Astronomy & Astrophysics manuscript no. note
April 15, 2014

Estimating extragalactic Faraday rotation

N. Oppermann*[1,2], H. Junklewitz[2,3], M. Greiner[2], T.A. Enßlin[2,4], T. Akahori[5], E. Carretti[6], B.M. Gaensler[5], A. Goobar[7], L. Harvey-Smith[6], M. Johnston-Hollitt[8], L. Pratley[8], D.H.F.M. Schnitzeler[9], J.M. Stil[10], and V. Vacca[2]

[1] Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto ON, M5S 3H8, Canada
[2] Max Planck Institute for Astrophysics, Karl-Schwarzschild-Str. 1, 85748 Garching, Germany
[3] Argelander-Institut für Astronomie, Auf dem Hügel 71, 52121 Bonn, Germany
[4] Ludwig-Maximilians-Universität München, Fakultät für Physik, D-80799 München, Germany
[5] Sydney Institute for Astronomy, School of Physics, The University of Sydney, NSW 2006, Australia
[6] CSIRO Astronomy and Space Science, PO Box 76, Epping, NSW 1700, Australia
[7] The Oskar Klein Centre, Physics Department, Stockholm University, Albanova University Center, SE 106 91 Stockholm, Sweden
[8] School of Chemical & Physical Sciences, Victoria University of Wellington, PO Box 600, Wellington 6014, New Zealand
[9] Max Planck Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany
[10] Department of Physics and Astronomy, The University of Calgary, 2500 University Drive NW, Calgary AB, T2N 1N4, Canada

April 15, 2014

ABSTRACT

Observations of Faraday rotation for extragalactic sources probe magnetic fields both inside and outside the Milky Way. Building on our earlier estimate of the Galactic contribution (Oppermann et al. 2012), we set out to estimate the extragalactic contributions. We discuss different strategies and the problems involved. In particular, we point out that taking the difference between the observed values and the Galactic foreground reconstruction is not a good estimate for the extragalactic contributions. We present a few possibilities for improved estimates using the foreground map of Oppermann et al. (2012), allowing for imperfectly described observational noise. In this context, we point out a degeneracy between the contributions to the observed values due to extragalactic magnetic fields and observational noise and comment on the dangers of over-interpreting an estimate without taking into account its uncertainty information. Finally, we develop an extended reconstruction algorithm based on the assumption that the observational uncertainties are accurately described for a subset of the data, which can overcome this degeneracy. We present a probabilistic derivation of the algorithm and demonstrate its performance using a simulation, yielding a high quality reconstruction of the Galactic Faraday rotation foreground, a precise estimate of the typical extragalactic contribution, and a well-defined probabilistic description of the extragalactic contribution for each data point. We then apply this reconstruction technique to the catalog of Faraday rotation observations for extragalactic sources of Oppermann et al. (2012) after adding 302 new sources. The analysis is done for several different scenarios, for which we consider the error bars of different subsets of the data to accurately describe the observational uncertainties. By comparing the results, we argue that a split that singles out only data near the Galactic poles is the most robust approach. We find that the dispersion of extragalactic contributions to observed Faraday depths is likely lower than 7 rad/m^2, in agreement with earlier results, and that the extragalactic contribution to an individual data point is poorly constrained by the data in most cases. Posterior samples for the extragalactic contribution to all data points and all results of our fiducial model are provided online.

Key words. Magnetic fields – Methods: data analysis – ISM: magnetic fields

1. Introduction

Polarized radiation from an astronomical source undergoes Faraday rotation as it travels through the magneto-ionic medium between the source and observer. For extragalactic sources, there are contributions from the Galactic interstellar medium, from any intergalactic magnetic fields, from intervening galaxies on the line of sight, as well as from magnetic fields in the source itself. In this work we attempt to estimate the contribution to the observed Faraday rotation of such sources that is due to magnetic fields outside of the Milky Way. This extragalactic contribution holds the potential for extracting information about cosmic magnetic fields on large scales, e.g. in galaxy clusters, galaxy filaments, or cosmic voids (Kolatt 1998; Blasi et al. 1999; Xu et al. 2006; Hammond et al. 2012; Bernet et al. 2012; Neronov et al. 2013; Joshi & Chand 2013).

For a hypothetical source of linear polarization that is point-like in all dimensions and situated at a physical distance r from the observer, the change in polarization angle is given by

$$\Delta_\chi = \phi\,\lambda^2, \tag{1}$$

where λ is the wavelength of the radiation and

$$\phi = \frac{\mathrm{e}^3}{2\pi\, m_\mathrm{e}^2\, c^4} \int_r^0 \mathrm{d}r' \, \frac{1}{(1+z(r'))^2}\, n_\mathrm{e}(r')\, B_r(r') \tag{2}$$

is the Faraday depth of the source (e.g. Burn 1966). In the last equation, n_e is the density of thermal electrons, B_r is the magnetic field vector projected onto the line of sight, z is the cosmological redshift, and the pre factor is a function of the electron charge e, the electron mass m_e, and the speed of light c. Theoretically, the line of sight integral in the definition of the Faraday depth can simply be split into an integral over the portion of the line of sight that lies within the Milky Way and the portion that

* e-mail: niels@cita.utoronto.ca

is outside the Milky Way. This allows one to write the total Faraday depth of the source as the sum of the Galactic Faraday depth in the direction of the source and an extragalactic contribution,

$$\phi = \phi_{\mathrm{g}} + \phi_{\mathrm{e}}. \tag{3}$$

For most extragalactic sources, the net rotation is dominated by the effect of the interstellar medium of the Milky Way (Leahy 1987; Schnitzeler 2010). This Galactic contribution has recently been estimated from a collection of observations of Faraday rotation of extragalactic sources (Oppermann et al. 2012). One way of estimating the extragalactic contributions is to subtract the estimate of the Galactic contribution from the observed values. However, we will argue that this is not a good estimate due to the presence of uncertainties both in the observations and in the foreground estimate.

Extracting the sub-dominant extragalactic part from the data is more difficult than extracting the Galactic part for two reasons. The first obvious reason is that, as it is smaller, it is more heavily obscured by observational noise in the data. In fact, for many of the data points that we use the expected extragalactic contribution is comparable to the observational uncertainty. The second reason is that the Galactic foreground contribution is spatially smooth to some extent, which enables the usage of correlation information and thus the transfer of information from many data points to each location on the sky. The extragalactic contributions, on the other hand, are expected to be mostly uncorrelated from source to source, given the typical source separation. Therefore, information on the extragalactic contribution to a data point is only contained in the other data points indirectly via the common Galactic foreground. Furthermore, the observational uncertainties are uncorrelated from source to source as well leading to a statistical degeneracy with the extragalactic contributions. This means that any split between extragalactic contributions and observational noise in the estimate can only be made according to the expected variances of these two components.

We will additionally argue that the statistical characterization of the observational uncertainties given by the error bars in the data catalogs we use may be incomplete in some cases. Therefore, there is an additional degree of freedom in the expected noise variance that further exacerbates the degeneracy between observational noise and extragalactic contributions.

While we will ultimately allow the expected extragalactic variance to be reconstructed from the data themselves, we use an initial estimate built on the work by Schnitzeler (2010), who studied the spread of the distribution of observed Faraday depths of extragalactic sources from the catalog of Taylor et al. (2009). Schnitzeler (2010) observed that this spread changes as a function of Galactic latitude even after the subtraction of a coarse foreground model. He then extracted an upper bound on the spread of the extragalactic contributions as the latitude independent part of this function.

In the following Sect. 2, we give precise definitions for the terminology that we use in the later discussion. Terms like noise and estimate are introduced and we explain why we regard the observational uncertainty itself as uncertain. In Sect. 3, we present a way to estimate the extragalactic contributions using the foreground map of Oppermann et al. (2012). We do this by making some crude assumptions about the expected variances of the extragalactic and noise contributions and we discuss in more detail the degeneracy problem alluded to before. We also discuss artifacts emerging in the estimate and point out possible pitfalls when over-interpreting an estimate without taking into account the corresponding uncertainty information. In Sect. 4, we derive an extension to the algorithm of Oppermann et al. (2012) that can overcome the degeneracies and thus yield estimates both for the Galactic foreground and the extragalactic contributions. We demonstrate the performance of the algorithm in a simulated scenario. In Sect. 5, we apply the algorithm to observational data and present the results. We perform different case studies to gauge the robustness of these results. Finally, we give a brief summary in Sect. 6.

2. Data model and terminology

The observed Faraday depth of the i-th source in a catalog, or data point d_i, is comprised of a Galactic contribution $\phi_{\mathrm{g},i}$, an extragalactic contribution $\phi_{\mathrm{e},i}$, and observational noise n_i,

$$d_i = \phi_{\mathrm{g},i} + \phi_{\mathrm{e},i} + n_i. \tag{4}$$

This equation holds for each data point, or equivalently, for all data points at once when one summarizes as vectors d and n the observational estimates and their uncertainties for each source and writes $\phi_{\mathrm{g/e}}$ as vectors containing the Galactic and extragalactic contributions to the Faraday rotation along all these lines of sight, respectively.

Before discussing possible ways of estimating individual constituents, we introduce the terminology that we will use in our discussion, explain why we need to allow for some uncertainty in the description of the observational uncertainties, and elaborate on the structure we will assume throughout this work for the correlations of the three constituents.

2.1. Terminology

For each source in nature, there are two individual numbers giving the Galactic and extragalactic contributions to the source's Faraday depth, respectively. We denote these numbers as $\phi_{\mathrm{g/e}}$. The *definition* of the noise term, n, is then simply the difference between the measured Faraday depth, d, and the sum of these two numbers,

$$n = d - \left(\phi_{\mathrm{g}} + \phi_{\mathrm{e}}\right). \tag{5}$$

The measurement and the noise contribution are each also given by a single number. Note, however, that the measurement itself is the only number that is known exactly. In fact, without further input, the three constituents ϕ_{g}, ϕ_{e}, and n are completely degenerate and therefore completely unknown.

Using additional input, which we will discuss in the following, it may be possible to construct reasonable ways to estimate the three numbers adding up to the measured number. We will use hatted variables to denote such estimated quantities, e.g. $\hat{\phi}_{\mathrm{g}}$ will be an estimate of the Galactic contribution. Distinguishing between the true numbers realized in nature and our estimates of these numbers is crucial.

The qualitative difference between the noise and the other two constituents is that we usually do not aim for an estimate of the noise itself. Noise can arise due to instrumental effects, features of the data processing, or the presence of any other physical effect that is not part of the sum $\left(\phi_{\mathrm{g}} + \phi_{\mathrm{e}}\right)$, such as ionospheric Faraday rotation. Since the exact contributions of these effects are unknown, we describe the noise via a probability density function (PDF), $\mathcal{P}(n_i|\Theta)$, which gives the probability for the noise contribution to the i-th data point to take on a certain value[1], given a set of assumptions that we make about

[1] Precisely, the PDF only becomes a probability once it is integrated over an interval of possible values for the noise contribution.

the processes generating the noise or an effective parameterization of these, denoted as Θ. Additionally, the definition of the Galactic and extragalactic contributions also influence this PDF. Specifically, the demand that we want to describe the Galactic and extragalactic contributions to each source's Faraday depth by a single number each, when in reality the source might itself be spread out over a range in Faraday depth, may force the PDF for the remainder of the data, the noise, to take on a complicated shape. This PDF immediately provides us with the *likelihood*, i.e., the probability to measure a certain value, assuming certain values for the Galactic and extragalactic contributions (plus the set of assumptions Θ),

$$\mathcal{P}(d_i|\phi_{\mathrm{g},i},\phi_{\mathrm{e},i},\Theta) = \mathcal{P}(n_i = d_i - \phi_{\mathrm{g},i} - \phi_{\mathrm{e},i}|\Theta). \tag{6}$$

This equality holds since, for fixed Galactic and extragalactic contributions, the measurement is completely determined by its noise contribution.

Modeling all effects that can contribute to the noise is often not practical. Therefore, one usually finds a few effective parameters that approximately describe the PDF for the noise contribution or the likelihood. The most common choice, arising e.g. from the central limit theorem or a maximum entropy argument (see e.g. Jaynes & Bretthorst 2003), is a Gaussian PDF with zero mean, i.e.,

$$\begin{aligned}\mathcal{P}(d_i|\phi_{\mathrm{g},i},\phi_{\mathrm{e},i},\sigma_i) &= \mathcal{G}(d_i - \phi_{\mathrm{g},i} - \phi_{\mathrm{e},i},\sigma_i^2)\\ &= \left(2\pi\sigma_i^2\right)^{-1/2}\exp\left(-\frac{\left(d_i - \phi_{\mathrm{g},i} - \phi_{\mathrm{e},i}\right)^2}{2\sigma_i^2}\right).\end{aligned} \tag{7}$$

Here, we have introduced the notation $\mathcal{G}(x,X)$ for a one-dimensional Gaussian distribution for a variable x with zero mean and variance X and have parameterized the PDF completely with the parameter σ_i. We will refer to the parameter σ_i, the standard deviation of the likelihood, as an error bar, since it gives the width of the PDF for the noise (or error) contribution to the measured value. In these formulas we have assumed that while we have some idea about the expected magnitude of the noise contribution, given by the standard deviation σ_i, the specific value of n_i is unknown. Note that we are not considering systematic errors that would lead to an offset in the observational values and thus a non-zero mean of the PDF for the noise.

2.2. Imperfectly described likelihood functions

Precise knowledge of the likelihood function is crucial for the inference algorithms that we will discuss later. However, the Gaussian model for the likelihood, parameterized by the standard deviation σ_i, may not be an accurate description of the unknown effect influencing the observations in all cases. For cases that cannot be described with sufficient precision by just this one parameter, the outcome of trying to describe them in this way can be misleading. For example, the likelihood could have a sharp peak that can be approximated as a narrow Gaussian, but significant sidelobes that are neglected in this description.

One possible cause for such a non-trivial likelihood function could be the inherent n-π ambiguity of polarization orientation measurements. This is a problem for Faraday depth estimates from a linear λ^2-fit, but not for RM synthesis studies. We illustrate this in Fig. 1. Assuming polarization angle measurements at a few frequencies and no other information still allows in principle for infinitely many discrete possible values for the source's Faraday depth. Any uncertainty in the measurement of the angles

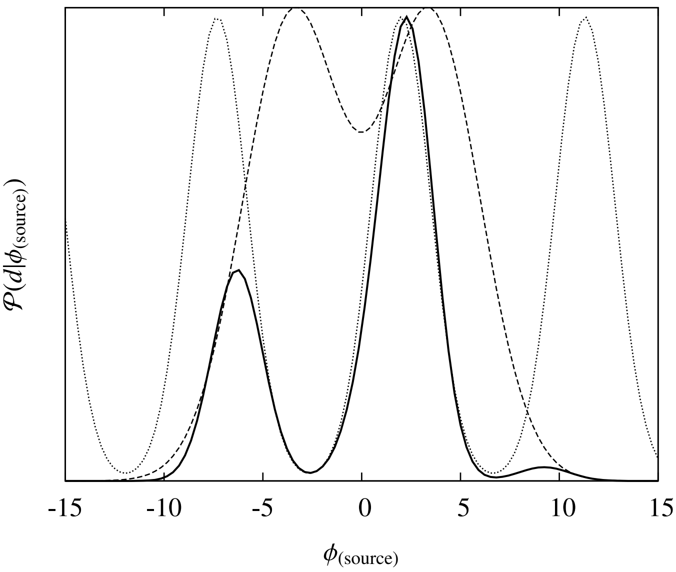

Fig. 1. Illustration of a possible likelihood curve for a single Faraday depth measurement. The dotted line shows the likelihood arising from the n-π ambiguity of position angle measurements and the uncertainty in these measurements. The dashed curve shows an additional constraint on the magnitude of the source's Faraday depth and the solid curve is the resulting combined likelihood curve, i.e., the product of the dotted and dashed curves. The plotted curves have been arbitrarily rescaled to show similar amplitudes.

(say, Gaussian) will turn these discrete possibilities into a series of equally likely peaks in the likelihood for the source's Faraday depth, as shown by the dotted line in Fig. 1. Additional data, like the degree of bandwidth depolarization (see e.g. Sunstrum et al. 2010) can lead to an additional constraint on the magnitude of the source's Faraday depth, but does not hold information on the sign of the Faraday depth. An example for such a constraint, formalized as a likelihood curve, is shown by the dashed line in Fig. 1. The combined likelihood, given by the product of the likelihood for the position angle measurements and the likelihood for the depolarization measurement, is shown by the solid curve. Evidently, the result can be highly non-Gaussian. The error bar that is quoted as a measure of the observational uncertainty could in this case correspond to the width of a single peak arising from the observational uncertainty in the measurement of the polarization angles. If this error bar is then interpreted as describing a Gaussian likelihood function, this likelihood function includes only the main peak and neglects any secondary peaks such as the ones visible in Fig. 1.

Usually, the error bar on an observational estimate of a source's Faraday depth is estimated as being inversely proportional to the signal to noise ratio of the polarized intensity observation, as well as the width of the frequency coverage in λ^2-space. This relation is based on linear Gaussian error propagation from the observations of the individual Stokes parameters to a polarization angle and to the slope of a straight line fit to the polarization angle as a function of λ^2, as shown by Brentjens & de Bruyn (2005) in their Appendix A. The result of this formula can be seen as an estimate for the width of the main peak shown in Fig. 1. However, it does not allow for the presence of n-π ambiguities and the ensuing non-Gaussianity shown by the solid line in Fig. 1. Furthermore, the Gaussian approximation to the observational uncertainty of a derived polarization angle is not perfect, as pointed out by Wardle & Kronberg (1974), and the estimation of the polarimetric noise in the first place can also be erroneous. These are effects that can lead to a general underestimation of the widths of the likelihood peaks, i.e., σ_i, as was detected by Stil et al. (2011) for the RM catalog of Taylor et al. (2009).

Furthermore, even though the sources that are used here are compact, it is not necessarily guaranteed that their emission as a function of Faraday depth is perfectly described by a single component. More complicated Faraday spectra due to, e.g., the internal structure of the source or differential Faraday rotation in the foreground within the telescope beam can lead to complicated effects on the observational estimates of Faraday depth, especially when a linear fit of polarization angle versus λ^2 is performed (see Farnsworth et al. 2011). Note that the usual formula used for the observational uncertainty described above implicitly assumes that the relationship between polarization angle and squared wavelength is linear. If polarized emission happens over an extended range of Faraday depths, the questions of what the Faraday depth of the source is and how large the intrinsic Faraday rotation of the source is become somewhat ill-defined. It is then not clear what the PDF for the noise should be if the noise is defined as the difference between the observed Faraday depth and two numbers characterizing the Galactic foreground contribution and the extragalactic contribution.

To accomodate the possibility of having such imperfect error information (among other things), Oppermann et al. (2012) allowed their algorithm to widen the likelihood functions of individual data points considerably. This was warranted because even when the uncertainty information for the vast majority of the data points is reliable, a few poorly described outliers can greatly influence the reconstruction. This was shown by Oppermann et al. (2011) for the method employed by Oppermann et al. (2012) and is still true for the algorithm we are deriving in Sect. 4, although to a lesser extent due to the spectral smoothness prior that we will employ.

2.3. Probability densities and covariance matrices

A generalization of the one-dimensional Gaussian model discussed so far is a multi-dimensional Gaussian model with correlations, described by

$$\begin{aligned}\mathcal{P}(n|N) &= \mathcal{G}(n,N)\\ &= \frac{1}{|2\pi N|^{1/2}}\exp\left(-\frac{1}{2}n^{\dagger}N^{-1}n\right).\end{aligned} \tag{8}$$

Here, n is to be read as a vector containing the noise contributions to all measurements and the dagger denotes a transposed quantity. The covariance matrix N contains the variances of the noise contributions to the individual measurements on its diagonal, $N_{ii} = \sigma_i^2$, and their correlations as off-diagonal entries. In the limit of uncorrelated noise contributions, i.e., $N_{ij} = 0$ for $i \neq j$, the one-dimensional Gaussian for each measurement is recovered.

So far, we have only discussed the likelihood. However, since we are trying to infer the Galactic and extragalactic contributions from the measurements and the likelihood is the PDF for the former quantity under the assumption of fixed values for the latter, it is clearly not the PDF that we are interested in. In order to turn the argument around, we make use of Bayes' theorem to construct the *posterior* PDF,

$$\mathcal{P}(\phi_{\mathrm{g}},\phi_{\mathrm{e}}|d,N) \propto \mathcal{P}(d|\phi_{\mathrm{g}},\phi_{\mathrm{e}},N)\,\mathcal{P}(\phi_{\mathrm{g}},\phi_{\mathrm{e}}), \tag{9}$$

which is the PDF for the Galactic and extragalactic contributions to all measurements, given the measured data and the assumptions about the noise covariance. The last PDF on the right hand side is the *prior* PDF for the Galactic and extragalactic contributions, i.e., a summary of knowledge we have about these constituents before taking into account the measurement data. For example, a prior could encode information about the expected variability or spatial smoothness of the Galactic and extragalactic contributions and thus serve to break the degeneracy between the two.

The posterior encodes all the knowledge that is available about the quantities of interest, both from the measurement data and from prior assumptions. Therefore, the posterior should be the main result of an analysis. However, it may be practical to summarize the information. To this end, we will approximate the posterior in our analysis as a Gaussian. This Gaussian is described by the *posterior mean*

$$\hat{\phi}_{\mathrm{g/e}} = \left\langle \phi_{\mathrm{g/e}} \right\rangle_{(\phi_{\mathrm{g/e}}|d)}, \tag{10}$$

a weighted average of all possible contributions, and the posterior covariance,

$$D_{\mathrm{g/e}} = \left\langle \left(\phi_{\mathrm{g/e}} - \hat{\phi}_{\mathrm{g/e}}\right)\left(\phi_{\mathrm{g/e}} - \hat{\phi}_{\mathrm{g/e}}\right)^{\dagger} \right\rangle_{(\phi_{\mathrm{g/e}}|d)}, \tag{11}$$

encoding some information on the uncertainty of the estimate. We use the angle bracket notation to denote integrals over all possible configurations of the quantity given in the index,

$$\langle f(x)\rangle_{(x|X)} = \int \mathcal{D}x\, f(x)\,\mathcal{P}(x|X). \tag{12}$$

As a prior, we will use a multi-dimensional Gaussian distribution with no linear correlation between any pair of the three constituents. Note that this does not exclude a correlation between, e.g., the noise *variance* and the Galactic contribution. For the extragalactic contribution, we choose a covariance matrix that is diagonal, i.e., the extragalactic contributions to the individual measurements are regarded as uncorrelated. In principle, a coherent magnetic field in the intergalactic medium within the Local Group or on cosmological scales could cause a correlated extragalactic Faraday rotation structure (e.g. Kolatt 1998; Blasi et al. 1999; Xu et al. 2006). However, Akahori & Ryu (2011) have shown that this contribution is expected to be about 3 to 4 orders of magnitude smaller than the contributions intrinsic to the sources and due to large scale structure filaments.

For the Galactic contribution, on the other hand, we allow for spatial correlations[2]. In this sense, we statistically define any correlated features to be "Galactic" and any uncorrelated features to be non-Galactic. It will therefore be possible for extended extragalactic features, e.g. from satellites of the Milky Way, to appear in our estimate of the Galactic contribution and for uncorrelated features from within the Milky Way to appear in our estimate of the extragalactic contribution or in the noise budget. For the estimate of the extragalactic contributions this means that a correlated offset is possible. This should, however, affect all sources in a similar way, unless sources behind a structure that causes this correlated extragalactic contribution are compared to sources without such a structure in front of them.

[2] To be precise, we assign a Gaussian prior with position independent correlation function to a dimensionless quantity that is linearly related to the Galactic contribution. See Sect. 4 or Oppermann et al. (2012) for details.

The fact that we assume no correlations both for the noise and extragalactic contributions further strengthens the degeneracy of the decomposition of the data. While the estimate of the Galactic contribution to a data point will be influenced by surrounding data points as well, due to the assumed spatial correlations, the remainder will simply be split between the estimates for the noise and extragalactic contribution according to the ratio of the expected variances for this particular data point. Furthermore, we will regard both expected variances to be uncertain. Without further constraints it is then impossible to decide what ratio should be chosen for the split. While we will also assume the correlation structure of the Galactic contribution to be uncertain, this can be extracted from the data unambiguously, except for a possible part with no spatial correlations, which would suffer from the same degeneracy with the noise and extragalactic contributions. We will discuss this problem in more detail in the next section.

