A low frequency study of linear polarization in radio galaxies

V. H. Mahatma1,2*, M. J. Hardcastle1, J. Harwood1, S. P. O’Sullivan3, G. Heald4, C. Horellou5 and D.J.B. Smith1

1Centre for Astrophysics Research, School of Physics, Astronomy & Mathematics, University of Hertfordshire, College Lane, Hatfield AL10 9AB, UK
2Thüringer Landessternwarte, Sternwarte 5, 07778 Tautenburg, Germany
3School of Physical Sciences and Centre for Astrophysics & Relativity, Dublin City University, Glasnevin, D09 W6Y4, Ireland.
4CSIRO Astronomy and Space Science, PO Box 1130, Bentley, WA 6102, Australia
5Chalmers University of Technology, Dept of Space, Earth and Environment, Onsala Space Observatory, SE-43992 Onsala, Sweden

Accepted 2020 December 22. Received 2020 December 11; in original form 2020 August 19

ABSTRACT
Radio galaxies are linearly polarized – an important property that allows us to infer the properties of the magnetic field of the source and its environment. However at low frequencies, Faraday rotation substantially depolarizes the emission, meaning that comparatively few polarized radio galaxies are known at low frequencies. Using the LOFAR Two Metre Sky Survey at 150 MHz and at 20 arcsec resolution, we select 342 radio galaxies brighter than 50 mJy and larger than 100 arcsec in angular size, of which 67 are polarized (18 per cent detection fraction). These are predominantly Fanaroff Riley type II (FR-II) sources. The detection fraction increases with total flux density, and exceeds 50 per cent for sources brighter than 1 Jy. We compare the sources in our sample detected by LOFAR to those also detected in NVSS at 1400 MHz, and find that our selection bias toward bright radio galaxies drives a tendency for sources depolarized between 1400 and 150 MHz to have flatter spectra over that frequency range than those that remain polarized at 150 MHz. By comparing observed rotation measures with an analytic model we find that we are preferentially sensitive to sources in low mass environments. We also infer that sources with one polarized hotspot are inclined by a small angle to the line of sight, while sources with hotspots in both lobes lie in the plane of the sky. We conclude that low frequency polarization in radio galaxies is related to a combination of environment, flux density and jet orientation.

Key words: polarization – galaxies: active – galaxies: jets – techniques: polarimetric – radiation mechanisms: non-thermal

1 INTRODUCTION

1.1 Polarized emission of radio-loud AGN
The synchrotron radiation by which we observe the jets and lobes of radio-loud active galactic nuclei (RLAGN) arises from relativistic electrons gyrating in magnetic fields. As a consequence of this process, the radiation is intrinsically linearly polarized. RLAGN, which can have a maximum degree of polarization of up to ~70 per cent (Pacholczyk 1970), are strong sources of polarized radiation which can be observed by radio telescopes (see reviews by Saikia & Salter 1988; Wielebinski 2012). Information on the polarized intensity from RLAGN is important to obtain for the following reasons:

- Since the position angle of the electric field vector of the radiation we observe is perpendicular to the projected magnetic field direction in the plane of the sky, polarization observations, if calibrated correctly, can directly give information on the structure of magnetic fields in the plane of the sky. This has led to studies of the magnetic field structure in the jets, lobes and hotspots of RLAGN, and their surrounding environment, on parsec scales (e.g. Gabuzda et al. 1992; Homan 2005) to kiloparsec scales (e.g. Laing 1980; Hardcastle et al. 1997; Laing et al. 2008; O’Sullivan & Gabuzda 2009; Guidetti et al. 2011 and see review by Bridle & Perley 1984). The hotspots of RLAGN, which are thought to contain compressed and ordered magnetic fields (Laing 1980; Hughes et al. 1989), are expected to be prime locations for polarized emission, so observations of polarization may enhance our understanding of particle acceleration processes.
- A lack of detectable polarization for high surface brightness objects gives evidence for substantial depolarization – a combination of factors such as the finite telescope beam, inhomogeneous magnetic field structures in the lobes or in their surrounding environment and Faraday rotation will reduce the observed polarization (as described below). Measurements of depolarization in the lobes can, in principle, trace their magnetoionic properties and their ther-
mal particle content, or that of their environment (Dreher et al. 1987; O’Sullivan et al. 2017, 2019; Knuttila et al. 2019).

Effects caused by Faraday rotation results in frequency-dependent depolarization: as linearly polarized emission travels through birefringent magnetised media (the intergalactic medium, for example), a difference in the phase velocity occurs for the right and left circular polarization constituents of the linear polarization. This manifests as a wavelength-dependent rotation of the polarization angle as

$$\chi = \chi_0 + RMA^2 \quad [\text{rad}] \quad (1)$$

where $\chi$ is the observed polarization angle (in radians), $\chi_0$ is the intrinsic polarization angle, $R$ is the rotation measure (in rad m$^{-2}$) and $A$ is the wavelength (in m). The $RM$ is related to the properties of the line-of-sight magnetised media by

$$RM = 0.81 \int_{\text{source}} n_e B || \cdot d\ell \quad [\text{rad m}^{-2}] \quad (2)$$

where $n_e$ is the electron density (in cm$^{-3}$), $B ||$ is the line of sight magnetic field strength (in $\mu$Gauss) and the integral is taken with respect to the path length $d\ell$ (in parsecs) through all intervening material between the source and the telescope. Differential Faraday rotation, and/or inhomogeneous magnetic field structures in the source, lead to different polarization angles across the telescope beam, which are then vector-averaged and lead to depolarization. Further depolarization can occur when these effects apply significantly within the observing bandwidth, since Faraday rotation is frequency-dependent.

Depolarization$^1$ of linearly polarized emission from RLAGN confirms the presence of magnetic fields in thermal plasma along lines of sight through their environments. Significant depolarization is generally attributed to environments local to the source that cause large $RM$s (e.g. in the interstellar or intracluster medium; Hardcastle 2003 and Carilli & Taylor 2002, respectively, or in the shocked gas surrounding radio lobes; Hardcastle et al. 2012), and hence $RM$s are useful in inferring properties of the environment which are otherwise difficult to obtain. For RLAGN environments well described by hot ($T \sim 10^7$ K) plasma that radiates in X-rays due to thermal bremsstrahlung, sensitive X-ray maps allow a measure of the gas density $n_e$ (e.g. Croston et al. 2008; Hicks et al. 2013; Mahatma et al. 2020), but may be expensive to obtain for large samples of radio galaxies. $RM$ maps can give valuable (but indirect) information on the surrounding environment of RLAGN, while giving constraints on the structure of magnetic fields.

In general, the interpretation of observed $RM$s is difficult, as they are in general a superposition of Faraday rotation from; the Earth ionosphere, the magnetized plasma in the Galaxy, the intergalactic medium, the intracluster/intragroup medium and within the source itself. Radio lobes in particular carry entrained thermal gas from their surroundings (Bicknell 1984), and source-intrinsic Faraday rotation can also be a significant contributor to observed $RM$s when the Galactic contributions have been subtracted. Precise $RM$ information can help to disentangle effects from different line of sight contributions if their respective $RM$s are found. In order to accurately quantify the $RM$ and polarization properties of RLAGN in general, large-sample statistics are needed.

1.2 Low frequency polarization

Low frequency (~ 100 MHz) linearly polarized source detections, particularly in a statistical study, are scarce. Due to the $J^2$ factor in Equation 1, Faraday rotation, and by extension depolarization, is much more important at low frequencies. RLAGN samples with polarization information exist in surveys at 1.4 GHz or greater (e.g. Taylor et al. 2007; Hales et al. 2014; Rudnick & Owen 2014), where depolarization is less significant in general, although this is also due to the fact that there are many more completed large-area radio surveys at GHz frequencies. Taylor et al. (2009) produced an $RM$ map of the sky using the NRAO VLA Sky Survey (NVSS; Condon et al. 1998), a 1400 MHz survey, detecting 37,543 polarized radio sources at declinations $\gtrsim -40^\circ$. However at lower frequencies the total flux density for any steep-spectrum RLAGN ($\alpha \approx -0.7$ where $S \propto \nu^\alpha$) is much higher$^2$, which may give adequate polarized signal to noise for low surface brightness regions that are undetected at higher frequencies. Moreover, and more crucially, since the $RM$ precision depends on the interval in $J^2$, low frequency instruments can out-perform centimetre-wave instruments by a few orders of magnitude in $RM$ precision. Larger source counts at these frequencies will test the robustness of the previously mentioned polarization studies, after understanding the detection statistics and possible selection biases of large samples at low frequencies. In sampling this new low-frequency parameter space, it is crucial to have radio telescopes with the ability to perform wide-area surveys of the sky combined with the required sensitivity to observe large samples of RLAGN – past surveys such as 3CRR (Laing et al. 1983) are severely biased towards the most luminous sources such as Fanaroff-Riley type-II objects (FR-II; Fanaroff & Riley 1974). Recently, Riseley et al. (2020) presented the POLarised GaLactic and Extragalactic All-Sky MWA Survey-the POLarised GLEAM Survey (POGS) in the frequency range 160-231 MHz at a resolution of 3 arcmin (at the highest), with 484 polarized RLAGN detected in the entire southern sky. However, instruments with the capability to perform sub-arcmin resolution surveys are more ideal in resolving the different components of RLAGN (core and lobes) and also mitigate the effects of beam depolarization in small angular size sources.

