		3-5
Broten[2]	various[b]		various[b]
Brown CGPS[3]	DRAO ST	CGPS[4]	4
Brown SGPS[5]	ATCA	SGPS[6]	12
Clarke[7]	VLA		4,6
Clegg[8]	VLA		6
Feain[9]	ATCA	Cent. A[10]	24
Gaensler[11]	ATCA	SGPS test	9
Hammond[12]	ATCA		23
Heald[13]	WSRT	WSRT-SINGS[14]	1024
Hennessy[15]	VLA		4
Johnston-Hollitt A[16]	ATCA		23
Johnston-Hollitt B[17]	ATCA		4
Kato[18]	Nobeyama		4[d]
Kim[19]	various[e]		various[e]
Klein[20]	VLA & Effelsberg	B3/VLA[21]	4
Lawler[22]	various[f]	[23]	various[f]
Mao SouthCap[24]	ATCA		32
Mao NorthCap[24]	WSRT		16
Mao LMC[25]	ATCA		14
Mao SMC[26]	ATCA		14
Minter[27]	VLA		4
Oren[28]	VLA		4,6
O'Sullivan[29]	ATCA		100
Roy[30]	ATCA & VLA		4 and more
Rudnick[31]	VLA		2
Schnitzeler[32]	ATCA		12
Simard-Normandin[33]	various[j]		various[j]
Tabara[34]	various[k]		various[k]
Taylor[35]	VLA	NVSS[36]	2
Van Eck[37]	VLA		14
Wrobel[38]	VLA		6

Table 4.2: Details of the data sets used for the map reconstruction (footnotes on page 100).

identifyer	frequency range / MHz	method	# data points
Bonafede[1]	various[a]	λ^2-fit	7
Broten[2]	various[b]	λ^2-fit	121+3/2[c]
Brown CGPS[3]	1 403-1 438	λ^2-fit	380
Brown SGPS[5]	1 332-1 436	λ^2-fit	148
Clarke[7]	1 365-4 885	λ^2-fit	125
Clegg[8]	1 379-1 671	λ^2-fit	56
Feain[9]	1 280-1 496	RM synthesis	281
Gaensler[11]	1 334-1 430	λ^2-fit	18
Hammond[12]	1 332-1 524	RM synthesis	88
Heald[13]	1 300-1 763	RM synthesis	57
Hennessy[15]	1 362-1 708	λ^2-fit	17
Johnston-Hollitt A[16]	1 292-1 484	RM synthesis	68
Johnston-Hollitt B[17]	1 384-6 176	λ^2-fit	12
Kato[18]	8 800-10 800[d]	λ^2-fit	1
Kim[19]	various[e]	λ^2-fit	20+1/2[c]
Klein[20]	1 400-10 600	λ^2-fit	143
Lawler[22]	various[f]	λ^2-fit	3
Mao SouthCap[24]	1 320-2 432	RM synthesis	329
Mao NorthCap[24]	1 301-1 793	RM synthesis	400
Mao LMC[25]	1 324-1 436	RM synthesis	188
Mao SMC[26]	1 324-1 436	λ^2-fit	62
Minter[27]	1 348-1 651	λ^2-fit	98
Oren[28]	various[g]	λ^2-fit	51+4/2[c]
O'Sullivan[29]	1 100-2 000	RM synthesis	46
Roy[30]	various[h]	λ^2-fit	67
Rudnick[31]	1 440-1 690	λ^2-fit	17+2/2[c]
Schnitzeler[32]	1 320-11 448[i]	RM synthesis	178
Simard-Normandin[33]	various[j]	λ^2-fit	535+6/2[c]
Tabara[34]	various[k]	λ^2-fit	62+3/2[c]
Taylor[35]	1 344-1 456	λ^2-fit	37 543
Van Eck[37]	1 353-1 498	RM synthesis[l]	194
Wrobel[38]	1 373-1 677	λ^2-fit	5+1/2[c]

[1] Bonafede et al. (2010)   [2] Broten et al. (1988)   [3] Brown et al. (2003a)
[4] Taylor et al. (2003)   [5] Brown et al. (2007)   [6] Haverkorn et al. (2006b);
McClure-Griffiths et al. (2005)   [7] Clarke et al. (2001); Clarke (2004)
[8] Clegg et al. (1992)   [9] Feain et al. (2009)   [10] Feain et al. (2011)
[11] Gaensler et al. (2001)   [12] Hammond (private communication)
[13] Heald et al. (2009)   [14] Braun et al. (2007)   [15] Hennessy et al. (1989)
[16] Johnston-Hollitt (in prep.)   [17] Johnston-Hollitt (2003); Johnston-Hollitt & Ekers (2004)   [18] Kato et al. (1987)   [19] Kim et al. (1991)
[20] Klein et al. (2003)   [21] Gregorini et al. (1998); Vigotti et al. (1999)
[22] Lawler & Dennison (1982)   [23] Dennison (1979)   [24] Mao et al. (2010)
[25] Mao (in prep.); Gaensler et al. (2005)   [26] Mao et al. (2008)
[27] Minter & Spangler (1996)   [28] Oren & Wolfe (1995)
[29] O'Sullivan (private communication)   [30] Roy et al. (2005)   [31] Rudnick & Jones (1983)   [32] Schnitzeler (in prep.)   [33] Simard-Normandin et al. (1981)   [34] Tabara & Inoue (1980)   [35] Taylor et al. (2009)
[36] Condon et al. (1998)   [37] Van Eck et al. (2011)   [38] Wrobel (1993)
[a] Three different frequency ranges, 4 510-8 490 MHz, 4 510-8 300 MHz, 1 340-4 960 MHz, were used.
[b] Compilation of several previously published data sets.
[c] Data points that seem to be duplications of the same observations, appearing in two different catalogs, are used only once and denoted as half data points for both catalogs, so that the sum of the data points is the total number of data points used.
[d] Additional data from Tabara & Inoue (1980) used in the fit.
[e] Compilation of several earlier data sets, including an unpublished one for which no details are provided.
[f] Compilation of several earlier data sets and the one described in Dennison (1979).
[g] Three different frequency ranges, 1 360-1 690 MHz, 1 373-4 898 MHz, 1 373-4 990 MHz, were used.
[h] The frequency range is 4 736-8 564 MHz for the ATCA observations and 4 605-8 655 MHz for the VLA observations.
[i] The frequency range is shifted to lower frequencies by up to 40 MHz for some sources.
[j] RMs calculated from previously published and unpublished data, as well as new measurements with various instruments.
[k] RMs calculated from previously published polarization data.
[l] λ^2-fits were also produced and found to agree with the synthesis results.

Chapter 5

Reconstruction of Gaussian and log-normal fields with spectral smoothness

Note: This chapter, as well as Appendix E, has been submitted for publication in Physical Review E *(Oppermann et al. 2012b).*

5.1 Introduction

Reconstructing continuous fields from a finite and noisy data set is a problem encountered often in all branches of physics and the geo-sciences. In this paper we develop a method to reconstruct a log-normal field, i.e. a field whose logarithm can be modeled as a Gaussian random field, defined on an arbitrary manifold. We show simple one-dimensional examples as well as planar and spherical two-dimensional cases.

The log-normal model is well suited to describe many physical fields. Its main features are that the field values are guaranteed to be positive, that they may vary across many orders of magnitude, and that spatial correlations may exist. One prominent example from the discipline of astrophysics that exhibits these three features and is therefore often modeled as a log-normal field is the matter density in today's universe (see e.g. Coles & Jones 1991). The simplest way to observationally estimate the matter density is to count the galaxies per volume. The relationship between the matter density and the galaxy number counts can be approximated as a Poisson process (e.g. Kitaura & Enßlin 2008; Kitaura et al. 2010; Jasche et al. 2010a; Jasche & Wandelt 2012), thus making the data likelihood Poissonian as well with the expected number of galaxies per volume element given as a function of the underlying log-normal density field. In the statistical literature, such a combination of a log-normal field with Poissonian statistics is known as a log Gaussian Cox process (Møller et al. 1998). This model has been applied in fields as diverse as finance (Basu & Dassios 2002), agriculture (Brix 2001), atmospheric studies (Pati et al. 2011), and epidemiology (Brix & Diggle 2001; Beneš et al. 2002).

The main motivation for our work, however, comes from astronomical observations of radiative processes. The intensity of radiation coming from different directions in such observations can vary over many orders of magnitude while being spatially correlated. Thus applying a log-normal model for its statistical description seems natural. While photon counts bring with them a Poissonian likelihood, the observational uncertainty can be approximatively regarded to be Gaussian in the limit of large photon numbers. Therefore, we restrict ourselves to cases in which the measurement noise can be assumed to be additive and Gaussian.

Apart from the noise contribution, we will assume a deterministic linear relationship between the log-normal field and the data. This model is general enough to accomodate a large variety of observational settings, such as targeted point-wise observations across limited areas, convolutions of the log-normal field, interferometric observations leading to a Fourier transformed version of the log-normal field, or any combination of these effects.

The inclusion of spatial correlations in the field model is necessary for an accurate statistical description and will aid in the reconstruction. If the spatial correlation structure is known, knowledge of the field values at some locations can be used to extrapolate the field into regions where the field value has not been measured or the observations have a higher uncertainty. However, in general it is not known a priori how strongly the field is correlated across a given distance. In the case of a statistically homogeneous field, the correlation structure is described by the field's power spectrum. So if one wants to make use of the spatial correlations during the reconstruction, one needs to simultaneously infer the power spectrum. Several techniques have been developed to solve this problem for Gaussian fields (e.g Wandelt et al. 2004; Jasche et al. 2010b). One such technique was derived within the framework of *information field theory* (Enßlin et al. 2009; Lemm 2003) by Enßlin & Frommert (2011) and later rederived by Enßlin & Weig (2010), where the formalism of minimum Gibbs free energy was introduced and employed to tackle this problem. In the same paper, the problem of inferring a log-normal field with unknown power spectrum in the presence of Poissonian noise was briefly discussed. Kitaura et al. (2012) use a Gibbs sampling procedure to reconstruct a log-normal field along with its power spectrum from Poissonian distributed data.

Here, we use the formalism of minimum Gibbs free energy to derive filter equations for the estimation of a log-normal field from data including a Gaussian noise component. The filter we present results in an estimate of the field, its power spectrum, and the respective uncertainties.

A problem that often arises when reconstructing power spectra is that the reconstructed power spectra do not turn out smooth, especially on scales with little redundancy, despite the true underlying spectra being smooth. One reason why this is not desirable is that sharp kinks and drop-offs in the power-spectrum can lead to ringing effects in the reconstructed field estimate. Another reason is that many of the fields occuring in nature are actually expected to have smoothly varying power over different scales since neighboring scales interact. One prominent example is the power spectrum of a viscous turbulent fluid (Kolmogorov 1941). Smoothness of the power spectrum is often enforced by an ad-hoc smoothing step in the power spectrum reconstruction (Oh et al. 1999; Oppermann et al.

2011a, 2012a). Here, we follow an idea presented by Enßlin & Frommert (2011) and show how to enforce spectral smoothness by introducing an appropriate prior for the power spectrum that punishes non-smooth spectra. We demonstrate the feasibility of this approach using one specific example for such a smoothness prior both for the reconstruction of Gaussian and log-normal fields and discuss the range of applicability of the chosen spectral smoothness prior. Our approach also allows for the estimation of a complete uncertainty matrix for the estimated power spectrum.

We first develop the formalism of the spectral smoothness prior for the case of Gaussian signal fields in Sec. 5.2, where we show how to derive the filter formulas of Enßlin & Frommert (2011) and Enßlin & Weig (2010) in a different way that easily accomodates an additional smoothness prior. After demonstrating the workings of the spectral smoothness prior, we use the Gibbs free energy formalism to derive filter formulas for the log-normal case and transfer the spectral smoothness results to this case in Sec. 5.3. We demonstrate the performance of our log-normal reconstruction algorithm in a variety of test cases in Sec. 5.3.2 and conclude in Sec. 5.4.

5.2 Reconstructing Gaussian fields with spectral smoothness

First, we need to introduce some basic assumptions and notation. We mainly follow the notation that is used throughout the literature on *information field theory* (e.g. Enßlin et al. 2009).

Throughout the paper, we assume that we are analyzing a set of data $d = (d_1, \ldots, d_r) \in \mathbb{R}^r$ that depends on an underlying signal field $s : \mathcal{M} \longrightarrow \mathbb{R}$, subject to additive noise $n = (n_1, \ldots, n_r) \in \mathbb{R}^r$,

$$d = f(s) + n. \tag{5.1}$$

Here, $\mathcal{M}$ is the discrete or continuous space or manifold that the signal is defined on. In Sec. 5.3 we will discuss the examples $\mathcal{M} = \mathcal{S}^1$, $\mathcal{S}^2$, and $\mathcal{T}^2$. In this section, we assume the relationship between signal field and data to be linear, described by a response operator R, so that

$$d = Rs + n. \tag{5.2}$$

In most applications, the response operator will include some instrumental or observational effects such as observations in specific locations of $\mathcal{M}$, convolutions of s with an instrumental response function, or a Fourier transform of the signal field. The only restriction that we make here is that the operation that generates the data from the signal has to be linear in the signal. Finally, we restrict ourselves to real-valued signals only in the interest of notational simplicity. All our results can be straightforwardly generalized to complex-valued signal fields.

For the noise n, we assume Gaussian statistics,

$$n \hookleftarrow \mathcal{G}(n, N), \tag{5.3}$$

described by a not necessarily diagonal covariance matrix N. Here, $\mathcal{G}(\phi, \Phi)$ denotes a multi-variate Gaussian probability distribution function with covariance Φ,

$$\mathcal{G}(\phi, \Phi) = \frac{1}{|2\pi\Phi|^{1/2}} \exp\left(-\frac{1}{2}\phi^\dagger \Phi^{-1} \phi\right). \tag{5.4}$$

We use the $\dagger$-symbol to denote a transposed (and in general complex conjugated) quantity and take the limit of an infinite-dimensional Gaussian whenever the argument is a continuous field, with the appropriate limit of the scalar product

$$\phi^\dagger \psi = \int_\mathcal{M} \mathrm{d}x\, \overline{\phi(x)}\, \psi(x) \ \ \forall\, \phi, \psi : \ \mathcal{M} \longrightarrow \mathbb{R}. \tag{5.5}$$

In this section we will deal with the case that the signal field s can also be regarded – or at least approximated – as a zero-mean Gaussian random field with covariance S,

$$s \hookleftarrow \mathcal{G}(s, S). \tag{5.6}$$

It can be straightforwardly shown (e.g. Enßlin et al. 2009) that the posterior mean m of the signal field under these assumptions is given by

$$m = \int \mathcal{D}s\, s\, \mathcal{P}(s|d) = Dj. \tag{5.7}$$

Here, $\int \mathcal{D}s$ denotes an integral over the configuration space of all possible signal realizations. The operator

$$D = \left(S^{-1} + R^\dagger N^{-1} R\right)^{-1} \tag{5.8}$$

is known as the *information propagator* and the field

$$j = R^\dagger N^{-1} d \tag{5.9}$$

is called *information source*.

In these formulas, the presence of the Gaussian signal prior described by the signal covariance serves as a means of regularization of the sought continuous signal field reconstruction which otherwise is under-determined when only constrained by the finite and noisy data set. However, in most physical applications, the signal covariance is not known a priori. The problem of reconstructing Gaussian fields with unknown signal covariances has in principle been solved (Enßlin & Frommert 2011; Enßlin & Weig 2010; Wandelt et al. 2004; Jasche et al. 2010b), and even an unknown noise covariance can be overcome (Oppermann et al. 2011b).

Enßlin & Weig (2010) use the formalism of minimum Gibbs free energy to derive filter formulas to be applied to the data set when the signal's covariance is unknown. We will recapture this formalism briefly in Sec. 5.3, where we employ it to reconstruct log-normal signal fields.

Under the assumption of statistical homogeneity and isotropy, the unknown signal covariance becomes diagonal in the harmonic eigenbasis, i.e. the Fourier basis for signals defined on Euclidean space and the spherical harmonics basis for signals defined on the sphere. In the following, we denote as $\vec{k}$ the vector of parameters determining one mode in the harmonic decomposition, i.e. $\vec{k} = (k_1, \ldots, k_n)$ for n-dimensional Euclidean space and $\vec{k} = (\ell, m)$ for the two-sphere, where ℓ is the angular momentum quantum number and m the azimuthal one. Furthermore, k shall stand for the scale of the harmonic component, i.e. $k = \sqrt{k_1^2 + \cdots + k_n^2}$ and $k = \ell$ for $\mathbb{R}^n$ and $\mathcal{S}^2$, respectively.

Due to the isotropy assumption, the diagonal of the signal covariance matrix, the signal's power spectrum P_k, depends only on the scale k and not on the specific mode $\vec{k}$, i.e.

$$S_{\vec{k}\vec{k}'} = \delta_{\vec{k}\vec{k}'} P_k. \tag{5.10}$$

Under these symmetry assumptions, the filter formulas derived by Enßlin & Weig (2010) can be written as

$$m = Dj, \tag{5.11}$$

$$P_k = \frac{1}{\alpha_k - 1 + \frac{\rho_k}{2}} \left(q_k + \frac{1}{2} \sum_{\{\vec{k}'|k'=k\}} \left(|m_{\vec{k}'}|^2 + D_{\vec{k}'\vec{k}'} \right) \right). \tag{5.12}$$

Note that the first of these equations is the same as in the case of a known signal covariance, Eq. (5.7), a generalized Wiener filter (Enßlin et al. 2009). Here, α_k and q_k are parameters used to determine the priors for the spectral coefficients P_k (see Sec. 5.3 for details) and

$$\rho_k = \sum_{\{\vec{k}'|k'=k\}} 1 \tag{5.13}$$

is the number of degrees of freedom on the scale k.

This formalism has been successfully applied for astrophysical reconstructions (Oppermann et al. 2011a, 2012a) and in a more abstract computational setting (Selig et al. 2012). In these applications, Eqs. (5.11) and (5.12) were simply iterated. However, they were supplemented with an additional ad-hoc smoothing step for the power spectrum as part of each iteration.

To illustrate the usefulness of a smoothing step, we set up a simple one-dimensional test-case. Here, we assume that our signal field is defined on the one-sphere $\mathcal{S}^1$, i.e. the interval $[0, 1)$ with periodic boundary conditions, which we discretize into 100 pixels. We make up a power spectrum for the signal of the form

$$P_k = P_0 \left(1 + \left(\frac{k}{k_0} \right)^2 \right)^{-\gamma/2}, \tag{5.14}$$

where we choose $P_0 = 0.2$, $k_0 = 5$, and $\gamma = 4$. Figures 5.1 and 5.2 show the power spectrum and the signal realization randomly drawn from it, respectively.
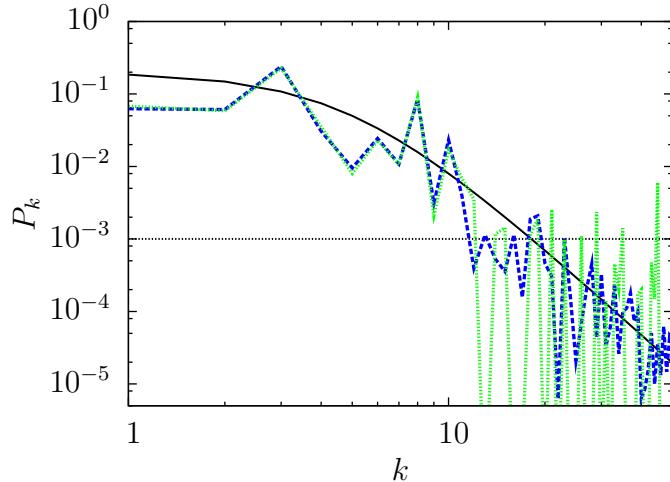
Figure 5.1: Power spectra of the one-dimensional test-case. The black solid line shows the theoretical power spectrum, the blue dashed line the power in the randomly drawn signal realization studied here, and the green dotted line shows the power spectrum reconstructed without smoothing. The horizontal dotted line indicates the noise power, given by $\sigma_n^2/100$.

Furthermore, we assume that we have measured the signal field in each of the 100 pixels once, subject to homogeneous and uncorrelated Gaussian noise with variance $\sigma_n^2 = 0.1$. In the formalism of Eq. (5.2), this corresponds to $R = \mathbb{1}$ and $N = \sigma_n^2 \mathbb{1}$. The data realization is also shown in Fig. 5.2.

We overplot in Figs. 5.1 and 5.2 the power spectrum and map reconstruction, respectively, obtained by iterating Eqs. (5.11) and (5.12) without smoothing. While the signal reconstruction suffers from some too pronounced small-scale fluctuations, the reconstructed power spectrum fluctuates wildly for $k > 10$ and is therefore not trustworthy at all on these scales. The scales on which the reconstructed power by chance peaks above the noise level are the ones that are especially misrepresented in the reconstructed signal field.

This example should serve to illustrate that some kind of spectral smoothing is necessary, especially in cases where one is interested in the power spectrum itself. In Fig. 5.2 we also show a comparison to the Wiener filter reconstruction, i.e. Eq. (5.7), using the true power spectrum. In the bottom panel of Fig. 5.2 it can be seen that the residuals, i.e. the differences between the true signal and its reconstruction, are reduced if a power spectrum is used that is closer to the true one. Smoothness is one aspect of the true power spectrum that can be used to constrain its reconstruction.

The two-point correlation function of the signal provides another way of looking at the smoothness property of its power spectrum. For a statistically homogeneous signal, the power spectrum is the Fourier transform of the correlation function and vice versa. Thus, a power spectrum that exhibits fluctuations on arbitrarily small scales in k-space corresponds to a signal that exhibits correlations over arbitrarily large scales in position space. Turning this argument around, any signal with correlations only over a finite range in position space or at least correlations that are decreasing rapidly with distance will have
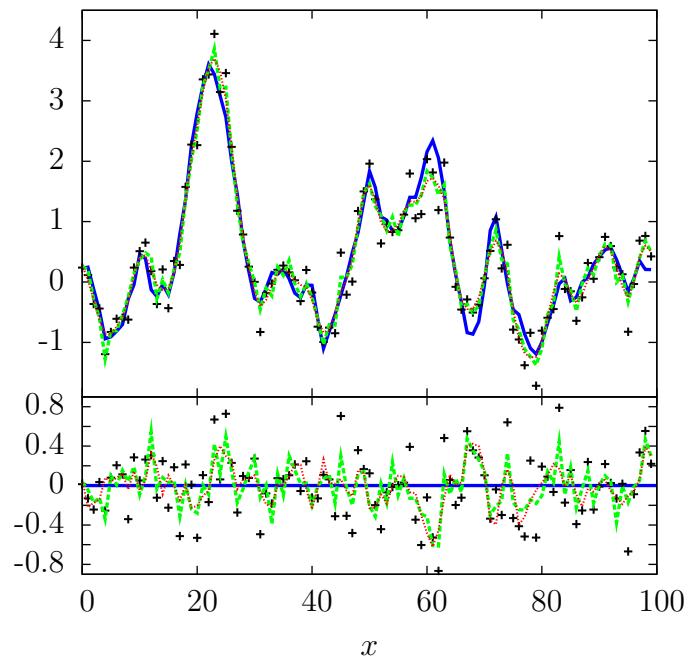
Figure 5.2: Signal reconstruction in the one-dimensional test-case. The blue solid line shows the randomly drawn signal realization, the crosses are noisy data points used to do the reconstruction, the green dashed line shows the signal reconstruction obtained without spectral smoothing and the red dotted line is the Wiener filter reconstruction. The lower panel shows the same with the true signal substracted.

a power spectrum that does not exhibit any features on arbitrarily small scales in k-space. Thus, the power spectrum can be expected to be smooth on the reciprocal scale of the typical correlation length scale.

In the remainder of this section, we will show how to incorporate a prior enforcing spectral smoothness into the formalism presented thus far.

5.2.1 Reconstruction as a combination of *posterior mean* and *maximum a posteriori*

Before incorporating a spectral smoothness prior into the derivation of Eqs. (5.11) and (5.12), we will rederive them in a way different from the one presented by Enßlin & Frommert (2011); Enßlin & Weig (2010).

