







# Identifying active galactic nuclei via brightness temperature with sub-arcsecond international LOFAR telescope observations

Leah K. Morabito <sup>1,2</sup>★ F. Sweijen,<sup>3</sup> J. F. Radcliffe <sup>4,5</sup> P. N. Best,<sup>6</sup> Rohit Kondapally <sup>6</sup>, Marco Bondi,<sup>7</sup> Matteo Bonato <sup>7,8</sup>, K. J. Duncan <sup>6</sup>, Isabella Prandoni <sup>7</sup>, T. W. Shimwell,<sup>3,9</sup> W. L. Williams,<sup>3</sup> R. J. van Weeren,<sup>3</sup> J. E. Conway<sup>10</sup> and G. Calistro Rivera<sup>11</sup>

<sup>1</sup>Centre for Extragalactic Astronomy, Department of Physics, Durham University, Durham DH1 3LE, UK

<sup>2</sup>Institute for Computational Cosmology, Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK

<sup>3</sup>Leiden Observatory, Leiden University, P.O. Box 9513, NL-2300 RA Leiden, the Netherlands

<sup>4</sup>Department of Physics, University of Pretoria, Lynnwood Road, Hatfield, Pretoria 0083, South Africa

<sup>5</sup>Jodrell Bank Centre for Astrophysics, University of Manchester, Oxford Road, Manchester M13 9PL, UK

<sup>6</sup>Institute for Astronomy, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK

<sup>7</sup>INAF - Istituto di Radioastronomia, Via P. Gobetti 101, I-40129 Bologna, Italy

<sup>8</sup>Italian ALMA Regional Centre, Via Gobetti 101, I-40129 Bologna, Italy

<sup>9</sup>ASTRON, Netherlands Institute for Radio Astronomy, Oude Hoogeveensedijk 4, NL-7991 PD Dwingeloo, the Netherlands

<sup>10</sup>Department of Space, Earth and Environment, Chalmers University of Technology, Onsala Space Observatory, SE-439 92 Onsala, Sweden

<sup>11</sup>European Southern Observatory, Karl-Schwarzschild-Straße 2, D-85748 Garching bei München, Germany

Accepted 2022 July 26. Received 2022 July 26; in original form 2022 May 31

## ABSTRACT

Identifying active galactic nuclei (AGNs) and isolating their contribution to a galaxy’s energy budget is crucial for studying the co-evolution of AGNs and their host galaxies. Brightness temperature ( $T_b$ ) measurements from high-resolution radio observations at GHz frequencies are widely used to identify AGNs. Here, we investigate using new sub-arcsecond imaging at 144 MHz with the International LOFAR Telescope to identify AGNs using  $T_b$  in the Lockman Hole field. We use ancillary data to validate the 940 AGN identifications, finding 83 percent of sources have AGN classifications from SED fitting and/or photometric identifications, yielding 160 new AGN identifications. Considering the multiwavelength classifications, brightness temperature criteria select over half of radio-excess sources, 32 percent of sources classified as radio-quiet AGNs, and 20 percent of sources classified as star-forming galaxies. Infrared colour–colour plots and comparison with what we would expect to detect based on peak brightness in 6 arcsec LOFAR maps imply that the star-forming galaxies and sources at low flux densities have a mixture of star-formation and AGN activity. We separate the radio emission from star-formation and AGN in unresolved,  $T_b$ -identified AGNs with no significant radio excess and find the AGN comprises  $0.49 \pm 0.16$  of the radio luminosity. Overall, the non-radio excess AGNs show evidence for having a variety of different radio emission mechanisms, which can provide different pathways for AGNs and galaxy co-evolution. This validation of AGN identification using brightness temperature at low frequencies opens the possibility for securely selecting AGN samples where ancillary data are inadequate.

**Key words:** acceleration of particles – radiation mechanisms: non-thermal – galaxies: active – galaxies: jets – radio continuum: galaxies.

## 1 INTRODUCTION

Understanding how active galactic nuclei (AGNs) interact with their host galaxies is one of the main challenges in astronomy today. The fact that some sort of feedback links the growth of galaxies to the supermassive black hole which powers the AGNs is widely accepted. Tight empirical relations between observed host galaxy properties (e.g. velocity dispersion, bulge mass; Ferrarese & Merritt 2000; Gebhardt et al. 2000) and the mass of the central supermas-

sive black hole are a strong, but indirect, argument for feedback. Cosmological simulations require AGN feedback to suppress the growth of the most massive galaxies and reproduce the stellar mass function we observe in the Universe today (e.g. Bower et al. 2006; Croton et al. 2006). Although it is clear from both observations and theory that AGN feedback is an important component in a galaxy’s growth, we do not yet understand the details of how AGN feedback works.

A significant limitation in understanding how AGN feedback works is what it operates on: star formation. AGN feedback can stimulate star formation (e.g. De Young 1981; Silk 2013; Zhuang, Ho & Shangqian 2021) by creating the dense, turbulent gas conditions

\* E-mail: [leah.k.morabito@durham.ac.uk](mailto:leah.k.morabito@durham.ac.uk)

which lead to star formation. Feedback could also suppress/quench star formation by heating or removing the gas from which stars would form (see e.g. Bower et al. 2006; Greene et al. 2011; King & Pounds 2015). Either way, it can be difficult to pick apart observational evidence to understand which is happening (Ward et al. 2022). provided excellent evidence for AGN feedback, particularly on large scales. High-redshift radio-loud AGNs exhibit an alignment between the optical and radio axes, implying that the radio jets may be stimulating star formation (Chambers, Miley & van Breugel 1987; Best, Longair & Rottgering 1996; Dey et al. 1997; Nesvadba et al. 2020). In more local radio-loud AGNs in the centres of clusters, we observe a balance between the radio jet power and the work it would take to create the co-spatial cavities in the hot gas seen in X-ray observations (e.g. McNamara et al. 2000; Birzan et al. 2004; Wise et al. 2007). Galaxy-scale radio jets can also have internal lobe energy comparable to the energy in the interstellar medium (ISM; Webster et al. 2021), indicating that the ability of radio jets to impact a galaxy’s evolution extends down to smaller scales.

While it is easy to identify large-scale radio jets, once on sub-galactic scales the problem compounds: both AGN activity and star formation generate radio emission, making it difficult to understand their interplay. It is also not clear what the source of radio emission is in radio-quiet AGNs; it may be due to small-scale jets, winds, or even star formation (for a review, see Panessa et al. 2019). The jury is even still out on whether the radio emission in radio-quiet AGNs is linked to AGN activity (e.g. Zakamska & Greene 2014; White et al. 2015; Macfarlane et al. 2021) or star formation (e.g. Condon et al. 2013; Padovani et al. 2015; Gürkan et al. 2018; Radcliffe et al. 2021a). Multiwavelength observations can help identify sources with radio excess above that expected from star formation (e.g. Wilson 1988; Drake et al. 2003; Del Moro et al. 2013; Calistro Rivera et al. 2017), but many AGNs lie on the radio to far-infrared correlation (e.g. Sopp & Alexander 1991; Bonzini et al. 2015) and cannot be identified in this way.

The only way to unambiguously identify radio emission from AGN activity and simultaneously separate it from star formation is via brightness temperature,  $T_b$ . This is defined as the temperature of the blackbody which would produce the observed surface brightness (flux density per solid angle) at the observed radio frequency. There is a limit to the amount of radio flux density per solid angle that can be generated by star formation, even in the most luminous starburst galaxies (Condon 1992) and particularly at redshifts  $z > 0.1$  (Kewley et al. 2000). Therefore, a value of brightness temperature above this limit provides a secure AGN classification, although the converse is not true: we cannot securely say there is no AGN in sources below this brightness temperature limit, although it is likely they are dominated by star formation.

Brightness temperature measurements are often used in very long baseline interferometry (VLBI) experiments, where the high resolution can measure the appropriate parameter space to distinguish between star formation and AGN activity. At GHz frequencies, wide-field VLBI has shifted from processing the entire primary beam for a single correlated phase centre (Garrett et al. 2001) to the more computationally feasible processing of multiple phase centres correlated independently across a larger field of view (Deller et al. 2011; Morgan et al. 2011). This has allowed surveys to follow up radio sources with brightness temperature measurements in targeted observations of select fields (e.g. Muxlow et al. 2005; Chi, Barthel & Garrett 2013; Radcliffe et al. 2018), up to areas of  $\sim 2 \text{ deg}^2$  (Herrera Ruiz et al. 2017). However, the limiting value of  $T_b$  depends not only on resolution but also frequency (see Section 3.2).

The International LOFAR Telescope (ILT; van Haarlem et al. 2013) is a phased array operating at 10–240 MHz, with stations spread across Europe. With baselines up to  $\sim 2000 \text{ km}$ , the ILT can achieve sub-arcsecond resolution at MHz frequencies. Sub-arcsecond imaging with the full ILT shares many common challenges with high-frequency VLBI, but with the addition of its own unique challenges at low frequencies (for more details, see Morabito et al. 2022). Imaging of individual sources has yielded significant results (e.g. Varenus et al. 2015; Morabito et al. 2016; Ramírez-Olivencia et al. 2018; Groeneveld et al. 2022; Kukreti et al. 2022; Timmerman et al. 2022), but these studies have remained in the realm of bright sources which can be self-calibrated. The real power of the ILT is in its wide field of view, allowing access to fainter sources. Sweijen et al. (2022) recently published the first full field-of-view image covering  $6.6 \text{ deg}^2$  at the full resolution of the ILT. In this  $\sim 7$  billion pixel image, there are 2316 sources detected with  $5\sigma$  significance (after removing duplicate sources).

In this paper, we exploit this ILT data set to study the brightness temperatures of the low-frequency radio population, for the first time. We begin by describing the data in Section 2. Section 3.2 opens with a discussion on brightness temperature at low frequencies, followed by the sample selection. Results are presented in Section 4 followed by the separation of star formation and AGNs in Section 5. Discussion and conclusions are in Sections 6 and 7, respectively. Throughout the paper, we assume the WMAP9 cosmology (Hinshaw et al. 2013) in `ASTROPY` and define radio flux density as  $S_\nu \propto \nu^\alpha$ , where  $\alpha$  is the radio spectral index. The code which generated the plots and results is publicly available at [https://github.com/lmorabit/ILT\\_AGNdetect](https://github.com/lmorabit/ILT_AGNdetect).