3. Estimators

As we already pointed out, an exact and definitive separation of d into its three constituents is not possible. However, under reasonable assumptions, we can formulate estimates for these quantities and draw conclusions about the statistics of the extragalactic contributions. In this section, we will discuss methods to estimate the Galactic and extragalactic contributions without any specific prior assumptions. In the next section, we will present a full Bayesian analysis with the goal of extracting the extragalactic contributions.

3.1. Wiener filter

To estimate the Galactic contribution to Faraday rotation, Oppermann et al. (2012) used an algorithm based on the Wiener filter, which calculates the linear estimate, $\hat{\phi}_{\mathrm{g}} = Fd$, that minimizes the expectation value of the square-norm of the residual $r = \phi_{\mathrm{g}} - \hat{\phi}_{\mathrm{g}}$. Here, the expectation value is calculated over the joint PDF of all constituents of the data, so that F minimizes the expression

$$\left\langle \left(\phi_{\mathrm{g}} - Fd\right)^{\dagger} \left(\phi_{\mathrm{g}} - Fd\right) \right\rangle_{(\phi_{\mathrm{g}},\phi_{\mathrm{e}},n)}. \tag{13}$$

The Wiener filter reconstruction Fd indeed yields the optimal estimate in this square-norm sense in cases in which all priors are Gaussian, since it corresponds to the posterior mean of ϕ_{e} in this case (see e.g. Enßlin & Weig 2010). In all other cases, the Wiener filter is still the optimal *linear* filter.

Solving for the filter matrix F yields (see e.g. Zaroubi et al. 1995)

$$F = \left\langle \phi_{\mathrm{g}} d^{\dagger} \right\rangle_{(\phi_{\mathrm{g}},\phi_{\mathrm{e}},n)} \left\langle d d^{\dagger} \right\rangle^{-1}_{(\phi_{\mathrm{g}},\phi_{\mathrm{e}},n)}. \tag{14}$$

We use $\dagger$ to denote transposed vectors, so that the expression in Eq. (13) is a scalar and the expression in Eq. (14) consists of two matrices. Assuming all three constituents of the data to be mutually linearly uncorrelated, the expectation values in the last equation simplify to

$$\left\langle \phi_{\mathrm{g}} d^{\dagger} \right\rangle_{(\phi_{\mathrm{g}},\phi_{\mathrm{e}},n)} = \left\langle \phi_{\mathrm{g}} \phi_{\mathrm{g}}^{\dagger} \right\rangle_{(\phi_{\mathrm{g}})} \equiv G \tag{15}$$

and

$$\left\langle d d^{\dagger} \right\rangle_{(\phi_{\mathrm{g}},\phi_{\mathrm{e}},n)} = \left\langle \phi_{\mathrm{g}} \phi_{\mathrm{g}}^{\dagger} \right\rangle_{(\phi_{\mathrm{g}})} + \left\langle \phi_{\mathrm{e}} \phi_{\mathrm{e}}^{\dagger} \right\rangle_{(\phi_{\mathrm{e}})} + \left\langle n n^{\dagger} \right\rangle_{(n)} \equiv G+E+N, \tag{16}$$

where we have introduced the covariance matrices G, E, and N of the Galactic, extragalactic, and noise contributions, respectively. With these abbreviations, the Wiener filter estimate of the Galactic contribution becomes

$$\hat{\phi}_{\mathrm{g}} = G\,(G + E + N)^{-1}\, d. \tag{17}$$

If, on the other hand, the goal is to estimate the extragalactic contribution ϕ_{e}, a straightforward idea is to take the data d and subtract the optimal estimate of the Galactic contribution. However, this is *not* the optimal estimate of the extragalactic contribution. Rather, it is the optimal estimate of the difference between the data and the Galactic contribution, i.e., of the sum of the extragalactic contribution and the noise,

$$d - \hat{\phi}_{\mathrm{g}} = (E + N)\,(G + E + N)^{-1}\, d = \hat{\phi}_{\mathrm{e}} + \hat{n}. \tag{18}$$

For the optimal estimate of the extragalactic contributions, one simply has to exchange the roles of ϕ_{g} and ϕ_{e} in the Wiener filter, leading to

$$\hat{\phi}_{\mathrm{e}} = E\,(G + E + N)^{-1}\, d. \tag{19}$$

The difference with respect to Eq. (18) is a weighting with the ratio of the expected extragalactic variance to the expected non-Galactic variance, i.e., the sum of the extragalactic and noise variances,

$$\hat{\phi}_{\mathrm{e}} = E\,(E + N)^{-1} \left(d - \hat{\phi}_{\mathrm{g}}\right). \tag{20}$$

Similarly, one could build an optimal estimator of the measurement noise by exchanging the roles of ϕ_{e} and n. This would lead to the estimate

$$\hat{n} = N\,(E + N)^{-1} \left(d - \hat{\phi}_{\mathrm{g}}\right). \tag{21}$$

The main problem with all these considerations is that the matrices G, E, and N are not necessarily known to sufficient precision. Oppermann et al. (2012) have overcome this problem for their estimate of the Galactic contributions by estimating the matrices from the same data set. For this, they summarized the non-Galactic contributions E and N, i.e., they did not attempt to differentiate between extragalactic and noise contributions. Furthermore, they made some simplifying assumptions about the structure of the matrices, namely diagonality in the spherical harmonics basis for a transformed version[3] of G and diagonality of $(E + N)$. Since the information contained in the data points is local to the sources' positions on the sky, diagonality of $(E + N)$ corresponds to a lack of spatial correlations. This difference in correlation structure allows the reconstruction of G and of the sum of N and E. For the latter, the diagonal of $(E+N)$ was assumed to consist of the sum of the observational uncertainty variance, σ^2, and an estimate of the typical variance of the extragalactic contributions, $\sigma_{\mathrm{e}}^2 = \left(6.6\,\mathrm{rad/m^2}\right)^2$, taken from Schnitzeler (2010) and consistent with the earlier study of Leahy (1987), multiplied with a correction factor η,

$$\mathrm{diag}\,(E + N) = \eta \left(\sigma_{\mathrm{e}}^2 + \sigma^2\right). \tag{22}$$

The correction factor was introduced to account for error bars that describe the likelihood with insufficient precision, the potential of extra multiples of π in the rotation of the polarization

[3] Oppermann et al. (2012) write $G = RSR^{\dagger}$, where R is an operator that projects an all-sky field onto the source-positions and multiplies with a Galactic latitude-dependent profile function, and assume diagonality in the spherical harmonics basis for S. See also Sect. 4, where we follow the same argument.

plane, and sources for which the extragalactic contribution is significantly larger. Note that η and σ are different for each data point, while σ_e is the same for all. With these assumptions, the Galactic contribution can be reconstructed according to Eq. (17). This involves the determination of the correction factors η and the matrix G from the same data set. We will recapitulate the underlying assumptions and mathematics in Sect. 4, where we present a similar, but refined, algorithm to reconstruct both the Galactic and extragalactic contributions.

For an optimal estimate of the extragalactic contribution according to Eq. (20) the extragalactic covariance matrix E needs to be known individually, not just the sum $(E + N)$. A separation of E and N using the same recipe used by Oppermann et al. (2012) for the separation of G and $(E+N)$ is, however, not feasible, since E and N are degenerate under the assumption that both of them are diagonal. In other words, it is impossible to decide on statistical grounds for individual data points, how much of the correction factor η is due to extragalactic contributions and how much of it is due to observational effects. In this section, we present a simple analysis making use only of the available results of Oppermann et al. (2012) and a crude assumption about the split between E and N to demonstrate some of the statistical features to be expected from the result. In Sect. 4, we will make use of further physically and observationally motivated assumptions to partially overcome the degeneracy within a full analysis that includes N and E individually.

Oppermann et al. (2012) found that the correction factors η tend to be larger at lower absolute Galactic latitude (see their Fig. 7), attributing this effect at least partly to multiples of π in the rotation angle measurements that were wrongly determined. Another reason could be that the Galactic plane contains much more structure in Galactic Faraday depth on the scale of the typical source separation than the polar regions. As a consequence, the algorithm would wrongly attribute this structure to noise. In any case, on physical grounds the observed latitude dependence is not expected to be due to extragalactic contributions. A reasonable first approximation could therefore be to assign the correction factors of Eq. (22) solely to observational effects and write

$$\mathrm{diag}\,(E) = \sigma_e^2 \tag{23}$$

and

$$\mathrm{diag}\,(N) = \eta\sigma^2 + (\eta - 1)\,\sigma_e^2. \tag{24}$$

Since many of the correction factors derived by Oppermann et al. (2012) are actually smaller than one, the last equation can lead to reduced or even negative error variances. To prevent this from happening, we replace all correction factors that are smaller than one with one here, recalculate the map of the Galactic contribution using these correction factors, and use these estimates in the following when we write η and $\hat{\phi}_g$.

With these assumptions and approximations, the posterior mean estimate for the extragalactic contributions is straightforwardly calculated according to Eq. (20). This estimate is shown in the top left panel of Fig. 2 as a function of Galactic latitude for the data from the NVSS RM catalog (Taylor et al. 2009). Here we use only the NVSS RM catalog because of its large number of sources and in order to be able to point out some statistical features in the result without being influenced by a combination of different catalogs with varying properties. In the full analysis presented in Sect. 4, we will use all observational catalogs available to us.

3.2. Features in the extragalactic estimates

The most striking feature of the extragalactic Faraday rotation contributions calculated under these crude assumptions is the dependence of their distribution on Galactic latitude. As a measure of the spread of extragalactic contributions, we calculate the sample standard deviation in bins of $\sin(b)$ of width 0.2, where b denotes Galactic latitude. The resulting standard deviations are shown in Fig. 3. This plot seems to suggest that there is something unexpected happening, potentially brought about by the assumption made in Eqs. (23) and (24). This is further confirmed by histograms of the distribution in latitude bins, plotted in the top right panel of Fig. 2. Evidently, the shape of the distributions is less peaked nearer to the Galactic poles than at intermediate latitudes. The distribution becomes very flat near its center in bins close to the Galactic poles.

It may be surprising that these distributions depend on Galactic latitude so strongly, given that they are supposed to show the extragalactic contributions. Before discussing the reasons for this behavior, we will make a few general comments regarding its implications.

It is important to note that the quantities plotted in the top half of Fig. 2 are *not* the extragalactic contributions, but only *estimates* of these. As we pointed out earlier, given the degeneracy of the problem and the uncertainties involved, a definitive determination of the extragalactic contributions is not possible. Any separation of the observed Faraday rotation measurements into Galactic, extragalactic, and noise contributions is a trade-off between the three and therefore each of the three estimates affects the other two. Any result will have to be probabilistic in nature, no matter how sophisticated the analysis method or how good the data set. The statistical degeneracy is, however, less severe if a data set is considered for which the noise covariance is known with certainty.

To illustrate how the split of the data into the three components is done in this case, we investigate the case of only two Faraday rotation measurements at two different locations on the sky. In this case the data vector is two-dimensional, $d = (d_{1,}\, d_2)^\dagger$, and can easily be plotted as a vector, see Fig. 4. We assume the data according to Eq. 4 to be the result of three independent stochastic processes, namely Galactic and extragalactic Faraday rotation at these locations plus measurement noise. We model these to be Gaussian processes with covariances

$$G = \begin{pmatrix} (10)^2 & 60 \\ 60 & (10)^2 \end{pmatrix} \mathrm{rad}^2/\mathrm{m}^4, \tag{25}$$

$$E = \begin{pmatrix} (6.6)^2 & 0 \\ 0 & (6.6)^2 \end{pmatrix} \mathrm{rad}^2/\mathrm{m}^4, \text{ and} \tag{26}$$

$$N = \begin{pmatrix} (6.0)^2 & 0 \\ 0 & (2.0)^2 \end{pmatrix} \mathrm{rad}^2/\mathrm{m}^4. \tag{27}$$

The noise of the two measurements is independent and is assumed in our example to have a standard deviation of $6\,\mathrm{rad/m}^2$ and $2\,\mathrm{rad/m}^2$ for the two different observations, respectively. The extragalactic contributions have the same variance everywhere, with a standard deviation of $6.6\,\mathrm{rad/m}^2$. The Galactic components on the two measurement locations are assumed to be correlated, leading to a non-diagonal correlation matrix G. In Fig. 4 we show the component vectors drawn to generate the data as well as their reconstruction from the data according to Eqs. (17), (20), and (21) under the assumption of known covariance matrices. These estimates add up to the data vector, $\hat{\phi}_g + \hat{\phi}_e + \hat{n} = d = \phi_g + \phi_e + n$, without being identical to the original signals. Due to the different structure of the covariance

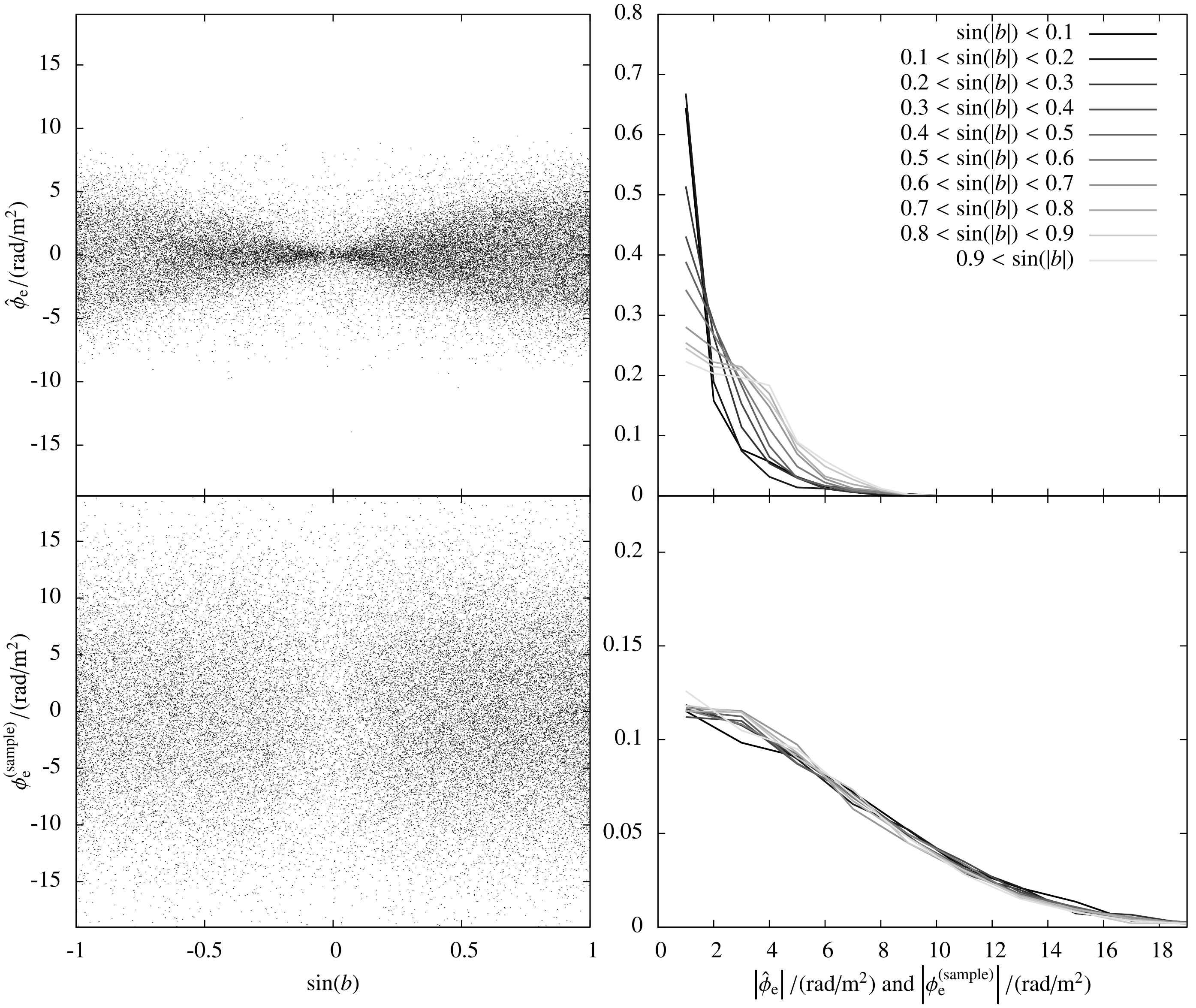

Fig. 2. *Top*: Estimates for the extragalactic contributions to the Faraday depths of the 37 543 sources in the NVSS RM catalog (Taylor et al. 2009), calculated according to Eqs. (20), (23), and (24). The left panel shows the estimate for each of the data points, plotted against Galactic latitude. The right panel shows normalized histograms of the distribution of these estimates in bins of $\sin(|b|)$. Each bin has a width of 0.1. *Bottom*: The same for a random sample drawn from the posterior PDF, Eq. (29), for the extragalactic contributions around the mean plotted in the upper panels.

matrices it is possible to construct an optimal and non-trivial partition of the data vector in non-parallel components.

This example shows that the reconstruction is able to capture some aspects of the components, but not all. In particular it illustrates that the reconstructed quantities have less variance than the original ones. Note also that the estimates $\hat{\phi}_e$ and $\hat{n}$ simply split the difference $d - \hat{\phi}_g$ according to the ratio of the variances, $(6.6)^2/(6.0)^2$ and $(6.6)^2/(2.0)^2$ for each of the two data points, respectively. This is due to both E and N being diagonal. Furthermore, Fig. 4 also illustrates that using $d - \hat{\phi}_g$ as an estimator for ϕ_e is suboptimal, since it contains some of the noise, as we have $d - \hat{\phi}_g = \hat{\phi}_e + \hat{n}$.

The amount of variance missing from the estimates is known statistically and can be characterized by covariance matrices. These are $D_g = \left(G^{-1} + (N+E)^{-1}\right)^{-1}$, $D_e = \left(E^{-1} + (N+G)^{-1}\right)^{-1}$, and $D_n = \left(N^{-1} + (G+E)^{-1}\right)^{-1}$ for the Galactic, extragalactic, and noise components, respectively.

Since the whole analysis presented here and by Oppermann et al. (2012) relies on the description of the statistics in terms of second moments or covariance matrices, it does not make any prediction regarding the shape of the resulting sample distributions. Therefore, even if the reasons for the change in shape with Galactic latitude are not immediately apparent, there is also no reason to expect that the shape of the distribution of the *estimate* $\hat{\phi}_e$ should not change with Galactic latitude.

Under the assumption of Gaussian priors for all three constituents of the data, i.e.,

$$\mathcal{P}(\phi_g, \phi_e, n|G, E, N) = \mathcal{G}(\phi_g, G)\,\mathcal{G}(\phi_e, E)\,\mathcal{G}(n, N), \tag{28}$$

the posterior for each of them individually is again Gaussian. Specifically, the posterior for the extragalactic contributions is given by

$$\mathcal{P}(\phi_e|d, G, E, N) = \mathcal{G}\Big(\phi_e - \hat{\phi}_e, D_e = \left(E^{-1} + (G+N)^{-1}\right)^{-1}\Big), \tag{29}$$

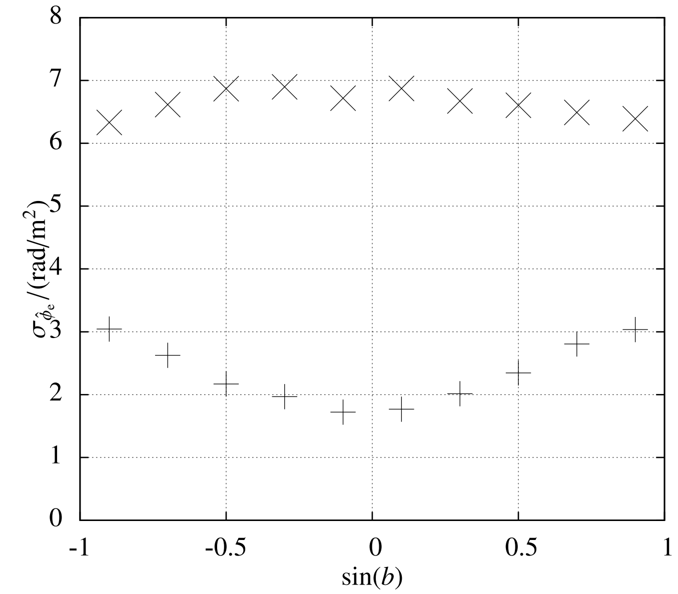

Fig. 3. Sample standard deviation of the estimated extragalactic contributions plotted in the top left panel of Fig. 2 (+-symbols), as well as of the random sample drawn from the posterior PDF for the same quantity, plotted in the bottom left panel of Fig. 2 (×-symbols), in bins of $\sin(b)$. Each bin has a width of 0.2.

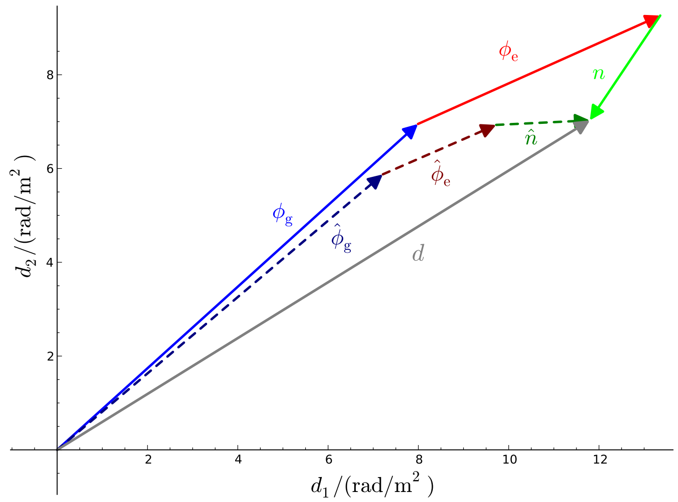

Fig. 4. Visualization of the data space in the case of only two Faraday rotation measurements. Each coordinate denotes the possible values for the two measurements d_1 and d_2 at two sky positions. The data vector, $d = (d_{1,} d_2)^\dagger$ is the sum of a vector of Galactic Faraday depth, $\phi_g = (\phi_{g,1}, \phi_{g,2})^\dagger$, of a vector of extragalactic Faraday depth, $\phi_e = (\phi_{e,1}, \phi_{e,2})^\dagger$, and of a measurement noise vector, $n = (n_1, n_2)^\dagger$. The reconstructions of these three components, $\hat{\phi}_g$, $\hat{\phi}_e$, and $\hat{n}$, are shown as dashed arrows. Their sum is also equal to the data, however they differ from the correct components due to the impossibility to uniquely separate one data vector into three statistically independent components.

with the Wiener filter estimate as mean and a covariance D_e describing the variance that is missing from the estimate itself (see e.g. Enßlin et al. 2009). Near the Galactic plane, the variance of the Galactic Faraday depth is greatly enhanced. Therefore, the entries of G corresponding to lines of sight at low absolute latitudes are comparatively large. The same is true to a different extent for the entries of the noise covariance N. This is almost entirely due to our usage of the noise variance correction factors, which tended to be larger for data points near the Galactic plane in the study of Oppermann et al. (2012). The two most likely reasons for this behavior of the correction factors are the presence of previously faulty multiples of π in rotation angle measurements and the assignment of small-scale structure in the Galactic contribution to the noise budget by Oppermann et al. (2012). This could happen if the amount of structure on small scales in the Galactic contribution is much larger near the Galactic plane than elsewhere, since an assumption of isotropic statistics modulo an overall scaling of the values with Galactic latitude was made for the Galactic contribution.