As was already mentioned, Eq. (5.11) corresponds to the posterior mean of the signal under the assumption of a known covariance matrix, i.e. a known power spectrum. Since the posterior probability distribution is Gaussian in this case, its mean also maximizes the probability, so that m is the posterior mean and the maximum a posteriori solution at the same time,

$$m = \int \mathcal{D}s \, s \, \mathcal{P}(s|d, P) = \operatorname*{argmax}_{s} \left\{ \mathcal{P}(s|d, P) \right\}. \tag{5.15}$$

Here, we will show that Eq. (5.12) can be written as a maximum a posteriori solution as well, considering the signal-marginalized posterior

$$\mathcal{P}(P|d) = \int \mathcal{D}s \, \mathcal{P}(s, P|d). \tag{5.16}$$

In order to do this, we first have to define a prior for the power spectrum P. In accordance with Enßlin & Frommert (2011); Enßlin & Weig (2010), we choose independent inverse-gamma distributions for each spectral component P_k,

$$
\begin{aligned}
\mathcal{P}(P) &= \prod_k \mathcal{P}_{\mathrm{IG}}(P_k) \\
&= \prod_k \frac{1}{q_k \Gamma(\alpha_k - 1)} \left(\frac{P_k}{q_k} \right)^{-\alpha_k} \exp\left(-\frac{q_k}{P_k} \right).
\end{aligned}
\tag{5.17}
$$

Here, $\Gamma(\cdot)$ denotes the gamma function, q_k is a parameter defining the location of an exponential cut-off at low values of P_k, and α_k defines the slope of a power-law for large values of P_k. By tuning these parameters, the prior can be narrowed or widened according to the a priori knowledge about the power spectrum. Taking the limit $q_k \to 0$ and $\alpha_k \to 1$ turns the inverse-gamma distribution into Jeffreys prior which is flat on a logarithmic scale. In the examples presented in this paper we always take this limit in the final filter formulas.

In the following, we will work with the logarithmic spectral components

$$p_k = \log P_k. \tag{5.18}$$

The corresponding prior for these can be straightforwardly derived from the conservation of probability under variable transformations and reads

$$\mathcal{P}(p) = \mathcal{P}(P) \left| \frac{\mathrm{d}P}{\mathrm{d}p} \right|$$

$$= \prod_k \frac{q_k^{\alpha_k - 1}}{\Gamma(\alpha_k - 1)} \mathrm{e}^{-(\alpha_k - 1)p_k} \exp\left(-q_k \mathrm{e}^{-p_k}\right). \tag{5.19}$$

Using this prior, we can calculate the signal-marginalized probability of data and power spectrum, $\mathcal{P}(d, p)$, and its negative logarithm, the Hamiltonian

$$H(d, p) = -\log \mathcal{P}(d, p)$$

$$= -\log \int \mathcal{D}s\, \mathcal{G}(d - Rs, N)\, \mathcal{G}(s, S)\, \mathcal{P}(p)$$

$$= \frac{1}{2}\mathrm{tr}\left(\log S\right) - \frac{1}{2}\mathrm{tr}\left(\log D\right) - \frac{1}{2}j^\dagger D j$$

$$+ \sum_k \left((\alpha_k - 1)\, p_k + q_k \mathrm{e}^{-p_k}\right) + \mathrm{const.}, \tag{5.20}$$

where we have made use of the definitions (5.8) and (5.9) and have collected all terms that do not depend on the power spectrum into an unimportant additive constant. Using the spectral dependence of the signal covariance matrix in a statistically homogeneous and isotropic setting, Eq. (5.10), we can take the derivative of the Hamiltonian with respect to one p_k and equate it to zero, thus maximizing the posterior probability of the logarithmic power spectrum. The resulting equation is exactly Eq. (5.12) for $P_k = \mathrm{e}^{p_k}$ if one makes the identification $m = Dj$.

Thus we have shown that the filter formulas, Eqs. (5.11) and (5.12), can be derived as a combination of posterior mean for the signal reconstruction and maximum a posteriori for the power spectrum. This effectively means that we have made the approximation

$$m = \int \mathcal{D}s\, s\, \mathcal{P}(s|d)$$

$$= \int \mathcal{D}s \int \mathcal{D}p\, s\, \mathcal{P}(s|d, p)\mathcal{P}(p|d)$$

$$\approx \int \mathcal{D}s \int \mathcal{D}p\, s\, \mathcal{P}(s|d, p)\, \delta\left(p - p^{(\mathrm{MAP})}\right), \tag{5.21}$$

i.e. we have approximated the posterior probability distribution for the power spectrum with a delta distribution centered on its maximum.

It may be worth noting that in the formalism of the maximum a posteriori solution for the power spectrum, it is straightforward to derive a rough uncertainty estimate as well. The Hessian of the Hamiltonian gives the curvature of the posterior probability distribution and its inverse can thus be regarded as an uncertainty matrix. For the Hamiltonian given

in Eq. (5.20) we obtain

$$
\left.\frac{\partial^2 H(d,p)}{\partial p_k \partial p_{k'}}\right|_{p=p^{(\mathrm{MAP})}} = \left(\alpha_k - 1 + \frac{\rho_k}{2}\right)\delta_{kk'}
$$

$$
-\frac{1}{2}\mathrm{e}^{-p_k-p_{k'}}\sum_{\substack{\{\vec{q}\,|q=k\}\\\{\vec{q}'\,|q'=k'\}}}\left(2\,\Re\!\left(m_{\vec{q}}m_{\vec{q}'}^{*}D_{\vec{q}\vec{q}'}\right) + |D_{\vec{q}\vec{q}'}|^2\right)\Bigg|_{p=p^{(\mathrm{MAP})}},
\tag{5.22}
$$

where $\Re(\cdot)$ denotes the real part of a complex number.

5.2.2   Spectral smoothness priors

Here, we show how to incorporate a spectral smoothness prior into the formalism developed in the previous section. We do this by augmenting the inverse-gamma distributions previously assumed as the spectral prior with a probability distribution that enforces smoothness of the power spectrum, so that

$$
\mathcal{P}(p) = \mathcal{P}_{\mathrm{sm}}(p)\prod_k \mathcal{P}_{\mathrm{IG}}(p_k).
\tag{5.23}
$$

As an example, we choose a smoothness-enforcing prior of the shape

$$
\mathcal{P}_{\mathrm{sm}}(p) \propto \exp\left(-\frac{1}{2\sigma_p^2}\int \mathrm{d}(\log k)\left(\frac{\partial^2 \log P_k}{\partial(\log k)^2}\right)^2\right).
\tag{5.24}
$$

This prior is maximized by any power-law power spectrum and punishes deviations from such a shape. The strength of the punishment is determined by the parameter σ_p. Other useful shapes for a smoothness prior could contain the first logarithmic derivative, punishing steep spectra, or simple derivatives with respect to k, punishing abrupt changes in the power spectrum. To illustrate the meaning of such smoothness priors and point out possible caveats, we discuss a few specific cases in App. E.1.

The spectral smoothness prior, Eq. (5.24), can be written as a Gaussian in $p = \log P$,

$$
\mathcal{P}_{\mathrm{sm}}(p) \propto \exp\left(-\frac{1}{2}p^{\dagger}Tp\right),
\tag{5.25}
$$

where the linear operator T includes both the second derivative and the scaling constant σ_p^2. The detailed form of the operator T that we use in our calculations can be found in App. E.2.

Introducing this prior corresponds to adding the term $\frac{1}{2}p^{\dagger}Tp$ to the Hamiltonian in Eq. (5.20). Taking its derivative with respect to one p_k and equating the result with zero, we now get

$$
\mathrm{e}^{p_k} = \frac{q_k + \frac{1}{2}\sum_{\{\vec{k}'\,|k'=k\}}\left(\left|m_{\vec{k}'}\right|^2 + D_{\vec{k}'\vec{k}'}\right)}{\alpha_k - 1 + \frac{\rho_k}{2} + (Tp)_k}.
\tag{5.26}
$$
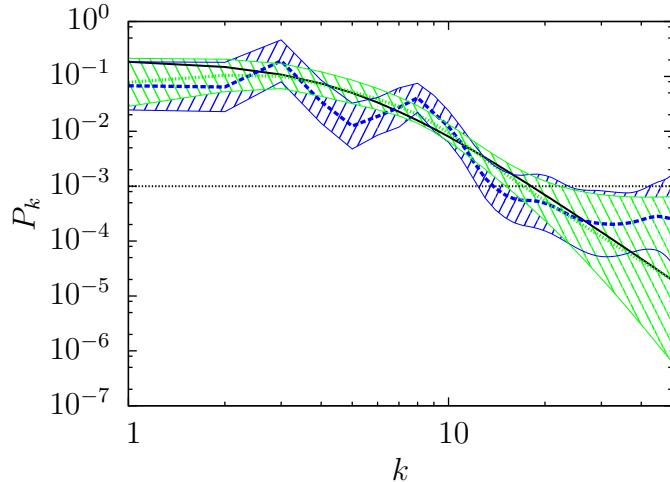
Figure 5.3: Power spectra of the one-dimensional test-case with smoothness prior. The black solid line shows the theoretical power spectrum, the blue dashed line the power spectrum reconstructed with a smoothness prior with $\sigma_p^2 = 1000$, and the green dotted line the reconstruction with a stiffer smoothness prior with $\sigma_p^2 = 10$. The hatched regions around the dashed and dotted lines are the corresponding one-sigma uncertainty regions estimated from the inverse Hessian of the Hamiltonian. The horizontal dotted line indicates the noise level.

The only point in which this equation differs from Eq. (5.12) is the extra term Tp in the denominator. As before, we can calculate the inverse Hessian of the Hamiltonian as an approximate uncertainty matrix. The Hessian is

$$\left. \frac{\partial^2 H(d,p)}{\partial p_k \partial p_{k'}} \right|_{p=p^{(\text{MAP})}} = \left(\alpha_k - 1 + \frac{\rho_k}{2} + (Tp)_k \right) \delta_{kk'} + T_{kk'}$$

$$\left. - \frac{1}{2} e^{-p_k - p_{k'}} \sum_{\substack{\{\vec{q}|q=k\} \\ \{\vec{q}'|q'=k'\}}} \left(2 \Re\left(m_{\vec{q}} m_{\vec{q}'}^* D_{\vec{q}\vec{q}'} \right) + |D_{\vec{q}\vec{q}'}|^2 \right) \right|_{p=p^{(\text{MAP})}} . \tag{5.27}$$

5.2.3 Test-cases

Using the same one-dimensional test-case as shown in Figs. 5.1 and 5.2, we perform a reconstruction using the spectral smoothness prior given in Eq. (5.24). The resulting power spectra using a moderate strength for the prior with $\sigma_p^2 = 1000$ and a strict smoothness prior with $\sigma_p^2 = 10$ are shown in Fig. 5.3. Clearly, the new reconstructions are a much better approximation to the shape of the theoretical power spectrum than the one shown in Fig. 5.1. Also, as expected, the power spectrum obtained when using the strict smoothness prior turns out smoother and more closely resembles a power law.

Also shown in Fig. 5.3 is an uncertainty estimate for the reconstructed power spectra, obtained from the Hessian of the Hamiltonian. Taking a one-sigma uncertainty estimate
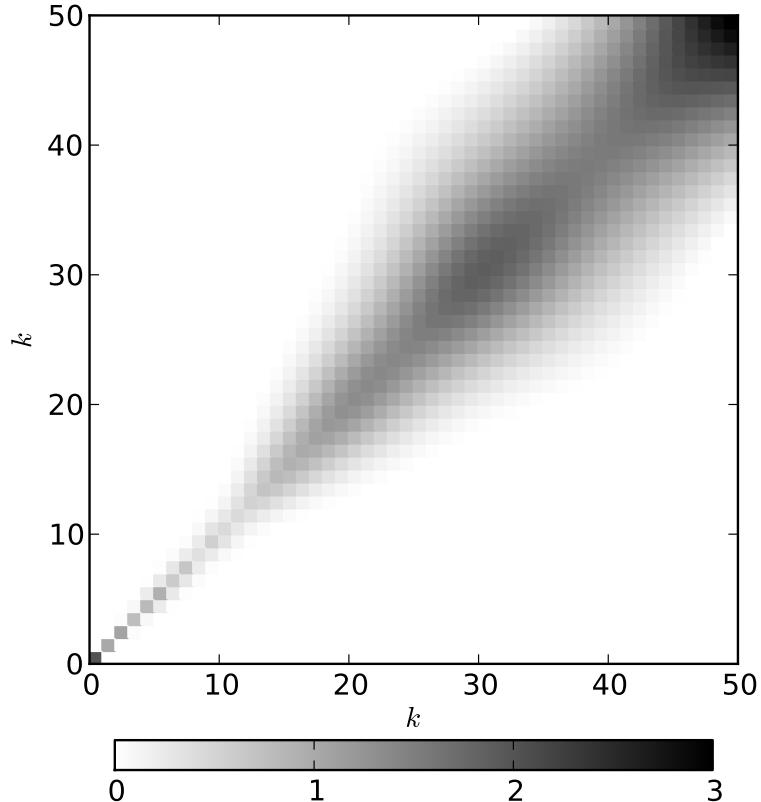
Figure 5.4: Inverse Hessian of the Hamiltonian with a smoothness prior with $\sigma_p^2 = 1000$. Plotted is the full matrix given by the inverse of Eq. (5.27), evaluated at $p = p^{(\mathrm{MAP})}$. The diagonal of this matrix can be translated into the uncertainty interval plotted in Fig. 5.3, however, that does not take into account the correlations across different k-modes that are visible in this plot.
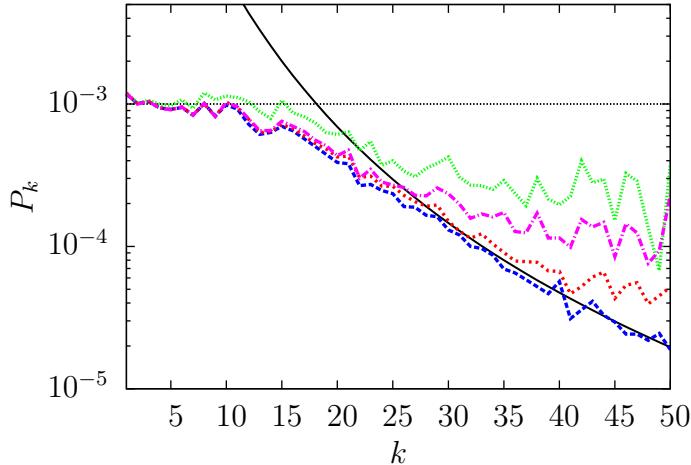
Figure 5.5: Residual power between true signals and their reconstructions averaged over 100 different realizations. The black solid line shows the theoretical power spectrum of the signals, given by Eq. (5.14) with $P_0 = 0.2$, $k_0 = 5$, and $\gamma = 4$. The dotted horizontal line indicates the noise level, corresponding to $\sigma_n^2 = 0.1$. The remaining lines show the residual power for different reconstruction schemes. From top to bottom, these are a reconstruction without any spectral smoothing (green dotted line), with an ad-hoc convolution of the power spectrum (magenta dash-dotted line), with the smoothness prior given by Eq. (5.24) with $\sigma_p^2 = 100$ (red short-dashed line), and using the correct power spectrum (blue long-dashed line).

for the logarithmic power spectrum components,

$$\Delta_p = \left[\mathrm{diag} \left(\left(\left. \frac{\delta^2 H(d,p)}{\delta p \, \delta p^\dagger} \right|_{p=p^{(\mathrm{MAP})}} \right)^{-1} \right) \right]^{1/2}, \tag{5.28}$$

we plot the uncertainty interval for the power spectra as $P_k = \exp\left(p_k^{(\mathrm{MAP})} \pm \Delta_p\right)$. The uncertainty interval of the power spectrum estimates calculated in this way can, however, be misleading. It should be noted that the uncertainty of the power on the different k-modes is correlated. To illustrate this, we plot the complete uncertainty matrix, i.e. the inverse Hessian of the Hamiltonian, for the case with $\sigma_p^2 = 1000$ in Fig. 5.4. It can be seen from this figure that the correlation of the uncertainty on different k-modes is especially large for small scales. Furthermore, the inverse Hessian is only a rough approximation to the uncertainties in the power spectrum estimation, which typically underestimates these.

In all following examples we will use a spectral smoothness prior with an intermediate stiffness, given by $\sigma_p^2 = 100$.

To study the improvement that the spectral smoothness prior brings for the reconstruction of the signal field, we draw 100 different signal realizations from the same power spectrum and add 100 different noise realizations. For each of these data sets, we perform the full reconstruction and then calculate the power of the difference between the recon-
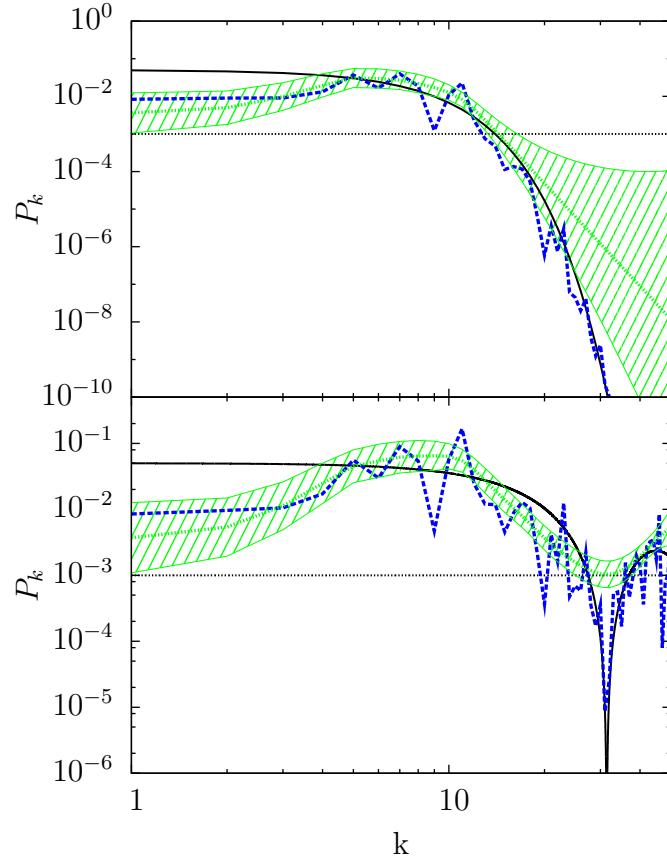
Figure 5.6: Power spectra of signals with correlation functions given by Eq. (E.4) (top panel) and Eq. (E.6) (bottom panel). The black solid lines show the theoretical power spectra, the blue dashed lines the power in the random field realization drawn from the theoretical power spectra, and the green dotted lines the reconstructed power spectra using the spectral smoothness prior given in Eq. (5.24) with $\sigma_p^2 = 100$. The parameters substituted into Eq. (E.4) are $\sigma = 0.2$ and $C_0 = 1/(4\sqrt{2\pi})$ and the ones used in Eq. (E.6) are $L = 0.2$ and $C_0 = 0.25$. The noise variance in both cases is $\sigma_n^2 = 0.1$. The hatched area indicates the uncertainty of the reconstructed power spectrum estimated from the inverse Hessian and the horizontal dotted line the noise level.
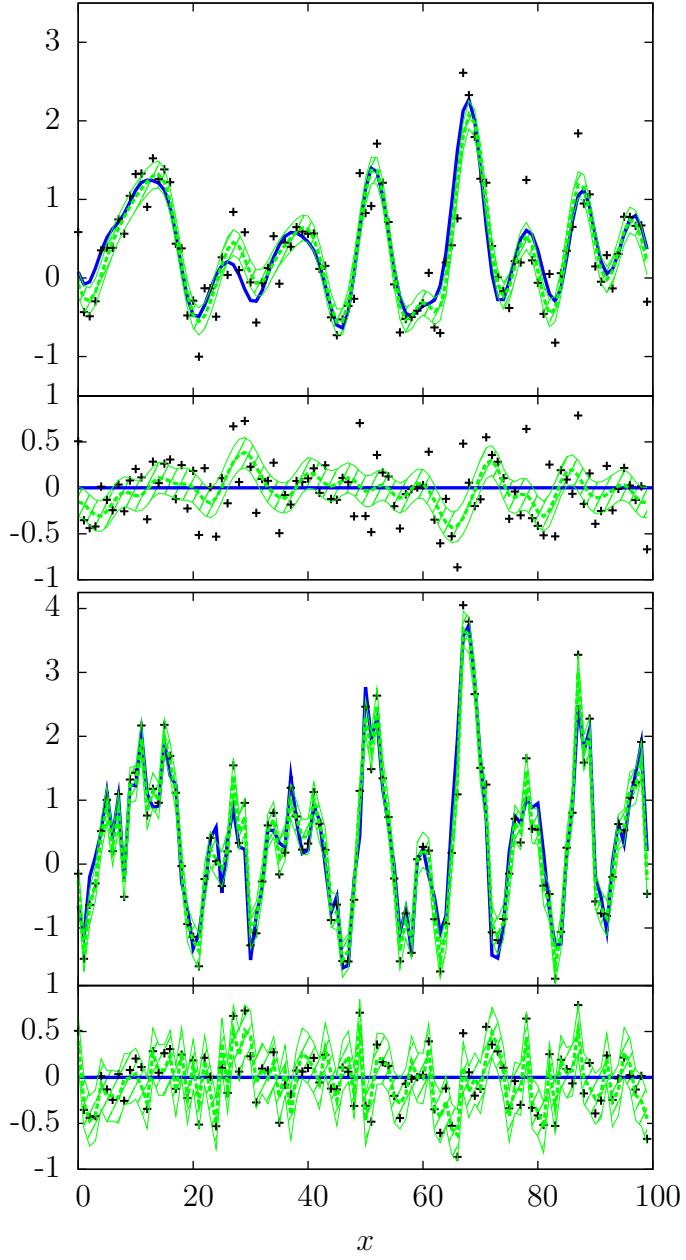
Figure 5.7: Signal fields corresponding to the power spectra shown in Fig. 5.6. The top panel shows a signal with a Gaussian two-point correlation function, Eq. (E.4), the bottom panel shows one with a triangular correlation function, Eq. (E.6). Shown in each plot are the signal realization s (blue solid line), the data d (crosses), the signal reconstruction m (green dashed line), and an estimate of the local one-sigma uncertainty of the reconstruction, given by $m \pm \mathrm{diag}(D)^{1/2}$ (hatched area). At the bottom of each panel, the same is plotted with the true signal subtracted.

structed field m and the true signal field s. In Fig. 5.5 we plot this power, averaged over the 100 realizations, for different reconstruction schemes. Under the assumption that the correct power spectrum is known a priori, the residual power is essentially given by the noise level on scales for which the signal is dominating and by the signal power on scales for which the noise is dominating. As can be seen from the plot, reconstructing the power spectrum without any spectral smoothing leads to a significantly increased residual power on the noise-dominated scales, while the quality of the reconstructions including the smoothness prior is close to the one of the Wiener filter reconstructions, showing only a slight excess in residual power on the smallest scales. We also plot the average residual power for reconstructions in which an ad-hoc smoothing step is applied to the power spectrum after each iteration of Eq. (5.12). This is implemented as a simple convolution with a Blackman window of width $\Delta_k = 9$. As can be seen in Fig. 5.5, this ad-hoc smoothing partly alleviates the problems of the power spectrum reconstruction but is clearly outperformed by the rigorous application of a smoothness prior.

Finally, we consider signal fields of the types discussed in App. E.1.2 and E.1.3, i.e. signals with Gaussian and triangular correlation functions given by Eqs. (E.4) and (E.6), respectively. As discussed in the appendix, these correlation functions lead to theoretical power spectra that are strongly suppressed by our spectral smoothness prior. Here, we investigate how serious this unwanted effect is in practice.

Shown in Figs. 5.6 and 5.7 are the results of the reconstruction for the power spectra and signal fields, respectively. We again use signals defined on an interval of length one, which we divide into 100 pixels, and choose $\sigma = 0.2$ and $C_0 = 1/(4\sqrt{2\pi})$ for the Gaussian correlation function, Eq. (E.4), and $L = 0.2$ and $C_0 = 0.25$ for the triangular correlation function, Eq. (E.6). In both cases, the noise variance is $\sigma_n^2 = 0.1$.