## 2 DATA

### 2.1 Lockman Hole data

We use the catalogue from Sweijen et al. (2022). This catalogue is drawn from the first high-resolution ( $0.3 \text{ arcsec} \times 0.4 \text{ arcsec}$ ) wide-field image made with LOFAR. This image, which covers  $6.6 \text{ deg}^2$  of the Lockman Hole, has a median rms noise level (across the rms map) of  $34 \mu\text{Jy beam}^{-1}$ . We will refer to this image as the *high-resolution image*. We refer the reader to Sweijen et al. (2022) for full details on how this image was generated, but briefly summarize relevant information here.

The wide-field image is made up of 25 ‘facets’ following the technique developed by van Weeren et al. (2016, 2021). Corrections for the point spread function (PSF) and astrometry were applied per facet, and a final flux density scale correction applied for the whole field. The Python Blob Detector and Source Finder (PYBDSF; Mohan & Rafferty 2015) was run individually on the 25 facet images before the outputs were combined into a single catalogue, where duplicate sources were removed. Removal of duplicate sources was done first using simple positional matching and followed up with visual identification. The resulting catalogue contains 2316 unique sources where the peak brightness,  $S_p$ , is  $\geq 5\sigma$  (using the rms noise for the associated island). The median rms noise in the full image is  $34 \mu\text{Jy beam}^{-1}$ , although this varies radially and is lower in the centre.

This high-resolution catalogue was matched to the 6 arcsec resolution image made as part of the LOFAR Surveys Deep Fields Data Release 1 (Tasse et al. 2021). We will refer to this image as the *standard-resolution image*. All of the sources detected at  $5\sigma$  in the high-resolution image have a match in the standard-resolution image, which is unsurprising as it uses  $\sim 100 \text{ h}$  of data to reach  $\lesssim 23 \mu\text{Jy beam}^{-1}$  in the inner part of the field, rather than the 8 h

used for the high-resolution image. This provides information on angular scales  $\geq 20$  times larger than the high-resolution catalogue, and in particular, is more sensitive to diffuse emission.

As part of the Deep Fields Data Release 1, science-ready catalogues were produced that provide extensive information at infrared (IR) through X-ray bands. Kondapally et al. (2021) present the identification of counterparts with radio sources, and the careful compilation of available multiwavelength information, for 31 162 sources in the standard-resolution image of the Lockman Hole field. This includes some flags for the identification of AGNs from optical, IR, and X-ray information. The catalogues also include photometric redshifts from Duncan et al. (2021), as only  $\sim 10$  percent of sources have spectroscopic redshifts. For the analysis presented in this paper, we remove sources which have no redshift information available ( $< 5$  percent), leaving 2214 sources from the high-resolution image.

## 2.2 Parameters from SED fitting

Source classifications, stellar masses, and star formation rates were calculated by Best et al. (in preparation) using SED fitting from the far-IR to the ultraviolet (UV). The SED fitting was carried out on all galaxies detected in the standard-resolution image using four different SED fitting software. The first two, MAGPHYS (da Cunha, Charlot & Elbaz 2008) and BAGPIPES (Carnall et al. 2018, 2019), operate on an energy balance approach, where the energy absorbed by dust in the optical and UV bands must match the thermal emission from dust in the far-IR through sub-mm. While these two codes generally give consistent results for high signal-to-noise data, they do not include any AGN component. CIGALE (Burgarella, Buat & Iglesias-Páramo 2005; Noll et al. 2009; Boquien et al. 2019) also operates on the energy balance principle, and includes an AGN component. The final code, AGNFITTER (Calistro Rivera et al. 2016), is designed to work in cases where the assumption of energy balance may not hold if the UV light and IR emission are spatially distinct. This is often the case for AGNs.

Best et al. (in preparation) modelled all of the galaxies using each of the four codes, including twice for CIGALE, using a different suite of models for the AGN component each time. This provided radiative mode AGN identifications. Radio AGNs were identified using the relationship between radio luminosity at 150 MHz,  $L_{150}$ , and the derived SFRs. Best et al. (in preparation) defined radio-excess identified AGNs as lying further than 0.7 dex from the ridge line of this relationship. For all fitted galaxies, the consensus stellar masses and SFRs were derived (Section 5; Best et al. in preparation) using a Chabrier initial mass function (Chabrier 2003).

## 3 SAMPLE SELECTION

### 3.1 Brightness temperatures at low frequencies

Frequently, in the literature one will see that a source is reliably considered to be an AGN if it exceeds  $T_b > 10^5$  K. However, the exact limiting value of  $T_b$  depends on the observing frequency, non-thermal spectral index, the redshift, and the temperature of the gas. It is worth revisiting this in light of these new LOFAR observations.

We start from brightness temperature, which is defined (in the Rayleigh–Jeans approximation) as

$$T_b = \frac{c^2}{2k_b} \frac{S_\nu}{\nu^2 \Omega} \text{ K}, \quad (1)$$

where  $c$  is the speed of light,  $k_b$  is the Boltzmann constant,  $S_\nu$  is the flux density,  $\nu$  is the observing frequency, and  $\Omega$  is the source solid

angle. For a Gaussian source,

$$\Omega = \frac{\pi \theta_1 \theta_2}{4 \ln 2}, \quad (2)$$

where  $\theta_1, \theta_2$  are the fitted Major and Minor axes of a 2D Gaussian. Rearranging equation (1), we arrive at:

$$\frac{S_\nu}{\Omega} = \frac{T_b \nu^2}{1.38 \times 10^{24}}, \quad (3)$$

in units of Jy arcsec $^{-2}$  (flux density per solid angle).<sup>1</sup>

Starting from the simplest case, the brightness temperature from a normal star-forming galaxy (SFG) (Condon 1992) is given by

$$T_b = T_e [1 - \exp(-\tau)] \left( 1 + 10 \left( \frac{\nu}{\text{GHz}} \right)^{0.1+\alpha} \right), \quad (4)$$

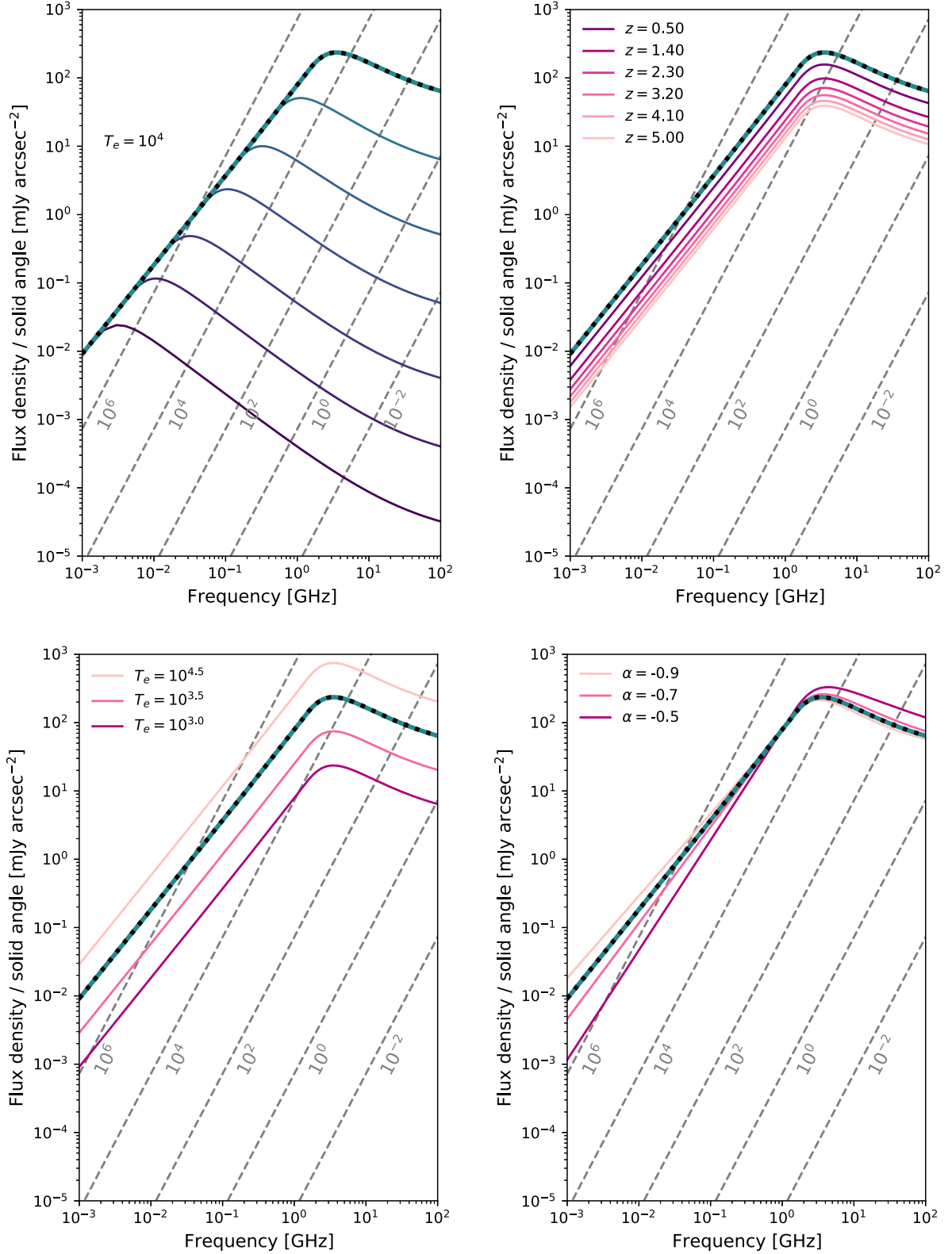
with  $\tau = \tau_0 \left( \frac{\nu}{\nu_0} \right)^{-2.1}$ ,

where  $\tau_0 = 1$  at frequency  $\nu_0$ . In equation (4), we can see that the brightness temperature depends on the frequency at which the optical depth reaches unity,  $\nu_0$ , and the electron temperature of the gas,  $T_e$ . It also depends on the synchrotron spectral index,  $\alpha$ . For this initial study, we choose to use this simple model which assumes a single-phase ISM, and build on this in future work.