The higher variance of the Galactic and noise contributions near the Galactic plane is therefore the cause for the smaller absolute values of the estimates plotted in the top left panel of Fig. 2 in this region. However, the covariance matrix describing the posterior Gaussian for the extragalactic contributions consequently also encodes a higher uncertainty of the estimate near the Galactic plane. So while the estimated extragalactic contributions tend to be smaller in modulus near the Galactic plane, the uncertainty of the estimate is higher. This effect can be seen by drawing random realizations from the posterior PDF. Each realization represents one possible configuration that is not ruled out by the data. By drawing the realizations from the posterior, one more often draws configurations that are well supported by the data than the ones that are marginally possible. While it is true that the most probable configuration is given by the posterior maximum, $\hat{\phi}_e$, this does not need to be a typical configuration in any sense. For example, while the configuration plotted in the top left panel of Fig. 2 with its small absolute values near the Galactic plane is more probable than any other specific realization, there are many more realizations with typically higher absolute values near the plane that are also not ruled out by the posterior, Eq. (29). Due to the higher posterior uncertainty near the Galactic plane, different realizations drawn from the posterior distribution will vary wildly for data points at low absolute Galactic latitude.

One random sample realization drawn from the posterior PDF is shown in the bottom left panel of Fig. 2. Evidently, the strong dependence on Galactic latitude has been reduced. This is further confirmed by the lower right panel of Fig. 2 and by the ×-symbols in Fig. 3, which show less variation with latitude than the +-symbols. This confirms that our analysis is based on plausible, albeit somewhat crude, assumptions, since the apparent artifact of the distribution in the top panels of Fig. 2 disappears almost completely when the posterior uncertainty is added.

4. Filtering Galactic and extragalactic contributions simultaneously

In this section we will extend the analysis of Sect. 3 to a reconstruction of the Galactic and extragalactic contributions at the same time. For this, we will introduce the prior variance of the extragalactic contribution as a free parameter right from the start. In the first subsection, we lay out the necessary assumptions. We then present the derivation of the algorithm from probabilistic considerations, demonstrate its performance in a simulated scenario, and finally show the results for the analysis of real-world data in Sect. 5.

4.1. Assumptions and covariance matrices

In the previous section we have worked with linear estimates based on the usage of covariance matrices. Here, we will go further and explicitly assume Gaussian priors for all three constituents as we already did in Eq. (28), i.e., we make use of only the covariance matrices to describe the prior statistics (see e.g. Jaynes & Bretthorst 2003, for a discourse on the appropriateness of Gaussian priors in such a case).

4.1.1. Galactic covariance

For the prior covariance of the Galactic contribution, we follow the argument of Oppermann et al. (2012). We model the Galactic Faraday depth as an isotropic Gaussian random field multiplied with a latitude-dependent profile function. This profile function serves to remove the largest-scale, most obvious, anisotropy, namely the presence of the Galactic disk, from the definition of the random field. This enables us to treat the random field, which we will try to reconstruct, approximatively as an isotropic Gaussian field. In short, we write

$$\phi_{\mathrm{g}}(l,b) = p(b)\, s(l,b), \tag{30}$$

where $p(b)$ is the Galactic profile function, $s(l,b)$ is the dimensionless isotropic random field, and l and b are Galactic longitude and latitude, respectively.

Modeling the field s as an isotropic field means that its two-point correlation function (or covariance) depends only on the distance of the two points. Equivalently, one can write the covariance matrix as diagonal in the basis of spherical harmonics components according to

$$S_{(\ell m),(\ell' m')} = \left\langle s_{\ell m} s^{*}_{\ell' m'} \right\rangle_{(s|S)} = \delta_{\ell\ell'}\, \delta_{mm'}\, C_\ell. \tag{31}$$

Here, δ is the Kronecker delta symbol, the asterisk denotes complex conjugation, and C_ℓ denotes the angular power spectrum on scale $\Delta_\theta \approx 180°/\ell$.

The field $s(l,b)$ is defined on the complete celestial sphere. However, observational estimates of Faraday depth are only available in the directions in which polarized sources have been studied at several frequencies. We make use only of extragalactic sources in order to avoid the inclusion of distance information in our analysis. In order to formalize the relationship between the all-sky field s and the data, we need to include both a projection of the field onto the direction in which the observations were made and a multiplication with the Galactic profile function, as well as the addition of the extragalactic contribution and noise. In total, the i-th data point can be written as

$$\begin{aligned} d_i &= p(b_i)\, s(l_i,b_i) + \phi_{\mathrm{e},i} + n_i \\ &= \int_{\mathcal{S}^2} \mathrm{d}\Omega\, R_i(\hat{r})\, s(\hat{r}) + \phi_{\mathrm{e},i} + n_i. \end{aligned} \tag{32}$$

We use (l,b) and $\hat{r}$ interchangably to denote positions on the celestial sphere with (l_i,b_i) being the position of the i-th data point. In the last line we have summarized the multiplication with the profile function and projection onto the appropriate direction in a response operator R. With this we can write the covariance matrix of the Galactic contribution, regarded as a quantity defined only in the locations of observed sources, as

$$G = RSR^{\dagger}. \tag{33}$$

The only thing left undetermined is the angular power spectrum. Inferring it from the data will be part of our reconstruction algorithm.

We constrain the angular power spectrum with a prior consisting of two parts, following Oppermann et al. (2013). The first part is an independent inverse-Gamma prior for each component,

$$\mathcal{P}_{\mathrm{IG}}(C_\ell) \propto C_\ell^{-\alpha_\ell} \exp\left(-\frac{q_\ell}{C_\ell}\right), \tag{34}$$

where q_ℓ and α_ℓ are parameters that vary the constraining power of the prior. In our application, we will use the limit of $q_\ell \to 0$ and $\alpha_\ell \to 1$, turning the inverse-Gamma prior into a Jeffreys prior, a probability density that is flat on a logarithmic scale. This choice is made to get a prior that is very broad and will not confine the reconstruction to a too narrow range. The second part of our power spectrum prior is a term that couples different scales and enforces spectral smoothness. This term is given by

$$\begin{aligned} \mathcal{P}_{\mathrm{sm}}(C) &\propto \exp\left(-\frac{1}{2\sigma_{\mathrm{sm}}^2} \int \mathrm{d}(\log \ell) \left(\frac{\partial^2 \log C_\ell}{\partial(\log \ell)^2}\right)^2\right) \\ &= \exp\left(-\frac{1}{2}(\log C)^{\dagger} T (\log C)\right), \end{aligned} \tag{35}$$

where the second derivative and integral are to be read as short-hands for finite-difference expressions and T is a matrix that performs the same operations. This prior favors angular power spectra that are close to power laws. Its strength is regulated by the parameter σ_{sm}. In total, the prior for the angular power spectrum is the product of the two terms,

$$\mathcal{P}(C) \propto \mathcal{P}_{\mathrm{sm}}(C) \prod_\ell \mathcal{P}_{\mathrm{IG}}(C_\ell). \tag{36}$$

We refer the reader to Oppermann et al. (2013) for a detailed discussion of this spectral prior.

4.1.2. Extragalactic and noise covariances

The extragalactic and noise contributions can both be regarded as quantities that are defined only for the source positions; hence the index i in Eq. (32). For the noise this choice is obvious since there is no noise if there is no measurement. For the extragalactic contribution the data have no constraining power at any other locations with the prior assumptions that we are about to make and we therefore only consider the extragalactic contributions at the source positions for the sake of simplicity.

We consider the covariance matrices E and N both to be diagonal, i.e., we do not allow for correlations between extragalactic or noise contributions of different data points. As mentioned before, this assumption is only strictly valid as long as there are no coherent components to the extragalactic contributions, e.g. from an intervening galaxy cluster or a coherent magnetic field in the intergalactic medium.

Since these two matrices are diagonal in the same basis, their entries are degenerate if both matrices are unknown, as we have discussed in Sect. 3.1. To break this degeneracy, we will make the assumption here that some of the entries of the noise covariance matrix are in fact known with certainty a priori, which will enable us to extract the variance of the extragalactic contributions from the data points corresponding to these entries. For this, we split the data set into two categories. The first category contains all data points for which we cannot be certain that the quoted error bar describes the likelihood function accurately enough. We will refer to these data as *Somewhat Informative Points* (SIP). For these data, we model the sum of the two covariance matrices as

$$(E+N)_{ii} = \eta_i \left(\sigma_{\mathrm{e}}^2 + \sigma_i^2\right), \tag{37}$$

i.e., we do not attempt to differentiate between insufficient error information and an under-estimated variance of the extragalactic contribution. For the other category of data, for which we assume that the noise statistics are sufficiently described by the Gaussian error bars, we model the covariance matrices as

$$N_{jj} = \sigma_j^2 \tag{38}$$

and

$$E_{jj} = \eta_e \sigma_e^2, \tag{39}$$

i.e., we do not allow for corrections of the error bars but use these data points to correct our initial assumption about the typical variance of the extragalactic contributions. We will refer to these data as *Very Informative Points* (VIP). Note that the error variance correction factors in Eq. (37) are individually determined for each data point, whereas the factor η_e in Eq. (39) is the same for each data point in the VIP category, since we assume that the statistics of the extragalactic contributions are homogeneous over the whole sky.

In the following, we denote as $d^{(\mathrm{SIP})}$, $\phi_e^{(\mathrm{SIP})}$, $n^{(\mathrm{SIP})}$ the vectors containing the observed values, extragalactic contributions, and noise contributions of the data points in the SIP category, described by Eq. (37), and as $d^{(\mathrm{VIP})}$, $\phi_e^{(\mathrm{VIP})}$, $n^{(\mathrm{VIP})}$ the same for the VIP category of data, described by Eqs. (38) and (39). Deciding in which category a specific data point should be is of course a non-trivial task. In Sect. 5 we will present several possible choices for such splits of the data.

For the free parameters η_i and η_e, we choose again inverse-Gamma priors,

$$\mathcal{P}_{\mathrm{IG}}(\eta_{i/e}) \propto \eta_{i/e}^{-\beta_{i/e}} \exp\left(-\frac{r_{i/e}}{\eta_{i/e}}\right). \tag{40}$$

We will discuss our choice of parameters $r_{i/e}$ and $\beta_{i/e}$ in the following section. Naturally, we do not enforce smoothness of the η-values in any sense.

4.2. Reconstruction strategy

Here we make use of the combined methodology of Oppermann et al. (2011) and Oppermann et al. (2013). In the first of these papers a method for the reconstruction of a Gaussian signal field, its power spectrum, and the η-factors was presented, the so-called *extended critical filter*. In the second paper, the inclusion of a spectral smoothness prior according to Eq. (35) was discussed, albeit without allowing for the η-factors. Combining these two methods is straightforward, using the ansatz of the empirical Bayes method (e.g. Robbins 1955). We approximate the posterior mean for the dimensionless signal field and the extragalactic contributions as

$$\begin{aligned} m &= \int \mathcal{D}s \int \mathcal{D}\tilde{C} \int \mathcal{D}\tilde{\eta}\; s\,\mathcal{P}(s|d,\tilde{C},\tilde{\eta})\,\mathcal{P}(\tilde{C},\tilde{\eta}|d) \\ &\approx \int \mathcal{D}s \int \mathcal{D}\tilde{C} \int \mathcal{D}\tilde{\eta}\; s\,\mathcal{P}(s|d,\tilde{C},\tilde{\eta})\,\delta(\tilde{C}-\hat{\tilde{C}})\,\delta(\tilde{\eta}-\hat{\tilde{\eta}}), \end{aligned} \tag{41}$$

$$\begin{aligned} \hat{\phi}_e &= \int \mathcal{D}\phi_e \int \mathcal{D}\tilde{C} \int \mathcal{D}\tilde{\eta}\; \phi_e\,\mathcal{P}(\phi_e|d,\tilde{C},\tilde{\eta})\,\mathcal{P}(\tilde{C},\tilde{\eta}|d) \\ &\approx \int \mathcal{D}\phi_e \int \mathcal{D}\tilde{C} \int \mathcal{D}\tilde{\eta}\; \phi_e\,\mathcal{P}(\phi_e|d,\tilde{C},\tilde{\eta})\,\delta(\tilde{C}-\hat{\tilde{C}})\,\delta(\tilde{\eta}-\hat{\tilde{\eta}}), \end{aligned} \tag{42}$$

where we have chosen to work with the logarithmic quantities

$$\tilde{C} = \log C \tag{43}$$

and

$$\tilde{\eta} = \log \eta. \tag{44}$$

We choose as estimators for the auxiliary quantities, $(\hat{\tilde{C}}_\ell)_\ell$, $(\hat{\tilde{\eta}}_i)_i$, and $\hat{\tilde{\eta}}_e$ the numbers that maximize the PDFs

$$\mathcal{P}\Big((\tilde{C}_\ell)_\ell \Big| d, (\tilde{\eta}_i)_i = (\hat{\tilde{\eta}}_i)_i, \eta_e = \hat{\eta}_e\Big), \tag{45}$$

$$\mathcal{P}\Big((\tilde{\eta}_i)_i | d, (\tilde{C}_\ell)_\ell = (\hat{\tilde{C}}_\ell)_\ell, \eta_e = \hat{\eta}_e\Big), \tag{46}$$

and

$$\mathcal{P}\Big(\tilde{\eta}_e \Big| d, (\tilde{C}_\ell)_\ell = (\hat{\tilde{C}}_\ell)_\ell\Big), \tag{47}$$

respectively. This leads to estimates m for the signal field and $\hat{\phi}_e$ for the extragalactic contributions that depend on the estimates $\hat{\tilde{C}}$ and $\hat{\tilde{\eta}}$, which are themselves again dependent on each other and on the estimates m and $\hat{\phi}_e$. A system of equations arises that needs to be solved self-consistently. Our estimate for the Galactic contributions to the observed values will then be $\hat{\phi}_g = Rm$.

Putting in the priors that we described in the previous section, we can calculate the PDFs needed for the estimates, after making a few approximations. The detailed calculations and filter formulas are discussed in Appendix A.

The choice of estimators made here represents a trade-off between statistical optimality and practical ability to calculate the relevant PDFs. Ideally, we would estimate each quantity by marginalizing over all other unknown quantities and averaging the resulting posterior distribution. However, these marginalizations are in general not possible analytically. Rather than resorting to computationally expensive sampling techniques, we use some unmarginalized PDFs. In the next section, we show a simulated example calculation that demonstrates the quality of the results obtained with our approximations.

In all this, we have assumed the Galactic latitude profile to be known. Another global iteration step will be needed to include this as a quantity to be reconstructed. In a first step, we calculate the profile function simply as the root mean square of all the data values in latitude bins. In doing this, we subtract the noise variance and smooth the squares with a Gaussian kernel of 4° full width at half maximum (FWHM). After the iteration of the filter equations derived in the appendix has converged using this profile function, we calculate an approximative map of the posterior mean for the squared Galactic Faraday depth according to

$$\left\langle \phi_g^2 \right\rangle_{(\phi_g|d)} \approx \left\langle (ps)^2 \right\rangle_{\mathcal{G}(s-m,D)} = p^2 m^2 + p^2 \mathrm{diag}(D), \tag{48}$$

where p is the profile function used in the iteration of the equations listed in the appendix. We then smooth this map again with a Gaussian kernel of 4° FWHM and average over Galactic latitude bins to obtain a new profile function. This procedure is repeated until the profile function has converged as well. As we will see in Sect. 4.3, a few of these global iteration steps suffice to achieve convergence.

We fix the remaining prior parameters according to the following scheme. For the reconstruction of the angular power spectrum of the Galactic contributions, we use the limit $q \to 0$ and $\alpha \to 1$. This turns the inverse-gamma prior for each parameter C_ℓ into a Jeffreys prior, and makes the prior for $\log C_\ell$ flat. For the strength of the spectral smoothness prior, we choose $\sigma_{\mathrm{sm}}^2 = 10$, entering the equations via the matrix T. This is a rather weak smoothness prior, allowing for a change in slope of $\sqrt{10}$ per e-folding in ℓ on a 1σ-level. The reconstruction of the power spectrum will therefore be largely data-driven. For the prior for the noise variance correction factors we choose $\beta_i = 2$, making this inverse-gamma prior more informative than the one used for the angular power spectrum. This is done to account for the expectation that most of the data points in the SIP category will have error bars that describe the likelihood sufficiently well and therefore will not need large η_i-factors. For the cutoff-parameters r_i, we adopt the values

$$r_i = \frac{3}{2} \max \left\{ 1, \frac{\sigma_i^2 + \eta_e \sigma_e^2}{\sigma_i^2 + \sigma_e^2} \right\}. \tag{49}$$

The lower threshold of $3/2$ is introduced to make sure that after all approximations made in the derivation, the noise variance correction factors never *decrease* the uncertainty in the measurement, setting $\eta_i = 1$ as a lower limit. This threshold is increased whenever it becomes possible for the pure noise variance of a data point,

$$N_{ii} = (N + E)_{ii} - E_{ii} = \eta_i \left(\sigma_i^2 + \sigma_e^2 \right) - \eta_e \sigma_e^2, \tag{50}$$

to decrease with respect to the initial value σ_i^2.

Finally, for the correction factor for the extragalactic variance, we again use Jeffreys prior, i.e., $\beta_e \to 1$ and $r_e \to 0$. This prior is broader than needed, since the order of magnitude of the extragalactic variance is already known, e.g. from Schnitzeler (2010). However, we expect the extragalactic variance to be sufficiently constrained by the data in the VIP category, so that we do not need to constrain it with the prior. As in the case with the spectral smoothness prior, we choose to forgo the use of a stricter prior in favor of a more data-driven analysis.

4.3. Simulation

To investigate the properties and limitations of the algorithm we developed in the previous section, we will apply it to a simulation of the Faraday sky.

4.3.1. Simulation setup

We model the Galactic Faraday depth as a dimensionless, isotropic, correlated Gaussian random field multiplied with a latitude-dependent profile function. The profile function and angular power spectrum that we use are shown in Fig. 5 and 6, respectively, with a thick solid line. Our choice of profile function is modeled on the one found by Oppermann et al. (2012) and as a power spectrum we choose a simple broken power-law,

$$C_\ell \propto \left(1 + \left(\frac{\ell}{\ell_0} \right)^2 \right)^{-\gamma/2}, \tag{51}$$

with breaking point $\ell_0 = 5$ and spectral index $\gamma = 2.5$, as an arbitrary model that is not too far removed from what we expect. We choose the normalization of the power spectrum in such a way that the resulting dimensionless field has variance one. This field and the corresponding map of the simulated Galactic Faraday depth are shown in the top two panels of Fig. 7.

The data points we simulate are at the source locations of the catalog used by Oppermann et al. (2012), which we will again use in Sect. 5. For the extragalactic contribution to these data points, we draw independent zero-centered Gaussian random numbers with a dispersion of $\sigma_e^{(\mathrm{true})} = 10\,\mathrm{rad/m^2}$. This is higher than the fiducial value of $\sigma_e = 6.6\,\mathrm{rad/m^2}$ that we use

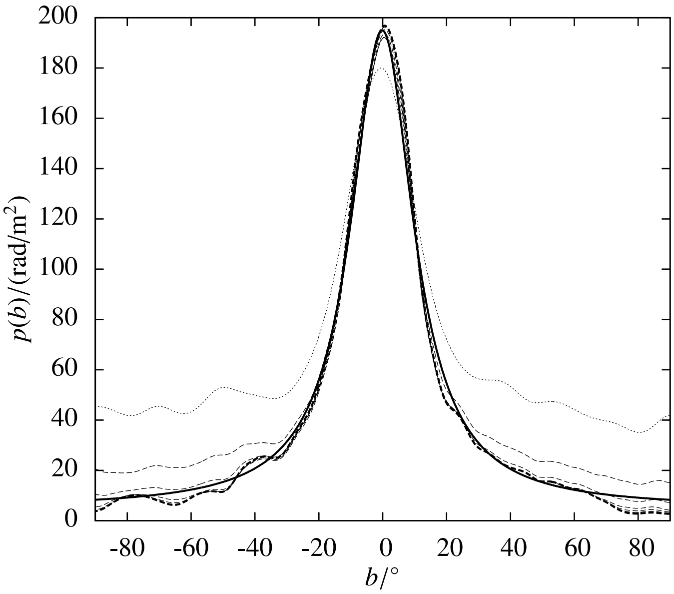

Fig. 5. Galactic latitude profile used in the simulation and its reconstruction. The thick solid line is the simulated profile, the thin dotted line is the profile initially used in the reconstruction and calculated directly from the simulated data. Subsequent reconstructions of the profile function are shown as thin dashed lines with earlier iterations lying higher in the plot. The final reconstruction is depicted by the thick dashed line.

in our reconstruction. Checking whether the algorithm is able to pick up this discrepancy is one of the main purposes of this study.

Finally, we assume uncorrelated Gaussian noise contributions for all data points. The individual noise variances are given by the error bars of the catalog in most cases. However, we do split the data points into two categories, as described earlier, and increase the error variance for a randomly chosen subset of 5% of the data points in the SIP category by an arbitrary factor of 400 (i.e., we increase the error bar by a factor 20), without informing the reconstruction algorithm about the magnitude and locations of this effect. We perform the split of the data into the two categories according to the methodology used in the calculation of the data values, putting all data points derived via a λ^2-fit into the SIP category and all data points derived using the RM-synthesis technique into the VIP category. Table 1 of Oppermann et al. (2012) includes this information for all data points. The resulting distribution of data points of categories one and two, as well as the location of the data points with increased noise variance, is shown in Fig. 8.

4.3.2. Results

The different iterations of the Galactic latitude profile are shown in Fig. 5. As is evident, the iteration converges quickly. We therefore stop our reconstruction after six iterations of the Galactic profile function. Figure 6 shows the reconstructed angular power spectra at the end of each iteration with a given latitude profile. As can be seen, the correct power spectrum is reconstructed rather well at the end. Note that the data contain information only on scales that are larger than the angular separation of sources. This separation is of course different in different regions of the sky. Typical source separations are of the order 1°, however, the data are still dominated by noise and extragalactic contributions on this scale. This can explain the mismatch between the true

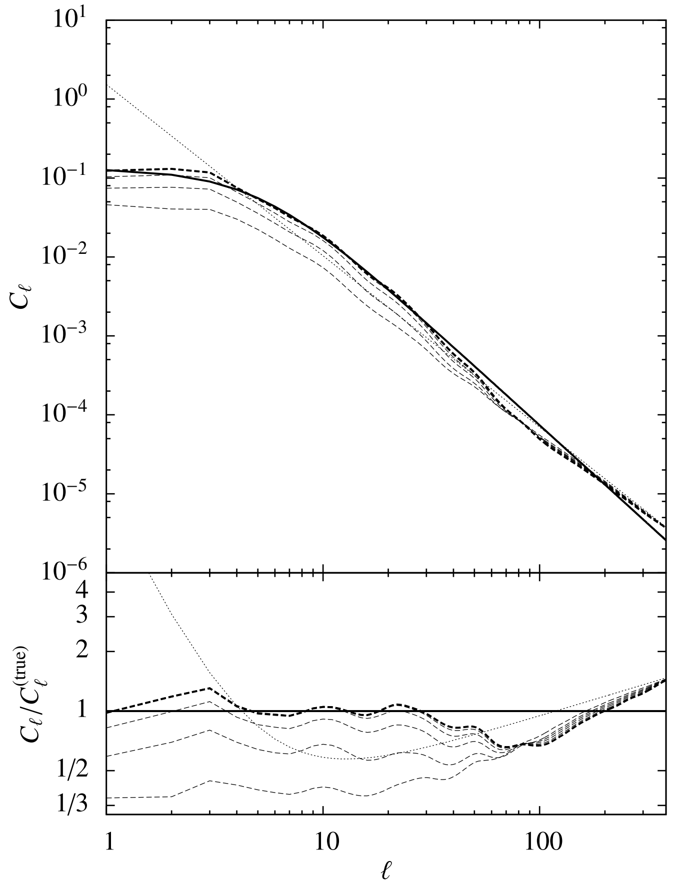

Fig. 6. Angular power spectrum of the simulated dimensionless version of the Galactic Faraday depth and its reconstructions. The thick solid line is the simulated power spectrum, the thin dotted line shows the initial guess, the thin dashed lines are subsequent reconstrucions, and the thick dashed line is the final reconstruction. Shown are the reconstructions at the end of each iteration with fixed Galactic latitude profile. In the bottom panel, we show the ratio of the reconstructed power spectra and the one used in the simulation.

and reconstructed power spectra even on scales that are a factor of a few larger than the typical source separation.

