

Faraday Tomography Tutorial

Shinsuke Ideguchi ^{1,*}, Yoshimitsu Miyashita ² and George Heald ³ 

¹ Department of Astrophysics/IMAPP, Radboud University Nijmegen, P.O. Box 9010, NL-6500 GL Nijmegen, The Netherlands

² Graduate School of Science and Technology, Kumamoto University, 2-39-1, Kurokami, Kumamoto 860-8555, Japan; randallmiya@gmail.com

³ CSIRO Astronomy and Space Science, P.O. Box 1130, Bentley, WA 6102, Australia; George.Heald@csiro.au

* Correspondence: s.ideguchi@astro.ru.nl

Received: 16 November 2018; Accepted: 12 December 2018; Published: 14 December 2018



Abstract: The capabilities of wide-band polarization datasets that are now becoming available from precursors/pathfinders to the Square Kilometre Array (SKA), and eventually from the SKA itself, make it possible to use the Faraday tomography technique to facilitate the study of cosmic magnetism. While many programs enabling Faraday tomography have been developed by various authors and it is now becoming easier to apply the required techniques, the interpretation of the results is not straightforward. This is not only because of the lack of a one-to-one relation between the Faraday depth and the physical depth, and observational artifacts such as instrumental polarization, but also because the choice of the method that is used and its settings can be reflected in the results. Thus, it is essential to understand how the various methods enabling Faraday tomography are suited for the efficient application of the technique. In the workshop “The Power of Faraday Tomography”, we organized a Faraday tomography tutorial to help the participants understand the required tools. In this article, we summarize the basics of the techniques, and provide an overview of the tutorial.

Keywords: magnetic fields; polarization; techniques: polarimetric

1. Introduction

Polarization observations provide us with information about magnetic fields embedded along a line of sight (LOS) through the Faraday effect. The Faraday effect is a phenomenon where the plane of polarization is rotated when the polarization passes through the magnetoionic medium. The degree of the rotation ($\Delta\theta$) depends on the squared-wavelength (λ^2) of the polarization and the Rotation Measure (RM), which are expressed as

$$\Delta\theta = \text{RM} \lambda^2, \quad (1)$$

$$\text{RM} = \frac{e^3}{2\pi m_e^2 c^4} \int_{\text{source}}^{\text{observer}} n_e(r') B_{\parallel}(r') dr', \quad (2)$$

where e is the elementary charge, m_e is the electron rest mass, c is the speed of light, n_e is the thermal electron number density, $B_{\parallel} = \vec{B} \cdot \hat{r}$ is the LOS component of the intervening magnetic field, and r' is the physical distance. The order of the limits of integration is chosen to preserve the sign convention that RM is positive for a magnetic field pointing toward the observer. The coefficient of the right hand side in Equation (2) is ~ 0.812 when n_e in units of cm^{-3} , B_{\parallel} in μG , and r' in parsec are chosen. We can estimate RM from the slope of a line which is obtained by linear-fitting to the θ at multiple λ^2 . The RM provides us with the information of the integrated (or averaged) B_{\parallel} along the LOS which is weighted with the n_e . The wide-band data that are now available thanks to Square Kilometre Array (SKA) precursors/pathfinders such as the Low Frequency Array (LOFAR; [1]), the Murchison Widefield

Array (MWA; [2]), the Australian SKA Pathfinder (ASKAP; [3]), the MeerKAT (originally known as the Karoo Array Telescope; [4]), the Karl G. Jansky Very Large Array (VLA) and eventually the SKA itself makes it possible to use the Faraday effect to create a tomographic reconstruction of magnetized structures along the LOS. This is a technique known as Faraday tomography. Faraday tomography is believed to provide us with richer information about cosmic magnetic fields compared with the traditional determination and interpretation of RMs. Readers interested in the importance and specific examples of the application of Faraday tomography for cosmic magnetism are referred to the many other articles in this special issue. This article provides a summary of the Faraday tomography tutorial that was held in the workshop “The Power of Faraday Tomography”¹.

While Faraday tomography is believed to be an important technique for the study of cosmic magnetism, it is not straightforward to extract physical information describing the target objects from the results of the underlying data processing techniques, which are typically encapsulated in the so-called Faraday spectrum (or Faraday dispersion function, FDF). One reason is because of the lack of a one-to-one relation between the Faraday depth and the physical depth as we describe in Section 2. This makes it complicated to understand what kind of physical information can be concluded on the basis of the Faraday spectrum. Another reason is our still-limited frequency coverage, and the appearance of observational artifacts in the data processing (e.g., instrumental polarization). These make it impossible to fully reconstruct the intrinsic Faraday spectrum in practice. An additional possible reason is that the reconstructed Faraday spectrum can depend on the different data processing methods used to support Faraday tomography, and their settings. Since the observable frequency band is limited as mentioned above, some assumption must be included when Faraday tomography is implemented. The assumptions depend on the processing methods and this creates the dependence of the methods and their settings on the results. Nowadays, we can find various codes/scripts to run procedures enabling Faraday tomography and to easily apply these techniques to observed polarization data. However, to maximize the potential of the technique, it is essential to first understand how the methods work. The tutorial in the workshop focuses on the third difficulty and aims to help radio astronomers to understand how the methods work, with a particular focus on two methods: RM Synthesis + RM CLEAN, which is the most commonly applied method, and QU-fitting with Markov Chain Monte Carlo (MCMC), which is (possibly) the second-most applied. Below, we briefly explain each method in Section 2. Then, we introduce the contents of the tutorial in Section 3 and the summary follows in Section 4.

2. Basics of Faraday Tomography

2.1. Faraday Tomography

We define “Faraday tomography” to mean the process of determining the distribution and properties of magnetoionic material along the line of sight toward, and within, astronomical objects. This is distinct from computational techniques that are commonly employed to categorize and quantify the properties of the polarized emission along the corresponding lines of sight, for example Rotation Measure (RM) Synthesis, RM CLEAN, and QU-fitting. We begin by discussing the challenges and capabilities of Faraday tomography, and then proceed to describe the required computational techniques in Sections 2.2–2.5.