Inserting equation (4) into equation (3) yields a beam-independent model of flux per solid angle, and we arrive at Fig. 1, which shows the flux per solid angle versus observing frequency. The top left panel is a reproduction of fig. 4 from Condon (1992), with a range of values for  $\nu_0$ , for a typical electron temperature  $T_e = 10^4$  K. As  $\nu_0$  increases, the frequency at which the flux per solid angle peaks also increases. At an observed frequency of 1.4 GHz, the maximum value allowed out of all of the curves is  $\sim 10^5$  K. Of course, if  $\nu_0$  causes the flux per solid angle to peak at a lower frequency, the limiting value of  $T_b$  will drop. As the value of  $\nu_0$  is virtually impossible to measure accurately for a distant galaxy, the most restrictive curve (i.e.  $\nu_0 = 3$  GHz) must be assumed for secure AGN identification. Assuming this most restrictive value of  $\nu_0$ , the three remaining panels of Fig. 1 show how this changes with redshift, electron temperature, and synchrotron spectral index. As the redshift increases, the curve shifts down, and therefore the limiting value of  $T_b$  drops. Lower values of  $T_e$  will also shift the curve down, while higher values shift it up. Synchrotron spectral index changes the shape of the curve, with steeper spectral indices reaching higher limiting values of  $T_b$  at low frequencies (but lower at high frequencies).

It is important to remember that  $T_b$  is not a physical temperature, and that the limiting value which divides star formation and AGN activity can change based on several parameters, e.g. with observed frequency: for  $z = 0$ ,  $T_e = 10^4$  K the limiting value is  $T_b \sim 10^5$  K at 1.4 GHz, but closer to  $T_b \sim 10^6$  K at 144 MHz. The limiting value of  $T_b$  should therefore be calculated per galaxy using all information available. For this work, we assume  $T_e = 10^4$  K, use the redshift of each individual source, and assume a spectral index of  $\alpha = -0.8$ . This last assumption is conservative: we know that compact sources tend to have flatter spectral indices (e.g. Jackson et al. 2022), which would lead to a lower limiting value of  $T_b$  (see bottom right of Fig. 1). In the absence of complementary spatially resolved spectral index information (the integrated spectral index in the standard resolution image may not be the same as the compact component), we opt to fix the spectral index for all sources. An ongoing survey with

<sup>1</sup>Note that the solid angle as defined in equation (2) is rarely explicitly used; a commonly seen version of equation (3) will absorb the constants and use  $\theta_1 \theta_2$  directly, resulting in a constant of  $1.22 \times 10^{24}$  instead.



**Figure 1.** The top left panel shows flux density per solid angle versus frequency (reproduction of fig. 4 in Condon 1992). The dashed grey lines show constant brightness temperatures, as labelled in the plot. The coloured lines show the limiting flux per solid angle as a function of frequency, for  $\nu_0$  (the frequency at which the optical depth is unity) of 3 MHz, 10 MHz, 30 MHz, 100 MHz, 300 MHz, 1 GHz, and 3 GHz (from purple to green, bottom to top). The line with  $\nu_0 = 3$  GHz is thicker in width and marked with black dots, and this is plotted in the other three panels for reference. For  $\nu_0 = 3$  GHz, the top right panel shows how this changes with redshift. The bottom left panel shows how it changes with electron temperature,  $T_e$ , while the bottom right panel shows the change with spectral index,  $\alpha$ . In all figures the  $z = 0$ ,  $T_e = 10^4$  K,  $\alpha = -0.8$  line is shown as the thicker line dotted with black.

e-MERLIN and the EVN (PI: McKean) will, in the future, provide appropriate spectral index information to be able to measure the spectral index per source. This is more appropriate than comparison with a previous VLBA study (Middelberg et al. 2013), which is discussed in Appendix A.

### 3.2 Selecting AGNs using brightness temperature

We select AGNs by calculating the limiting brightness temperature assuming  $\nu_0 = 3$  GHz,  $T_e = 10^4$  K,  $\alpha = -0.8$ , and the specific redshift per source. We use spectroscopic redshifts where they exist, supplemented by photometric redshifts for the rest of the sample. Practically, we need the flux density information,  $S$ , and the area over which it is spread,  $\Omega$ . For the area over which the flux density is spread, ideally we would use the real angular size of the source. This information is not directly available, so we use the deconvolved sizes from PYBDSF. Where the deconvolved major or minor axis is smaller than the limiting resolution (0.4 arcsec and 0.3 arcsec, respectively), we replace the size with the limiting resolution (see Lobanov 2005; Radcliffe et al. 2018). This will lead to an underestimation of  $T_b$ , and therefore we expect the AGN identifications to be secure. We treat the major and minor axes separately, and this impacts a total of 950 unique sources.

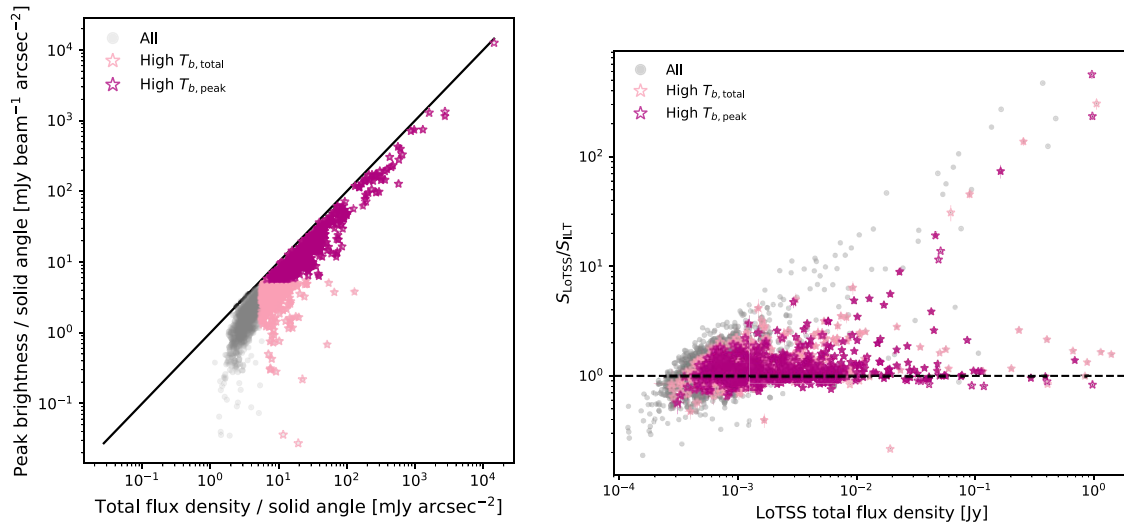
Using deconvolved sizes is not ideal, as the actual PSF (as opposed to the restoring beam, which is the 2D Gaussian fit to the PSF used by PYBDSF) is the convolution of several kernels: (i) the  $u$ - $v$  sampling function (dirty beam), (ii) the autocorrelation function (ACF) of any residual calibration errors, and (iii) the PSF distortions due to smearing. However, the fitted Gaussian restoring beam will be smaller than the actual PSF, meaning that when PYBDSF uses the restoring beam to calculate the deconvolved sizes they will therefore be an *overestimate* of the true size, which leads to an *underestimate* of the brightness temperature. The results in this paper are therefore conservative. To properly calculate the brightness temperature accounting for smearing and other intensity losses, one would need to simulate a dirty beam from the  $u$ - $v$  coverage of a particular observation, convert it to a residual image (this can be done as the dirty beam is the ACF of the noise, see Schreiber et al. 2021 Appendix A for details), compare that to the actual residuals to find the calibration error PSF, and finally calculate the smearing PSF. This procedure is computationally expensive, and we leave it to future work; the imaging strategy is still undergoing optimization and the final procedure is not set in stone. For now, we note that our  $T_b$  estimates are lower limits, and in the future we may be able to access more sources with high values of  $T_b$ . To be clear, this means we expect our current AGN identifications to be secure, and we may be able to add more AGN identifications in the future.

For the flux density we make a calculation from both the peak brightness and the total flux density in the high-resolution catalogue, resulting in a total of 946 unique sources selected. Of these sources, 486 were identified as AGNs using the peak brightness, while all 946 were identified using the total flux density. A comparison of the peak brightness and total flux density per solid angle are shown in the left-hand panel of Fig. 2, while the right-hand panel shows the ratio of integrated flux densities in the standard and high-resolution images as a function of the integrated flux density in the standard resolution image. Sources which were identified by their peak brightness temperature appear to be largely unresolved: they lie close to the line where the peak brightness to total flux density per solid angle ratio is unity. There does not seem to be a preference for sources identified by their total flux density per solid angle to be brighter than those identified by peak brightness per solid angle.

