

Magnetic Fields in High Velocity Clouds and Rotation Measure Foreground Removal Techniques

Olivia Walters

Supervisors: Craig Anderson, Naomi McClure-Griffiths

An honours thesis submitted for
The Australian National University
Research School of Astronomy and Astrophysics

September 2024

© Olivia Walters 2024

Except where otherwise indicated, this thesis is my own original work.

Olivia Walters
11 September 2024

Acknowledgments

I would like to thank my primary and secondary supervisors, Dr. Craig Anderson and Prof. Naomi-McClure Griffiths for being the main body of support, data, and ideas when writing this thesis. I would also like to acknowledge Prof. Brian Gaensler and Dr. Lyla Jung for providing further aid in the process of analysis, and their support in regards to POSSUM itself. The ASKAP POSSUM team more broadly has come as a great aid and source of community during the production of this thesis.

Fellow radio astronomer Callum Lynn has also been a great source of advice and help in constructing this thesis for a more general astronomical audience, along with his guidance about the honours and PhD process.

The ANU SSN Consultancy Group was a crucial part in the validation of the statistical techniques listed in this paper, including the suggestion to use the Weighted ANOVA and Tukey HSD test.

I would also like to thank my grandmother, Hilda Filipovic, for being a very important person in my life, and a rock on which I relied on in the process of completing my honours year.

Abstract

NB: This abstract is old! It should not be looked at!

High Velocity Clouds (HVCs) are a proposed solution to how extragalactic gas enters into star-forming galaxies. However, the presence of magnetic fields is required to ensure that HVCs can travel through the halo without being torn apart by ram pressure. This report aims to measure the strength of the magnetic fields present within the Milky Way's halo and surrounding circumgalactic HVCs, to determine if they can support HVCs as they travel through the Galactic halo. This report uses data from the ASKAP Polarisation Sky Survey of the Universe's Magnetism (POSSUM) to pull a series of rotation measures (RMs) across the southern sky. These RMs are then converted into magnetic fields using an algorithm, developed using previously collected data from the Smith Cloud. Statistical analysis tools are applied to confirm if hypothesised magnetic fields can support HVCs. The report finds that [Summary of discussion and conclusion] ...

Contents

Acknowledgments	iii
Abstract	iv
1 Introduction	1
1.1 High Velocity Clouds	1
1.1.1 Chemical Properities and Emission	2
1.2 Magnetic Fields	3
1.2.1 Draping	3
1.2.2 Faraday Rotation	4
1.2.2.1 Noise Interference	5
1.3 Smith Cloud	7
1.4 Report Outline and Objectives	7
2 Data Collection	10
2.1 The Australian Square Kilometer Array Pathfinder (ASKAP)	10
2.2 Observational Data	11
2.2.1 RM Sky Interpolation	11
2.2.2 Other Data Sources	12
2.3 Selection of HVCs	12
2.4 Data Collation	14
2.4.1 HVC Image Overlays	14
3 Foreground Subtraction	16
3.1 Interpolation	16
3.2 Annulus Subtraction	17
3.3 Fast Fourier Transforms (FFTs)	18
3.3.1 Non-Uniform Fast Fourier Transforms (NUFFTs)	19
3.3.2 Bandpass Filtering	20
3.3.3 Kernel Filtering	22
3.4 Characterising the RM data	23
3.5 Statistical Comparison of Foreground Removal Methods	26
3.6 Other Methods	28

4	Magnetic Field Derivation	29
4.1	Assumptions	29
4.2	Derivation of Line-of-Sight Magnetic Fields	30
4.3	KS Testing	33
4.4	Mathematical methods to evaluate HVC Magnetic Fields	34
4.4.1	KS-EDF Method	36
4.4.2	Uncertainty Subtraction	37
4.5	Magnetic Field Results	38
4.5.1	Uncertainties	39
4.5.2	Statistical Comparison of Methods	39
5	Discussion	41
5.1	Foreground Removal	41
5.2	Magnetic Field Derivation	43
5.2.1	Data Collection	44
5.2.2	Necessity of Assumptions	45
5.2.3	Validity of Detection Methods	46
5.2.4	Validity of Derivation Methods	48
6	Conclusions	50
	Bibliography	57
	Appendix	58
A	Developed code and data	58
B	All HVCs	58
C	Planck Mission Cosmic Microwave Background	59
D	PyNUFFT Python Module	59
E	Statistical ANOVA Tests	60

List of Figures

1.1	From Konz et al. [2002], figure 2. An example of a typical HVC shape and structure, specifically that of HVC125+41-207. The contour lines and shading indicate HI column density.	2
1.2	From Loi et al. [2019], figure 4. A graph of the relationship between the sensitivity at 1.4 GHz (x-axis) and the minimum number of RM sample points per square degree (y-axis). The black points are not relevant to the report, but the purple line and equation describe the determined relationship.	6
1.3	From Macquart et al. [2012], figure 1. A graph displaying the effect of observational (stokes parameters) signal-to-noise ratio the resultant faraday depth on a sample set of observations.	8
1.4	From Lockman et al. [2008], figure 1. A HI image of the Smith Cloud taken from the Green Bank Telescope at a local standard of velocity rest of 100 km s^{-1} . The purpose of the arrows are not meaningful to this paper.	9
2.1	An Aitoff projection of all portions of the RM sky observed by ASKAP for the POSSUM survey. The map is up to date as of May 2024 and all RMs in this map are used in the production of this report. The faraday depth of these RMs are indicated by colour.	11
2.2	Cartesian plots of peripheral data maps (left): the HI sky from the Westmeier [2018] modification of the HI4PI survey (top); the H-alpha sky from Finkbeiner [2003] (middle); and the Hutschenreuter map from Hutschenreuter and Enßlin [2020]; Hutschenreuter et al. [2022] (bottom). Uncertainty maps for H-alpha and the Hutschenreuter map are displayed respectively (right). The HI uncertainty is not displayed due to it simply being a scalar multiple of the HI map.	13
2.3	(Left) A graph of all ~ 180000 RMs plotted against their corresponding Galactic latitude; (Right) The corresponding graph of all Hutschenreuter map faraday depths matched to the POSSUM RMs. Both graphs represent a significant level of scatter present in RMs collected near the Galactic midplane.	15

3.1	An Aitoff projection of all portions of the RM sky observed by ASKAP for the POSSUM survey similar to 2.1, filtered through a proprietary fixed-sampling algorithm with inner radius of 0.4" and a sampling constant of 50.	18
3.2	(Right) A cropped and grayscale image of the CMB. (Left) The same image after being fed through a forward then adjoint NUFFT. The amount of sample points total to 27000. Both plots share a common intensity colourbar for reference.	20
3.3	Crosshatch-Bandpassed versions of the interpolated RM sky at various opacity gradings - where "original" means 0% opacity. The term "opacity" refers to the effect of the bandpass e.g. 100% opacity means the bandpass completely removes selected frequencies and 50% opacity means that the bandpass halves the presence of selected frequencies. . .	21
3.4	Cartesian plots of the interpolated RM sky (Top) compared against the Annulus-Bandpassed version of the interpolated RM sky (Bottom). The annulus kernel used sums to unity, making it act like a unitary operator, standard for blurring techniques [Pulfer, 2019].	22
3.5	Residual histogram plots of the actual POSSUM RMs compared to the corresponding various foreground removal methods: interpolation (Top left); crosshatch-bandpassed (Top right); annulus-bandpassed (Bottom left); and annulus-convolved (Bottom right).	24
3.6	Three corner plots of all RM sampling points, describing the relationship between the several associated RM values and Galactic latitude indicated by colour. Here, a comparison between actual RMs, interpolated RMs, and Crosshatch-Bandpassed RMs is displayed.	25
3.7	Three corner plots of all RM sampling points, describing the relationship between the several associated RM values and Galactic latitude indicated by colour. Here, a comparison between actual RMs, interpolated RMs, and Annulus-Convolved RMs is displayed.	25
3.8	Three corner plots of all RM sampling points, describing the relationship between the several associated RM values and Galactic latitude indicated by colour. Here, a comparison between actual RMs, Annulus-Bandpassed RMs, and Annulus-Convolved RMs is displayed.	26
3.9	The sample of all RMs corrected with the Hutschenreuter Interpolation (Blue) and Crosshatch-Bandpassed Interpolation (Red). A bootstrap fitted t-distribution is drawn on the graph for each respective residual RM set.	27

4.1	All 13 HVCs used in the analysis of the primary outcome. The HI column density is represented using a greyscale image background. The RMs are represented by circular markers, their size equal to the magnitude and the colour representative of their sign with Red being positive. The black circle is the deliniation between in-HVC and out-HVC populations, and the black 'x' indicates the centre of the HVC.	31
4.2	The same set of HVCs in HI with RM overlays, displayed the same as in 4.1. However, instead of using the interpolation as a method of correction, the mean uncorrected RM value is subtracted out.	32
4.3	Two histograms describing the RM sample within a certain distance to the 13 sample HVCs. The red histogram uses uncorrected RMs. The green histogram uses corrected RMs.	34
4.4	A diagram representing the physical evaluation of the KS-EDF method. The graph shows the EDFs of inside and outside populations for an example HVC. The lines drawn on the diagram represent the statistic value and the KS-EDF magnetic field value.	36
4.5	A boxplot representing the absolute-value magnetic field of each HVC for both derivational methods. The red and blue dotted lines indicate two upper bounds, one set by the Grønnow et al., simulations, and the other by the Smith Cloud respectively.	39

List of Tables

3.1	(Top) Student's t-distribution fit parameters for the residual histograms as displayed in figure 3.9. (Middle) Results of the χ^2 test comparing residual histograms with the t-distribution. (Bottom) R^2 statistics comparing the Hutschenreuter map with the other methods detailed in this section.	27
4.1	A table describing the KS test results for each HVC.	35
4.2	A table describing the magnetic field derivations for each HVC. HVCs from the sample of 13 that had no significant KS test detection, or an invalid variance subtraction result are removed.	38
4.3	A table describing the Tukey Multiple Comparison of Means test results for the HVCs with magnetic field detection.	40

Introduction

The question of how gas is accreted into galaxies fuelling star-formation is a puzzle that has perplexed Astronomers for decades. Due to the complexities in the structures of star-forming galaxies, there are many factors involved in the process of, and potential sources of accretion. What Astronomers do know, at least, is that star-forming galaxies require a continuous supply of fresh gas to continue their star formation.

Due to observational constraints, Astronomers are required to attempt to answer this question by examining the behaviours of our own Milky Way and Local Group environment, assuming the Milky Way is typical of a star-forming galaxy.

A major factor to consider when answering this question is where fresh pristine gas comes from, and by what mechanism it takes to enter the disks of star-forming galaxies. High Velocity Clouds (HVCs) have been a suggested mechanism for Galactic gas accretion and this report aims to evaluate the viability of this suggestion.

1.1 High Velocity Clouds

HVCs are clouds of gas found in the Milky Way's Circumgalactic Medium (CGM) and Galactic halo. They have a high peculiar velocity relative to the Galactic Standard of Rest (GSR), typically $70\text{--}90 \text{ km s}^{-1}$ [Wakker, 1991; Wakker and van Woerden, 1997; Blitz et al., 1999]. As will be shown in section 1.2.1, this increased speed, and its interaction with halo magnetic fields, is hypothesized to allow the HVC to survive as it travels through the CGM and halo so it can reach the Galactic disk and Interstellar Medium (ISM) of the Milky Way.

There are a few hypotheses as to where HVCs originate. Blitz et al. [1999] suggests that HVCs originate from the Intergalactic Medium (IGM) surrounding the local group. However, there is also a belief that some HVCs likely 'tore off' from satellites like the Magellanic Clouds, due to the presence of their own dark matter subhaloes, and the existence of HVCs in the Magellanic Stream and Leading Arm [Kaczmarek et al., 2017; McClure-Griffiths et al., 2010]. Recent simulations from TNG50 have indicated that HVCs mainly have their origins in the warm-hot CGM [Lucchini et al., 2024].

HVCs typically have a neutral mass gas content of $3 \times 10^7 M_{\odot}$ and a dark matter content of $3 \times 10^8 M_{\odot}$ [Blitz et al., 1999]. They are generally shaped like comets, with



Figure 1.1: From Konz et al. [2002], figure 2. An example of a typical HVC shape and structure, specifically that of HVC125+41-207. The contour lines and shading indicate HI column density.

a primary bulb that is approximately 0.5-25 kpc in diameter - a value that is highly dependent on distance to the Galactic midplane, which can range from approx. 50 kpc - 1 Mpc [Blitz et al., 1999; Konz et al., 2002]. Furthermore, HVCs have tail-like structures that account for one eighth the baryonic mass of the HVC [Konz et al., 2002]. These tails leave behind long streams of gas that remain after collision with the Galactic disk [Putman et al., 2012]. This size-distance correlation, comet-like structure, and long streaming tails suggest quite conclusively that HVCs shed a lot of material as they make their journey to the ISM.

Figure 1.1 is from Konz et al. [2002], which provides a typical example of what a HVC looks like in HI, specifically using the example of HVC125+41-207.

1.1.1 Chemical Properties and Emission

Due to their hypothesised origins in extragalactic gas, HVCs contain mostly hydrogen gas such as HI, which can be seen with 21 cm emission [Wakker, 1991; Wakker and van Woerden, 1997; Westmeier, 2018]. The proportion of ionised gas in HVCs is still heavily debated. HVCs also typically contain molecular gas that emits H-alpha, however due to extinction effects, it is difficult to observe H-alpha emission correctly [Bland-Hawthorn and Maloney, 1999; Finkbeiner, 2003]. Another difficulty for H-alpha analysis is that HVCs typically do not contain CO in significant quantities, CO a common tracer for H-alpha [Blitz et al., 1999].

Despite HVCs supposed gas purity, there is evidence that HVCs can contain alpha group elements. Hill et al. [2009]; Madsen et al. [2006] found the presence of [NII] 6583Å, [SII] 6716Å, and [OIII] 5007Å emission lines – with a conclusion that Nitrogen abundance is 0.15-0.44 times solar abundance levels. The observation of these emis-

sion lines can help constrain the metallicity of any HVC [Hill et al., 2009]. Metallicity is important in both answering the capacity for HVCs to supply fresh gas for star formation, and the mechanism by which HVCs can survive [Grønnow et al., 2018]; more on the latter in section 1.2.1. While HVCs can contain heavier elements Hayakawa and Fukui [2024] finds that these concentrations are low enough that HVCs can remain as viable candidates for fuelling star formation via gas accretion.

HVCs additionally have a temperature relationship with proximity to the Galactic midplane, with an average HVC temperature of 10000 K and a range of temperatures ranging from 8000 – 12000 K [Hill et al., 2009; Madsen et al., 2006]. The temperature of a HVC is in the region in which atomic hydrogen transitions from neutral (HI) to ionised (HII), suggesting that HVCs may be partly ionised [Hill et al., 2009; Madsen et al., 2006; Kawaguchi, 1952]. This temperature relationship is dependent on their position with respect to the Galactic midplane, with HVCs closer to the midplane generally being cooler [Madsen et al., 2006].

1.2 Magnetic Fields

The primary issue facing HVCs as an explanation for gas accretion is its capacity to survive as it travels through the CGM and Galactic halo. As discussed, in section 1.1, HVCs can loose a lot of size and mass as it approaches towards the Galactic disk, with the long trails it leaves behind being evidence for ram-pressure stripping as the HVC collides with the gas present in the halo [Jones et al., 1996; Grønnow et al., 2017, 2022]. Heitsch and Putman [2009] demonstrates that without anything to counter this effect, HVCs with masses under $10^{4.5} M_{\odot}$ would completely disperse within 10 kpc of halo travel.

Additionally, HVCs are subject to Kelvin-Helmholtz (K-H) instabilities, which is triggered by the nature of the HVC being a cloud of warm gas travelling at high speeds through a medium. These K-H instabilities are a significant factor that would lead a to HVC collapsing before it reaches the Galactic disk [Jones et al., 1996; Grønnow et al., 2017, 2022].

1.2.1 Draping

The proposed solution to handle this problem is magnetic fields. The Galactic halo is magnetised to some degree in a generally turbulent manner [Mao et al., 2010; Han and Qiao, 1994; Jung et al., 2023; Beck et al., 2012]. It is hypothesised that HVCs accumulate these existing magnetic fields in the Galactic halo, causing them to cloak the HVC with a shield that protects against ram pressure stripping and supresses K-H instabilities [Dursi and Pfrommer, 2008; Jones et al., 1996; Konz et al., 2002; Grønnow et al., 2017, 2018; Jung et al., 2022]. This phenomenon is referred to as ‘magnetic

draping’.

