INTERNAL REPORT

TESTING A NEW ALGORITHM FOR BAYESIAN INFERENCE OF FARADAY ROTATION IMAGES

N.Pintus and M.Murgia

Report N. 11, released: 12/10/2011

Reviewer: F.Govoni

OAC Osservatorio Astronomico di Cagliari

Preface

These pages collect the experience gained by N.Pintus at Osservatorio Astronomico Cagliari, under the supervision of M.Murgia, during a stage lasting from February to July 2011 in the framework of the second level degree course in Mathematics at University of Cagliari. The general plan of the stage was to fit the observed radio data using computational and graphic tools (described in Chapter 1) and to test a new algorithm used to solve the Bayes' theorem (as shown in Section 2.2).

This work was organized into two parts: the first was dedicated to learn the basics of these tools and, especially, the basics of the theory they dealt with. The rest of the stage was focused on the test of the new algorithm on the analysis of polarization images of the cluster of galaxies A2255.

Contents

Introduction

The problem we dealt with is understanding magnetic fields in clusters of galaxies. In particular, we want to estimate field strength and structure. In this work, we have considered a technique based on the analysis of the polarized emission from radio-galaxies, which are types of active galaxy that are very luminous at radio wavelengths, with luminosities up to 10^{39} W between 10 MHz and 100 GHz.
The Faraday effect (or Faraday rotation) is a magneto-optical phenomenon, that is, an interaction between light and a magnetic field in a medium. The Faraday effect causes a rotation of the plane of polarization which is linearly proportional to the component of the magnetic field in the direction of propagation. The polarized synchrotron radiation incoming from radio sources located inside or behind a galaxy cluster, experiences the Faraday rotation of the plane of polarization as it passes through the magnetized and ionized intracluster medium:

$$\Psi_{Obs}(\nu) = \Psi_{Int} + (c/\nu)^2 \times RM \tag{1}$$

where $\Psi_{Obs}(\nu)$ is the observed polarization angle at a frequency ν and Ψ_{Int} is the intrinsic polarization angle.
The RM is related to the thermal electron density, n_e, and magnetic field along the line of sight, $B_{||}$, through the cluster by the equation:

$$RM = 812 \int_0^L n_e B_{||} dl \qquad rad\, m^{-2} \tag{2}$$

where $B_{\|}$ is measured in μG, n_e in cm^{-3} and L is the depth of the screen in kpc.
The intracluster medium is filled with a tenuous plasma of ionized hydrogen at very high temperatures of $10^7 - 10^8 K$. This hot gas radiates in the X-ray band through the thermal free-free process. The electron density of the plasma is generally not known, but it can be estimated under the hypothesis that the intracluster medium is in hydrostatic equilibrium. In this condition,

it is often assumed that the gas density distribution can be described by the standard β-model:

$$n_e(r) = n_0(1 + r^2/r_C^2)^{-3\beta/2} \tag{3}$$

where r, n_0 and r_C are the distance from the cluster center, the central electron density, and the cluster core radius, respectively. The parameters of the β-model can be determined through X-ray satellite observations and hence the electron gas density in the cluster can be reconstructed.

A simple single-scale magnetic field model

The intracluster magnetic field is likely turbulent, with a characteristic scale much smaller than overall size of the cluster. Indeed, we expect that the magnetic field inverts its direction many times along the line-of-sight. The random nature of the intracluster magnetic field makes difficult to solve Eq. 2, even for the simple case of the spherically symmetric β-model in Eq. 3. We can, however, address the problem from a statistical perspective. Let assume that the field is uniform in cells of size Λ_C whose orientation in space is completely random. If $\Lambda_C \ll L$, the RM is given by a random walk process involving many cells along the line-of-sight. Indeed, the expectation value of the Faraday rotation is null, $\langle RM \rangle = 0$, while its dispersion is:

$$\langle RM^2 \rangle^{1/2} \equiv \sigma_{RM} = RM_{cell}\sqrt{N},$$

where $N = L/\Lambda_C$ while $RM_{cell} = 812 n_e B_0/\sqrt{3} \cdot \Lambda_C$. Here we assumed for isotropy $B_\parallel = B_0/\sqrt{3}$, with $B_0 = \sqrt{B_x^2 + B_y^2 + B_z^2}$. We indeed expect that the RM distribution is Gaussian distributed around zero with a dispersion of

$$\sigma_{RM} = 812 n_e B_0/\sqrt{3} \cdot \sqrt{\Lambda_C}.$$

This basic derivation consider constant gas density and magnetic field profiles along the cluster. However, we must consider that n_e decreases with increasing radius according to Eq. 3. Moreover, also the magnetic field strength may have a similar scaling. Indeed we assume that

$$B(r) = B_0(n_e/n_0)^\eta \tag{4}$$

here, B_0 is the magnetic field strength at the cluster center while the index η is expected in the range $0.5 - 1.0$.
The correct relation for the RM dispersion with the projected distance from the cluster, $r_\perp$, is obtained by integrating:

$$\langle RM^2\rangle = \frac{812^2}{3}\Lambda_C \int B^2 n_e^2 dl = \frac{812^2}{3} B_0^2 \Lambda_C n_0^2 \int \left(1+\frac{r^2}{r_C^2}\right)^{-3\beta(1+\eta)} dl$$
$$= \frac{812^2}{3} B_0^2 \Lambda_C n_0^2 \int \frac{1}{\left(1+\frac{r^2}{r_C^2}\right)^{3\beta(1+\eta)}} dl$$
$$= 2\frac{812^2}{3} B_0^2 \Lambda_C n_0^2 \int_0^\infty \frac{1}{\left(1+\frac{r_\perp^2}{r_C^2}+\frac{l^2}{r_C^2}\right)^{3\beta(1+\eta)}} dl;$$

involving the Abel Transform:

$$\int_0^\infty \frac{l^a dl}{(m+l^b)^c} = \frac{m^{\frac{a+1-bc}{b}}}{b}\Big[\frac{\Gamma(\frac{a+1}{b}\Gamma(c-\frac{a+1}{b})}{\Gamma(c)}\Big]; \qquad \big(a>-1, b>0, m>0, c>\frac{a+1}{b}\big)$$

we obtain:

$$\langle RM^2\rangle = \sigma_{RM}^2 - \langle RM\rangle^2 = 2\frac{812^2}{3} B_0^2 \Lambda_C n_0^2 \int_0^\infty \frac{dl}{\left(1+\frac{r_\perp^2}{r_C^2}+\frac{l^2}{r_C^2}\right)^{3\beta(1+\eta)}} =$$
$$= 2\frac{812^2}{3} B_0^2 \Lambda_C n_0^2 r_C \frac{1}{2}(1+\frac{r_\perp^2}{r_C^2})^{\frac{1-6\beta(1+\eta)}{2}} \frac{\Gamma[3\beta(1+\eta)-\frac{1}{2}]\Gamma(\frac{1}{2})}{\Gamma[3\beta(1+\eta)]}$$
$$= \frac{812^2}{3} B_0^2 \Lambda_C n_0^2 r_C (1+\frac{r_\perp^2}{r_C^2})^{\frac{1-6\beta(1+\eta)}{2}} \frac{\Gamma[3\beta(1+\eta)-\frac{1}{2}]\Gamma(\frac{1}{2})}{\Gamma[3\beta(1+\eta)]}$$
$$= \frac{812^2}{3} B_0^2 \Lambda_C n_0^2 r_C \sqrt{\pi}(1+\frac{r_\perp^2}{r_C^2})^{\frac{1-6\beta(1+\eta)}{2}} \frac{\Gamma[3\beta(1+\eta)-\frac{1}{2}]}{\Gamma[3\beta(1+\eta)]}$$

So,

$$\sigma_{RM}(r_\perp) = K B_0 \Lambda_C^{\frac{1}{2}} n_0 r_C^{\frac{1}{2}} \frac{1}{(1+r_\perp^2/r_C^2)^{\frac{6\beta(1+\eta)-1}{4}}} \sqrt{\frac{\Gamma[3\beta(1+\eta)-\frac{1}{2}]}{\Gamma[3\beta(1+\eta)]}} \qquad (5)$$

where $K = \frac{812}{\sqrt{3}}\sqrt[4]{\pi}$ is a constant depending on the integration path over the gas density distribution: $K = 624$ if the source lies completely beyond the cluster, and $K = 441$ if the source is halfway the cluster.

Chapter 1

A short description of Abell 2255

The existence of magnetic fields in the intracluster medium in clusters of galaxies is established through different methods of analysis, and the most important evidence for the presence of cluster magnetic fields comes from radio observations. Magnetic fields are revealed through the synchrotron emission of cluster-wide diffuse sources, such as "radio halo" or "relics", and from studies of the RM of polarized radio galaxies.

In any case, the presence of a magnetized plasma between an observer and a radio source changes the properties of the polarized emission from the radio source. Radio halo and relics are extended but very faint sources on the contrary radio galaxies are compact but can be very bright. Therefore, the latter are suitable for the RM analysis along specific line-of-sights so that a set of information on clusters magnetic fields can be determined in conjunction with X-ray observations of the hot gas.

We considered the cluster A2255, which is a rich cluster with signs of undergoing a merger event. It is particularly suitable to study the intracluster magnetic field because it is characterized by the presence of a diffuse radio halo source at the cluster center, a relic source at the cluster periphery, and several embedded radio galaxies. The radio halo of A2255 shows filaments of strong polarized emission: the distribution of the polarization angles in the filaments indicates that the cluster magnetic field fluctuates up to scales of 400 kpc in size.