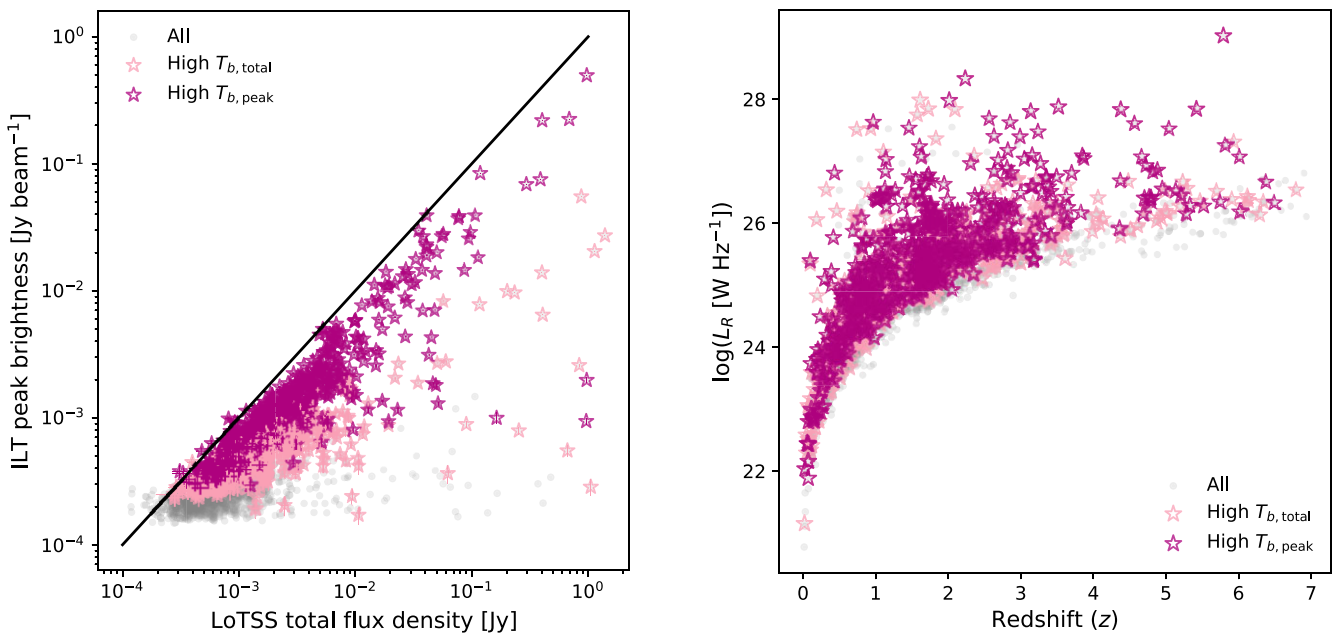
It is interesting to note, in the right-hand panel of Fig. 2, that at high flux densities the sources divide into two clear categories: some sources show a high ratio of integrated flux densities from the standard image ( $S_{\text{LoTSS}}$ ) to the integrated flux density in the high-resolution image ( $S_{\text{ILT}}$ ), while others show very similar flux densities at the two resolutions. There are no sources with intermediate flux density ratios. To explore this further, we consider the sources brighter than 50 mJy (all of which are classified as AGNs via their radio excess and are therefore securely AGNs) and separate them into sources with high ( $>7$ ) and low ( $<7$ ) flux density ratios. We find that, as expected, the high ratio sample show resolved structure in both the high-resolution and standard-resolution images (except for one source that is compact in both), such that the low-resolution image encapsulates more flux. On the other hand, the low ratio sample is divided fairly evenly across three categories: those which are compact in both the high-resolution and standard-resolution images (8/31 sources), those with resolved structure in the high-resolution image but are compact or show only marginal extension in the standard-resolution image (11/31 sources), and those which show resolved structure in both images (12/31). Although the low ratio sample appear to be more compact sources in general, the high-ratio sample have smaller reported major axis values in the high-resolution image; this implies that the catalogued high-resolution size is measuring the size of just a central component (e.g. the radio core) and is not representative of the full size of the source. Given the range of properties seen, this investigation provides no clear insight as to why there are no high flux density sources with intermediate flux density ratios; future imaging of the Lockman Hole at intermediate resolutions (1–2 arcsec) will help answer this question.

As a check of the quality of the catalogues, we compare the peak brightness in the high-resolution image with the integrated flux density in the standard-resolution image in Fig. 3. We find that six of the  $T_b$ -identified sources have peak brightness values higher than their integrated flux density values, but overall the source population behaves as expected. The six outliers are all at low flux densities, and four of them agree with a peak brightness to integrated flux density ratio of unity within their uncertainties. The final two agree within  $3\sigma$ ; they could potentially be variable sources. Although the standard and high-resolution images were made from the same data, the standard-resolution image uses more observations, and sources in the observation selected for the high-resolution image could vary slightly from the average. We do not investigate this further here, but simply remove all six sources from the rest of our analysis, although keeping them in would not impact any of the outcomes. The final sample comprises 481 sources identified as AGNs via both peak brightness and total flux density, and a further 459 identified only by total flux density (a total of 940 sources). We refer to these sources as either  $T_{b, \text{peak}}$  or  $T_{b, \text{total}}$  in the text. Note that all sources identified as AGNs via  $T_{b, \text{peak}}$  are also identified as AGNs via  $T_{b, \text{total}}$ ; the sources flagged as  $T_{b, \text{total}}$  are those identified only via  $T_{b, \text{total}}$ .

Using redshift information and assuming a spectral index of  $\alpha = -0.8$ , Fig. 4 shows the integrated radio luminosity (calculated from the standard-resolution catalogue) versus redshift for both the whole high-resolution sample, and subsamples of high- $T_b$  AGNs. This shows that we are identifying AGNs using their brightness temperature across all redshifts, with a wide range of radio powers. This selection method of course relies on a higher flux per beam threshold than allowed by the image sensitivity, and results in a slightly higher effective flux limit than the flux limit of the survey.



**Figure 2.** Observed properties of the high- $T_b$  selected sample. On the left, we show the peak brightness per solid angle versus total flux density per solid angle, for all sources in the high-resolution image (grey circles). Sources identified as AGNs via  $T_{b, \text{total}}$  are marked with pink stars, with sources also identified via  $T_{b, \text{peak}}$  marked with magenta stars. A line of unity is drawn to guide the eye. The high- $T_b$  sources identified by their peak brightness lie close to this line, indicating that they are unresolved. The high- $T_b$  sources identified only by their total flux density predictably have more points which lie further from this line. The right-hand panel shows where these sources lie in terms of their total flux density measured from the standard resolution image (the  $x$ -axis), with the ratio of total flux density from the standard and high-resolution images shown as the dependent variable. The dashed line at unity indicates where the total flux density in the standard resolution image is the same as measured in the high-resolution image.



**Figure 3.** The peak brightness measured from the high-resolution image, compared to the integrated flux density from the standard-resolution image. The solid black line shows where these two quantities are equal. The grey circles show the entire high-resolution sample, with the pink stars marking  $T_{b, \text{total}}$  selected sources, and the magenta stars marking  $T_{b, \text{peak}}$  selected sources.

## 4 RESULTS

### 4.1 Source classifications

Thus far we have only used the flux density and measured sizes to identify sources as AGNs. We can check these identifications against those provided as part of the LOFAR Deep Fields data release, based

**Figure 4.** Radio luminosity versus redshift for the whole sample. Error bars are not shown, because the median uncertainty on the radio power is smaller than the point size. The grey circles show the entire high-resolution sample, with the pink stars marking  $T_{b, \text{total}}$  selected sources, and the magenta stars marking  $T_{b, \text{peak}}$  selected sources.

on both photometric AGN identification (Kondapally et al. 2021) and identification from detailed SED fitting (Best et al., in preparation).

Optically identified AGNs come from the Million Quasar Catalogue (Flesch 2019) compilation, IR identified AGNs satisfy the Donley et al. (2012) criteria, and X-ray identified AGNs have an X-ray counterpart. In Best et al. (in preparation) AGNs are classified using two criteria, which when combined provide three AGN classes.

**Table 1.** Source classification. The final column shows the number of unique AGN identifications from  $T_b$  selection.

Identified with peak and total $T_b$				
SED class	#	Opt. AGNs	IR AGNs	Unique $T_b$ AGNs
SFG	45	0	6	39
Unclass	19	1	4	15
RQAGN	31	11	21	0
LERG	292	0	8	0
HERG	94	30	44	0
<b>Total</b>	<b>481</b>	<b>42</b>	<b>83</b>	<b>54</b>
Identified with total $T_b$ only				
SED class	#	Opt. AGNs	IR AGNs	Unique $T_b$ AGNs
SFG	99	0	5	94
Unclass	14	1	2	12
RQAGN	47	17	35	0
LERG	241	0	2	0
HERG	58	8	28	0
<b>Total</b>	<b>459</b>	<b>26</b>	<b>72</b>	<b>106</b>

The first criteria evaluates whether or not there is significant radiative output from an AGN; this classification is derived from the results of the four different SED-fitting codes, in particular through comparing the results of codes which include AGNs in the fitting against those which do not. The second criteria evaluates the presence of a radio excess based on the distance from the  $L_R$ –SFR relation. Best et al. (in preparation) found a ridge line value for this relation of  $\log_{10}(L_{144} [\text{W Hz}^{-1}]) = 22.24 + 1.08 \times \log_{10}(\text{SFR} [\text{M}_{\odot} \text{yr}^{-1}])$ , and define galaxies to have a radio excess when they lie 0.7 dex (equivalent to  $3\sigma$ ) above this relation. The combination of these two criteria yields three classes of AGNs:

- (i) High-excitation radio galaxies (HERGs) are radiative AGNs with a radio excess;
- (ii) Low-excitation radio galaxies (LERGs) are non-radiative AGNs with a radio excess;
- (iii) Radio-quiet AGNs (RQAGN) are radiative AGNs without a radio excess.

If a classification in either category of criteria is unclear or ambiguous, the AGN identification is *Unclassified*. Sources which do not fall into any of the above AGN classes are classified as SFGs.

Table 1 shows the classifications of the high- $T_b$  identified AGNs,<sup>2</sup> split by whether the sources were identified via  $T_{b, \text{peak}}$  and  $T_{b, \text{total}}$  or  $T_{b, \text{total}}$  only. For the  $T_{b, \text{peak}}$  sample, 89 percent (427/481) also have AGN identifications from the multiwavelength data, while this drops to 77 percent (353/459) for the  $T_{b, \text{total}}$  sample (and is 83 percent overall considering both  $T_{b, \text{peak}}$  and  $T_{b, \text{total}}$ ). This shows that our  $T_b$  criterion reliably selects sources which have AGN identifications from other methods. The SED fitting may yield the wrong classification in some cases, as there is a lot of scatter around the classification lines. The classifications are also weighted by what is the bulk source of the ultraviolet through infrared emission, and in composite systems sources may be classified as an SFG when a low-luminosity AGN is present. Combining the SED and photometric AGN classifications, we also report the number of unique AGN identifications from  $T_b$ . There are a total of 160 new AGN identifications.

<sup>2</sup>No AGNs were classified via X-ray luminosity, and this criterion is not included.

## 4.2 Redshift distributions

Only 191 ( $\sim 9$  percent) of the sources have spectroscopic redshifts. If photometric redshifts are inaccurate then the limiting  $T_b$  value for AGN identification can also be inaccurate, although using the redshift rather than assuming  $z = 0$  when calculating the limiting value of  $T_b$  only adds 27 sources (3 percent) to the sample, so this is not expected to be a large effect. Duncan et al. (2021) shows that the photometric redshifts are reasonably accurate out to  $z \sim 4$ , with  $\sigma_{\text{NMAD}} = 0.017$  and 0.077 for galaxy/host-dominated and AGNs, respectively. Fig. 5 shows that some of the Unclassified, SFG, and RQAGN sources have primary redshift solutions beyond this, but out of the 2214 sources which had SED fitting performed this is a small number, only 112 ( $\sim 5$  percent). What is clear is that the Unclassified and SFG sources may be misidentified or not identified because their redshift is insecure; there is a dearth of  $z_{\text{spec}} > 0.5$  sources in these classes, which is not true for the RQAGN class. The long tails of photometric redshifts for the Unclassified and SFG categories may change if followed up to acquire spectroscopic redshifts, which is an interesting prospect for a future study.