The reconstructed map of the dimensionless isotropic Gaussian random field is shown in the left column of Fig. 7, along with a map quantifying its uncertainty per pixel, given by $\mathrm{diag}(D)$. In this simulated scenario, where we know what the map is that we are trying to reconstruct, we can quantify this uncertainty information by checking the number of pixels for which the true value lies within the interval given by the reconstruction plus or minus the uncertainty, i.e., $m \pm \sqrt{\mathrm{diag}(D)}$. We find in our example that this is the case for 66% of the pixels, hinting that the Gaussian approximation to the posterior PDF that we are calculating is likely not too far from the true posterior. The right column of Fig. 7 shows the same for the physical Faraday depth. Sixty-three percent of the pixels have an estimated Galactic Faraday depth that lies within the approximate 1σ interval around the simulated value. These maps demonstrate that, with the data that we have simulated, the reconstruction and the true map agree on large and intermediate scales. Only the small-scale features are missing in the reconstructed map. This effect is of course more prominent in regions of the sky where the data density is lower.

To study the reconstructed error variance correction factors, $(\eta_i)_i$, we plot a histogram of these in Fig. 9. Clearly, the noise variance was increased for a significant fraction of the data points in the SIP category. The mean value for η_i is 6.2 and the median is 2.0. Also plotted in Fig. 9 is a histogram only for the data points for which the error variance was indeed increased in the simulation. The mean and median of these η_i-factors is 82.7 and 31.2, respectively. Taking the mean and median for all data points in the SIP category for which the error variance was *not* increased in the simulation, on the other hand, yields 2.3 and 2.0, respectively. So while the algorithm tends to increase all error bars slightly, there is clearly a trend for the error bars of the right data points to be increased much more severely.

The factor η_e that corrects the assumed extragalactic variance is reconstructed in this example to be 2.6, corresponding to a standard deviation of

$$\hat{\sigma}_\mathrm{e} = \sqrt{\eta_\mathrm{e}\, \sigma_\mathrm{e}^2} \approx 10.7\,\mathrm{rad/m^2}, \tag{52}$$

which is close to the value of $\sigma_\mathrm{e}^{(\mathrm{true})} = 10\,\mathrm{rad/m^2}$ that we put into the simulation. Therefore we can conclude that with our algorithm we are able to reconstruct the variance of the extragalactic contribution with high precision. In principle, we could quantify the uncertainty of this estimate by taking the second derivative of the PDF given by Eq. (A.22) at its maximum. However, we expect that the error in this estimate will in reality be dominated by the limitations of our assumptions, i.e., the fact that the extragalactic Faraday rotation is not exactly an isotropic, uncorrelated, Gaussian random field, and the ambiguity of the categorization of the data points, and not so much by the statistical information content of the data. We therefore believe that the quantification of the statistical uncertainty would be potentially misleading and therefore is not worth the computational effort. Even in this simulated scenario, the difference between the reconstructed $10.7\,\mathrm{rad/m^2}$ and the $10\,\mathrm{rad/m^2}$ is probably mostly due to the approximations made in the derivation of our filter formulas.

Finally, in Fig. 10, we plot the resulting extragalactic contributions for each data point versus Galactic latitude. Just like for the estimators discussed in Sect. 3, some artifacts of the filtering procedure are obviously present in the resulting estimate. In the top panel of Fig. 10, we plot the estimates for the data points of the two different categories in different colors. The plot shows that the data points in the SIP category end up with estimates that have a clear latitude dependence and a rather sharp cut-off around $\left|\hat{\phi}_\mathrm{e}\right| \approx 8\,\mathrm{rad/m^2}$. The estimates for the data points of the VIP category also show a dependence on Galactic latitude, however, their spread is generally larger and their distribution does not exhibit a sharp cut-off.

As with the other estimators, it should be made absolutely clear that what we plot here is only an *estimate* of the extragalactic contributions. In our analysis, this estimate is calculated in a different way for the VIP and SIP data. Therefore, it is not entirely surprising that the distribution of the resulting values is also qualitatively different for the two categories. In essence, the estimation of the extragalactic contributions is easier when the Galactic and noise contributions are more tightly constrained. Therefore, the algorithm will find seemingly large extragalactic contributions more trustworthy for data points for which this is the case. This can likely explain the tendency for larger estimates for the data points of the VIP category, for which the noise variance cannot be increased. It can also explain the latitude dependence of the estimates for the data points of the SIP category, for which the Galactic contributions are less well constrained near the Galactic plane. The sharp cut-off for the estimates for the data points of the SIP category can be interpreted as a threshold beyond which our assumptions make it more believable that

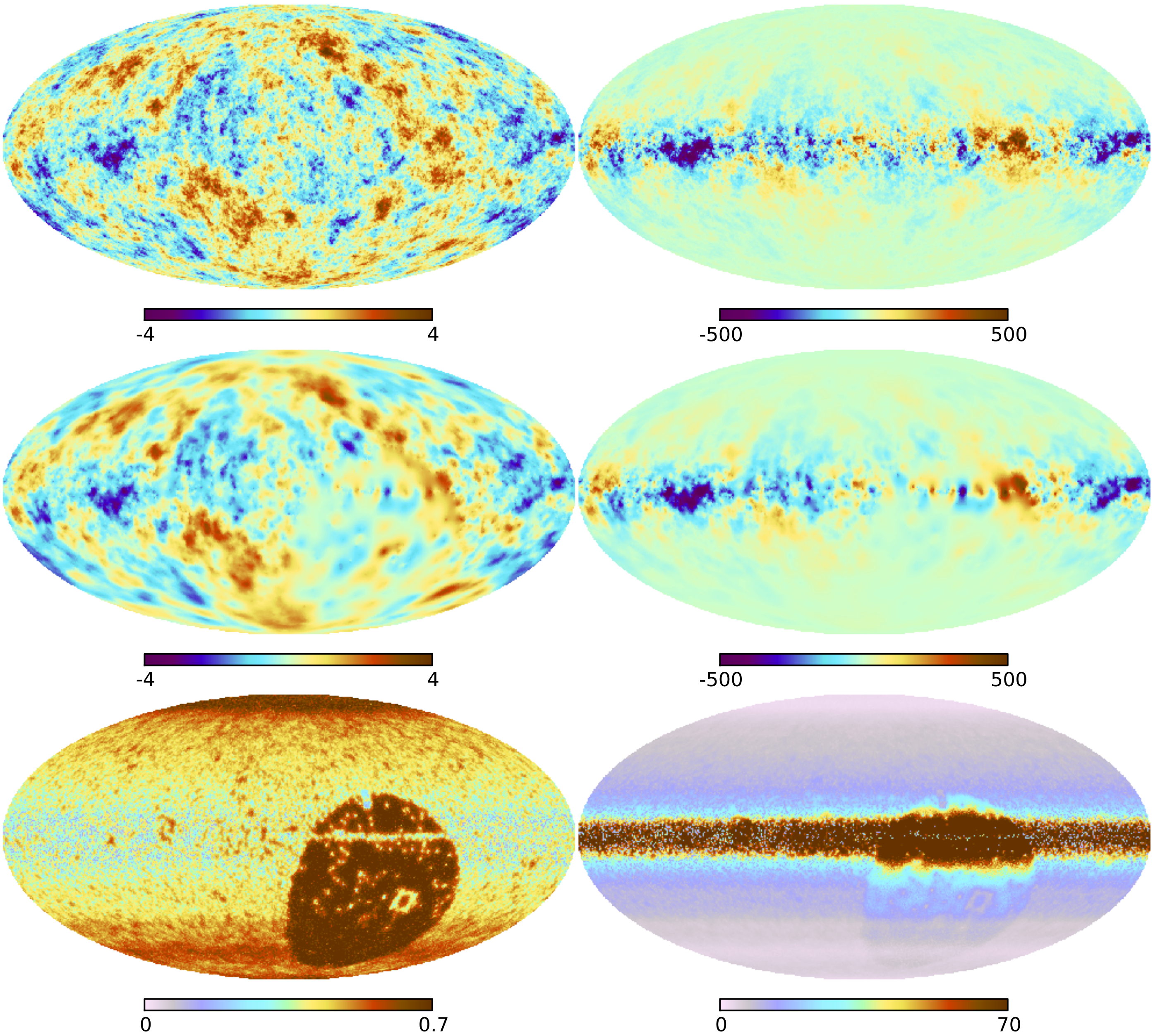

Fig. 7. Simulation and reconstruction of the Galactic contribution to Faraday rotation. The left column shows the dimensionless isotropic Gaussian random field used in our statistical model. The right column shows the resulting physical Galactic Faraday depth in units of rad/m^2. The top row shows the simulation, the middle row its reconstruction, and the bottom row the reconstruction's uncertainty per pixel.

the noise variance should be increased than that the extragalactic contribution is larger.

As before, we should take a look at the uncertainty of this estimate as well, given by the covariance matrix D_e that describes the approximate Gaussian posterior PDF, Eq. (A.6). Note that this equation is the same as Eq. (29), which describes the posterior if we fix the estimates for the prior covariance matrices G, E, and N for the three contributions to the data. As in Sect. 3, we draw a random sample from this distribution to demonstrate the effect of the remaining uncertainty. This is plotted in the lower panel of Fig. 10. Obviously, the artifacts visible in the estimate of the extragalactic contributions are mostly compensated by the uncertainties of these estimates. As is true for any scientific analysis, our state of knowledge about the extragalactic contributions to Faraday rotation after analyzing the data set cannot be described completely by an estimate, but only by a probability distribution. The random sample drawn from this distribution shows that this distribution does not exhibit any crude artifacts of the analysis, only the attempted summary in a single estimate does.

One possible point of curiosity with regards to our methodology could be that, quite naturally, we have assumed that the prior PDFs for the Galactic and extragalactic contributions are centered on zero. This is motivated by the conjecture that there is no reason to assume that a preferred direction for magnetic fields exists, both in the Milky Way and on larger scales in the universe. Therefore, a Faraday depth of $0\,\mathrm{rad/m^2}$ is effectively treated slightly differently from large absolute Faraday depths, even though observationally, they are usually treated in the same way. To investigate the effect that a mean different from zero in the Galactic component has on our reconstruction, we studied

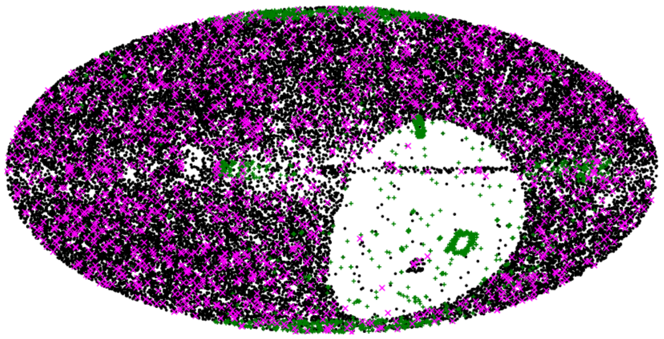

Fig. 8. Locations of the simulated data points on the sky. The magenta ×-signs denote data points of the SIP category for which the noise variance has been increased. Other data points of the SIP category are shown as black dots, while data points of the VIP category are shown as green symbols.

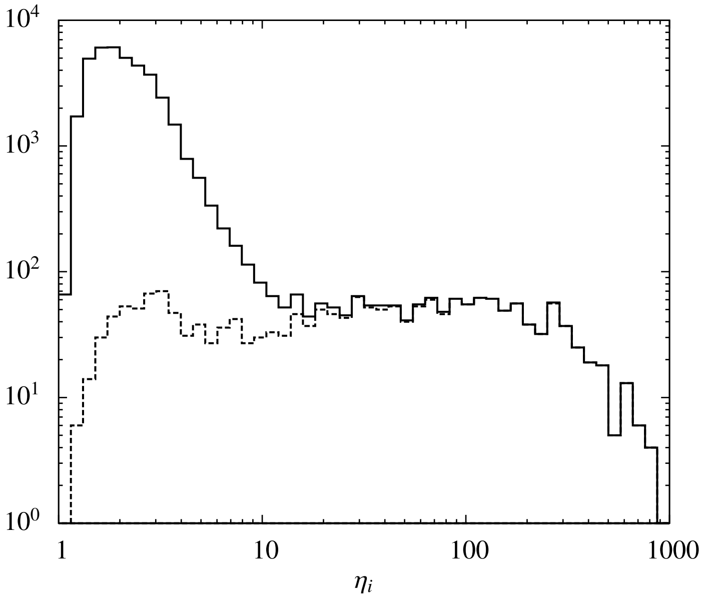

Fig. 9. Histogram of the error variance correction factors η_i for the data points of the SIP category in the simulation discussed in Sect. 4.3. The solid line shows the histogram for all data points in this category, while the dashed line shows the histogram only for the data points for which the noise variance was indeed increased in the simulation. Note the logarithmic scale on both axes.

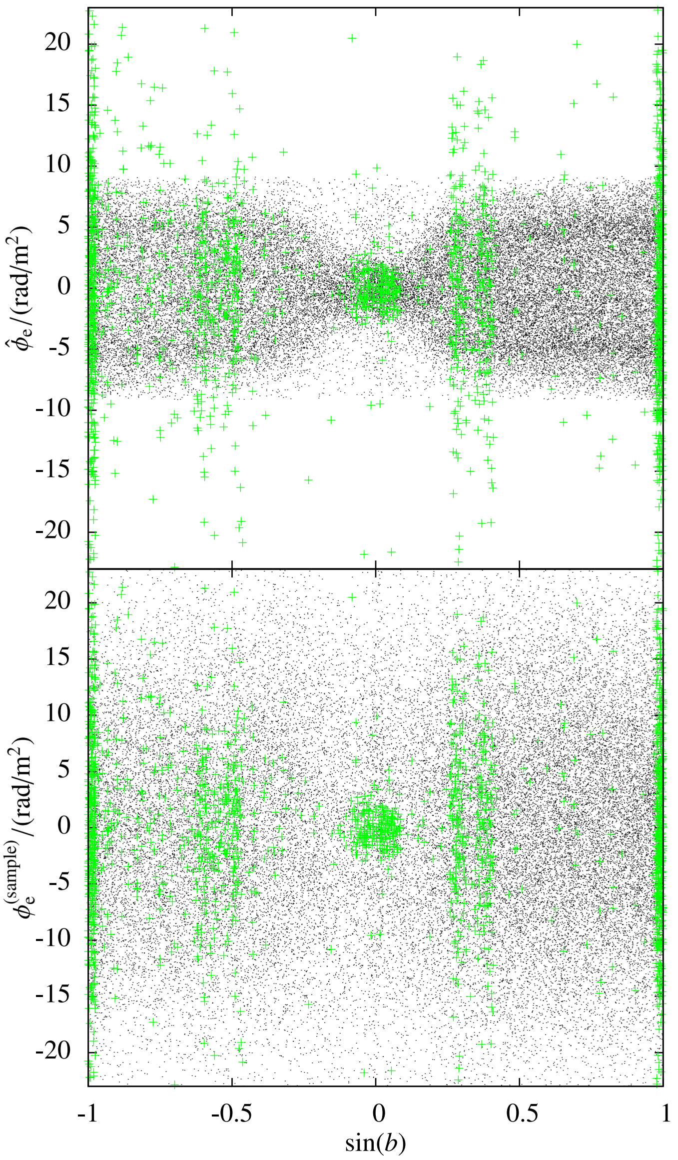

Fig. 10. Extragalactic contribution to each source's observed Faraday depth versus Galactic latitude for the simulated scenario discussed in Sect. 4.3. The upper panel shows the approximate posterior mean estimate as calculated by our algorithm, the lower panel shows a random sample drawn from the posterior PDF. Data points of the SIP category, i.e., with noise variance correction factors, are plotted in black, data points of the VIP category, i.e., without noise variance correction factors, are plotted in green.

the same simulation with a constant offset of 5 rad/m^2 added to each data point. In this case we find an extragalactic dispersion of $\hat{\sigma}_e = 9.7$ and get essentially the same reconstruction with the Galactic foreground offset by 5 rad/m^2. This constant offset was chosen for simplicity to represent a non-zero mean that might marginally be hidden in the data without being apparent a priori, although a constant Faraday depth in all directions, corresponding to magnetic field lines converging near the solar system, is of course unlikely. We note that for the data set we will use in Sect. 5, a simple average of the observed Faraday depths gives -1.8 rad/m^2. A large constant offset is therefore also unrealistic from an observational point of view.

Another point of concern might be that a single data point in our formalism enters both the estimation of the Galactic foreground and the estimation of the extragalactic contribution. This seems to be using the same data point twice, if one views the analysis as a two-step procedure in which first the Galactic foreground map is calculated and then the difference between the data value and the Galactic foreground is post-processed. However, our algorithm performs a simultaneous reconstruction of the Galactic foreground and the extragalactic contributions, even though the actual implementation is iterative. It is therefore only natural to use the entirety of the data in each step of this iteration. In Appendix B, we discuss the differences between using each data point in the reconstruction of the Galactic foreground and excluding the data point in whose extragalactic contribution one is interested for a simplistic foreground reconstruction technique.

Finally, we performed another simulation along the same lines as the one discussed in this section, but with a true variance of the extragalactic contributions that was assumed to be lower than the initial guess. The statements we have made about the Galactic foreground reconstruction are equally true in this case and the difference between the reconstructed extragalactic variance and the one used in the simulation is roughly the same as in the case that we discussed in this section.

5. Application to real data

5.1. Description of the data

In the following, we make use of the data catalogs assembled by Oppermann et al. (2012) and described in their Table 1. We add the new catalog of Mao et al. (2012), which has the same specifications as the catalog of Van Eck et al. (2011), detailed in the table, except for the number of sources and their locations. Altogether, this data set consists of 41 632 observationally estimated Faraday depths for extragalactic sources. The extragalactic nature of the sources is not entirely guaranteed for the NVSS rotation measure catalog (Taylor et al. 2009). While it is possible that a few of the data points from this catalog correspond to pulsars in the Milky Way, we note that the overwhelming majority of the sources has to be extragalactic. Pulsars, for which not the complete line of sight to the outer edge of the Milky Way is probed by the observations, provide one more reason to attempt a reconstruction that is robust against an incomplete description of the observational uncertainties.

The data set is rather inhomogeneous both spatially, with a relatively sparse source population in the southern equatorial hemisphere, and with a view on the observational parameters, ranging from linear fits to polarization angle measurements in two adjacent frequency bands to RM synthesis studies over wide bands in λ^2-space.

In the following, we multiply the published error bars of Taylor et al. (2009) by a factor 1.22, according to Sect. 4.2.1 of Stil et al. (2011).

5.2. Possible splits

The algorithm presented in the previous section hinges on the assumption that we can split the data set into a subset for which the likelihood is fully described by the published Gaussian error bars (VIP data) and another subset for which this is not necessarily the case (SIP data). How to judge whether a data point should be in the VIP category or the SIP category is not clear. Aspects that are to be considered in this decision are that a continuous frequency coverage eliminates the risk of a polarization angle rotation by a multiple of π between bands and that a wider coverage in λ^2-space leads to a higher resolution in Faraday depth space and therefore a lower risk of misleading results occurring from several emission components within the same beam, as described by Farnsworth et al. (2011) and Kumazaki et al. (2014). This demands a large fractional band-width. Furthermore, the estimation of the variance of the extragalactic contribution relies mostly on the data points that we assign to the VIP category. As we pointed out earlier, this estimation is complicated by large contributions from the Milky Way and large noise contributions. It is therefore desirable to have at least some data points of the VIP category away from the Galactic plane. Finally, it is good to split the data in a conservative way. Assigning to the SIP category a data point for which the likelihood is well described by the given error bar will not bias the result, only increase the posterior uncertainty. Assigning to the VIP category a data point for which the likelihood is insufficiently described, however, will in most cases lead to an overestimated variance of the extragalactic contribution and thus influence all other results of the reconstruction.

Instead of arguing for a single definitive split, we will explore a set of different possibilities. This will enable us to make statements about the reliability of the results. The following ways of splitting the data set will be used:

1. Five catalogs with a wide frequency coverage at particularly low frequencies are regarded as data of the VIP category. These are the catalogs referred to as O'Sullivan (O'Sullivan, private communication, 2011), Heald (Heald et al. 2009), Schnitzeler (Schnitzeler, private communication, 2011), as well as Mao SouthCap and Mao NorthCap (Mao et al. 2010) in Table 1 of Oppermann et al. (2012). All other data are considered part of the SIP category. We will refer to this split as the 'bandwidth' split.
2. Only the two catalogs consisting entirely of data points near the Galactic poles, i.e., Mao SouthCap and Mao NorthCap, are considered as the VIP category. These combine the demand for large coverage in λ^2-space with a low foreground region in the sky. All other data are considered SIP category data. We will refer to this split as the 'polar caps' split.
3. Only the Mao NorthCap and Mao SouthCap data are used. These are considered data of the VIP category. All other data are completely ignored, i.e., there is no SIP category of data. This means that the reconstruction is completely insensitive to anything that happens at low Galactic latitudes and therefore our assumption of approximate isotropy for the Galactic foreground can be expected to be rather accurate in this case. We will refer to this split as 'polar caps only'.
4. The data of the O'Sullivan, Heald, and Schnitzeler catalogs are regarded as being of the VIP category, all other data are in the SIP category. This is done to see the effect that having VIP data in regions with a large foreground may have on the result. We will refer to this split as the 'complement' split, as it regards as VIP data the points that are regarded as VIP under the 'bandwidth' condition but not under the 'polar caps' condition.
5. Only data points with Galactic latitudes that satisfy $|b| > 45°$ are considered at all. Of these, the data that stem from any of the O'Sullivan, Heald, Mao NorthCap, Mao SouthCap, or Schnitzeler catalogs and additionally satisfy $|b| > 55°$ are considered as comprising the VIP category. The last condition is introduced to avoid any potential boundary effects on the reconstruction of the extragalactic variance. Otherwise, this is essentially an extension of the 'polar caps only' ansatz, which adds a few data points of the VIP category and a significant number of data points in the SIP category. We will refer to this split as the 'around polar caps' split.
6. All data points are considered part of the VIP category. The SIP category is empty, i.e., the observational uncertainty is regarded as precisely reliable for each and every data point. We regard this split as a cross-check to see whether the algorithm behaves in the expected way if we contaminate the VIP data category with data points for which the uncertainty information is incomplete. We will refer to this assumption as 'all VIP'.
7. Only the data from the Mao NorthCap catalog are considered part of the VIP category. All other data points are considered SIP category data. We will refer to this split as 'north polar'.
8. Only the data from the Mao SouthCap catalog are considered part of the VIP category. All other data points are considered SIP category data. We will use this split and the one before

as consistency checks for the results of the 'polar caps' split. This split will be referred to as 'south polar'.

9. 10 000 randomly chosen data points are assigned to the VIP data category. The rest of the data (i.e., 31632 measurements) are used as SIP category data. We will refer to this as the 'random' split.

Table 1 gives an overview of the data splits we consider and Fig. 11 shows the locations of the VIP and SIP data points in the first six splits. Note that the majority of the data points (37 543 points) stems from the NVSS RM catalog (Taylor et al. 2009). These are either regarded as part of the SIP category or neglected completely in our splits, except for the 'all VIP' and 'random' cases.

5.3. Results and discussion

In this section, we will first discuss the results of the first six splits in detail in Sect. 5.3.1 and then introduce the most important aspects of the remaining splits. We will argue for adopting the 'polar caps' split as a reasonable fiducial model and use the 'north polar' and 'south polar' splits as cross-checks for the reliability of the results derived under this split in Sect. 5.3.2. Finally, we will present detailed results for the 'polar caps' split in Sect. 5.3.3.

5.3.1. The first six data splits

Figure 12 shows the reconstructions of the Galactic contribution in the first six cases. Naturally, the reconstruction in the 'polar caps only' and 'around polar caps' cases suffers from the scarcity of data. The 'all VIP' reconstruction shows small scale structure, especially in the plane (note for example the Galactic center), that is washed out in the other reconstructions, again as expected. The three other reconstructions that make use of the entirety of the data are rather similar. Some details, however, do differ. Note for example the blob of positive Galactic Faraday depth at $l \approx 275°$, $b \approx 10°$ that is present in the 'bandwidth' and 'complement' reconstructions but not in the 'polar caps' reconstruction. The reasons for these differences are not immediately apparent, and are most likely due to the interplay of all involved quantities and possibly an instability with respect to numerical inaccuracies.