There is not enough observational evidence to support the magnetic draping hypothesis. Past research, instead, was focused on: technological improvements to surveys [Gaensler et al., 2010, 2024; Vanderwoude et al., 2024; Moss et al., 2013; Westmeier, 2018; Taylor et al., 2009; Finkbeiner, 2003; Hutschenreuter and Enßlin, 2020; Hutschenreuter et al., 2022]; the analysis of physical phenomena such as ram pressure stripping [Jones et al., 1996; Grønnow et al., 2017, 2022]; the development of magnetic field derivation techniques [Betti et al., 2019; Grønnow et al., 2017; Mao et al., 2010; Jones et al., 1996; Hill et al., 2013; Schnitzeler, 2010]; and simulations of HVCs [Konz et al., 2002; Grønnow et al., 2017, 2018, 2022; Jung et al., 2022] – all of which lay the groundwork for a proper investigation of magnetic draping.

Previous and recent simulations involve reports produced by Grønnow et al. [2017, 2018, 2022] (henceforth referred to as the “Grønnow et al., simulations”) that provide detailed insight onto how a magnetic draping protects HVCs from collapse, complementing earlier work by Konz et al. [2002]; Jones et al. [1996]. It is shown from the Grønnow et al., simulations that magnetic fields of about 0.3-1 μG can provide stability to HVCs.

However, increasing magnetic field strength beyond a certain threshold can result in less stability; magnetic fields can accelerate the effects of Rayleigh-Taylor (R-T) instabilities and the magnetic pressure applied by the draped fields can also slow down a HVC to the point that it no longer is fast enough to sweep up these magnetic fields. From the Grønnow et al., simulations, the upper threshold where these effects start increasing instability is about 1 μG (specifically stating a maximum of approx. 3 μG), thus HVCs should ideally have a “Goldilocks” magnetic field strength on the order of magnitude of 0.1 μG , with 1 μG being too high, and 0.01 μG being too low.

The effectiveness of magnetic draping is affected by the morphology of these fields and the physical properties of the HVC. The Grønnow et al., simulations state that both the orientation of the magnetic field with respect to the direction of motion of the HVC, and where the magnetic field is located are important considerations. It is expected that HVC is not entirely covered in a magnetic field, only the part that is front facing in the direction of travel. While it is possible to draw conclusions about the survivability of a HVC from the strength of the magnetic field, modelling is required to confirm the accuracy of such conclusions [Betti et al., 2019]. The Grønnow et al., simulations also predict that metallicity can affect the HVC’s survivability, with high-density metal-rich clouds and low-density metal-poor clouds being more unstable than their counterparts.

1.2.2 Faraday Rotation

Magnetic fields cannot directly be imaged by a telescope. Instead, researchers can use the phenomenon of Faraday Rotation to quantify the line-of-sight magnetic field

strength. Low-frequency polarised radiation tends to rotate as it travels through a medium with a magnetic field present. Thus, by recording the stokes parameters of incoming light from distant radio sources, one can derive the Rotation Measure (RM) of incoming radiation, which is a statistical quantifier of Faraday Rotation [Vanderwoude et al., 2024; Brentjens and de Bruyn, 2005].

Vanderwoude et al. [2024] describes the method by which this report's main source, the Polarisation Sky Survey of the Universe's Magnetism (POSSUM), obtained its RMs from raw stokes parameters - which involves a complex fourier analysis technique to evaluate the rotation measure. There is a direct connection between RM and line-of-sight magnetic field strength, quantified by the below equation [Betti et al., 2019; Vanderwoude et al., 2024; Hill et al., 2013; Kaczmarek et al., 2017; Mao et al., 2010].

$$\text{RM} = 0.812 \int_{s_{\text{observer}}}^{s_{\text{source}}} \frac{n_e(s)}{\text{cm}^{-2}} \frac{B_{\parallel}}{\mu\text{G}} \frac{ds}{\text{pc}} \text{radm}^{-2} \quad (1.1)$$

In which, B_{\parallel} is the magnetic field strength, RM is the Faraday Depth a.k.a. the value of the RM^1 , and n_e is the electron density of the medium as a function of line-of-sight distance s . The analysis of this equation, its solutions, and the use of it in calculating the magnitude of draped magnetic fields is discussed in section 4.

Faraday Rotation also occurs in the ISM, due to the slight magnetisation of the ISM [Ferrière, 2001; Mao et al., 2010; Schnitzeler, 2010]. This magnetisation is anti-symmetric with galactic longitude [Mao et al., 2010]. Hence it is also important to remove the foreground from RM observations.

1.2.2.1 Noise Interference

When making radio observations of RMs, a principal factor to consider is signal to noise ratio and detector sensitivity. Radio sources tend to appear in the field after high exposure times as point-like sources. These point-like sources are then collated into an "RM grid" which has a particular density measured in sample points per square degree. The sensitivity of a detector determines the number of source points observed as seen in figure 1.2 [Loi et al., 2019]. At grid densities that are too high, RM sources can bleed into each other, leading to confusion noise. However, no modern radio telescope can reach the confusion limit [Loi et al., 2019].

Signal to noise is of primary concern when measuring the effect of Faraday Rotation. At low enough signal-to-noise ratios, it is possible to encounter 'phantom RMs' which do not accurately represent the real RM [Macquart et al., 2012]. Figure 1.3 gives a visual illustration of this phenomenon. The issue in question is that any observation of RM grids is going to dip below the signal-to-noise threshold of approx. 6, which

¹For the purposes of this report, rotation measure and faraday depth are treated as equivalent, despite their subtle differences

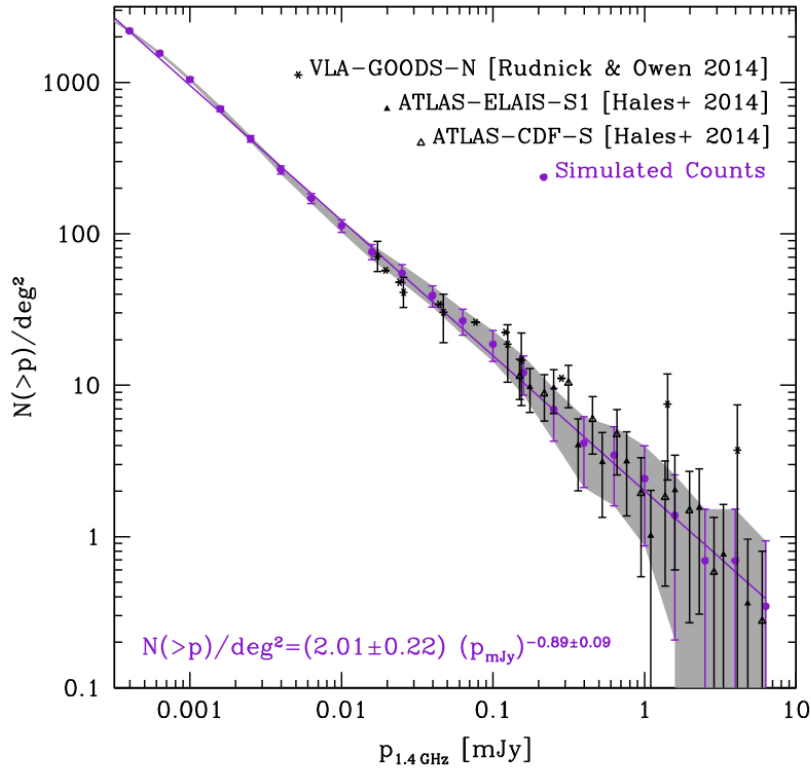


Figure 1.2: From Loi et al. [2019], figure 4. A graph of the relationship between the sensitivity at 1.4 GHz (x-axis) and the minimum number of RM sample points per square degree (y-axis). The black points are not relevant to the report, but the purple line and equation describe the determined relationship.

can introduce an intrinsic scatter in collected RM grid data. There is not much that can be done outside of avoiding regions where the signal-to-noise is likely to be lower (i.e. the Galactic midplane), without modifying the RM grid in a manner that can interfere with its statistical qualities. However, POSSUM itself obtains RM grid data that has signal-to-noise above 8 [Gaensler et al., 2024].

1.3 Smith Cloud

The Smith Cloud is a large HVC that is in the process of colliding with the Galactic disk [Lockman et al., 2008; Tepper-García and Bland-Hawthorn, 2017; Lockman, 2008]. Unlike most HVCs it is quite large in both mass (at least $10^6 M_\odot$ in HI mass) and angular size (the main bulb covering an area of approx. 144 square degrees) [Lockman et al., 2008; Tepper-García and Bland-Hawthorn, 2017; Lockman, 2008]. It has a predicted physical size of 3 square kpc, which is large for a HVC close to the Galactic disk [Lockman et al., 2008]. Due to its proximity and size, the Smith Cloud has been used as a source point of analysis for most of the properties already discussed. For example, metallicity tracers and alpha-group elements in Madsen et al. [2006]; Hill et al. [2009] were determined by analysing the Smith Cloud. Figure 1.4 provides an image of the Smith Cloud in HI from Lockman et al. [2008].

The Smith Cloud has additionally already had its magnetic draping effect analysed. Simulations by Grønnow et al. [2017] and observations by Hill et al. [2013] both agree on an effective magnetic field of $\sim 8 \mu\text{G}$. Note that this number is well above the Goldilocks zone mentioned in section 1.2.1 – indicating the potential exceptionality of the Smith Cloud.

While there are other HVCs that have been sampled in the past, the Smith Cloud has been the main source of HVC information. This is a problem, as the Smith Cloud is an outlier amongst most HVCs – evidenced by its unusual size and magnetic field strength, both factors being related. There is a necessity to analyse more typical nearby HVCs to gain an understanding of the effects of magnetic fields.

1.4 Report Outline and Objectives

The primary objective of this report is to (a) construct a rudimentary algorithm for evaluating the presence of magnetic draping in HVCs found in the CGM and halo and, (b) use this base algorithm to come up with a *very* rough estimate for the magnetic field strength surrounding typical HVCs. The primary source of data will be POSSUM, which will be used to obtain the primary outcome of the report. Due to the previous lack of robust RM grid data, along with other periphery data, and a large level of required assumptions, it is not possible to come up with an estimate of field strength for any HVC, that is accurate to within an order of magnitude. The



Figure 1.3: From Macquart et al. [2012], figure 1. A graph displaying the effect of observational (stokes parameters) signal-to-noise ratio the resultant faraday depth on a sample set of observations.

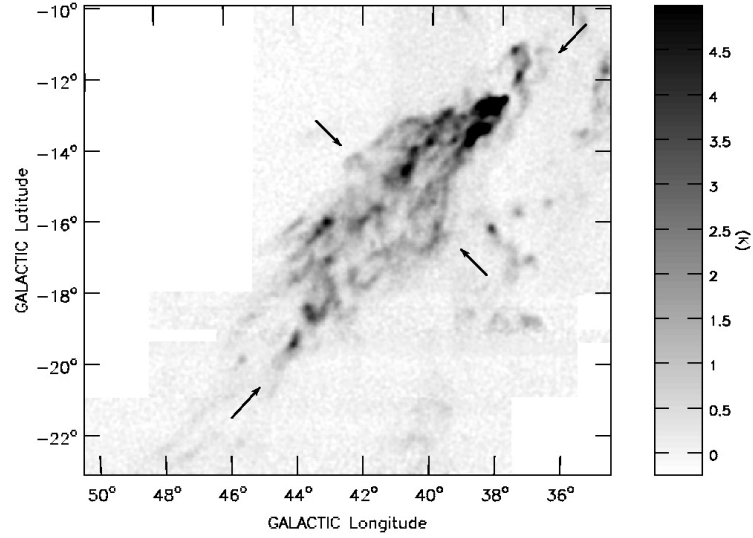


Figure 1.4: From Lockman et al. [2008], figure 1. A HI image of the Smith Cloud taken from the Green Bank Telescope at a local standard of velocity rest of 100 km s^{-1} . The purpose of the arrows are not meaningful to this paper.

best-case scenario is an estimate that is accurate within a single order of magnitude. Hence, why the aim of this report is to lay the foundations for a future, more detailed analysis, of more numerous HVCs (up to 1693 entries as discussed in section 2.3).

There is a secondary objective that is integral to completing the primary objective: that of foreground removal. Past research, specifically within the analysis of HVCs, has relied on the technique of interpolation to obtain a RM foreground to use in corrections [Schnitzeler, 2010; Moss et al., 2013; Hill et al., 2013]. However, as researchers move onto measurements of magnetism that demand more accuracy, foreground removal needs to equally match that that growing need for accuracy – thus this report also aims to investigate avenues for improving foreground subtraction techniques.

This report is split into six sections plus an appendix. Section 2 describes the process of how and which data was obtained to achieve the research question, and how this data was collated together. Section 3 summarises the investigation into foreground removal, which is the secondary aim of the research. Section 4 describes how the magnetic fields for HVCs were derived, and the level to which morphology was involved in the analysis. Section 5 discusses the viability of the methods described in the report, along with broader scientific and statistical considerations. Lastly, section 6 concludes and outlines the many possible directions for future research. Appendix A lists information on how to obtain the program data and algorithms used in research.

Data Collection

To achieve the goals listed in 1.4, several sources of data are required. Furthermore, work needs to be done to collate the data together to measure the magnetic field, quantify the efficacy of foreground removal techniques, and to eliminate problematic data.

2.1 The Australian Square Kilometer Array Pathfinder (ASKAP)

ASKAP represents a recent development in the progress of the field of radio astronomy. It is a part of a new generation of southern-hemisphere telescopes built with the aim to build towards the Square-Kilometer Array (SKA) [Hotan et al., 2021; Gaensler et al., 2010, 2024]. ASKAP aims to allow for the progression of research in the “Pre-SKA” era. As mentioned in section 1.4, the importance of preliminary research in the Pre-SKA era is meant to allow for a more efficient process of astronomy in the coming years when the SKA is fully operational. This means that the data and methodology is simultaneously rudimentary and advanced compared to previous projects.

POSSUM is an ongoing project using ASKAP with the aim of measuring the RM southern sky. The benefit of using ASKAP as opposed to previous RM sky surveys is the RM grid density ASKAP can provide [Gaensler et al., 2010, 2024; Hotan et al., 2021]. Previous surveys such as the NRAO VLA Sky Survey (NVSS) were only able to record an RM grid density of 1 sampling point per square degree, with POSSUM providing a density 30 times greater [Vanderwoude et al., 2024; Gaensler et al., 2010, 2024; Hotan et al., 2021; Taylor et al., 2009]. The higher grid density allows for the analysis of regular-sized HVCs in the CGM, as opposed to only the largest HVCs like the Smith Cloud.

Additionally, POSSUM is set to cover a region of the southern sky that has seldom been recorded properly in previous RM grid surveys, primarily due to the lack of RM radio astronomy in the southern sky [Hutschenreuter and Enßlin, 2020; Hutschenreuter et al., 2022; Gaensler et al., 2010, 2024]. This allows for the analysis of HVCs which otherwise would not be analysed under legacy data.

Generally, Gaensler et al. [2010, 2024] are the references for current POSSUM data. All POSSUM data is recorded using the ASKAP radio frequency band 1 at 880-1088



Figure 2.1: An Aitoff projection of all portions of the RM sky observed by ASKAP for the POSSUM survey. The map is up to date as of May 2024 and all RMs in this map are used in the production of this report. The faraday depth of these RMs are indicated by colour.

MHz [Vanderwoude et al., 2024; Gaensler et al., 2010, 2024]. The POSSUM grid data used in this report represents all collected and processed RMs by the POSSUM survey as of May 2024, with a total RM source count of 188842. Figure 2.1 represents the entire sample on an Aitoff projection.

2.2 Observational Data

While the primary source of data used was from POSSUM, several other sources of data were employed in the process of analysis.

2.2.1 RM Sky Interpolation

The process of estimating RM foreground contributions typically requires the use of an interpolation between RM grid point measurements. While more modern research has been done in interpolating the POSSUM data set, for example the recent paper by Khadir et al. [2024], due to POSSUM's lack of complete sky coverage, it was determined as better to use a whole-sky RM interpolation - this is due to: the uniformity of the map across the whole sky; the level of sky coverage (especially required for section 3); the reliability of using well-developed data; the accuracy of an interpolation of a complete RM sky. The most recent RM sky interpolation came from Hutschenreuter and Enßlin [2020]; Hutschenreuter et al. [2022], which combined all previous sources

of RM grid data with free-free emission from the Planck survey. This survey will be referred to as the “Hutschenreuter map”. Most notably, the Hutschenreuter map has a large ‘blind spot’ near the terrestrial southern pole, with the use of free-free emission to constrain the data better [Hutschenreuter and Enßlin, 2020; Hutschenreuter et al., 2022]. This does not disqualify the merits of using Hutschenreuter map, but should encourage further research into improvements on existing interpolations of the RM sky.

2.2.2 Other Data Sources

Both HI and H-alpha maps of the sky were obtained in the process of data collection. Both maps were used in previous literature to detect HVCs and construct interpolations of the RM sky. In this report specifically, their primary purpose in including them in the collated data is due to illustrative and future application purposes.