Faraday tomography presupposes the availability of a reliable method to estimate the Faraday spectrum, $F(\phi)$, from the observed polarization spectrum, $P(\lambda^2)$, through the equation

$$P(\lambda^2) = Q(\lambda^2) + iU(\lambda^2) = \int_{-\infty}^{+\infty} F(\phi) e^{2i\phi\lambda^2} d\phi, \quad (3)$$

¹ http://ska-jp.org/ws/SKAJP_MAGWS2018/index.html.

where P in units of Jy and F in units of Jy/(rad m⁻²) express the complex polarized flux density, Q and U are the linear Stokes parameters, and λ is the wavelength in meters. The Faraday depth, $\phi(r)$, has units of rad m⁻² and is defined as

$$\phi(r) = \frac{e^3}{2\pi m_e^2 c^4} \int_r^0 n_e(r') B_{\parallel}(r') dr' \quad (4)$$

where the physical quantities are the same as those in Equation (2). In the Faraday tomography regime, RM is redefined as the slope of a polarization angle θ versus λ^2 plot:

$$\text{RM} = \frac{d\theta(\lambda^2)}{d\lambda^2}, \quad (5)$$

where

$$\theta = \frac{1}{2} \tan^{-1} \frac{U}{Q}. \quad (6)$$

Note that Faraday depth is equal to RM in the simplest situation, in which the $F(\phi)$ is composed of a single delta function.

From Equation (4), we can see one of the primary challenges related to Faraday tomography. The sign of B_{\parallel} can change along the line of sight, which means that the value of ϕ does not vary monotonically with r . In other words, *Faraday depth is not equivalent to physical depth*. This can make it hard to order polarized emission components identified at different Faraday depths along the line of sight. To resolve this issue, other information is usually brought to bear: for example, considerations of the angular size of polarized features at particular values of Faraday depth; the morphology and angular scales of depolarization features; and correspondence with objects and features identified in other portions of the electromagnetic spectrum (e.g., H α). Studies that have employed these techniques to build a tomographic model include those presented by [5,6].

As noted above, the starting point for Faraday tomography is to quantify the polarized emission at various values of the Faraday depth, i.e., to estimate the function $F(\phi)$. Different techniques can be used, depending on whichever is most suitable to the problem at hand. We now summarize the most commonly used techniques for achieving this step. Pros and cons of the fundamental methods are discussed in Section 2.5.

2.2. RM Synthesis

RM Synthesis was initially introduced by Burn [7] and was further developed by Brentjens and de Bruyn [8]. It is the simplest and the most commonly applied technique to estimate $F(\phi)$. This is achieved by means of the inverse Fourier transform of Equation (3):

$$F(\phi) = \frac{1}{\pi} \int_{-\infty}^{+\infty} P(\lambda^2) e^{-2i\phi\lambda^2} d\lambda^2. \quad (7)$$

Note that perfect reconstruction of $F(\phi)$ is impossible since negative values of λ^2 are non-physical, and the observable wavelength range of any realistic observation is limited. Thus, we instead consider the reconstructed Faraday spectrum, $\tilde{F}(\phi)$, which is expressed as

$$\tilde{F}(\phi) = \frac{1}{\pi} \int_{-\infty}^{+\infty} W(\lambda^2) P(\lambda^2) e^{-2i\phi\lambda^2} d\lambda^2 \quad (8)$$

where the window function, $W(\lambda^2)$, is usually set to 1 for observable wavelengths and to 0 otherwise (uniform weights). Alternatively, $W(\lambda^2)$ can be specified as the reciprocal squared of the rms noise in

each frequency channel (inverse variance weighting). Through the use of the convolution theorem, Equation (8) can be rewritten as

$$\tilde{F}(\phi) = \frac{1}{K} F(\phi) * R(\phi), \quad (9)$$

$$R(\phi) = K \int_{-\infty}^{+\infty} W(\lambda^2) e^{-2i\phi\lambda^2} d\lambda^2, \quad (10)$$

$$K = \left(\int_{-\infty}^{+\infty} W(\lambda^2) d\lambda^2 \right)^{-1}. \quad (11)$$

The function $R(\phi)$ is commonly referred to as the rotation measure spread function (RMSF). The factor K is introduced to normalize the peak of $R(\phi)$ to 1. Equation (9) shows us that the reconstructed Faraday spectrum $\tilde{F}(\phi)$ is the convolution of the intrinsic $F(\phi)$ with the RMSF, which imposes sidelobe structure on the reconstruction. The $\tilde{F}(\phi)$ is thus often called the “dirty Faraday spectrum” in analogy to the case of aperture synthesis imaging.

In practice, since the observational data are obtained with discrete channels, Equations (9)–(11) can be written as sums,

$$\tilde{F}(\phi) \simeq K \sum_{j=1}^N P(\lambda_j^2) e^{-2i\phi(\lambda_j^2 - \lambda_0^2)}, \quad (12)$$

$$R(\phi) \simeq K \sum_{j=1}^N W(\lambda_j^2) e^{-2i\phi(\lambda_j^2 - \lambda_0^2)}, \quad (13)$$

$$K = \left(\sum_{j=1}^N W(\lambda_j^2) \right)^{-1}, \quad (14)$$

$$\lambda_0^2 = \frac{\sum_{j=1}^N W(\lambda_j^2) \lambda_j^2}{\sum_{j=1}^N W(\lambda_j^2)} \quad (15)$$

where the factor λ_0 is typically added for convenience and for aesthetics, and can take any value (including zero). The particular value provided in Equation (15), namely the weighted average of the observation wavelengths, is often adopted in order to avoid rapid variation with ϕ of the RMSF in the complex plane.