The reconstructed angular power spectra of the dimensionless Galactic signal fields in the first six cases are shown in Fig. 13. Evidently, the resulting spectra are all very similar. In the 'polar caps only' and 'around polar caps' cases, in which a large fraction of the data were ignored, the result is closer to a pure power law, since the spectral smoothness prior becomes more important in these cases. A power law with a spectral index of -2.17 is a good fit to these spectra, as was already found by Oppermann et al. (2012).

Figure 14 shows the variance profiles for the Galactic contribution introduced in Eq. (30) that result from the six different data splits. Obviously, in the 'polar caps' and 'around polar caps' cases, the profile function is only reconstructed well near the poles, as all the low latitude data are ignored. Among the other reconstructions, the profile functions do not differ greatly, with the main differences appearing near the Galactic plane. The 'all VIP' ansatz leads to a higher variance near the Galactic plane, whereas a smaller fraction of VIP data leads to a more heavily smoothed Galactic reconstruction and therefore less variance and a slightly lower profile function, as exemplified by the 'polar caps' and 'complement' splits.

The reconstructed values of the extragalactic dispersion, σ_e, are presented in Table 1. Note that the 'polar caps', 'polar caps only', and 'around polar caps' numbers are rather similar. For all of these reconstructions, the VIP category of data is dominated by the Mao NorthCap and Mao SouthCap catalogs. The number yielded by the 'bandwidth' reconstruction is not very different either. In this case the Mao SouthCap and Mao NorthCap catalogs still comprise more than 70% of the VIP data category. The 'complement' number differs significantly, indicating that the assumptions made for the VIP category of data points are probably not met by all of the data points in the O'Sullivan, Heald, and Schnitzeler catalogs. Another factor here may be the fact that in the 'complement' split, the data of the VIP category are rather few and far in between (a total of 281 data points, mostly in the southern equatorial hemisphere, some in the northern hemisphere). Obviously, the 'all VIP' scenario yields a number that is even more in disagreement.

Figure 15 shows a comparison of the estimates for the extragalactic contributions under the six different assumptions. Black points are for data points that are in the VIP category under both assumptions that are being compared in each individual panel. As can be seen in the figure, the estimates for the extragalactic contributions are basically the same for data points that are in the VIP category for both compared scenarios if the estimate of σ_e is similar in the two scenarios. A significantly larger estimate of σ_e, however, leads to increased estimates of ϕ_e as well, as can be seen most clearly in the top row of the figure. The black dots in these panels still follow a linear relationship, but the slope deviates from one.

Red points in Fig. 15 are data points that are in the SIP category under both of the two assumptions that are being compared. These red dots show essentially the same effect as the black dots, namely that the estimates are the same if the estimate of σ_e is the same and a higher estimate of σ_e results in a higher estimate of ϕ_e. The latter effect is visible whenever the 'complement' estimate is part of the comparison. Overall, the red dots have a smaller dispersion than the black ones. This is expected, as part of the discrepancy between data and Galactic reconstruction can be explained by increased error bars in case of data in the SIP category.

Finally, the green points in Fig. 15 show data points that are in the SIP category under one of the assumptions that are being compared and in the VIP category under the other. These points in the figure show that allowing the noise variance to be corrected upward for a data point will keep the estimate of its extragalactic contribution small. Note the bifurcation in the green points in the comparisons between the 'polar caps' and 'complement' estimates and between the 'around polar caps' and 'complement' estimates. This is due to data points that are in the SIP category under the 'complement' split on the one hand and data points being in the SIP category under the 'polar caps' and 'around polar caps' splits on the other hand. Overall, all the trends exhibited by the estimates shown in Fig. 15 follow the expectation.

The effect of interpreting the error bars as a complete description of the likelihood functions can be seen clearly in the first row of panels in Fig. 15. For many data points the observed Faraday depth cannot be explained by the Galactic foreground reconstruction and the published noise variance alone. The algorithm will therefore increase the dispersion σ_e of the extragalactic contributions until it agrees with the dispersion of the remaining differences. This leads to the high reconstructed value of σ_e and the large estimates for the values of ϕ_e seen in Fig. 15 in the 'all VIP' case. If, on the other hand, only a subset of the error

Table 1. Overview of the different data splits considered for the analysis of the observational data. The first column gives the name by which the split is referred to, the second one describes the criterion by which data are assigned to the VIP category, the third column lists the data catalogs whose data points are assigned to the VIP category, the fourth column gives the criterion for a data point to be part of the SIP category. The resulting estimate for the dispersion of the extragalactic contributions is given in the second to last column and the last column lists the figures that show results obtained under the data split in question.

data split	condition for VIP data	VIP catalogs	condition for SIP data	$\hat{\sigma}_{\mathrm{e}}/(\mathrm{rad/m}^2)$	Figs.
'bandwidth'	high bandwidth	O'Sullivan[a], Heald[b], Schnitzeler[c], Mao SouthCap[d], Mao NorthCap[d]	remaining data	7.7	11 - 16
'polar caps'	high bandwidth & low foreground	Mao SouthCap[d], Mao NorthCap[d]	remaining data	6.4	11 - 19, C.1 - C.3
'polar caps only'	high bandwidth & low foreground	Mao SouthCap[d], Mao NorthCap[d]	-	7.1	11 - 15
'complement'	high bandwidth & not low foreground	O'Sullivan[a], Heald[b], Schnitzeler[c]	remaining data	16.0	11 - 16
'around polar caps'	high bandwidth & $\|b\| > 55°$	Mao SouthCap[d], Mao NorthCap[d], part of O'Sullivan[a], Heald[b], Schnitzeler[c]	$\|b\| > 45°$	7.0	11 - 16
'all VIP'	all data	all	-	38.8	11 - 15
'random'	10 000 random data points	from all catalogs	remaining data	30.4	-
'north polar'	high bandwidth & near Galactic north pole	Mao NorthCap[d]	remaining data	6.4	-
'south polar'	high bandwidth & near Galactic south pole	Mao SouthCap[d]	remaining data	6.0	-

Notes. [(a)] O'Sullivan, private communication (2011) [(b)] Heald et al. (2009) [(c)] Schnitzeler, private communication (2011) [(d)] Mao et al. (2010)

bars are assumed to accurately describing the likelihood functions, the estimate of σ_{e} will be dominated by this subset of the data points (i.e., the VIP data). Further, if this assumption is indeed true for the chosen subset, the estimate of σ_{e} will naturally be lower. Consequently, the algorithm will explain large differences between observed Faraday depths for the SIP data and the Galactic foreground reconstruction as largely due to a likelihood function that is wider than described by the error bar, i.e., an error variance correction factor η_i that is significantly larger than one. This keeps the estimated extragalactic contributions relatively small, as plotted on the horizontal axes of the first row of panels in Fig. 15.

Figure 16 compares the error variance correction factors, η_i, for data that are in the SIP category under two sets of assumptions. This comparison can only be done for the four splits of the six under consideration that do have data of the SIP category, namely the 'bandwidth', 'polar caps', 'complement', and 'around polar caps' splits. This shows that the η_i factors are almost unaffected by the differences between the 'bandwidth', 'polar caps', and 'around polar caps' assumptions. The 'complement' split, however, leads to systematically larger correction factors. This is most likely due to a contamination of the VIP data category by data for which the likelihood is not well described by the observational error bars. Such a data point in the VIP category will lead to an increase of the estimated variance of the extragalactic contribution. However, the data points that were appropriately placed into the VIP category will prevent this estimate from growing enough to completely explain the faulty data point. As a result, the faulty data point will still have a strong influence on the reconstruction of the Galactic foreground in its vicinity, thus necessitating increased noise variances for other data points to be consistent with this reconstruction. To check whether this is indeed a generic effect of a contaminated VIP data category, we now study the 'random' split, for which we randomly assign 10 000 of the 41 632 data points to the VIP category and run our reconstruction algorithm under these assumptions. We find as expected a high estimate for the extragalactic dispersion of $\sigma_{\mathrm{e}} = 30.4\,\mathrm{rad/m}^2$. Furthermore, we find the same trend for increased error variance factors as for the 'complement' split. We therefore conclude that these enlarged error variances are indeed a sign for a contamination of the VIP data category with data that should have been put into the SIP category.

5.3.2. Cross-checks for the 'polar caps' split

From the splits that we have studied so far, we regard as most reliable the ones that use only the Mao SouthCap and Mao NorthCap data as data of the VIP category, i.e., the 'polar caps' and the 'polar caps only' splits. To check whether this is indeed the case or whether the values for $\hat{\sigma}_{\mathrm{e}}$ calculated under the splits that are dominated by these data sets are only similar by chance, we consider two further subsets of the data for the VIP category. We perform reconstructions using the entirety of the data with only the data from the Mao NorthCap catalog and only the data from the Mao SouthCap catalog assigned to the VIP data cate-

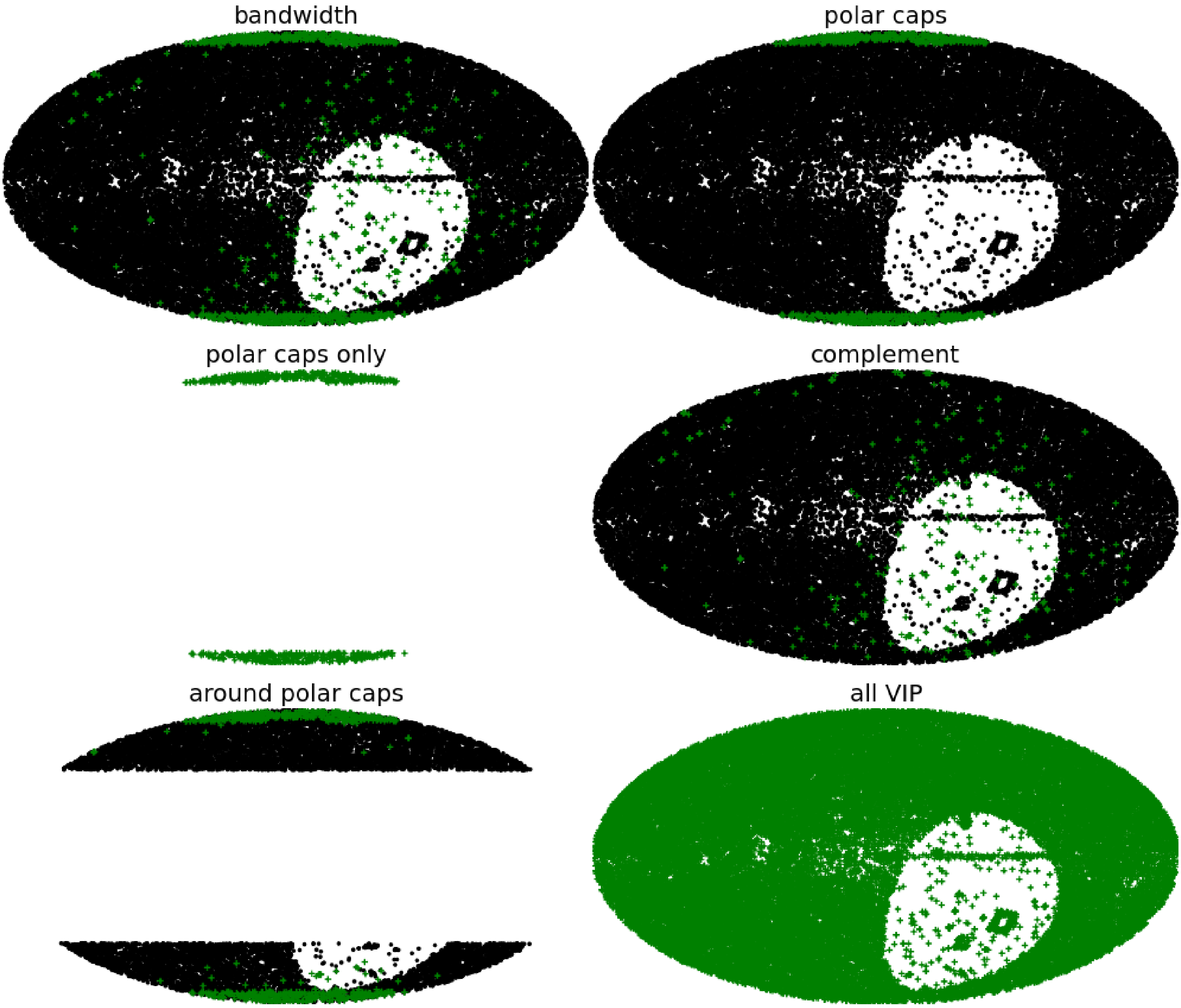

Fig. 11. Locations of the data points in the sky for the different data splits. The black dots show the locations of data that are considered part of the SIP category, while VIP data points are marked in green. The labels refer to the first six data splits described in Sect. 5.2 and in Table 1.

gory, the 'north polar' and 'south polar' splits mentioned before. The resulting values for the estimated extragalactic dispersion are $\hat{\sigma}_{\rm e} = 6.4\,{\rm rad/m^2}$ and $\hat{\sigma}_{\rm e} = 6.0\,{\rm rad/m^2}$, respectively, close to the value found in the 'polar caps' split. We therefore conclude that the value for the extragalactic dispersion in reality is likely to lie somewhere in the interval between $6.0\,{\rm rad/m^2}$ and $7.1\,{\rm rad/m^2}$, as spanned by the results of the different splits using the Mao NorthCap and Mao SouthCap data as the only ones for which the likelihood is assumed to be precisely described by the published Gaussian error bars. It should also be noted that the extragalactic dispersion estimated in this way will at least partially include any contribution from the Milky Way that is uncorrelated on the scales probed by current observations, i.e., typically around 1°. Therefore, the purely extragalactic dispersion may be even lower.

5.3.3. Results for the 'polar caps' split

We therefore adopt the 'polar caps' split as our fiducial model. The left panels of Fig. 17 show the estimated extragalactic contribution for this reconstruction as a function of latitude, as well as a random sample drawn from the posterior PDF for this quantity. The right panels of Fig. 17 show histograms of the distributions of the values plotted in the left panels in bins of $\sin(|b|)$. Note that again the distribution of the posterior mean estimates shows a latitude dependence, as was the case for the results in our simulated scenario shown in Fig. 10. This is due to the latitude dependence of the Galactic contribution and the estimated noise variances, as discussed in Sect. 3.2. However, as pointed out before, this is only an *estimate* for the extragalactic contributions and as such should always be regarded together with its uncertainty. To demonstrate this, we draw a random sample from the posterior PDF for the extragalactic contributions and show it in the lower panels of Fig. 17. The sample includes the spread that is due to the remaining uncertainty of the estimate. This is larger

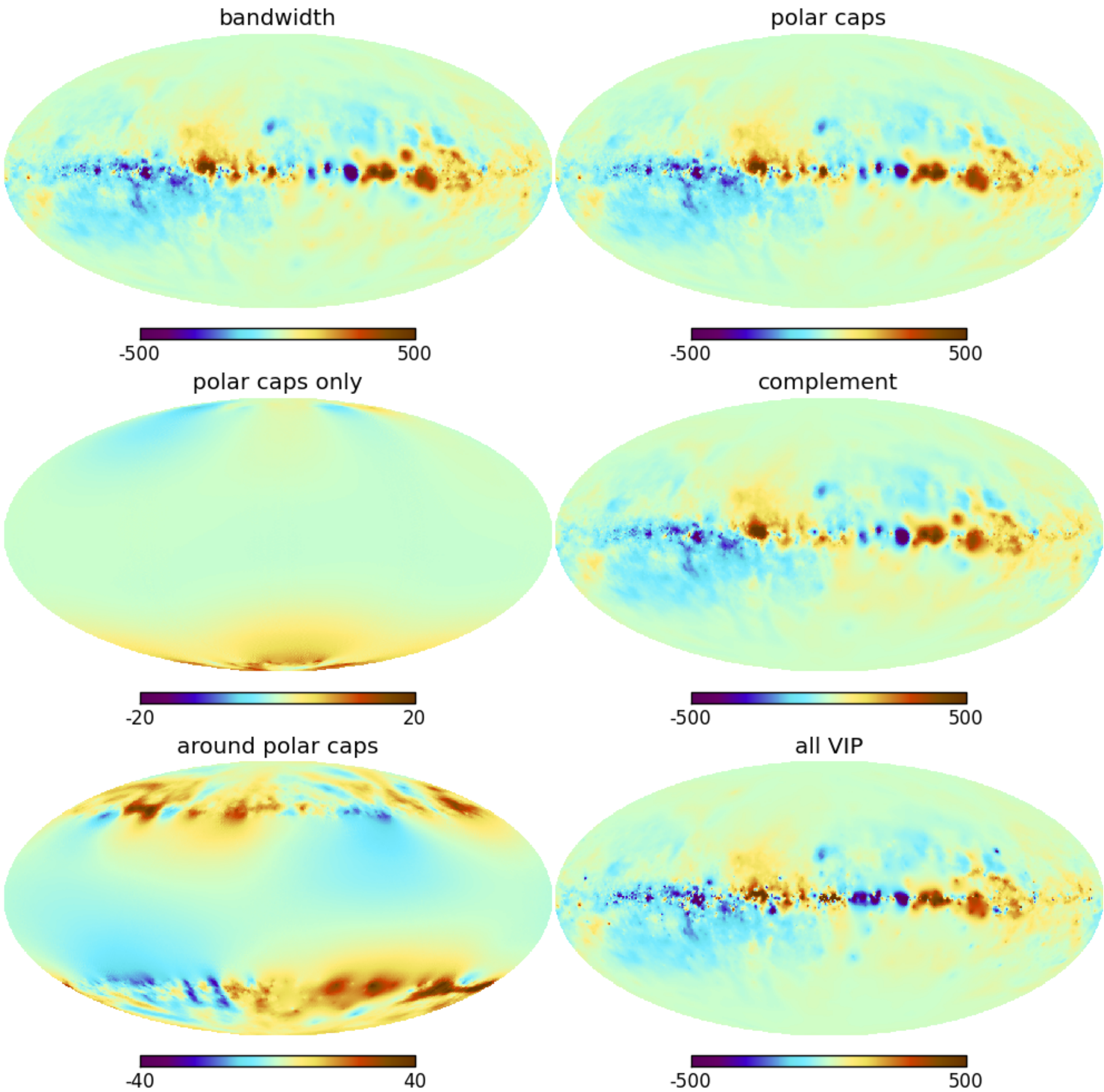

Fig. 12. Reconstructions of the Galactic foreground in the first six scenarios. The labels refer to the first six data splits described in Sect. 5.2 and Table 1. The units are rad/m^2. Note the different color scale in the 'polar caps only' and 'around polar caps' splits.

near the Galactic plane and mostly lets the latitude dependence vanish in the lower two panels of the figure. Note, however, that there is still a slight difference between the distributions for the VIP and SIP categories. We have created a website[4] where we provide such posterior samples for use in studies of extragalactic Faraday rotation along with the other results of the 'polar caps' reconstruction. Note that the samples are again not drawn from the full posterior PDF, but from the approximate one we use in our derivations. This approximate posterior is effectively the posterior for the extragalactic contributions after fixing their variance, the error variance correction factors, the angular power spectrum of the Galactic contribution, and the Galactic latitude profile, i.e., the uncertainty due to the uncertainty of these reconstructed quantities is not represented by the samples. See Appendix C for details on the provided files and the usage of the samples, as well as a discussion of numerical artifacts present in the samples.

We repeat the analysis of Fig. 3 for these new results in Fig. 18. Plotted are the sample standard deviations in bins of $\sin(b)$ with and without taking into account the posterior uncertainty. We see that even after including the posterior uncertainty, a slight latitude dependence remains (×-symbols), albeit much

[4] See `http://www.mpa-garching.mpg.de/ift/faraday/`

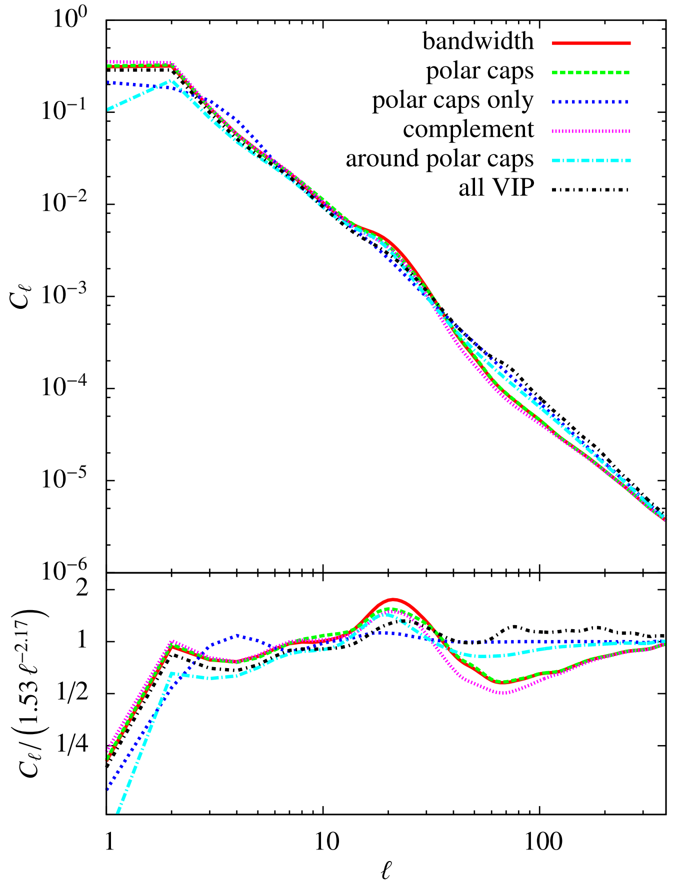

Fig. 13. Reconstructed angular power spectra of the dimensionless Galactic signal field. The lower panel shows the ratio of the reconstructed spectra and a power law fit, $C_\ell = 1.53\,\ell^{-2.17}$.

less pronounced than in the pure posterior mean estimates (+-symbols). The remaining latitude dependence is presumably due to the approximations made in our calculation, such as the assumption of approximate isotropy for the Gaussian random field

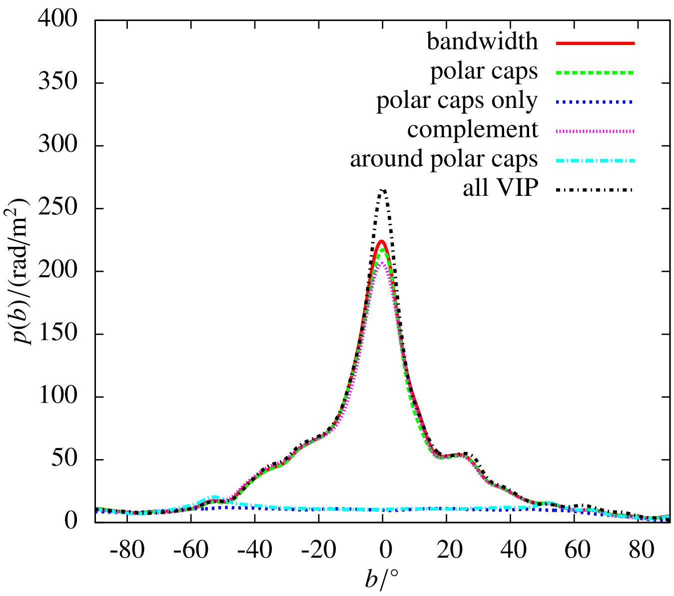

Fig. 14. Reconstructed Galactic latitude profiles, $p(b)$, describing the relationship between the dimensionless Galactic signal field s and the physical Galactic Faraday depth ϕ_g, for the six different data splits.

related to the Galactic Faraday depth by the latitude dependent profile function. Note also that the latitude dependence of the sample standard deviation for the posterior mean estimate (+ symbols) is a bit more pronounced than in Fig. 3. However, it must be noted that the outermost bins are now influenced by the VIP data points for which the spread of the posterior mean estimates is much larger.