As expected, the features in the power spectra that are discouraged by the spectral smoothness prior are not well reconstructed. From Fig. 5.6 it is obvious that in the case with a Gaussian correlation function, the reconstructed and true power spectra deviate quite strongly on small scales. While the true power spectrum drops off rapidly, the reconstructed one stays comparatively flat. In the case of a triangular correlation function, the same can be said about the finite k-value where the true power spectrum drops to zero. While the drop is indeed represented by the signal realization, as can be seen from the dashed line in the lower panel, the reconstructed power spectrum stays relatively level.

However, the features in the power spectra that the filter fails to reconstruct are below the noise level, whereas the rise of the power spectrum in the bottom panel of Fig. 5.6 on the smallest scales, which is above the noise level, is indeed represented in the reconstructed power spectrum as well. Thus, the signal reconstruction does not suffer significantly, as can be seen in Fig. 5.7. An accurate reconstruction of the power spectrum is important mainly at the scales on which the signal-response and noise are of comparable magnitude. If the signal power is dominant, the reconstruction will follow the data closely, while it will smooth the data heavily in the opposite case of dominating noise, irrespective of the exact shape of the power spectrum. Thus, the reconstruction algorithm will in general perform well even in cases in which certain features in the power spectra are not allowed by the spectral smoothness prior. However, if the objective is an accurate reconstruction of the

power spectrum and any such features are suspected to be present, the spectral smoothness prior needs to be adapted to this situation.

5.3   Reconstructing log-normal fields

Now we turn to the problem of reconstructing a log-normal field. We define our signal field s to be the logarithm of this log-normal field, so that the prior probability distribution for s is again a Gaussian, described by a mean and a covariance S. For simplicity, we assume the prior mean to be zero. The data model then becomes

$$d = Re^s + n, \tag{5.29}$$

where we have again included additive Gaussian noise n and the application of the exponential function to the signal field is to be interpreted pointwise. Due to the non-linearity of the exponential function, the posterior of s,

$$\mathcal{P}(s|d, S) \propto \mathcal{G}(d - Re^s, N)\,\mathcal{G}(s, S), \tag{5.30}$$

is highly non-Gaussian, even when the signal covariance S is known. Adding a prior for S and marginalizing over it only makes the problem more complex,

$$\mathcal{P}(s|d) \propto \mathcal{G}(d - Re^s, N) \int\! \mathcal{D}S\, \mathcal{G}(s, S)\,\mathcal{P}(S). \tag{5.31}$$

The path we pursue here is to treat this probability distribution approximatively. In order to do so, we employ the formalism of minimum Gibbs free energy presented by Enßlin & Weig (2010). The basic idea is to approximate the posterior $\mathcal{P}(s|d)$ with a Gaussian, described by a mean m and a covariance D. An approximate Gibbs free energy can be calculated as a function of these quantities and it was shown by Enßlin & Weig (2010) that the optimal Gaussian approximation – in the sense of minimum Kullback-Leibler divergence – can be obtained by minimizing the approximate Gibbs free energy with respect to m and D.

The approximate Gibbs free energy is defined as

$$\tilde{G}(m, D) = \tilde{U}(m, D) - TS_{\mathrm{B}}(D). \tag{5.32}$$

The last term in this definition is the Boltzmann entropy

$$S_{\mathrm{B}}(D) = \frac{1}{2}\mathrm{tr}(1 + \log{(2\pi D)}), \tag{5.33}$$

which depends only on the covariance D. The other term in the approximate Gibbs free energy is the approximate internal energy

$$\tilde{U}(m, D) = \langle H(s, d)\rangle_{\mathcal{G}(s-m, D)}. \tag{5.34}$$

Here, $\langle \cdot \rangle_{\mathcal{G}(s-m,D)}$ denotes an expectation value calculated with respect to the Gaussian posterior approximation and $H(s,d) = -\log \mathcal{P}(s,d)$ is the Hamiltonian of the problem. Finally, the temperature T can be regarded as a tuning parameter that regulates the importance that is given to the region around the maximum of the full posterior and regions removed from the maximum. The limiting value $T = 0$ leads to the maximum a posteriori solution for the signal field. In the following, we choose the default value of $T = 1$ and refer the reader to Enßlin & Weig (2010); Iatsenko et al. (2012); Enßlin & Weig (2012) for an in-depth discussion of the temperature parameter.

Assuming independent inverse-gamma distributions as priors for the power spectrum components P_k, as we had already done in Eq. (5.17), the calculation of the Hamiltonian yields

$$
\begin{aligned}
H(s,d) &= -\log \left(\mathcal{P}(d|s)\, \mathcal{P}(s|P)\, \mathcal{P}(P) \right) \\
&= -\log \left(\mathcal{G}(d - R\mathrm{e}^s, N)\, \mathcal{G}(s, S) \prod_k \mathcal{P}_{\mathrm{IG}}(P_k) \right) \\
&= -j^\dagger \mathrm{e}^s + \frac{1}{2} \left(\mathrm{e}^s \right)^\dagger M \mathrm{e}^s \\
&\quad + \sum_k \left(\alpha_k - 1 + \frac{\rho_k}{2} \right) \log \left(q_k + \frac{1}{2} \sum_{\left\{ \vec{k}' | k' = k \right\}} \left| s_{\vec{k}'} \right|^2 \right) \\
&\quad + \mathrm{const.}
\end{aligned}
\tag{5.35}
$$

Here, we have again collected s-independent terms in an additive constant and introduced the abbreviation

$$
M = R^\dagger N^{-1} R.
\tag{5.36}
$$

In order to calculate the Gaussian expectation value of this Hamiltonian analytically, we expand the logarithm appearing in this expression in a power series around its expectation value

$$
\begin{aligned}
\tilde{q}_k &= \left\langle q_k + \frac{1}{2} \sum_{\left\{ \vec{k}' | k' = k \right\}} \left| s_{\vec{k}'} \right|^2 \right\rangle_{\mathcal{G}(s-m,D)} \\
&= q_k + \frac{1}{2} \sum_{\left\{ \vec{k}' | k' = k \right\}} \left(\left| m_{\vec{k}'} \right|^2 + D_{\vec{k}'\vec{k}'} \right)
\end{aligned}
\tag{5.37}
$$

as

$$\log \left(q_k + \frac{1}{2} \sum_{\{\vec{k}'|k'=k\}} |s_{\vec{k}'}|^2 \right)$$

$$\approx \log \left(\tilde{q}_k \right)$$

$$- \sum_{i=1}^{i_{\max}} \frac{(-1)^i}{i \tilde{q}_k^i} \left(q_k + \frac{1}{2} \left(\sum_{\{\vec{k}'|k'=k\}} |s_{\vec{k}'}|^2 \right) - \tilde{q}_k \right)^i. \tag{5.38}$$

We truncate this expansion after the first order, i.e. $i_{\max} = 1$. Note that our choice of $\tilde{q}_k$ ensures that the first-order term itself vanishes.

With this simplification, we can calculate the approximate Gibbs free energy to be

$$\tilde{G}(m, D) = \int_{\mathcal{M}} \mathrm{d}x \, j_x \mathrm{e}^{m_x + \frac{1}{2} D_{xx}}$$

$$+ \int_{\mathcal{M}} \mathrm{d}x \int_{\mathcal{M}} \mathrm{d}y \, \frac{1}{2} M_{xy} \mathrm{e}^{m_x + m_y + \frac{1}{2} D_{xx} + \frac{1}{2} D_{yy} + D_{xy}}$$

$$+ \sum_k \left(\alpha_k - 1 + \frac{\rho_k}{2} \right) \log \left(\tilde{q}_k \right)$$

$$- \frac{1}{2} \mathrm{tr}(1 + \log \left(2\pi D \right)). \tag{5.39}$$

To avoid confusion, we write out explicitly all integrals appearing here and in the following. Taking the functional derivatives with respect to m and D and equating them with zero yields two filter equations that determine m and D. In these equations, the right hand side expression of Eq. (5.12) appears. If we reidentify this expression with the spectral components P_k and write $S_{\vec{k}\vec{k}'} = \delta_{\vec{k}\vec{k}'} P_k$, the two filter equations become

$$m_x = \int_{\mathcal{M}} \mathrm{d}y \, S_{xy} \left[j_y \mathrm{e}^{m_y + \frac{1}{2} D_{yy}} \right.$$

$$\left. - \int_{\mathcal{M}} \mathrm{d}z \, M_{yz} \mathrm{e}^{m_y + m_z + \frac{1}{2} D_{yy} + \frac{1}{2} D_{zz} + D_{yz}} \right] \tag{5.40}$$

and

$$\left(D^{-1} \right)_{xy} = -j_x \mathrm{e}^{m_x + \frac{1}{2} D_{xx}} \delta_{xy}$$

$$+ \int_{\mathcal{M}} \mathrm{d}z \, M_{xz} \mathrm{e}^{m_x + m_z + \frac{1}{2} D_{xx} + \frac{1}{2} D_{zz} + D_{xz}} \delta_{xy}$$

$$+ M_{xy} \mathrm{e}^{m_x + m_y + \frac{1}{2} D_{xx} + \frac{1}{2} D_{yy} + D_{xy}}$$

$$+ \left(S^{-1} \right)_{xy}. \tag{5.41}$$

Together with Eq. (5.12), the last two equations fully determine the Gaussian approximation to the posterior. Solving them self-consistently gives an estimate m for the signal field and an estimate D for the corresponding uncertainty matrix. The corresponding approximate posterior mean estimate of the exponentiated field values then is

$$\langle \mathrm{e}^{s_x} \rangle_{\mathcal{P}(s|d)} \approx \langle \mathrm{e}^{s_x} \rangle_{\mathcal{G}(s-m,D)} = \mathrm{e}^{m_x + \frac{1}{2} D_{xx}}. \tag{5.42}$$

Of special interest for many applications is the case in which the matrix $M = R^\dagger N^{-1} R$ is diagonal, e.g. when the noise contributions to the individual data points are uncorrelated and the response is purely local. In this case the filter equations simplify somewhat to

$$m_x = \int_{\mathcal{M}} \mathrm{d}y\, S_{xy} \left[j_y \mathrm{e}^{m_y + \frac{1}{2} D_{yy}} - M_{yy} \mathrm{e}^{2m_y + 2D_{yy}} \right] \tag{5.43}$$

and

$$\left(D^{-1} \right)_{xy} = \left[-j_x \mathrm{e}^{m_x + \frac{1}{2} D_{xx}} + 2 M_{xx} \mathrm{e}^{2m_x + 2D_{xx}} \right] \delta_{xy}$$
$$+ \left(S^{-1} \right)_{xy}. \tag{5.44}$$

5.3.1   Spectral smoothness in the log-normal case

In Sec. 5.2.1 we had seen that in the Gaussian case the full reconstruction with unknown power spectrum can be regarded as a combination of the posterior mean reconstruction under the assumption of a known power spectrum and the estimation of this power spectrum as the one that maximizes its posterior probability. In the case of a log-normal field, the posterior mean cannot be calculated analytically, even if the power spectrum is assumed to be known. However, it is interesting to note that employing the formalism of minimum Gibbs free energy to the log-normal reconstruction problem with known power spectrum, one arrives exactly at the formulas derived in the previous section, Eqs. (5.40) and (5.41).

Thus, the full reconstruction, consisting of Eqs. (5.40), (5.41), and (5.12) can again be regarded as a combination of the calculation of the posterior mean for the signal under the assumption of a power spectrum and the estimation of the power spectrum according to Eq. (5.12). From this viewpoint, the inclusion of a smoothness prior for the power spectrum is trivial. We simply replace the power spectrum estimation step according to Eq. (5.12) with the one derived in Sec. 5.2.2, i.e. with Eq. (5.26), just as we had done in the Gaussian case.

5.3.2   Test cases

In this section, we present some test cases for the theory developed so far on the reconstruction of log-normal fields. We study one-dimensional tests with differing degrees of non-linearity and differing noise-levels, as well as two two-dimensional test-cases. For simplicity, we assume an ideal local response, $R = \mathbb{1}$, except in the last example, where we study the effects of an observational mask. The noise is assumed to be uncorrelated and
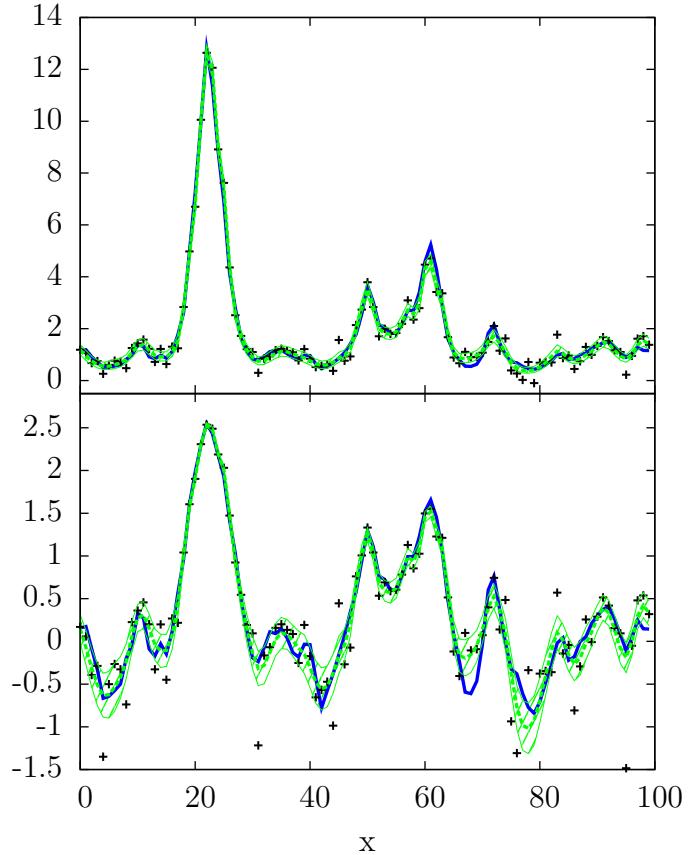
Figure 5.8: One-dimensional log-normal reconstruction in a mildly non-linear low-noise case. The power spectrum from which the signal realization was drawn is given by Eq. (5.14) with $P_0 = 0.1$, $k_0 = 5$, and $\gamma = 4$. The noise variance is $\sigma_n^2 = 0.1$. The top panel shows the exponentiated signal field e^s (blue solid line), the data d (crosses), the reconstruction $\mathrm{e}^{m+\frac{1}{2}\mathrm{diag}(D)}$ (green dashed line), and the uncertainty interval of the reconstruction, given by $\mathrm{e}^{m+\frac{1}{2}\mathrm{diag}(D)\pm(\mathrm{diag}(D))^{1/2}}$ (hatched region). The lower panel shows the signal field s (blue solid line), the logarithm of the data $\log(d)$ (crosses), the reconstruction m (green dashed line), and the uncertainty interval for m, given by $m \pm (\mathrm{diag}(D))^{1/2}$. In the lower panel, only data points for which $\log d$ is greater than -1.5 are shown.
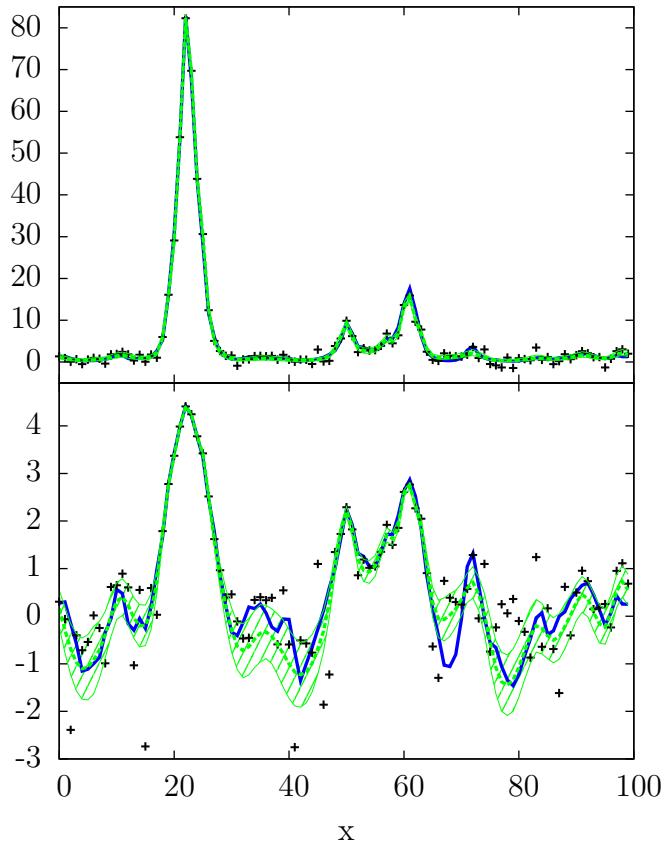
Figure 5.9: One-dimensional log-normal reconstruction in a highly non-linear low-noise case. The plotted quantities are the same as in Fig. 5.8. The parameters describing the power spectrum are $P_0 = 0.3$, $k_0 = 5$, and $\gamma = 4$ and the noise variance is $\sigma_n^2 = 1$. In the lower panel, only data points for which $\log d$ is greater than -3 are shown.
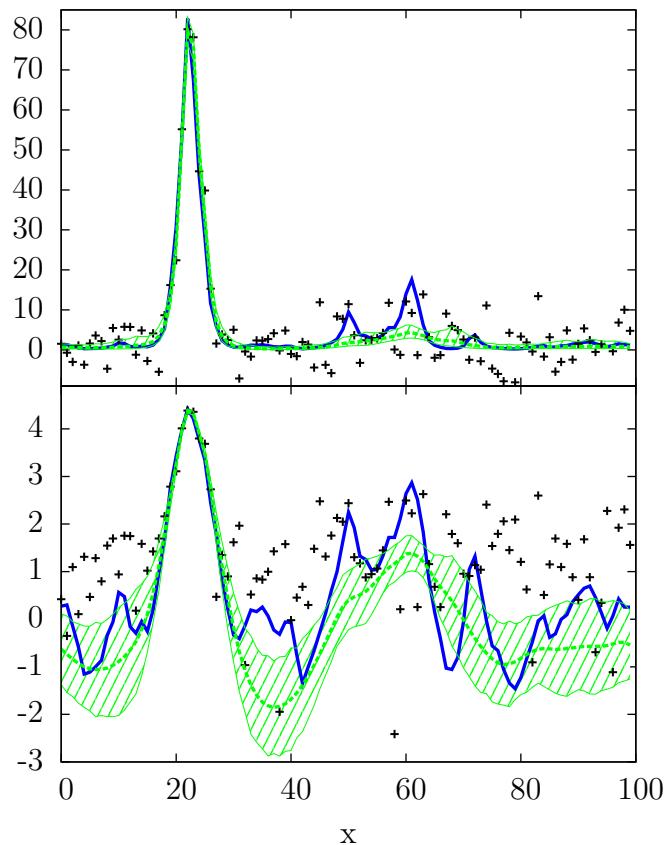
Figure 5.10: One-dimensional log-normal reconstruction in a highly non-linear high-noise case. The plotted quantities are the same as in Fig. 5.8. The parameters describing the power spectrum are $P_0 = 0.3$, $k_0 = 5$, and $\gamma = 4$ and the noise variance is $\sigma_n^2 = 25$. In the lower panel, only data points for which $\log d$ is greater than -3 are shown.
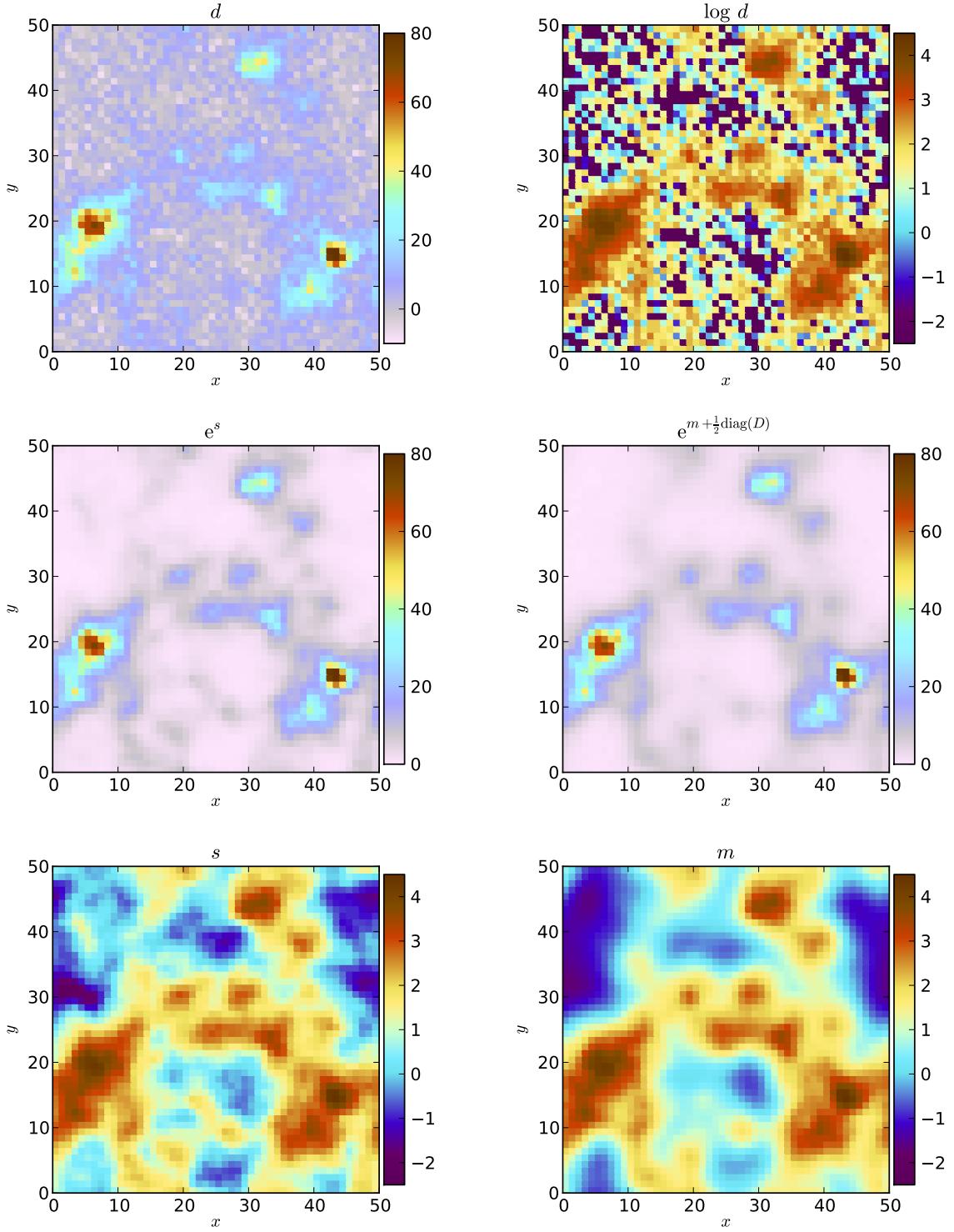
Figure 5.11: Log-normal reconstruction example in a toroidal setting. The signal's power spectrum is given by Eq. (5.14) with $P_0 = 0.3$, $k_0 = 2$, and $\gamma = 5$. The noise variance is $\sigma_n^2 = 10$. The top row shows the data set d (left) and its logarithm (right). In the logarithmic version, pixels with negative data values are plotted as dark blue. The middle row shows the exponentiated signal field e^s (left) and its reconstruction, given by $e^{m+\frac{1}{2}\text{diag}(D)}$, (right). The corresponding non-exponentiated quantities s (left) and m (right) are shown in the bottom row.
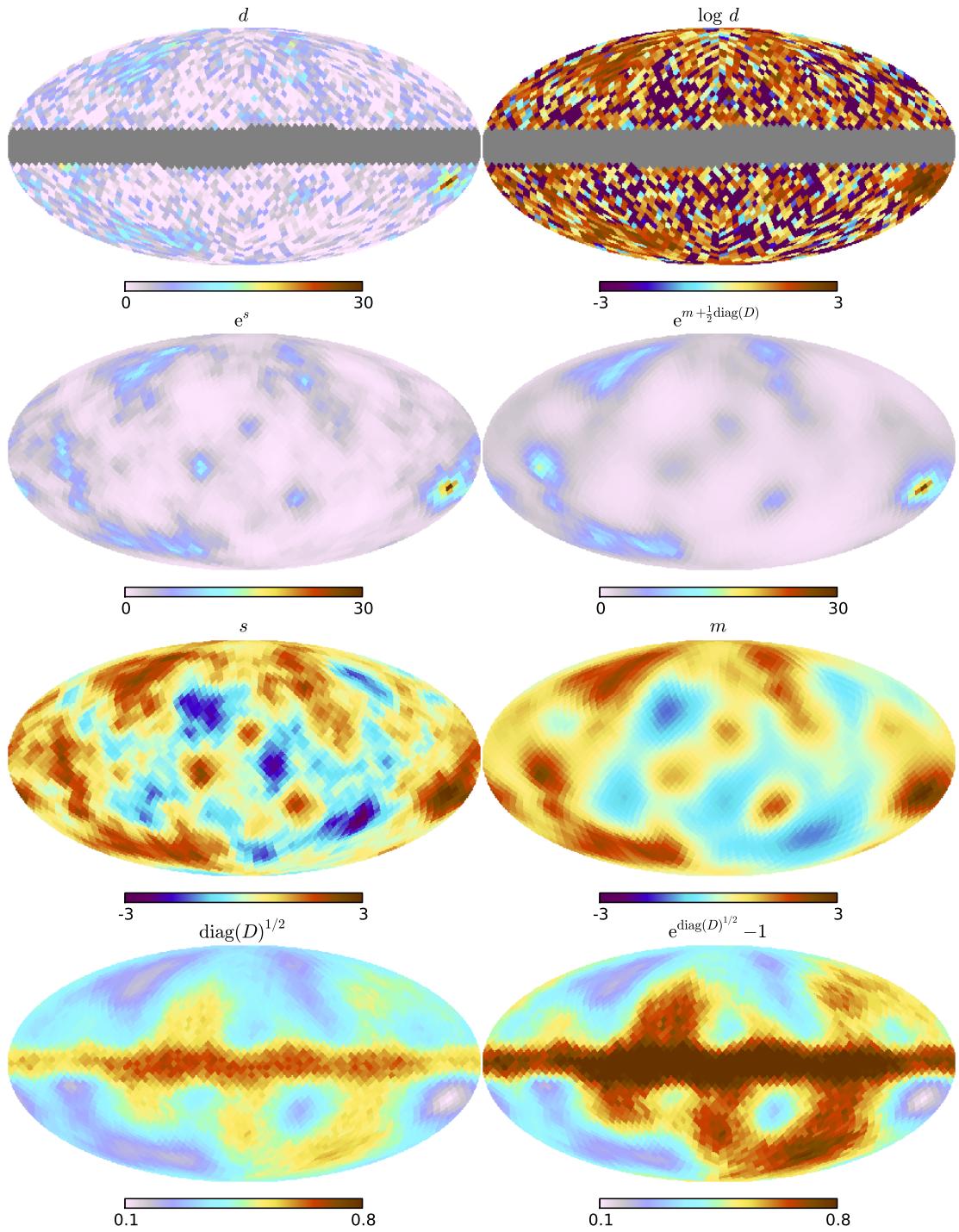
Figure 5.12: Log-normal reconstruction example on a two-sphere. The quantities plotted in the top three rows are the same as in Fig. 5.11. The bottom row shows the uncertainty of the signal reconstruction given by $\mathrm{diag}(D)^{1/2}$ (left side) and the fractional uncertainty of the reconstruction of the exponentiated field given by $\mathrm{e}^{\mathrm{diag}(D)^{1/2}} - 1$ (right side). The power spectrum parameters are $P_0 = 0.3$, $k_0 = 5$, and $\gamma = 4$ and the noise variance is $\sigma_n^2 = 10$. In the top row, pixels without measurements are plotted gray.