The LOw Frequency ARray (LOFAR; van Haarlem et al. 2013) is one such instrument, giving an angular resolution of 6 arcsec at 150 MHz. LOFAR is able to obtain a Faraday depth resolution (ability to resolve structures in Faraday depth space, where Faraday depth is the more generalised form of $RM$ in Equation 1) of $\sim 1$ rad m$^{-2}$ at 150 MHz, significantly better than that obtained by higher-frequency instruments (a factor of 200 better than the upcoming VLA Sky Survey, VLASS; Lacy et al. 2019). Additionally, LOFAR’s mixture of long and short baselines and its sensitivity to large extended structures (such as the lobes of nearby FR-I and FR-II radio galaxies) enables straightforward selection of RLAGN. The LOFAR Two-Metre Sky Survey (LoTSS; Shimwell et al. 2019), an on-going survey of the northern hemisphere enables large samples of RLAGN to be obtained for polarization studies. The first data release (DR1; Shimwell et al. 2019) covered the area of the Hobby-Eberly Telescope Dark Energy xperiment (HETDEX; Hill et al. 2008) Spring field: over 420 square degrees on the sky within $161^\circ < RA < 231^\circ$ and $45.5^\circ < DEC < 57^\circ$ observed at 6 arcsec resolution with a median sensitivity of $\sim 70 \mu$Jy beam$^{-1}$. With a large low-frequency sample of polarized RLAGN, in combination with measurements at higher frequencies, we may start to answer

$^2$ Down until the self-absorption regime which affects the radio continuum at frequencies lower than $\sim 100$ MHz (e.g. Scheuer & Williams 1968)
Polarization properties of radio galaxies

2 OBSERVATIONS

2.1 RLAGN sample

LoTSS DR2 (scheduled for public release in early 2021) will have a northern sky coverage of over 5700 deg$^2$, including the DR1 area which covered 424 deg$^2$. While DR1 does not contain polarization information, DR2 contains Stokes QU cubes (at 20 arcsec angular resolution) as data products, enabling polarization information to be extracted in this sky area. However, DR2 does not contain optical IDs for radio sources at the time of writing. For the purposes of our study we required a sample of radio galaxies with physical information such as radio luminosities. Hence, we use the DR2 polarization products to find polarized radio galaxies in the DR1 catalogue, which is publicly available and includes a value added catalogue with optical identifications (Shimwell et al. 2019; Duncan et al. 2019; Williams et al. 2019). This is particularly important since a radio galaxy catalogue from the DR1 sources has been made (Hardcastle et al. 2019), which we use to select sources for this study.

The DR1 radio galaxy selection details are given by Hardcastle et al. (2019), but we briefly describe them here. Sources were selected as having an optical ID from either Pan-STARRS (Chambers et al. 2016) or the Wide Infrared Survey Explorer (WISE; Wright et al. 2010), and either a spectroscopic redshift or a photometric redshift with a fractional error < 10%. From this sample of 71,955 sources, star-forming galaxies (SFG) were identified using the HETDEX region at an angular resolution of 4.3 arcmin, with their sources with the giant radio galaxy catalogue of Hardcastle et al. (2019), we selected sources with total flux density $S_{144} > 50$ mJy (as are all sources detected in polarization in the study by Stuardi et al. 2020) and with angular size $L > 100$ arcsec$^2$. These criteria resulted in a total of 382 sources in the DR1 area of 424 deg$^2$, from which we study the bulk polarization properties in the rest of the paper.

2.2 Polarized emission detection

To produce polarization and $RM$ maps of our sample, we utilised the $RM$ synthesis technique (Brentjens & de Bruyn 2005), using pyrmsynth$^4$, a Python script developed primarily for LOFAR Stokes Q and U cubes. The complex polarization $(P = Q + iU)$ can be written as

$$P(l^2) = \int_{-\infty}^{\infty} F(\phi) e^{2\pi i \phi l^2} d\phi$$

(Burn 1966), where $P(l^2)$ is the polarized intensity as a function of wavelength $(l)$ squared and $F(\phi)$ is the Faraday spectrum (polarized intensity as a function of the Faraday depth $\phi$). $RM$ synthesis transforms the cubes from frequency space into Faraday

\[\text{https://github.com/sabourke/pyrmsynth_lite}\]

---

3 \text{https://www.sdss.org/dr15/data_access/value-added-catalogs/value_id=mpa-jhu-stellar-masses}\n
4 Various other values for these criteria were tested, with the result that lower cut-off values resulted in a large number of undetected and unresolved sources in polarization for the purposes of this study.

5 \text{https://github.com/sabourke/pyrmsynth_lite}
depth space by inverting Equation 3 so that (an approximated reconstruction of) the Faraday spectrum is calculated. The reconstructed Faraday spectrum ($F(\phi)$) is then used to measure the peak polarized signal for each pixel in the image. The value of $\phi$ at which a peak in the Faraday spectrum is found is taken as the $RM$ of each pixel$^6$. This technique has already been extensively applied to LOFAR data (Van Eck et al. 2018, 2019; O’Sullivan et al. 2018, 2019; Stuardi et al. 2020; O’Sullivan et al. 2020). We extracted the (unCLEANed) Stokes Q and U cubes of all sources in our sample, which were spatially masked with a $3\sigma$ cut-off based on the Stokes I image of the source at 20 arcsec resolution. We inputted the QU cubes for each source into Pymsynth, using the rmclean tool (Heald et al. 2009) to deconvolve the Faraday spectrum using a maximum of 1000 iterations, fitting a Gaussian to the peak of the reconstructed CLEANed Faraday spectrum, resulting in linearly polarized intensity and $RM$ maps. Note that this procedure implicitly assumes only one peak in the Faraday spectrum of each source, but we verified that multiple and equally strong peaks were not present (except in the case of leakage – see below) by inspecting the spectra. We limited the Faraday depth range to $-150 < \phi < 150$ (rad m$^{-2}$) to search for polarized emission, with increments of $\delta \phi = 0.3$ rad m$^{-2}$ for LOFAR, with initial analysis we found that we do not detect peaks in the Faraday spectra outside the range stated above. With this spectral setup of RM synthesis we are sensitive to scales $\lesssim 1$ rad m$^{-2}$ in Faraday space. As no corrections were made for leakage signal, which can dominate near Faraday depths close to zero, we further exclude the range $-3 < \phi < 1.5$ rad m$^{-2}$ from the fitting of the peak in the Faraday spectrum. Hence, we are only Faraday-depth-complete outside this range. The linear polarization and $RM$ maps of six sources in our sample are shown in Figure 1 and Figure 2. It should be noted that the typical Galactic RM in the HETDEX region is $0 \lesssim RM_{\text{Galactic}}$ (rad m$^{-2}$) $\lesssim +23$ (Oppermann et al. 2015).

To determine which sources have detections of polarized emission, given the relatively low resolution of the maps and the expectancy of low $S/N$ polarization, we use a simple island-finding method by masking non-detected pixels$^7$. First, we remove background noise pixels by masking pixels which have surface brightnesses less than the mean pixel value in the image plus $3\sigma_p$, where $\sigma_p$ is the standard deviation of the pixel brightnesses in the image. For some sources that were clearly not polarized (sporadic regions of high pixel intensity, mostly off-source) but still had unmasked regions, particularly those with large angular size, we manually disallowed a detection. Detections were also manually disallowed of high pixel intensity, mostly off-source) but still had unmasked regions of nearby polarized sources in the map unrelated to the source (i.e. background quasars). For some sources generally with small angular sizes and hence a low number of background pixels present in the maps (as the cubes were constructed near the source to only contain on-source pixels), mean pixel surface brightnesses and $\sigma_p$ were generally overestimated as the $3\sigma_p$ threshold prevented detections of clearly polarized hotspots. We therefore implemented a procedure to iteratively mask $>5\sigma_p$ regions in the polarized intensity maps and re-calculate $\sigma_p$, until the fractional difference between $\sigma_p$ in the current and last iteration became ($\sigma_p_{\text{last}} - \sigma_p_{\text{current}})/\sigma_p_{\text{current}} < 1 \times 10^{-6}$. For some source images still containing too few background pixels, where the image is dominated by bright emission, we reduce the masking criterion to pixels $>3\sigma_p$.