A2255 hosts several radio galaxies, four of them J1712.4+6401, J1713.3+6347, J1713.5+6402 and J1715.1+6402, have been selected for the analysis on the basis of their high flux density, extension, and the presence of polarized emission [1]. Now, we make a short description of each source.
J1712.4+6401: the source has a narrow-angle tail structure and it is located

Table 1.1: The radio sources analyzed in this work

Source	Nickname	$r_\perp$ (kpc)	σ_{RM} (rad/m^2)	$\langle ERR \rangle$ (rad/m^2)	nbeam
J1712.4+6401	B	315	86.8	2.3	24.3
J1713.5+6402	E	450	66.3	1.7	11.1
J1715.1+6402	F	1398	3.3	2.2	8.3
J1713.3+6347	D	1570	6.9	2.6	18.7

in projection quite near to the cluster center (nickname B).

J1713.3+6347: this narrow-angle tail radio galaxy is located in the southern part of the cluster far from the center (nickname D).

J1713.5+6402: the source has a total extension of about 35" and a double structure with an unresolved core (nickname E).

J1715.1+6402: this wide-angle tail source, located in the periphery of the cluster (nickname F).

For each source we calculated the RM image and the relative uncertainty image by fitting the observed polarization angle as a function of frequency (see Fig. 1.1). Our polarization data has been taken at six frequencies: 1385, 1465, 4535, 4885, 8085, and 8465 MHz. A few plots showing the RM fit are presented in Fig. 1.2. A summary of the RM images is reported in Table 1.1. The RM images have a resolution of 4 arcsecond and they are sampled at 4 pixels per beam, i.e. the image cellsize is 1 arcsecond. At the distance of A2255, 1 arcsecond corresponds to a linear scale of 1.5 kpc.

An analysis of the rotation measure of radio sources sampling different lines-of-sight across the cluster, together with an X-ray observation of the intracluster gas, can be used to derive information on the strength and structure of the cluster magnetic field.

We made use of the `FARADAY` tool [2]. This dedicated C++ code permits to investigate cluster magnetic fields by comparing the observations with the expectations of different magnetic field models. Applying this method to A2255, we aim to find the optimal magnetic field strength and structure which best describe the data.

For the distribution of the thermal electron gas density we assumed a standard β-model profile with parameters: $r_C = 432$ kpc, $n_0 = 2.05 \times 10^{-3}$ cm^{-3}, and $\beta = 0.74$.