The HI sky was taken from the HI4PI 21 cm survey, with modifications done by Westmeier [2018] to filter for high-velocity HI sources (above a column density of $2 \times 10^{18} \text{cm}^{-2}$). This modification allows for the better detection of HVCs in the sky, however it eliminates the ability for RMs to directly be analysed using the real HI column density, due to the removal of low-velocity modes – a factor which will be accounted for in section 4. Unlike all other sources of data, Westmeier [2018] does not provide uncertainties for its HI emissions. It is assumed that the HI uncertainty is approximately equivalent to the Poisson noise i.e. the logarithmic column density is multiplied by one half.

The H-alpha sky was taken from Finkbeiner [2003], which was a collage of three smaller H-alpha surveys. Unfortunately, this map is also limited by the same problems as the Hutschenreuter map, with a notable lack of coverage near the terrestrial south pole. It was decided that the H-alpha map would be included in the process of collation for the purposes of future research potential.

Both maps, the Hutschenreuter map, and their respective error maps are displayed in figure 2.2. The only exception being the HI error, which is only altered by a scalar. All data sources collected were first converted to a FITS file under the cartesian projection. The location of individual point-values will not be affected by a cartesian projection, however it can introduce distortions at a high Galactic latitude. This does need to be accounted for.

2.3 Selection of HVCs

All HVC data was obtained from Moss et al. [2013] (hereafter referred to as the “Moss catalogue”) – a catalogue of all HVCs found using the Galactic All Sky Survey (GASS). The Moss catalogue is a primary source of data, for the location and size analysis of



Figure 2.2: Cartesian plots of peripheral data maps (left): the HI sky from the Westmeier [2018] modification of the HI4PI survey (top); the H-alpha sky from Finkbeiner [2003] (middle); and the Hutschenreuter map from Hutschenreuter and Enßlin [2020]; Hutschenreuter et al. [2022] (bottom). Uncertainty maps for H-alpha and the Hutschenreuter map are displayed respectively (right). The HI uncertainty is not displayed due to it simply being a scalar multiple of the HI map.

HVCs.

The Moss catalogue includes a total of 1693 HVCs, of which most are not viable candidates for analysis. There are several reasons why a particular HVC may introduce significant analysis errors. The first consideration is size. From section 1.1, HVCs globally have incredibly variable sizes, but locally HVCs in the halo are more consistent in size. Exceptions to this rule cannot be included in HVC analysis as they may not be representative of a typical HVC. For example, the Smith Cloud, due to its size, has an abnormally large corresponding magnetic field. Thus, HVCs that were not in an angular size range of $(1, \pi)$ degrees were be masked out. The lower limit exists to guarantee that there are enough RMs covering the cloud itself. This reduces the sample to 151 HVCs.

Other considerations made when filtering HVCs were their overlap with the current POSSUM RM grids. Not every HVC is properly covered by the current RM grid. HVCs were filtered out if their centres (obtained from the Moss catalogue) were more than one degree separated from the nearest RM sampling point. This further reduced the sample size to 26 HVCs.

There is a major increase in scatter with POSSUM RMs and interpolated RMs closer to the Galactic midplane [Schnitzeler, 2010]. This is explored in later sections; however, figure 2.3 represents this scatter. Because of this, HVCs close to the Galactic midplane must be eliminated to reduce scatter – specifically HVCs located with Galactic Latitudes $|b| < 20^\circ$ were excluded. This reduces the sample size to 15.

The distortions introduced by cartesian projections need to be accounted for. HVCs located at Galactic Latitudes $|b| > 80^\circ$ were also eliminated from the sample set. This removes one HVC to give a sample size of 14.

Lastly, it was clear that HVC G282.3-38.3+117 had not detectable HI emission according to the provided HI data. It was removed on the basis that it could not be analysed properly. This reduces the final sample to 13 HVCs.

2.4 Data Collation

Once the data was obtained, calculations were done using the `astropy.wcs` pipeline. For every RM point in the sky, the estimated foreground RM from the Hutschenreuter map, the HI column density, and the H-alpha flux, including all associated errors were attached to that particular RM sampling point.

2.4.1 HVC Image Overlays

For each HVC, the HI, H-alpha, Hutschenreuter map, and RM grid was ‘cropped’ according to a field twice the size of the maximum HVC source x and y extents. This is to allow for analysis of both RMs in the HVC and surrounding the HVC.



Figure 2.3: (Left) A graph of all ~ 180000 RMs plotted against their corresponding Galactic latitude; (Right) The corresponding graph of all Hutschenreuter map faraday depths matched to the POSSUM RMs. Both graphs represent a significant level of scatter present in RMs collected near the Galactic midplane.

Foreground Subtraction

Foreground subtraction remains an open question in the field of RM radio astronomy. Due to the inherent magnetisation of the Galactic halo and the ISM, as discussed in sections 1.2.1 and 1.2.2, there are significant contributions to observed RMs from the Galactic foreground across the entire sky. To correctly determine the magnetic field surrounding objects of interest, one must first remove this source of systematic error.

A prime example of the consequences of not correctly accounting for foreground RM contributions is the paper by McClure-Griffiths et al. [2010], which attempted to estimate the magnetic field strength surrounding HVCs in the Leading Arm. However, as from Jung et al. [2021], this result is possibly erroneous due to the obstruction of the nearby Antila supernova remnant region. While the analysis of multiple HVCs is more likely to prevent these errors from compounding to invalidate the conclusions of the report, it is still useful to account for these contributions as much as possible.

3.1 Interpolation

All previous work on magnetic field analysis of HVCs involve the use of interpolation [Moss et al., 2013; Betti et al., 2019; Westmeier, 2018; Hill et al., 2013]. Interpolations are very beneficial due to its ability to convert a discrete distribution of RM grid points into a continuous distribution of the RM sky. Interpolations also benefit from a ‘smoothing’ effect; that interpolated maps can smooth out small-scale imperfections in the RM grid that may not correspond to actual foreground objects [Hutschenreuter and Enßlin, 2020; Hutschenreuter et al., 2022; Khadir et al., 2024]. This smoothing can occur because of the interpolation algorithm itself, or the lack of RM grid density to resolve objects on a particular scale.

The primary issue with interpolation is that they are too effective a technique at reconstructing the RM foreground. Despite the smoothing effects that they can provide, there is no way one can confirm that the RM sky has not included objects of interest as well; in other words, an interpolated RM sky could contain the profiles of objects of interest. This has the effect of making interpolated maps sourced from high-density RM grids redundant, as the original intention of interpolation is to subtract out foreground contributors to RMs – not the objects of interest themselves.

Thus, with the increased RM grid density afforded by the POSSUM survey, and

future SKA-era projects, it is of high importance that discussions on how to account for these issues can be solved.

3.2 Annulus Subtraction

The immediate alternative is annulus subtraction. This is a method employed across all fields of astronomy, including radio astronomy, being most applicable to single-object analysis. The method generally involves selecting a series of RM sampling points surrounding any given central RM grid point, averaging the selected RMs, and subtracting the average from the central RM grid point.

There are two sub-methods to consider when performing annulus subtraction: fixed-size annulus subtraction and fixed-sampling annulus subtraction. Fixed-size annulus subtraction involves defining an annulus with a constant inner and outer radius and averaging the RMs exclusively within this radial range. Fixed-sampling annulus subtraction involves defining an inner radius and then selecting a fixed amount of RM grid points that are closest to the central point, but still outside the inner radius. Assuming a constant grid density everywhere in the field, a relationship between the two methods can be quantified, as in the following equation:

$$\begin{aligned} R &= \sqrt{r^2 + \frac{N}{\pi n}} \\ N &= \pi n (R^2 - r^2) \end{aligned} \tag{3.1}$$

Where n is the RM grid density in deg^{-2} , r is the inner radius of the annulus in degrees, R is the outer radius of the annulus in degrees (which is directly fixed under the fixed-size regime), and N is the number of RM grid points used (which is directly fixed under the fixed-sampling regime) and is unitless. This means that under a constant RM grid density, these two methods should be approximately equivalent.

There are benefits to both methods. On one hand, fixed-size methods can be described mathematically as convolutions, making them linear. However, they can run into measurement and calculation errors when there is a low amount of RM grid points surrounding the central RM. On the other hand, fixed-sampling methods guarantee a consistent uncertainty and the existence of an average. However, this method is both non-linear and prone to including RM grid points very far away from the central point.

The primary issue with annulus subtraction is determining the size of the annulus, or the amount of RM sampling points to select i.e. what counts as the “foreground”. This issue is what leads to many of the above-mentioned errors in both methods. The logical response is either to select a large annulus that completely removes the objects’



Figure 3.1: An Aitoff projection of all portions of the RM sky observed by ASKAP for the POSSUM survey similar to 2.1, filtered through a proprietary fixed-sampling algorithm with inner radius of $0.4''$ and a sampling constant of 50.

RM contributions (in the case of HVCs this would correspond to an annulus of $1 - \pi$ degrees in radius), or to select a small radius with numerous RM points to capture the foreground contributions both overlapping the object and isolated in the field. Both methods will be analysed in this paper, with the former being discussed in section 3.3.3 and the latter being shown in figure 3.1¹. The specific choice of parameters in the latter method being an inner radius of $0.4''$ and a sample size of 50 grid points – corresponding to an outer radius of approx. 0.728° .

3.3 Fast Fourier Transforms (FFTs)

Many of the methods for foreground subtraction beyond interpolation appear to have a common intersection point in the form of image-based signal processing. Thus, the introduction of Fourier Transforms (FTs) may be a very useful direction for analysis. The benefit of FTs is their linearity, which has several benefits including: the trivialisation uncertainty calculations (quantified in equation 3.3); the linear combination of several kernel techniques; signal processing in separate orthogonal dimensions; and consistent scaling relationships. FTs also can utilise both convolutional blurring and bandpassing separately, with convolutions already being discussed with annulus subtractions.

FFTs extend the benefits of FTs by providing a FT algorithm of $O(N \log N)$ com-

¹The fixed-sampling annulus subtraction was proprietary data collected from the supervisor. However, the quantitative and qualitative analysis of the viability of this method is my own work.

plexity and allowing FTs to be performed over discrete sets of data. This allows for the analysis of high-definition pixellated images, which is not unlike the standard format and use-case of a FITS file, especially when using a cartesian projection of the sky. Thus, by applying 2-dimensional FFTs to provided interpolated RM sky images, it is possible to solve the problem introduced by interpolated high-density RM grids.

3.3.1 Non-Uniform Fast Fourier Transforms (NUFFTs)

FFTs can further be extended to the analysis of non-uniform data sets. Standard FFTs rely on the assumption that the grid of sampling points is uniform, and resultant output uniform-density frequency distributions. NUFFTs do not require the assumption of uniformity, nor do they need to output uniform-density frequency distributions [Bagchi and Mitra, 1996; Greengard and Lee, 2004]. This means that instead of relying on interpolations at all, FTs can be applied directly to the RM grid itself. The primary sources for NUFFTs used in this paper are Bagchi and Mitra [1996]; Greengard and Lee [2004], with heavy reliance on the python module PyNUFFT (see appendix D for more).

There are three types of NUFFT: Forward, Adjoint, and "True" ². The forward and adjoint types are inverses of each other – forward NUFFTs take a uniform image and a set of sampling points and return a non-uniform frequency distribution and adjoint NUFFTs reverses that process. True NUFFTs take a non-uniform distribution and output a non-uniform frequency distribution. True NUFFTs are not generally useful for the purposes of this report.

Applying a forward and then an adjoint NUFFT to a set of RM grid sampling points should perform the same task as creating an interpolation. From there, the intermediary step of a bandpass or kernel can be applied to the frequency distribution to produce an interpolation with objects of a particular scale removed from the field.

The important first step in determining if this method can create a reliable interpolation on its own. First, an image was selected, specifically a grayscale and cropped image of the Cosmic Microwave Background (CMB) from the Planck mission (see appendix C for more). This was chosen as the test image due to the CMB being able to replicate a noisy and 'blobby' structure, the CMB has also been analysed using FFTs for unrelated cosmological purposes.

Then, a random set of sampling points were selected and treated as the 'mock RMs', with the colour of the background corresponding to the intensity of the RM at that point. The image and the sampling points were then given to the PyNUFFT module and transformed in and out of the frequency domain. Figure 3.2 represents the outcomes of this analysis, performed on a sample of simulated RM grids with

²There is no generally accepted nomenclature for type 3 NUFFTs that align with the single-adjective terminology. So "True" is used because it is an accurate descriptor for the type, involving both non-uniform x and k spaces.



Figure 3.2: (Right) A cropped and grayscale image of the CMB. (Left) The same image after being fed through a forward then adjoint NUFFT. The amount of sample points total to 27000. Both plots share a common intensity colourbar for reference.

size $30^\circ \times 30^\circ$. Ignoring the grid-like structure in the recreated image (a consequence of the random point generation algorithm), even with a very high sampling point density or large field, the image is still very low-quality.

This does not disqualify the NUFFT as an analysis technique. Instead, it means that this technique can only work on a very large continuously connected RM set i.e. a complete or partially complete RM grid map of the sky. However, due to the lack of POSSUM data in its early stages, this is a method that must be investigated in the future.

3.3.2 Bandpass Filtering

A simpler method is to directly alter the Hutschenreuter map itself using normal FFTs. First, a 2-D FFT was applied to the Hutschenreuter map. A crosshatch-shaped bandpass was created. This crosshatch imitates a bandpass commonly applied to 1-dimensional signals, where objects of a particular angular size are eliminated by removing all frequencies corresponding to that angular size in the k-space. The equation below quantifies the relationship between frequency and angular size:

$$k_{\text{HVC}} = \frac{1}{2\theta_{\text{HVC}}R} \quad (3.2)$$

Where k_{HVC} is the spatial frequency in deg^{-1} , θ_{HVC} is the angular size of the HVC in degrees, and R is the pixel resolution of the axis, in pixels per degree. Assuming all RM sky images exist in a 2:1 cartesian space, due to the range of Galactic latitude and longitude, the value of R is constant across the two axes.

The crosshatch is shaped such that, when multiplied by the original k-space, objects of a particular size are either eliminated or reduced in prevalence. This method



Figure 3.3: Crosshatch-Bandpassed versions of the interpolated RM sky at various opacity gradings - where "original" means 0% opacity. The term "opacity" refers to the effect of the bandpass e.g. 100% opacity means the bandpass completely removes selected frequencies and 50% opacity means that the bandpass halves the presence of selected frequencies.

also guarantees the linearity of the crosshatch 'function'. After this, the inverse FFT is applied to give a resulting foreground map, seen in figure 3.3.

However, bandpassing introduces ripples into the Hutschenreuter map. This effect is expected but undesirable. There are two methods to remove this: either to apply the crosshatch at a certain 'opacity' i.e. the crosshatch is not eliminating all the k-space in its region, but instead is reducing those frequencies by a percentage; or using a more complex window than a Top Hat, such as a Tukey window or Gaussian window. The effects of the former are seen in figure 3.3. The latter was not investigated due to time constraints.

When applying FTs to any interpolated image, it is important to maintain the corresponding uncertainty map's accuracy. This is where one can take advantage of linearity. The following formula below determines how uncertainties can be calculated:

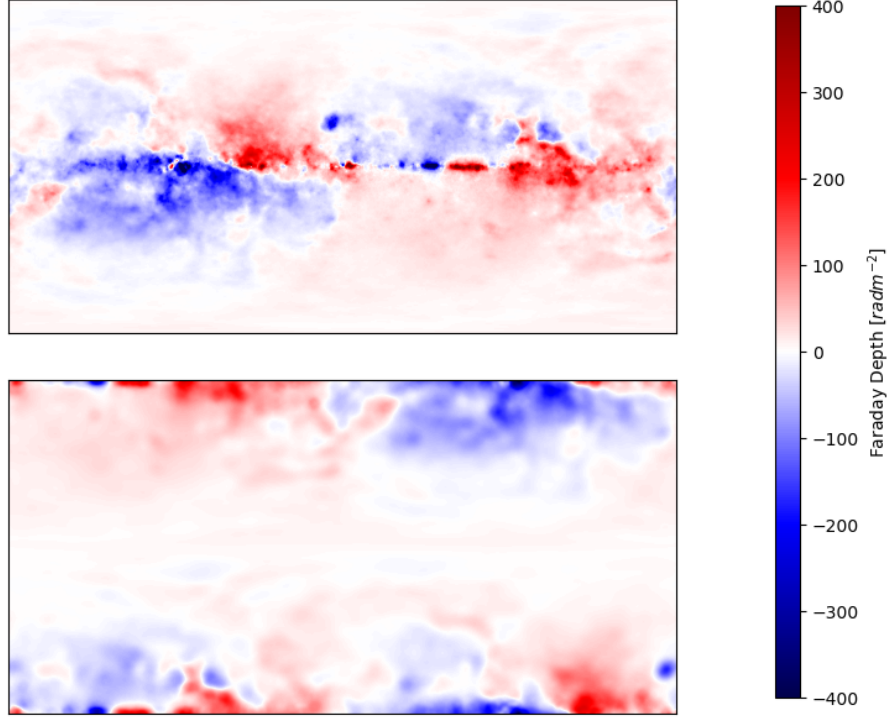


Figure 3.4: Cartesian plots of the interpolated RM sky (Top) compared against the Annulus-Bandpassed version of the interpolated RM sky (Bottom). The annulus kernel used sums to unity, making it act like a unitary operator, standard for blurring techniques [Pulfer, 2019].

$$\sigma_{\text{output}} = \mathfrak{F}^{-1} \left[B(\mathfrak{F}[\sigma_{\text{original}}]) \right] \quad (3.3)$$

Where $B: \sigma \rightarrow \sigma$ is the bandpass function, σ_{output} and σ_{original} is the uncertainty images for the output and input respectively in radm^{-2} , and \mathfrak{F} is the FT.