We then label as detections of polarized emission where groups of unmasked pixels are in a $2 \times 2$ configuration or larger, so as to prevent single pixel detections which we regard as insufficient as all of our sources are resolved at 20 arcsec in Stokes I. We measured polarized flux densities as the sum of the detected pixel surface brightnesses divided by the beam area. The uncertainty on the measured polarized flux densities are quoted as $3\sigma_{\text{QU}}$, where $\sigma_{\text{QU}}$ is given as the mean of the detected pixels in the linearly polarized rms map output from $RM$ synthesis$^8$. For the $RM$ of each source, we take a weighted mean pixel value from the $RM$ map as

$$\langle RM \rangle = \frac{\sum_i w_i RM_i}{\sum_i w_i} \tag{4}$$

where $w_i$ is the normalised pixel brightness of pixel $i$ of all detected pixels in the polarized intensity map. The $RM$ error is given by

$$\sigma_{\text{RM}} = \frac{\text{RMSF FWHM}}{2 \times S/N} \tag{5}$$

(Brentjens & de Bruyn 2005), where the RMSF FWHM (Full Width Half Maximum) is $1.16$ rad m$^{-2}$ for LoTSS pointings, and the signal to noise ratio $S/N$ is the peak pixel brightness over the rms value in that pixel. An additional, more dominant error arises from the systematic error of the ionospheric $RM$ correction included in $RM$ synthesis, and results in uncertainties of $0.1-0.3$ rad m$^{-2}$ (Sotomayor-Beltran et al. 2013). We give conservative error estimates by adding in quadrature the error from Equation 5 and a maximum systematic error of $0.3$ rad m$^{-2}$ for each source. We also correct our polarized intensities for Ricean bias, which can be significant at low signal to noise, using Equation 5 of George et al. (2012). No corrections have been made for the dependence of the derived Faraday spectrum on spectral index, but it does not affect the peak of the Faraday spectrum and hence the corrections are minimal (Brentjens & de Bruyn 2005).

We then identified sources where we find evidence of leakage signal being manifested as detections – the Faraday depth range we excluded from $RM$ synthesis ($-3 < \phi < 1.5$ rad m$^{-2}$) is not precisely centred on zero due to the ionospheric $RM$ correction(s) shifting the leakage signals. It is possible for this exclusion range to not capture all sources with leakage signal and some of our detected sources which have $RM$s below $-3$ rad m$^{-2}$ and above $+1.5$ rad m$^{-2}$ may show strong leakage, particularly at the locations of the sidelobes of the leakage signal. Since leakage signals will generally have low fractional polarization ($\Pi \lesssim 1\%$), we removed sources in our sample with $\Pi \lesssim 1\%$ that were detected at $-3.5 < \phi < -3$ rad m$^{-2}$ and $1.5 < \phi < 2$ rad m$^{-2}$ in their Faraday spectra. This resulted in the removal of seven sources. Faraday spectra of three detected sources in our sample that were not excluded are shown in Figure A1.

Applying our methods we finally obtain a reliable polarization detection in 67 out of 382 sources – a detection fraction of 18 per cent at 150 MHz. We regard the non-detected sources as depolarized. This represents a polarized radio galaxy surface density within the HETDEX field, for sources brighter than 50 mJy and larger than 100 arcsec in angular size, of $0.16\pm0.02\text{deg}^{-2}$ (errors quoted here

---

$^6$ Note that this value for the $RM$ applies in the case of a delta function for the Faraday spectrum.

$^7$ Each map has a pixel size of 4.5 arcsec.

$^8$ The rms is given by the average spread of the ‘wings’ of the Faraday spectrum over Stokes Q and U, limited by our Faraday depth range of $-150 \lesssim \phi$ (rad m$^{-2}$) $\lesssim 150$. 
Figure 1. 150 MHz polarized FR-Is from our RLAGN sample. Coloured pixels represent 20 arcsec polarization detections: left panels are the polarized intensity and the right panels are the $RMs$, while black contours display the 6 arcsec Stokes I emission (beam sizes shown as magenta and black circles in the lower left of each image, respectively). Red circles show the position of the catalogued host galaxy in LoTSS DR1.
Figure 2. 150 MHz polarized FR-IIs from our RLAGN sample. See Figure 1 for image descriptions.
represent Poisson statistics). As a comparison with recent studies from the LoTSS data in the same area, Van Eck et al. (2018) found 92 point sources at a surface density of 0.16±0.02 deg$^{-2}$ at 4.3 arcmin resolution, while Mulcahy et al. (2014) and Neld et al. (2018) found 6 securely-detected polarized sources in a single LOFAR pointing at the same resolution as ours, finding a surface density of 0.30±0.12 deg$^{-2}$, demonstrating consistency between LOFAR studies and confirming that RLAGN are the predominant source of polarization at 150 MHz. Subsets of our detected sources are presented in Figure 1 and Figure 2.

### 2.3 Caveats affecting our polarized sample

The most important limitation to our analysis is the use of un-CLEANED Stokes QU cubes that are used to produce polarized intensity (dirty) maps. We cannot reliably CLEAN the cubes prior to $RM$ synthesis due to the low signal to noise in each channel of the cubes. The effect of this is higher noise and artefacts in the resulting polarized intensity maps, particularly around bright point sources, than in the case of CLEANed maps. Since the aims of our study do not rely on accurate astrometry and analysis of spatial structure, the low image fidelity due to the use of dirty QU cubes does not affect our analysis or results.

We also note the issue of using $RM$ synthesis and the associated rms map in polarized intensity to determine detection rates: imperfect imaging and calibration (as is the case here) can lead to non-Gaussian tails in the Stokes Q and U Faraday spectra, whence the rms is calculated. While this will overestimate errors, it may also cause high false detection rates in low signal to noise pixels (George et al. 2012) where relatively low detection thresholds are used ($3\sigma_{QU}$, as used here). While George et al. (2012) propose $8\sigma_{QU}$ to serve as a detection threshold, we note that our detection method involved visual inspection, and in the case of polarized emission not associated with a core, lobe or hotspot (i.e. structures of high intensity in the Stokes I image), sources were discarded as false detections. A $5\sigma_{QU}$ threshold was tested, giving the result that low signal to noise sources that had clearly polarized emission (e.g. in hotspots) were undetected with our method.

### 3 ANALYSIS

In this section we present a statistical analysis of the polarization properties of our RLAGN sample. Unless otherwise stated, to distinguish the characteristics of two distributions we quote the $p$-value from a Wilcoxon-Mann-Whitney test (Mann & Whitney 1947), using a 95 per cent confidence level (i.e. a $p$-value < 0.05 means we can reject the null hypothesis that the two samples have identical median values).

#### 3.1 Observational properties

In the top panel of Figure 3 we plot the distributions in 150 MHz flux density and angular size of our detected and non-detected sources. We see that, statistically, our polarized sources (hatched beige) are significantly brighter compared to those that are depolarized (grey), as expected. Further, the detection fraction is greater than 50 per cent for sources brighter than 1 Jy. On the other hand, there are statistically similar medians in angular size between detected and non-detected sources, meaning that the detectability of polarization amongst RLAGN in our sample is driven primarily by flux density rather than, or in addition to, angular size. Similar statistics are seen with physical properties as shown in the bottom panel of Figure 3, where the polarized sources are more luminous but similar in physical size. This is consistent with the idea that brighter and more luminous RLAGN have a preference for detectable polarization.

In the left panel of Figure 4 we plot the 6 arcsec total flux density against the 20 arcsec polarized flux density for our polarized sources. We do not see any clear correlation, meaning that although the polarized detection fraction increases with flux density (as seen in Figure 3), the amount of polarized emission is not entirely driven by total flux density. Rather, a range of polarized emission is detected in RLAGN for a large range in total flux density at 150 MHz. This is consistent with a model in which the level of polarized emission seen in RLAGN is strongly related to the characteristics of the associated Faraday screen and the attributed depolarization, some components of which are unrelated to the source (e.g. the foreground IGM). We also plot the fractional polarization $\Pi$ (ratio of polarized emission to total emission in the polarization-detected pixels) against the total flux density measured in the 20 arcsec Stokes I image, in the right panel of Figure 4 (the red dashed line indicates the approximate sensitivity to polarization by taking an average of $3\sigma_{QU}$ from all sources). As expected, since the latter observable is the denominator of the former, the fractional polarization decreases with increasing total flux density. However there is a large scatter, particularly for $>1$ Jy sources, which is likely driven by a combination of different jet properties and environmental Faraday depolarization in the line of sight at 150 MHz.