homogeneous, $N = \sigma_n^2 \mathbb{1}$. For the smoothness prior, we use the one discussed in Sec. 5.2.2 and choose $\sigma_p^2 = 100$.

First, we discuss again the case of a signal on the one-sphere $\mathcal{S}^1$, i.e. the interval $[0, 1)$ with periodic boundary conditions, here discretized into 100 pixels. Shown in Figs. 5.8-5.10 are three different cases. In each of these cases, the signal realization is the same as in Sec. 5.2, only with a differing normalization of the power spectrum. In Fig. 5.8 we show a mildly non-linear case with $P_0 = 0.1$ and in Figs. 5.9 and 5.10 a highly non-linear case with $P_0 = 0.3$. The noise variance σ_n^2 is 0.1 in Fig. 5.8, 1 in Fig. 5.9, and 25 in Fig. 5.10. In the lower panels of these three figures, we show the signals and their reconstructions and in the upper panels the exponentiated fields, e^s, as well as our posterior mean estimate for these, $\langle e^s \rangle_{\mathcal{P}(s|d)} \approx e^{m + \frac{1}{2} \text{diag}(D)}$.

In the mildly non-linear case the reconstruction is a reasonably good approximation to the true signal both in regions of high signal values and regions of low signal values. However, it is apparent that the quality of the reconstruction is slightly higher in the former regions. This is due to the homogeneity of the noise statistics that we have assumed. Since our signal response, given by Re^s, depends non-linearly on the signal, the noise impact is lower in regions of higher signal values and hence the signal inference is less demanding in these regions. From Fig. 5.8 it is apparent that this effect is well represented by the point-wise uncertainty of the reconstruction, given by $\text{diag}(D)$. As can be seen from Fig. 5.9, the effect becomes more pronounced for higher degrees of non-linearity, i.e. larger signal variances on a logarithmic scale.

The highly nonlinear case with high noise level, depicted in Fig. 5.10, exhibits a point-wise uncertainty estimate for the reconstruction that can clearly not be interpreted as a 68% confidence interval. This is due to a known problem of the filter discussed in Sec. 5.2 and by extension also of the filter derived in this section. As discussed by Enßlin & Frommert (2011), the filter exhibits a perception threshold. This means that if the signal-response-to-noise ratio is lower than a certain threshold on a given scale, then the filter will not reconstruct any power on this scale. Our usage of the spectral smoothness prior partly alleviates this porblem in that it prevents the power of individual scales to drop to zero. However, the reconstructed power on the noise-dominated scales will in general still be too low. This directly affects the estimate for the reconstruction's uncertainty, given by $D = (S^{-1} + M)^{-1}$, which tends toward zero in the limit of zero power on all scales.

Furthermore, the lack of power on all but the few signal-dominated scales can lead to ringing effects, i.e. prominent signal-dominated features are well reconstructed and extrapolated periodically into the noise-dominated regions. The deep trough in the signal reconstruction that can be seen in the lower panel of Fig. 5.10 around pixel number 35 is most likely due to this effect. The high degree of non-linearity acts to reinforce this effect. As can be seen in the top panel of Fig. 5.10, the effect of the trough onto the exponentiated reconstruction is almost intangible.

It is certainly mandatory to keep the potential problems of ringing and an underestimated uncertainty in mind when reconstructing a field which is swamped by noise in the better part of its domain. However, as can be seen from Fig. 5.10, the main features of the signal field will still be reconstructed reliably even in such an unfavorable case.

In these test cases, we have chosen the one-dimensional interval as domain of the log-normal signal field mainly for illustrative purposes. The algorithm is, however, versatile in that it can operate on virtually any space[1]. To illustrate this point, we close this section with two two-dimensional examples.

In Fig. 5.11, we show an example for a signal defined on the two-torus $\mathcal{T}^2 = \mathcal{S}^1 \times \mathcal{S}^1$, i.e. a segment of $\mathbb{R}^2$ with periodic boundary conditions. The periodicity here and in the earlier one-dimensional examples is a consequence of our use of Fast Fourier Transform routines. It can be approximately overcome by artificially extending the interval or the section of $\mathbb{R}^2$ for more than the typical correlation length of the signal while leaving the data unchanged when doing the reconstruction. Thus, the influence of the reconstruction on one side of the interval or rectangle will have negligible influence on the reconstruction on the other side.

We discretize the two-torus and its corresponding Fourier space into 50×50 pixels. The rectangular orientation of the Fourier pixels leads to many different scales k being represented, each, however, only by a few pixels, corresponding to a few different Fourier vectors $\vec{k}$. To avoid dealing with all these scales individually, we bin the scales logarithmically into 17 bins. We replace all summations over Fourier components with a certain scale appearing in the filter formulas with summations over all Fourier components whose scale falls within one bin. Thus, we reconstruct the power for each bin instead of for each scale that is represented in the rectangular Fourier grid. The signal is again drawn from the power spectrum given by Eq. (5.14) with $P_0 = 0.3$, $k_0 = 2$, and $\gamma = 5$. The noise level is chosen as $\sigma_n^2 = 10$. Signal, data, and reconstruction are shown in Fig. 5.11.

The toroidal example exhibits essentially the same features that we had already seen in the one-dimensional case. While the reconstruction is generally in good agreement with the true underlying signal, it is less accurate in the regions of small signal values where the signal-response-to-noise ratio is far worse due to the quite high degree of non-linearity.

In Fig. 5.12, we show an example for a reconstruction on the two-sphere $\mathcal{S}^2$. This example has special relevance for astronomical applications since one can interpret the sky as a two-sphere and therefore any astrophysical signals without distance information will be defined on this manifold. We generate a mock signal from the power spectrum given in Eq. (5.14) with $P_0 = 0.3$, $k_0 = 5$, and $\gamma = 4$. The noise level is $\sigma_n^2 = 10$. We discretize the sphere using the HEALPIX[2] package (Górski et al. 2005) with a resolution parameter of $N_{\text{side}} = 16$ leading to $3\,072$ pixels in total. The power spectrum components P_k are in this case the components of an angular power spectrum, often denoted as C_ℓ.

In this last case study, we replace the trivial response $R = \mathbb{1}$ with a projection onto part of the sphere, effectively masking 600 pixels around the equator for which we assume that no measurements have been taken. This resembles a typical situation in extragalactic astronomy, where observations through the Galactic plane are not possible due to the obscuration by the Milky Way.

---

[1]We are using the NIFTY package (Selig in prep.) in our implementation. This makes changing the domain of the signal field trivial, requiring only minuscule changes to the code.

[2]The HEALPIX package is available from `http://healpix.jpl.nasa.gov/`.

Fig. 5.12 shows the signal field, data, and reconstruction both in the exponentiated and in the linear version. In the panels showing the reconstructed signal field, it can be nicely seen how the algorithm is able to extrapolate into the gap region from the data on the boundary. This is possible due to the knowledge of the correlation structure that was inferred from the same data set. Also shown in Fig. 5.12 are the pixel-wise uncertainty estimate of the signal field's reconstruction, given by $\mathrm{diag}(D)^{1/2}$, and the fractional uncertainty of the reconstructed exponentiated signal, given by $\mathrm{e}^{\mathrm{diag}(D)^{1/2}} - 1$, which is approximated well by $\mathrm{diag}(D)^{1/2}$ in most regions. It can be seen that the uncertainty tends to be higher in the regions of low signal values, as was the case in the one-dimensional examples, and in the region around the equator which lacks observations. This is to be expected, since the only constraint on the signal in this region comes from extrapolations from neighboring regions using the signal's correlation structure inferred from the data.

5.4   Summary and Conclusions

We have developed an algorithm to infer log-normal random fields from noisy measurement data. The log-normal model was chosen due to its wide range of applications in astrophysics and other fields. The reconstruction method uses the correlation structure of the log-normal field to differentiate between features in the data that are due to noise and such that are due to variations in the true underlying field. This correlation structure, determined by the field's power spectrum, is, however, in general not known a priori. We have therefore extended the theory for simultaneous reconstructions of a field and its power spectrum that was developed and applied successfully for Gaussian random fields in the past (Enßlin & Frommert 2011; Enßlin & Weig 2010; Oppermann et al. 2011a,b, 2012a; Selig et al. 2012) to log-normal fields.

An additional feature of our reconstruction method is the use of a smoothness prior for the power spectrum. We have suggested to employ a prior based on the second double-logarithmic derivative of the power spectrum and shown that it is well suited to handle a large variety of cases. Having investigated possible pitfalls associated with the usage of such a prior from theoretical as well as practical viewpoints, we should stress that the derivation of the filter formulas laid out here does not depend on the specific form of the spectral smoothness prior. In cases in which the prior we employed here cannot be expected to yield satisfactory results it should simply be replaced by a different one.

The algorithm we have derived depends in no way on the space on which the signal field is to be reconstructed. We have demonstrated this by showing examples of reconstructions of mock signals on a one-dimensional interval, a flat two-dimensional space, and a spherical space. We have discussed the performance of the algorithm in these scenarios and pointed out possible caveats when dealing with very low signal-response-to-noise ratios.

In these application examples, we have assumed that the observational data simply represent the underlying log-normal field, subject to additive noise. Furthermore, we have illustrated the ability of the algorithm to extrapolate from the given data in a test case in which these data were assumed to be incomplete, thus demonstrating the power of the usage

of the correlation information contained in the data. The derivation of the filter formulas, however, is even more general. It allows for an arbitrary linear relationship between log-normal field and data, described by a response matrix R. The resulting formulas include the general response matrix which can e.g. represent an incomplete observation, a convolution, or a Fourier transformation. We have also allowed for a general noise covariance matrix, thus including cases of heteroscedastic or correlated noise.

This makes the algorithm widely applicable. Applications that we have in mind include for example the study of diffuse Galactic emission components at radio frequencies and reconstructions of emissivity fields across galaxy clusters from interferometric observations.

Chapter 6

Conclusions and outlook

We have developed several new statistical inference techniques and applied them in studies of the Galaxy. Our studies of magnetic helicity are hindered by the degeneracy between magnetic features and features in the density of free thermal electrons. Thus, we are only able to conclude that either the Galactic magnetic field is non-helical, or the thermal electron density varies on scales that are small enough to make a detection of magnetic helicity with the employed technique impossible.

After developing the *extended critical filter* technique that enables reconstructions in the presence of an uncertain amount of uncertainty in the data, we have applied it to the largest available catalog of Faraday rotation observations of extragalactic sources, which we have newly assembled here. The resulting map is not only the most detailed and most highly resolved all-sky map of the astrophysical Faraday effect that exists so far, it is also the only one in which the extragalactic contributions are largely filtered out. Thus, it is ideally suited to constrain the Galactic magnetic field configuration and to subtract the Galactic part of the Faraday rotation in cases in which the extragalactic contributions are of interest. The statistical results of our reconstruction have opened up multiple pathways for Galactic magnetic field studies as well.

Finally, we have left the statistical safe haven of Gaussianity to explore more complicated fields. With a view on Galactic emission maps, we have developed a technique to reconstruct log-normal fields with unknown power spectra. The promising results are now awaiting application to observational data.

One future direction of research is the application of the log-normal reconstruction technique for emission from different physical processes within the Milky Way. The different processes can be disentangled by using their different frequency dpendence. Using observations at different wavelengths one can thus separate many emission components from each other. Our newly developed technique will improve this methodology by making use of the correlation structure of each component and by the realistic description of the different emission components with a log-normal probability distribution. Making use also of cross-correlations between the individual components will be the focus of future work. Such a technique will be of paramount importance for observations of the cosmic microwave background and its foregrounds as conducted e.g. by the *Planck* satellite.

On the magnetic field front, the relationship between different observables, such as polarized synchrotron emission, polarized dust emission, and Faraday rotation, and the opportunities that it opens for studies of statistical aspects of the Galactic magnetic field clearly warrants more research. Ultimately, however, one will always be limited by the degeneracy of the magnetic field and the density of the particles that cause the observable effect. To obtain a truly reliable model of the magnetic field of the Milky Way, one will therefore have to do inference on all these density distributions and the magnetic field at the same time. This is of course a herculean task that will only become feasible with the flood of new highly precise data that will become available thanks to a large number of radio astronomical surveys that are currently being conducted or planned.

Appendix A

Power spectra for real and complex fields

A.1 Definition of power spectrum

The power spectrum of a signal s is the diagonal of its covariance matrix S in its eigenbasis. For statistically isotropic signals on the sphere $\mathcal{S}^2$ this eigenbasis is given by the spherical harmonics and the angular power spectrum C_ℓ is independent of m, i.e.

$$S_{(\ell m)(\ell' m')} = \langle s_{\ell m} s^*_{\ell' m'} \rangle = \delta_{\ell\ell'} \delta_{mm'} C_\ell. \tag{A.1}$$

Here, the signal components in the spherical harmonics basis, $s_{\ell m}$, are complex numbers and the asterisk denotes complex conjugation.

Here, we discuss fields on $\mathcal{S}^2$. In Euclidean space, analogous thoughts are valid if the spherical harmonics components are replaced by Fourier components, i.e. $\ell \longrightarrow |k|$ and $m \longrightarrow \frac{k}{|k|}$.

A.2 Complex signals

Instead of representing the signal by complex coefficients $s_{\ell m}$, one can equally well represent it by twice as many real numbers $s_{\ell m r}$, where $\ell = 0, 1, \ldots$, $m = -\ell, \ldots, \ell$, and $r \in \{\text{"real"}, \text{"imaginary"}\}$. Again, the signal covariance matrix is diagonal in this basis,

$$S_{(\ell m r)(\ell' m' r')} = \langle s_{\ell m r} s_{\ell' m' r'} \rangle = \delta_{\ell\ell'} \delta_{mm'} \delta_{rr'} \alpha_\ell \tag{A.2}$$

The complex phases of $s_{\ell m}$ have to be uniformly distributed for a statistically isotropic signal, therefore the real and imaginary parts are uncorrelated and their variances are the same. In order to calculate the diagonal entries α_ℓ, we substitute Eq. (A.2) into Eq. (A.1),

yielding

$$
\begin{aligned}
S_{(\ell m)(\ell' m')} &= \langle (\mathrm{Re}\,(s_{\ell m}) + i\mathrm{Im}\,(s_{\ell m})) \, (\mathrm{Re}\,(s_{\ell' m'}) - i\mathrm{Im}\,(s_{\ell' m'})) \rangle \\
&= \langle \mathrm{Re}\,(s_{\ell m}) \, \mathrm{Re}\,(s_{\ell' m'}) \rangle + \langle \mathrm{Im}\,(s_{\ell m}) \, \mathrm{Im}\,(s_{\ell' m'}) \rangle \\
&\quad - i \, \langle \mathrm{Re}\,(s_{\ell m}) \, \mathrm{Im}\,(s_{\ell' m'}) \rangle + i \, \langle \mathrm{Im}\,(s_{\ell m}) \, \mathrm{Re}\,(s_{\ell' m'}) \rangle \\
&= \delta_{\ell \ell'} \delta_{m m'} 2 \alpha_\ell \\
&= \delta_{\ell \ell'} \delta_{m m'} C_\ell .
\end{aligned}
\tag{A.3}
$$

From this we can read off $\alpha_\ell = \frac{C_\ell}{2}$.

A.3 Real signals

For real signals, the condition $s_{\ell,-m} = s_{\ell m}^*$ holds. Therefore the eigenbasis of the covariance matrix is changed with respect to the complex case. The correlation of the coefficients $(s_{\ell m r})_{\ell m r}$ becomes

$$
S_{(\ell m r)(\ell' m' r')} = \delta_{\ell \ell'} \left(\delta_{m m'} \delta_{r r'} + \delta_{m,-m'} \delta_{r r'} \delta_{r,\text{``real''}} - \delta_{m,-m'} \delta_{r r'} \delta_{r,\text{``imag''}} \right) \beta_\ell .
\tag{A.4}
$$

Substituting this into Eq. (A.1) gives

$$
\begin{aligned}
S_{(\ell m)(\ell' m')} &= \delta_{\ell \ell'} \left(\delta_{m m'} + \delta_{m,-m'} + \delta_{m m'} - \delta_{m,-m'} \right) \beta_\ell \\
&= \delta_{\ell \ell'} \delta_{m m'} 2 \beta_\ell \\
&= \delta_{\ell \ell'} \delta_{m m'} C_\ell .
\end{aligned}
\tag{A.5}
$$

So again, we read off $\beta_\ell = \frac{C_\ell}{2}$.

This signal covariance matrix has the two Eigenvalues $2\beta_\ell = C_\ell$ and 0. The corresponding eigenspaces are spanned by vectors of the type

$$
v_{\ell m}^{(1)} = \hat{e}_{\ell, m, \text{``real''}} + \hat{e}_{\ell, -m, \text{``real''}}
$$

and

$$
v_{\ell m}^{(2)} = \hat{e}_{\ell, m, \text{``imag''}} - \hat{e}_{\ell, -m, \text{``imag''}}
$$

for the eigenvalue C_ℓ and by vectors of the type

$$
v_{\ell m}^{(3)} = \hat{e}_{\ell, m, \text{``real''}} - \hat{e}_{\ell, -m, \text{``real''}}
$$

and

$$
v_{\ell m}^{(4)} = \hat{e}_{\ell, m, \text{``imag''}} + \hat{e}_{\ell, -m, \text{``imag''}}
$$

for the eigenvalue 0, where $\ell = 0, 1, \dots$ and $m = 0, \dots, \ell$.

Note that the eigenspaces of the second kind exist only for complex fields. If we restrict ourselves to real fields, there is only one eigenspace for each ℓ which has eigenvalue C_ℓ. Furthermore, in this case the eigenspaces are spanned by the vectors

$$
v_{\ell m}^{(1)'} = \hat{e}_{\ell, m, \text{``real''}}
$$

and

$$v_{\ell m}^{(2)'} = \hat{e}_{\ell, m, \text{``imag''}},$$

where still $\ell = 0, 1, \ldots$ and $m = 0, \ldots, \ell$.

In order to multiply a real isotropic signal field with its covariance matrix, one therefore has to simply multiply each complex spherical harmonics coefficient $s_{\ell m}$, or equivalently both its real and imaginary part individually, with the corresponding coefficient of the angular power spectrum, C_ℓ, and *not* with $\beta_\ell = \frac{C_\ell}{2}$.

Appendix B

Biased variance estimate from posterior samples

B.1 Setting

Assuming a linear data model,

$$d = Rs + n, \tag{B.1}$$

with Gaussian signal and noise,

$$s \hookleftarrow \mathcal{G}(s, S), \qquad n \hookleftarrow \mathcal{G}(n, N), \tag{B.2}$$

the posterior for the signal will also be a Gaussian

$$\mathcal{P}(s|d) = \mathcal{G}(s - m, D), \tag{B.3}$$

described by a mean m and a covariance D which takes on the form

$$D = \left(S^{-1} + M \right)^{-1}, \tag{B.4}$$

where $M = R^{\dagger} N^{-1} R$. The uncertainty of the posterior mean m can be quantified by the local variance $\hat{D}$. It is our goal to estimate this variance in a situation where we are able to compute the matrix-vector product Ds for any signal configurations but are not provided with the entries of the matrix D explicitly. One way of doing this is to draw samples from the Gaussian and calculate the variance of these samples.[1] It will be shown that inaccuracies in the calculation of the product Ds can lead to a positive bias for the resulting variance estimate.

---

[1] Note that we are dealing with a Gaussian posterior for definiteness only. Any posterior with the same first and second connected moments can be treated in exactly the same way.