## 4.3 Infrared colour–colour plots

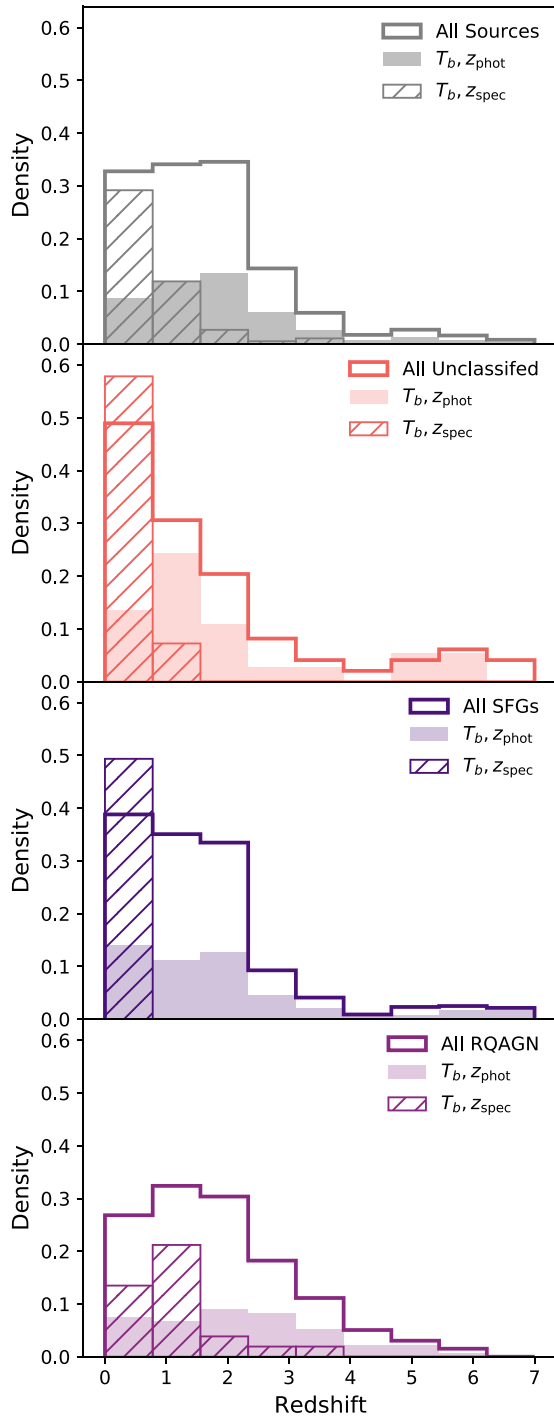
We construct the infrared colour–colour plots using the multiwavelength catalogue from Kondapally et al. (2021), for sources with  $>2\sigma$  detections in all four IRAC channels (3.6, 4.5, 5.8, and 8.0  $\mu\text{m}$ ). Fig. 6 shows where they lie in relation to the wedges from Donley et al. (2012) and Lacy et al. (2004) overplotted. Unsurprisingly, given that the mid-IR region of the SED is an important driver of the classifications of Best et al. (in preparation), the RQAGN fall mostly in the wedges which identify them as radiative AGNs, as do the HERGs. The LERG population extends down to the lower left of the colour–colour plot. The bulk of the SFG population seems to hover around the edge of the wedges, suggesting that perhaps they are composite SFG/AGN sources. The Unclassified category splits into two groups: those which fall securely in the AGN wedges, and those which lie below the wedges. It is worth keeping in mind that the position of sources on these plots will vary with redshift (Radcliffe et al. 2021b).

## 4.4 The $L_R$ –SFR relation

We can also look at the source classifications in terms of the  $L_R$ –SFR relation, in Fig. 7. Although the sources unclassified by the SED fitting do not have estimated SFR values, we place them on the plot to the right-hand side to show their  $L_R$  values. It is clear that in general the high- $T_b$  identified sources run the range of SFR and  $L_R$  values. For non-radio excess sources, there is avoidance of the ridge line from Best et al. (in preparation): 87 percent of the high- $T_b$  SFGs (125/144) and 91 percent of the high- $T_b$  RQAGN (71/78) are above the ridge line of the  $L_R$ –SFR relation (but, by definition, not 0.7 dex above it, which would classify them as radio-excess sources). This may suggest the presence of AGN-related radio emission, even if not formally defined as radio excess sources.

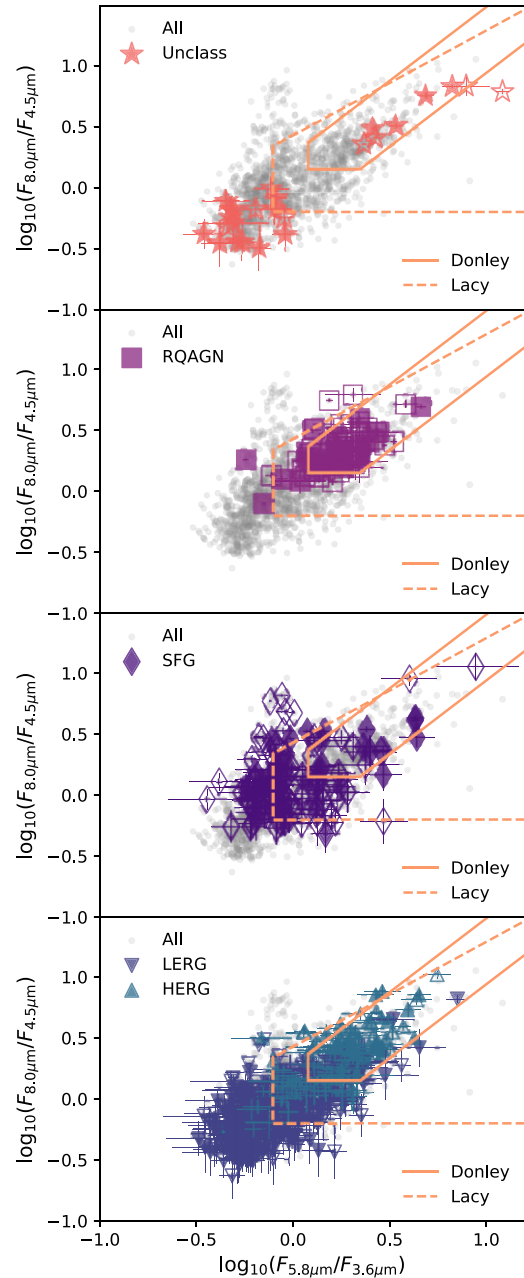
## 4.5 Detectability and source fractions

In the standard-resolution image, there are 23 734 sources with  $\text{SNR} > 5$  which are within the coverage of the high-resolution image. Not all of these will be detectable at high resolution, as they may be extended low-surface brightness sources. To determine the number of sources which are detectable, we compare the peak brightness from



**Figure 5.** The overall redshift distribution for all sources in a class (all, unclassified, SFG, or RQAGN, from top to bottom panel) is plotted as a solid line. The hatched and shaded regions show the spectroscopic and photometric redshift distributions, respectively, for sources which are identified as AGNs via  $T_b$ . The  $T_b$  distributions are normalized to the overall density of all sources in each subclass: e.g. if the hatched and shaded regions are added together the area under the resulting distribution would be unity.

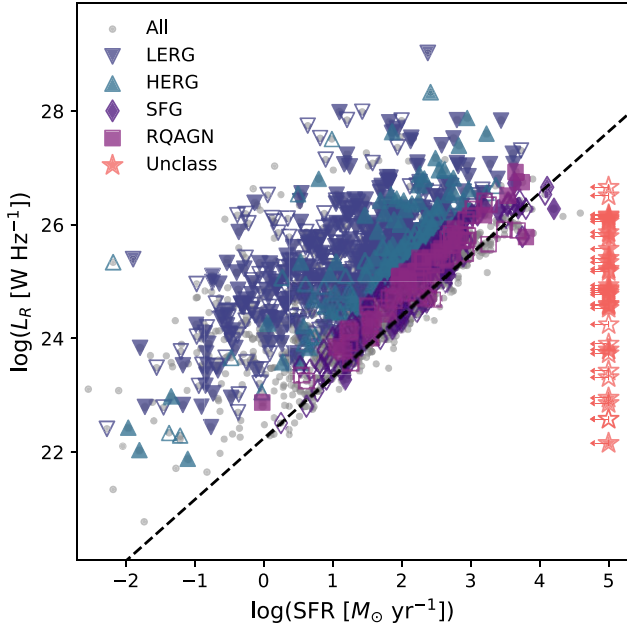
the standard-resolution image to the local rms in the high-resolution image. There are 13 439 sources for which the peak brightness is larger than five times the rms noise at that location in the rms map, which are potentially detectable. Only 16 percent (2214) of these sources are detected in the high-resolution image, which is similar to



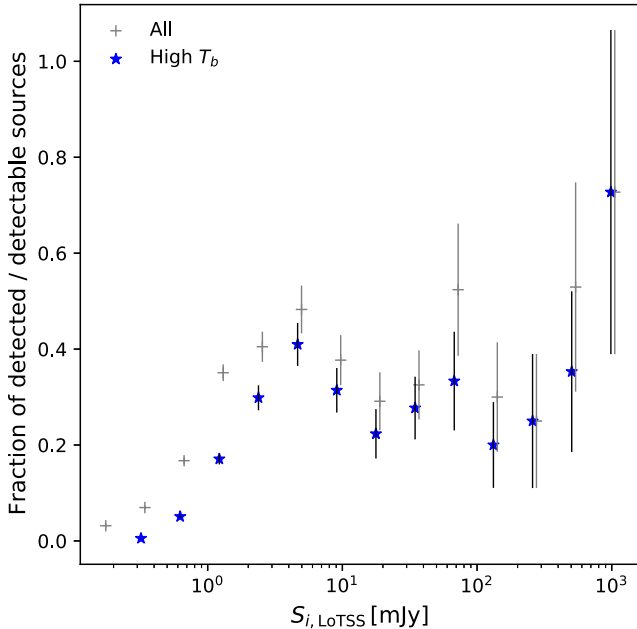
**Figure 6.** Infrared colour–colour plots. The grey points show all sources detected in the high-resolution image, while the coloured points show the  $T_b$ -identified AGNs, using the SED-fitting classifications. The filled points are identified via  $T_{b, \text{peak}}$ , while the unfilled points are identified only via  $T_{b, \text{total}}$ . Note that only sources with  $>2\sigma$  detections in all four bands are plotted. Wedges which are often used to identify AGNs from their infrared colours are shown in orange.