We also repeat here three important parameters characterizing the ability of RM Synthesis to reconstruct the true Faraday spectrum. As introduced in Brentjens and de Bruyn [8], these are:

$$\delta\phi \simeq \frac{2\sqrt{3}}{\lambda_{\max}^2 - \lambda_{\min}^2}, \quad (16)$$

$$\text{max-scale} \simeq \frac{\pi}{\lambda_{\min}^2}, \quad (17)$$

$$||\phi_{\max}|| \simeq \frac{\sqrt{3}}{\delta\lambda^2}, \quad (18)$$

where λ_{\min} and λ_{\max} are the shortest and the longest squared wavelengths, respectively, and $\delta\lambda^2$ is the channel width in squared wavelengths space. Here, $\delta\phi$ is the Full Width at Half Maximum (FWHM) of the RMSF and quantifies the resolution in ϕ space. The parameter ‘max-scale’ refers to the fact that RM Synthesis cannot fully reconstruct Faraday spectra that contain polarized emission present over a wide range of ϕ , and provides the representative scale in ϕ space for which sensitivity has dropped by half. Finally, $||\phi_{\max}||$ estimates the maximum Faraday depth at which the observations provide more than half sensitivity despite bandwidth depolarization within individual channels. The first two parameters demonstrate that while observational coverage at lower frequencies provides higher

precision in ϕ space, lower frequency observations also provide less sensitivity to broad structure in ϕ space (commonly referred to as “Faraday thick” structures). The last parameter shows that higher spectral resolution (expressed in λ^2 space) is necessary for properly studying polarized sources with large absolute values of Faraday depth.

2.3. RM CLEAN

The RM CLEAN algorithm was proposed by [9], implemented by [10], and described in further detail by Heald [11]. The algorithm is conceptually similar to the CLEAN deconvolution that was developed for imaging of visibility data from aperture synthesis radio telescopes. RM CLEAN is usually performed to improve the dirty Faraday spectrum, $\tilde{F}(\phi)$, as follows. First, the location of ϕ at which $||\tilde{F}(\phi)||$ has its peak, ϕ_p , is searched. The values of the real and imaginary part of $\tilde{F}(\phi_p)$ are multiplied with the real-valued loop gain parameter γ , and are stored as CLEAN components. Then, the shifted and scaled RMSF, $\gamma\tilde{F}(\phi)R(\phi - \phi_p)$, is subtracted from $\tilde{F}(\phi)$. The residuals of $||\tilde{F}(\phi)||$ are searched for a new peak, and the above two steps are repeated until the peak of the residual is under a threshold or until the loop has reached a maximum number of iterations. Finally, the stored CLEAN components are accumulated and convolved with a real-valued Gaussian function whose FWHM is the same as that of the RMSF, and the residual of $\tilde{F}(\phi)$ is added. This is often referred to as the “CLEANed Faraday spectrum”. The loop gain parameter γ is often chosen as 0.1 due to the optimal balance between reducing the number of iterations, which is equivalent to reducing the computational costs, and minimizing the residuals. Through the use of this technique, the sidelobes that occur due to a limited set of observable wavelengths are largely removed, and the true components of the Faraday spectrum are better reconstructed.

RM Ambiguity

Farnsworth et al. [12] report that a CLEANed Faraday spectrum of $F(\phi)$ with two closely-spaced ϕ components sometimes conspires to produce a third, artificial component between the two even when the gap between the pair in ϕ space is as large as a value that is similar to the FWHM of the RMSF. When the number of components of input $F(\phi)$ and that of the dominant accumulated CLEAN components are different, it is called an RM ambiguity. Basically, this occurs when $\tilde{F}(\phi)$ is dominated by a single peak that has a centroid between the two true components, due to constructive interference between considerable (complex-valued) sidelobes corresponding to each component. If one applies the RM CLEAN algorithm to such a $\tilde{F}(\phi)$, the procedure regards the single intervening peak as the most reliable component and creates a CLEAN component at a value of ϕ where there is actually no polarized emission. At that point, RM CLEAN has started with an incorrect model and can never recover to properly reconstruct the true $F(\phi)$. The RM ambiguity can lead to substantial error in the estimation of ϕ values at which polarized emission is present in $F(\phi)$.

2.4. QU-Fitting with MCMC

The alternative method to estimate $F(\phi)$ is the model fitting method, which is often called “QU-fitting” (e.g., [12,13]). This approach is usually computationally more expensive than RM Synthesis + RM CLEAN. Moreover, the method requires an assumption to be made for the fitting model, unlike with RM Synthesis + RM CLEAN. On the other hand, Sun et al. [14] show that QU-fitting can be more powerful than RM Synthesis + RM CLEAN provided that an appropriate fitting model is selected. The QU-fitting method proceeds as follows. First, the model Faraday spectrum given the model parameter(s), w , is assumed as $F_{\text{mod}}(\phi; w)$. Then the model polarized spectrum, $P_{\text{mod}}(\lambda^2; w)$, is calculated from $F_{\text{mod}}(\phi; w)$ and it is compared with the observed spectrum, $P(\lambda^2)$. The above steps are repeated by changing w until $P_{\text{mod}}(\lambda^2; w)$ has been identified as the best fit to $P(\lambda^2)$. Usually,

the noise on $P(\lambda^2)$, $\sigma(\lambda^2)$, is also used to apply weights to each measured value. Thus, the fitting is equivalent to find w that satisfies:

$$\min_w [\chi^2(w)], \quad (19)$$

$$\chi^2(w) = \sum_{j=1}^N \frac{(P(\lambda_j^2) - P_{\text{mod}}(\lambda_j^2; w))^2}{\sigma^2(\lambda_j^2)}, \quad (20)$$

where N is the number of frequency channels and j refers to the individual channels.