B.2   Drawing samples

One way of drawing samples $s^{(\alpha)}$ from a Gaussian probability distribution described by a mean m and a covariance D was described by Jasche et al. (2010b). The idea is to add residuals to the mean m. These residuals are created by drawing signal and noise realizations from the respective priors, combining them into a data vector, and subtracting the filtered data vector from the signal realization, according to

$$\tilde{s} \hookleftarrow \mathcal{G}(s, S), \tag{B.5a}$$

$$\tilde{n} \hookleftarrow \mathcal{G}(n, N), \tag{B.5b}$$

$$\tilde{d} = R\tilde{s} + \tilde{n}, \tag{B.5c}$$

$$y = \tilde{s} - DR^{\dagger}N^{-1}\tilde{d}, \tag{B.5d}$$

$$s^{(\alpha)} = m + y. \tag{B.5e}$$

Averaging over many different samples is then used as a proxy of calculating the expectation value,

$$\frac{1}{A}\sum_{\alpha=1}^{A} f(s^{(\alpha)}) \approx \left\langle f(s^{(\alpha)}) \right\rangle_{\mathcal{G}(s,S)\mathcal{G}(n,N)}. \tag{B.6}$$

The expectation value of the samples is obviously m. In order to calculate the expected variance, first consider the residual itself,

$$\begin{aligned} s^{(\alpha)} - m &= y \\ &= \tilde{s} - DR^{\dagger}N^{-1}\tilde{d} \\ &= (1 - DM)\,\tilde{s} - DR^{\dagger}N^{-1}\tilde{n} \\ &= D\left(D^{-1} - M\right)\tilde{s} - DR^{\dagger}N^{-1}\tilde{n} \\ &= DS^{-1}\tilde{s} - DR^{\dagger}N^{-1}\tilde{n}. \end{aligned} \tag{B.7}$$

This leads to the expectation value for the covariance

$$\begin{aligned} \left\langle \left(s^{(\alpha)} - m\right)\left(s^{(\alpha)} - m\right)^{\dagger} \right\rangle_{\mathcal{G}(s,S)\mathcal{G}(n,N)} &= DS^{-1}SS^{-1}D + DR^{\dagger}N^{-1}NN^{-1}RD \\ &= DS^{-1}D + DMD \\ &= DD^{-1}D \\ &= D. \end{aligned} \tag{B.8}$$

B.3   Effect of numerical inaccuracies

Now assume that the numerical application of the covariance matrix D' differs from the analytical multiplication with D by some matrix $\Delta = D - D'$. The numerically calculated

residual vector y therefore becomes

$$
\begin{aligned}
s^{(\alpha)} - m &= y \\
&= \tilde{s} - D' R^\dagger N^{-1} \tilde{d} \\
&= \left(1 - D'M\right) \tilde{s} - D' R^\dagger N^{-1} \tilde{n} \\
&= \left(1 - (D - \Delta) M\right) \tilde{s} - (D - \Delta) N^{-1} \tilde{n} \\
&= \left(D \left(D^{-1} - M\right) + \Delta M\right) \tilde{s} - (D - \Delta) N^{-1} \tilde{n} \\
&= \left(DS^{-1} + \Delta M\right) \tilde{s} - (D - \Delta) N^{-1} \tilde{n}.
\end{aligned}
\tag{B.9}
$$

This leads to an altered expectation value for the covariance,

$$
\begin{aligned}
&\left\langle \left(s^{(\alpha)} - m\right) \left(s^{(\alpha)} - m\right)^\dagger \right\rangle_{\mathcal{G}(s,S)\mathcal{G}(n,N)} \\
&= \left(DS^{-1} + \Delta M\right) S \left(S^{-1} D + \Delta M\right)^\dagger + (D - \Delta) M (D - \Delta)^\dagger \\
&= DS^{-1}D + \Delta M S M \Delta^\dagger + DM\Delta^\dagger + \Delta M D + DMD + \Delta M \Delta^\dagger - DM\Delta^\dagger - \Delta M D \\
&= DD^{-1}D + \Delta \left(MSM + M\right) \Delta^\dagger \\
&= D + \Delta \left(MSM + M\right) \Delta^\dagger.
\end{aligned}
\tag{B.10}
$$

The matrix $(MSM + M)$ is positive definite. If the matrix Δ has full rank, the product $\Delta (MSM + M) \Delta^\dagger$ will therefore also be positive definite (see Petersen & Pedersen 2008) and lead to an over-estimation of the variance.[2]

B.4   Alternative, unbiased methods

An alternative strategy to estimate any matrix diagonal is the method of "matrix probing" described e.g. by Selig et al. (2012). A Bayesian extension of the same method was introduced in the same paper. The results of this appendix make clear that these methods should be preferred whenever the posterior variance needs to be estimated. In cases in which other expectation values are of interest, as indicated in Eq. (B.6), the posterior sampling is however often the only feasible approach.

---

[2]If Δ does not have full rank, the product will in general be positive semi-definite and the conclusion will still hold. A formal proof, however, is beyond the scope of this appendix.

Appendix C

Higher order solutions for the *extended critical filter*

Here we briefly list the results for nontrivial temperatures up to second order, i.e. considering terms up to $i = 2$ in Eqs. (3.24) and (3.25). Since the first order terms are zero for our choice of $\tilde{q}_k$ and $\tilde{r}_j$, we list only the resulting filter formulas for the zeroth and second order internal energy.

C.1 Zeroth order

The zeroth order solution with arbitrary temperature is rather similar to the one with $T = 1$ presented in Sect. 3.3. It is given by

$$
m = D'j, \tag{C.1}
$$

$$
j = \sum_j \frac{\delta_j}{\tilde{r}_j} R^\dagger N_j^{-1} d, \tag{C.2}
$$

$$
D' = \left(\sum_k \frac{\gamma_k}{\tilde{q}_k} S_k^{-1} + \sum_j \frac{\delta_j}{\tilde{r}_j} R^\dagger N_j^{-1} R \right)^{-1} \tag{C.3}
$$

$$
D = TD'. \tag{C.4}
$$

The mean m is completely unchanged. However, the covariance D of the Gaussian approximation is now T times the information propagator D', i.e. the Gaussian approximation becomes wider at higher temperature.

C.2   Second order

The second order solution is given by

$$m = D'j, \tag{C.5}$$

$$j = \sum_j \frac{\delta_j}{\tilde{r}_j} Y_j R^\dagger N_j^{-1} d, \tag{C.6}$$

$$D' = \left(\sum_k \frac{\gamma_k}{\tilde{q}_k} X_k + \sum_j \frac{\delta_j}{\tilde{r}_j} Y_j R^\dagger N_j^{-1} R \right)^{-1}, \tag{C.7}$$

$$X_k = 1 + \frac{1}{\tilde{q}_k^2} \mathrm{tr} \left(\left(mm^\dagger + \frac{1}{2} D \right) S_k^{-1} D S_k^{-1} \right) - \frac{1}{\tilde{q}_k} S_k^{-1} D, \tag{C.8}$$

$$Y_j = 1 + \frac{1}{\tilde{r}_j^2} \mathrm{tr} \left(\left((d - Rm)(d - Rm)^\dagger + \frac{1}{2} R D R^\dagger \right) N_j^{-1} R D R^\dagger N_j^{-1} \right)$$
$$- \frac{1}{\tilde{r}_j} R^\dagger N_j^{-1} R D, \tag{C.9}$$

$$D = T \left(D'^{-1} - \sum_k \frac{\gamma_k}{\tilde{q}_k^2} S_k^{-1} \left(mm^\dagger \right) S_k^{-1} \right.$$
$$\left. - \sum_j \frac{\delta_j}{\tilde{r}_j^2} R^\dagger N_j^{-1} \left((d - Rm)(d - Rm)^\dagger \right) N_j^{-1} R \right)^{-1}. \tag{C.10}$$

Again, the only effect of the temperature is to broaden the approximate Gaussian. However, in the second order solution the operators X_k and Y_j appear, destroying the one-to-one correspondence between the terms in these expressions for D, D', and j and the Wiener filter formula Eq. (3.8). Therefore, the values of the parameters p_k and η_j are not immediately determined by these equations. Note, however, that the goal was not to determine the signal and noise covariance matrices, but to find the optimal Gaussian approximation to the signal posterior, given by m and D.

Appendix D

The three-dimensional Fourier power spectrum of the differential Faraday depth and the resulting angular power spectrum

Defining the *differential Faraday depth*

$$\tilde{\varphi}(\vec{x}) = n_{\mathrm{e}}(\vec{x}) B_r(\vec{x}) \tag{D.1}$$

as the product of the thermal electron density and the line of sight component of the magnetic field, we assume that its functional form factorizes into a known radial dependence and a remaining term

$$\tilde{\varphi}(\vec{x}) = \bar{\varphi}(r)\varphi(\vec{x}), \tag{D.2}$$

as we had done in Eq. (4.15). We assume further that the remaining term φ can be modeled as a statistically homogeneous and isotropic random field whose two-point statistics are completely described by its power spectrum, i.e.

$$\left\langle \varphi(\vec{k})\varphi^*(\vec{k}') \right\rangle = (2\pi)^3 \, \delta^{(3)}(\vec{k} - \vec{k}') P_\varphi(k). \tag{D.3}$$

Now, consider the two-point correlation function of the Faraday depth ϕ, represented

by its spherical harmonics components $\phi_{\ell m}$,

$$
\begin{aligned}
\langle \phi_{\ell m}\phi_{\ell'm'}\rangle &= \left\langle \int_{\mathcal{S}^2} \mathrm{d}\Omega\, Y_{\ell m}^*(\hat{n})\phi(\hat{n}) \int_{\mathcal{S}^2} \mathrm{d}\Omega'\, Y_{\ell'm'}(\hat{n}')\phi(\hat{n}') \right\rangle \\
&= \int_{\mathcal{S}^2} \mathrm{d}\Omega \int_{\mathcal{S}^2} \mathrm{d}\Omega'\, Y_{\ell m}(\hat{n})Y_{\ell'm'}(\hat{n}')\,\langle \phi(\hat{n})\phi(\hat{n}')\rangle \\
&= \int_{\mathcal{S}^2} \mathrm{d}\Omega \int_{\mathcal{S}^2} \mathrm{d}\Omega'\, Y_{\ell m}(\hat{n})Y_{\ell'm'}(\hat{n}') \left\langle \int_0^\infty \mathrm{d}r\, \bar{\varphi}(r)\varphi(r\hat{n}) \int_0^\infty \mathrm{d}r'\, \bar{\varphi}(r')\varphi(r'\hat{n}') \right\rangle \\
&= \int_{\mathcal{S}^2} \mathrm{d}\Omega \int_{\mathcal{S}^2} \mathrm{d}\Omega'\, Y_{\ell m}(\hat{n})Y_{\ell'm'}(\hat{n}') \int_0^\infty \mathrm{d}r \int_0^\infty \mathrm{d}r'\, \bar{\varphi}(r)\bar{\varphi}(r')\,\langle \varphi(r\hat{n})\varphi(r'\hat{n}')\rangle \\
&= \int_{\mathcal{S}^2} \mathrm{d}\Omega \int_{\mathcal{S}^2} \mathrm{d}\Omega'\, Y_{\ell m}(\hat{n})Y_{\ell'm'}(\hat{n}') \int_0^\infty \mathrm{d}r \int_0^\infty \mathrm{d}r'\, \bar{\varphi}(r)\bar{\varphi}(r') \\
&\quad \times \left\langle \frac{1}{(2\pi)^3}\int_{\mathbb{R}^3} \mathrm{d}^3k\, e^{i\vec{k}\cdot\hat{n}r}\varphi(\vec{k}) \frac{1}{(2\pi)^3}\int_{\mathbb{R}^3} \mathrm{d}^3k'\, e^{-i\vec{k}'\cdot\hat{n}'r'}\varphi^*(\vec{k}') \right\rangle \\
&\propto \int_{\mathcal{S}^2} \mathrm{d}\Omega \int_{\mathcal{S}^2} \mathrm{d}\Omega'\, Y_{\ell m}(\hat{n})Y_{\ell'm'}(\hat{n}') \int_0^\infty \mathrm{d}r \int_0^\infty \mathrm{d}r'\, \bar{\varphi}(r)\bar{\varphi}(r') \\
&\quad \times \int_{\mathbb{R}^3} \mathrm{d}^3k \int_{\mathbb{R}^3} \mathrm{d}^3k'\, e^{i\vec{k}\cdot\hat{n}r}e^{-i\vec{k}'\cdot\hat{n}'r'} \left\langle \varphi(\vec{k})\varphi^*(\vec{k}') \right\rangle \\
&\propto \int_{\mathcal{S}^2} \mathrm{d}\Omega \int_{\mathcal{S}^2} \mathrm{d}\Omega'\, Y_{\ell m}(\hat{n})Y_{\ell'm'}(\hat{n}') \int_0^\infty \mathrm{d}r \int_0^\infty \mathrm{d}r'\, \bar{\varphi}(r)\bar{\varphi}(r') \\
&\quad \times \int_{\mathbb{R}^3} \mathrm{d}^3k\, e^{i\vec{k}\cdot(\hat{n}r-\hat{n}'r')}P_\varphi(k).
\end{aligned}
\tag{D.4}
$$

Using

$$
e^{i\vec{k}\cdot\hat{n}r} = \sum_{\ell=0}^\infty (2\ell+1)\, i^\ell j_\ell(kr) P_\ell(\hat{k}\cdot\hat{n}),
\tag{D.5}
$$

where j_ℓ is a spherical Bessel function and P_ℓ is a Legendre polynomial (e.g. Abramowitz

& Stegun 1964), this becomes

$$
\begin{aligned}
\langle \phi_{\ell m} \phi_{\ell' m'} \rangle &\propto \int_{\mathbb{R}^3} \mathrm{d}^3 k \int_0^\infty \mathrm{d}r \int_0^\infty \mathrm{d}r' \; P_\varphi(k) \bar{\varphi}(r) \bar{\varphi}(r') \\
&\quad \times \int_{\mathcal{S}^2} \mathrm{d}\Omega \; Y_{\ell m}^*(\hat{n}) \sum_{\ell''=0}^\infty (2\ell''+1) \, i^{\ell''} j_{\ell''}(kr) P_{\ell''}(\hat{k} \cdot \hat{n}) \\
&\quad \times \int_{\mathcal{S}^2} \mathrm{d}\Omega' \; Y_{\ell' m'}(\hat{n}') \sum_{\ell'''=0}^\infty (2\ell'''+1) \, i^{-\ell'''} j_{\ell'''}(kr') P_{\ell'''}(\hat{k} \cdot \hat{n}') \\
&\propto \int_{\mathbb{R}^3} \mathrm{d}^3 k \int_0^\infty \mathrm{d}r \int_0^\infty \mathrm{d}r' \; P_\varphi(k) \bar{\varphi}(r) \bar{\varphi}(r') \\
&\quad \times \sum_{\ell''=0}^\infty \sum_{m''=-\ell''}^{\ell''} i^{\ell''} j_{\ell''}(kr) \int_{\mathcal{S}^2} \mathrm{d}\Omega \; Y_{\ell m}^*(\hat{n}) Y_{\ell'' m''}(\hat{n}) Y_{\ell'' m''}^*(\hat{k}) \\
&\quad \times \sum_{\ell'''=0}^\infty \sum_{m'''=-\ell'''}^{\ell'''} i^{-\ell'''} j_{\ell'''}(kr') \int_{\mathcal{S}^2} \mathrm{d}\Omega' \; Y_{\ell' m'}(\hat{n}') Y_{\ell''' m'''}^*(\hat{n}') Y_{\ell''' m'''}(\hat{k}),
\end{aligned} \tag{D.6}
$$

where we have used

$$
P_\ell(\hat{k} \cdot \hat{n}) = \frac{4\pi}{2\ell+1} \sum_{m=-\ell}^\ell Y_{\ell m}(\hat{k}) Y_{\ell m}^*(\hat{n}). \tag{D.7}
$$

Performing the integrals over $\hat{n}$ and $\hat{n}'$ yields

$$
\begin{aligned}
\langle \phi_{\ell m} \phi_{\ell' m'} \rangle &\propto \int_{\mathbb{R}^3} \mathrm{d}^3 k \int_0^\infty \mathrm{d}r \int_0^\infty \mathrm{d}r' \; P_\varphi(k) \bar{\varphi}(r) \bar{\varphi}(r') j_\ell(kr) Y_{\ell m}^*(\hat{k}) j_{\ell'}(kr') Y_{\ell' m'}(\hat{k}) \\
&= \int_0^\infty \mathrm{d}r \int_0^\infty \mathrm{d}r' \int_0^\infty \mathrm{d}k \; k^2 P_\varphi(k) \bar{\varphi}(r) \bar{\varphi}(r') j_\ell(kr) j_{\ell'}(kr') \int_{\mathcal{S}^2} \mathrm{d}\Omega_{\hat{k}} \; Y_{\ell m}(\hat{k}) Y_{\ell' m'}(\hat{k}) \\
&= \int_0^\infty \mathrm{d}r \int_0^\infty \mathrm{d}r' \int_0^\infty \mathrm{d}k \; k^2 P_\varphi(k) \bar{\varphi}(r) \bar{\varphi}(r') j_\ell(kr) j_{\ell'}(kr') \delta_{\ell\ell'} \delta_{mm'},
\end{aligned} \tag{D.8}
$$

which proves the statistical homogeneity and isotropy of the Faraday depth itself under our assumptions. The angular power spectrum now is

$$
C_\ell \propto \int_0^\infty \mathrm{d}r \; \bar{\varphi}(r) \int_0^\infty \mathrm{d}r' \; \bar{\varphi}(r') \int_0^\infty \mathrm{d}k \; k^2 P_\varphi(k) j_\ell(kr) j_\ell(kr'). \tag{D.9}
$$

We assume further that the power spectrum of the field φ is given by a power law,

$$
P_\varphi(k) \propto k^{-\alpha}. \tag{D.10}
$$

Using the relation

$$
j_\ell(z) = \sqrt{\frac{\pi}{2z}} J_{\ell+1/2}(z), \tag{D.11}
$$

where $J_{\ell+1/2}$ is a Bessel function of the first kind, and the integral

$$
\int_0^\infty \mathrm{d}t \, \frac{J_\mu(at) J_\nu(bt)}{t^\lambda} = \frac{a^\mu \Gamma\left(\frac{\mu+\nu-\lambda+1}{2}\right)}{2^\lambda b^{\mu-\lambda+1} \Gamma(\mu+1)\Gamma\left(\frac{\nu-\mu+\lambda+1}{2}\right)}
$$
$$
\times \; {}_2F_1\left(\frac{\mu+\nu-\lambda+1}{2}, \frac{\mu-\nu-\lambda+1}{2}; \mu+1; \left(\frac{a}{b}\right)^2\right) \qquad \text{(D.12)}
$$

for $\mu+\nu-\lambda+1 > 0$, $\lambda > -1$, and $0 < a < b$, where ${}_2F_1$ is Gauss' hypergeometric function (e.g. Abramowitz & Stegun 1964), we can solve the k-integral in Eq. (D.9),

$$
\int_0^\infty \mathrm{d}k \, k^{2-\alpha} j_\ell(kr) j_\ell(kr') = \frac{\pi}{2} \int_0^\infty k^{1-\alpha} J_{\ell+1/2}(kr) J_{\ell+1/2}(kr')
$$
$$
= \frac{\pi}{2} \frac{r^{\ell+\frac{1}{2}} \Gamma\left(\ell+1-\frac{\alpha}{2}\right)}{2^{\alpha-2} (r')^{\ell+2-\alpha} \Gamma(\ell+1)\Gamma\left(\frac{\alpha}{2}\right)}
$$
$$
\times \; {}_2F_1\left(\ell+1+\frac{\alpha}{2}, 1-\frac{\alpha}{2}; \ell+1; \left(\frac{r}{r'}\right)^2\right) \qquad \text{(D.13)}
$$

for cases in which $2\ell+2-\alpha > 0$, $\alpha-1 > 0$, and $0 < r < r'$. To enforce the last condition, we replace

$$
\int_0^\infty \mathrm{d}r \int_0^\infty \mathrm{d}r' \longrightarrow 2 \int_0^\infty \mathrm{d}r \int_r^\infty \mathrm{d}r', \qquad \text{(D.14)}
$$

which is possible since the integrand in Eq. (D.9) is symmetric in r and r'.

Altogether, Eq. (D.9) becomes

$$
C_\ell \propto \frac{\Gamma\left(\ell+1-\frac{\alpha}{2}\right)}{\Gamma(\ell+1)} \int_0^\infty \mathrm{d}r \int_r^\infty \mathrm{d}r' \, \bar\varphi(r)\bar\varphi(r') \frac{r^{\ell+1}}{(r')^{\ell+2-\alpha}} \, {}_2F_1\left(\ell+1-\frac{\alpha}{2}, 1-\frac{\alpha}{2}; \ell+1; \left(\frac{r}{r'}\right)^2\right).
$$
$$
\text{(D.15)}
$$

Using the power series formula

$$
{}_2F_1(a,b;c;z) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \sum_{n=0}^\infty \frac{\Gamma(a+n)\Gamma(b+n)}{\Gamma(c+n)n!} z^n, \qquad \text{(D.16)}
$$

which has convergence radius 1 (Abramowitz & Stegun 1964), we can expand the integral in Eq. (D.15), yielding

$$
C_\ell \propto \frac{1}{\Gamma\left(1-\frac{\alpha}{2}\right)} \int_0^\infty \mathrm{d}r \int_r^\infty \mathrm{d}r' \, \bar\varphi(r)\bar\varphi(r') \frac{r^{\ell+1}}{(r')^{\ell+2-\alpha}} \sum_{n=0}^\infty \frac{\Gamma\left(\ell+1-\frac{\alpha}{2}+n\right)\Gamma\left(1-\frac{\alpha}{2}+n\right)}{\Gamma(\ell+1+n)\,n!} \left(\frac{r}{r'}\right)^{2n}
$$
$$
\propto \sum_{n=0}^\infty \frac{\Gamma\left(\ell+1-\frac{\alpha}{2}+n\right)\Gamma\left(1-\frac{\alpha}{2}+n\right)}{\Gamma(\ell+1+n)\,n!} \int_0^\infty \mathrm{d}r \, \bar\varphi(r) r^{2n+\ell+1} \int_r^\infty \mathrm{d}r' \, \bar\varphi(r')\,(r')^{\alpha-2n-\ell-2},
$$
$$
\text{(D.17)}
$$

where we have exchanged the integration and the sum.

To arrive at the result presented in Sec. 4.4.3, we assume now that the part of the differential Faraday depth containing the known radial dependence is constant within a sphere of radius r_0 and zero outside,

$$\bar{\varphi}(r) = \begin{cases} \varphi_0 & \text{if } r < r_0 \\ 0 & \text{else} \end{cases} . \tag{D.18}$$

Thus, we can solve the remaining integrals, arriving at the final expression for the ℓ-dependence of the angular power spectrum

$$
\begin{aligned}
C_\ell &\propto \sum_{n=0}^{\infty} \frac{\Gamma\left(\ell+1-\frac{\alpha}{2}+n\right)\Gamma\left(1-\frac{\alpha}{2}+n\right)}{\Gamma\left(\ell+1+n\right) n!} \int_0^{r_0} \mathrm{d}r \ r^{2n+\ell+1} \int_r^{r_0} \mathrm{d}r' \ (r')^{\alpha-2n-\ell-2} \\
&= \sum_{n=0}^{\infty} \frac{\Gamma\left(\ell+1-\frac{\alpha}{2}+n\right)\Gamma\left(1-\frac{\alpha}{2}+n\right)}{\Gamma\left(\ell+1+n\right) n!} \\
&\qquad \times \int_0^{r_0} \mathrm{d}r \ r^{2n+\ell+1} \frac{1}{\alpha-2n-\ell-1}\left(r_0^{\alpha-2n-\ell-1}-r^{\alpha-2n-\ell-1}\right) \\
&= \sum_{n=0}^{\infty} \frac{\Gamma\left(\ell+1-\frac{\alpha}{2}+n\right)\Gamma\left(1-\frac{\alpha}{2}+n\right)}{\Gamma\left(\ell+1+n\right) n! \left(\alpha-2n-\ell-1\right)} \int_0^{r_0} \mathrm{d}r \ \left(r_0^{\alpha-2n-\ell-1}r^{2n+\ell+1}-r^\alpha\right) \\
&= \sum_{n=0}^{\infty} \frac{\Gamma\left(\ell+1-\frac{\alpha}{2}+n\right)\Gamma\left(1-\frac{\alpha}{2}+n\right)}{\Gamma\left(\ell+1+n\right) n! \left(\alpha-2n-\ell-1\right)} \left(\frac{1}{2n+\ell+2}r_0^{\alpha+1}-\frac{1}{\alpha+1}r_0^{\alpha+1}\right) \\
&= r_0^{\alpha+1} \sum_{n=0}^{\infty} \frac{\Gamma\left(\ell+1-\frac{\alpha}{2}+n\right)\Gamma\left(1-\frac{\alpha}{2}+n\right)}{\Gamma\left(\ell+1+n\right) n! \left(\alpha-2n-\ell-1\right)} \left(\frac{1}{2n+\ell+2}-\frac{1}{\alpha+1}\right) \\
&\propto \sum_{n=0}^{\infty} \frac{\Gamma\left(\ell+1-\frac{\alpha}{2}+n\right)\Gamma\left(1-\frac{\alpha}{2}+n\right)}{\Gamma\left(\ell+1+n\right) n! \left(\alpha-2n-\ell-1\right)} \left(\frac{1}{2n+\ell+2}-\frac{1}{\alpha+1}\right) .
\end{aligned}
\tag{D.19}
$$

The integration over r' as performed here is only valid if $0 \neq \alpha-2n-\ell-2 \neq -1$, which is guaranteed to be satisfied for $\ell > \alpha-1$. For these ℓ-values, the first condition needed for Eq. (D.13) is also satisfied. Note that the radius r_0 only affects the normalization of the power spectrum, not its shape.