what is found in studies carried out at higher radio frequencies. High-resolution radio follow-up with VLBI arrays of deep radio samples have shown a detection rate of 20–30 percent, depending on the sample (e.g. Herrera Ruiz et al. 2017; Radcliffe et al. 2018). Fig. 8 shows the fraction of detected to detectable sources as a function of integrated flux density from the standard-resolution image. The high-resolution sample is of course biased towards compact sources, as sources with low surface brightness are less likely to be detected in the high-resolution image. While the high- $T_b$  sources (we remind





**Figure 7.** The  $L_R$ –SFR relation. The line above which sources are classified as radio excess is not explicitly drawn, but clear from the lower limit of where LERGs/HERGs fall on the plot.  $T_b$ -identified AGNs run the full range of star formation rates and luminosities. The unclassified sources do not have estimated SFR values, and are shown on the right-hand side of the plot to illustrate their range of  $L_R$ . The ridge line from Best et al. (in preparation) described in Section 5 is shown as a dashed black line.



**Figure 8.** Fraction of detected sources (of those which are detectable) as a function of  $S_{i, \text{LotSS}}$ . The grey symbols show this fraction for all sources, while the blue stars show this for the  $T_b$ -selected sources. The distribution increases up to  $\sim 5$  mJy, then remains roughly constant (within uncertainties) at higher flux densities. This could correspond to the switch between the SFG- and AGN-dominated populations. Uncertainties are estimated assuming  $\sqrt{N}$  for  $N$  sources in a bin, and propagated in the standard manner.

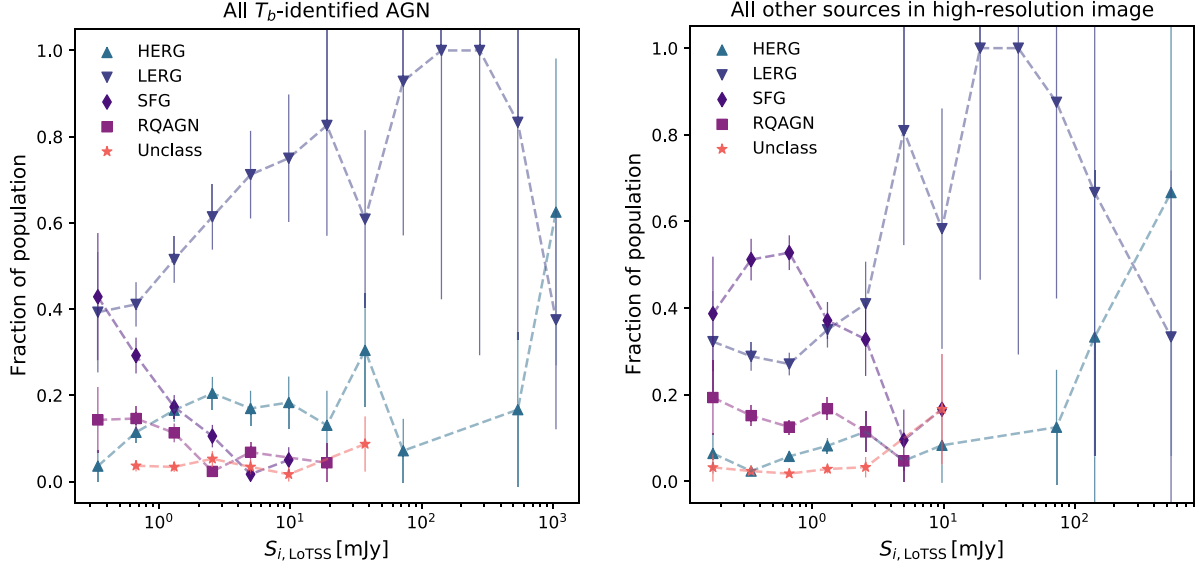
the reader that this includes all  $T_{b, \text{peak}}$  and  $T_{b, \text{total}}$  sources) do not differ from the overall population of detected sources (within the uncertainties) above  $\sim 5$  mJy, there is a clear separation of the high- $T_b$  sources below this. Sources with higher flux densities are more likely to be radio-loud, and therefore AGN dominated, whereas we expect star formation to dominate at fainter flux densities (e.g. Ibar et al. 2009; Ocraan et al. 2017; Prandoni et al. 2018).

Finally, we investigate the nature of the high- $T_b$  sample, and compare it to the non- $T_b$  selected sources in the high-resolution image (for comparison with the full sample detected in the standard-resolution image, see fig. 9 of Best et al. in preparation). For each sample,  $T_b$ -identified AGNs and non- $T_b$  selected sources, we split them into their subpopulations as a function of flux density. As above, we use the total flux density measured from the standard-resolution image. We calculate the number of high- $T_b$  sources in each flux density bin, and find the fraction of each subpopulation. The results are shown in the left ( $T_b$ -identified AGNs) and right (non- $T_b$  selected sources) panels of Fig. 9. It is clear that the LERGs dominate the both samples, particularly at higher flux densities, with the HERG contribution increasing at the highest flux densities. This is the same behaviour seen in Best et al. (in preparation) (their fig. 9). In the  $T_b$ -identified AGN sample, the SFG population is different: in Best et al. (in preparation) this dominates below  $\sim 1$  mJy, whereas here the SFG population fraction remains low for most of the flux density range, only rising to 0.4 in the lowest flux density bin. This is because the SFGs should all be excluded from the current sample by the high- $T_b$  identification. We see that the SFGs are contributing the most to the non- $T_b$  selected sources below flux densities of  $\sim 2$  mJy, which is consistent with Best et al. (in preparation).

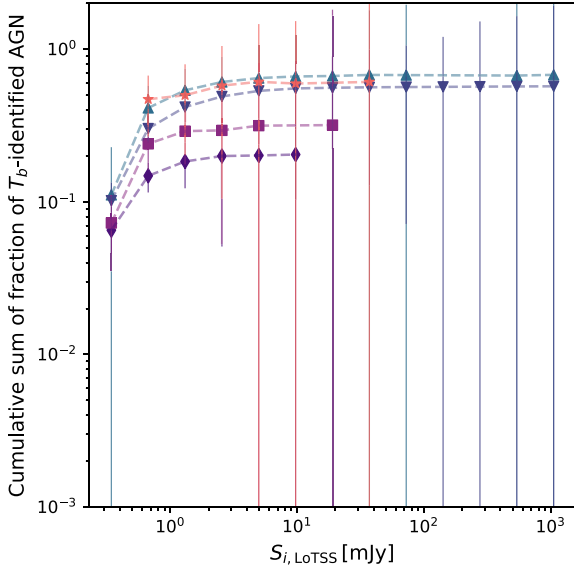
Fig. 10 shows the cumulative sum of the fraction of high- $T_b$  sources in the respective subpopulation. The approximate completeness of the  $T_b$  identification is apparent: the cumulative sum of the fractions asymptote above  $\sim 1$  mJy. The overall fraction of high- $T_b$  sources in each overall subpopulations is: HERGs – 68 percent; LERGs – 57 percent; Unclassified – 61 percent; RQAGN – 32 percent; SFG – 20 percent. The HERG, LERG, and unclassified populations have the largest fractions of high- $T_b$  sources, which is a further indication that the unclassified sources are indeed AGNs. The RQAGN and SFG populations follow with smaller fractions of their subpopulations comprising high- $T_b$  sources. For the RQAGN, this implies that the population may not be dominated by one type of radio emission mechanism, but instead may comprise sources with different types of radio emission (e.g. jets, winds, star formation; see Panessa et al. 2019). It is also possible that some sources identified as SFGs are actually RQAGN, which would increase the number of  $T_b$ -identified RQAGN here (although this number is expected to be a lower limit, see Section 3.2), leaving a very small fraction of genuine SFGs potentially hosting radio jets. Either way, it seems clear that high- $T_b$  AGN identifications are preferentially associated with radiative AGNs (RQAGN) than SFGs.

## 5 SEPARATING STAR FORMATION AND AGN ACTIVITY

We turn to separating radio emission from star formation and AGN activity, which relies on understanding how star formation contributes to the brightness temperature, as described in Section 3.2. We do this only for  $T_b$ -identified sources which are unresolved in the high-resolution image and do not have a radio excess, i.e. the SFG, RQAGN, and Unclass populations. This avoids sources which may have extended or resolved emission due to AGN activity. We determine the unresolved population in the same way as Shimwell



**Figure 9.** Here, we plot the composition of the  $T_b$ -identified AGNs (left-hand panel) and all other sources (right-hand panel) as a function of integrated flux density measured from the standard-resolution image. In each flux density bin, the fraction of each subpopulation of the total number of sources in the bin is plotted.



**Figure 10.** The cumulative sum of the fraction of  $T_b$ -identified AGNs in each subpopulation, as a function of integrated flux density as measured from the standard-resolution image. Symbols and lines are the same as in Fig. 9. The lines flatten out at the total fraction of each subpopulation which are  $T_b$ -identified AGNs.

et al. (2019, 2022), by fitting an upper envelope at the 99.9th percentile to the sources fit with a single Gaussian component (code ‘ $S'$ ’). This was done using the high-resolution image. A total of 53  $T_b$ -identified sources (31  $T_{b, \text{peak}}$ , 22  $T_{b, \text{total}}$ ) were determined to be resolved and therefore removed.