Finally, we show in Fig. 19 the results for the Galactic reconstruction using the 'polar caps' data split, i.e., the reconstructed dimensionless signal field and physical Galactic Faraday depth, as well as their pixel-wise uncertainties, given by the square-root of the diagonal of the posterior covariance in position space representation. Figure 19 also shows a comparison to the results of Oppermann et al. (2012). As can be seen, differences are most pronounced near the Galactic plane, in the region of scarce data in the southern equatorial hemisphere, and near the poles. These differences are due to the differences of the reconstruction algorithms. The biggest difference is of course our special treatment of a subset of the data that will influence the Galactic reconstruction as well. Additionally, the spectral smoothness prior that we used here prevents the angular power spectrum from dropping off steeply at the smallest scales, as was found by Oppermann et al. (2012), and thus will lead to more small-scale fluctuations in the Galactic map. Finally, we no longer allow the error variance correction factors to become smaller than one. This will in general have a suppressing effect on the small scale structure and seems to dominate over the opposing effect, as we see more small scale structure in the third row of the figure than in the first. While the bottom panels in Fig. 19 highlight the differences with respect to the old reconstruction, it should be noted that overall the differences are small. The estimate of the Galactic Faraday depth in the 'polar caps' scenario lies within the uncertainty range published by Oppermann et al. (2012) for 95% of the pixels.

We believe that the new assumptions that we made in this work are well motivated and therefore regard the reconstruction presented here as an improvement over the reconstruction of Oppermann et al. (2012). The results of Oppermann et al. (2012) should only be used in cases in which one explicitly does not want to be influenced by one of the assumptions we made here, like the spectral smoothness prior, the assumption that error variances should not be lower than quoted in the observational catalogs, or the explicit split of the data into the two categories.

6. Summary

We have studied the contributions to the observed Faraday rotation of extragalactic sources that are due to the Galactic interstellar medium, due to extragalactic magnetic fields, and due to observational noise. Extracting any of these three contributions is non-trivial, as they are superimposed on every line of sight. Another complication is that the observational error bars do not in every case describe the data likelihood accurately. This makes even a probabilistic analysis of the fractions of the data values due to the three different constituents challenging.

If the observations were noise-less, the extragalactic contributions could be estimated by simply subtracting an estimate of the Galactic foreground from the data values. However, in reality the observations are noisy and an estimate of the extragalactic contributions calculated in this way will contain this noise as well. Simply subtracting a Galactic foreground from the data is therefore not a good way of estimating extragalactic contributions. Furthermore, any estimate of the Galactic foreground will itself be uncertain and this uncertainty, when not taken prop-

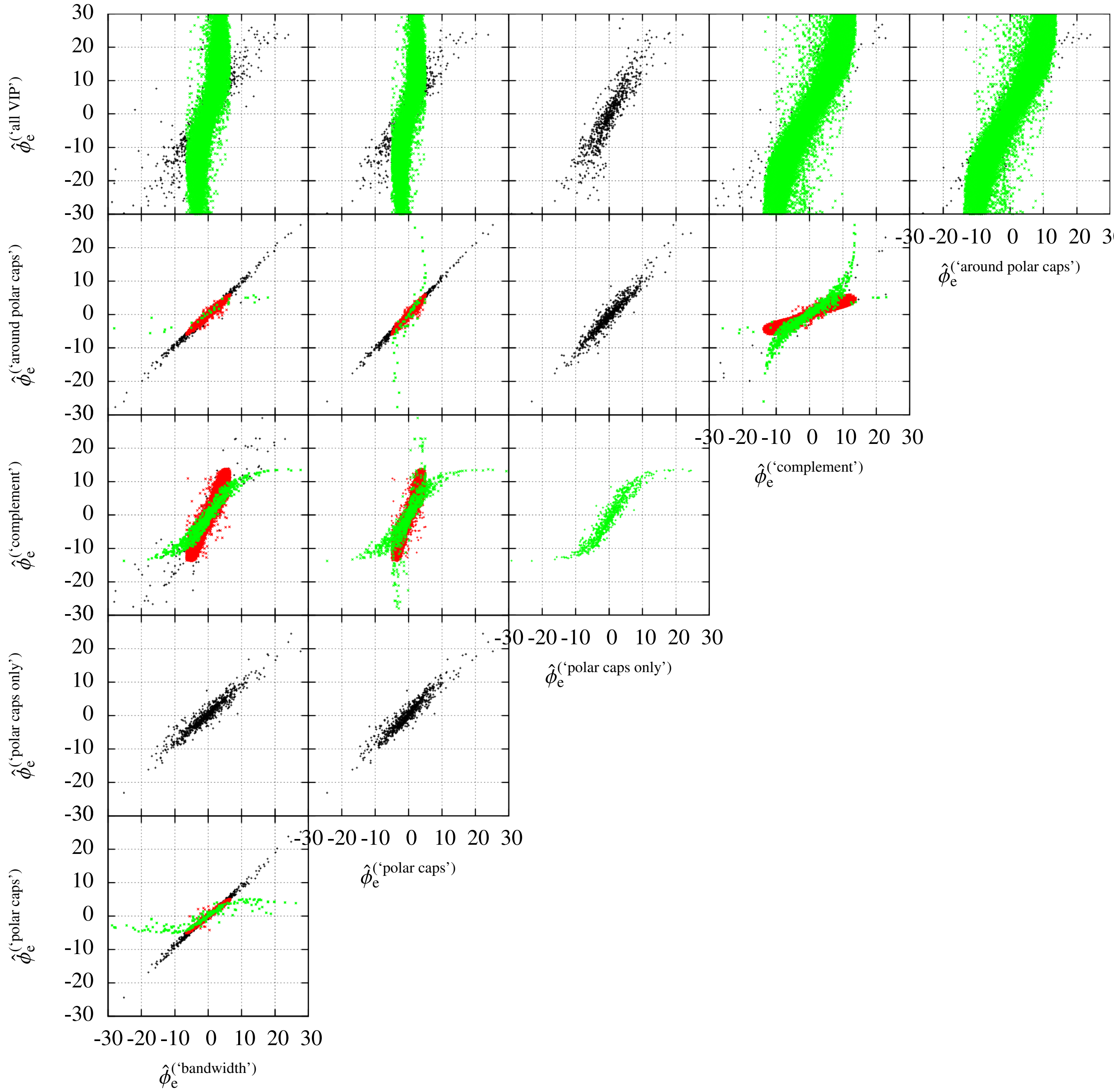

Fig. 15. Comparison of the estimates for the extragalactic contributions, $\hat{\phi}_e$, for the six different sets of assumptions described in the text. Black points are for data points that are in the VIP category under both assumptions that are being compared, green data points are in the SIP category under one of the assumptions and in the VIP category under the other, and red points are in the SIP category under both assumptions. The units of all axes are rad/m^2.

erly into account, will introduce artifacts in the extragalactic estimate.

In our considerations, we strictly made the distinction between a physical quantity and an estimate of this quantity. The latter aims to equal the former, but, even if calculated correctly, there is always uncertainty involved and artifacts in the estimate may result. Taking into account the uncertainty of the estimate, however, should remove the artifacts. An example of such an artifact is the latitude dependence that we observed in the estimates of the extragalactic contributions to the observed Faraday rotation values that we calculated. This latitude dependence vanishes once the uncertainty is taken into account.

To treat the complete problem of estimating the amount of both Galactic and extragalactic Faraday rotation from observations, we extended the algorithm of Oppermann et al. (2012). This extended algorithm is based on a split of the data into a subset for which the observational error bars describe the data likelihood sufficiently and another subset for which this is not the case. It includes the estimation of the angular power spectrum of the Galactic foreground, assumed to be statistically isotropic up to a single latitude-dependent modulation, the estimation of this latitude-dependent function, the estimation of corrected noise variances for the subset of the data for which this is deemed necessary, and the estimation of the variance of the extragalac-

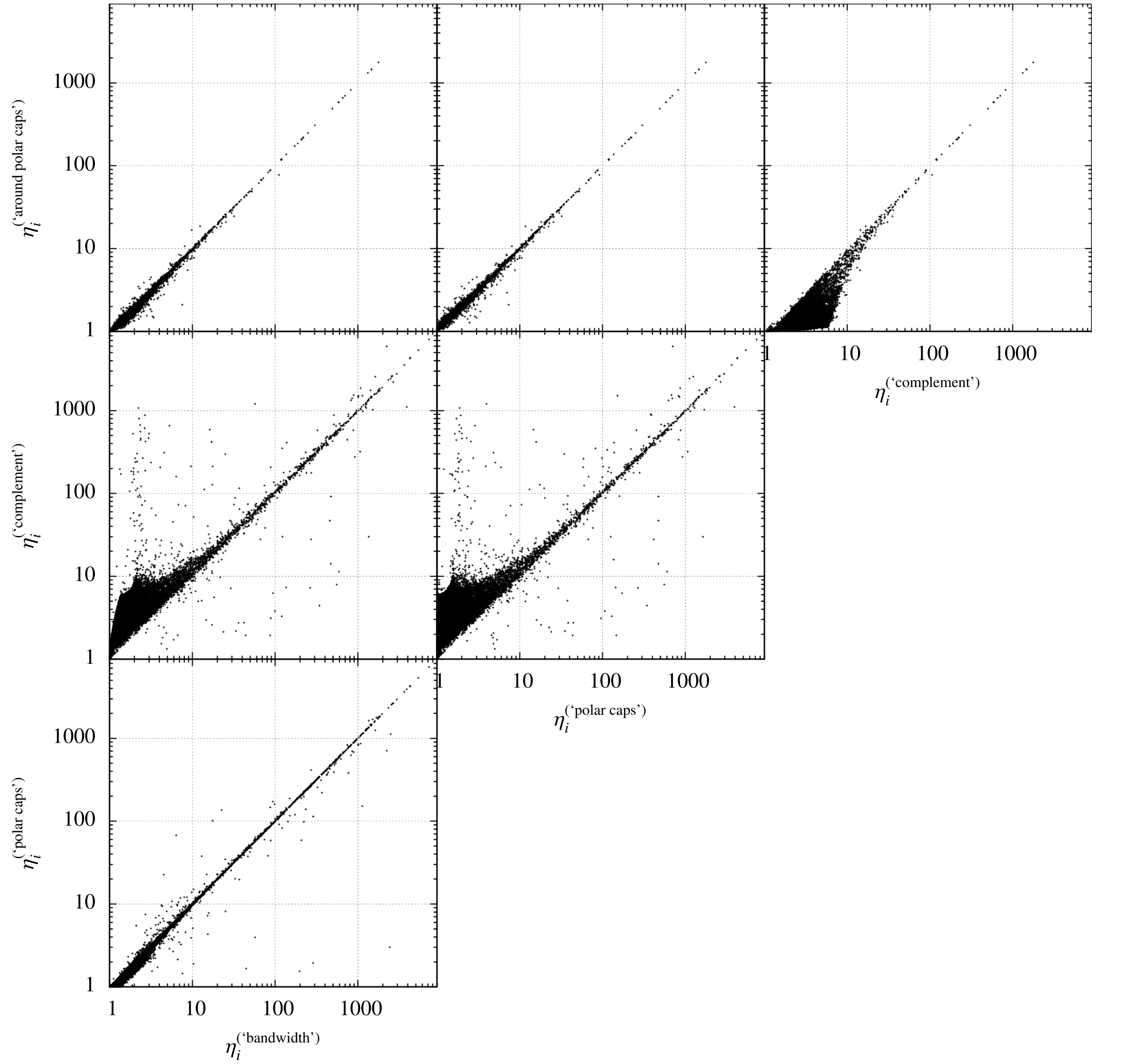

Fig. 16. Comparison of the error variance correction factors for the data of the SIP category under the four splits of the six under consideration that have data points of the SIP category.

tic contributions. We showed in a simulated scenario that all of these quantities are accurately reconstructed by our algorithm if our statistical model, including the split of the data, is correct.

For the application to observational data, we have considered several different ways to split the data into the two categories. We find that the most robust outcomes are achieved with splits that only regard a small fraction of the data (we use 1.75% of the data points) situated near the Galactic poles as not afflicted by potential problems in the description of the data likelihood. In these cases we find extragalactic dispersions between 6.0 rad/m^2 and 7.1 rad/m^2. These numbers agree remarkably with the ones derived by Schnitzeler (2010) by splitting the dispersion of observed Faraday rotation values into a latitude-dependent part and a constant offset, deemed to be extragalactic in origin. Strictly speaking, both analyses only produce upper limits on the dispersion of the extragalactic contributions, but for slightly different reasons. While the estimate of Schnitzeler (2010) may be increased due to a latitude-independent Galactic contribution, our estimate may be increased due to a Galactic contribution that is spatially uncorrelated on the scales probed by the observations.

We provide the derived estimates for all the involved quantities online at `http://www.mpa-garching.mpg.de/ift/faraday/`. The foreground products can be seen as updated versions of the results of Oppermann et al. (2012) that should be used preferentially, except in special cases where one of the assumptions we made in this paper is at question. We also provide 1 000 samples of extragalactic contributions to the observed Faraday rotation, drawn from the posterior PDF for this quan-

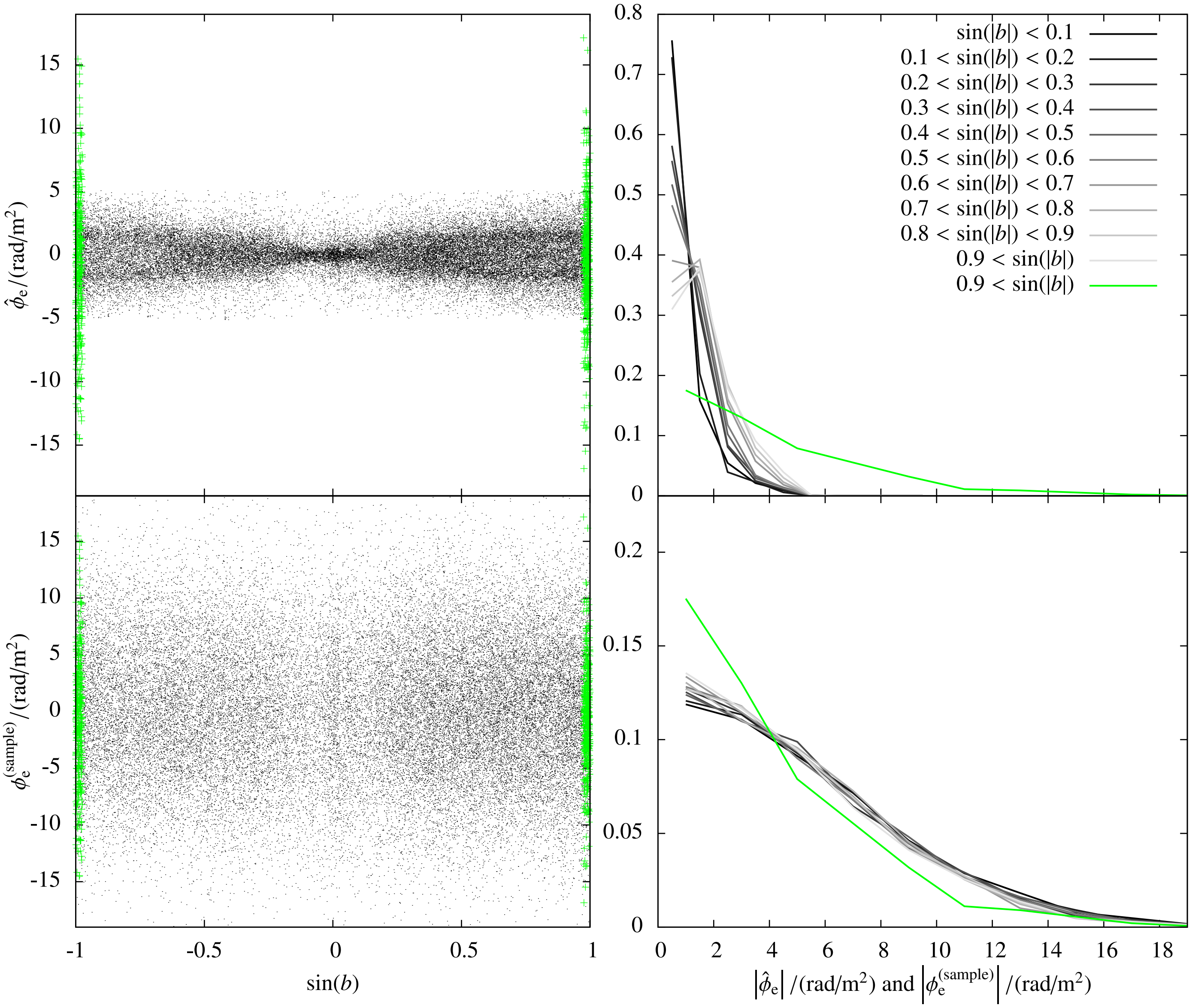

Fig. 17. *Top*: Estimates for the extragalactic contribution to each source's observed Faraday depth for the 'polar caps' analysis. The left panel shows the estimate versus Galactic latitude. The right panel shows normalized histograms of the distribution of these estimates in bins of $\sin(|b|)$. Each bin has a width of 0.1. *Bottom*: The same for a random sample drawn from the posterior PDF for the extragalactic contributions around the mean plotted in the upper panels. In all panels, data points of the SIP category, i.e., with noise variance correction factors, are plotted in black, data points of the VIP category, i.e., without noise variance correction factors, are plotted in green.

tity. This will enable future studies of extragalactic Faraday rotation to take into account the full probability distribution for these values, by performing any analysis on the set of samples rather than only on the posterior mean estimate. It should be noted that, within the framework of our assumptions, the extragalactic contributions are not very well constrained by the data. This is to some extent due to allowing the observational error bars of sources to get increased during the reconstruction, which increases the uncertainty of all reconstructed quantities. Also, sources for which such an increase of the error bar can happen in our reconstruction, will not have large estimates of the extragalactic contribution.

All of our considerations point toward the importance of understanding the uncertainties of observational Faraday rotation measurements. For future surveys, this means that not only should the largest possible interval in λ^2-space be covered, but, as already pointed out by Farnsworth et al. (2011) and Farnes et al. (2014), all the available information should be used in the data reduction, including the behavior of polarization fraction with frequency, as this can help avoid some of the rather poorly understood effects in RM synthesis studies that can lead to faulty estimates.

Acknowledgements. The authors would like to thank Shane P. O'Sullivan for providing his unpublished rotation measure synthesis data and S. Ann Mao for providing her newest data set in digital form. The results in this publication have been derived using the `NIFTy`[5] package (Selig et al. 2013), as well as the `HEALPix`[6] package (Górski et al. 2005). This research has made use of NASA's Astrophysics Data System. It was supported by the DFG Forschergruppe 1254 "Magnetisation of Interstellar and Intergalactic Media: The Prospects of Low-Frequency Radio Observations". T.A. is supported by the Japan Society for the Promotion of Science (JSPS). B.M.G. acknowledges the support of the Australian Research Council through grant FL100100114.

[5] `http://www.mpa-garching.mpg.de/ift/nifty/`
[6] `http://healpix.sf.net`

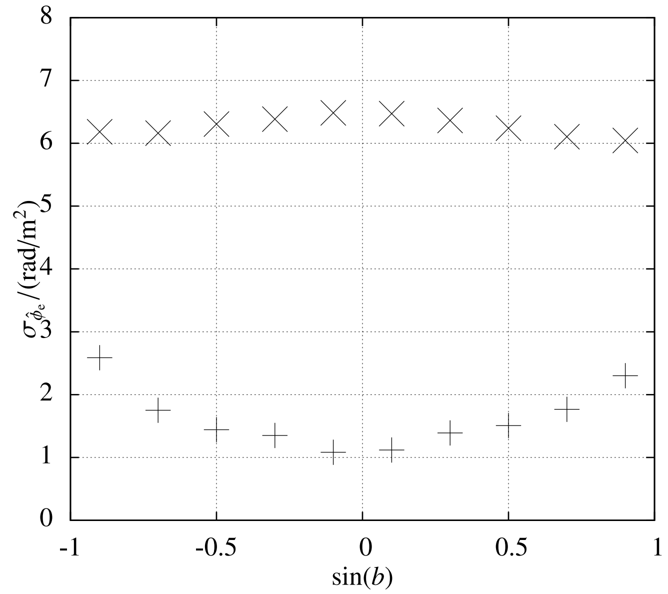

Fig. 18. Sample standard deviation of the estimated extragalactic contributions plotted in the top left panel of Fig. 17 (+-symbols), as well as for random samples drawn from the posterior PDF for the same quantity, such as plotted in the lower left panel of Fig. 17, averaged over 1 000 samples (×-symbols), in bins of sin(b). Each bin has a width of 0.2.

References

Akahori, T. & Ryu, D., Faraday Rotation Measure due to the Intergalactic Magnetic Field. II. The Cosmological Contribution, ApJ **738** (Sep. 2011) 134, arXiv:1107.0142 [astro-ph.CO]

Bernet, M. L., Miniati, F., & Lilly, S. J., The Interpretation of Rotation Measures in the Presence of Inhomogeneous Foreground Screens, ApJ **761** (Dec. 2012) 144, arXiv:1209.2410 [astro-ph.CO]

Blasi, P., Burles, S., & Olinto, A. V., Cosmological Magnetic Field Limits in an Inhomogeneous Universe, ApJ **514** (Apr. 1999) L79–L82, astro-ph/9812487

Brentjens, M. A. & de Bruyn, A. G., Faraday rotation measure synthesis, A&A **441** (Oct. 2005) 1217–1228, astro-ph/0507349

Burn, B. J., On the depolarization of discrete radio sources by Faraday dispersion, MNRAS **133** (1966) 67, ADS

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]

Farnes, J. S., Gaensler, B. M., & Carretti, E., A Broadband Polarization Catalog of Extragalactic Radio Sources, ArXiv e-prints (Mar. 2014), arXiv:1403.2391 [astro-ph.CO]

Farnsworth, D., Rudnick, L., & Brown, S., Integrated Polarization of Sources at λ ˜ 1 m and New Rotation Measure Ambiguities, AJ **141** (Jun. 2011) 191, arXiv:1103.4149 [astro-ph.CO]

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, astro-ph/0409513

Hammond, A. M., Robishaw, T., & Gaensler, B. M., A New Catalog of Faraday Rotation Measures and Redshifts for Extragalactic Radio Sources, ArXiv e-prints (Sep. 2012), arXiv:1209.1438v3 [astro-ph.CO]

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]

Jaynes, E. & Bretthorst, G. 2003, Probability Theory: The Logic of Science, , Cambridge University Press

Joshi, R. & Chand, H., Dependence of Residual Rotation Measure (RRM) on Intervening MgII Absorbers at Cosmic Distances, ArXiv e-prints (Jul. 2013), arXiv:1307.2678 [astro-ph.CO]

Kolatt, T., Determination of the Primordial Magnetic Field Power Spectrum by Faraday Rotation Correlations, ApJ **495** (Mar. 1998) 564, astro-ph/9704243

Kumazaki, K., Akahori, T., Ideguchi, S., Kurayama, T., & Takahashi, K., Intrinsic polarization angle ambiguity in Faraday tomography, ArXiv e-prints (Feb. 2014), arXiv:1402.0612 [astro-ph.IM]

Leahy, J. P., Small-scale variations in the Galactic Faraday rotation, MNRAS **226** (May 1987) 433–446, ADS

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., McClure-Griffiths, N. M., Gaensler, B. M., et al., New Constraints on the Galactic Halo Magnetic Field Using Rotation Measures of Extragalactic Sources toward the Outer Galaxy, ApJ **755** (Aug. 2012) 21, arXiv:1206.3314 [astro-ph.GA]

Neronov, A., Semikoz, D., & Banafsheh, M., Magnetic Fields in the Large Scale Structure from Faraday Rotation measurements, ArXiv e-prints (May 2013), arXiv:1305.1450 [astro-ph.CO]

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., 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, Phys. Rev. E **87** no. 3, (Mar. 2013) 032136, arXiv:1210.6866 [astro-ph.IM]

Robbins, H., An Empirical Bayes Approach to Statistics, Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability Volume 1: Contributions to the Theory of Statistics (1955) 157–163, http://projecteuclid.org/DPubS?service=UI&version=1.0&verb=Display&handle=euclid.bsmsp/1200501653

Schnitzeler, D. H. F. M., The latitude dependence of the rotation measures of NVSS sources, MNRAS **409** (Nov. 2010) L99–L103, arXiv:1011.0737 [astro-ph.GA]