It is important to test whether there are correlations within the detected sources due to morphology (i.e. different jet or lobe properties). It is well known that FR-I sources are expected to have higher depolarization than FR-II sources for the following reasons: given the same external environment, the bright cores of FR-I sources would typically be associated with higher Faraday depths due to the increasing radial profile towards the centre of the thermal plasma in which they are embedded. On the other hand, FR-II sources are typically polarized at their hotspots (e.g. O’Sullivan et al. 2018), which are at larger projected distances from the center of their environments. In addition, FR-I sources are generally found in richer environments than FR-Hs, leading to higher Faraday depths in their line of sight. FR-II sources are also thought to significantly entrain dense material from their surrounding medium (e.g. Laing & Bridle 2002; Croston et al. 2008; Croston & Hardcastle 2014), which in general would lead to more internal depolarization over that in FR-II sources. Brighter polarization might also be expected in FR-II sources as they tend to be brighter in total flux density in general than FR-I sources – in our sample polarized FR-Is and FR-Hs have median total flux densities of 513 mJy and 626 mJy, respectively, the difference being statistically significant ($p$-value < 0.05). While FR-II hotspots are the predominant source of polarization at 150 MHz (O’Sullivan et al. 2018), it is important to quantify the detection fractions between FR-Is and FR-Hs, as well as to compare their fractional polarization.

We use the code of Mingo et al. (2019), which categorizes sources into FR-I and FR-II based on whether the peak radio emission is located close to our away from the centre of the source, respectively, to morphologically categorize the polarized sources in our sample. While visual inspection may be used to classify our sources, an automated classification based on the definition of the

---

We note that the flux density cut in selecting sources in our sample introduces a selection bias, and our results do not necessarily apply for sources with $S_{150} < 50$ mJy.
FR dichotomy (as used by Mingo et al. 2019) represents a systematic morphological stratification of our sample without cognitive bias. The 382 objects in our sample were separated based on the code into three categories; FR-I, FR-II and indeterminate. The latter category represents the case where the code cannot clearly determine the morphology, and this is usually the case for more compact objects or objects with non-symmetric lobes where there may be FR-I-like lobes on one side of the jet and FR-II-like lobes on the other, which are likely due to projection effects in many cases (Harwood et al. 2020). Due to the ability of LOFAR to observe both compact and extended structures of RLAGN, we expected the code to classify a large number of sources which can be visually identified as either FR-I or FR-II as indeterminate. Some of the authors (VHM, MJH and JH) visually checked the 6 arcsec total intensity maps of the indeterminate sources in our sample, and re-classified each source as either FR-I or FR-II where appropriate. Around ~ 10 per cent of the indeterminate sample were sources that could be clearly identified as an FR-I or an FR-II, and were moved to those categories. We also checked the sources in the original FR-I and FR-II samples, to conservatively check for obvious contaminants from either class. Only a small number (< 5 per cent of sources) from each category were moved to the other category. In all scenarios, sources were declassified as indeterminate on a conservative basis in order to form robust FR-I and FR-II samples. Our morphological analysis uses the FR-I and FR-II samples only.

Table 1 lists the detection fractions of our RLAGN sample with morphology. We see that our RLAGN sample of 382 sources is categorized as 122 FR-I sources, 146 FR-II sources and 114 indeterminate-morphology sources. Comparing the polarization statistics, we see that FR-II sources have more than twice the detection fraction of FR-I (and that of indeterminate sources). These quantities robustly confirm findings by OS18, that FR-II radio galaxies are much more likely to be detected in polarization than FR-I sources.
Figure 4. 150 MHz linearly polarized flux density (left) and fractional polarization (right) against total flux density for our detected sources. Note that for the right panel we only plot the total flux density for the regions of the source that are polarized. Uncertainties are shown as 3σ error bars, but are not visible due to the logarithmic scale. Dashed lines give an indicative estimate of our sensitivity to polarized emission, plotted as 3σ_{mean,QU}, where σ_{mean,QU} is the average rms error of polarized fluxes in our sample.

Figure 5. Same as Figure 4, for sources classed as FR-I (beige) and FR-II (blue) according to the criteria of Mingo et al. (2019). Note that indeterminate-morphology sources are not plotted.

| Sample      | Counts | Sample detection fraction (%) |
|-------------|--------|-------------------------------|
| RLAGN       | 382    | 17.5                          |
| FR-I        | 122    | 3.4                           |
| FR-II       | 146    | 10.2                          |
| Indeterminate | 114  | 3.9                           |

Table 1. Polarization detection fractions for our RLAGN sample, and the FR-I, FR-II and indeterminate subsets based on morphology. Counts refer to the total number of sources in that sample.

We reproduce Figure 4 in Figure 5, now labelling the sources as FR-I (beige) and FR-II (blue). We see that in general, at a given 150 MHz total flux density, FR-II sources tend to be brighter in polarization than FR-I sources, although with large scatter. In Figure 6 we plot the distribution in fractional polarization Π_{150} for FR-I and FR-II sources, showing a statistically higher median Π_{150} for FR-IIs, where FR-IIs solely dominate at Π_{150} > 4 per cent. This is likely due to the presence of bright hotspots in FR-IIs, that are clearly dominating our statistics, compared to FR-Is mostly with polarized cores.

3.2 Host galaxy properties

The host galaxy properties of RLAGN populations give valuable information on the drivers of AGN activity. It is important to deter-
mine if there are differences in detection fractions as a function of host galaxy type. In particular, we test the hypothesis that polarized and depolarized RLAGN at 150 MHz can be driven by the same type of host galaxy.

In Figure 7 we plot the distributions of host galaxy optical $K_s$ and $r$-band absolute magnitudes, available in the LoTSS DR1 catalogue, for our polarized (yellow) and depolarized (grey) sources. We see that the distributions in both optical bands are similar – the $p$ values (quoted above both figures) from a two-sample Kolmogorov-Smirnov (KS) test are both $> 0.05$, indicating that we cannot reject the hypothesis that both polarized and depolarized subsets have similar distributions, at a confidence level of 95 per cent. The optical and near-IR intrinsic brightness of the host galaxies that drive radio jets in our sample are not in general associated with a detection of polarized emission.

In terms of polarized radio sources, Banfield et al. (2014) find that quasar-type galaxies typically host sources with lower fractional polarization than quiescent galaxies. O’Sullivan et al. (2015) obtain similar results and find differences in the fractional polarization between High Excitation Radio Galaxies (HERGs) and Low Excitation Radio Galaxies (LERGs), finding that LERGs can achieve higher intrinsic degrees of polarization at GHz frequencies, and O’Sullivan et al. (2017) relate this to LERGs having more intrinsically ordered magnetic fields in the radio plasma. We instead test the low frequency detectability in polarization by comparing the detection rates of HERGs and LERGs using the WISE colour-colour plot, given by WISE mid-infrared apparent magnitudes at 3.4, 4.6, 12, and 22 $\mu$m (W1, W2, W3, W4 bands). While this does not give a direct classification of the HERG and LERG status of a particular galaxy, the majority of LERGs tend to have lower values of $W1$-$W2$ and $W2$-$W3$, while HERGs tend to have higher values of colour, with higher levels of dust-obscuration (which may arise due to the presence of an optically thick torus surrounding the accretion system). The position of a particular object in the WISE colour-colour plot can therefore give information on the nuclear properties of a galaxy.

In the top panel of Figure 8 we present the WISE colour-colour plot for our RLAGN sample. We find that the hosts of the polarized sources in our sample have significantly higher values of $W1$-$W2$ and $W2$-$W3$ than depolarized sources, shown by the larger fraction of polarized objects in the upper-right hand side of the plot (statistics were tested for distributions in $W1$-$W2$ and $W2$-$W3$ between polarized and depolarized sources, with $p$–values < 0.05). We have over-plotted the ‘AGN wedge’, defined by Mateos et al. (2012), who select luminous AGN selected in X-rays. This region is typically occupied by obscured AGN and quasars (i.e. HERGs), which, at low redshift, tend to be associated with FR-II radio galaxies (Zirbel 1997). In the bottom panel of Figure 8 we separate our sample into FR-I and FR-II objects, and it is clearly seen that the majority of the sources with higher values of WISE colours are FR-II sources (while there are still a comparable amount of FR-II sources where FR-Is are situated). Further, in Figure 9, we show the fractional polarization $\Pi_{150}$ as a function of the $W1$-$W2$ colour. We see that there is a tendency for FR-Is to have a higher $W1$-$W2$ colour than FR-Is, for a given $\Pi_{150}$, though with large overlap. We therefore have insufficient evidence to suggest

Figure 6. Distribution of fractional polarization between FR-I (hashed orange) and FR-II (blue) radio galaxies. Dashed lines represent median values.