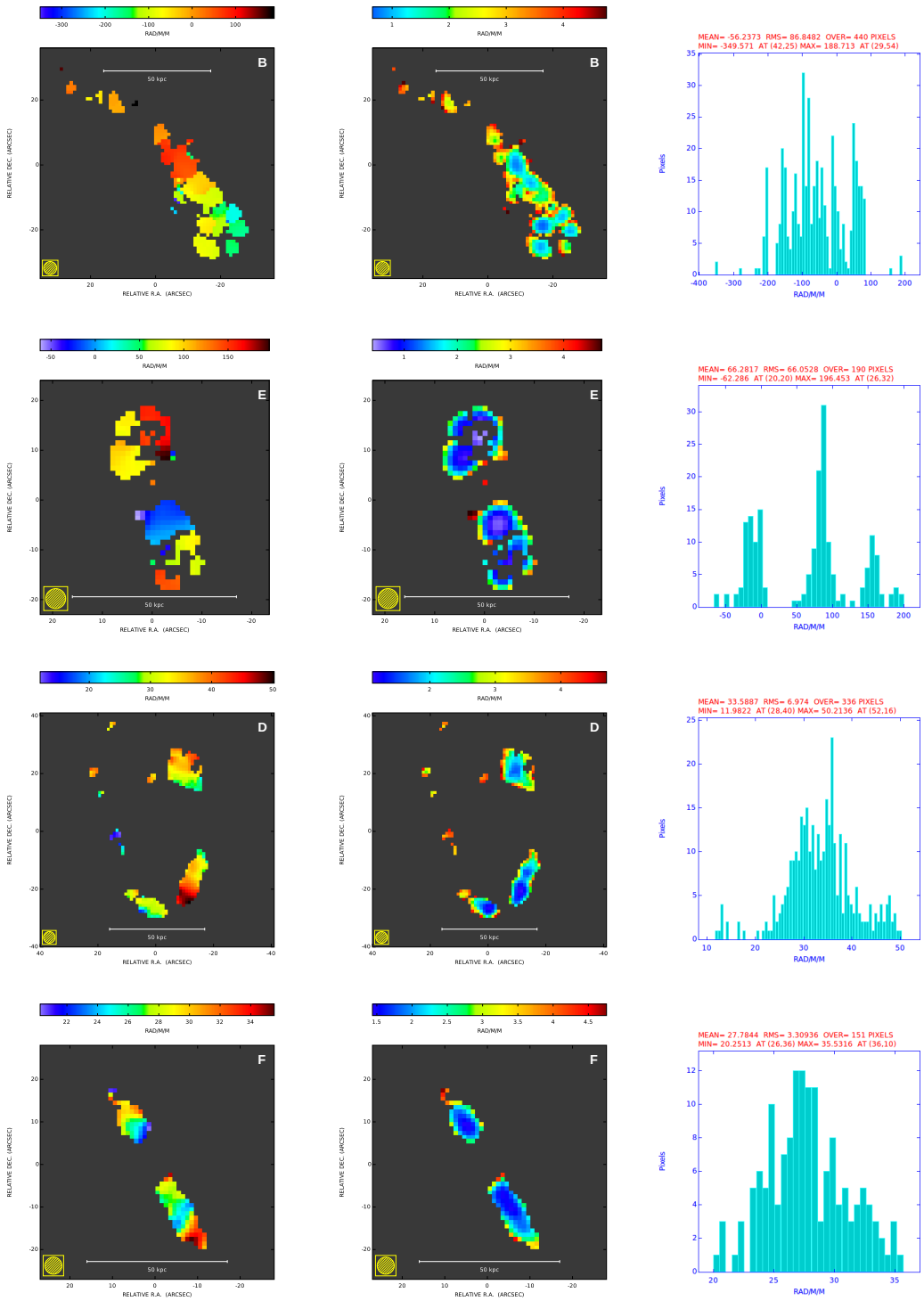

Figure 1.1: Images of the RM (left column), RM error (central column), and RM distribution (right column) for the four radio sources.

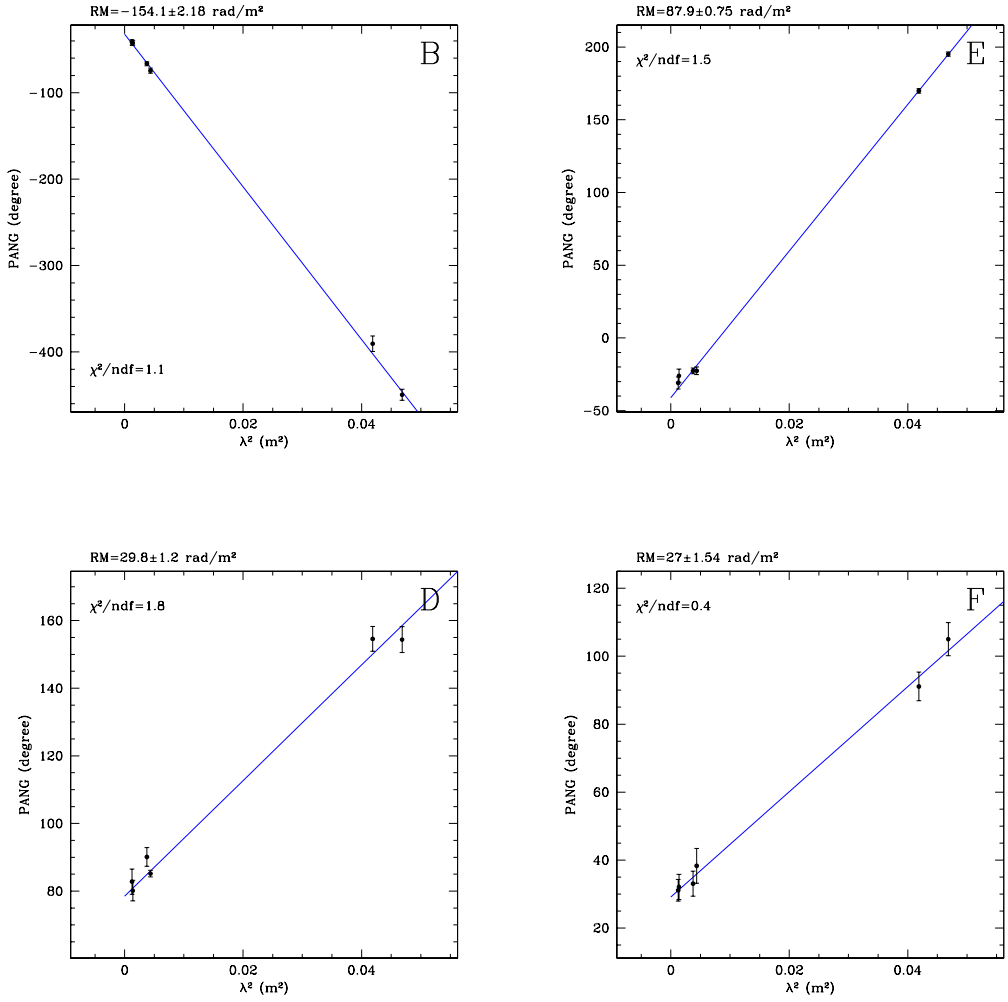

Figure 1.2: Examples of fit of the polarization angle as a function of $\lambda^2 = (c/\nu)^2$ for the four radio sources. The RM represents the slope of the law in Eq. 2 while the intercept is the intrinsic polarization angle.