The choice of the fitting method is important for the efficient sampling of the parameter(s). Among various available methods, we select the Metropolis algorithm of the Markov Chain Monte Carlo (MCMC) as the fitting method employed here, because it is relatively easy to implement and effective. The MCMC is executed as follows. First, the candidate parameter(s), w' , are generated from the normal distribution with the previous parameter(s), w_t , as the average and the given step width as the variance. The candidate is accepted with the probability u ,

$$u = \min \left(1, \frac{L(w')}{L(w_t)} \right), \quad (21)$$

and the parameter(s) are updated as $w_{t+1} = w'$, or otherwise the candidate is rejected and the parameter(s) are unchanged as $w_{t+1} = w_t$. Here, $L(w)$ is the likelihood given the parameter(s) and is expressed using χ^2 as

$$L(w) \propto \exp \left(-\frac{1}{2} \chi^2(w) \right). \quad (22)$$

Since minimizing $\chi^2(w)$ is equivalent to maximizing $L(w)$, the ultimate goal of the fitting is to find the peak or the global maximum of $L(w)$. However, the fitting sometimes suffers from the local maximum problem in which the parameter(s) are trapped in one of the local maxima of $L(w)$. MCMC often resolves the local maximum problem by allowing the parameter(s) candidate that provides a smaller $L(w)$ to be updated. The fitting is usually terminated when the parameter chain has converged, since this occurs when the parameter(s) are located in some maximum. Although it is not easy to figure out if the parameter(s) are located in a local maximum or the global maximum, the resulting χ^2 value can be an indicator of this. This is because the condition $\chi^2 = N - M$, where M is the number of model parameters, indicates that the resulting differences between the observations P and model estimates P_{mod} are in accordance with the measurement errors σ .

Another important procedure for the model fitting method is the selection of a model. It is often good practice to try fitting a single dataset with several models and then to select the model that provides the best goodness-of-fit to the data. In general, the χ^2 value decreases (improves) with an increasing number of parameters. On the other hand, selecting a model with too many parameters can result in so-called overfitting: the model has too much freedom to closely track the incidental variations of the data due to e.g., observational noise. One solution to find the optimum model is to use the information criteria. Among many measures to assess the goodness-of-fit, we use the Bayesian Information Criterion (BIC), which is expressed as

$$BIC = -2 \ln(L(w_{bf})) + k \ln(N) \propto \chi^2(w_{bf}) + k \ln(N), \quad (23)$$

where w_{bf} are the parameter(s) that give the best fit to the data and k is the number of parameters. The first term shows the lowest χ^2 value that is obtained from the result of MCMC, and the second term is the penalty for the number of model parameters. The penalty discourages overfitting. The model with the minimum BIC value is selected as the optimum model.

2.5. Comparison of Methods

Here, we briefly mention the pros and cons of the two techniques described above (see Table 1). The first factor to be mentioned is whether a model selection procedure is required. The process of RM Synthesis + RM CLEAN basically entails the Fourier transform (FT) of $P(\lambda^2)$, and thus the assumption of a model is not explicitly necessary. (However, RM CLEAN entails the same sort of *implicit* assumption as the CLEAN technique when applied to aperture synthesis imaging: namely, that the domain is largely empty and contains a few unresolved sources.) On the other hand, we do require the explicit selection of fitting models in order to apply QU-fitting. In this sense, RM Synthesis + RM CLEAN is a more generic approach that is therefore much easier to apply for an arbitrary dataset. However, as soon as appropriate models can be assumed and properly selected, the QU-fitting approach provides a better estimation of $F(\phi)$ as compared to the combination of RM Synthesis + RM CLEAN. In addition, we can estimate in a straightforward manner the errors of model parameter(s) from the sampled MCMC chains.

Finally, we present some basic considerations regarding the computational costs of these techniques. Here, we approximate the required number of basic arithmetic operations. Although the computational complexity of these processes is quite difficult to assess, we can already understand some basic differences from the scaling with a few key parameters. For example, RM Synthesis needs $O(N^2)$ calculations if the discrete FT is used, or $O(N \log_2 N)$ if the Fast FT is used (e.g., [15]), where N is the number of frequency channels in the observation. A more detailed analysis and the opportunities presented by GPU acceleration are presented by Sridhar et al. [16]. RM CLEAN involves an iterative loop for subtracting the factor-multiplied RMSF from $\tilde{F}(\phi)$, and thus the computational cost is $O(N \times N_{\text{CLEAN}})$, where N_{CLEAN} is the number of required RM CLEAN iterations. Although N_{CLEAN} depends on the exact situation, it is usually $O(10)$ for a typical choice of the CLEAN gain parameter ($\gamma \approx 0.1$). In the case of QU-fitting with MCMC, the computational cost is $O(N \times N_{\text{MCMC}})$. N_{MCMC} is the length of MCMC parameter chain. Experimentally, $O(10^4)$ is needed to fit models with a number of parameters of $O(1)$, while larger N_{MCMC} may be necessary for more complex fitting models. Note that usually we need to perform QU-fitting with several models, and then select the best result on the basis of the BIC, so the total computation cost also depends on the number of models that are employed. The same scaling holds when the fitting model can be Fourier transformed analytically so that a numerical FT is not needed.

Table 1. Comparison of pros (○) and cons (×) of $F(\phi)$ estimation techniques.

	RM Synthesis + RM CLEAN	QU-fitting
Model assumption	○ Not explicitly necessary	× Needed
Estimation of $F(\phi)$	⊗ Not bad	○ Good (If appropriate models are given)
Estimation of error	× Impossible	○ Possible
Computational cost	○ Cheap	× Expensive

3. Tutorial

In this section, we describe the mock observational data and the scripts that are used in the tutorial, and we detail the contents of the practical exercises. The data and scripts, as well as the document describing the exercises, can be downloaded from the workshop website².