Figure 7. Distributions in rest frame absolute magnitudes in the optical $K_s$ band (top) and $r$-band (bottom) of the host galaxies of our polarization detected sources (hashed beige) and our non-detected sources (grey).
Polarization properties of radio galaxies

that quasars with higher values of WISE colour tend to drive radio lobes with higher polarization. Rather, FR-II sources are more likely to be detected due to their more powerful jets than FR-IIs, and tend to be hosted by dust-obscured galaxies with a torus, also associated with radiatively efficient accretion in HERGs (Laing et al. 1994; Evans et al. 2006; van der Wolk et al. 2010; Gürkan et al. 2014). While this may present an apparent contradiction with the results of Banfield et al. (2014), O’Sullivan et al. (2015) and O’Sullivan et al. (2017), in that we have a higher detection rate of polarization for quasar/HERG-type radio galaxies (mostly FR-IIs as seen in Figure 8), we emphasise their results are based on the modelled intrinsic fractional polarization at high frequencies, while ours is based on observed polarization at low frequencies. Further, and on a related note, it is very likely that LERG FR-IIs which may have high intrinsic degrees of polarization are generally depolarized in our sample due to high Faraday dispersions in the centres of their hot gas environments, which are not likely to be detected with LOFAR at 150 MHz (see Section 3.3). Hence, we suggest that the 150 MHz polarization in radio galaxies is not driven by a particular type of optical brightness or nuclear emission in the host galaxy relative to the population of radio galaxies, but an association exists between polarization and WISE colour due to the association between the FR-class and WISE colour.

3.3 Frequency dependence

Any frequency dependence of the measured fractional polarization can give evidence for depolarization for any given source. To investigate this, we compared the polarization properties of our sample at 150 MHz with the polarization measured for the same objects with NVSS at 1.4 GHz (Taylor et al. 2009), as was done by Van Eck et al. (2018) for their sample at 4.3 arcmin resolution. Taylor et al. (2009) present a 1.4 GHz sample of polarized sources in the sky at declination north of δ = −40°, which also covers the LoTSS area. The NVSS completeness limit of 2.5 mJy gives an ideal comparison survey, though the restoring beam at 45 arcsec is more than a factor of two larger than that of our sample. Van Eck et al. (2018) find that their LOFAR-detected sample contains sources with most of their polarized emission in broad Faraday-thick components (as their sources have higher fractional polarization with NVSS than with LOFAR, with LOFAR only being sensitive to structures ≤1 rad m$^{-2}$ in Faraday depth). We apply some similar comparisons with our sample to investigate the line-of-sight environments of RL AGN.

We identify our sources with the Taylor et al. (2009) sample by using a positional cross-match criterion with a separation limit of 135 arcsec (three NVSS beams). Lower separation limits of one or two NVSS beams resulted in fewer correct cross-matches since our sources are centered on the optical ID, whereas the NVSS polarized sources are centred on the polarized emission, such as hotspots, which can be more than two NVSS beams from the optical ID. To ensure our cross-matching criteria selected the correct NVSS source, we convolved our LOFAR images with a 45 arcsec beam, and compared each image to the cross-matched NVSS source in Stokes I and in polarization, by downloading Stokes IQU cubes using the

![Figure 9. Fractional polarization against host galaxy W1-W2 colour for FR-I (yellow) and FR-II (blue) radio galaxies.](image-url)
NRAO postage-stamp server. We also ensured that sources that did not have an NVSS counterpart were not missed by our cross-matching criteria (i.e. polarized hotspots in NVSS that are more than three NVSS beams from the optical ID). We found that one source (ILT105715.33+484108.6) was missed due to its large angular size, and included it as a correct NVSS counterpart.

Using this method we find 58 NVSS counterparts to our parent RLAGN sample of 382 sources. Out of these sources, 24/58 are detected in polarization in our LOFAR sample, giving approximately a 40% LOFAR detection rate. However, we note that out of the 67 polarized sources in our sample, there is an absence of polarized NVSS counterparts in 44/67. This is surprising as we would naively expect all 67 LOFAR detections to also be detected with NVSS since the fractional polarization should be higher at higher frequencies and since our sources, being larger than 100 arcsec, are all resolved with NVSS. Beam depolarization could be more prominent in NVSS due to its factor of two larger beam – comparing the polarized intensity NVSS and LOFAR images at 45 arcsec, we find only one source with significant depolarization in the NVSS images, which we regard as insufficient to explain the lack of NVSS detections. We instead attribute the lack of polarized NVSS counterparts to the fact that LOFAR is more sensitive to steep-spectrum sources relative to NVSS, so that many steep-spectrum sources that suffer little depolarization (so that they are detected at 150 MHz) are undetected in NVSS. As a check, we calculated the expected Faraday dispersion $\sigma_{RM}$ assuming external Faraday rotation using our LOFAR and NVSS polarized counterparts using

$$\Pi_{150}/\Pi_{1400} = \exp(-2\sigma_{RM}^2 \alpha_{150}^4 + 2\sigma_{RM}^2 \alpha_{1400}^4)$$

(Sokoloff et al. 1998). The average $\sigma_{RM}$ in our sample is 0.24 rad m$^{-2}$. For comparison, the median $\sigma_{RM}$ for 20 double radio galaxies measured by O’ Sullivan et al. (2017) is 12.5 rad m$^{-2}$, implying that the sources detected by LOFAR have very little depolarization. Moreover, Equation 3 implies that the Faraday dispersion function must be narrow for less depolarization at long wavelengths.

In order to confirm if the lack of polarized NVSS counterparts is due to their steep spectra, we further cross-matched the sources in our sample with no NVSS polarized counterparts (using the same criteria as above) with the NVSS source catalogue (Condon et al. 1998). In Figure 10 we display the distributions in 150 MHz flux density, 1400 MHz flux density and spectral index between those two frequencies (corrected for the difference in beam sizes). The top panel indicates that the polarized NVSS sources which are also polarized with LOFAR (blue) are brighter at 150 MHz than those that are not polarized with LOFAR (beige hatched, note the $p$–value above the figure), as expected. However in the middle panel, there is no statistical evidence for different average flux densities at 1400 MHz, implying LOFAR is more sensitive to steep-spectrum sources at low frequencies. In both panels, the black dashed histogram denotes those sources polarized with LOFAR with no polarized NVSS counterpart, showing that they are fainter than polarized NVSS sources both at 150 MHz and 1400 MHz. In the bottom panel, affirming our earlier inferences, we see that while the average spectral indices of the NVSS polarized sources with LOFAR polarization are steeper than those without LOFAR polarization as expected, the sources polarized with LOFAR but not with NVSS are even steeper (WMU tests performed between all three distributions have $p$–values $< 0.05$). This confirms that the lack of NVSS polarized counterparts is due to their steep spectral indices.

---

10 https://www.cv.nrao.edu/nvss/postage.shtml

---

Figure 10. Distributions in total flux density at 150 MHz (top), 1400 MHz (middle) and spectral index (bottom) for our NVSS-LOFAR cross-matched sources. NVSS polarized sources that are not polarized at 150 MHz are shown in hatched beige and those that are polarized at 150 MHz are shown in blue. Sources that are polarized at 150 MHz but not at 1400 MHz are shown in black dashed lines. Note the $p$–values refer to the comparison between blue and hatched beige distributions only.
so that their flux densities have decreased beyond detection at 1400 MHz, while still being polarized at 150 MHz. Note that the median error in the spectral indices in our sample is $\Delta \alpha_{1400}^{150} = 0.11$, taking into account $3\sigma$ errors on the total flux density measurements.

In Figure 11 we plot the fractional polarization at 150 MHz against that at 1400 MHz for our cross-matched sources. We see that all sources, except one, have a lower fractional polarization at 150 MHz, showing depolarization at low frequencies. For the source not depolarized at 150 MHz relative to 1400 MHz (at $\Pi_{150} = -0.025$), we attribute its higher $\Pi_{150}$ to beam depolarization in NVSS – upon visual inspection the LOFAR data show resolved components in the lobes which are shown as a single component in the NVSS image.