Selig, M., Bell, M. R., Junklewitz, H., et al., NIFTY - Numerical Information Field Theory. A versatile PYTHON library for signal inference, A&A **554** (Jun. 2013) A26, arXiv:1301.4499 [astro-ph.IM]

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]

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

Taylor, A. R., Stil, J. M., & Sunstrum, C., A Rotation Measure Image of the Sky, ApJ **702** (Sep. 2009) 1230–1236

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]

Wardle, J. F. C. & Kronberg, P. P., The linear polarization of quasi-stellar radio sources at 3.71 and 11.1 centimeters, ApJ **194** (Dec. 1974) 249–255

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, astro-ph/0509826

Zaroubi, S., Hoffman, Y., Fisher, K. B., & Lahav, O., Wiener Reconstruction of the Large-Scale Structure, ApJ **449** (Aug. 1995) 446, arXiv:astro-ph/9410080

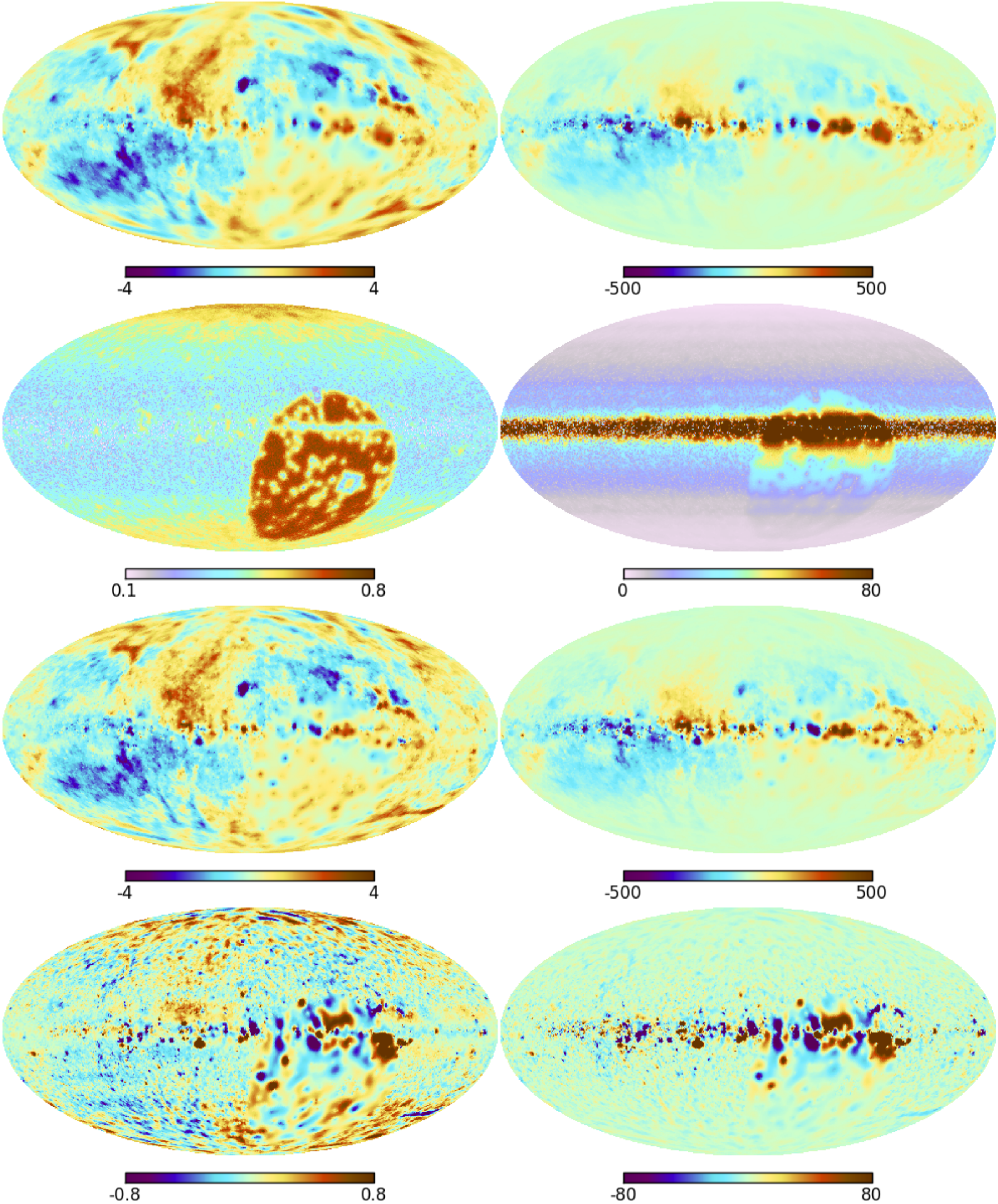

Fig. 19. Estimates of the Galactic contribution under the 'polar caps' data split and comparison to the results of Oppermann et al. (2012). The left column illustrates the dimensionless Galactic signal field s, the right column the physical Galactic Faraday depth ϕ_{g} in units of rad/m^2. The top row shows the posterior mean estimates derived with the 'polar caps' split, the second row shows the pixel-wise uncertainty of this estimate, the third row shows the result of Oppermann et al. (2012), and the bottom row shows the result of subtracting the third row from the top row.

Appendix A: Derivation of the filter formulas

Here we discuss the filter formulas we use, their derivation, and the necessary approximations, following the strategy outlined in Sect. 4.2. Throughout, we assume that the covariance matrices have the structure described in Sect. 4.1.

Appendix A.1: Estimating the Galactic contribution

In order to estimate the Galactic contribution to Faraday rotation, we first have to estimate the dimensionless signal field $s = \frac{\phi_\mathrm{e}}{p}$, for which we calculate the mean over the PDF

$$\mathcal{P}\left(s \,\middle|\, d, (\tilde{C}_\ell)_\ell = (\hat{\tilde{C}}_\ell)_\ell, (\tilde{\eta}_i)_i = (\hat{\tilde{\eta}}_i)_i, \tilde{\eta}_\mathrm{e} = \hat{\tilde{\eta}}_\mathrm{e}\right). \tag{A.1}$$

Using the zero-mean Gaussian priors for this signal field, the extragalactic contributions, as well as the noise contribution with covariances S, E, and N, respectively, this PDF is again a Gaussian with covariance

$$D = \left(S^{-1} + R^\dagger \left(N + E\right)^{-1} R\right)^{-1} \tag{A.2}$$

and mean

$$m = DR^\dagger \left(N + E\right)^{-1} d. \tag{A.3}$$

This m therefore becomes our estimate for the dimensionless signal field and the diagonal of the matrix D a measure for its pixel-wise uncertainty. The corresponding estimate for the Galactic contribution is obtained simply by multiplying with the Galactic latitude profile,

$$\hat{\phi}_\mathrm{g} = pm, \tag{A.4}$$

and its uncertainty accordingly as

$$\mathrm{diag}\left(D_{\phi_\mathrm{g}}\right) = p^2 \mathrm{diag}(D). \tag{A.5}$$

Of course, the operators D and $(N + E)$ depend on our estimates of the unknown quantities $(C_\ell)_\ell$, $(\eta_i)_i$, and η_e and necessitate that we estimate these in separate steps.

Appendix A.2: Estimating the extragalactic contribution

For the extragalactic contribution, we repeat the analysis done for the Galactic contribution and simply swap the roles of the Galactic and extragalactic contributions. We therefore find again a Gaussian posterior

$$\mathcal{P}\left(\phi_\mathrm{e} \,\middle|\, d, (\tilde{C}_\ell)_\ell = (\hat{\tilde{C}}_\ell)_\ell, (\tilde{\eta}_i)_i = (\hat{\tilde{\eta}}_i)_i, \tilde{\eta}_\mathrm{e} = \hat{\tilde{\eta}}_\mathrm{e}\right) = \mathcal{G}\left(\phi_\mathrm{e} - \hat{\phi}_\mathrm{e} \,\middle|\, D_\mathrm{e}\right), \tag{A.6}$$

where the covariance is given by

$$D_\mathrm{e} = \left(E^{-1} + \left(RSR^\dagger + N\right)^{-1}\right)^{-1} \tag{A.7}$$

and the mean and our estimate by

$$\hat{\phi}_\mathrm{e} = D_\mathrm{e} \left(RSR^\dagger + N\right)^{-1} d = E \left(E + N\right)^{-1} \left(d - \hat{\phi}_\mathrm{g}\right). \tag{A.8}$$

Appendix A.3: Estimating the angular power spectrum of the dimensionless auxiliary field

To estimate the angular power spectrum, we maximize the PDF

$$\mathcal{P}\left((\tilde{C}_\ell)_\ell \,\middle|\, d, (\tilde{\eta}_i)_i = (\hat{\tilde{\eta}}_i)_i, \eta_\mathrm{e} = \hat{\eta}_\mathrm{e}\right), \tag{A.9}$$

where a tilde denotes a logarithmic quantity, i.e., $\tilde{C}_\ell = \log\left(C_\ell\right)$. This PDF is calculated straightforwardly by multiplying the Gaussian likelihood function

$$\mathcal{P}\left(d \,\middle|\, s, (\tilde{C}_\ell)_\ell, (\tilde{\eta}_i)_i = (\hat{\tilde{\eta}}_i)_i, \eta_\mathrm{e} = \hat{\eta}_\mathrm{e}\right) = \mathcal{G}(d - Rs, E + N) \tag{A.10}$$

with the Gaussian signal prior and the prior for the angular power spectrum, given by Eqs. (34)-(36), and marginalizing over s. The result is

$$\mathcal{P}\left((\tilde{C}_\ell)_\ell \,\middle|\, d, (\tilde{\eta}_i)_i = (\hat{\tilde{\eta}}_i)_i, \eta_\mathrm{e} = \hat{\eta}_\mathrm{e}\right) \propto |S|^{-1/2} \, |D|^{1/2} \left(\prod_\ell C_\ell^{-\alpha_\ell + 1} \mathrm{e}^{-\frac{q_\ell}{C_\ell}}\right) \exp\left\{-\frac{1}{2}\tilde{C}^\dagger T \tilde{C} + \frac{1}{2} d^\dagger \left(E + N\right)^{-1} RDR^\dagger \left(E + N\right)^{-1} d\right\}, \tag{A.11}$$

where we have dropped all factors that are independent of the angular power spectrum. Equating the derivative of this function with respect to $\tilde{C}_\ell$ with zero leads to the equation for our estimate of the angular power spectrum,

$$\hat{C}_\ell = \frac{q_\ell + \frac{1}{2}\mathrm{tr}\Big(\big(mm^\dagger + D\big)\, S_{(\ell)}\Big)}{\rho_\ell/2 + \alpha_\ell - 1 + \Big(T\hat{\hat{C}}\Big)_\ell}. \tag{A.12}$$

Here, $S_{(\ell)}$ denotes an operator that projects a field on the sphere onto its ℓ-th multipole and $\rho_\ell = 2\ell + 1$ is the number of degrees of freedom of the ℓ-th multipole.

Appendix A.4: Estimating the noise variance correction factors

Similar to the estimation of the angular power spectrum, we multiply the likelihood

$$\mathcal{P}\Big(d\,\Big|\,s, (\tilde{\eta}_i)_i, (\tilde{C}_\ell)_\ell = (\hat{\tilde{C}}_\ell)_\ell, \eta_\mathrm{e} = \hat{\eta}_\mathrm{e}\Big) = \mathcal{G}(d - Rs, N + E) \tag{A.13}$$

with the signal prior and the prior for the noise variance correction factors, given by Eq. (40), and marginalize over s, resulting in

$$\mathcal{P}\Big((\tilde{\eta}_i)_i |\, d, (\tilde{C}_\ell)_\ell = (\hat{\tilde{C}}_\ell)_\ell, \eta_\mathrm{e} = \hat{\eta}_\mathrm{e}\Big) \propto |(E+N)|^{-1/2}\, |D|^{1/2} \left(\prod_i \eta_i^{-\beta_i+1} \exp\left(-\frac{r_i}{\eta_i}\right)\right)$$
$$\times \exp\left\{\frac{1}{2} d^\dagger\,(E+N)^{-1}\, RDR^\dagger\,(E+N)^{-1}\, d - \frac{1}{2} d^\dagger\,(E+N)^{-1}\, d\right\}. \tag{A.14}$$

We have again dropped all factors that are independent of η_i. After differentiating with respect to $\tilde{\eta}_i$ and equating to zero we find our estimate

$$\hat{\eta}_i = \frac{r_i + \frac{1}{2}\left((d - Rm)_i^2 + \left(RDR^\dagger\right)_{ii}\right)}{\beta_i - 1/2}. \tag{A.15}$$

Appendix A.5: Estimating the extragalactic variance correction factor

The calculation for the extragalactic variance correction factor is slightly more involved than the ones for the other estimators. We begin again by multiplying the likelihood

$$\mathcal{P}\Big(d\,\Big|\,s, (\tilde{\eta}_i)_i = (\hat{\tilde{\eta}}_i)_i, (\tilde{C}_\ell)_\ell = (\hat{\tilde{C}}_\ell)_\ell, \eta_\mathrm{e}\Big) \tag{A.16}$$

with the priors for the signal s and for the variance correction factors, Eq. (40). However, now we marginalize first over the error variance correction factors $(\eta_i)_i$. This leads to

$$\mathcal{P}\Big(s, \log(\eta_\mathrm{e})\,\Big|\,d, (\tilde{C}_\ell)_\ell = (\hat{\tilde{C}}_\ell)_\ell\Big) \propto \left(\prod_{j\in(\mathrm{VIP})} \left(\eta_\mathrm{e}\sigma_\mathrm{e}^2 + \sigma_j^2\right)^{-1/2}\right) \eta_\mathrm{e}^{-\beta_\mathrm{e}+1} \exp\left(-\frac{r_\mathrm{e}}{\eta_\mathrm{e}}\right) \left(\prod_{i\in(\mathrm{SIP})} \left(r_i + \frac{1}{2}\frac{(d-Rs)_i^2}{\sigma_i^2 + \sigma_\mathrm{e}^2}\right)^{1/2-\beta_i}\right)$$
$$\times \exp\left\{-\frac{1}{2}\,(d - Rs)^\dagger\,(E+N)^{-1}_{(\mathrm{VIP})}\,(d - Rs)\right\} \mathcal{G}(s, S), \tag{A.17}$$

where the first product is to be taken over all data points in the VIP category and the second product over all data points in the SIP category. $(E + N)_{(\mathrm{VIP})}$ denotes the combination of the diagonal operator $(E + N)$ and the projection onto the data points of the VIP category. Here and in the following, we use the superscript -1 to denote the inverse for regular operators and the pseudo-inverse for singular operators, such as $(E + N)_{(\mathrm{VIP})}$.

Marginalizing this PDF over the dimensionless signal field s amounts to calculating the Gaussian integral

$$\int \mathcal{D}s\, \left(\prod_{i\in(\mathrm{SIP})} \left(r_i + \frac{1}{2}\frac{(d-Rs)_i^2}{\sigma_i^2 + \sigma_\mathrm{e}^2}\right)^{1/2-\beta_i}\right) \mathcal{G}(s - m_{(\mathrm{VIP})}, D_{(\mathrm{VIP})}), \tag{A.18}$$

where we have defined

$$D_{(\mathrm{VIP})} = \left(S^{-1} + R^\dagger\,(E+N)^{-1}_{(\mathrm{VIP})}\,R\right)^{-1} \tag{A.19}$$

and

$$m_{(\mathrm{VIP})} = D_{(\mathrm{VIP})} R^\dagger\,(E+N)^{-1}_{(\mathrm{VIP})}\, d. \tag{A.20}$$

These would be the posterior covariance and mean for the dimensionless signal field if only the data points of the VIP category existed. Calculating this integral analytically for a positive value of β_i is not possible. We therefore Taylor-expand the product in $\left(\frac{(d-Rs)_i^2}{\sigma_i^2+\sigma_\mathrm{e}^2}\right)_i$ up to first order around its expectation value, given by

$$\left\langle \frac{(d-Rs)_i^2}{\sigma_i^2+\sigma_\mathrm{e}^2} \right\rangle_{\mathcal{G}(s-m_{(\mathrm{VIP})},D_{(\mathrm{VIP})})} = \frac{(d-Rm_{(\mathrm{VIP})})_i^2 + \left(RD_{(\mathrm{VIP})}R^\dagger\right)_{ii}}{\sigma_i^2+\sigma_\mathrm{e}^2}. \tag{A.21}$$

After the integration, the first order expansion term vanishes by definition and we are left with the zero-order term. Altogether, the PDF we are maximizing becomes

$$\begin{aligned}\mathcal{P}\left(\log(\eta_\mathrm{e}) \middle| d, (\tilde{C}_\ell)_\ell = (\hat{\tilde{C}}_\ell)_\ell\right) \propto \left|D_{(\mathrm{VIP})}\right|^{1/2} \left(\prod_{j\in(\mathrm{VIP})} \left(\eta_\mathrm{e}\sigma_\mathrm{e}^2+\sigma_j^2\right)^{-1/2}\right) \eta_\mathrm{e}^{-\beta_\mathrm{e}+1} \mathrm{e}^{-\frac{r_\mathrm{e}}{\eta_\mathrm{e}}} \left(\prod_{i\in(\mathrm{SIP})} \left(r_i + \frac{1}{2}\frac{(d-Rm_{(\mathrm{VIP})})_i^2 + \left(RD_{(\mathrm{VIP})}R^\dagger\right)_{ii}}{\sigma_i^2+\sigma_\mathrm{e}^2}\right)^{1/2-\beta_i}\right) \\ \times \exp\left\{-\frac{1}{2}d^\dagger\,(E+N)^{-1}_{(\mathrm{VIP})}\,d + \frac{1}{2}d^\dagger\,(E+N)^{-1}_{(\mathrm{VIP})}\,RD_{(\mathrm{VIP})}R^\dagger\,(E+N)^{-1}_{(\mathrm{VIP})}\,d\right\}.\end{aligned} \tag{A.22}$$

The value of η_e that maximizes this function fulfills

$$\hat{\eta}_\mathrm{e} = \frac{A+B}{C}, \tag{A.23}$$

where

$$A = r_\mathrm{e} + \frac{\eta_\mathrm{e}^2\sigma_\mathrm{e}^2}{2}\left(\sum_{j\in(\mathrm{VIP})} \frac{(d-Rm_{(\mathrm{VIP})})_j^2 + \left(RD_{(\mathrm{VIP})}R^\dagger\right)_{jj}}{\left(\eta_\mathrm{e}\sigma_\mathrm{e}^2+\sigma_j^2\right)^2}\right), \tag{A.24}$$

$$B = \sum_{i\in(\mathrm{SIP})} \frac{\eta_\mathrm{e}^2\sigma_\mathrm{e}^2\left(\beta_i-\frac{1}{2}\right)}{2} \frac{2\,(d-Rm_{(\mathrm{VIP})})_i \left(RD_{(\mathrm{VIP})}R^\dagger\,(E+N)^{-2}_{(\mathrm{VIP})}\,(d-Rm_{(\mathrm{VIP})})\right)_i - \left(RD_{(\mathrm{VIP})}R^\dagger\,(E+N)^{-2}_{(\mathrm{VIP})}\,RD_{(\mathrm{VIP})}R^\dagger\right)_{ii}}{r_i\left(\sigma_i^2+\sigma_\mathrm{e}^2\right) + \frac{1}{2}\,(d-Rm_{(\mathrm{VIP})})_i^2 + \frac{1}{2}\,(RD_{(\mathrm{VIP})}R^\dagger)_{ii}}, \tag{A.25}$$

and

$$C = \beta_\mathrm{e} - 1 + \frac{1}{2}\left(\sum_{j\in(\mathrm{VIP})} \frac{\eta_\mathrm{e}\sigma_\mathrm{e}^2}{\eta_\mathrm{e}\sigma_\mathrm{e}^2+\sigma_j^2}\right). \tag{A.26}$$

Appendix A.6: Implementation

We start our reconstruction with a starting guess for the latitude-dependent profile function and the angular power spectrum. The initial profile is calculated directly from the data as described in Sect. 4.2. For the angular power spectrum, we choose as a starting guess a simple power law,

$$C_\ell = 1.53\,\ell^{-2.17}, \tag{A.27}$$

based on the results of Oppermann et al. (2012). We then iterate the following steps until convergence:

- Calculate a new estimate of the Galactic Faraday depth according to Eq. (A.3).
- As an auxiliary field, calculate the estimate of the Galactic Faraday depth using only the data points of the VIP category, $m_{(\mathrm{VIP})}$, according to Eq. (A.20).
- Use these current estimates to update the estimates for the error variance correction factors $(\eta_i)_i$ and the correction factor for the extragalactic variance η_e according to Eqs. (A.15) and (A.23).
- Use these to update the estimate of the angular power spectrum according to Eq. (A.12).

After this iteration has converged, we calculate a new profile function as described in Sect. 4.2 and repeat the whole procedure until the profile function has converged as well. At any point, the estimate for the extragalactic contributions can be calculated via Eq. (A.8).

Due to the high dimensionality of the involved vector spaces (41 330 data points in the simulation, 41 632 observational data points, 196 608 pixels in our maps), we avoid treating the involved operators as explicit matrices. Operator inversions are performed with a conjugate gradient routine and whenever diagonal elements are needed explicitly, these are estimated via the technique of operator probing as implemented in the `NIFTy` package (Selig et al. 2013). These methods yield only approximate solutions that can lead to artifacts in the results, such as the increased posterior variance discussed in Appendix C.2.1.

Appendix B: Effects of excluding data points from the estimation of the Galactic foreground

Here, we discuss the effects that the usage of a data point in simultaneously estimating its Galactic foreground contribution *and* its extragalactic contribution has vis-à-vis excluding a data point from the foreground estimation when estimating its extragalactic contribution. To focus on the difference in perspective that could lead to either of these two procedures, we discuss a simplistic foreground reconstruction algorithm that merely consists of averaging observed data from different sources.

Assume that we have $N + 1$ data points of the form

$$d_i = \phi_\mathrm{g} + \phi_{\mathrm{e},i}. \tag{B.1}$$

Here, the Galactic contribution to the data point, ϕ_g, is assumed to be a fixed constant (one may think of data points that are very close together) and the extragalactic contribution $\phi_{\mathrm{e},i}$ is different for each data point. We are assuming noiseless observations.

Assume further that the extragalactic contributions to the different data points are uncorrelated, i.e.,

$$\left\langle \phi_{\mathrm{e},i}\, \phi_{\mathrm{e},j} \right\rangle_{(\phi_\mathrm{e})} = \delta_{ij}\, \sigma_\mathrm{e}^2, \tag{B.2}$$

where σ_e^2 is the extragalactic variance, also assumed to be the same for each source.

Now assume that we are interested specifically in the extragalactic contribution to the 0-th data point, $\phi_{\mathrm{e},0}$. There are two ways of simply averaging the data to build an estimator for this quantity; one can either subtract an average of all data points,

$$\hat{\phi}^{(\mathrm{A})}_{\mathrm{e},0} = d_0 - \frac{1}{N+1} \sum_{i=0}^{N} d_i; \tag{B.3}$$

we will call this estimator A, or one can exclude the 0-th data point,

$$\hat{\phi}^{(\mathrm{B})}_{\mathrm{e},0} = d_0 - \frac{1}{N} \sum_{i=1}^{N} d_i; \tag{B.4}$$

we will call this estimator B.

Now one may ask, which is better. The answer depends on what exactly is meant with ‘better’. One way to judge the quality of an estimator is to look at the expectation value for its squared error,

$$\left\langle \left(\hat{\phi}_{\mathrm{e},0} - \phi_{\mathrm{e},0} \right)^2 \right\rangle_{(\phi_\mathrm{e})}, \tag{B.5}$$

where we take the expectation value with respect to all $\phi_{\mathrm{e},i}$ contributions (remember that we have simplified things by fixing the Galactic contribution and neglecting noise).