3.3.3 Kernel Filtering

The same techniques from above can be applied to convolutions, where the aim is to convolve the Hutschenreuter map with a defined kernel. Two-dimensional convolutions have a time complexity of $O(n^4)$, depending on the kernel size, whereas the FFT has a complexity $O(n^2 \log^2 n)$. By performing a FFT on both the kernel and the Hutschenreuter map separately, then multiplying the two k-spaces together, and applying an inverse FFT, the result is a faster application of a convolution with a kernel. This was the chosen method to demonstrate the large fixed-size annulus subtraction method. Figure 3.4 displays the results of this method.

The main problem, as apparent from this method, is that the annulus kernel acts

as an edge detection algorithm [Pulfer, 2019]. This causes defects at higher absolute Galactic latitudes due to the effect of distortions from the cartesian distribution. Another problem is that this subtraction method is being applied to the Hutschenreuter map, instead of the actual RM grid. This secondary issue can be solved by relying on the NUFFT of a larger RM grid. As it stands, this method appears to be incompatible with the goals of foreground removal.

3.4 Characterising the RM data

As seen from figures 1.3 and 2.3, there are several ways in which the RM grid can have ‘bad data’ – most notably in a lack of signal-to-noise and the inherent scatter when observing near the Galactic midplane. Thus, the final step of this chapter is to both characterise the RM sample set and to compare the subtraction methods against each other.

Figure 3.5 represents a simple residual histogram comparison between all the methods discussed in this chapter. The desired result is seen in the residuals between the interpolated or crosshatch-bandpassed RMs and the actual RMs - appearing as a distribution centred at zero (which does not look gaussian). This is opposed to the two annulus methods, which do not appear to interact with the RM grid in a desirable manner.

Figure 3.6 demonstrates the similarities between the crosshatch-bandpass and unaltered Hutschenreuter map, with them being related to each other in a linear manner, specifically with a gradient of approximately unity. This is ideal, as it means that the crosshatch-bandpass method is not deviating significantly from the Hutschenreuter map, only altering it subtly. The figure also demonstrates how scattered the RMs become near the Galactic midplane, hence it being plotted for colour to delineate between RMs near and far away from the midplane.

Figure 3.7 compares the unaltered Hutschenreuter map with the annulus-convolved method (a.k.a. the fixed-sampling annulus method). There is a somewhat linear relationship between the actual RMs and the annulus-convolved RMs, ignoring the heavy scatter closer to the midplane. This gives credibility to this method, and the choice of having a small-sized annulus.

Lastly, figure 3.8 compares the two annulus methods together. The annulus-bandpass method (a.k.a. the fixed-size annulus method). From the bottom two graphs, the annulus bandpass does not interact properly with the Galactic midplane, sending most RMs to near-zero. Despite this, RMs at higher Galactic latitudes ‘blow up’. This implies that the utilised annulus-bandpass method is not a viable for foreground correction.

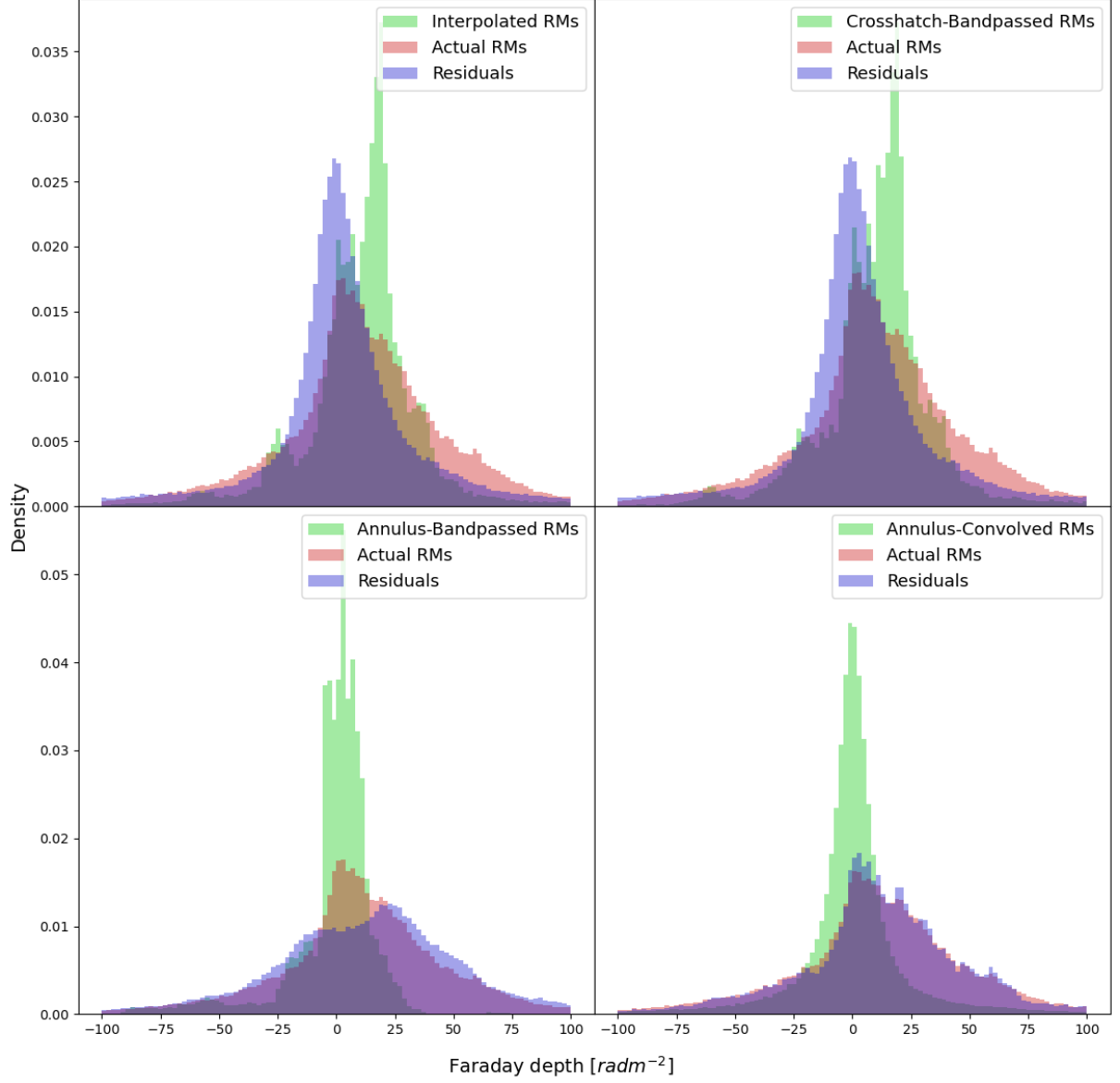


Figure 3.5: Residual histogram plots of the actual POSSUM RMs compared to the corresponding various foreground removal methods: interpolation (Top left); crosshatch-bandpassed (Top right); annulus-bandpassed (Bottom left); and annulus-convolved (Bottom right).

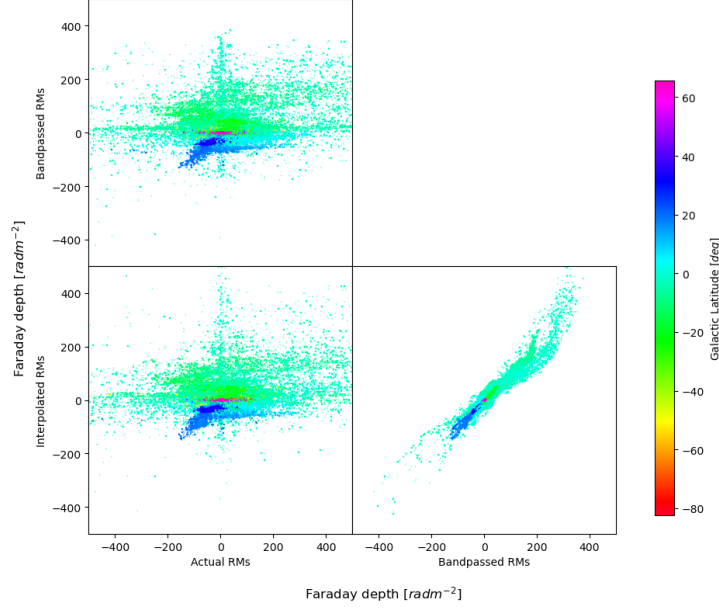


Figure 3.6: Three corner plots of all RM sampling points, describing the relationship between the several associated RM values and Galactic latitude indicated by colour. Here, a comparison between actual RMs, interpolated RMs, and Crosshatch-Bandpassed RMs is displayed.

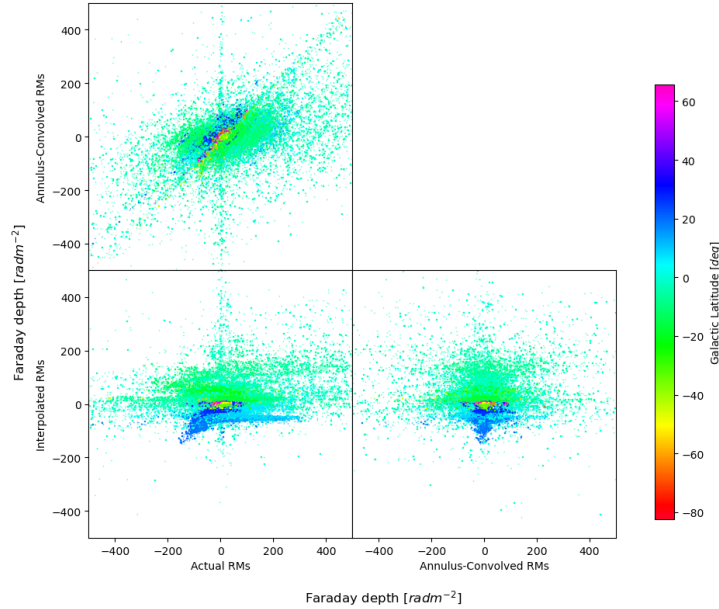


Figure 3.7: Three corner plots of all RM sampling points, describing the relationship between the several associated RM values and Galactic latitude indicated by colour. Here, a comparison between actual RMs, interpolated RMs, and Annulus-Convolved RMs is displayed.



Figure 3.8: Three corner plots of all RM sampling points, describing the relationship between the several associated RM values and Galactic latitude indicated by colour. Here, a comparison between actual RMs, Annulus-Bandpassed RMs, and Annulus-Convolved RMs is displayed.

3.5 Statistical Comparison of Foreground Removal Methods

There is a lack of an ability to quantify the utility of the discussed foreground removal methods. Hence, why most analysis in this section is visual or qualitative. However, there is still benefit in deriving numerical statistics.

The histograms displayed in figure 3.5 do not follow a Gaussian distribution or a Cauchy distribution, but do follow a Student's t-distribution. This has been seen before in past analysis of residual RMs [Anderson et al., 2024]. By eye, it is quite clear that this is not the case for the other two distributions, hence why no statistical analysis was conducted on them. A chi-squared test can be used to compare the residual RM distributions with the t-distribution, with the t-distribution being fitted via a bootstrapping method. 1000 resamplings with replacement were used.

Table 3.1 displayed the resulting t-distribution fit and Pearson's chi-squared statistic. The p-values for both were extremely small, despite the clear visual incongruency in the model and data set in figure 3.9. This is most likely caused by the t-distribution fitting very well with the tails and mid-section of the data histogram.

The Pearson R^2 test can be performed in comparing the Hutschenreuter map with its corrected versions. Ideally, the R^2 statistic should be somewhat close to 1, as not

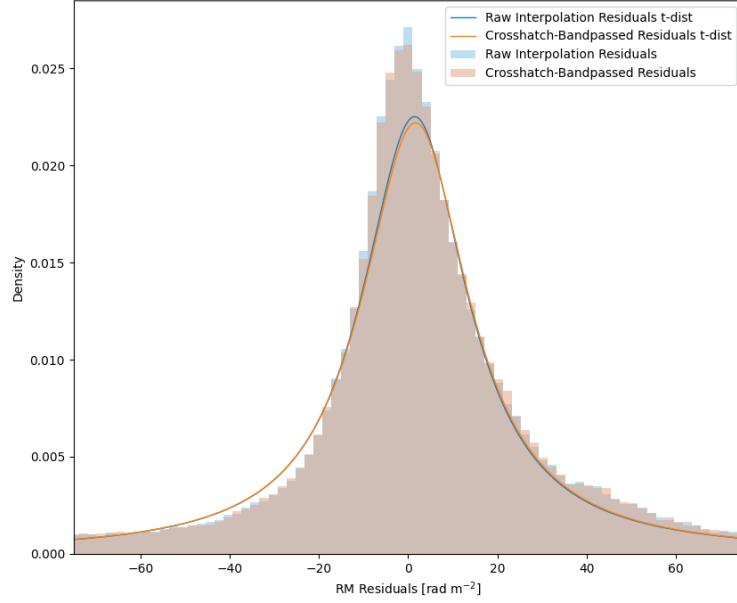


Figure 3.9: The sample of all RMs corrected with the Hutschenreuter Interpolation (Blue) and Crosshatch-Bandpassed Interpolation (Red). A bootstrap fitted t-distribution is drawn on the graph for each respective residual RM set.

Residual Set	t-Distribution Fit		
	Centre	d.o.f.	Scale
Straight Interpolation	1.0395 ± 0.0039	1.4583 ± 0.1828	14.241 ± 0.0587
Crosshatch-Bandpassed	1.0463 ± 0.0038	1.6425 ± 0.0694	14.461 ± 0.0583

Residual Set	χ^2 Results	
	Statistic	p-value
Straight Interpolation	0.01432	1.042E-166
Crosshatch-Bandpassed	0.01459	2.618E-166

Methods	R^2 Statistic
Interpolation vs. Crosshatch-Bandpassed	0.9279
Interpolation vs. Annulus-Bandpassed	0.007611
Interpolation vs. Annulus-Convolved	0.003866
Annulus-Bandpassed vs. Annulus-Convolved	1.3894E-06

Table 3.1: (Top) Student's t-distribution fit parameters for the residual histograms as displayed in figure 3.9. (Middle) Results of the χ^2 test comparing residual histograms with the t-distribution. (Bottom) R^2 statistics comparing the Hutschenreuter map with the other methods detailed in this section.

a lot of sub-degree variations are expected (due to the single-degree scales of the Hutschenreuter map's sources). Table 3.1 presents the R^2 statistics for a series of compared cases.

3.6 Other Methods

There are other methods for foreground removal that, while researched, were not considered viable options or of enough importance to numerically analyse in this report. However, they may still offer useful methods for future researchers. The first is median filtering, which is like previously discussed convolutional blurring methods. Huang et al. [1979] provides a description of a fast median filtering algorithm for two-dimensional images. Arias-Castro and Donoho [2009] is also a reference describing the potential benefits of median filtering, including a more robust removal of noise due to the statistical properties of the median, and the preservation of edges. It was disregarded in this report due to its inherent non-linearity, and the debatable nature of whether this disadvantage is worth the advantages it can bring. Linear decomposition of line-of-sight RMs are also possible, attempted in Schnitzeler [2010]. However, it is also quite mathematically complex and may not be conducive to a generalised algorithm, hence the lack of focus on this technique in this report.

Magnetic Field Derivation

In constructing a rudimentary algorithm to quantify the effect of magnetic draping on HVCs, as in section 1.4, it is important to pay attention to the validity of the methods being employed and less the accuracy. As an inaccurate methodology can be improved using better data, removing assumptions, and more precision, however an invalid methodology cannot provide a proper estimate whatsoever.

4.1 Assumptions

Due to the significant limitations in available data, the use of assumptions is necessary to begin the process of rough evaluation. Several assumptions were employed in the process of evaluation, some of which are less than satisfactory at contributing an accurate answer to the research question. Again, it is highly important to note that goal is to ensure the validity of the method is sound and to ensure the method's manipulability, but not necessarily provide a guarantee of accuracy and precision.

The first major assumption is that the ionisation fraction of the Galactic halo is approximately constant, i.e. gas in the halo is well-mixed. This does not account for major sources of heterogeneity caused by HVCs (a desired outcome) or other ionised regions such as the Magellanic Clouds (an undesired outcome). It is not necessarily unreasonable to assume this, especially if the region of analysis per HVC is under 4 square degrees. As there is less likelihood of interference from one or more sources.