Chapter 2

Bayesian inference of Faraday rotation images

We want to interpret the Faraday rotation images of the radio galaxies in A2255 using the single-scale model outlined in Chapter 1. Specifically, we want to estimate the model free parameters: Λ_C, the cell's width, B_0, the magnetic field at the cluster center, and η, the slope of the magnetic field scaling with gas density. The remaining model parameters, β, r_C, n_0, and L, are considered fixed to the following values: $\beta = 0.74$, $r_C = 432$ kpc, $n_0 = 2.05 \times 10^{-3}$ cm^{-3}, and while L, i.e. the position along the line-of-sight, is valued for all the four sources at the cluster mid-plane.

In our attempt, we will made use of the Bayes' theorem, which relates the conditional and marginal probabilities of model M given the data D, provided that the probability of D does not equal zero:

$$P(M|D) = \frac{P(D|M)P(M)}{P(D)}. \tag{2.1}$$

The theorem express the *posterior probability* $P(M|D)$ of the model given the data in terms of the product between the *likelihood function*, $P(D|M)$ and the *prior probability* of the model, $P(M)$. The term $P(D)$ is called the *evidence* and acts as a normalizing constant.

The likelihood is a function of the model parameters so that, given some observed outcomes, is equal to the probability of those observed outcomes given those parameter values[1]. The prior represents our belief about the probability of a given configuration of model parameters before the data has been taken. The evidence can be viewed as the average likelihood weighted

[1]Note, however, that the likelihood function is not a probability density function, i.e. in general its integral is not equal to 1.

by the prior over the parameter space:

$$P(D) = \int P(D|M)P(M)dM. \tag{2.2}$$

In our problem the data D is represented by the four RM images with their relative error images while the model M is represented by any combinations of the free parameters Λ_C, B_0, and η. Hence, in order to apply the Bayes theorem we must specify the likelihood function and the prior probability distribution function (pdf).

2.1 The two-points RM likelihood function

In the framework of the single-scale model we expect that the probability of a certain RM value at a given position $\vec{r}_\perp = (x, y)$ is normal distributed around zero with a dispersion given by Eq.5.

Lets consider to probability of the Faraday rotation values at two distinct position in the cluster, RM_1 at position (x_1, y_1) and RM_2 at position (x_2, y_2). More specifically, we want to consider the probability of the difference between the RM at these location, $P(RM_1 - RM_2)$. We can identify to extreme situations. If the separation between the two points $dr = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ is much smaller than Λ_C, we expect that probability for the RM difference is:

$$dr \ll \Lambda_c \rightarrow P(RM_1 - RM_2) = 0,$$

since both line-of-sights will likely intersect the same set of magnetic field cells across the cluster. On the other hand, if the separation between the couple of points is much larger than the cell size, $dr \gg \Lambda_C$, we expect that the two RM values originate from distinct random walks and hence the probability for their difference is still normal:

$$P(RM_1 - RM_2) = \frac{1}{\sqrt{2\pi\sigma_{12}^2}} e^{-\frac{(RM_1 - RM_2)^2}{2\sigma_{12}^2}}$$

where $\sigma_{12}^2 = \sigma_{RM1}^2 + err_{RM1}^2 + \sigma_{RM2}^2 + err_{RM2}^2$.

Now we consider all the four RM images in A2255 and we extract a large set of couples of pixels with baselines both inside each source both between different sources (see Fig.2.1). These set of couple forms our data D, and we consider the likelihood function as the product of probabilities for the difference of the RM between all these locations:

$$P(D|M) = \Pi_{i \neq j}\{P(RM_i - RM_j)\}. \tag{2.3}$$

2.2 The n-simplex Bayesian solver

In this section we describe the algorithm we used to solve the Bayes' theorem in Eq. 2.2.

The `FARADAY` software package implements a Bayesian solver based on the griding of the parameters space with a finite set of n-simplexes. Specifically, an n-simplex is an n-dimensional polytope which is the convex hull of its $n+1$ vertexes. For example, a 2-simplex is a triangle while a 3-simplex is a tetrahedron, etc.. The algorithms exploits two stages: an initial griding and a refining stage.

2.2.1 Initial griding

The algorithm starts with an initial sampling of the posterior over the parameters space. A number of points is extracted randomly and a n-dimensional Delaunay triangulation is performed [3, 4]. Given a set of points in the space, a Delaunay triangulation is such that no point of the set is inside the hyper-circumsphere of the n-simplexes composing the grid.

The product of the likelihood and the prior, f_i, is evaluated at the vertexes of each n-simplexes, and the δ-evidence of the simplex is calculated by the "trapezoidal rule":

$$\delta P(D) = V \times \frac{1}{n+1} \sum_{i=1}^{n+1} f_i, \tag{2.4}$$

where V is the volume of the n-simplex. A first, approximated, value for the evidence is then obtained by adding the contributes of all the n-simplexes in the grid.