Plugging in the formulas for the estimators and using the fact that the expectation value for $\phi_{\mathrm{e},i}$ is zero (independent of i), as well as the second moment, given by Eq. (B.2), we can straightforwardly calculate

$$\left\langle \left(\hat{\phi}^{(\mathrm{A})}_{\mathrm{e},0} - \phi_{\mathrm{e},0} \right)^2 \right\rangle_{(e)} = \frac{\sigma_\mathrm{e}^2}{N+1} \tag{B.6}$$

and

$$\left\langle \left(\hat{\phi}^{(\mathrm{B})}_{\mathrm{e},0} - \phi_{\mathrm{e},0} \right)^2 \right\rangle_{(e)} = \frac{\sigma_\mathrm{e}^2}{N}. \tag{B.7}$$

Thus, one would conclude that estimator A is superior. This argument follows the Bayesian logic of calculating expectation values over all unknown quantities that are involved.

However, one might be inclined to take the expectation value only with respect to $\phi_{\mathrm{e},i}$ for $i = 1, \ldots, N$ and regard $\phi_{\mathrm{e},0}$ as a fixed value. This would allude to the frequentist way of doing statistics, of regarding whichever quantity one is interested in as fixed and only averaging over other quantities, regarded as noise. In this case the expectation value of $\phi_{\mathrm{e},0}$ is of course no longer zero and the expected squared errors become

$$\left\langle \left(\hat{\phi}^{(\mathrm{A})}_{\mathrm{e},0} - \phi_{\mathrm{e},0} \right)^2 \right\rangle_{(\phi_{\mathrm{e},i=1,\ldots,N})} = \frac{N}{(N+1)^2}\sigma_\mathrm{e}^2 + \frac{1}{(N+1)^2}\phi_{\mathrm{e},0}^2 \tag{B.8}$$

and

$$\left\langle \left(\hat{\phi}^{(\mathrm{B})}_{\mathrm{e},0} - \phi_{\mathrm{e},0} \right)^2 \right\rangle_{(\phi_{\mathrm{e},i=1,\ldots,N})} = \frac{\sigma_\mathrm{e}^2}{N}. \tag{B.9}$$

Thus, one would conclude that estimator B can be superior for small N. The term depending on $\phi_{\mathrm{e},0}$ makes estimator A slightly unintuitive. However, the idea to treat $\phi_{\mathrm{e},0}$ different from the other extragalactic contributions when calculating expectation values may seem unintuitive as well.

Another interesting point is to look at the expectation values of the estimates themselves. If the expectation value is taken by averaging over all extragalactic contributions, it is zero for both estimators,

$$\left\langle \hat{\phi}^{(\mathrm{A})}_{\mathrm{e},0} \right\rangle_{(\phi_\mathrm{e})} = 0, \tag{B.10}$$

$$\left\langle \hat{\phi}^{(\mathrm{B})}_{\mathrm{e},0} \right\rangle_{(\phi_\mathrm{e})} = 0, \tag{B.11}$$

as one might expect. When calculating the expectation value only with respect to the extragalactic contributions to data points 1 to N, the result changes to

$$\left\langle \hat{\phi}^{(\mathrm{A})}_{\mathrm{e},0} \right\rangle_{(\phi_{\mathrm{e},i=1,\ldots,N})} = \frac{N}{N+1}\phi_{\mathrm{e},0}, \tag{B.12}$$

$$\left\langle \hat{\phi}^{(\mathrm{B})}_{\mathrm{e},0} \right\rangle_{(\phi_{\mathrm{e},i=1,\ldots,N})} = \phi_{\mathrm{e},0}, \tag{B.13}$$

which is another intuitive reason for estimator B.

In conclusion, excluding the 0-th data point from the subtracted average only seems superior because one implicitly replaces the average over the other data points with an ensemble average without taking the next step of extending the ensemble average also to the extragalactic contribution of the data point in question. If one does this, including the 0-th data point in the subtracted average slightly improves the resulting estimator.

Appendix C: Online access and usage of the results

At `http://www.mpa-garching.mpg.de/ift/faraday/` we provide the results of our study in the ‘polar caps’ data split described in Sect. 5. All results are provided in binary format both as `hdf5` files and `fits` files.

Appendix C.1: Foreground products

We provide one file containing the results of the reconstruction for the Galactic foreground. It contains the angular power spectrum as plotted in Fig. 13, the Galactic variance profile as plotted in Fig. 14, maps of the reconstructed auxiliary field, m, and of the reconstructed Galactic Faraday depth, $\hat{\phi}_{\mathrm{g}}$, as well as of the pixel-wise uncertainties of these maps as plotted in the top two rows of Fig. 19. The last five quantities are stored as `HEALPix` maps in `RING` ordering scheme at a resolution of $N_{\mathrm{side}} = 128$. The variance profile is stored as a map for easier use. The angular power spectrum is a simple list of 384 numbers corresponding to the values of C_ℓ for $\ell = 0, \ldots, 383$.

Appendix C.2: Samples for the extragalactic contributions

For the extragalactic contributions, we provide 1 000 samples drawn from the Gaussian approximation to their posterior probability distribution, given by Eq. (29). For testing purposes, we also provide a smaller file containing only the first 100 samples. Each data set in the `hd5f` files and each table column in the `fits` files contains 41 632 entries, corresponding to the 41 632 data points used in Sect. 5.

Table C.2 shows the first few rows of the file containing the samples. The first seven columns in the table, corresponding to a group of data sets within the `hdf5` file and a binary table extension in the `fits` file, give information about the sources. The first column gives the Faraday rotation catalog from which the data point is taken. Here, the notation of Table 1 of Oppermann et al. (2012) is used, except for the newly added data from Mao et al. (2012), for which the identifier 'Mao 2012' is used. The second and third columns give the Galactic longitude and latitude of the sources, respectively. Note that the large number of decimal places in this column is due to a coordinate conversion from the original catalogs. The fourth and fifth columns give the observed value of Faraday rotation and its error bar, for sources for which this information is published. The error bars of the NVSS rotation measure catalog have been multiplied by a factor 1.22 as discussed in Sect. 5.1. The sixth and seventh columns provide the estimated Galactic contribution and its uncertainty at the source's location.

Further 1 000 columns are stored in a second group of data sets in the `hdf5` file and a second and third binary table extension in the `fits` file. These contain the samples drawn from the posterior probability distribution for the extragalactic contributions. Each row in this table corresponds to one data point and each column to a possible configuration of the extragalactic contributions.

The second-to-last and last columns in the table, corresponding to a final binary table extension in the `fits` file and a final group of data sets in the `hdf5` file, give a summary of the posterior PDF by providing the mean and standard deviation of the sample values.

Appendix C.2.1: How to use the posterior samples

A range of values that is less likely appears less often in the samples and vice versa. Thus, the frequency with which the sample values lie within a certain interval gives the posterior probability for the true extragalactic contribution to lie within that interval.

When calculating a quantity as a function of the extragalactic Faraday contribution for one or several sources,

$$f(\phi_{\mathrm{e},1}, \phi_{\mathrm{e},2}, \ldots, \phi_{\mathrm{e},41\,632}), \tag{C.1}$$

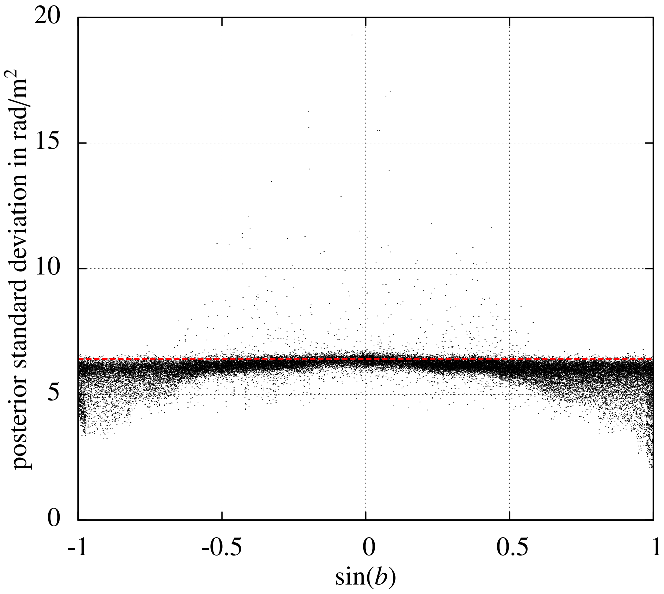

Fig. C.1. Standard deviation of the posterior for the extragalactic contribution to each data point versus $\sin(b)$. Plotted is the estimate calculated from the samples as the square root of Eq. (C.6). The horizontal dashed line shows the prior standard deviation of $\sigma_{\mathrm{e}} = 6.4\,\mathrm{rad/m^2}$.

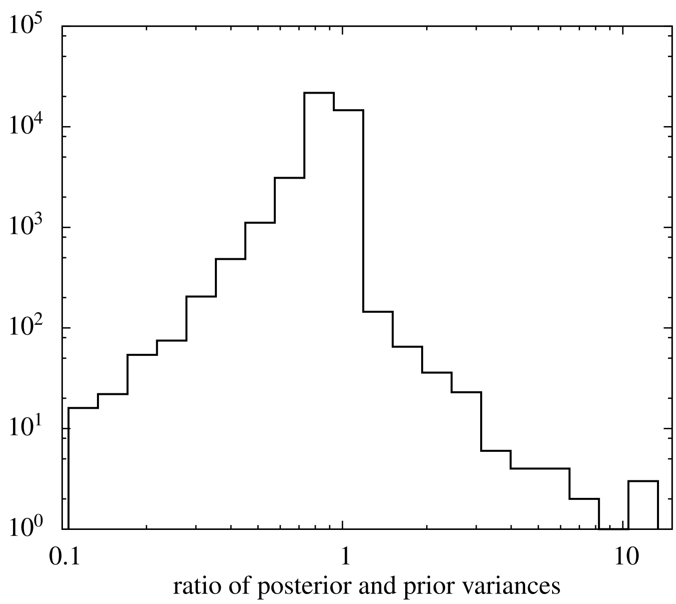

Fig. C.2. Histogram of the ratios of the posterior variance for the extragalactic contribution to each data point, as calculated from the samples via Eq. (C.6), to the prior variance. Note the logarithmic scales on both axes.

this function should be evaluated for each of the samples. This will yield 1 000 different answers,

$$f^{(k)} = f\Big(\phi^{(k)}_{\mathrm{e},1}, \phi^{(k)}_{\mathrm{e},2}, \ldots, \phi^{(k)}_{\mathrm{e},41\,632}\Big), \quad k = 0, \ldots, 999, \tag{C.2}$$

where $\phi^{(k)}_{\mathrm{e},i}$ is the value for the extragalactic contribution to the i-th source according to the k-th sample. In the limit of infinitely many samples, the distribution of these answers gives the posterior distribution for the quantity of interest f, given the data and assumptions that we have used and the approximations that we have made. In practice, a finite number of samples has to be used. The more samples are used, the more accurate the resulting distribution.

Finally the probability density for f approximated thusly can again be summarized, e.g., by calculating its mean

$$\langle f \rangle_{(f|d)} \approx \frac{1}{1000} \sum_{k=0}^{999} f^{(k)} \tag{C.3}$$

and its (co)variance

$$\left\langle \left(f - \langle f \rangle_{(f|d)}\right)\left(f - \langle f \rangle_{(f|d)}\right)^{\dagger} \right\rangle_{(f|d)}$$
$$\approx \frac{1}{1000} \sum_{k=0}^{999} \left(f^{(k)} - \frac{1}{1000} \sum_{k'=0}^{999} f^{(k')} \right)\left(f^{(k)} - \frac{1}{1000} \sum_{k'=0}^{999} f^{(k')} \right)^{\dagger}. \tag{C.4}$$

These formulas are equally true for scalar functions f and vector-valued functions f.

Thus, we can for example calculate the posterior mean for the extragalactic contribution to the i-th data point as

$$\langle \phi_{\mathrm{e},i} \rangle_{(\phi_{\mathrm{e},i}|d)} \approx \frac{1}{1000} \sum_{k=0}^{999} \phi_{\mathrm{e},i}^{(k)} \tag{C.5}$$

and the posterior variance for the i-th data point as

$$\left\langle \left(\phi_{\mathrm{e},i} - \langle \phi_{\mathrm{e},i} \rangle_{(\phi_{\mathrm{e},i}|d)}\right)^2 \right\rangle_{(\phi_{\mathrm{e},i}|d)} \approx \frac{1}{1000} \sum_{k=0}^{999} \left(\phi_{\mathrm{e},i}^{(k)} - \frac{1}{1000} \sum_{k'=0}^{999} \phi_{\mathrm{e},i}^{(k')} \right)^2. \tag{C.6}$$

In the last two columns of the provided files, we give this mean and the standard deviation, i.e., the square root of the last expression.

The posterior mean is also plotted in the top panel of Fig. 17 and we show the posterior standard deviations in Fig. C.1. Here, a word of warning is warranted. The approximate posterior we use in the calculation of the posterior mean estimate and in the drawing of the samples corresponds to a Gaussian posterior after fixing the prior covariances for the Galactic and extragalactic contributions and for the noise, i.e., to Eq. (29). Consequently, the uncertainty due to the uncertain reconstruction of the angular power spectrum $(C_\ell)_\ell$, the error variance correction factors $(\eta_i)_i$, and the correction factor for the extragalactic variance η_e, is no longer represented by this PDF. One logical consequence is that the posterior standard deviations for the extragalactic contributions, which give a measure of our uncertainty after considering the data, should in every case be smaller than the corresponding prior standard deviation, which we have reconstructed to be $\sigma_\mathrm{e} = 6.4\,\mathrm{rad/m^2}$ in the 'polar caps' split that is used here. From Fig. C.1 it is clear that this is not the case for a few data points at low and intermediate Galactic latitude. This effect must be the result of numerical inaccuracies, likely brought about by the usage of approximate iterative schemes for matrix inversion in the sampling procedure. We provide the posterior standard deviations so that it becomes easy to remove the sources that are afflicted by this problem in any further analysis.

Apart from this issue, the posterior standard deviations plotted in Fig. C.1 behave as expected, being in general slightly lower than the prior standard deviation, and more so nearer to the poles, where the sensitivity to the extragalactic contributions is largest.

In Fig. C.2, we plot a histogram of the ratio of the posterior variance as estimated from the samples and the prior variance,

$$\left(\left\langle \left(\phi_{\mathrm{e},i} - \langle \phi_{\mathrm{e},i} \rangle_{(\phi_{\mathrm{e},i}|d)}\right)^2 \right\rangle_{(\phi_{\mathrm{e},i}|d)} \right) / \sigma_\mathrm{e}^2. \tag{C.7}$$

Table C.1. Specifications of the four sources for which sample values are plotted in Fig. C.3.

panel	source	catalog	$l/°$	$b/°$
top	A	Taylor	-3.5608535	-5.6028647
top	B	Taylor	-3.5593824	-5.5575824
bottom	C	O'Sullivan	-45.865494	-27.910206
bottom	D	O'Sullivan	-45.973431	-28.065784

This ratio can be roughly interpreted as a measure for the constraining power of the data, since it compares the uncertainty after considering the data to the uncertainty before. Note, however, that in our reconstruction, the prior variance was itself reconstructed from the data, so we have actually extracted more information from the data. A smaller ratio in Fig. C.2 means more constraining power, with a ratio of 1 meaning no new constraint at all. This figure contains two interesting aspects. First, it is evident that the overwhelming majority of points does not exhibit a mathematically impossible ratio larger than one (note the two orders of magnitude between the height of the peak at a ratio of 1 and the next bin to the right) and is therefore probably not gravely affected by numerical inaccuracies. Second, the ratio is still quite close to 1 for most of the sources, meaning that the data do not constrain the extragalactic contribution to any individual source much.

Appendix C.2.2: Correlations

It should also be noted that correlations are present in the posterior distribution for the extragalactic contributions. To illustrate this, we plot sample values for two pairs of sources that are nearby one another in Fig. C.3. The sources are described in Table C.1. Although both panels of the figure show a pair of sources that is very close, only one of them shows significant correlations. In conclusion, for some sources the posterior uncertainty of the extragalactic contributions is strongly correlated, for others not. This complicated correlation structure is automatically included when an analysis is performed using the samples as described in this appendix but is lost completely if only posterior mean and variance for each source individually are considered.

Appendix C.2.3: Change of prior

The samples discussed here describe the posterior PDF for the extragalactic contributions, which depends on the prior we have used for these, i.e., the uncorrelated Gaussian distribution with a standard deviation of $\sigma_\mathrm{e} = 6.4\,\mathrm{rad/m^2}$. We will denote this prior as $\mathcal{P}(\phi_\mathrm{e}|\sigma_\mathrm{e} = 6.4\,\mathrm{rad/m^2})$ in the following. However, the samples can even be used to calculate expectation values of a function f with respect to a posterior distribution $\mathcal{P}(\phi_\mathrm{e}|d, X)$ that is based on a new prior $\mathcal{P}(\phi_\mathrm{e}|X)$. This can be seen from a simple application of Bayes' theorem.

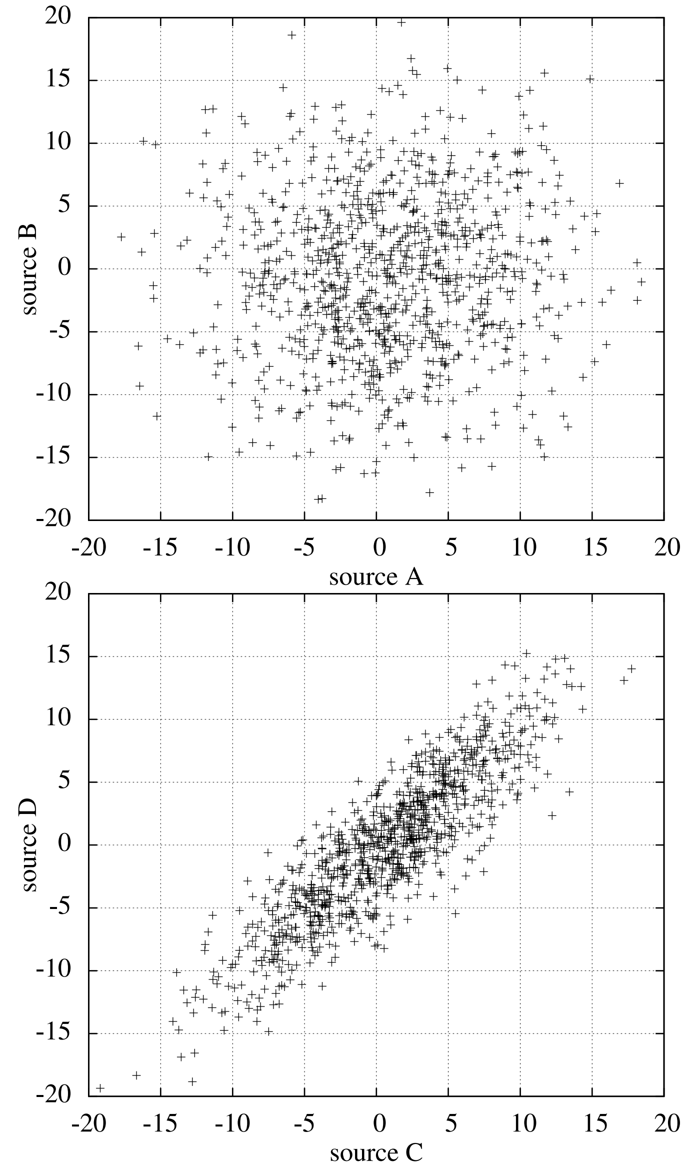

Fig. C.3. Posterior sample values for the extragalactic contributions for the four sources listed in Table C.1. All axes are in rad/m^2.

We can write the expectation value with respect to the new posterior as

$$\begin{aligned}
&\int \mathcal{D}\phi_{\rm e}\, f(\phi_{\rm e})\, \mathcal{P}(\phi_{\rm e}|d, X) \\
&\propto \int \mathcal{D}\phi_{\rm e}\, f(\phi_{\rm e})\, \mathcal{P}(d|\phi_{\rm e})\, \mathcal{P}(\phi_{\rm e}|X) \\
&= \int \mathcal{D}\phi_{\rm e}\, f(\phi_{\rm e})\, \mathcal{P}(d|\phi_{\rm e})\, \mathcal{P}(\phi_{\rm e}|\sigma_{\rm e} = 6.4\,\mathrm{rad/m^2}) \\
&\quad \times \frac{\mathcal{P}(\phi_{\rm e}|X)}{\mathcal{P}(\phi_{\rm e}|\sigma_{\rm e} = 6.4\,\mathrm{rad/m^2})} \\
&= \int \mathcal{D}\phi_{\rm e}\, f(\phi_{\rm e})\, \mathcal{P}(\phi_{\rm e}|d, \sigma_{\rm e} = 6.4\,\mathrm{rad/m^2})\, \frac{\mathcal{P}(\phi_{\rm e}|X)}{\mathcal{P}(\phi_{\rm e}|\sigma_{\rm e} = 6.4\,\mathrm{rad/m^2})},
\end{aligned} \tag{C.8}$$

so as an expectation value of the function $f(\phi_{\rm e})\, \frac{\mathcal{P}(\phi_{\rm e}|X)}{\mathcal{P}(\phi_{\rm e}|\sigma_{\rm e}=6.4\,\mathrm{rad/m^2})}$ with respect to the original posterior. In practice this means that one has to calculate a weighted average of the function f evaluated on the samples,

$$\int \mathcal{D}\phi_{\rm e}\, f(\phi_{\rm e})\, \mathcal{P}(\phi_{\rm e}|d, X) \approx \frac{1}{W} \sum_{k=0}^{999} f^{(k)}\, w^{(k)}, \tag{C.9}$$

where the weights are given by the prior ratios

$$w^{(k)} = \frac{\mathcal{P}(\phi_{\rm e}^{(k)}|X)}{\mathcal{P}(\phi_{\rm e}^{(k)}|\sigma_{\rm e} = 6.4\,\mathrm{rad/m^2})} \tag{C.10}$$

and

$$W = \sum_{k=0}^{999} w^{(k)}. \tag{C.11}$$

Table C.2. The first 10 rows of the file containing the 1 000 samples of the extragalactic contributions to the sources' Faraday rotation.

source information							extragalactic sample values in rad/m^2			sample	sample
catalog	$l/°$	$b/°$	$\phi^{(\mathrm{observed})a}$	$\sigma^{(\mathrm{observed})a}$	$\hat{\phi}_{\mathrm{g}}{}^{a}$	$\sigma_{\mathrm{g}}{}^{a}$	sample 0^a	$\cdots$	sample 999^a	mean^a	stand. dev.^{a}
Bonafede	58.10523605	88.01133728	-256.	303.	3.20471859	2.24866343	9.62054348	$\cdots$	-1.98793721	-0.08555332	6.37027788
Bonafede	54.58750916	88.20271301	372.	154.	3.36967444	2.3799901	-10.78033829	$\cdots$	4.84859753	0.02512168	6.42885876
Bonafede	57.3413353	88.07319641	-120.	166.	3.20471859	2.24866343	3.57049346	$\cdots$	2.29499125	0.0852802	6.24923992
Bonafede	62.86000824	87.48761749	21.	65.	2.6139226	2.36245131	0.3949866	$\cdots$	-5.09730434	-0.09238409	5.97149372
Bonafede	78.20828247	88.37332916	6.	56.	3.75813174	2.49758863	-4.15771627	$\cdots$	3.49833322	-0.22671212	6.36542845
Bonafede	41.70134354	87.35212708	32.	27.	2.54712152	2.2230804	-3.73729539	$\cdots$	0.07162303	1.52458452	6.27055311
Bonafede	59.23501587	87.68695831	51.	4.	2.6139226	2.36245131	-10.76599312	$\cdots$	-1.93891513	2.234744	6.29180908
Heald	144.25799561	32.43564987	-18.	2.	-15.78512192	6.85237551	7.51348639	$\cdots$	1.15834486	-1.06253238	5.58011627
Heald	140.46893311	43.54379654	-12.	3.	-17.27035713	6.21389675	10.73939037	$\cdots$	7.52750111	3.30266675	4.97111082
Heald	140.34072876	43.51054382	24.	3.	-17.27035713	6.21389675	2.55306602	$\cdots$	1.62762988	-3.9213904	4.90966082
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$

Notes. $^{(a)}$ In rad/m^2.