### 3.4 Rotation measure analysis

In this section we analyse the $RM$s of our sample. The $RM$ arises from the superposition of the line of sight from contributions from the Galaxy, the intergalactic medium, the intragroup/intracluster medium and the source itself. We are more interested in the source and local environment properties, and so we subtract the Galactic $RM$ measured at the location of each source in our sample. To do this we positionally cross match the sources in our sample with the Galactic Faraday sky map made by Oppermann et al. (2015). The pixel size in this map is around 30 arcmin, and hence all our sources are spatially coincident within single pixel regions in this map. We subtract the pixel $RM$ value from the $RM$ value we measure at 150 MHz, as well as from the $RM$s at 1400 MHz from the Taylor et al. (2009) catalogue, giving Galaxy-subtracted values. For each source we propagate the LOFAR and NVSS $RM$ errors (Section 2.2) with the $RM$ errors catalogued by Oppermann et al. (2015).

In Figure 12 we plot the distribution in observed $RM$ at 150 MHz and the Galaxy-subtracted $RM$ for our sample. The average $RM$ in our sample is positive, with a mean\(^\dagger\) observed value of $+12.6 \pm 0.15$ rad m\(^{-2}\), and the mean Galaxy subtracted value is $+1.0 \pm 0.14$ rad m\(^{-2}\) (with an rms of 7.56 rad m\(^{-2}\)). This shows that

\(^\dagger\) Uncertainties quoted are the standard errors of the mean.

Figure 11. LOFAR fractional polarization at 150 MHz against NVSS fractional polarization at 1400 MHz for 28 sources in our sample with both LOFAR and NVSS detections. The red line shows the line of equality. NVSS error bars are taken from the catalogue of Taylor et al. (2009).

Figure 12. Rotation measure distribution of our sample. The yellow histogram shows our observed rotation measures, and the blue histogram shows the same data after being corrected for the Galactic contribution. Dashed lines represent medians.

the bulk of our sample do not have large magnitudes of $RM$ from non-Galactic Faraday screens. It is likely then that the dominant contributor to our $RM$s comes from multiple Faraday screens with multiple magnetic field reversals, which acts to reduce the $RM$ magnitude, or that our sources are preferentially located in low density environments which lead to lower depolarization, or a combination of both factors.

We also compare the distributions in Galaxy-subtracted $RM$ at 150 MHz and 1400 MHz in Figure 13. In the top panel we show the 1400 MHz $RM$s of those sources that are polarized (blue) and depolarized (hatched beige) with LOFAR. We see that there are similar distributions and medians in 1400 MHz subtracted $RM$s between LOFAR polarized and depolarized sources, with a similar range. This suggests, on a statistical level, that the depolarization of RLAGN at low frequencies is not solely driven by the magnitude of Faraday rotation, as expected. Rather, it is the large spatial and/or line of sight dispersion of Faraday rotation that causes depolarization at low frequencies, and since small-scale variations (i.e. internal Faraday rotation) will cause depolarization, it is more likely that the sources we detect with little $RM$ variation across the detected regions have significant contributions from the foreground medium local to the source (i.e an ICM). Moreover, as shown by the bottom panel of Figure 13, the Galaxy subtracted $RM$s have similar peaks between LOFAR and NVSS data for those sources polarized with LOFAR, while there is a much larger spread in NVSS $RM$ as expected due to the lower $RM$ resolution. This also shows a general consistency between Faraday screens in the foreground responsible for depolarization in our sample. These results imply that the absence of polarized emission with LOFAR for these sources is mostly driven by the combination of their total flux density at 150 MHz and their individual Faraday-rotating media. In Figure 14, where we plot the $RM$s at 1400 MHz against those at 150 MHz for the cross-matched LOFAR polarized sources, the red diagonal line of equality shows that there is a general lack of a correlation, even with the large errors (calculated using propagation of errors during subtraction). The lack of a clear correlation implies that we may be tracing components of different Faraday screens, rather than dif-
different components of the same Faraday screen. Since we infer that our RM's are sensitive to external Faraday screens such as the ICM, we may predict RM's toward realistic ICM environments around RLAGN and compare them with our observations.

3.5 Environment modelling

X-ray data have long been used to determine environmental properties of radio galaxies (e.g. Croston et al. 2008, 2011; Hardcastle et al. 2016). However, these are difficult to obtain for large samples, particularly in probing the high-redshift regime (\(z > 1\)). RM maps can provide an independent method of determining the line-of-sight environmental properties that may be used to infer the environments of large samples of radio galaxies. While it is difficult to infer the environment from the RM without information on the electron density, magnetic field strength or the magnetic field reversal scale (the typical physical length between magnetic field direction reversals in the line of sight), we may instead use a model to predict RM's based on realistic radio galaxy environments. In particular, assuming purely external Faraday rotation due to a local environment, with plausible assumptions about thermal gas distributions, magnetic field strengths, reversal scales and geometry, we test whether it is possible to reproduce the RM distribution that we observe with LOFAR. If such a distribution is obtained (by way of a two-sample KS test between modelled and observed RM distributions), we may then compare the physical environmental properties of models that LOFAR is sensitive to and those that LOFAR is not sensitive to. We also compare the fraction of models that have RM's in our observed range against our polarized detection fraction, and discuss any discrepancies between the two.

3.5.1 RM prediction model

We create an analytic model which predicts RM's using Equation 2, which relies on the electron density \(n_e\) and the magnetic field \(\mathbf{B}\) through the line of sight toward each polarized source in our sample. A calculation of these properties requires knowledge of the physical environment of each source. Recent studies have shown that the radio lobe properties can be reliable indicators of the ICM pressure at a fixed distance (Ineson et al. 2017; Croston et al. 2017), but such associations based on large samples of the overall RLAGN population have large uncertainties when predicting the properties of any given source. Instead, to determine the physical information needed to predict RM's, we draw cluster/group masses from a distribution appropriate for radio galaxy environments. We also include prescriptions for the magnetic field reversal scale (which affects whether the incremental RM is positive or negative) and the orientation of the radio source (which affects the line of sight path length), both of which are unknown, but we use appropriate probability distributions for these input parameters and sample them as a Monte Carlo simulation. Our RM prediction model is as follows:

- We generate distributions of group/cluster masses using the

![Figure 13. RM's for our LOFAR-NVSS cross-matched sample. Top panel shows the 1400 MHz RM's between those polarized with LOFAR (blue) and those depolarized with LOFAR (hatched beige). Bottom panel shows the LOFAR (hatched beige) and NVSS (blue) RM's for the sources polarized with LOFAR. Medians are shown as dashed lines.](image1)

![Figure 14. 150 MHz RM against 1400 MHz RM for our cross-matched sample, after Galaxy RM subtraction. Calculation of error bars are stated in Section 2.2. The red line is the line of equality, and not a regression line fit to the data.](image2)
mass function of Girardi & Giuricin (2000), who show a good agreement between a single Schechter function at \( z = 0 \) and the local mass function for groups and clusters. We generate a Schechter function (with \( \alpha = -1.5 \)) for group/cluster masses in the range \( 10^{13} - 10^{15} M_{\odot} \), which gives a bias towards masses of groups/clusters which tend to host RLAGN based on optical and X-ray studies (Hill & Lilly 1991; Hardcastle & Worrall 1999; Best 2004; Ineson et al. 2015). For each polarized source in our sample, we draw a sample of 1000 values from this function.