2.2.2 Refinement stage

After the initial griding is completed, we have a rough knowledge of the posterior. Furthermore, we can calculate the probability for each region of the parameters space enclosed in each n-simplex. We then proceed by refining the grid until the evidence value converges to the desired fractional precision. `FARADAY` offers different refinement schemes. We experimented the "*maximum probability*" and the "*random*" schemes. In the former, the n-simplex with the highest probability is selected for refinement. In the latter,

a random extraction is performed from the cumulative distribution function. In both cases, a new point is inserted in the grid at the position of the geometric center of the selected simplex. The Delaunay triangulation is updated and a refined posterior is calculated. The two schemes are characterized by different pro and cons. The "max probability" refinement offers a faster convergence but it is more susceptible to the presence of secondary peaks. The "random" scheme is slower but less susceptible to the presence of secondary peaks since it explores also low-probability regions of the posterior distribution.

A 2-dimensional example of the n-simplex Bayesian solver is shown in Fig. 2.2. In this test we fixed the value of the magnetic field cell size to $\Lambda_c = 68$ kpc and we find the posterior distribution of the parameters B_0 and η. The initial griding (performed with 5 points) is shown in the top-left panel. The refined gridding is shown in the bottom-right panel. The 1-dimensional marginalization for the model parameters are shown in green. Here we used the "random" refinement scheme in which the simplexes to be re-sampled are extracted randomly from the posterior cumulative distribution shown in the insets of each panel.

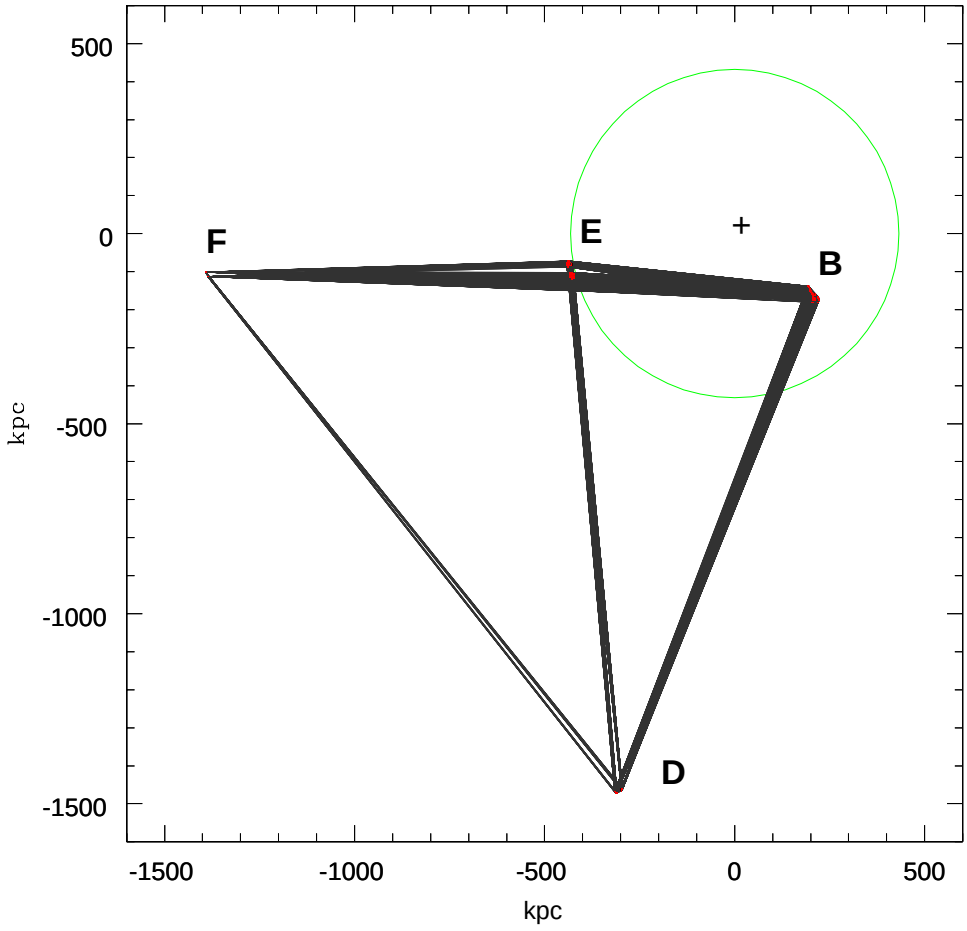

Figure 2.1: Location of the radio sources with respect to the cluster center (cross). The lines represent the trans-source baselines between the couple of pixels (red points) used to compute the RM likelihood. The green circle is the cluster core radius.