3.1. Data

We calculate $P(\lambda^2)$ from the assumed $F(\phi)$ using Equation (3). In order to concentrate on understanding the behaviors of the methods themselves, we choose simple models describing $F(\phi)$:

² http://ska-jp.org/ws/SKAJP_MAGWS2018/information.html.

one delta function ($d1$), two delta functions ($d2$), one Gaussian function with three different values of the standard deviation ($g1a$, $g1b$ and $g1c$), and two Gaussian functions ($g2$). The parameters of the components are summarized in Table 2. We also use these simple $F(\phi)$ as the fitting models employed for QU-fitting in order to avoid numerical Fourier transform, which in general increases computational costs. We prepare the data from 100 MHz to 9 GHz with 10^{-3} m^2 spacing (i.e., the channels are equally-spaced in λ^2 space), and add random Gaussian noise with a mean value of 0 and a variance of 1 per channel.

Table 2. The parameters of the assumed $F(\phi)$ models used to generate the mock data. Amp. is the amplitude, θ is the polarization angle and $\delta\phi$ is the standard deviation of the component for the Gaussian function. Note that the amplitude for Gaussians is the integrated value in ϕ space. For two-component $F(\phi)$, namely $d2$ and $g2$, parameters for both components are shown.

$F(\phi)$	ϕ [rad m ⁻²]	Amp. [arb. units]	θ [deg.]	$\delta\phi$ [rad m ⁻²]
$d1$	20	6	30	-
$d2$	20	6	30	-
	42	7.5	-60	-
$g1a$	20	6	30	11
$g1b$	20	6	30	22
$g1c$	20	6	30	44
$g2$	20	6	30	22
	42	7.5	-60	11

3.2. Scripts

The scripts performing RM Synthesis and RM CLEAN are fully written in Python³. As described in Section 2, RM Synthesis is basically the Fourier transform, and RM CLEAN is an iterative loop that subtracts the factor-multiplied RMSF from the complex dirty Faraday spectrum. These are executed using the NumPy library [17]. The QU-fitting script is written in Python for the I/O part and the post-processing part such as calculation of information criteria and plotting, while the MCMC part is, for the purpose of faster computation, written in Fortran90 and is called from Python. These scripts are executed with parameters such as the name of the data file, the minimum and maximum frequencies as well as the necessary parameters for each method. The figures are plotted using matplotlib [18].

3.3. Practical Exercises

In the tutorial, we held practical exercises in order for the participants to interpret the behaviors of each Faraday tomography method. Here, we introduce the contents of the exercises. One can use a Jupyter notebook [19], which can be found in the downloaded file, to follow the exercises below and to check the results.

3.3.1. RM Synthesis

The RM Synthesis part starts with the reconstruction of the model $d1$. We check how the reconstructed (dirty) Faraday spectrum, $\tilde{F}(\phi)$, changes with different frequency band widths that determine the FWHM. Since the model $F(\phi)$ consists of only one delta function, this is consistent with the check of the dependency of the RMSF on different frequency bands except that the peak of the RMSF is normalized to 1. Here, let us try three band widths, 700–1800 MHz (e.g., ASKAP full band, case I), 800–1800 MHz (e.g., ASKAP without the lowest 100 MHz, case II) and 120–240 MHz

³ <https://github.com/gheald/RMtoolkit>.

(e.g., LOFAR High Band Antenna, case III). As described in Section 2, the $\tilde{F}(\phi)$ has the peak at the same ϕ with the input $F(\phi)$ but also has sidelobes that causes uncertainty in estimating the position of component(s). We can see that the width of the peak is smaller with larger bandwidth in λ^2 space such that case III < case I < case II.

As the next step, we perform the same procedure with the same three bandwidths using the model *d2*. We check that the two components can be resolved when the gap between the two is larger than or similar to the FWHM. Cases I (left panel of Figure 1) and III correspond to this situation. On the other hand, the two components cannot be resolved when the gap is smaller than the FWHM (case II, right panel of Figure 1).

Finally, the model *g1a* is used to check how the results change depending on the relation between the widths of the FWHM and that of the Gaussian function. Since the higher frequency is more sensitive to diffuse components in ϕ space as shown in Equation (17), here we consider bands with higher frequency, 1000–10,000 MHz (case IV) and 600–10,000 MHz (case V), that lead to the FWHM of ~ 39 and ~ 14 rad m $^{-2}$, respectively. Using these bands, we apply RM Synthesis to the model *g1a*, a Gaussian with a width of 11 rad m $^{-2}$. When the width of components is larger than the FWHM (case IV), the peak of $\tilde{F}(\phi)$ has the width of \sim FWHM since a Faraday structure smaller than the FWHM basically cannot be resolved. On the other hand, when the width of components is similar to the FWHM (case V), the width of the peak of $\tilde{F}(\phi)$ is larger than the FWHM and we can guess that there may be a diffuse component.

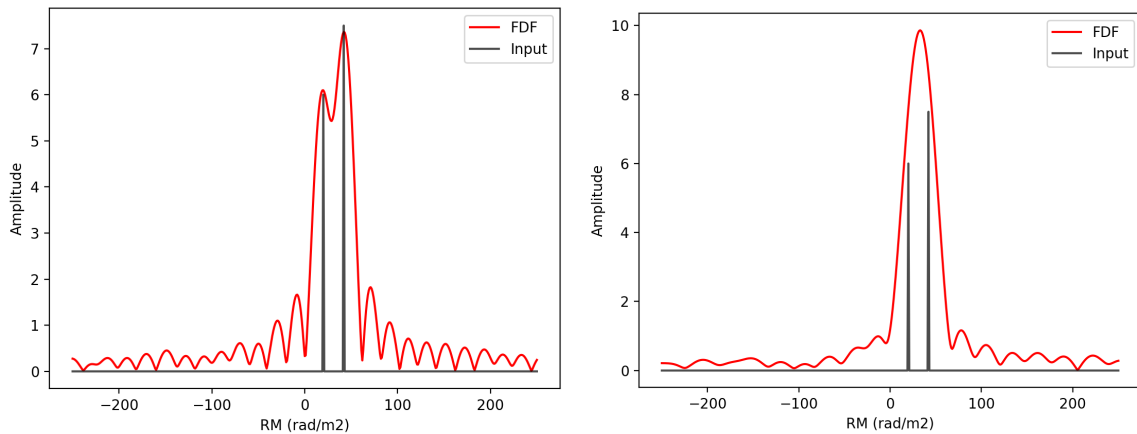


Figure 1. Examples of the results of the rotation measure (RM) Synthesis. The left and right figures show the results for the model *d2* using the frequency band case I and case II, respectively.