Figure D.1 shows the reconstructed angular power spectrum, the power law fit $C_\ell \propto \ell^{-2.17}$, and the result of a numerical evaluation of Eq. (D.15) for $\alpha = 2.17$.

D.1 Statistical homogeneity and isotropy of differential Faraday rotation

As a thought experiment, assume that both the electron density n_e and the line of sight component of the magnetic field B_r are statistically homogeneous and isotropic Gaussian random fields with covariances

$$\left\langle n_\mathrm{e}(\vec{k})n_\mathrm{e}^*(\vec{k}')\right\rangle = (2\pi)^3 \delta^{(3)}(\vec{k}-\vec{k}')P_{n_\mathrm{e}}(k), \tag{D.20}$$
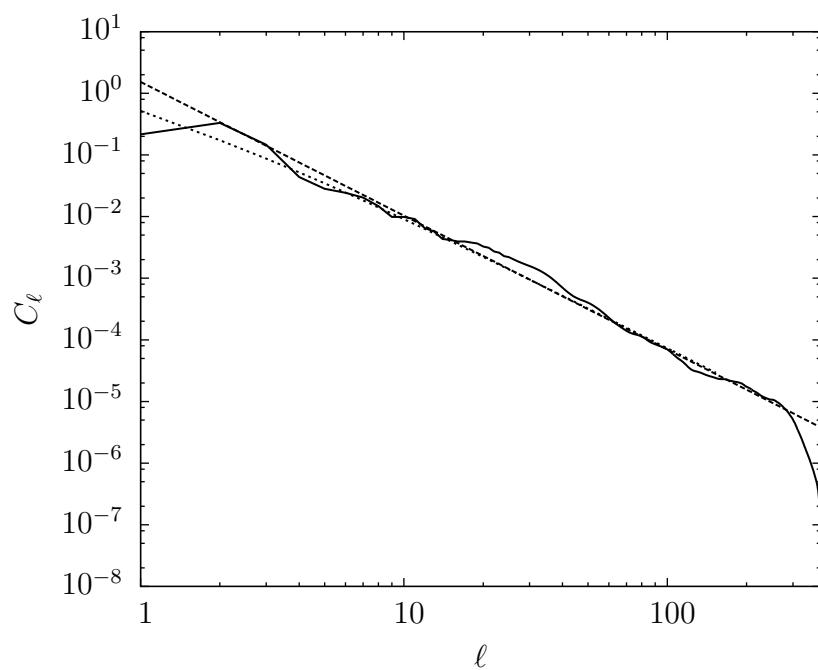
Figure D.1: Angular power spectrum of the Faraday depth. Shown is the reconstructed power spectrum (solid line), a power law fit (dashed line), and the prediction from $P_\varphi(k) \propto k^{-2.17}$.

$$\left\langle B_r(\vec{k})B_r^*(\vec{k'}) \right\rangle = (2\pi)^3\, \delta^{(3)}(\vec{k}-\vec{k'})P_{B_r}(k), \tag{D.21}$$

and

$$\left\langle B_r(\vec{k})n_{\mathrm{e}}^*(\vec{k'}) \right\rangle = 0. \tag{D.22}$$

The last assumption is justified since the line of sight component of the magnetic field will not have a preferred direction that is correlated with the electron denisty.

Now, consider the autocorrelation of the differential Faraday rotation $\tilde\varphi$ in Fourier space,

$$
\begin{aligned}
\left\langle \tilde\varphi(\vec{k})\tilde\varphi^*(\vec{k'}) \right\rangle &= \int_{\mathbb{R}^3} \mathrm{d}^3x \int_{\mathbb{R}^3} \mathrm{d}^3y\ \mathrm{e}^{-i(\vec{k}\cdot\vec{x}-\vec{k'}\cdot\vec{y})}\,\langle \tilde\varphi(\vec{x})\tilde\varphi(\vec{y})\rangle \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3x \int_{\mathbb{R}^3} \mathrm{d}^3y\ \mathrm{e}^{-i(\vec{k}\cdot\vec{x}-\vec{k'}\cdot\vec{y})}\,\langle n_{\mathrm{e}}(\vec{x})B_r(\vec{x})n_{\mathrm{e}}(\vec{y})B_r(\vec{y})\rangle \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3x \int_{\mathbb{R}^3} \mathrm{d}^3y\ \mathrm{e}^{-i(\vec{k}\cdot\vec{x}-\vec{k'}\cdot\vec{y})}\,C^{(n_{\mathrm{e}})}\left(|\vec{x}-\vec{y}|\right)C^{(B_r)}\left(|\vec{x}-\vec{y}|\right) \\
&= \int_{\mathbb{R}^3} \mathrm{d}^3x \int_{\mathbb{R}^3} \mathrm{d}^3r\ \mathrm{e}^{-i(\vec{k}-\vec{k'})\cdot\vec{x}}\mathrm{e}^{-i\vec{k'}\cdot\vec{r}}C^{(n_{\mathrm{e}})}(r)C^{(B_r)}(r) \\
&\propto \delta^{(3)}(\vec{k}-\vec{k'}) \int_{\mathbb{R}^3} \mathrm{d}^3r\ \mathrm{e}^{-i\vec{k'}\cdot\vec{r}}C^{(n_{\mathrm{e}})}(r)C^{(B_r)}(r) \\
&\propto \delta^{(3)}(\vec{k}-\vec{k'}) \int_{\mathbb{R}^3} \mathrm{d}^3r\ \mathrm{e}^{-i\vec{k'}\cdot\vec{r}} \int_{\mathbb{R}^3} \mathrm{d}^3k''\ \mathrm{e}^{i\vec{k''}\cdot\vec{r}}P_{n_{\mathrm{e}}}(k'') \int_{\mathbb{R}^3} \mathrm{d}^3k'''\ \mathrm{e}^{i\vec{k'''}\cdot\vec{r}}P_{B_r}(k''') \\
&= \delta^{(3)}(\vec{k}-\vec{k'}) \int_{\mathbb{R}^3} \mathrm{d}^3k'' \int_{\mathbb{R}^3} \mathrm{d}^3k'''\ P_{n_{\mathrm{e}}}(k'')P_{B_r}(k''') \int_{\mathbb{R}^3} \mathrm{d}^3r\ \mathrm{e}^{-i(\vec{k'}-\vec{k''}-\vec{k'''})\cdot\vec{r}} \\
&\propto \delta^{(3)}(\vec{k}-\vec{k'}) \int_{\mathbb{R}^3} \mathrm{d}^3k'' \int_{\mathbb{R}^3} \mathrm{d}^3k'''\ P_{n_{\mathrm{e}}}(k'')P_{B_r}(k''')\delta^{(3)}(\vec{k'}-\vec{k''}-\vec{k'''}) \\
&= \delta^{(3)}(\vec{k}-\vec{k'}) \int_{\mathbb{R}^3} \mathrm{d}^3k''\ P_{n_{\mathrm{e}}}(k'')P_{B_r}(|k-k''|) \\
&\propto \delta^{(3)}(\vec{k}-\vec{k'}) \int_0^\infty \mathrm{d}k''\ k''^2 P_{n_{\mathrm{e}}}(k'')P_{B_r}(|k-k''|), \tag{D.23}
\end{aligned}
$$

where the autocorrelation functions in position space, $C^{(n_{\mathrm{e}})}\left(|\vec{x}-\vec{y}|\right) = \langle n_{\mathrm{e}}(\vec{x})n_{\mathrm{e}}(\vec{y})\rangle$ and $C^{(B_r)}\left(|\vec{x}-\vec{y}|\right) = \langle B_r(\vec{x})B_r(\vec{y})\rangle$ were used and a relative coordinate $\vec{r} = \vec{x}-\vec{y}$ was introduced in the fourth line. This result shows that the differential Faraday rotation is a statistically homogeneous and isotropic field under the assumption that both the thermal electron density and the line of sight component of the magnetic field have these properties.

Appendix E

Spectral smoothness in special cases and its practical implementation

E.1 Case studies for the spectral smoothness prior

In this appendix, we discuss a few possible correlation structures and power spectra and examine the suitability of a spectral smoothness prior of the form described by Eq. (5.24).

E.1.1 Power law spectra

If the signal's power spectrum is a broken power law of the shape given by Eq. (5.14), the first double-logarithmic derivative is

$$\frac{\partial p_k}{\partial \log k} = -\frac{\gamma}{1 + \left(\frac{k_0}{k}\right)^2}, \tag{E.1}$$

which behaves like $-\gamma \left(\frac{k}{k_0}\right)^2$ for $k \ll k_0$ and tends toward $-\gamma$ for $k \to \infty$. The second double-logarithmic derivative is

$$\frac{\partial^2 p_k}{\partial \left(\log k\right)^2} = -\frac{2\gamma}{\left(\frac{k}{k_0} + \frac{k_0}{k}\right)^2}, \tag{E.2}$$

which tends toward zero both for $k \to 0$ and $k \to \infty$. The second derivative takes on its extremum, given by $-\frac{\gamma}{2}$, at the knee frequency k_0 where the spectral index of the power spectrum changes from 0 to γ. Thus neither a prior punishing large values for the first logarithmic derivative, nor one punishing large values for the second logarithmic derivative prevents the true solution from being found, provided the values of $-\gamma$ and $-\frac{\gamma}{2}$, respectively, are well within the range of values that are allowed for the derivative by the prior. In fact, choosing $\sigma_p = \gamma/2$ would allow such a change in spectral index roughly once per e-folding of the k-value. Choosing $\sigma_p = \frac{\gamma}{2h}$ would turn such a kink into a less common h-sigma event.

The case of a broken power-law contains the special cases of a pure power-law, i.e. $k_0 \to 0$, such as arises for example for Brownian motion of a particle, and of an exponential two-point correlation function, i.e.

$$C(r) = S_{xy} = C_0 e^{-\beta r}, \tag{E.3}$$

where $r = |x - y|$ is the distance between two points or time-instances x and y. Such a correlation function arises for example from the Ornstein-Uhlenbeck process (Uhlenbeck & Ornstein 1930). Fourier transforming this correlation function in a one-dimensional space yields the power spectrum which takes on the form given by Eq. (5.14) with $P_0 = \frac{2C_0}{\beta}$, $k_0 = \beta$, and $\gamma = 2$.

However, there are also cases that do not lead exclusively to small values of the double-logarithmic derivatives. Two such scenarios will be studied in the remainder of this appendix.

E.1.2   Gaussian correlations

Consider a two-point correlation function of Gaussian shape for a field on a one-dimensional space, i.e.

$$C(r) = C_0 e^{-\frac{r^2}{2\sigma^2}}. \tag{E.4}$$

A field whose statistics are described by such a two-point correlation function can be regarded as a stationary and spatially uncorrelated field convolved with a Gaussian. Calculating the corresponding power spectrum via Fourier transformation yields

$$P_k = \sqrt{2\pi}\sigma C_0 e^{-\frac{\sigma^2 k^2}{2}}. \tag{E.5}$$

This power spectrum drops quickly with increasing k, due to the flatness of the correlation function around $r = 0$. Therefore, both the first and second double-logarithmic derivatives grow unbounded as $k \to \infty$. Thus, by employing a spectral smoothness prior that punishes large values for these derivatives, one prevents in principle the reconstruction of the true power spectrum and supresses the small-scale correlations in the reconstructed field that are in reality more pronounced. In cases in which Gaussian correlations are expected, it might therefore be advisable to choose the strength of the spectral smoothness prior, given by σ_p, k-dependent or choose a different smoothness prior altogether.

E.1.3   Triangular correlations

Another case in which the double-logarithmic derivatives of the power spectrum can become divergent is a correlation function with finite support. As an example, we consider a signal on a one-dimensional space with correlations only over the finite distance $2L$, given by

$$C(r) = \begin{cases} C_0\left(1 - \frac{r}{L}\right) & \text{for } r < L \\ 0 & \text{else} \end{cases}. \tag{E.6}$$

This triangular correlation function describes a stationary and spatially uncorrelated field that has been convolved with a top-hat kernel. Fourier transforming it yields the power spectrum

$$P_k = \frac{2C_0}{k^2 L} \left(1 - \cos\left(kL\right)\right).$$

(E.7)

This power spectrum becomes exactly zero at finite k-values. Its logarithm, and therefore also the double-logarithmic derivatives, are divergent at these locations, so not even a k-dependent value of σ_p can ensure the correct reconstruction of the power spectrum in this case. The oscillatory behavior of this power spectrum is a generic feature of signal fields that are correlated only over a finite distance.

Note that both in the case of only locally correlated fields and in the case of Gaussian correlations, the double-logarithmic derivatives can be kept finite by adding a constant floor to the power spectrum, i.e. introducing an additive part to the signal that is spatially uncorrelated. In many cases, the variance of the uncorrelated addition needed to satisfy the spectral smoothness prior, Eq. (5.24), will be small enough so as not to influence the signal reconstruction significantly. Note that in practice, any field variations are restricted by the finite pixel size and only k-values up to a finite $k_{\max}$ will be considered. In Sec. 5.2.3, we investigate the effect that the spectral smoothness prior given in Eq. (5.24) has in practice on the recunstruction of a signal field exhibiting the two potentially problematic two-point correlations that we discussed here.

In conclusion, using a spectral smoothness prior that punishes large values for the first or second double-logarithmic derivative of the power spectrum can lead to the introduction of spurious small-scale variations in cases in which the true two-point correlation function is flat around $r = 0$ or has only finite support. In the latter case, it can also introduce spurious large-scale correlations. This will, however, in general only be a problem for the reconstruction if some feature that mimics these large-scale correlations is present in the data, i.e. caused by the noise. Another case in which the employment of any spectral smoothness prior is obviously a bad idea is that of a signal that exhibits prominent periodicities. The detection of such spectral lines would only get hindered by the usage of a spectral smoothness prior.

E.2 Discretization of the spectral smoothness prior

In our implementation we use discretized values $(k_i)_{i=0,\ldots,i_{\max}}$ for the length scales that are represented on the computational grid, or in case of the two-torus, bins thereof. We approximate the integral in the exponent of the spectral smoothness prior, Eq. (5.24), with a sum according to

$$\int \mathrm{d}(\log k) \left(\frac{\partial^2 \log P_k}{\partial \left(\log k\right)^2}\right)^2 \approx \sum_{i=1}^{i_{\max}-1} \delta_i \left(\Delta p\right)_i^2.$$

(E.8)

Here, we use the abbreviations

$$\delta_i = \frac{\log k_{i+1} - \log k_{i-1}}{2} \tag{E.9}$$

for integer values of i and

$$\delta_i = \log k_{i+1/2} - \log k_{i-1/2} \tag{E.10}$$

for half-integer values of i. Note that we have excluded the boundaries at k_0 and $k_{i_{\mathrm{max}}}$ from the sum to avoid numerical problems at these locations.

We approximate the second logarithmic derivative as

$$(\Delta p)_i = \sum_j \Delta_{i,j} p_j = \frac{\frac{p_{i+1}-p_i}{\delta_{i+1/2}} - \frac{p_i-p_{i-1}}{\delta_{i-1/2}}}{\delta_i}, \tag{E.11}$$

so that we can represent it as a matrix with the entries

$$\Delta_{i,i} = -\frac{1}{\delta_i}\left(\frac{1}{\delta_{i+1/2}} + \frac{1}{\delta_{i-1/2}}\right), \tag{E.12}$$

$$\Delta_{i,i\pm 1} = \frac{1}{\delta_i \delta_{i\pm 1}}, \tag{E.13}$$

acting on the vector p. All other entries of the matrix Δ are zero.

We can now write the exponent of Eq. (5.24) as

$$-\frac{1}{2\sigma_p^2}\int d(\log k)\left(\frac{\partial^2 \log P_k}{\partial(\log k)^2}\right)^2 = -\frac{1}{2}\sum_{i,j=0} p_i T_{i,j} p_j, \tag{E.14}$$

where the matrix T is given by

$$T_{i,j} = \frac{1}{\sigma_p^2}\sum_l \Delta_{l,i}\delta_l\Delta_{l,j}. \tag{E.15}$$

Bibliography

Abramowitz, M. & Stegun, I. A. 1964, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, Dover Publications, New York

Ackermann, M., Ajello, M., Atwood, W. B., et al., Fermi LAT observations of cosmic-ray electrons from 7 GeV to 1 TeV, Phys. Rev. D **82** no. 9, (Nov. 2010) 092004, `arXiv:1008.3999 [astro-ph.HE]`

Agafonov, G. I., Ruzmaikin, A. A., & Sokolov, D. D., Magnetic Field in the Perseus Arm, Soviet Ast. **32** (Jun. 1988) 268–+, ADS

Andreasian, R. R., Structure of the galactic magnetic field, Astrofizika **16** (Oct. 1980) 707–713, ADS

Andreasian, R. R., Investigation of the magnetic field of the galaxy, Astrofizika **18** (Apr. 1982) 255–262, ADS

Armstrong, J. W., Rickett, B. J., & Spangler, S. R., Electron density power spectrum in the local interstellar medium, ApJ **443** (Apr. 1995) 209–221

Basu, S. & Dassios, A., A Cox process with log-normal intensity, Insurance: mathematics and economics **31** no. 2, (October 2002) 297–302, © 2002 Elsevier, `http://eprints.lse.ac.uk/16375/`

Batchelor, R. A., Caswell, J. L., Haynes, R. F., et al., Galactic plane H2O masers - A southern survey, Australian Journal of Physics **33** (Mar. 1980) 139–157, ADS

Battaglia, G., Helmi, A., Morrison, H., et al., The radial velocity dispersion profile of the Galactic halo: constraining the density profile of the dark halo of the Milky Way, MNRAS **364** (Dec. 2005) 433–442, `arXiv:astro-ph/0506102`

Battaglia, G., Helmi, A., Morrison, H., et al., Erratum: The radial velocity dispersion profile of the Galactic halo: constraining the density profile of the dark halo of the Milky Way, MNRAS **370** (Aug. 2006) 1055–1056

Beck, R., Cosmic Magnetic Fields: Observations and Prospects, in American Institute of Physics Conference Series, Vol. 1381, American Institute of Physics Conference Series,

ed. F. A. Aharonian, W. Hofmann, & F. M. Rieger. 2011, 117–136, `arXiv:1104.3749` `[astro-ph.CO]`

Beck, R. & Hoernes, P., Magnetic spiral arms in the galaxy NGC6946, Nature **379** (Jan. 1996) 47–49

Beneš, V., Bodlák, K., Møller, J., & Waagepetersen, R. P. 2002, Bayesian analysis of log Gaussian Cox processes for disease mapping, Tech. Rep. R-02-2001, Department of Mathematical Sciences, Aalborg University

Bennett, C. L., Hill, R. S., Hinshaw, G., et al., First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Foreground Emission, ApJS **148** (Sep. 2003) 97–117, `arXiv:astro-ph/0302208`

Bhat, N. D. R. & Gupta, Y., Pulsar Scintillation in the Local Interstellar Medium: Loop I and Beyond, ApJ **567** (Mar. 2002) 342–353, `arXiv:astro-ph/0111034`

Boggess, N. W., Mather, J. C., Weiss, R., et al., The COBE mission - Its design and performance two years after launch, ApJ **397** (Oct. 1992) 420–429

Bonafede, A., Feretti, L., Murgia, M., et al., The Coma cluster magnetic field from Faraday rotation measures, A&A **513** (Apr. 2010) A30+, `arXiv:1002.0594` `[astro-ph.CO]`

Brandenburg, A. & Subramanian, K., Astrophysical magnetic fields and nonlinear dynamo theory, Phys. Rep. **417** (Oct. 2005) 1–209, `arXiv:astro-ph/0405052`

Braun, R., Heald, G., & Beck, R., The Westerbork SINGS survey. III. Global magnetic field topology, A&A **514** (May 2010) A42, `arXiv:1002.1776` `[astro-ph.CO]`

Braun, R., Oosterloo, T. A., Morganti, R., Klein, U., & Beck, R., The Westerbork SINGS survey. I. Overview and image atlas, A&A **461** (Jan. 2007) 455–470, `arXiv:astro-ph/0610114`

Brentjens, M. A. & de Bruyn, A. G., Faraday rotation measure synthesis, A&A **441** (Oct. 2005) 1217–1228, `arXiv:astro-ph/0507349`

Brix, A., Space-time Multi Type Log Gaussian Cox Processes with a View to Modelling Weeds, Scandinavian Journal of Statistics **28** no. 3, (2001) 471–488, `http://EconPapers.repec.org/RePEc:bla:scjsta:v:28:y:2001:i:3:p:471-488`

Brix, A. & Diggle, P. J., Spatiotemporal prediction for log-Gaussian Cox processes, Journal of the Royal Statistical Society Series B **63** no. 4, (2001) 823–841, `http://EconPapers.repec.org/RePEc:bla:jorssb:v:63:y:2001:i:4:p:823-841`

Broten, N. W., MacLeod, J. M., & Vallée, J. P., Catalogue of unambiguous (Faraday-thin, one-component, spectrum-selected) rotation measures for galaxies and quasars, Ap&SS **141** (Feb. 1988) 303–331

Brown, A. G. A., Hartmann, D., & Burton, W. B., The Orion OB1 association. II. The Orion-Eridanus Bubble., A&A **300** (Aug. 1995) 903, `arXiv:astro-ph/9503016`

Brown, J. C., Haverkorn, M., Gaensler, B. M., et al., Rotation Measures of Extragalactic Sources behind the Southern Galactic Plane: New Insights into the Large-Scale Magnetic Field of the Inner Milky Way, ApJ **663** (Jul. 2007) 258–266, `arXiv:0704.0458`

Brown, J. C. & Taylor, A. R., The Structure of the Magnetic Field in the Outer Galaxy from Rotation Measure Observations through the Disk, ApJ **563** (Dec. 2001) L31–L34

Brown, J. C., Taylor, A. R., & Jackel, B. J., Rotation Measures of Compact Sources in the Canadian Galactic Plane Survey, ApJS **145** (Apr. 2003) 213–223

Brown, J. C., Taylor, A. R., Wielebinski, R., & Mueller, P., On Large-Scale Magnetic Field Reversals in the Outer Galaxy, ApJ **592** (Jul. 2003) L29–L32

Burigana, C., Polarized synchrotron emission, in CMB and Physics of the Early Universe. 2006, `arXiv:astro-ph/0607469`

Burkhart, B., Falceta-Gonçalves, D., Kowal, G., & Lazarian, A., Density Studies of MHD Interstellar Turbulence: Statistical Moments, Correlations and Bispectrum, ApJ **693** (Mar. 2009) 250–266, `arXiv:0811.0822`

Burn, B. J., On the depolarization of discrete radio sources by Faraday dispersion, MNRAS **133** (1966) 67–+, ADS

Caticha, A., Lectures on Probability, Entropy, and Statistical Physics, ArXiv e-prints (Jul. 2008), `arXiv:0808.0012 [physics.data-an]`

Caticha, A. 2012, Entropic Inference and the Foundations of Physics, , Brazilian Chapter of the International Society for Bayesian Analysis – ISBrA (Sao Paulo, Brazil)

Caticha, A. & Preuss, R., Maximum entropy and Bayesian data analysis: Entropic prior distributions, Phys. Rev. E **70** no. 4, (Oct. 2004) 046127, `arXiv:physics/0307055`

Clarke, T. E., Faraday Rotation Observations of Magnetic Fields in Galaxy Clusters, Journal of Korean Astronomical Society **37** (Dec. 2004) 337–342, `arXiv:astro-ph/0412268`