Simulation data can be used to analyse the problematic nature of assuming a constant ionisation fraction [Grønnow et al., 2017]. However, it is difficult to apply simulation data to observational data due to the amount of variables that need to be controlled for. For now, it is 'good enough' to rely on the assumption from Grønnow et al. [2017] that the ionisation fraction is high and constant.

The second major assumption is that the weighted average of the HI column density on a HVC is linearly proportional to, and biased towards the peak column density in the centre of the HVC. This is an assumption that only works under the consideration of averages, as HI column density is certainly not uniform within clouds, as mentioned in Heitsch et al. [2021]. It also necessitates some level of correlation between line-of-sight magnetic fields and their uncertainties (a consequence of Poisson noise).

Both above assumptions can, and should, be taken to be unreasonable for providing an accurate estimate of magnetic field strength. However, the result of applying both assumptions is that the line-of-sight magnetic field strength is linearly proportional to the faraday depth with a scaling constant unique to each HVC – as demonstrated in later section 4.2. This greatly simplifies the analysis, without loss of validity, primarily due to the method of calculation explained by Kaczmarek et al. [2017] and in section 4.2.

The third core assumption is that HVCs will appear as approximately circular within the field. Meaning that to distinguish between the surrounding medium and HVC RM grids, a simple circular method can be applied. As seen from the HI plots in figure 4.1, this assumption appears to only be correct in particular cases. But do note that the HI background discriminates for high-VLSR HI emission.

All three assumptions additionally meet the desired goal of an alterable method. As discussed in the discussion and conclusions sections 5.2.2 and 6, these three assumptions can be dropped in the presence of more detailed data or more rigorous methodologies manipulated from the one described in this paper.

Lastly, a very typical assumption made is that the RM contribution along a line of sight is averaged across said line of sight. This is not necessarily correct; however, it is a fair assumption that most past literature has made to deconstruct the integral as seen in Kaczmarek et al. [2017]; Moss et al. [2013]; Hill et al. [2013]. This makes the fourth assumption.

4.2 Derivation of Line-of-Sight Magnetic Fields

As from equation 1.1, there is a relationship between the line-of-sight magnetic field and the faraday depth. By removing foreground contributions to the faraday depth, one can isolate the specific contribution of the line-of-sight magnetic field made by the HVC. At the same time, the fourth assumption can be used to solve evaluate the integral, resulting as follows [Hill et al., 2013]:

$$\frac{\langle B_{\parallel} \rangle}{\mu\text{G}} = \frac{\langle \Phi_{\text{cor}} \rangle}{0.81 \langle n_e \rangle L_{H^+}} \quad (4.1)$$

Where $\langle B_{\parallel} \rangle$ is the average line-of-sight magnetic field in μG , Φ_{cor} is the corrected faraday depth in radm^{-2} , n_e is the electron density in cm^{-3} , and L_{H^+} is the HVC path length in pc. From this equation, Kaczmarek et al. [2017] derives a simplified equation that eliminates the need to calculate the electron density and path length:

$$\frac{B_{\parallel}}{\mu\text{G}} = 3.80 \times 10^{18} \frac{\Phi_{\text{cor}}}{X \langle N_{\text{HI}} \rangle} \quad (4.2)$$



Figure 4.1: All 13 HVCs used in the analysis of the primary outcome. The HI column density is represented using a greyscale image background. The RMs are represented by circular markers, their size equal to the magnitude and the colour representative of their sign with Red being positive. The black circle is the deliniation between in-HVC and out-HVC populations, and the black 'x' indicates the centre of the HVC.



Figure 4.2: The same set of HVCs in HI with RM overlays, displayed the same as in 4.1. However, instead of using the interpolation as a method of correction, the mean uncorrected RM value is subtracted out.

Where X is the unitless ionisation fraction, and $\langle N_{HI} \rangle$ is the weighted average HI column density over the region being analysed, measured in cm^{-2} . This equation can only be derived under the first assumption. Additionally, the use of the average HI column density is intrinsically tied to the second assumption.

Applying both assumptions one and two and utilizing this equation demonstrates that the line-of-sight magnetic field is linearly proportional to the faraday depth. Thus, a ‘master RM’ can be constructed - denoting the faraday depth associated with any given HVC, and any given background surrounding the HVC. These two master RM values can be subtracted from each other and converted into a magnetic field via multiplication of a constant factor. The method can be performed in that order due to the second assumption. The methods used to create this ‘master RM’ are described in 4.4.

Despite the detailed analysis conducted in section 3, it was decided that the non-altered interpolated sky map be used to correct RM contributions. This is due to most already-existing literature relying on this method. From figure 4.3, which displays histograms of the corrected and uncorrected line-of-sight magnetic field points surrounding HVCs, it is demonstrated that the Hutschenreuter map can remove the anisotropy of the ISM by turning an uncorrected bimodal symmetric distribution into a normal distribution centred at zero. However, it also is apparent that the correction can introduce a significantly higher variance in the data, shown in the ridge-like structures in figure 4.1, which presents HI images of all 13 HVCs, the overlapping RM grids, and the cropped HI fields. The filtered Moss catalogue, including all 13 HVCs is displayed in appendix B.

4.3 KS Testing

Before evaluating the effect of magnetic draping on HVCs, it is important to first confirm that there is a detectable magnetic field to analyse. To do this, the KS test was employed. The set of RM grid points for each HVC collated in section 2.4.1 include any grid point within $2 - 2\pi$ degrees of the HVC centre, depending on HVC size. Employing assumption three, two separate populations can be formed by delineating RMs inside and outside a certain circular distance from the HVC centre. The specific radius of delineation was the average value between the Moss catalogue’s ‘dx’ and ‘dy’ values (i.e. its cartesian dimensions). All sampling points within the defined circle is apart of the “HVC RMs” population, and the ones outside the circle are apart of the “Background RMs” population.

A two-sample, two-tailed, KS test was performed comparing both the background and HVC RM distributions. A critical p-value of 0.001 was assumed to determine if a significant magnetic field was detected. Table 4.1 shows the KS test results applied to each HVC. Out of the 13 HVCs tested, 10 HVCs (77%) had a significant difference

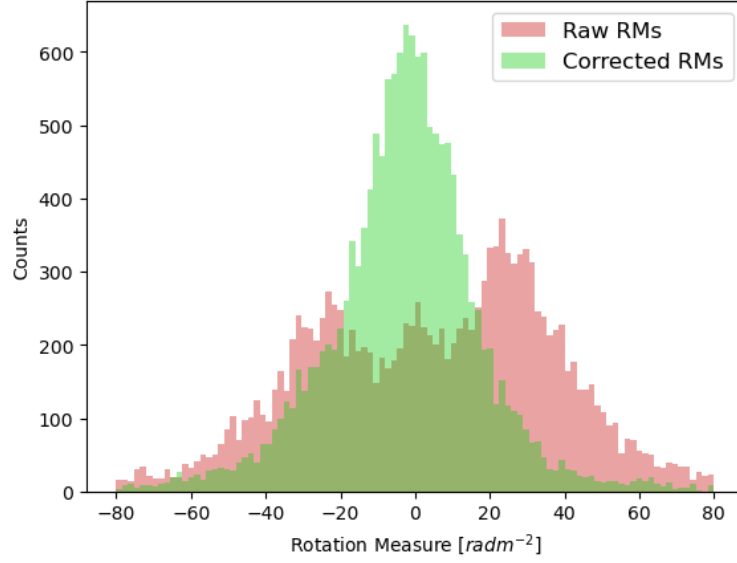


Figure 4.3: Two histograms describing the RM sample within a certain distance to the 13 sample HVCs. The red histogram uses uncorrected RMs. The green histogram uses corrected RMs.

between both populations in the corrected RM case. This is a promising confirmation of the hypothesis presented by the literature, that magnetic draping is an existing phenomenon.

For uncorrected RMs, the KS test determined that 12 HVCs had a significant difference between both populations. No HVC had both a significant detection for the uncorrected case and a non-significant detection for the corrected case. This implies that using the corrected RMs makes the KS test more selective.

There appears to be a very weak correlation between the KS statistic and the Galactic latitude for each HVC, with a calculated R^2 statistic of 0.6039. When determining correlation with respect to p-value instead of KS statistic, the R^2 statistic drops to 0.4489. The KS test does not appear to be correlated with VLSR or VGSR, having an R^2 statistic of 0.03023 and 0.008367 respectively.

4.4 Mathematical methods to evaluate HVC Magnetic Fields

For the HVCs which did have detectable magnetic fields, the next step is to estimate the strength of the magnetic field draping over each HVC. Previous analyses of the Smith Cloud, such as in Betti et al. [2019]; Hill et al. [2013], utilises a weighted average to determine the magnetic field strength. However, this method may not be transferable to a generalised algorithm. In Betti et al. [2019]; Hill et al. [2013], due to analysing

Name	Corrected RMs		Raw RMs	
	Statistic	p-value	Statistic	p-value
G038.7-37.3-241	0.207	3.52E-09	0.198	1.98E-08
G038.8-33.2-261	0.311	1.39E-08	0.516	8.69E-24
G043.4-30.4-230	0.241	3.61E-11	0.174	5.20E-06
G089.0-64.7-311	0.374	1.51E-02	0.499	2.96E-04
G133.5-75.6-294	0.461	4.57E-08	0.485	6.34E-09
G248.9+36.8+181	0.081	4.53E-01	0.110	1.32E-01
G261.5-34.5+237	0.289	5.96E-08	0.324	6.28E-10
G262.4-26.2+143	0.213	1.22E-06	0.221	4.48E-07
G264.6-53.3+224	0.320	4.07E-05	0.346	6.04E-06
G267.1-38.4+263	0.343	2.32E-11	0.348	1.14E-11
G271.6-48.4+297	0.125	5.09E-05	0.155	1.69E-07
G271.7-53.1+254	0.227	3.40E-04	0.246	7.62E-05
G323.0+25.6+122	0.099	1.51E-02	0.187	4.79E-08

Table 4.1: A table describing the KS test results for each HVC.

one single and large object, it is possible to divide the object into several smaller sections for analysis. It is not possible to do this especially given the limitations created by assumption three.

While the background RM population will generally follow the expected normal distribution, the in-HVC population can come in multiple varieties depending on the orientation of the HVC with respect to the observer. These may appear ‘coffee stain’ RM fields, similar to the superbubbles found in the Magellanic clouds as shown in Jung et al. [2024].

The more complex HVC RM distribution combined with the oversimplified spherical morphological model means that using a weighted average is dubious in validity at best - the same applies to most forms of modal analysis including gaussian fitting i.e. any method that can not account for antimodal or multimodal analysis. Henceforth, this report proposes two new methods to determine the magnetic field strength surrounding HVCs under the simplified spherical morphological model. Sections 4.4.1 and 4.4.2 attempts to justify the mathematical validity of these two methods, and section 4.5 shows both the results of using these methods, with some further post-hoc justifications demonstrated via statistical analysis in 4.5.2.

Regardless of the hypothesis that the weighted average is inadequate, it will still be included as a method in the analysis of the results, for the point of either disproving or validating said hypothesis.



Figure 4.4: A diagram representing the physical evaluation of the KS-EDF method. The graph shows the EDFs of inside and outside populations for an example HVC. The lines drawn on the diagram represent the statistic value and the KS-EDF magnetic field value.

4.4.1 KS-EDF Method

The first proposed method relies on both the results of the KS test and the Empirical Distribution Functions (EDFs) of the in-HVC and out-HVC populations required to evaluate the KS statistic. The x-axis location of the KS statistic is where the distribution of the two populations differs the greatest. And thus, it is not unreasonable to state that the magnetic field strength corresponds to this location in some manner. This is shown in figure 4.4, where the red line denotes the KS statistic's length and x-axis position.

The second step in this method is to further remove background interference as explained in section 4.2. This can be done by finding the x-axis location at which the density in the background population EDF equals that of the density of the HVC population EDF at the KS-derived x-axis location. This step has the effect of removing any potential deviation from zero that the magnetic field in the background has due the ridge-like structures, which will appear in the EDF as either (i) any deviation from a standard normal distribution, or (ii) a deviation from the centre of the normal distribution as a function of its uncertainty.

This is shown in figure 4.4 with the black line – the length of the black line representing the final assessed magnetic field value. In this case, the background population in orange is approximately normal. So, case (ii) from the paragraph above is demonstrated.

The second step is not always valid, due to the inherent variability in the EDFs and only being loosely correlated with meaningful population variables. Hence the necessity of further statistical analysis.

4.4.2 Uncertainty Subtraction

A more legitimate method involves the reverse-propagation of uncertainties. An assumption can be made that for some HVCs, there is an inherent variation in the magnetic field as one looks at different locations in the defined morphological circle, appearing as a 'ring' of changing numerical parity. This inherent variance is denoted as σ_{true}^2 .

Additionally, each RM measurement comes with a measurement uncertainty, denoted as σ_{meas} . Propagating these two uncertainties would give what should be the observed variance, σ_{obs}^2 :

$$\sigma_{\text{obs}}^2 = \sigma_{\text{true}}^2 + \sigma_{\text{meas}}^2. \quad (4.3)$$

The observed uncertainty can be calculated by taking the statistical standard deviation of the population, and the measurement uncertainty can be averaged to give a whole-population measurement uncertainty approximation. This allows for equation 4.3 can be reversed to derive the 'true' variance in the magnetic field:

$$\sigma_{\text{true}} = \sqrt{\text{Var}(B_{\parallel, \text{virt}}) + \langle \sigma_{B_{\parallel, \text{virt}}} \rangle^2} \quad (4.4)$$

By definition of standard deviation, the corresponding true variance can be equated to the average separation between RMs in the population, thus quantifying the inherent variation in the HVC magnetic field, i.e. the magnetic field strength.

The same can be done to the background population, instead to detect the potential interferences of the ridge-like structures of the magnetic field. This can be used to subtract out the background interferences that were included in the derivation of the true in-HVC variance.

A clear potential source of error in this method comes in the fact that the standard deviation is positive, affecting the validity in two separate ways: (i) cases where the average measurement uncertainty is larger than the statistical standard deviation, resulting in a complex uncertainty; (ii) cases where the magnetic field strength *should* be negative but is calculated as positive (before the true background variance is subtracted). Case (i) does not invalidate the method outright but does mean that 'bad data' might have to be thrown out. This happened in only one case as seen in section 4.5. Case (ii) is irrelevant in the conclusion, as the parity of the magnetic field is not a factor being analysed. However, the if the main concern is if the background and HVC true variances differ in sign. This is not of concern because if there is a sign difference, the background ridge-like structures would cancel-out the presence of HVC magnetic

Name	Abs. Magnetic Field $ B_{\parallel} $ (μG)		
	Wgt. Avg.	KS-EDF	σ -Sub.
G038.7-37.3-241	2.56 ± 1.88	1.15 ± 1.12	1.17 ± 1.07
G038.8-33.2-261	1.25 ± 1.26	0.36 ± 0.75	1.13 ± 0.70
G043.4-30.4-230	2.30 ± 1.75	0.31 ± 0.93	1.61 ± 0.93
G133.5-75.6-294	1.79 ± 1.54	1.24 ± 1.29	0.76 ± 1.27
G261.5-34.5+237	7.08 ± 8.70	3.88 ± 4.55	3.39 ± 4.39
G262.4-26.2+143	5.01 ± 6.13	10.93 ± 8.64	3.94 ± 2.77
G264.6-53.3+224	2.12 ± 3.44	0.11 ± 2.58	2.18 ± 2.26
G267.1-38.4+263	9.41 ± 6.69	3.84 ± 3.73	3.48 ± 3.29
G271.6-48.4+297	0.59 ± 1.98	0.13 ± 1.24	0.08 ± 1.21
G271.7-53.1+254	2.40 ± 3.60	0.57 ± 2.43	0.02 ± 2.22

Table 4.2: A table describing the magnetic field derivations for each HVC. HVCs from the sample of 13 that had no significant KS test detection, or an invalid variance subtraction result are removed.

fields, resulting in every invalidating occurrence of case (ii) necessarily leading to case (i).

To reiterate, the caveats and assumptions described in this derivation may appear as mathematically or intuitively valid, however it is worthwhile to test things via statistical analysis and modelling (see chapter 6).

4.5 Magnetic Field Results

The results of each method are compiled in table 4.2.

The distributions of the results are displayed in the boxplots of figure 4.5, which includes the mean (green triangle) and median (orange line) of each set. The Smith Cloud magnetic field as discussed in section 1.3 is marked with a red line, and the upper bound as discussed by the Grønnow et al., simulations in section 1.2.1 is marked in blue. Both values at 8 and 3 μG respectively act as upper bounds for the strength of these HVCs.

From figure 4.5, the magnetic field strengths are overall higher compared to the hypothesised estimate of being on the order of magnitude of 0.1 μG . The three methods appear to visually agree with each other approximately, with the KS-EDF overestimating the magnetic field value, the Variance Subtraction having the greatest spread, and the Weighted Average having the largest tail. It is important to note that due to the logarithmic scale used, the mean is inherently biased to visually appear at higher values. However, the use of a logarithmic scale is still justified due to the order-of-magnitude levels of error.