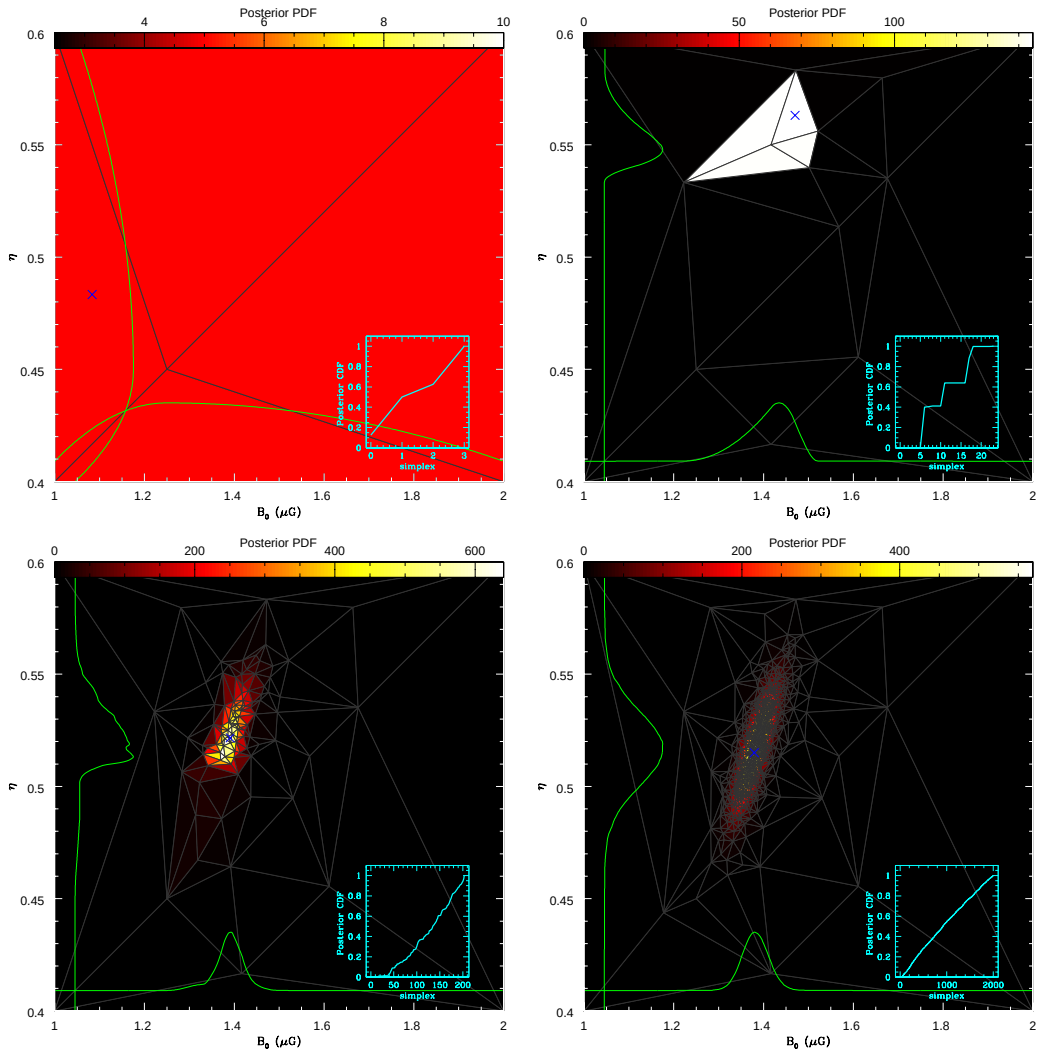

Figure 2.2: Example of 2-dimensional Delanauy triangulation of the $B_0 - \eta$ plane. The cross indicates the position of the maximum posterior. The green lines are the 1-dimensional marginalization for the posterior distribution of the model parameters. The cumulative distribution function of the simplexes probabilities is shown in the insets. The panels represent four different level of grid refinement.

Chapter 3

Results

We used the algorithm described in the previous section to derive the posterior distribution of the model parameters: B_0, η, and Λ_C. The RM likelihood is calculated between 1964 pixel pairs. We selected one pixel every five to ensure that the RM values are effectively independent (the resolution of the radio images is sampled at 4 pixel per beam). The final 3-dimensional grid is composed of about 13000 tetrahedrons.

We assumed uniform priors for all the three free model parameters. The maximum posterior parameters are reported in Table 3.1. The reported uncertainties refer to the boundary of the region of the posterior that includes the 68% probability of the model parameters.

The magnetic field value at the cluster center is of about 1.3 μG. Since $\eta \simeq 0.5$, the magnetic field energy scales as the gas density ($B^2 \propto n_{gas}$) hence the field weakens in the cluster periphery. At the distance of source D, namely 1570 kpc from the cluster centre, the intracluster magnetic field strenght is as low as 0.28 μG.