Clarke, T. E., Kronberg, P. P., & Böhringer, H., A New Radio-X-Ray Probe of Galaxy Cluster Magnetic Fields, ApJ **547** (Feb. 2001) L111–L114, `arXiv:astro-ph/0011281`

Clegg, A. W., Cordes, J. M., Simonetti, J. M., & Kulkarni, S. R., Rotation measures of low-latitude extragalactic sources and the magnetoionic structure of the Galaxy, ApJ **386** (Feb. 1992) 143–157

Coles, P. & Jones, B., A lognormal model for the cosmological mass distribution, MNRAS **248** (Jan. 1991) 1–13, ADS

Condon, J. J., Cotton, W. D., Greisen, E. W., et al., The NRAO VLA Sky Survey, AJ **115** (May 1998) 1693–1716

Cordes, J. M. & Lazio, T. J. W., NE2001.I. A New Model for the Galactic Distribution of Free Electrons and its Fluctuations, ArXiv Astrophysics e-prints (Jul. 2002), `arXiv:astro-ph/0207156`

Cordes, J. M. & Lazio, T. J. W., NE2001. II. Using Radio Propagation Data to Construct a Model for the Galactic Distribution of Free Electrons, ArXiv Astrophysics e-prints (Jan. 2003), `arXiv:astro-ph/0301598`

de Avillez, M. A. & Breitschwerdt, D., The Generation and Dissipation of Interstellar Turbulence: Results from Large-Scale High-Resolution Simulations, ApJ **665** (Aug. 2007) L35–L38, `arXiv:0707.1740`

Dennison, B., On intracluster Faraday rotation. I - Observations, AJ **84** (Jun. 1979) 725–729

Dickey, J. M., McClure-Griffiths, N., Gibson, S. J., et al., GASKAP – The Galactic ASKAP Survey, ArXiv e-prints (Jul. 2012), `arXiv:1207.0891 [astro-ph.GA]`

Dineen, P. & Coles, P., A Faraday rotation template for the Galactic sky, MNRAS **362** (Sep. 2005) 403–410, `arXiv:astro-ph/0410636`

Draine, B. T., Interstellar Dust Grains, ARA&A **41** (2003) 241–289, `arXiv:astro-ph/0304489`

Draine, B. T. & Fraisse, A. A., Polarized Far-Infrared and Submillimeter Emission from Interstellar Dust, ApJ **696** (May 2009) 1–11, `arXiv:0809.2094`

Draine, B. T. & Fraisse, A. A., Erratum: "Polarized Far-infrared and Submillimeter Emission from Interstellar Dust" ¡A href="/abs/2009ApJ...696....1D"¿(2009, ApJ, 696, 1)¡/A¿, ApJ **757** (Sep. 2012) 106

Draine, B. T. & Lazarian, A., Diffuse Galactic Emission from Spinning Dust Grains, ApJ **494** (Feb. 1998) L19, `arXiv:astro-ph/9710152`

Draine, B. T. & Lazarian, A., Electric Dipole Radiation from Spinning Dust Grains, ApJ **508** (Nov. 1998) 157–179, `arXiv:astro-ph/9802239`

Enßlin, T. A., Does circular polarisation reveal the rotation of quasar engines?, A&A **401** (Apr. 2003) 499–504, `arXiv:astro-ph/0212387`

Enßlin, T. A. & Frommert, M., Reconstruction of signals with unknown spectra in information field theory with parameter uncertainty, Phys. Rev. D **83** no. 10, (May 2011) 105014–+, `arXiv:1002.2928 [astro-ph.IM]`

Enßlin, T. A., Frommert, M., & Kitaura, F. S., Information field theory for cosmological perturbation reconstruction and nonlinear signal analysis, Phys. Rev. D **80** no. 10, (Nov. 2009) 105005–+, `arXiv:0806.3474`

Enßlin, T. A. & Weig, C., Inference with minimal Gibbs free energy in information field theory, Phys. Rev. E **82** no. 5, (Nov. 2010) 051112–+, `arXiv:1004.2868 [astro-ph.IM]`

Enßlin, T. A. & Weig, C., Reply to "Comment on 'Inference with minimal Gibbs free energy in information field theory' ", Phys. Rev. E **85** no. 3, (Mar. 2012) 033102

Fauvet, L., Macías-Pérez, J. F., Aumont, J., et al., Joint 3D modelling of the polarized Galactic synchrotron and thermal dust foreground diffuse emission, A&A **526** (Feb. 2011) A145, `arXiv:1003.4450 [astro-ph.CO]`

Fauvet, L., Macías-Pérez, J. F., & Désert, F. X., Model of the polarized foreground diffuse Galactic emissions from 33 to 353 GHz, Astroparticle Physics **36** (Aug. 2012) 57–63, `arXiv:1204.3659 [astro-ph.GA]`

Feain, I. J., Cornwell, T. J., Ekers, R. D., et al., The Radio Continuum Structure of Centaurus A at 1.4 GHz, ApJ **740** (Oct. 2011) 17, `arXiv:1104.0077 [astro-ph.CO]`

Feain, I. J., Ekers, R. D., Murphy, T., et al., Faraday Rotation Structure on Kiloparsec Scales in the Radio Lobes of Centaurus A, ApJ **707** (Dec. 2009) 114–125, `arXiv:0910.3458 [astro-ph.CO]`

Finkbeiner, D. P., A Full-Sky Hα Template for Microwave Foreground Prediction, ApJS **146** (Jun. 2003) 407–415, `arXiv:astro-ph/0301558`

Finkbeiner, D. P., Davis, M., & Schlegel, D. J., Extrapolation of Galactic Dust Emission at 100 Microns to Cosmic Microwave Background Radiation Frequencies Using FIRAS, ApJ **524** (Oct. 1999) 867–886, `arXiv:astro-ph/9905128`

Fletcher, A., Beck, R., Shukurov, A., Berkhuijsen, E. M., & Horellou, C., Magnetic fields and spiral arms in the galaxy M51, MNRAS **412** (Apr. 2011) 2396–2416, `arXiv:1001.5230 [astro-ph.CO]`

Francis, C. & Anderson, E., Galactic Spiral Structure, ArXiv e-prints (Jan. 2009), `arXiv:0901.3503 [physics.gen-ph]`

Frick, P., Stepanov, R., Shukurov, A., & Sokoloff, D., Structures in the rotation measure sky, MNRAS **325** (Aug. 2001) 649–664, `arXiv:astro-ph/0012459`

Gabuzda, D. C., Murray, É., & Cronin, P., Helical magnetic fields associated with the relativistic jets of four BL Lac objects, MNRAS **351** (Jul. 2004) L89–L93, `arXiv:astro-ph/0405394`

Gaensler, B. M., Dickey, J. M., McClure-Griffiths, N. M., et al., Radio Polarization from the Inner Galaxy at Arcminute Resolution, ApJ **549** (Mar. 2001) 959–978, `arXiv:astro-ph/0010518`

Gaensler, B. M., Haverkorn, M., Staveley-Smith, L., et al., The Magnetic Field of the Large Magellanic Cloud Revealed Through Faraday Rotation, Science **307** (Mar. 2005) 1610–1612, `arXiv:astro-ph/0503226`

Gaensler, B. M., Madsen, G. J., Chatterjee, S., & Mao, S. A., The Vertical Structure of Warm Ionised Gas in the Milky Way, PASA **25** (Nov. 2008) 184–200, `arXiv:0808.2550`

Gardner, F. F., Morris, D., & Whiteoak, J. B., The linear polarization of radio sources between 11 and 20 cm wavelength. III. Influence of the Galaxy on source depolarization and Faraday rotation, Australian Journal of Physics **22** (Dec. 1969) 813–+, ADS

Gardner, F. F. & Whiteoak, J. B., The Polarization of Cosmic Radio Waves, ARA&A **4** (1966) 245

Gasperini, M., Giovannini, M., & Veneziano, G., Primordial magnetic fields from string cosmology, Physical Review Letters **75** (Nov. 1995) 3796–3799, `arXiv:hep-th/9504083`

Ginzburg, V. L. & Syrovatskii, S. I., Cosmic Magnetobremsstrahlung (synchrotron Radiation), ARA&A **3** (1965) 297

Górski, K. M., Hivon, E., Banday, A. J., et al., HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere, ApJ **622** (Apr. 2005) 759–771, `arXiv:astro-ph/0409513`

Greaves, J. S., Holland, W. S., Minchin, N. R., Murray, A. G., & Stevens, J. A., Submillimetre polarization and constraints on dust grain alignment, A&A **344** (Apr. 1999) 668–674, ADS

Gregorini, L., Vigotti, M., Mack, K.-H., Zoennchen, J., & Klein, U., Multi-frequency study of the B3-VLA sample I. 10.6-GHz data, A&AS **133** (Nov. 1998) 129–147, `arXiv:astro-ph/9808357`

Han, J. L., Manchester, R. N., Berkhuijsen, E. M., & Beck, R., Antisymmetric rotation measures in our Galaxy: evidence for an A0 dynamo., A&A **322** (Jun. 1997) 98–102, ADS

Han, J. L., Manchester, R. N., Lyne, A. G., Qiao, G. J., & van Straten, W., Pulsar Rotation Measures and the Large-Scale Structure of the Galactic Magnetic Field, ApJ **642** (May 2006) 868–881, `arXiv:astro-ph/0601357`

Han, J. L. & Qiao, G. J., The magnetic field in the disk of our Galaxy, A&A **288** (Aug. 1994) 759–772, ADS

Harvey-Smith, L., Madsen, G. J., & Gaensler, B. M., Magnetic Fields in Large-diameter H II Regions Revealed by the Faraday Rotation of Compact Extragalactic Radio Sources, ApJ **736** (Aug. 2011) 83, `arXiv:1106.0931 [astro-ph.GA]`

Haverkorn, M., Brown, J. C., Gaensler, B. M., & McClure-Griffiths, N. M., The Outer Scale of Turbulence in the Magnetoionized Galactic Interstellar Medium, ApJ **680** (Jun. 2008) 362–370, `arXiv:0802.2740`

Haverkorn, M., Gaensler, B. M., Brown, J. C., et al., Enhanced Small-Scale Faraday Rotation in the Galactic Spiral Arms, ApJ **637** (Jan. 2006) L33–L35, `arXiv:astro-ph/0512456`

Haverkorn, M., Gaensler, B. M., McClure-Griffiths, N. M., Dickey, J. M., & Green, A. J., The Southern Galactic Plane Survey: Polarized Radio Continuum Observations and Analysis, ApJS **167** (Dec. 2006) 230–238, `arXiv:astro-ph/0609010`

Haverkorn, M., Katgert, P., & de Bruyn, A. G., Characteristics of the structure in the Galactic polarized radio background at 350 MHz, A&A **403** (Jun. 2003) 1045–1057, `arXiv:astro-ph/0303644`

Heald, G., Braun, R., & Edmonds, R., The Westerbork SINGS survey. II Polarization, Faraday rotation, and magnetic fields, A&A **503** (Aug. 2009) 409–435, `arXiv:0905.3995 [astro-ph.GA]`

Heiles, C., An almost complete survey of 21-centimeter line radiation for galactic latitudes of at least 10 deg. VI - Energetic expanding H I shells, ApJ **208** (Sep. 1976) L137–L139

Heiles, C., A survey of recent magnetic field measurements in H I features of the Galaxy, in IAU Symposium, Vol. 140, Galactic and Intergalactic Magnetic Fields, ed. R. Beck, R. Wielebinski, & P. P. Kronberg. 1990, 35–40, ADS

Heiles, C., 9286 Stars: An Agglomeration of Stellar Polarization Catalogs, AJ **119** (Feb. 2000) 923–927, `arXiv:astro-ph/9910303`

Heiles, C. & Robishaw, T., Zeeman splitting in the diffuse interstellar medium-The Milky Way and beyond, in IAU Symposium, Vol. 259, IAU Symposium, ed. K. G. Strassmeier, A. G. Kosovichev, & J. E. Beckman. 2009, 579–590

Heiles, C. & Troland, T. H., The Millennium Arecibo 21 Centimeter Absorption-Line Survey. IV. Statistics of Magnetic Field, Column Density, and Turbulence, ApJ **624** (May 2005) 773–793, `arXiv:astro-ph/0501482`

Hennessy, G. S., Owen, F. N., & Eilek, J. A., Faraday rotation from Abell clusters of galaxies, ApJ **347** (Dec. 1989) 144–151

Hoang, T. & Lazarian, A., Spinning Dust Emission from Wobbling Grains: Important Physical Effects and Implications, ArXiv e-prints (Oct. 2012), `arXiv:1211.0036` `[astro-ph.GA]`

Iatsenko, D., Stefanovska, A., & McClintock, P. V. E., Comment on "Inference with minimal Gibbs free energy in information field theory", Phys. Rev. E **85** no. 3, (Mar. 2012) 033101

Inoue, M. & Tabara, H., Structure of the Galactic Magnetic Field in the Solar Neighborhood, PASJ **33** (1981) 603–+, ADS

Isserlis, L., On a Formula for the Product-Moment Coefficient of any Order of a Normal Frequency Distribution in any Number of Variables, Biometrika **12** no. 1/2, (1918) pp. 134–139, `http://www.jstor.org/stable/2331932`

Jaffe, T. R., Leahy, J. P., Banday, A. J., et al., Modelling the Galactic magnetic field on the plane in two dimensions, MNRAS **401** (Jan. 2010) 1013–1028, `arXiv:0907.3994` `[astro-ph.GA]`

Jansson, R. & Farrar, G. R., A New Model of the Galactic Magnetic Field, ApJ **757** (Sep. 2012) 14, `arXiv:1204.3662 [astro-ph.GA]`

Jansson, R. & Farrar, G. R., The Galactic Magnetic Field, ApJ **761** (Dec. 2012) L11, `arXiv:1210.7820 [astro-ph.GA]`

Jansson, R., Farrar, G. R., Waelkens, A. H., & Enßlin, T. A., Constraining models of the large scale Galactic magnetic field with WMAP5 polarization data and extragalactic rotation measure sources, J. Cosmology Astropart. Phys. **7** (Jul. 2009) 21–+, `arXiv:0905.2228 [astro-ph.GA]`

Jasche, J., Kitaura, F. S., Li, C., & Enßlin, T. A., Bayesian non-linear large-scale structure inference of the Sloan Digital Sky Survey Data Release 7, MNRAS **409** (Nov. 2010) 355–370, `arXiv:0911.2498 [astro-ph.CO]`

Jasche, J., Kitaura, F. S., Wandelt, B. D., & Enßlin, T. A., Bayesian power-spectrum inference for large-scale structure data, MNRAS **406** (Jul. 2010) 60–85, `arXiv:0911.2493` `[astro-ph.CO]`

Jasche, J. & Wandelt, B. D., Bayesian physical reconstruction of initial conditions from large scale structure surveys, ArXiv e-prints (Mar. 2012), `arXiv:1203.3639` `[astro-ph.CO]`

Johnston, S., Nicastro, L., & Koribalski, B., Scintillation parameters for 49 pulsars, MNRAS **297** (Jun. 1998) 108–116

Johnston-Hollitt, M., e. a., in prep.

Johnston-Hollitt, M. 2003, PhD thesis, University of Adelaide

Johnston-Hollitt, M. & Ekers, R. D., Faraday Rotation Measures through the Cores of Southern Galaxy Clusters, ArXiv Astrophysics e-prints (Nov. 2004), `arXiv:astro-ph/0411045`

Johnston-Hollitt, M., Hollitt, C. P., & Ekers, R. D., Statistical Analysis of Extra-galactic Rotation Measures, in The Magnetized Interstellar Medium, ed. B. Uyaniker, W. Reich, & R. Wielebinski. 2004, 13–18, `arXiv:astro-ph/0410659`

Jones, C. & Dickey, J. M., Kinematic Distance Assignments with H I Absorption, ApJ **753** (Jul. 2012) 62

Junklewitz, H. & Enßlin, T. A., Imprints of magnetic power and helicity spectra on radio polarimetry statistics, A&A **530** (Jun. 2011) A88, `arXiv:1008.1243 [astro-ph.IM]`

Kahniashvili, T., Gogoberidze, G., & Ratra, B., Polarized Cosmological Gravitational Waves from Primordial Helical Turbulence, Physical Review Letters **95** no. 15, (Oct. 2005) 151301–+, `arXiv:astro-ph/0505628`

Kahniashvili, T. & Ratra, B., Effects of cosmological magnetic helicity on the cosmic microwave background, Phys. Rev. D **71** no. 10, (May 2005) 103006–+, `arXiv:astro-ph/0503709`

Kahniashvili, T. & Vachaspati, T., Detection of magnetic helicity, Phys. Rev. D **73** no. 6, (Mar. 2006) 063507–+, `arXiv:astro-ph/0511373`

Kato, T., Tabara, H., Inoue, M., & Aizu, K., Extragalactic radio sources with very large Faraday rotation, Nature **329** (Sep. 1987) 223–+

Kerr, F. J., The Large-Scale Distribution of Hydrogen in the Galaxy, ARA&A **7** (1969) 39

Kim, K.-T., Tribble, P. C., & Kronberg, P. P., Detection of excess rotation measure due to intracluster magnetic fields in clusters of galaxies, ApJ **379** (Sep. 1991) 80–88

Kissmann, R., Kleimann, J., Fichtner, H., & Grauer, R., Local turbulence simulations for the multiphase ISM, MNRAS **391** (Dec. 2008) 1577–1588, `arXiv:0901.4495 [astro-ph.GA]`

Kitaura, F. S. & Enßlin, T. A., Bayesian reconstruction of the cosmological large-scale structure: methodology, inverse algorithms and numerical optimization, MNRAS **389** (Sep. 2008) 497–544, `arXiv:0705.0429`

Kitaura, F.-S., Gallerani, S., & Ferrara, A., Multiscale inference of matter fields and baryon acoustic oscillations from the Lyα forest, MNRAS **420** (Feb. 2012) 61–74, `arXiv:1011.6233 [astro-ph.CO]`

Kitaura, F.-S., Jasche, J., & Metcalf, R. B., Recovering the non-linear density field from the galaxy distribution with a Poisson-lognormal filter, MNRAS **403** (Apr. 2010) 589–604, `arXiv:0911.1407 [astro-ph.CO]`

Klein, U., Mack, K.-H., Gregorini, L., & Vigotti, M., Multi-frequency study of the B3-VLA sample. III. Polarisation properties, A&A **406** (Aug. 2003) 579–592

Kogut, A., Anomalous Microwave Emission, in Astronomical Society of the Pacific Conference Series, Vol. 181, Microwave Foregrounds, ed. A. de Oliveira-Costa & M. Tegmark. 1999, 91, `arXiv:astro-ph/9902307`

Kogut, A., Banday, A. J., Bennett, C. L., et al., Microwave Emission at High Galactic Latitudes in the Four-Year DMR Sky Maps, ApJ **464** (Jun. 1996) L5, `arXiv:astro-ph/9601060`

Kolmogorov, A., The Local Structure of Turbulence in Incompressible Viscous Fluid for Very Large Reynolds' Numbers, Akademiia Nauk SSSR Doklady **30** (1941) 301–305, ADS

Koltchinskii, V. I., Empirical Geometry of Multivariate Data: A Deconvolution Approach, The Annals of Statistics **28** no. 2, (2000) 591–629, `http://www.jstor.org/stable/2674043`

Krause, M., Magnetic Fields and Star Formation in Galaxies of Different Morphological Types, in Astronomical Society of the Pacific Conference Series, Vol. 396, Formation and Evolution of Galaxy Disks, ed. J. G. Funes & E. M. Corsini. 2008, 147, `arXiv:0806.2047`

Kronberg, P. P. & Newton-McGee, K. J., Remarkable Symmetries in the Milky Way Disc's Magnetic Field, PASA **28** (Jun. 2011) 171–176, `arXiv:0909.4753 [astro-ph.GA]`

Kullback, S. & Leibler, R. A., On Information and Sufficiency, The Annals of Mathematical Statistics **22** no. 1, (1951) pp. 79–86, `http://www.jstor.org/stable/2236703`

Kuntz, K. D. & Danly, L., Intermediate-Velocity Gas in the North Galactic Hemisphere: H i Studies, ApJ **457** (Feb. 1996) 703

Lawler, J. M. & Dennison, B., On intracluster Faraday rotation. II - Statistical analysis, ApJ **252** (Jan. 1982) 81–91

Lazarian, A., Magnetic Fields via Polarimetry: Progress of Grain Alignment Theory, J. Quant. Spec. Radiat. Transf. **79** (2003) 881, `arXiv:astro-ph/0208487`

Lazarian, A. & Hoang, T., Alignment of Dust by Radiative Torque: Recent Developments, ArXiv e-prints (Dec. 2009), `arXiv:0901.0146`

Lemm, J. 2003, Bayesian Field Theory, , Johns Hopkins University Press

Lin, C. C. & Shu, F. H., On the Spiral Structure of Disk Galaxies., ApJ **140** (Aug. 1964) 646

Lockman, F. J., H I and Galactic Structure, in Astronomical Society of the Pacific Conference Series, Vol. 276, Seeing Through the Dust: The Detection of HI and the Exploration of the ISM in Galaxies, ed. A. R. Taylor, T. L. Landecker, & A. G. Willis. 2002, 107, `arXiv:astro-ph/0203210`

Lorimer, D. R., Binary and Millisecond Pulsars, Living Reviews in Relativity **11** no. 8, (2008), cited on 2012-12-06, `http://www.livingreviews.org/lrr-2008-8`

Mao, S. A., e. a., in prep.

Mao, S. A., Gaensler, B. M., Haverkorn, M., et al., A Survey of Extragalactic Faraday Rotation at High Galactic Latitude: The Vertical Magnetic Field of the Milky Way Toward the Galactic Poles, ApJ **714** (May 2010) 1170–1186, `arXiv:1003.4519 [astro-ph.GA]`

Mao, S. A., Gaensler, B. M., Stanimirović, S., et al., A Radio and Optical Polarization Study of the Magnetic Field in the Small Magellanic Cloud, ApJ **688** (Dec. 2008) 1029–1049, `arXiv:0807.1532`

Maoz, D. 2011, Astrophysics in a Nutshell, , Princeton University Press

Marshall, D. J., Robin, A. C., Reylé, C., Schultheis, M., & Picaud, S., Modelling the Galactic interstellar extinction distribution in three dimensions, A&A **453** (Jul. 2006) 635–651, `arXiv:astro-ph/0604427`

McClure-Griffiths, N. M., Dickey, J. M., Gaensler, B. M., et al., The Southern Galactic Plane Survey: H I Observations and Analysis, ApJS **158** (Jun. 2005) 178–187, `arXiv:astro-ph/0503134`

McClure-Griffiths, N. M., Pisano, D. J., Calabretta, M. R., et al., Gass: The Parkes Galactic All-Sky Survey. I. Survey Description, Goals, and Initial Data Release, ApJS **181** (Apr. 2009) 398–412, `arXiv:0901.1159 [astro-ph.GA]`

Men, H., Ferrière, K., & Han, J. L., Observational constraints on models for the interstellar magnetic field in the Galactic disk, A&A **486** (Aug. 2008) 819–828, `arXiv:0805.3454`

Minter, A. H. & Spangler, S. R., Observation of Turbulent Fluctuations in the Interstellar Plasma Density and Magnetic Field on Spatial Scales of 0.01 to 100 Parsecs, ApJ **458** (Feb. 1996) 194–+

Møller, J., Syversveen, A. R., & Waagepetersen, R. P., Log Gaussian Cox Processes, Scandinavian Journal of Statistics **25** no. 3, (1998), `http://dx.doi.org/10.2307/4616515`

Morris, D. & Berge, G. L., Direction of the Galactic Magnetic Field in the Vicinity of the Sun., ApJ **139** (May 1964) 1388–+

Nakanishi, H. & Sofue, Y., Three-Dimensional Distribution of the ISM in the Milky Way Galaxy: I. The H I Disk, PASJ **55** (Feb. 2003) 191–202, `arXiv:astro-ph/0304338`

Nakanishi, H. & Sofue, Y., Three-Dimensional Distribution of the ISM in the Milky Way Galaxy: II. The Molecular Gas Disk, PASJ **58** (Oct. 2006) 847–860, `arXiv:astro-ph/0610769`

Navarro, J. F., Frenk, C. S., & White, S. D. M., A Universal Density Profile from Hierarchical Clustering, ApJ **490** (Dec. 1997) 493, `arXiv:astro-ph/9611107`

Neal, R. M. 1993, Probabilistic Inference using Markov Chain Monte Carlo Methods, Tech. Rep. CRG-TR-93-1, Dept. of Computer Science, University of Toronto, 144pp.