3.3.2. RM CLEAN

As with the RM Synthesis part, we start with the reconstruction of the model *d1*. As a first step, we follow the iterative behavior of the RM CLEAN algorithm by looking at the CLEANing procedure step-by-step using the case I frequency band. The result shows that the CLEAN components appear cumulatively at the input ϕ , and the side lobes are gradually subtracted with each iteration. Finally, the CLEANed $F(\phi)$ shows up at the input ϕ with the FWHM width. Note that the small fluctuations around the peak of the CLEANed $F(\phi)$ are due to noise on $P(\lambda^2)$.

In the case of the model *d2*, we check that the two components can be clearly reconstructed if the gap between these two is larger than or comparable to the FWHM. Cases I (left panel of Figure 2) and III correspond to this situation. In addition, we confirm that the two components cannot be resolved even after RM CLEAN has been performed if these two are located within the FWHM (for case II, right panel of Figure 2). Here, there are three substantial CLEAN components although the input $F(\phi)$ consists of only two components. This demonstrates the RM ambiguity. We may mistakenly infer that there are three sources in this situation.

Finally, we perform RM CLEAN for *g1a*. In the case that the width of components is much smaller than the FWHM (case IV), there appears to only be a single CLEAN component (left panel of Figure 3). We may mistakenly suppose that there is one sharp component in ϕ space. On the other hand, when the width of the model components is similar to the FWHM (case V), we can see multiple CLEAN components which are densely distributed in ϕ space (right panel of Figure 3). In this situation, we can infer that there may be a Faraday-diffuse source.

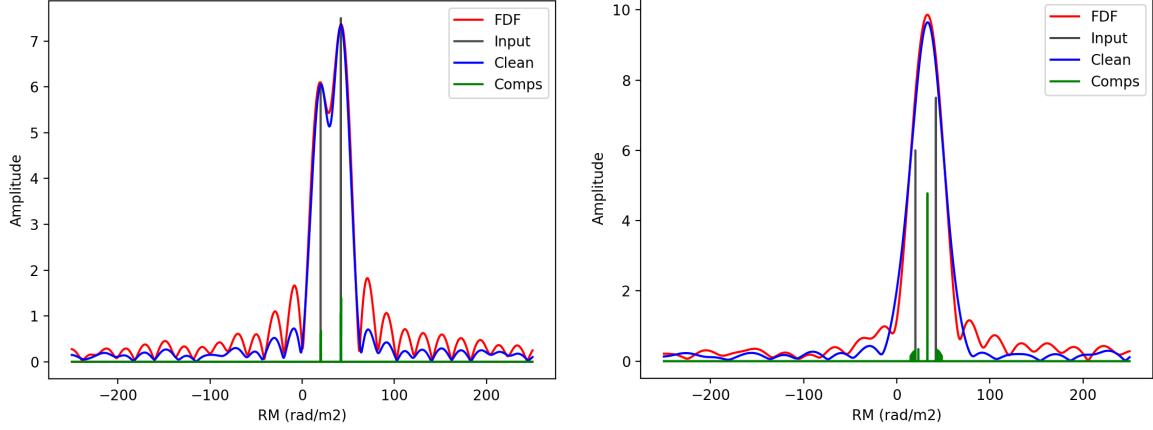


Figure 2. Examples of the results of the RM CLEAN. The left and right figures show the results for the model *d2* using the frequency band case I and case II, respectively.

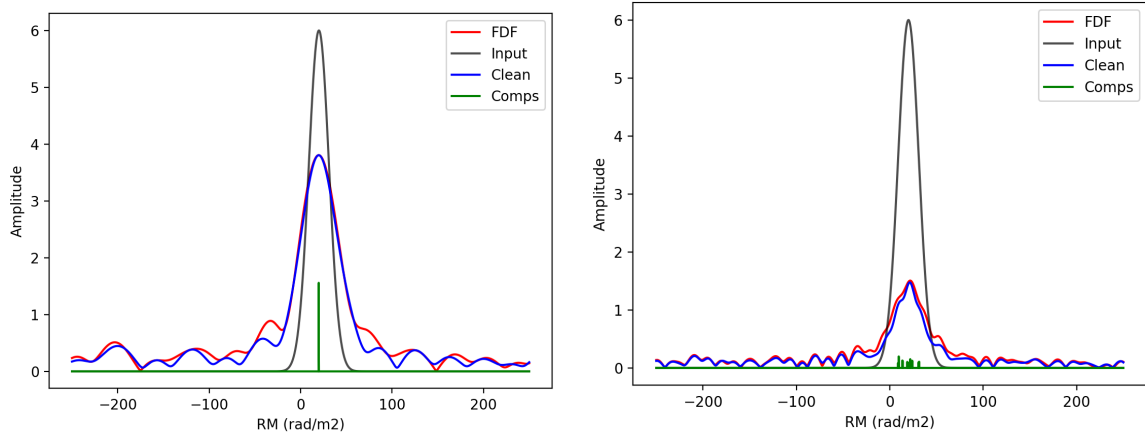


Figure 3. Examples of the results of the RM CLEAN. The left and right figures show the results for the model *g1a* using the frequency band case IV and case V, respectively.

3.3.3. QU-Fitting with MCMC

In the same manner as the above two methods, we again start with the model *d1* for QU-fitting. We first run QU-fitting with the use of the correct fitting model (i.e., one delta function model) that was used to generate the mock data. We can see that the fitting performs well in comparison to the data in frequency or λ^2 space, and the model parameters are also well estimated.