The magnetic field cell size is about 68 kpc. We note that the corresponding size of RM patches is $\Lambda_{RM} = \Lambda_C/\pi \simeq 22$ kpc (see [2]), i.e. comparable to size of the radio sources. This value is consistent with the large offset of $\langle RM \rangle$ observed for all the four radio sources.

Finally, in Fig. 3.1 we show the observed trend of σ_{RM} along with the prediction of the single-scale magnetic field model. The model prediction is obtained as a marginalization of the posterior distribution and provide a good description of the data.

Table 3.1: Maximum posterior parameters.

Parameter	Maximum PDF	units
B_0	1.34 ± 0.11	μG
η	0.52 ± 0.04	
Λ_C	68 ± 7.5	kpc

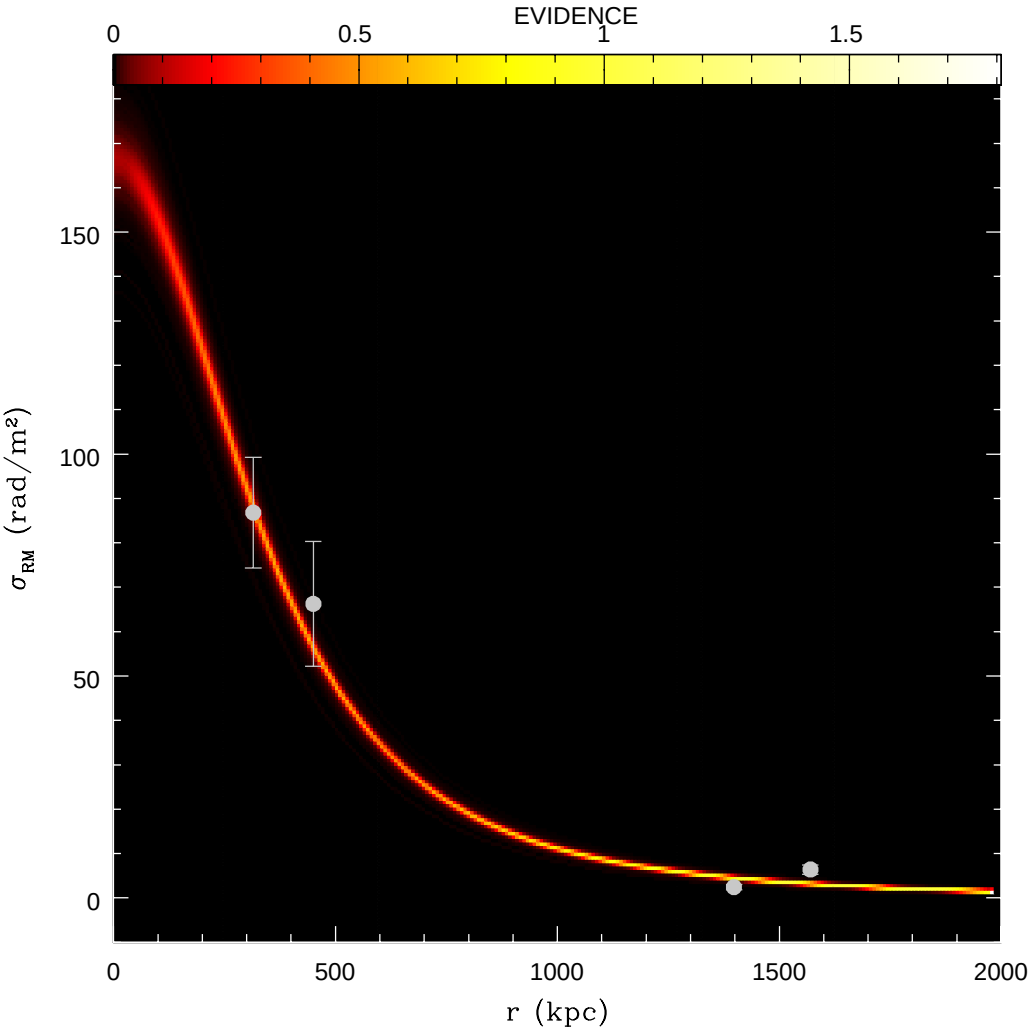

Figure 3.1: RM dispersion as a function of the projected distance from the cluster center for the four radio galaxies (dots). The shaded region represents the trend expected on the basis of the posterior distribution for the model parameters.

Bibliography

[1] Govoni, F., Murgia, M., Feretti, L., et al. 2006. *The intracluster magnetic field power spectrum in Abell 2255*, 460, 425

[2] Murgia, M., Govoni, F., Feretti, L., et al. 2004. *Magnetic fields and Faraday rotation in clusters of galaxies*, A&A, 424, 429

[3] Adrian Bowyer (1981). *Computing Dirichlet tessellations, The Computer Journal*, 24(2):162166

[4] David F. Watson (1981). *Computing the n-dimensional tessellation with application to Voronoi polytopes, The Computer Journal*, 24(2):167172.