Nicholson, D. R. 1983, Introduction to plasma physics, Wiley, New York

Nota, T. & Katgert, P., The large-scale magnetic field in the fourth Galactic quadrant, A&A **513** (Apr. 2010) A65+

Oh, S. P., Spergel, D. N., & Hinshaw, G., An Efficient Technique to Determine the Power Spectrum from Cosmic Microwave Background Sky Maps, ApJ **510** (Jan. 1999) 551–563, `arXiv:astro-ph/9805339`

Olsen, S. I., Estimation of noise in images: an evaluation, CVGIP: Graph. Models Image Process. **55** (July 1993) 319–323, `http://portal.acm.org/citation.cfm?id=167608.167619`

Oppermann, N., Junklewitz, H., Robbers, G., et al., An improved map of the Galactic Faraday sky, A&A **542** (Jun. 2012) A93, `arXiv:1111.6186 [astro-ph.GA]`

Oppermann, N., Junklewitz, H., Robbers, G., & Enßlin, T. A., Probing magnetic helicity with synchrotron radiation and Faraday rotation, A&A **530** (Jun. 2011) A89+, `arXiv:1008.1246 [astro-ph.IM]`

Oppermann, N., Robbers, G., & Enßlin, T. A., Reconstructing signals from noisy data with unknown signal and noise covariance, Phys. Rev. E **84** no. 4, (Oct. 2011) 041118, `arXiv:1107.2384 [astro-ph.IM]`

Oppermann, N., Selig, M., Bell, M. R., & Enßlin, T. A., Reconstruction of Gaussian and log-normal fields with spectral smoothness, ArXiv e-prints (Oct. 2012), `arXiv:1210.6866 [astro-ph.IM]`

Oren, A. L. & Wolfe, A. M., A Faraday rotation search for magnetic fields in quasar damped LY alpha absorption systems, ApJ **445** (Jun. 1995) 624–641

Page, L., Hinshaw, G., Komatsu, E., et al., Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Polarization Analysis, ApJS **170** (Jun. 2007) 335–376, `arXiv:astro-ph/0603450`

Pati, D., Reich, B. J., & Dunson, D. B., Bayesian geostatistical modelling with informative sampling locations, Biometrika **98** no. 1, (2011) 35–48, `http://EconPapers.repec.org/RePEc:oup:biomet:v:98:y:2011:i:1:p:35-48`

Petersen, K. B. & Pedersen, M. S. 2008, The Matrix Cookbook, version 20081110, `http://www2.imm.dtu.dk/pubdb/p.php?3274`

Ponthieu, N., Macías-Pérez, J. F., Tristram, M., et al., Temperature and polarization angular power spectra of Galactic dust radiation at 353 GHz as measured by Archeops, A&A **444** (Dec. 2005) 327–336, `arXiv:astro-ph/0501427`

Pshirkov, M. S., Tinyakov, P. G., Kronberg, P. P., & Newton-McGee, K. J., Deriving the Global Structure of the Galactic Magnetic Field from Faraday Rotation Measures of Extragalactic Sources, ApJ **738** (Sep. 2011) 192, `arXiv:1103.0814 [astro-ph.GA]`

Rand, R. J. & Kulkarni, S. R., The local Galactic magnetic field, ApJ **343** (Aug. 1989) 760–772

Rand, R. J. & Lyne, A. G., New Rotation Measures of Distant Pulsars in the Inner Galaxy and Magnetic Field Reversals, MNRAS **268** (May 1994) 497–+, ADS

Reynolds, R. J., High spatial and spectral resolution pictures of the faint galactic H alpha background, ApJ **236** (Feb. 1980) 153–159

Reynolds, R. J., Roesler, F. L., & Scherb, F., The Intensity Distribution of Diffuse Galactic Hα Emission, ApJ **192** (Sep. 1974) L53

Rickett, B. J., Interstellar scattering and scintillation of radio waves, ARA&A **15** (1977) 479–504

Rickett, B. J., Radio propagation through the turbulent interstellar plasma, ARA&A **28** (1990) 561–605

Robert, C. & Casella, G. 2004, Monte Carlo Statistical Methods, Springer Texts in Statistics, , Springer

Roy, S., Pramesh Rao, A., & Subrahmanyan, R., Magnetic field near the central region of the Galaxy: rotation measure of extragalactic sources, A&A **478** (Feb. 2008) 435–442, `arXiv:0712.0269`

Roy, S., Rao, A. P., & Subrahmanyan, R., Extragalactic sources towards the central region of the Galaxy, MNRAS **360** (Jul. 2005) 1305–1324, `arXiv:0712.0258`

Rudnick, L. & Jones, T. W., Rotation measures for compact variable radio sources, AJ **88** (Apr. 1983) 518–526

Ruzmaikin, A. A. & Sokolov, D. D., The interpretation of rotation measures of extragalactic radio sources, A&A **58** (Jun. 1977) 247–253, ADS

Ruzmaikin, A. A., Sokolov, D. D., & Kovalenko, A. V., Galactic magnetic field parameters determined from Faraday rotation of radio sources, Soviet Ast. **22** (Aug. 1978) 395–+, ADS

Rybicki, G. & Lightman, A. 2008, Radiative Processes in Astrophysics, Physics textbook, , Wiley

Schlegel, D. J., Finkbeiner, D. P., & Davis, M., Maps of Dust Infrared Emission for Use in Estimation of Reddening and Cosmic Microwave Background Radiation Foregrounds, ApJ **500** (Jun. 1998) 525, `arXiv:astro-ph/9710327`

Schlickeiser, R. 2002, Cosmic Ray Astrophysics, Astronomy and Astrophysics Library, , Springer

Schnitzeler, D. H. F. M., The latitude dependence of the rotation measures of NVSS sources, MNRAS(Oct. 2010) L160+, `arXiv:1011.0737 [astro-ph.GA]`

Schnitzeler, D. H. F. M., Modelling the Galactic distribution of free electrons, MNRAS **427** (Nov. 2012) 664–678, `arXiv:1208.3045 [astro-ph.GA]`

Schnitzeler, D. H. F. M., e. a., in prep.

Schwarz, M. & van Bellegem, S., Consistent density deconvolution under partially known error distribution, Statistics & Probability Letters **80** no. 3-4, (2010) 236 – 241, `http://www.sciencedirect.com/science/article/pii/S0167715209003988`

Selig, M., e. a., in prep.

Selig, M., Oppermann, N., & Enßlin, T. A., Improving stochastic estimates with inference methods: Calculating matrix diagonals, Phys. Rev. E **85** no. 2, (Feb. 2012) 021134, `arXiv:1108.0600 [astro-ph.IM]`

Seymour, P. A. H., Models of the galatic magnetic filed, MNRAS **134** (1966) 389–+, ADS

Seymour, P. A. H., Faraday Rotation and the Galactic Magnetic Field - a Review, QJRAS **25** (Sep. 1984) 293–+, ADS

Short, M. B., Higdon, D. M., & Kronberg, P., P., Estimation of Faraday Rotation Measures of the Near Galactic Sky Using Gaussian Process Models, Bayesian Analysis **2** (Jun. 2007) 665–680, ADS

Simard-Normandin, M. & Kronberg, P. P., New large-scale magnetic features of the Milky Way, Nature **279** (May 1979) 115–118

Simard-Normandin, M. & Kronberg, P. P., Rotation measures and the galactic magnetic field, ApJ **242** (Nov. 1980) 74–94

Simard-Normandin, M., Kronberg, P. P., & Button, S., The Faraday rotation measures of extragalactic radio sources, ApJS **45** (Jan. 1981) 97–111

Simonetti, J. H. & Cordes, J. M., Small-scale variations in the galactic magnetic field Rotation measure variations across extragalactic radio sources, ApJ **310** (Nov. 1986) 160–171

Simonetti, J. H., Cordes, J. M., & Spangler, S. R., Small-scale variations in the galactic magnetic field - The rotation measure structure function and birefringence in interstellar scintillations, ApJ **284** (Sep. 1984) 126–134

Smoot, G. F., Galactic Free-free and H-alpha Emission, ArXiv Astrophysics e-prints (Jan. 1998), `arXiv:astro-ph/9801121`

Sofue, Y. & Fujimoto, M., A bisymmetric spiral magnetic field and the spiral arms in our Galaxy, ApJ **265** (Feb. 1983) 722–729

Spitzer, L. 2007, Physical Processes in the Interstellar Medium, , Wiley-VCH Verlag GmbH

Stanev, T., Ultra–High-Energy Cosmic Rays and the Large-Scale Structure of the Galactic Magnetic Field, ApJ **479** (Apr. 1997) 290–+, `arXiv:astro-ph/9607086`

Starck, J.-L. & Murtagh, F., Automatic Noise Estimation from the Multiresolution Support, PASP **110** (Feb. 1998) 193–199

Stil, J. M. & Taylor, A. R., Polarization Shadows of Extragalactic Sources by the Local Magnetoionic Interstellar Medium, ApJ **663** (Jul. 2007) L21–L24, `arXiv:0705.2741`

Stil, J. M., Taylor, A. R., & Sunstrum, C., Structure in the Rotation Measure Sky, ApJ **726** (Jan. 2011) 4–+, `arXiv:1010.5299 [astro-ph.GA]`

Stroman, W. & Pohl, M., Radio Polarimetry Signatures of Strong Magnetic Turbulence in Supernova Remnants, ApJ **696** (May 2009) 1864–1870, `arXiv:0902.1701 [astro-ph.HE]`

Strong, A. W. & Moskalenko, I. V., Models for galactic cosmic-ray propagation, Advances in Space Research **27** (2001) 717–726, `arXiv:astro-ph/0101068`

Strong, A. W., Moskalenko, I. V., & Ptuskin, V. S., Cosmic-Ray Propagation and Interactions in the Galaxy, Annual Review of Nuclear and Particle Science **57** (Nov. 2007) 285–327, `arXiv:astro-ph/0701517`

Subramanian, K., Magnetic helicity in galactic dynamos., Bulletin of the Astronomical Society of India **30** (Dec. 2002) 715–721, `arXiv:astro-ph/0204450`

Sun, X. H. & Han, J. L., Structure Function Studies for Turbulent Interstellar Medium, in The Magnetized Interstellar Medium, ed. B. Uyaniker, W. Reich, & R. Wielebinski. 2004, 25–30, `arXiv:astro-ph/0402180`

Sun, X. H., Reich, W., Waelkens, A., & Enßlin, T. A., Radio observational constraints on Galactic 3D-emission models, A&A **477** (Jan. 2008) 573–592, `arXiv:0711.1572`

Sunstrum, C., Taylor, A. R., & Stil, J., Rotation Measures from Two Data Points - the NVSS Rotation Measure Catalogue, in Astronomical Society of the Pacific Conference Series, Vol. 438, Astronomical Society of the Pacific Conference Series, ed. R. Kothes, T. L. Landecker, & A. G. Willis. 2010, 296–+, ADS

Tabara, H. & Inoue, M., A catalogue of linear polarization of radio sources, A&AS **39** (Mar. 1980) 379–393, ADS

Taylor, A. R., Gibson, S. J., Peracaula, M., et al., The Canadian Galactic Plane Survey, AJ **125** (Jun. 2003) 3145–3164

Taylor, A. R., Stil, J. M., & Sunstrum, C., A Rotation Measure Image of the Sky, ApJ **702** (Sep. 2009) 1230–1236

Taylor, J. H. & Cordes, J. M., Pulsar distances and the galactic distribution of free electrons, ApJ **411** (Jul. 1993) 674–684

Tofflemire, B. M., Burkhart, B., & Lazarian, A., Interstellar Sonic and Alfvénic Mach Numbers and the Tsallis Distribution, ApJ **736** (Jul. 2011) 60–+, `arXiv:1103.3299 [astro-ph.GA]`

Uhlenbeck, G. E. & Ornstein, L. S., On the Theory of the Brownian Motion, Phys. Rev. **36** (Sep 1930) 823–841, `http://link.aps.org/doi/10.1103/PhysRev.36.823`

Vallée, J. P., The rotation measures of radio sources in selected celestial zones - The Perseus arm window, A&AS **51** (Jan. 1983) 127–129, ADS

Vallée, J. P., Pulsar-based Galactic Magnetic Map: A Large-Scale Clockwise Magnetic Field with an Anticlockwise Annulus, ApJ **619** (Jan. 2005) 297–305

Vallée, J. P., An Improved Magnetic Map of the Milky Way, with the Circularly Orbiting Gas and Magnetic Field Lines Crossing the Dusty Stellar Spiral Arms, ApJ **681** (Jul. 2008) 303–310

Vallee, J. P. & Bignell, R. C., Magnetic field structure in the GUM Nebula area, ApJ **272** (Sep. 1983) 131–136

Vallée, J. P. & Kronberg, P. P., Galaxy-Magnetic field in the galactic spiral arm, Nature **246** (Nov. 1973) 49–51, ADS

Van Eck, C. L., Brown, J. C., Stil, J. M., et al., Modeling the Magnetic Field in the Galactic Disk Using New Rotation Measure Observations from the Very Large Array, ApJ **728** (Feb. 2011) 97–+, `arXiv:1012.2938 [astro-ph.GA]`

Vigotti, M., Gregorini, L., Klein, U., & Mack, K.-H., Multi-frequency study of the B3-VLA sample. II. The database, A&AS **139** (Oct. 1999) 359–376, `arXiv:astro-ph/9909011`

Volegova, A. A. & Stepanov, R. A., Helicity detection of astrophysical magnetic fields from radio emission statistics, Soviet Journal of Experimental and Theoretical Physics Letters **90** (Jan. 2010) 637–641, `arXiv:1001.2857 [astro-ph.IM]`

Waelkens, A., Jaffe, T., Reinecke, M., Kitaura, F. S., & Enßlin, T. A., Simulating polarized Galactic synchrotron emission at all frequencies. The Hammurabi code, A&A **495** (Feb. 2009) 697–706, `arXiv:0807.2262`

Walsh, A. J., Breen, S. L., Britton, T., et al., The H_2O Southern Galactic Plane Survey (HOPS) - I. Techniques and H_2O maser data, MNRAS **416** (Sep. 2011) 1764–1821, `arXiv:1105.4663 [astro-ph.GA]`

Wandelt, B. D., Larson, D. L., & Lakshminarayanan, A., Global, exact cosmic microwave background data analysis using Gibbs sampling, Phys. Rev. D **70** no. 8, (Oct. 2004) 083511, `arXiv:astro-ph/0310080`

Weig, C. & Enßlin, T. A., Bayesian analysis of spatially distorted cosmic signals from Poissonian data, MNRAS **409** (Dec. 2010) 1393–1411, `arXiv:1003.1311 [astro-ph.CO]`

Weingartner, J. C. & Draine, B. T., Dust Grain-Size Distributions and Extinction in the Milky Way, Large Magellanic Cloud, and Small Magellanic Cloud, ApJ **548** (Feb. 2001) 296–309, `arXiv:astro-ph/0008146`

Weisberg, J. M., Cordes, J. M., Kuan, B., et al., Arecibo 430 MHz Pulsar Polarimetry: Faraday Rotation Measures and Morphological Classifications, ApJS **150** (Jan. 2004) 317–341, `arXiv:astro-ph/0310073`

Widrow, L. M., Origin of galactic and extragalactic magnetic fields, Reviews of Modern Physics **74** (2002) 775–823, `arXiv:astro-ph/0207240`

Wolleben, M., Fletcher, A., Landecker, T. L., et al., Antisymmetry in the Faraday Rotation Sky Caused by a Nearby Magnetized Bubble, ApJ **724** (Nov. 2010) L48–L52, `arXiv:1011.0341 [astro-ph.GA]`

Wrobel, J. M., Faraday rotation measures and intrinsic polarization position angles of very long baseline interferometry core-jet sources, AJ **106** (Aug. 1993) 444–454

Xu, Y., Kronberg, P. P., Habib, S., & Dufton, Q. W., A Faraday Rotation Search for Magnetic Fields in Large-scale Structure, ApJ **637** (Jan. 2006) 19–26, `arXiv:astro-ph/0509826`

Yan, H. & Lazarian, A., Cosmic-Ray Propagation: Nonlinear Diffusion Parallel and Perpendicular to Mean Magnetic Field, ApJ **673** (Feb. 2008) 942–953, `arXiv:0710.2617`

Zhang, H., Helicity of solar magnetic field from observations, in IAU Symposium, Vol. 264, IAU Symposium, ed. A. G. Kosovichev, A. H. Andrei, & J.-P. Roelot. 2010, 181–190

Acknowledgments

Many people have contributed to the success of this thesis. First and foremost, these are the members of the Enßlin group at MPA. Most of the results presented here were born in one way or another out of countless blackboard discussions. The following people deserve to be mentioned for their contributions to these discussions, as well as to life in- and outside of the institute:

- Henrik, who knows stuff about field theory and has accompanied me and my work from the very beginning and often late into the night.

- Mike, who actually knows astronomy and brought an international flavor to the group.

- Marco, who knows more about pixel volume factors than anyone else and who helped me and generations to come by thinking these questions through.

- Georg, who is in part responsible for the smooth start that I had at MPA and will forever be a reminder that there may be alternatives.

- Mona and Jens, who unfortunately left shortly after I came.

- Our undergrads, Helin, Max Uh., Max Ul., Lars, Maksim, and Sebastian, who provided many new questions and answered some old ones.

- The MPA *Planck* group, Jörg K., Martin, Wolfgang, Jörg R., Thomas, Theo, and Uwe, who taught me many things about the workings of the collaboration and were always helpful in times of computer crisis.

Most of all, however, I should thank my thesis advisor, Torsten Enßlin, who managed to supervise without patronizing, to guide without restricting, and to inspire without daydreaming (almost). It is hard to imagine an advisor who is more motivated or more willing to invest time into a student's project. Whether in front of the blackboard, in front of a computer screen, or in front of a glass of beer, his advice was always welcome and most of the time helpful. During the past years, he has guided me through the world of astrophysics, through large scientific collaborations, and through the Paris nightlife. All the time, he never got tired of promoting my work whenever an opportunity appeared.

University regulations have prevented his name from appearing at the front of this thesis, where it belongs.

Finally, I should thank all people at MPA for making these three short years possible and enjoyable. Thanks to the directors and the administration for giving me the opportunity to come to MPA and to leave MPA often for schools and conferences; thanks to Simon for acting as my official advisor; thanks to Marat for supporting me as a member of my TAC; thanks to too many students and post-docs to name here for the activities in our abundant spare time.

Curriculum Vitae

Personal data

Name:	Niels Oppermann
Date of birth:	October 5$^{\text{th}}$, 1985
Place of birth:	Göttingen, Germany
Address:	Liebigstraße 20 B, 80538 München, Germany

Contact:

phone:	+49-89-97891285
mobile:	+49-1577-3369243
office:	+49-89-30000-2269
e-mail:	niels@mpa-garching.mpg.de

Science and education

since Nov. 2011:	Member of the Planck mission as a Low Frequency Instrument Core Team member and a High Frequency Instrument associate
since Jan. 2010:	PhD student at the Max Planck Institute for Astrophysics in Garching, Germany, under the supervision of Dr. Torsten A. Enßlin and Prof. Dr. Simon D. M. White; project topic: "Signal inference in Galactic astrophysics"
Oct. 2009:	Received Diploma of Physics (with distinction) from Eberhard-Karls-Universität in Tübingen, Germany
2008 - 2009:	Diploma Thesis on "Gravitational Collapse of a Magnetized Star" in the Theoretical Astrophysics group at Eberhard-Karls-Universität in Tübingen, Germany, under the supervision of Prof. Dr. Konstantinos D. Kokkotas and Dr. Hajime Sotani
2004 - 2009:	Studies in physics at the Eberhard-Karls-Universität in Tübingen, Germany

Teaching experience

since 2012	Co-supervision of a Master's thesis on primordial non-Gaussianities
2011:	Co-supervision of a Bachelor's thesis on "A map of the symmetric part of the Faraday depth"
2011 & 2012:	Conducting the exercises for the lecture Information Theory and Signal Reconstruction
2007 - 2008:	Tutoring an exercise group for the lectures Mathematics for Physicists – I and – II

Refereed Publications

2012 The XENON100 exclusion limit without considering L_{eff} as a nuisance parameter
*J. H. Davis, C. Bœhm, **N. Oppermann**, T. A. Enßlin, and Th. Lacroix*
Physical Review D, Volume 86, Issue 1, 07/2012

An improved map of the Galactic Faraday sky
***N. Oppermann**, H. Junklewitz, G. Robbers, M. R. Bell, T. A. Enßlin, A. Bonafede, R. Braun, J.-A. C. Brown, T. E. Clarke, I. J. Feain, B. M. Gaensler, A. Hammond, L. Harvey-Smith, G. Heald, M. Johnston-Hollitt, U. Klein, P. P. Kronberg, S. A. Mao, N. M. McClure-Griffiths, S. P. O'Sullivan, L. Pratley, T. Robishaw, S. Roy, D. H. F. M. Schnitzeler, C. Sotomayor-Beltran, J. Stevens, J. M. Stil, C. Sunstrum, A. Tanna, A. R. Taylor, and C. L. Van Eck*
Astronomy & Astrophysics, Volume 542, id.A93, 06/2012

Improving stochastic estimates with inference methods: calculating matrix diagonals
*M. Selig, **N. Oppermann**, T. A. Enßlin*
Physical Review E, Volume 85, Issue 2, 02/2012

2011 Reconstructing signals from noisy data with unknown signal and noise covariance
***N. Oppermann**, G. Robbers, T. A. Enßlin*
Physical Review E, Volume 84, Issue 4, 10/2011

Probing magnetic helicity with synchrotron radiation and Faraday rotation
***N. Oppermann**, H. Junklewitz, G. Robbers, T. A. Enßlin*
Astronomy & Astrophysics, Volume 530, id.A89, 06/2011

Preprints

2012 Improved CLEAN reconstructions for rotation measure synthesis with maximum likelihood estimation
*M. R. Bell, **N. Oppermann**, A. Crai, T. A. Enßlin*
arXiv:1211.5105; submitted to Astronomy & Astrophysics

Reconstruction of Gaussian and log-normal fields with spectral smoothness
***N. Oppermann**, M. Selig, M. R. Bell, and T. A. Enßlin*
arXiv:1210.6866; submitted to Physical Review E

Presentations at scientific meetings and conferences

November 2012:   "The Galactic Faraday sky – What it is, how it's done, and why it's useful"
Talk given at the *workshop on Polarized Foreground for Cosmic Microwave Background* in Garching, Germany

July 2012:   "Reconstructing signals from noisy data with unknown signal and noise covariance"
Talk given at the *32nd International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering* in Garching, Germany
Proceeding in review

July 2012:   "The Galactic Faraday sky – What it is, how it's done, and why it's useful"
Talk given at the *annual meeting of the DFG research unit 1254: "Magnetisation of Interstellar and Intergalactic Media – The Prospects of Low-frequency Radio Observations"* in Mainz, Germany

February 2012:   "The Galactic Faraday sky – What it is, how it's done, and why it's useful"
Talk given at the conference on *Astrophysics from the Radio to the Sub-millimeter – Planck and Other Experiments in Temperature and Polarization* in Bologna, Italy

December 2011:   "The Faraday sky – mapmaking and helicity inference"
Invited seminar talk at the Institut d'Astrophysique Spatiale in Orsay, France

May 2011:   "Reconstructing CMB Foregrounds and the Faraday Sky"
Talk given at the conference on *Understanding Galactic and Extragalactic Foregrounds: a Road to Success for Cosmological Experiments* in Zadar, Croatia

May 2011:   "The Faraday sky – mapmaking and helicity inference"
Talk given at the *German-Russian Open Workshop on "Galactic Magnetism – Perspectives of Observation and Modeling"* in Pushchino, Russia

January 2011:   "Detecting magnetic helicity from polarization at CMB frequencies"
Poster presentation at the conference on *The Millimeter and Submillimeter Sky in the Planck Mission Era* in Paris, France