Next, we look at the effect of model selection with mock data generated from model *d2*. In this tutorial, we use the number of components to characterize the difference in fitting models. Our aim is to see whether it is possible to correctly identify the input number of components through model selection. For this purpose, we perform QU-fitting against model Faraday spectra with one to four delta functions. These models are fitted to the $P(\lambda^2)$ calculated from the model *d2*. We see that the model with only one delta function cannot fit well to $P(\lambda^2)$. The models with two to four delta functions produce fits that match visually well to $P(\lambda^2)$, but the BIC value is found to be minimized

for the two delta function model. Figure 4 shows an example of the result using the two delta functions model for the fit. Thus, we have verified that the model selection using BIC works properly.

We next try to apply QU-fitting to the situation where an RM ambiguity occurs. For this purpose, we use the model *d2* with the case II frequency band and perform QU-fitting with models containing one and two delta functions. The results show that the two delta function model provides a far better fit to the mock data as compared to the one delta function model, and is selected as the preferred model by BIC.

Finally, we apply QU-fitting to the diffuse source model *g1a* with case II frequency band. This is the situation where only one CLEAN component appeared. We perform a fit to the one delta function model and to the one Gaussian function model. The results show that the QU-fitting method correctly selects the Gaussian model, and moreover correctly estimates the width of source.

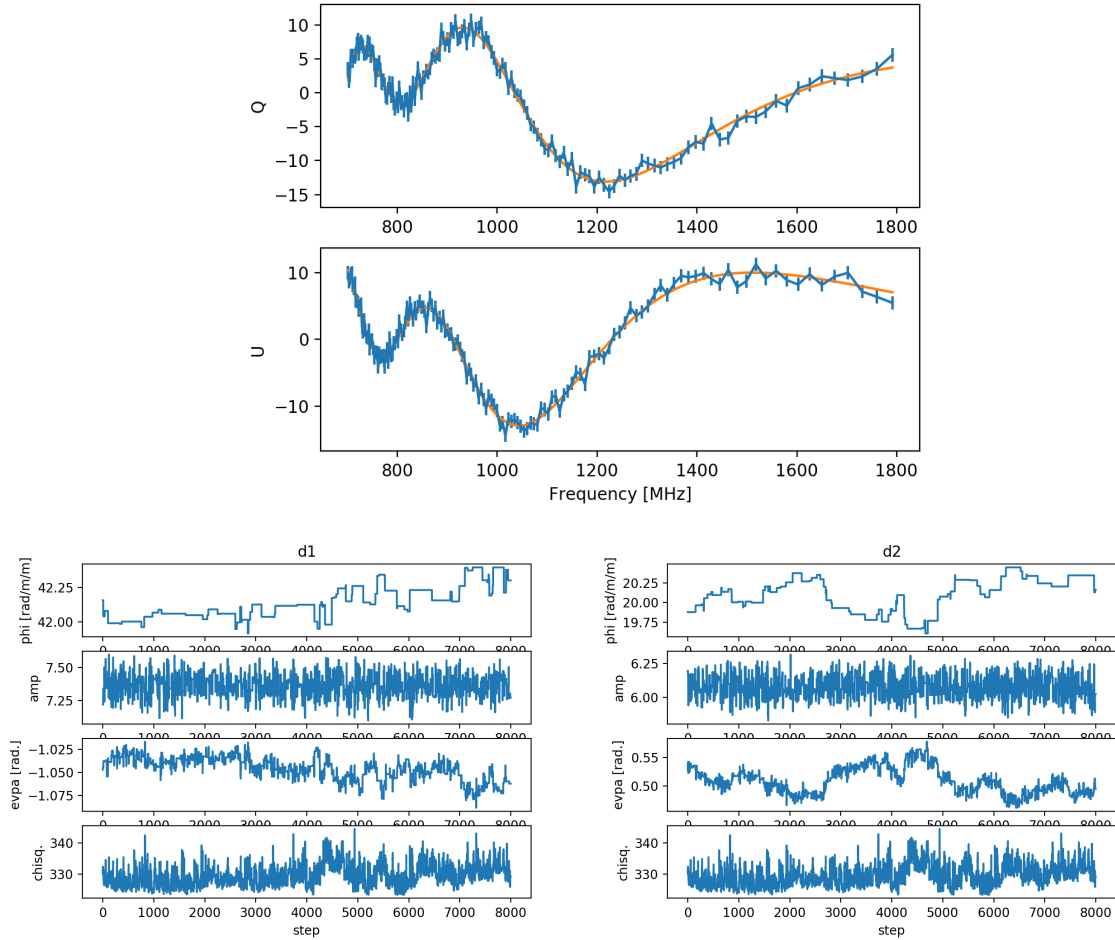


Figure 4. An example of the result of the QU-fitting. The figure shows the result when two delta functions are adopted in the fitting model, applied to mock data produced using the model *d2* with the frequency band case I. The top panel shows the Stokes Q and U spectra of the mock data (blue) and the fitted results (orange) in frequency space. The bottom left and right panels show the parameter chains for the first and second components of the fitting model and χ^2 value.

4. Summary

Today, the Faraday tomography technique is becoming necessary for the study of cosmic magnetism. On the other hand, the results of the technique are not straightforward to interpret. As a part of the workshop “The Power of Faraday Tomography”, we organized a Faraday tomography tutorial with the aim of understanding how the data processing methods work, which should assist the interpretation of Faraday tomography results. We chose the two most commonly-applied methods,

RM Synthesis + RM CLEAN and QU-fitting with MCMC, and demonstrated the changing behavior of the results depending on the frequency bands employed, as well as the different methods and their settings.

Author Contributions: Software, S.I., Y.M. and G.H.; Writing—original draft, S.I.; Writing—review & editing, G.H.