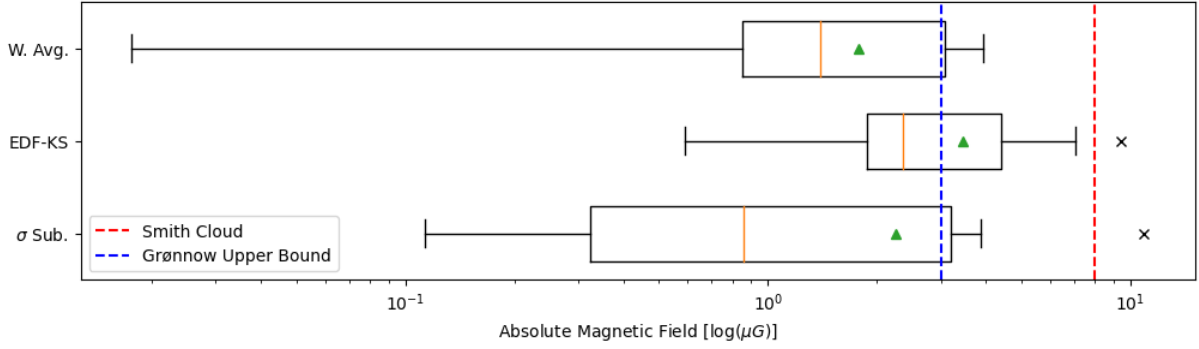


Figure 4.5: A boxplot representing the absolute-value magnetic field of each HVC for both derivational methods. The red and blue dotted lines indicate two upper bounds, one set by the Grønnow et al., simulations, and the other by the Smith Cloud respectively.

4.5.1 Uncertainties

With each method comes a unique derivation of uncertainties. The uncertainty in the weighted mean is given as trivial – calculated separately for the HVC and background populations, and then propagated.

When calculating RM uncertainties, an intrinsic uncertainty of 7 must be added. This value comes from Schnitzeler [2010], which details its importance as a means of accounting for extragalactic uncertainty contributions.

The uncertainty in the KS-EDF statistic reflects the inherent variability in the method and is calculated by the Euclidean addition of the average measurement uncertainty of both the HVC and background populations. This corresponds to the propagation of RM uncertainties shown in figure 4.4, where the black line is the source of determining the uncertainties to propagate.

The uncertainty in the variance subtraction method was calculated via bootstrap resampling with replacement. The number of samples generated per population is equal to the size of the population being analysed. Once the HVC and background population uncertainties were evaluated, the two uncertainties were propagated to produce the listed result.

4.5.2 Statistical Comparison of Methods

Two statistical tests were used to compare the three methods in-between each other. While typically, a chi-squared test may be used, because there is no ‘expected’ model to compare observations to, it is difficult to designate the variables in the chi-squared test.

A weighted ANOVA test was employed as it can compare multiple methods

Methods		Difference (μG)			p-value
		Mean	Upper Bd.	Lower Bd.	
Var. Sub.	KS-EDF	-1.199	-2.783	0.3853	0.8259
Wgt. Mean	KS-EDF	-1.675	-3.259	-0.09067	0.9643
Wgt. Mean	Var. Sub.	-0.476	-2.06	1.108	0.2455

Table 4.3: A table describing the Tukey Multiple Comparison of Means test results for the HVCs with magnetic field detection.

against each other, with the weights allowing for the accounting of uncertainties. The F-Statistic calculated from the weighted ANOVA test was 3.375, which corresponds to a p-value of 0.0388. This p-value is within a 95% confidence interval of similarity, but is still fairly weak.

The weighted ANOVA test results were also fed into a Tukey multiple comparisons of means test. The results of this test are shown in table 4.3. Note that here, the p-value is a measure of the likelihood that the two data sets are different. It is clear that an alternative hypothesis, that two given distributions are different, is not met to 95% confidence in all cases except when comparing the weighted mean and KS-EDF method. However it is clear that the KS-EDF method does not match well with the other two cases in-general.

The reason for the weighted ANOVA test as a whole giving a weak significance of similarity compared to only one of three values in the Tukey test giving strong evidence of similarity, is due to the ability of the weighted ANOVA test to 'bridge the gaps' created by two data points with no overlapping uncertainties and a third data point in the centre. This is why a three-way comparison is needed.

The primary three sources for the weighted ANOVA and Tukey test is provided in the papers by Kutner et al. [2005]; Tukey [1949]; Montgomery [2019]. The specific code to perform this analysis is additionally presented in E.

Discussion

There have been several caveats made in the explanation of the data collection, methodology, results, and analysis of results. The goal of this section is to summarize such information (involving a re-invoking of certain statements made in previous sections) and to build on that. Like the rest of the report structure, this section is split under the two different research questions investigated in this thesis, starting with the secondary goal of the report and finishing with the primary goals of the report.

5.1 Foreground Removal

A major shortfall of the foreground subtraction investigation is the lack of statistical analysis, which will be a common theme throughout this section. In some cases, this is justified, as statistical analysis does not need to be performed especially when the method visually appears as adequate. At the same time, this investigation is limited from a lack of characterisation of the true RM sky, including simulations to predict the distribution of corrected RMs and high-quality data to interpolate with.

This report is not meant to act as "closing the book" on research into foreground subtraction techniques, but to point out important considerations to make as research moves into the SKA era, and to conduct a preliminary analysis of how to tackle these predicted problems. This comes in the explanation in section 3.1, which highlights how higher-definition data will inevitably result in issues involving foreground correction for objects of a particular angular size. Thus, the need to extend beyond interpolation techniques.

The NUFFT method is one example of a method that can be assessed visually, and without the need for statistical analysis. Figure 3.2 clearly shows that even carrying out an interpolation using a NUFFT requires a large sample size to be adequate. It would not be appropriate to completely write-off NUFFTs as an interpolation and bandpass technique, as with a more complete POSSUM dataset, and future SKA RM projects, millions of RM points would be adequate to interpolate and bandpass using a NUFFT. There will simply not be enough data in the few years after this paper.

Given that the Hutschenreuter map interpolates data that is at a spatial sampling frequency of 1 sampling point per degree squared, attempting to remove objects on the scale of 1 degree should not alter the interpolation at all. However,

there are several factors that prevents this from being an accurate description of the Hutschenreuter map: H-alpha inclusion providing more detailed information; more RM sources being detected by telescopes close to the Galactic midplane; and distortion effects occurring on the edges of the cartesian map. The latter two can be accounted for by performing the filtering on a reduced form of the dataset that does not include extremes in Galactic latitude. Additionally, from figure 2.3, it is clear that RMs near the Galactic midplane can become quite scattered. Even though this is true for the Hutschenreuter map, it still is a demonstration of the scatter present near the Galactic midplane. However, filtering not performed with the statistical analysis of foreground removal techniques, as it is important to determine the effectiveness of a given method at accounting for these changes.

This is the justification for the use of a Pearson's R-squared statistic, as it can determine how similar the pre and post bandpass and convolution methods are at preserving the broader RM sky. With the expectation being that the R-squared statistic is close to unity but not necessarily exactly unity. One major downside is that there is no way to quantify the appropriate R-squared statistic that should be desired. Thus, only a high R-squared statistic is possible qualifier.

It is clear from the R-squared statistic that, even with a less-exact method of determining validity, only the Crosshatch-Bandpass method is viable as an interpolation correction method. It can be pointed out that it may be unfair to compare the Annulus-Convolved method to the raw interpolation. However, it is justified in the sense that the raw interpolation has been used as the primary method in past literature, and from figure 4.3, it is demonstrably useful at removing the anisotropy produced by the ISM.

Both annulus methods had major points of failure regarding the specific technique that they relied on to correct the RMs. When the RM sky is meant to be full of objects and obstructions of different sizes and overlapping positions, it can make it difficult to use a fixed size or RM count to correct. The RM distributions in figure 4.2 show visually that a method involving the averages of a set of RMs, and corrections based on that average are possible, but the effective 'goldilocks zone' for a good size or sampling count is small and would require large simulations to find the ideal size – even then, it may not be adequate for certain objects.

The Crosshatch-Bandpass method seems to be the most appropriate as a first step in SKA-era foreground interpolation. However, with the rippling effects shown in figure 3.3, there are improvements that can be made. The Fourier Transform of a top-hat function is inevitably going to introduce rippling. This may not be as much of the case with the Tukey or Gaussian windows, which is much more common in other signal processing analyses. The choice to reduce bandpass opacity is another option that clearly removes ripples, at the cost of dampening the intended goal of using FFT signal processing methods. Some combination of these may be useful for future research relying on RM-correction.

The interesting result provided in Anderson et al. [2024] clearly still shows with a repeated analysis on the Hutschenreuter map, and the improved Crosshatch-Bandpass method. There is still little reason for why corrected RMs take this shape, or if this is meant to be the true distribution of RMs post ISM correction. It is rather crucial that research be conducted into what the true distribution of ISM-corrected RMs should be. It is trivial to understand why the distribution should be symmetrical and centred at zero, as extragalactic polarisation is effectively random and not biased based on sign. The Student's t-distribution is one of many different distributions satisfying these properties that are viable candidates for representing corrected RMs, it needs to be explained why this specific distribution is the one that RMs follow.

Under the assumption that it is already known why and expected that the t-distribution is a correct reflection of reality, both the original and the Crosshatch-Bandpass filtered Hutschenreuter map align with the expected t-distribution incredibly well, with the parameters for the t-distribution fit being similarly identical. As mentioned in section 3.5, despite the clear visual incongruence, the important property of the t-distribution which appeared to not work with other distributions centred at zero, is the tailed prevalence. This ultimately gave such a small p-value in the Pearson's chi-squared test. Both distributions additionally are centred close to zero, with the systematic error being most likely a consequence of measurement bias or a consequence of the POSSUM dataset covering only a specific proportion of the sky. The latter can be seen in how the ISM's anisotropic distribution does not seem exactly symmetric in figure 4.3, connecting it with the amount of RMs recorded by POSSUM in a specific quadrant of the southern hemisphere, seen in figure 2.1.

As mentioned, the results in this report regarding foreground techniques are results which may not immediately show utility in the next year or two. However, with more accurate interpolation techniques, like the one seen in Khadir et al. [2024], the pre-SKA and SKA-era analyses of the RM sky will greatly benefit from the preliminary research direction opened by this report.

5.2 Magnetic Field Derivation

A key theme in this report is the distinction between validity and accuracy of a particular given method. As is common with first attempts as estimation, the lack of data and methodological development is apparent. This is the primary motivation for the report's focus on less on agreement and accuracy, but more on consideration of new methods and their mathematical validity. There is plenty of future work available, particularly in-terms of expanding upon the methodology of this report.

5.2.1 Data Collection

The current POSSUM sample size is a major hinderance to achieving the stated re-search objectives. With the survey in its early stages, a significant proportion of the sky is yet to be covered. This does not prevent viable analysis from being performed, however it does limit the scope of the analysis, and can affect the accuracy of the evaluation. A big proportion of documented HVCs could not be analysed as they will not or currently are not covered by the POSSUM dataset as of May 2024.

The use of the HI sky map from [SOURCE] is justified for the use of visual display and visual HVC detection, with the map being used to eliminate two HVCs for a lack of detection (HVC [NUM]) and cartesian map distortions (HVC [NUM]). However, the choice to use a HI map which filters out HI emission below a VLSR threshold (equivalent to a given column density threshold) can make it difficult to utilise the HI map for masking – a common method of isolating in-HVC and out-HVC RM populations. It is potentially difficult to do this even with a non-filtered HI map, as the boundary between what is and is not the HVC will always be difficult to discern. Masking would only be a first step, with the need to include RMs near but not on the HVC's HI boundary due to magnetic draping occurring outside the physical cloud itself. Any algorithm that can adapt itself to outline the specific shape of any given HVC and select the RMs within and around said outline would have to be highly complex and may even have to rely on new Artificial Intelligence techniques to account for the various shapes and sizes of HVCs.

The use of the Hutschenreuter map is justified given the use of interpolations in past research [SOURCE], and the analysis performed all throughout chapter [SE-CREF]. It is clear as well that the Hutschenreuter map can eliminate the most important source of foreground interference – that being the ISM. This is evidenced indirectly by figure [FIG], with the elimination of the ISM's key feature being the innate anisotropy. It is expected that what remains after correction is still rather variant, as the interpolation is likely to add its own level of standard measurement error, and that the halo itself is magnetised to some degree, which is shown in the ridge-like structures in figure [FIG].

Removing the ridge-like structures is much more difficult and must be done after or during the calculation of the master RMs or line-of-sight magnetic field strengths. In concerns with determining a significant detection however, it is not of concern. This is because the HVC contribution to the RMs effectively 'overlaps' and adds to the background RM field, meaning that a detection can still be made even if the background interferes with the HVC. The exception to this rule is if the background is so noisy or strong, that any variation of RMs created by the HVC is eliminated. While this is rare, it can potentially bias results towards HVCs with notably higher line-of-sight magnetic field strengths, an effective form of Malmquist bias. Despite this, there is an argument to be made that HVCs which do not have significant contributions

compared to the background should be thrown out in the calculation process, as there is naturally less evidence to verify if the HVC is experiencing magnetic draping, and that this phenomenon is calculable. Since most HVCs did have significant detections, however, it is not of concern to the outcome of specifically this report.

From sections [SECREf] and [SECREf], there is quite a variance in HVC size, albeit a smaller variance for HVCs in the CGM. The use of angular size to discriminate against HVCs that are abnormally large or small is reliant on the assumption that all HVCs are approximately the same distance away from the Earth as each other. Within an order of magnitude, this is assumedly true as all HVCs capable of being resolved are likely to be in the Galactic halo or nearby CGM [SOURCE]. Additionally, HVCs that are much further away from or closer to the Galaxy are likely to be malformed in shape [SOURCE] – key example being the Smith Cloud [SOURCE]. This means that these outlier HVCs are likely to have angular size distributions well outside [NUM] degrees.

All these methods constrain the total amount of HVCs being analysed. Which is how a resulting 13 HVCs were selected. This limits the sample size, which can directly affect the quality of statistical analysis. Hence, an additional reason to label the findings of this report as a rough estimation.

5.2.2 Necessity of Assumptions

As mentioned in section [SECREf], four major assumptions were employed in the analysis post-HVC selection and RM correction: (1) the Galactic halo is well-mixed; (2) the weighted average of a given HVC's HI column density is approximately equivalent to its peak value; (3) HVCs are approximately spherical; and (4) the RM contribution along the line of sight is an additive. There are both reasons for the necessity of these assumptions, and clear blind spots that they present.

While there is evidence that the Galactic halo has an ionisation fraction close to unity [SOURCE], the main reason for this assumption is that there is very little, and often conflicting evidence as to the true ionisation fraction of the halo, if it is homogeneous at all. There is also evidence to suggest that the gas surrounding HVCs themselves has very complex mixing patterns Heitsch et al. [2021]. The ridge-like structures found in the RM background of each HVC image is evidence to suggest a lack of homogeneity, however, it could be a consequence of turbulence captured at an instantaneous moment as well. Until there is better data on the ionisation fraction surrounding HVCs and in the halo itself, the results are likely to be inaccurate to within an order of magnitude. Note that by decreasing the ionisation fraction, the estimated magnetic field strengths become increasingly larger.

The assumption of homogeneity appears in the second assumption as well, with the core idea being that one can assume that the 'master RM' is effectively a measure of the RM passing through the centre of an ideally homogenous HVC. The need for

this assumption is due to the lack of a method to truly find the correct HI outline of the HVC and the use of the HI map from [SOURCE]. As well, the Moss catalogue only provides the peak HI column density, and all the RM data must be condensed into a single data point – thus a single-point analysis needs to be used regardless. Statistically speaking, the weighted average is likely to be biased towards higher HI column densities due to Poisson noise, and given the equation from [SOURCE], equation [EQREF] – which relies on and approximates using this weighted average - this assumption appears to be much more reasonable than it may initially seem. The shortfall of this assumption comes from the use of the background overlap as correction, and the use of a circular selection window. With the background overlap acting as a correction to the HVC’s ‘master RM’, it means that the HI column density calculation is also in-effect being applied to the RMs surrounding the HVC, which is certainly not an accurate representation of the HI column density in that region. The circular selection window also, as required, included RMs which surround but do not overlap with the HVC’s HI profile, meaning that these RMs are analysed with the same column density as the RMs overlapping the HVC. Both issues are mitigated by the fact that the RMs are only converted to line-of-sight magnetic fields once a ‘master RM’ is created – the ideal being that issues surrounding electron density are dealt with once a single RM value has been reached.

The third and fourth assumptions are more trivial in-terms of their purpose. Without the use of masking, and with the requirement of a generalised algorithm to distinguish in-HVC and out-HVC populations, the most logical first step is to pick a simple shape to act as the outline. Given that HVCs are comet-like structures [SOURCE], it is also reasonable to estimate the main bulb of HVCs as a sphere. The shortfall of this assumption is simply in how it may not correctly distinguish RM populations in more malformed HVCs, which as mentioned, are less likely to appear due to the method of HVC selection chosen. The last assumption is used very often in past research [SOURCE]. Thus, it is not of great concern in how it may affect the outcomes of this report, it is a fair approximation to make, given repeated amounts of correction for non-HVC interferences.