- We assume an equivalence between the group/cluster mass and \( M_{500} \), the mass enclosed in a sphere within which the mean density is 500 times the critical density at \( z = 0 \). For each \( M_{500} \) for each source, we determine a radial pressure profile \( p(r) \) of the environment, parametrised by \( M_{500} \) using the universal pressure profile of Arnaud et al. (2010). The physical size of the pressure profile is determined by calculating the distance from the polarized source at its redshift to \( z = 0 \).
- We determine a density profile \( n_{e}(r) \) for each environment by scaling the pressure profile with a single temperature \( kT = 5.0 \times \left( \frac{M_{500}}{M_{\odot}} \right)^{1/2} \times 10^{14} \) keV, based on the empirical relationship determined by Arnaud et al. (2010).
- For each model we assume a central peak magnetic field strength \( B_{0} \), following the prescriptions in the numerical radio galaxy simulations by Hardcastle & Krause (2014), of \( B_{0} = 7\sqrt{kT/\mu G} \) (calibrated by observations of groups and clusters, see e.g. Guidetti et al. 2012), with a randomly chosen direction (positive or negative). We then determine the magnetic field profile \( B(l) \) by scaling peak field strength with the density profile using \( B(l) \propto n_{e}(r) \gamma \), where we use \( \gamma = 0.9 \), as found by Dolag et al. (2001) and Dolag (2006) using correlations between rms \( RMs \) and X-ray surface brightnesses for groups and clusters. We then account for the fact that we only consider the line of sight component of the total magnetic field, so that \( B(l)_{\|} = B(l)/\sqrt{3} \).
- We perform the integral \( R_{M_{\text{model}}} = 0.812 \int_{L_{1}}^{L_{2}} n_{e} \tilde{B}_{\|} d\tilde{U} \) incrementally for each model to calculate the total \( R_{M} \), where \( L' \) is the distance from the polarized emission (either core or hotspot(s) for our sources) to the observer. We visually classified sources as having either: one polarized hotspot, two polarized hotspots or a polarized core. We assume in this model that the AGN is located at the centre of the ICM/IGM, and hence the location of a hotspot is half the linear size of the source in projection from the centre of the ICM/IGM, from where the radial profiles \( n_{e}(r') \) and \( B(r') \) are taken. For core-polarized sources we assume a jet with a polarized hotspot at an arbitrarily small distance of 1 pc from the center of the environment, from where the profiles are taken. Due to spherical symmetry of the environment the choice of hotspots (east or west), for core-polarized sources and for sources with one detected hotspot, does not change our results. For models with two polarized hotspots, we calculate the \( R_{M} \) from both hotspots and take the mean, as is done for our observations. In Figure 15 we display the distribution of projected physical distances of the polarized emission from the cluster/group centre (i.e half the projected physical size) for the sources in our sample. Note we do not plot the core-polarized sources as their polarized emission has been fixed at 1 pc from the centre of their environments. We immediately see that the sources where both hotspots are polarized are significantly larger (in projection) than sources with a polarized hotspot in one lobe (note the \( p \)-value from a WMU test in the figure heading). This implies that detecting polarized hotspots in both lobes requires larger sources where, assuming an environment with radially decreasing density (as in our model), the hotspots are located in a less dense medium where the effects of Faraday rotation are less severe.
- The values in the integral above depend on the field reversal scale and orientation \( \theta \). We sample a uniform distribution of scales between \( 10^3 \) pc and \( 10^5 \) pc. The choice in field reversal scales were chosen so that they sample the range of scales consistent with observations of groups and clusters (e.g. \( \sim 10^4 \) pc; Laing et al. 2008) and with cosmological magnetohydrodynamic simulations of clusters (e.g. 1 Mpc; Dolag et al. 2002). In each \( R_{M} \) integral calculated, the initial sign of \( \tilde{B}_{\|} \) is changed every time the incremental path length in the integral reaches the reversal scale. For the orientation angle \( \theta \) we draw 1000 values from the distribution \( p(\theta) = 1/2 \sin(\theta) \) within the range \( 0^\circ < \theta < 180^\circ \), with \( 0^\circ \) being perpendicular to the plane of the sky and towards the observer and \( 180^\circ \) being away from the observer.
- We then truncate each resulting distribution of modelled \( RMs \) to lie only in the range of those that we sample through \( RMs \) synthesis: \(-150 < RMs(\text{rad m}^{-2}) < +150 \). This allows the model to return \( RMs \) that LOFAR can be sensitive to in our analysis.

There are caveats to this model which we address before describing our results. The first is in the use of a uniform distribution of field reversal scales, the values of which are relatively unconstrained for the environments around the population of radio galaxies, though we reiterate that they are in approximate agreement with the few studies that do constrain them. Another caveat is that our observations have a finite beam size, whereas we have modelled our sources through single line(s) of sight towards a hotspot(s) with infinite resolution. This is likely to be a very minor affect on our results as the polarized emission we observe is predominantly unresolved, and we take pixel-averaged values as the observed \( RMs \) within the beam. The final caveat is that we have assumed all sources lie in the geometric centres of their environments. This is supported observationally by studies finding that the majority of X-ray-selected clusters host a central RLAGN (e.g. Magliocchetti & Brüggen 2007; Best et al. 2007), though it is possible that a small number of our sources do not lie in the geometric centers of their environments. In general, while our results (discussed below) are clearly related to our choice of input distributions, we use observationally calibrated
results where possible with currently available data. Furthermore our choice of 1000 values drawn from each unknown parameter distribution was made to ensure the models are stochastic.

We separated our models into those which lead to RM$\delta$s within the range of the observed galaxy-subtracted RM$\delta$ distribution of our sample (matched RM$\delta$s; $-20 < \text{RM} (\text{rad m}^{-2}) < +25$, see Figure 12) and those that are not (unmatched RM$\delta$s; outside our observed range but within those that we sample with RM$\delta$ synthesis and that LOFAR is sensitive to, i.e. $-150 < \text{RM} (\text{rad m}^{-2}) < +150$). In terms of the model statistics, we have $\sim 61000/67000$ models within our RM$\delta$ synthesis range, of which $\sim 32000$ are within our observed range, i.e 50 per cent of our models are matched to our observed RM$\delta$ range. Given our observed detection fraction of 18 per cent, the model predicts a factor of three higher detection fraction in the RM$\delta$ range that we observe. We partly associate this with our selection bias for our sample: our models have a matched-RM$\delta$ fraction of $\sim 58$ per cent for core-polarized sources, compared to our observed detection fraction for such sources of 3.4 per cent. While in reality these are mostly based on the FR-I radio galaxies in our sample, such modelled RM$\delta$s would also come from polarized blazars (e.g. as found by O’Sullivan et al. 2018 in LoTSS), which have been excluded from our sample through our selection of large angular size (>100 arcsec) sources. Hence, the incompleteness in our sample removes polarized sources that we may detect, whereas our models do not take into account any flux or angular size limit. However, more importantly, core-polarized sources in our models that have matched RM$\delta$s produce similar RM$\delta$s based on all orientation angles (since the 1 pc distance of their polarized emission at any orientation from the centre of the environment profile does not significantly affect the final aggregated RM$\delta$), our models are already biased towards a very high matched fraction. Further, since FR-I sources (which tend to have polarized cores) live in rich environments relative to FR-IIs, it is possible that the Schechter function we use for all models underestimates the cluster/group masses for FR-IIs as a population.

In top panel of Figure 16 we compare the resulting distribution of RM$\delta$ for matched models and our observations. We see a fairly good agreement between the distributions, with a KS test between both distributions having a p-value of $\sim 0.05$. Given this statistical similarity, we can analyse the observable and physical properties of the matched models to give inferences of the polarization detectability at low frequencies. For completeness we also show the distributions between matched and unmatched RM$\delta$s for our models (bottom panel of Figure 16), showing a clear peak for models in our observational range and a strong decrease in model counts beyond this range, highlighting that our model inputs are appropriate in predicting RM$\delta$s that we observe with LOFAR.

### 3.5.2 Model properties

To give an example of the effects of the RM$\delta$ with $M_{500}$ we display our models for ILTJ112543.21+543903.2, a polarized source in our sample with a projected size of 673 kpc and with one polarized hotspot, modelled with a field reversal scale of $10^{5.5}$ pc, color-coded by the orientation angle in each model in Figure 17. For this type of source we see that there is a preference for the jet with a polarized hotspot to be inclined toward the observer (≤ 90°) and as
Polarization properties of radio galaxies

In Figure 18, we plot the input physical parameters of our matched and unmatched models. Grey hexagons are our unmatched models (color-coded by density of counts), yellow points are matched models with a polarized core, orange points are matched models with one polarized hotspot and blue points are matched models with polarized hotspot from both lobes.