Funding: This research was financially supported in part by Grant-in-Aid from the Ministry of Education, Culture, Sports, and Science and Technology (MEXT) of Japan, No. 17J06936 (YM).

Acknowledgments: We thank K. Takahashi and T. Akahori for useful discussions about the structure of the tutorial.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Van Haarlem, M.P.; Wise, M.W.; Gunst, A.W.; Heald, G.; McKean, J.P.; Hessels, J.W.T.; de Bruyn, A.G.; Nijboer, R.; Swinbank, J.; Fallows, R.; et al. LOFAR: The LOw-Frequency ARray. *Astron. Astrophys.* **2013**, *556*, A2. [\[CrossRef\]](#)
2. Tingay, S.J.; Goeke, R.; Bowman, J.D.; Emrich, D.; Ord, S.M.; Mitchell, D.A.; Morales, M.F.; Booler, T.; Crosse, B.; Wayth, R.B.; et al. The Murchison Widefield Array: The Square Kilometre Array Precursor at Low Radio Frequencies. *Publ. Astron. Soc. Aust.* **2013**, *30*, e007. [\[CrossRef\]](#)
3. McConnell, D.; Allison, J.R.; Bannister, K.; Bell, M.E.; Bignall, H.E.; Chippendale, A.P.; Edwards, P.G.; Harvey-Smith, L.; Hegarty, S.; Heywood, I.; et al. The Australian Square Kilometre Array Pathfinder: Performance of the Boolardy Engineering Test Array. *Publ. Astron. Soc. Aust.* **2016**, *33*, e042. [\[CrossRef\]](#)
4. Jonas, J.; MeerKAT Team. The MeerKAT Radio Telescope. In Proceedings of the MeerKAT Science: On the Pathway to the SKA (MeerKAT2016), Stellenbosch, South Africa, 25–27 May 2016; p. 1.
5. Brentjens, M.A. Wide field polarimetry around the Perseus cluster at 350 MHz. *Astron. Astrophys.* **2011**, *526*, A9. [\[CrossRef\]](#)
6. Jelić, V.; de Bruyn, A.G.; Pandey, V.N.; Mevius, M.; Haverkorn, M.; Brentjens, M.A.; Koopmans, L.V.E.; Zaroubi, S.; Abdalla, F.B.; Asad, K.M.B.; et al. Linear polarization structures in LOFAR observations of the interstellar medium in the 3C 196 field. *Astron. Astrophys.* **2015**, *583*, A137. [\[CrossRef\]](#)
7. Burn, B.J. On the depolarization of discrete radio sources by Faraday dispersion. *Mon. Not. R. Astron. Soc.* **1966**, *133*, 67–83. [\[CrossRef\]](#)
8. Brentjens, M.A.; de Bruyn, A.G. Faraday rotation measure synthesis. *Astron. Astrophys.* **2005**, *441*, 1217–1228. [\[CrossRef\]](#)
9. Brentjens, M.A. Radio Polarimetry in 2.5D. Ph.D. Thesis, University of Groningen, Groningen, The Netherlands, 2007.
10. Heald, G.; Braun, R.; Edmonds, R. The Westerbork SINGS survey. II Polarization, Faraday rotation, and magnetic fields. *Astron. Astrophys.* **2009**, *503*, 409–435. [\[CrossRef\]](#)
11. Heald, G. The Faraday rotation measure synthesis technique. *Cosm. Magn. Fields* **2009**, *259*, 591–602. [\[CrossRef\]](#)
12. Farnsworth, D.; Rudnick, L.; Brown, S. Integrated Polarization of Sources at $\lambda \sim 1$ m and New Rotation Measure Ambiguities. *Astron. J.* **2011**, *141*, 191. [\[CrossRef\]](#)
13. O’Sullivan, S.P.; Brown, S.; Robishaw, T.; Schnitzeler, D.H.F.M.; McClure-Griffiths, N.M.; Feain, I.J.; Taylor, A.R.; Gaensler, B.M.; Landecker, T.L.; Harvey-Smith, L.; et al. Complex Faraday depth structure of active galactic nuclei as revealed by broad-band radio polarimetry. *Mon. Not. R. Astron. Soc.* **2012**, *421*, 3300–3315. [\[CrossRef\]](#)
14. Sun, X.H.; Rudnick, L.; Akahori, T.; Anderson, C.S.; Bell, M.R.; Bray, J.D.; Farnes, J.S.; Ideguchi, S.; Kumazaki, K.; O’Brien, T.; et al. Comparison of Algorithms for Determination of Rotation Measure and Faraday Structure. I. 1100–1400 MHz. *Astron. J.* **2015**, *149*, 60. [\[CrossRef\]](#)
15. Bell, M.R.; Enßlin, T.A. Faraday synthesis. The synergy of aperture and rotation measure synthesis. *Astron. Astrophys.* **2012**, *540*, A80. [\[CrossRef\]](#)
16. Sridhar, S.S.; Heald, G.; van der Hulst, J.M. cuFFS: A GPU-accelerated code for Fast Faraday rotation measure Synthesis. *Astron. Comput.* **2018**, *25*, 205–212. [\[CrossRef\]](#)

17. Walt, S.V.D.; Colbert, S.C.; Varoquaux, G. The NumPy Array: A Structure for Efficient Numerical Computation. *Comput. Sci. Eng.* **2011**, *13*, 22–30. [[CrossRef](#)]
18. Hunter, J.D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95. [[CrossRef](#)]
19. Kluyver, T.; Ragan-Kelley, B.; Pérez, F.; Granger, B.; Bussonnier, M.; Frederic, J.; Kelley, K.; Hamrick, J.; Grout, J.; Corlay, S.; et al. Jupyter Notebooks—A publishing format for reproducible computational workflows. In *Positioning and Power in Academic Publishing: Players, Agents and Agendas*; Loizides, F., Schmidt, B., Eds.; IOS Press: Amsterdam, The Netherlands, 2016; pp. 87–90.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).