5.2.3 Validity of Detection Methods

As mentioned, the Hutschenreuter map at mid-latitudes closely emulates the ideal of having variations on one square degree. This may be problematic, as it means that the Hutschenreuter map may introduce variations in the corrected RM grid on such a scale. In fact, the ridge-like structures of the HVCs may be caused by spatial variations in interpolated data instead of it representing the true non-ISM foreground. This is why a visual comparison is made between the corrected RMs and the RMs post-average subtraction in figures [FIG] and [FIG]. By taking an average of all the RMs within a neighbourhood of the HVC and subtracting it out and confirming that

it looks visually like the corrected RMs, it can be demonstrated that the Hutschenreuter map is like subtraction using a constant factor i.e. the Hutschenreuter introduces no spatial variations because of interpolation. Two statistical tests in this report can further validate the qualitative analysis. Given that the Crosshatch-Bandpassed, and Raw, Hutschenreuter map is correlated significantly (as seen with the R-Squared statistic in table [TABLE]), it means that the Hutschenreuter map is likely not creating additional spatial variations. It can also be seen that in the mid-latitudes in figure [FIG], the linear correlation is stronger. The subtraction of a constant value also is equivalent to the uncorrected RMs in terms of the output of the KS test. Thus, by showing that the KS test simply is more selective for corrected RMs than it is for uncorrected RMs, it quantitatively verifies that these spatial variations are not significantly affecting the results. Note that while the KS test is more selective for corrected RMs, this does not imply an issue with spatial variation, but instead that the Hutschenreuter map is better at correcting the foreground than no foreground correction. Only in the vice versa case where the corrected and uncorrected KS tests retrieve uncorrelated or weaker results does it imply a problem with the Hutschenreuter map itself, as it would mean that spatial variations are introducing correlated interference.

Due to the lack of POSSUM data, even surrounding HVCs that overlap with the existing POSSUM dataset, there is a potential issue created from the Missing Completely at Random (MCAR) data. With this missing data, the full picture of foreground interferences or ridge-like structure variations cannot be determined and can cause the KS test to flag detections where detections may otherwise not be. While there are statistical methods to quantify the effect of MCAR data, such as Little's MCAR test [SOURCE], it is far too speculative to essentially predict what the full POSSUM dataset may look like and is much more worth waiting a few more years to repeat the analysis performed by this report.

From section [SECREf], there appears to be a weak correlation between the results of the KS test and the HVC's Galactic latitude ([NUM]). Fortunately, this correlation is even weaker when comparing with the p-value ([NUM]). Any level of correlation implies that the KS test itself is biased towards HVCs of a particular Galactic latitude, which can damage the credibility of detections made by the KS test. Part of this could be due to a low sample size or a limited Galactic latitude range, which may artificially increase the correlation – in this case there would be no bias. Another possibility may be that due to RMs at mid-latitudes being 'tamer', the KS test is more robust at determining a difference i.e. there is less noise in the calculated background – this case is what may introduce bias. Due to the small sample size of HVCs, whether the KS test is correlated with latitude cannot be determined, but it is important for future researchers to pay attention to this problem. There does not appear to be a problem with other HVC properties like VLSR or VGSR.

A main concern is if the ridge-like structures are a consequence of noise or spatially

significant systematic error that is not caused by the interpolation itself. Or that the background noise may be different depending on the portion of sky being analysed. Typically, to remove the sky, astronomers would look at a random and isolated patch of sky to compare against. This method is not viable as the background population of each HVC is expected to be unique, and dissimilar to any arbitrarily sampled portion of the POSSUM dataset, due to the potential presence of RM foreground interferences overlapping with a specific HVC. Thus, there is no point in comparing the background population with a separate portion of RM sky.

5.2.4 Validity of Derivation Methods

To further distinguish the differences between accuracy and validity is the notion of mathematical, statistical, and experimental validity. Because a method may be valid in a mathematical abstract does not mean that it may be experimentally valid. Mathematical derivations necessarily introduce small assumptions, on top of the major four listed in the report. These assumptions may cause the method to fail when put into practice. Here, there is an expectation that all three derivation methods (Variance Subtraction, KS-EDF, Weighted Mean) have specific points of failure for any given scenario. For example, the weighted mean mathematically underestimates for situations where the HVC is face-on. This is to such a predicted effect that in section [SECREf], it was hypothesised that the weighted mean of RMs is likely to return bad data. This was confirmed false by the weighted mean having a reasonable magnetic field value range in figure [FIG] – although it does appear to return small values for some HVCs indicated by the lower quartile whisker of the box plot. The KS-EDF method overestimates due to the strongest deviation between populations likely occurring after a stationary point on the PDF, or inflexion point on the EDF, due to a delayed rise in the cumulative proportion. As a visual representation, note that the location of the x-statistic in figure [FIG] is off its nearest inflexion point by about 0.5-1 [UNIT]. All three tests in mathematically abstract and idealised situations work for most or some cases they may face. Which is why to statistically validate the results, a three-way method of analysis is required.

The use of the Weighted ANOVA test is crucial to the statistical validation of data. It allows for a three-way simultaneous comparison of the calculated values for each of the three methods while accounting for their uncertainties. If the Weighted ANOVA test returns a significance, it means that despite the apparent points of failure each method may have for a particular HVC, and the variations between each of the three methods, in general, the three methods agree with each other to some degree. Here, there is a weak agreement between each of the three methods – resulting in a 2-sigma confidence. This not as strong as a result as one may desire, but it is a starting point that demonstrates all three methods are headed in the right direction.

The use of the Tukey test is meant to compare each method with each other

[SOURCE]. This two-way comparison can provide greater insight into which methods are more likely to agree or disagree with each other. As from section [SECREf], it is quite apparent that the KS-EDF method is the worst performing of the three methods, which can be seen both visually and numerically. When comparing the KS-EDF method against the weighted mean, a 95% confidence of statistically significant difference is obtained – which is particularly problematic given that the weighted mean is sanctioned by research [SOURCE]. From this, it is reasonable to conclude that the KS-EDF method is the odd one out of the three and should be disregarded in future research. The variance subtraction, as far as this analysis is concerned, is statistically valid to some degree. However, more needs to be done to confirm this fact.

Despite the statistical confirmation of validity, there is ultimately no way to determine the experimental validity of each of these three methods without modelling. This is a recommended direction for cementing the results of the report.

As expected, from table [TABLE] the uncertainties are relatively large. There is no decent way at visually representing this statement due to how cluttered any given graph would be. The large uncertainties are expected. Certainly, for both variance subtraction and weighted mean, the calculation of uncertainties are mathematically justified. It was much more difficult to obtain KS-EDF uncertainties, leading to a potential over-estimation of error. This only further supports the conclusion that the KS-EDF method is a ‘dead end’.

The methodologies proposed in chapter [SECREf] play a role in how researchers will estimate the effect of magnetic draping, with the clear requirement of more techniques or ‘angles of attack’ to solving the problem of calculation for an arbitrary HVC, of arbitrary orientation, size, shape, and angular size. Effectively the goal of developing a generalised algorithm as discussed in section [SECREf] is necessary to allow for astronomers to mass-estimate magnetic draping. An important caveat being that the four major assumptions being reduced in their contribution to systematic or random errors.

Conclusions

E

Also comment on projects in current development. Including Tahani et al. [2024]; Li [2013]; Bron, Emeric et al. [2021].

Bibliography

- ANDERSON, C. S., MCCLURE-GRIFFITHS, N. M., RUDNICK, L., GAENSLER, B. M., O'SULLIVAN, S. P., BRADBURY, S., AKAHORI, T., BAIDOO, L., BRUGGEN, M., CARRETTI, E., DUCHESNE, S., HEALD, G., JUNG, S. L., KACZMAREK, J., LEAHY, D., LOI, F., MA, Y. K., OSINGA, E., SETA, A., STUARDI, C., THOMSON, A. J. M., ECK, C. V., VERNSTROM, T., AND WEST, J., (2024). Probing the magnetised gas distribution in galaxy groups and the cosmic web with possum faraday rotation measures. *Monthly Notices of the Royal Astronomical Society*, 533(4), 4068–4080. (cited on pages 26 and 43)
- ARIAS-CASTRO, E. AND DONOHO, D. L., (2009). Does median filtering truly preserve edges better than linear filtering? *The Annals of Statistics*, 37(3), 1172–1206. doi:10.1214/08-AOS604. (cited on page 28)
- BAGCHI, S. AND MITRA, S., (1996). The nonuniform discrete fourier transform and its applications in filter design. ii. 2-d. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 43(6), 434–444. doi:10.1109/82.502316. (cited on page 19)
- BECK, A. M., LESCH, H., DOLAG, K., KOTARBA, H., GENG, A., AND STASYSZYN, F. A., (2012). Origin of strong magnetic fields in milky way-like galactic haloes. *Monthly Notices of the Royal Astronomical Society*, 422(3), 2152–2163. doi:10.1111/j.1365-2966.2012.20759.x. (cited on page 3)
- BETTI, S. K., HILL, A. S., MAO, S. A., GAENSLER, B. M., LOCKMAN, F. J., MCCLURE-GRIFFITHS, N. M., AND BENJAMIN, R. A., (2019). Constraining the magnetic field of the smith high-velocity cloud using faraday rotation. *The Astrophysical Journal*, 871(2), 215. doi:10.3847/1538-4357/aaf886. (cited on pages 4, 5, 16, and 34)
- BLAND-HAWTHORN, J. AND MALONEY, P. R., (1999). The galactic halo ionizing field and h alpha distances to hvcs. *arXiv*. doi:10.48550/arXiv.astro-ph/9812297. (cited on page 2)
- BLITZ, L., SPERGEL, D. N., TEUBEN, P. J., HARTMANN, D., AND BURTON, W. B., (1999). High-velocity clouds: Building blocks of the local group. *The Astrophysical Journal*, 514(2), 818. doi:10.1086/306963. (cited on pages 1 and 2)
- BRENTJENS, M. A. AND DE BRUYN, A. G., (2005). Faraday rotation measure synthesis. *Astronomy & Astrophysics*, 441(3), 1217–1228. doi:10.1051/0004-6361:20052990. (cited on page 5)

- BRON, EMERIC, ROUEFF, EVELYNE, GERIN, MARYVONNE, PETY, JÉRÔME, GRATIER, PIERRE, LE PETIT, FRANCK, GUZMAN, VIVIANA, ORKISZ, JAN H., DE SOUZA MAGALHAES, VICTOR, GAUDEL, MATHILDE, VONO, MAXIME, BARDEAU, SÉBASTIEN, CHAINAIS, PIERRE, GOICOECHEA, JAVIER R., HUGHES, ANNIE, KAINULAINEN, JOUNI, LANGUIGNON, DAVID, LE BOURLOT, JACQUES, LEVRIER, FRANÇOIS, LISZT, HARVEY, ÖBERG, KARIN, PERETTO, NICOLAS, ROUEFF, ANTOINE, AND SIEVERS, ALBRECHT, (2021). Tracers of the ionization fraction in dense and translucent gas - i. automated exploitation of massive astrochemical model grids. *Astronomy & Astrophysics*, 645, A28. doi:10.1051/0004-6361/202038040. (cited on page 50)
- DURSI, L. J. AND PFROMMER, C., (2008). Draping of cluster magnetic fields over bullets and bubbles—morphology and dynamic effects. *The Astrophysical Journal*, 677(2), 993. doi:10.1086/529371. (cited on page 3)
- FERRIÈRE, K. M., (2001). The interstellar environment of our galaxy. *Reviews of Modern Physics*, 73(4), 1031–1066. doi:10.1103/RevModPhys.73.1031. (cited on page 5)
- FINKBEINER, D. P., (2003). A full-sky $h\alpha$ template for microwave foreground prediction. *The Astrophysical Journal*, 146(2), 407. doi:10.1086/374411. (cited on pages vii, 2, 4, 12, and 13)
- GAENSLER, B. M., HEALD, G. H., MCCLURE-GRIFFITHS, N. M., ANDERSON, C. S., ECK, C. L. V., WEST, J. L., THOMSON, A. J. M., LEAHY, J. P., RUDNICK, L., MA, Y. K., AKAHORI, T., GÜRKAN, G., LANDECKER, T. L., MAO, S. A., O’SULLIVAN, S. P., RAJA, W., SUN, X., VERNSTROM, T., BAIDOO, L., CARRETTI, E., TAYLOR, A. R., WILLIS, A., OSINGA, E., LIVINGSTON, J. D., AN, T., BRACCO, A., BRÜGGEN, M., ESWARAIHAH, C., ENSSLIN, T., UPASANA, G., IDEGUCHI, S., JUNG, S. L., KACZMAREK, J. F., KOTHES, R., LAZAREVIĆ, S., LEAHY, D., LOI, F., PANDHI, A., PRICE, J. M., SHEN, A. X., SETA, A., SOBEY, C., STUARDI, C., AND VELOVIĆ, V., (2024). The polarisation sky survey of the universe’s magnetism (possum). i. science goals and survey description. In preperation. (cited on pages 4, 7, 10, and 11)
- GAENSLER, B. M., LANDECKER, T. L., AND TAYLOR, A. R., (2010). Survey science with askap: Polarisation sky survey of the universes’s magnetism (possum). In *Bulletin of the American Astronomical Society*. (cited on pages 4, 10, and 11)
- GREENGARD, L. AND LEE, J.-Y., (2004). Accelerating the nonuniform fast fourier transform. *Society for Industrial and Applied Mathematics Reviews*, 46(3), 443–454. doi:10.1137/S003614450343200X. https://doi.org/10.1137/S003614450343200X. (cited on page 19)
- GRØNNOW, A., TEPPER-GARCÍA, T., AND BLAND-HAWTHORN, J., (2018). Magnetic fields in the galactic halo restrict fountain-driven recycling and accretion. *The Astrophysical Journal*, 865(1), 64. doi:10.3847/1538-4357/aada0e. (cited on pages 3 and 4)

-
- GRÖNNOW, A., TEPPER-GARCÍA, T., BLAND-HAWTHORN, J., AND FRATERNALI, F., (2022). The role of the halo magnetic field on accretion through high-velocity clouds. *Monthly Notices of the Royal Astronomical Society*, 509(4), 5756–5770. doi:10.1093/mnras/stab3452. (cited on pages 3 and 4)
- GRÖNNOW, A., TEPPER-GARCÍA, T., BLAND-HAWTHORN, J., AND MCCLURE-GRIFFITHS, N. M., (2017). Magnetized high velocity clouds in the galactic halo: A new distance constraint. *The Astrophysical Journal*, 845(1), 69. doi:10.3847/1538-4357/aa7ed2. (cited on pages ix, 3, 4, 7, 29, 38, and 39)
- HAN, J. L. AND QIAO, G. J., (1994). The magnetic field in the disk of our galaxy. *Astronomy & Astrophysics*, 288, 759–772. <https://ui.adsabs.harvard.edu/abs/1994AnA..288..759H>. (cited on page 3)
- HAYAKAWA, T. AND FUKUI, Y., (2024). Dust-to-neutral gas ratio of the intermediate- and high-velocity H I clouds derived based on the sub-mm dust emission for the whole sky. *Monthly Notices of the Royal Astronomical Society*, 529(1), 1–31. doi:10.1093/mnras/stae302. (cited on page 3)
- HEITSCH, F., MARCHAL, A., MIVILLE-DESCHÊNES, M.-A., SHULL, J. M., AND FOX, A. J., (2021). Mass, morphing, metallicities: the evolution of infalling high velocity clouds. *Monthly Notices of the Royal Astronomical Society*, 509(3), 4515–4531. doi:10.1093/mnras/stab3266. <https://doi.org/10.1093/mnras/stab3266>. (cited on pages 29 and 45)
- HEITSCH, F. AND PUTMAN, M. E., (2009). The fate of high-velocity clouds: Warm or cold cosmic rain? *The Astrophysical Journal*, 698(2), 1485. doi:10.1088/0004-637X/698/2/1485. (cited on page 3)
- HILL, A. S., HAFFNER, L. M., AND REYNOLDS, R. J., (2009). Ionized gas in the smith cloud. *The Astrophysical Journal*, 703(2), 1832. doi:10.1088/0004-637X/703/2/1832. (cited on pages 2, 3, and 7)
- HILL, A. S., MAO, S. A., BENJAMIN, R. A., LOCKMAN, F. J., AND MCCLURE-GRIFFITHS, N. M., (2013). Magnetized gas in the smith high velocity cloud. *The Astrophysical Journal*, 777(1), 55. doi:10.1088/0004-637X/777/1/55. (cited on pages 4, 5, 7, 9, 16, 30, and 34)
- HOTAN, A. W., BUNTON, J. D., CHIPPENDALE, A. P., WHITING, M., TUTHILL, J., MOSS, V. A., MCCONNELL, D., AMY, S. W., HUYNH, M. T., ALLISON, J. R., AND ET AL., (2021). Australian square kilometre array pathfinder: I. system description. *Publications of the Astronomical Society of Australia*, 38, e009. doi:10.1017/pasa.2021.1. (cited on page 10)