We make the assumption that star formation is widespread in a galaxy, producing a low surface brightness component that will not be detected in the high-resolution image, while the AGN component, identified via brightness temperature, will be a single compact element. Thus, the peak brightness in the high-resolution image will represent the AGN luminosity, although this may be underestimated

because of smearing (see Section 3.2). We use the peak brightness and not the total flux density to avoid contamination from potential low-surface brightness contributions (i.e. star formation). We use the standard-resolution image to estimate the contribution from star formation. Unless very nearby, galaxies will be unresolved in the standard-resolution image, and low-surface brightness radio emission from star formation will be above the limit for detection. Forward-modelling the problem using 3D models of galaxies that are ‘observed’ in the same way as these data will, in a future paper, allow us to place more secure constraints on the separation of star formation and AGN activity.

With these assumptions in hand, we calculate the following AGN and star formation contributions:

$$L_{\text{AGN}} = L_{p, \text{ILT}} \quad (5)$$

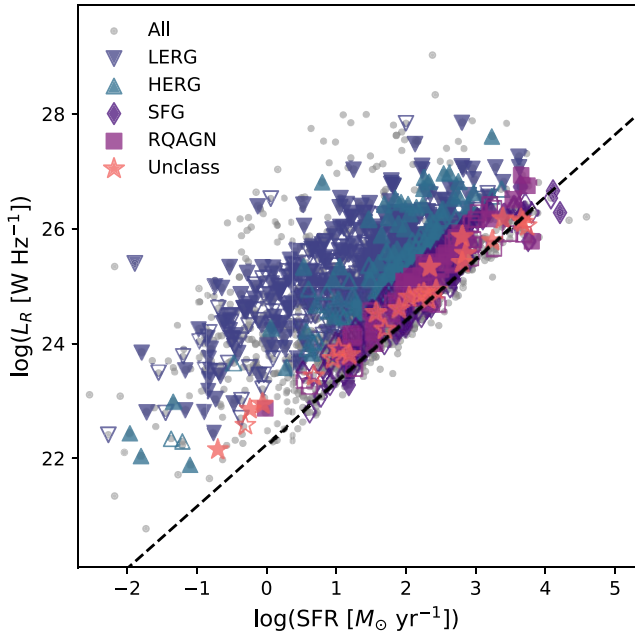
$$L_{\text{SF}} = L_{i, \text{LoTSS}} - L_{p, \text{ILT}}, \quad (6)$$

where  $L_{p, \text{ILT}}$  is the luminosity obtained using the peak intensity in the high-resolution image, and  $L_{i, \text{LoTSS}}$  is the luminosity obtained from the integrated flux density in the standard-resolution image. Errors are propagated in the standard way. The SFR contribution to the radio luminosity can be converted into SFR using the best-fitting parameters for eqn. 2 of Smith et al. (2021):

$$0.9 \times \log_{10}(\text{SFR}) = \log_{10}(L_{\text{SFR}}) - 22.22 - 0.33 \times \log_{10}(M_*), \quad (7)$$

where the units of SFR and  $M_*$  are  $M_{\odot} \text{ yr}^{-1}$  and  $10^{10} M_{\odot}$ , respectively. This equation assumes that there is no low-frequency absorption. This mass-dependent form is similar to the Best et al. (in preparation) ridge line. Practically, we calculate  $S_{\text{AGN}}$  and  $S_{\text{SF}}$  and convert the flux densities into rest-frame luminosities using the distance modulus, assuming a typical synchrotron spectrum of  $\alpha = -0.8$ .

Using the total radio luminosity and the measured SFR after subtracting the AGN component, we can now place our unclassified sources in the  $L_R$ -SFR parameter space, see Fig. 11. None of the unclassified sources shows a radio excess based on the SFR, although they lie on the upper half of the  $L_R$ -SFR relation, with 73 percent (22/30) of sources lying above the ridge line.



**Figure 11.** Same as Fig. 7 for all sources but the unclassified category. For these, we have only been able to place them on the plot after estimating their star formation rate by subtracting the  $L_{\text{AGN}}$  contribution to the total radio luminosity.

We investigate the relative contributions of the AGNs to the total radio luminosity, in Fig. 12. The top panel of this figure shows the ratio of  $L_{\text{AGN}}/L_{\text{SF}}$ , while the bottom panel shows  $L_{\text{AGN}}/L_{\text{total}}$ , both as a function of total radio luminosity. The three populations are well mixed, with no one population having significantly higher or lower fractions than the others. We report the median values in Table 2. We use the median (with the median absolute deviation for the uncertainties) rather than the mean as it is more robust against outliers. The values generally agree with each other: when  $L_{\text{AGN}}/L_{\text{total}}$  is  $\sim 0.5$ ,  $L_{\text{AGN}}/L_{\text{SF}} \sim 1$ , as expected. Although the ratio of  $L_{\text{AGN}}/L_{\text{total}}$  is slightly lower in the SFG category, all three types of sources have medians which agree within the uncertainties. This is consistent with what we would expect: an ILT detection means there must be a compact component, and high values of  $T_b$  mean there is an AGN component, so  $L_{\text{AGN}}/L_{\text{SF}}$  will always be above zero. Additionally, we have limited this part of the analysis to sources which do not have a radio excess, so the total emission will always be less than approximately five times that expected from SFR, meaning that the AGNs will not account for more than  $\sim 80$  percent of the total radio emission.

We also fit a linear model to  $L_{\text{AGN}}/L_{\text{total}} = m \times \log_{10}(L_R) + b$  to check for any dependence between radio and AGN luminosity. The fit results are reported in Table 3, which shows at most only very mild dependence, with relatively large uncertainties. This may be impacted by the effective flux limit of the  $T_b$  selection (see Fig. 4), imperfect separation between SFR and AGN radio emission, or due to the fact that we are probing a regime where neither AGN nor SFR radio luminosity dominates the total radio luminosity. Future work to more precisely quantify the uncertainties on the AGN luminosity (using forward modelling) will help us understand the picture here.

The SED fitting also calculates an AGN fraction, which we compare with  $L_{\text{AGN}}/L_{\text{total}}$  in Fig. 13. As there are multiple SED fits to find consensus values, we show the  $f_{\text{AGN}}$  for AGNFITTER and CIGALE, the latter with both the SKIRTOR (Stalevski et al. 2012, 2016) and

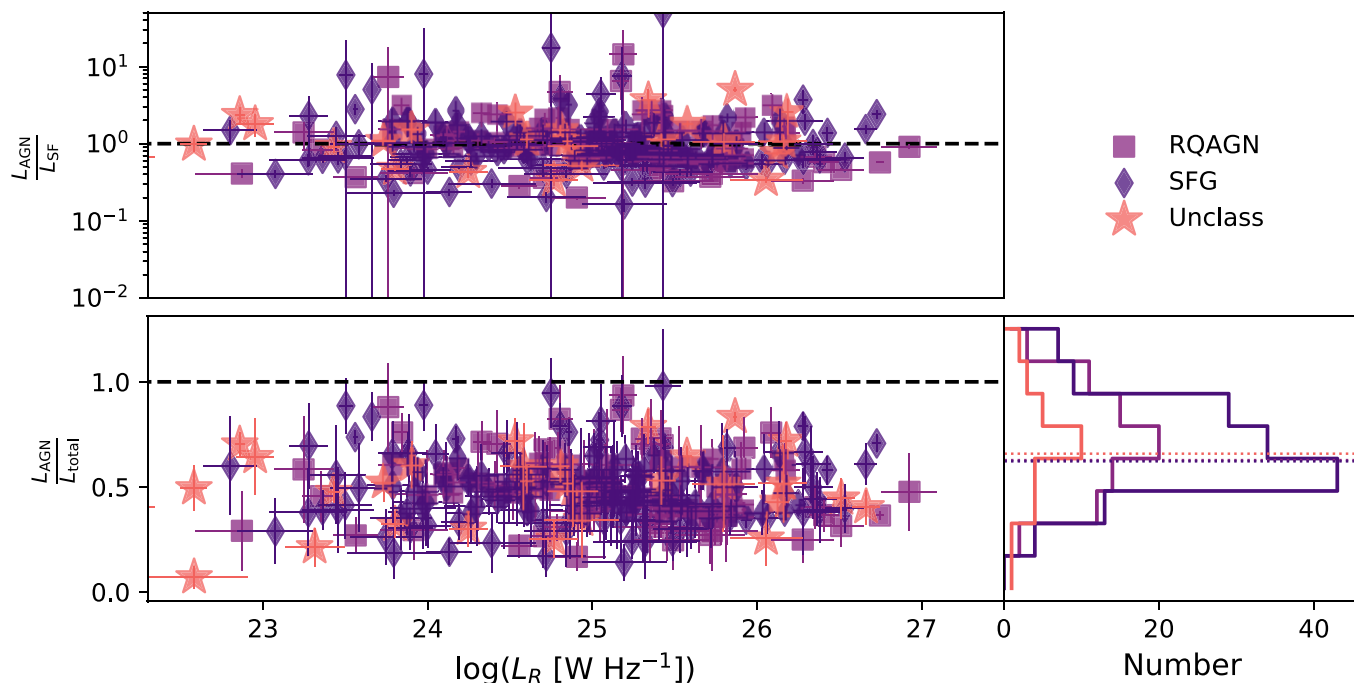
Fritz, Franceschini & Hatziminaoglou (2006) AGN models. It is clear that there is a wide range of  $f_{\text{AGN}}$  values regardless of which SED fitting method is used. This is partly dependent on the fact that  $f_{\text{AGN}}$  is defined differently for each SED fitting method. For AGNFITTER it is the fraction of emission in the 1–30 micron range from the AGN components. For CIGALE it is the total IR luminosity fraction, which naturally will give lower values than AGNFITTER. For a LERG-type source, where the dominant AGN output is in the radio band, SED fitting can yield low values of  $f_{\text{AGN}}$ ; this is consistent with at least some of the SFGs and Unclassified sources not being identified as LERGs because they are not above the radio excess definition, while the majority of RQAGN still have high  $f_{\text{AGN}}$ . It is clear that we do not (except for an outlier) measure  $L_{\text{AGN}}/L_{\text{total}} \lesssim 0.2$ , while the SED fitting measures values of  $f_{\text{AGN}}$  as small as 0.01; as  $L_{\text{AGN}}/L_{\text{total}}$  sample is biased towards AGNs selected via  $T_b$  in the radio band, this is not surprising. Aside from some outliers, the RQAGN population seems to show better agreement between  $f_{\text{AGN}}$  and  $L_{\text{AGN}}/L_{\text{total}}$  for all three SED fitting methods, which is not totally unexpected as they have been identified as AGNs via SED fitting. For AGNFITTER and CIGALE with the FRITZ models, there seem to be a population of SFGs with higher relative  $f_{\text{AGN}}$  values that correlate with  $L_{\text{AGN}}/L_{\text{total}}$ , although not with a one-to-one relation. Overall, the measured  $L_{\text{AGN}}/L_{\text{total}}$  value is higher than that estimated by the SED fitting code: we are either overestimating the AGN contribution or the SED fitting is underestimating the AGN contribution – however, this must be tempered by the fact that  $f_{\text{AGN}}$  and  $L_{\text{AGN}}/L_{\text{total}}$  are measured differently and may not be like comparisons. The AGN contribution to the radio emission may also be less direct in these non-radio excess sources (e.g. if it comes from secondary effects like shocks from AGN winds).