In Figure 18, we plot the input physical parameters of our matched and unmatched models. In the top panel we plot the source orientation against $M_{500}$ where we see that for any given orientation, there is a higher matched model fraction at masses below $10^{13.5} M_\odot$, i.e. sources tend to be depolarized in higher mass environments as the resulting $[RM]$ would be too high to detect in our observed range (and in reality they would cause higher $RM$ dispersion resulting in depolarization at 150 MHz). As a comparison with sources for which spatially-resolved $RM$ images have been obtained (at high frequencies), the clusters surrounding Cygnus A and Hydra A are at $M_{500} = 2.8 \times 10^{14} M_\odot$ (Wilson et al. 2002) and $M_{500} = 1.7 \times 10^{14} M_\odot$ (Zhang et al. 2017) respectively, while a less rich group such as that surrounding 3C31 has a mass of $6.3 \times 10^{13} M_\odot$ (Komossa & Böhringer 1999). This implies that LOFAR is preferentially sensitive to polarized sources in less rich clusters and poor environments, and is consistent with our earlier remarks that with LOFAR we are sensitive to polarized radio galaxies with low dispersion in Faraday depth. Interestingly we see that, for sources with a polarized hotspot in one lobe (orange circles), there is a strong preference for angles $\leq 90^\circ$, meaning that we are seeing the approaching jet which is inclined towards the observer (as seen in Figure 17 for one source). This is due to the fact that such sources will tend to have relatively low $|RM|$ due to smaller path lengths through the line of sight and they experience less Faraday rotation, such that they are within our observed $RM$ range with LOFAR. On the other hand, core-polarized sources (yellow) are populated at all orientation angles, since the jet orientation of their assumed polarized emission at a distance of 1 pc from the centre does not significantly affect the final aggregated $RM$ at $z = 0$. Double polarized sources (polarized hotspots in both lobes of an FR-II radio galaxy; blue points) seem to be almost exclusively populated at angles $\sim 90^\circ$ (i.e. on the plane of the sky), as would be expected since the hotspot from the receding jet at a larger angles from the plane of the sky can become more easily depolarized (Laing-Garrington effect; Laing 1988; Garrington et al. 1988). Hence, according to our model, LOFAR would tend to only detect both hotspots in polarization if the source is on the plane of the sky.

In summary, we find that, in a model with only external Faraday rotation from the local environment of a source, polarized radio galaxies at low frequencies are predominantly detected in cluster/group masses $\lesssim 10^{13.5} M_\odot$, while polarized hotspots are preferentially seen when the jets are on the plane of the sky, otherwise only one hotspot from the approaching jet is seen (Laing 1988; Garrington et al. 1988). We reiterate that our model distributions are a direct product of the assumed input parameter distributions, and in depth analyses of cluster $RM$s are needed to test the robustness of our inputs.

4 DISCUSSION & CONCLUSIONS

We have analyzed 20 arcsec resolution polarization data of radio galaxies from part of the upcoming LoTSS DR2. This statistical study of the bulk properties has extended the work of OS18 and
We find that at 150 MHz the polarization detection fraction increases with total flux density, as expected; however, the distributions in angular size between detected and non-detected sources are statistically indistinguishable for sources > 100 arcsec. This trend may be biased due to our selection criteria, and it is possible that the polarized detection fraction for RLAGN increases with smaller angular size due to the presence of blazars. We confirm the conclusions of OS18 that, in terms of resolved sources, the hotspots of FR-II radio galaxies are predominantly detected even at low frequencies. FR-II radio galaxies are not only brighter and more luminous, but they are known to reside in less rich environments than FR-Is, and so physical depolarization due to the ambient medium is less prominent, particularly since the brightest emission is in the hotspots which are far away from the densest part of the IGM/ICM, contrary to the case for FR-Is. The morphologically-classed FR-Is in our sample generally have a higher polarized flux and fractional polarization than the FR-I sources over the range in total flux density, though with large overlap.

The comparison of host galaxy photometry between polarized and depolarized sources further highlights the importance of morphology in polarization – accounting for morphology, host galaxies with higher values of WISE colour (more AGN-like on a WISE colour-colour diagram) seem to drive RLAGN with a higher detection fraction of polarization. The observed low frequency polarization is related to FR morphology rather than WISE colour, with the more powerful FR-Is having a high detection fraction, though this case may be unrelated for intrinsic polarization, for which past studies have found significant differences between HERGs and LERGs (which at low redshift tend to be associated with FR-Is and FR-Is, respectively). More dense cluster environments contributing to higher internal depolarization via entrainment of the thermal material is a possible explanation for the lack of polarized FR-Is, but our data provide no direct evidence for this hypothesis.

For the sources that have polarized counterparts to the sources in our sample at 1400 MHz, we find that they are de-polarized (weaker but detectable polarization) at 150 MHz for all but one source. This is further confirmed by Figure 10, which shows that the distribution in 150 MHz total flux density is significantly higher for those NVSS sources that are detected with LOFAR, whereas the non-detected sources are not bright enough to produce sufficient polarized emission to be detected. The spectral index distributions imply that the sources detected by LOFAR have significantly steeper spectral indices on average, explaining the lack of polarized NVSS counterparts to the polarized sources in our sample.

Modelling of the environments toward radio galaxies and their subsequently integrated $RMs$ shows that, for a range of cluster/group masses, field reversal scales and jet orientation angles, we would expect to preferentially observe polarized hotspots that are inclined towards the observer, for the case where a hotspot from one lobe is detected in polarization. For the case where hotspots in both lobes are detected, our models indicate that the jets are on the plane of the sky, consistent with the Laing-Garrington effect. Core-polarized sources are generally favoured at all orientation angles as a function $M_{bolo}$ and reversal scale (due to our model assumption that they have a compact polarized component at 1 pc from the centre). Our results generally imply that there is a very low chance of detecting a polarized radio galaxy at 150 MHz if it is in even a moderately rich environment ($M_{bolo} \gtrsim 10^{11.5} M_\odot$, depending on its orientation angle or physical size (Figure 15), as the $RMs$ would be too high to observe at low frequencies. We reiterate that our results are dependent on our input model parameters, which are based on empirical relationships and observations, but must be validated by more robust numerical modelling.

Our overall results imply that detecting polarized radio galaxies with LOFAR at 150 MHz is related to the combination of total flux density, environment and jet orientation. These results will be useful in determining the properties of polarized sources in the full LoTSS survey, which is expected to contain around ten million radio sources.

ACKNOWLEDGEMENTS
We thank Aritra Basu, Ettore Carretti, Rainer Beck, Błażej Nikiel-Wroczyński and the anonymous referee for helpful comments in improving this paper.

VHM thanks the University of Hertfordshire for a research studentship [ST/N504105/1]. MJH acknowledges support from the UK Science and Technology Facilities Council [ST/R000905/1].

This paper is based (in part) on data obtained with the International LOFAR Telescope (ILT). LOFAR (van Haarlem et al. 2013) is the Low Frequency ARray designed and constructed by ASTRON. It has observing, data processing, and data storage facilities in several countries, which are owned by various parties (each with their own funding sources), and are collectively operated by the ILT foundation under a joint scientific policy. The ILT resources have benefitted from the following recent major funding sources: CNRS-INSU, Observatoire de Paris and Université d’Orléans, France; BMBF, MIWF-NRW, MPG, Germany; Science Foundation Ireland (SFI), Department of Business, Enterprise and Innovation (DBEI), Ireland; NWO, The Netherlands; The Science and Technology Facilities Council, UK; Ministry of Science and Higher Education, Poland.

This research made use of the Dutch national e-infrastructure with support of the SURF Cooperative (e-infra 180169) and the LOFAR e-infra group. The Jülich LOFAR Long Term Archive and the German LOFAR network are both coordinated and operated by the Jülich Supercomputing Centre (JSC), and computing resources on the supercomputer JUWELS at JSC were provided by the Gauss Centre for Supercomputing e.V. (grant CHTB00) through the John von Neumann Institute for Computing (NIC).

This research made use of the University of Hertfordshire high-performance computing facility and the LOFAR-UK computing facility located at the University of Hertfordshire and supported by STFC [ST/P000096/1], and of the Italian LOFAR IT computing infrastructure supported and operated by INAF, and by the Physics Department of Turin university (under an agreement with Consorzio Interuniversitario per la Fisica Spaziale) at the C3S Supercomputing Centre, Italy.

DATA AVAILABILITY
The data underlying this article will be shared on reasonable request to the corresponding author, after the LoTSS DR2 data is made available to the public in early 2021.

REFERENCES
Arnaud M., Pratt G. W., Piffaretti R., Bohringer H., Croston J. H., Pointecouteau E., 2010, A&A, 517, A92

STUARDI ET AL. (2020) WITH THE USE OF AN OPTICALLY IDENTIFIED SAMPLE OF 382 CLASSIFIED RADIO GALAXIES, WITH 67 DETECTED IN POLARIZATION.

STUARDI ET AL. (2020) WITH THE USE OF AN OPTICALLY IDENTIFIED SAMPLE OF 382 CLASSIFIED RADIO GALAXIES, WITH 67 DETECTED IN POLARIZATION.
Figure A1. Polarized intensity maps (left; see Figure 1 for details) and the Faraday spectrum (right) of their peak polarized intensity pixel. The red cross denotes the $RM$ of the pixel as found using $RM$ synthesis (neglecting $-3 \leq \phi < \text{rad m}^{-2}$ for leakage signal).