-
- HUANG, T., YANG, G., TANG, AND G., (1979). A fast two-dimensional median filtering algorithm. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 27(1), 13–18. doi:10.1109/TASSP.1979.1163188. (cited on page 28)
- HUTSCHENREUTER, S., ANDERSON, C. S., BETTI, S., BOWER, G. C., BROWN, J.-A., BRÜGGEN, M., CARRETTI, E., CLARKE, T., CLEGG, A., COSTA, A., CROFT, S., ECK, C. V., GAENSLER, B. M., DE GASPERIN, F., HAVERKORN, M., HEALD, G., HULL, C. L. H., *, INOUE, M., JOHNSTON-HOLLITT, M., KACZMAREK, J., LAW, C., MA, Y. K., MACMAHON, D., MAO, S. A., RISELEY, C., ROY, S., SHANAHAN, R., SHIMWELL, T., STIL, J., SOBEY, C., O’SULLIVAN, S. P., TASSE, C., VACCA, V., VERNSTROM, T., WILLIAMS, P. K. G., WRIGHT, M., AND ENSSLIN, T. A., (2022). The galactic faraday rotation sky 2020. *Astronomy & Astrophysics*, 657(A43). doi:10.1051/0004-6361/202140486. (cited on pages vii, 4, 10, 11, 12, 13, and 16)
- HUTSCHENREUTER, S. AND ENSSLIN, T. A., (2020). The galactic faraday depth sky revisited. *Astronomy & Astrophysics*, 633(A150). doi:10.1051/0004-6361/201935479. (cited on pages vii, 4, 10, 11, 12, 13, and 16)
- JONES, T. W., RYU, D., AND AND, I. L. T., (1996). The magnetohydrodynamics of supersonic gas clouds: Mhd cosmic bullets and wind-swept clumps. *The Astrophysical Journal*, 473(1), 365. doi:10.1086/178151. (cited on pages 3 and 4)
- JUNG, S. L., GRØNNOW, A., AND MCCLURE-GRIFFITHS, N. M., (2022). Magnetic field draping around clumpy high-velocity clouds in galactic halo. *Monthly Notices of the Royal Astronomical Society*, 522(3), 4161–4180. doi:10.1093/mnras/stad1236. (cited on pages 3 and 4)
- JUNG, S. L., MCCLURE-GRIFFITHS, N. M., AND HILL, A. S., (2021). Distant probes of rotation measure structure: where is the faraday rotation towards the magellanic leading arm? *Monthly Notices of the Royal Astronomical Society*, 508(3), 3921–3935. doi:10.1093/mnras/stab2773. (cited on page 16)
- JUNG, S. L., MCCLURE-GRIFFITHS, N. M., PAKMOR, R., MA, Y. K., HILL, A. S., ECK, C. L. V., AND ANDERSON, C. S., (2023). Sampling the faraday rotation sky of tng50: Imprint of the magnetised circumgalactic medium around milky way-like galaxies. *Monthly Notices of the Royal Astronomical Society*, 526(1), 836–853. doi:10.1093/mnras/stad2811. (cited on page 3)
- JUNG, S. L., SETA, A., PRICE, J., MCCLURE-GRIFFITHS, N., LIVINGSTON, J., GAENSLER, B., MA, Y. K., TAHANI, M., ANDERSON, C., FEDERRATH, C., VAN ECK, C., LEAHY, D., O’SULLIVAN, S., WEST, J., HEALD, G., AND AKAHORI, T., (2024). Magnetised hi superbubbles in the small magellanic cloud revealed by the possum pilot survey. Submitted to MNRAS. (cited on page 35)

-
- KACZMAREK, J. F., PURCELL, C. R., GAENSLER, B. M., McCLURE-GRIFFITHS, N. M., AND STEVENS, J., (2017). Detection of a coherent magnetic field in the magellanic bridge through faraday rotation. *Monthly Notices of the Royal Astronomical Society*, 467(2), 1776–1794. doi:10.1093/mnras/stx206. (cited on pages 1, 5, and 30)
- KAWAGUCHI, I., (1952). On the excitation and ionization temperature of the hydrogen in the chromosphere. *Publications of the Astronomical Society of Japan*, 4, 131. (cited on page 3)
- KHADIR, A., PANDHI, A., HUTSCHENREUTER, S., GAENSLER, B. M., VANDERWOUDE, S., WEST, J. L., AND O’SULLIVAN, S. P., (2024). Choosing interpolation techniques for reconstructing galactic faraday rotation. Submitted to ApJ. (cited on pages 11, 16, and 43)
- KONZ, C., BRÜNS, C., AND BIRK, G. T., (2002). Dynamical evolution of high velocity clouds in the intergalactic medium. *Astronomy & Astrophysics*, 391(2), 713–723. doi:10.1051/0004-6361:20020863. (cited on pages vii, 2, 3, and 4)
- KUTNER, M. H., NACHTSHEIM, C. J., NETER, J., AND LI, W., (2005). *Applied Linear Statistical Models*. McGraw-Hill/Irwin, 5th edition. (cited on page 40)
- LI, C., (2013). Little’s test of missing completely at random. *The Stata Journal*, 13(4), 795–809. doi:10.1177/1536867X1301300407. (cited on page 50)
- LOCKMAN, F. J., (2008). *High-Velocity Clouds Merging with the Milky Way*, chap. 53, 239–242. Springer. ISBN 978-1-4020-6932-1. (cited on page 7)
- LOCKMAN, F. J., BENJAMIN, R. A., HEROUX, A. J., AND LANGSTON, G. I., (2008). The smith cloud: A high-velocity cloud colliding with the milky way. *The Astrophysical Journal*, 679(1), L21. doi:10.1086/588838. (cited on pages vii, 7, and 9)
- LOI, F., MURGIA, M., GOVONI, F., VACCA, V., PRANDONI, I., BONAFEDE, A., AND FERETTI, L., (2019). Simulations of the polarized radio sky and predictions on the confusion limit in polarization for future radio surveys. *Monthly Notices of the Royal Astronomical Society*, 485(4), 5285–5293. doi:10.1093/mnras/stz350. <https://doi.org/10.1093/mnras/stz350>. (cited on pages vii, 5, and 6)
- LUCCHINI, S., HAN, J. J., HERNQUIST, L., AND CONROY, C., (2024). On the origin of high velocity clouds in the galaxy. Accepted by ApJ. (cited on page 1)
- MACQUART, J.-P., EKERS, R. D., FEAIN, I., AND JOHNSTON-HOLLITT, M., (2012). On the reliability of polarization estimation using rotation measure synthesis. *The Astrophysical Journal*, 750(2), 139. doi:10.1088/0004-637X/750/2/139. <https://dx.doi.org/10.1088/0004-637X/750/2/139>. (cited on pages vii, 5, and 8)

-
- MADSEN, G. J., REYNOLDS, R. J., AND HAFFNER, L. M., (2006). A multiwavelength optical emission line survey of warm ionized gas in the galaxy. *The Astrophysical Journal*, 652(1), 401. doi:10.1086/508441. (cited on pages 2, 3, and 7)
- MAO, S. A., GAENSLER, B. M., HAVERKORN, M., ZWEIBEL, E. G., MADSEN, G. J., MCCLURE-GRIFFITHS, N. M., SHUKUROV, A., AND KRONBERG, P. P., (2010). A survey of extragalactic faraday rotation at high galactic latitude: The vertical magnetic field of the milky way toward the galactic poles. *The Astrophysical Journal*, 714(2), 1170. doi:10.1088/0004-637X/714/2/1170. (cited on pages 3, 4, and 5)
- MCCLURE-GRIFFITHS, N. M., MADSEN, G. J., GAENSLER, B. M., MCCONNELL, D., AND SCHNITZELER, D. H. F. M., (2010). Measurement of a magnetic field in a leading arm high-velocity cloud. *The Astrophysical Journal*, 725(1), 275. doi:10.1088/0004-637X/725/1/275. (cited on pages 1 and 16)
- MONTGOMERY, D. C., (2019). *Design and Analysis of Experiments*. Wiley, 10th edition. (cited on page 40)
- MOSS, V. A., MCCLURE-GRIFFITHS, N. M., MURPHY, T., PISANO, D. J., KUMMERFELD, J. K., AND CURRAN, J. R., (2013). High-velocity clouds in the galactic all sky survey. i. catalog. *The Astrophysical Journal*, 209(1), 12. doi:10.1088/0067-0049/209/1/12. (cited on pages 4, 9, 12, 16, and 30)
- PULFER, E.-M., (2019). *Different Approaches to Blurring Digital Images and Their Effect on Facial Detection*. Honour's thesis, University of Arkansas. (cited on pages viii, 22, and 23)
- PUTMAN, M., PEEK, J., AND JOUNG, M., (2012). Gaseous galaxy halos. *Annual Reviews of Astronomy & Astrophysics*, 50, 491–529. doi:10.1146/annurev-astro-081811-125612. (cited on page 2)
- SCHNITZELER, D. H. F. M., (2010). The latitude dependence of the rotation measures of nvss sources. *Monthly Notices of the Royal Astronomical Society: Letters*, 409(1), L99–L103. doi:10.48550/arXiv.1011.0737. (cited on pages 4, 5, 9, 14, 28, and 39)
- TAHANI, M., NGO, J. M., GLOVER, J., CLAIRMONT, R., ZARAZUA, G. M., AND PLUME, R., (2024). Mc-blos: Determination of the line-of-sight component of magnetic fields associated with molecular clouds. Submitted to ApJ. (cited on page 50)
- TAYLOR, A. R., STIL, J. M., AND SUNSTRUM, C., (2009). A rotation measure image of the sky. *The Astrophysical Journal*, 702(2), 1230. doi:10.1088/0004-637X/702/2/1230. (cited on pages 4 and 10)

-
- TEPPER-GARCÍA, T. AND BLAND-HAWTHORN, J., (2017). The Smith Cloud: surviving a high-speed transit of the Galactic disc. *Monthly Notices of the Royal Astronomical Society*, 473(4), 5514–5531. doi:10.1093/mnras/stx2680. <https://doi.org/10.1093/mnras/stx2680>. (cited on page 7)
- TUKEY, J. W., (1949). Comparing individual means in the analysis of variance. *Biometrics*, 5(2), 99–114. (cited on page 40)
- VANDERWOUDE, S., WEST, J. L., GAENSLER, B. M., RUDNICK, L., VANECK, C. L., THOMSON, A. J. M., ANDERNACH, H., ANDERSON, C. S., CARRETTI, E., HEALD, G. H., LEAHY, J. P., MCCLURE-GRIFFITHS, N. M., O’SULLIVAN, S. P., TAHANI, M., AND WILLIS, A. G., (2024). Prototype faraday rotation measure catalogs from the polarisation sky survey of the universe’s magnetism (possum) pilot observations. *The Astrophysical Journal*, 167(5). doi:10.3847/1538-3881/ad2fc8. (cited on pages 4, 5, 10, and 11)
- WAKKER, B., (1991). High-velocity clouds. *Symposium - International Astronomical Union*, 144, 27–40. doi:10.1017/S0074180900088884. (cited on pages 1 and 2)
- WAKKER, B. P. AND VAN WOERDEN, H., (1997). High-velocity clouds. *Annual Reviews of Astronomy & Astrophysics*, 35, 217–266. doi:10.1146/annurev.astro.35.1.217. (cited on pages 1 and 2)
- WESTMEIER, T., (2018). A new all-sky map of galactic high-velocity clouds from the 21-cm hi4pi survey. *Monthly Notices of the Royal Astronomical Society*, 474(1), 289–299. doi:10.1093/mnras/stx2757. (cited on pages vii, 2, 4, 12, 13, and 16)

Appendix

A Developed code and data

All developed code and data is found in a publically-available GitHub repository as shown here: <https://github.com/Olivex727/hvc-magnetic-honours-programs>.

B All HVCs

The filtered Moss catalogue is as below. The table has been modified from what was originally retrieved for display and formatting purposes. The actual Moss catalogue data can be found at VisieR: <https://vizier.cds.unistra.fr/viz-bin/VizieR-3?-source=J/ApJS/209/12>.

Name	RAJ200	DEJ2000	VLSR	VLSR Error	VGSR	V Dev.
G038.7-37.3-241	21 15 48.32	-11 50 19	-241.4	1.0	-131.9	-203.9
G038.8-33.2-261	21 01 03.36	-10 00 23	-261	1.0	-145.6	-223.5
G043.4-30.4-230	20 58 09.18	-05 16 07	-230.7	1.0	-100.4	-193.2
G089.0-64.7-311	23 56 11.31	-05 36 13	-311.7	1.0	-217.9	-274.2
G133.5-75.6-294	01 02 11.56	-13 00 21	-294.1	1.8	-254.6	-261.6
G248.9+36.8+181	10 08 09.38	-08 21 50	181.6	4.6	17.3	139.1
G261.5-34.5+237	05 22 03.94	-53 57 52	237	2.7	57.8	194.5
G262.4-26.2+143	06 19 11.01	-54 00 25	143.6	1.9	-52	101.1
G264.6-53.3+224	03 17 32.52	-51 33 15	224.2	2.2	93.3	186.7
G267.1-38.4+263	04 51 25.06	-58 06 39	263.9	11.5	91.9	226.4
G271.6-48.4+297	03 32 17.39	-57 45 07	297	1.1	151.1	259.5
G271.7-53.1+254	03 03 24.54	-55 14 30	254.3	2.6	122.4	216.8
G323.0+25.6+122	14 18 59.71	-33 43 42	122.2	1.3	3.1	84.7

Name	FWHM	T_b Fit	N_{HI}	N_{HI} Error	Area	dx	dy
G038.7-37.3-241	12.2	0.87	3.00E+19	1.00E+18	1.5	4.2	3.0
G038.8-33.2-261	20.7	0.94	4.00E+19	2.00E+18	2.2	4.9	3.9
G043.4-30.4-230	22.9	0.64	3.00E+19	2.00E+18	1.6	3.4	3.0
G089.0-64.7-311	18.9	0.5	2.00E+19	2.00E+18	1.4	3.4	3.1
G133.5-75.6-294	26.8	0.42	2.00E+19	3.00E+18	1.3	4.2	1.4
G248.9+36.8+181	22.4	0.13	1.00E+19	2.00E+18	1.1	2.1	1.6
G261.5-34.5+237	26.6	0.28	6.00E+18	3.00E+18	1.5	2.6	2.0
G262.4-26.2+143	19.5	0.28	1.00E+19	2.00E+18	1.6	3.1	2.6
G264.6-53.3+224	22.6	0.28	1.00E+19	2.00E+18	1.7	2.6	2.7
G267.1-38.4+263	30	0.07	9.00E+18	3.00E+18	1.9	3.0	3.0
G271.6-48.4+297	21.5	0.53	2.00E+19	2.00E+18	1.8	2.1	4.6
G271.7-53.1+254	26.4	0.28	1.00E+19	3.00E+18	1.2	2.2	1.8
G323.0+25.6+122	20.8	0.43	2.00E+19	2.00E+18	1.4	2.1	2.4

C Planck Mission Cosmic Microwave Background

The source for the Planck mission's CMB temperature map is available on this website:
https://www.esa.int/ESA_Multimedia/Images/2013/03/Planck_CMB/.

D PyNUFFT Python Module

The pyNUFFT Python Module, which was used in the investigation for foreground removal, has its main documentation page here: <https://pynufft.readthedocs.io/en/latest/index.html>

E Statistical ANOVA Tests

The R language was used to calculate all tests mentioned in 4.5.2, the file is available here: https://github.com/Olivex727/hvc-magnetic-honours-programs/blob/main/KS_confirmation/wgt_anova.R. This R-Code, in addition to the research of the Weighted ANOVA and Tukey HSD test was aided by the ANU SSN Consultancy Group. See the Acknowledgements for more.

The specific R code output when running this file is as follows:

```
> summary(base_V2_aov)
Df Sum Sq Mean Sq F value Pr(>F)
data_new$variable.x  2    44.7   22.349    3.375 0.0388 *
Residuals           87   576.1    6.622
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> tukey.test
Tukey multiple comparisons of means
95% family-wise confidence level

Fit: aov(formula = data_new$Estimate ~ data_new$variable.x,
weights = data_new$prescision)

$`data_new$variable.x`
      diff      lwr      upr      p adj
Var_Sub-KS_EDF -1.199 -2.783335  0.38533482 0.1741072
Wgt_Mean-KS_EDF -1.675 -3.259335 -0.09066518 0.0357301
Wgt_Mean-Var_Sub -0.476 -2.060335  1.10833482 0.7544765

>
```

A copy of this text output is available in the file: https://github.com/Olivex727/hvc-magnetic-honours-programs/blob/main/KS_confirmation/chisq.txt.