## 6 DISCUSSION

### 6.1 Caveats

Identification of AGNs using brightness temperature is a widely used method, but this is the first time it has been applied at low observing frequencies. One of the primary things to keep in mind is that the models to calculate the limiting value of brightness temperature (practically, here we use flux density per solid angle) assume only the simplest model of a free–free absorbing medium that is well mixed with the radio synchrotron-emitting plasma, and in which there is a constant ratio of these two components. Other low-frequency cut-off mechanisms can exist such as synchrotron self-absorption and low-energy electron cut-off. Additionally, the geometry of any free–free absorbing gas could be different from the assumptions of the Condon model, in particular it could be clumpy by being confined to H II regions (Lacki 2013; Ramírez-Olivencia et al. 2022) or could be foreground to the synchrotron emission (Varenius et al. 2015; Conway, Elitzur & Parra 2018). Such different geometries affect the predicted radio spectrum from a star-formation powered source and hence the maximum brightness temperature expected at a given frequency.

In addition to absorption with a covering factor less than unity across the whole source, AGN activity and star formation in galaxies which host both can experience different types of absorption in each. For example, if the compact AGN emission comes from jets which are experiencing synchrotron self-absorption, radio emission from star formation in the galaxy would be unaffected. In the case of free–free absorption of radio emission from star formation due to ionizing gas along the line of sight, whether or not it impacts the compact AGN emission depends on the geometry of the system



**Figure 12.** The top panel shows the ratio of radio luminosity from AGNs to radio luminosity from star formation, as a function of total radio luminosity. The bottom panel shows the ratio of AGNs to total radio luminosity, with the distributions of the AGN fraction for each population in the panel to the right (dotted lines show the medians reported in Table 2). The black dashed line in each main panel is placed at unity to help guide the eye.

**Table 2.** Median values of ratios from Fig. 12.

$L_{\text{AGN}}/L_{\text{SF}}$	
Class	Median
RQAGN	$0.98 \pm 0.67$
SFG	$0.96 \pm 0.55$
Unclass	$1.07 \pm 0.68$
<b>Total</b>	<b><math>0.98 \pm 0.60</math></b>
$L_{\text{AGN}}/L_{\text{total}}$	
Class	Median
RQAGN	$0.49 \pm 0.17$
SFG	$0.49 \pm 0.15$
Unclass	$0.52 \pm 0.17$
<b>Total</b>	<b><math>0.49 \pm 0.16</math></b>

**Table 3.** Fit parameters for  $L_{\text{AGN}}/L_{\text{total}} = m \times \log_{10}(L_R) + b$  from Fig. 12.

Class	$m$	$b$
All	$-0.00 \pm 0.01$	$0.52 \pm 0.29$
RQAGN	$-0.03 \pm 0.02$	$1.15 \pm 0.58$
SFG	$-0.00 \pm 0.02$	$0.57 \pm 0.40$
Unclass	$0.03 \pm 0.03$	$-0.22 \pm 0.63$

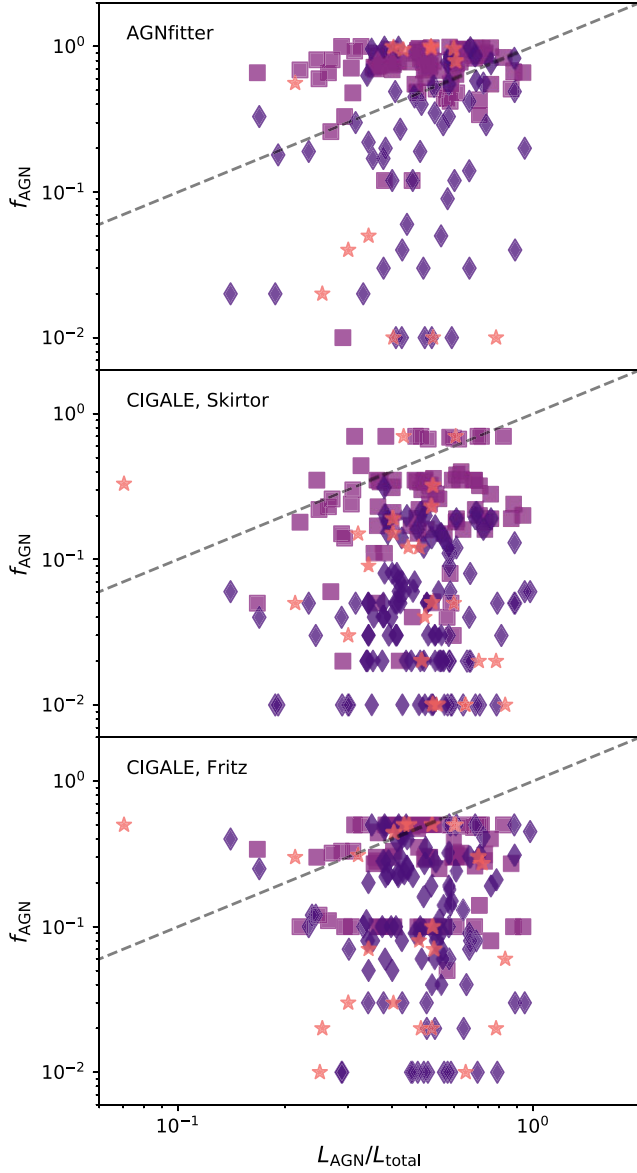
and the covering factor of the ionizing gas. Of course, the covering factor could be unity and provide a constant amount of absorption across the entire source, which means the *relative* contributions, if not the *absolute*, of star formation and AGN activity would remain the same. Regardless of its origin, low-frequency absorption would work to suppress the flux density per solid angle. This means that while our conservative limit ensures we are securely identifying AGNs, many other ILT sources may also be AGNs.

Another issue to consider is the beam size. In high-frequency observations, the spatial scales probed are smaller. For example, Middelberg et al. (2013) achieve a beam size (with natural weighting) of  $11.7 \times 9.4 \text{ mas}^2$ , which is a beam solid angle  $10^3$  smaller than that achieved with the ILT in Sweijen et al. (2022),  $0.3 \times 0.4 \text{ arcsec}^2$  (in principle, this could be pushed to  $\sim 0.2 \text{ arcsec}$ , at the cost of increased noise). So are we actually probing the relevant spatial scales? We are secure in our identifications of AGNs, but what process is actually producing the radio emission related to the AGNs could be on pc to kpc scales. Although we explore this in the next subsection, we will likely require ancillary information to help determine what the source of radio emission is in these  $T_b$ -identified AGNs to make further progress.

## 6.2 What is driving the radio emission?

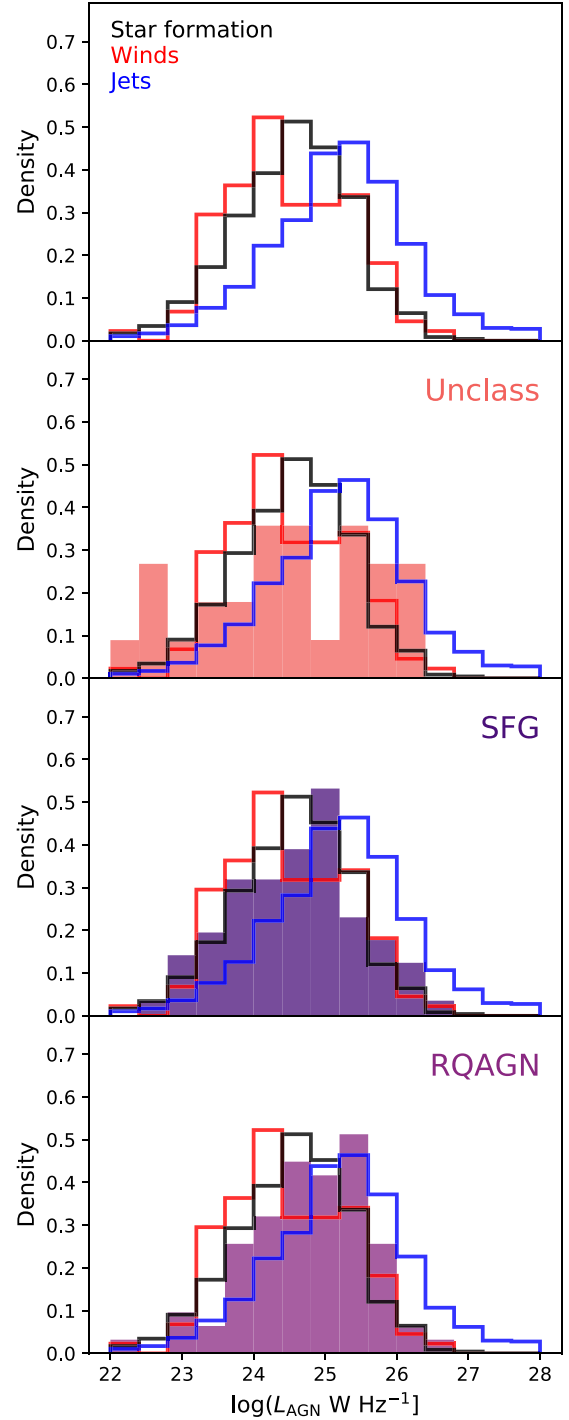
For HERGs and LERGs, the dominant radio emission mechanism is jets. But in radio-quiet AGNs, we see that only 32 percent of the population are identified via their  $T_b$ . In these sources, we can separate the AGN and star formation luminosities (see Section 5) and try to investigate the origin of the radio luminosity due to AGNs. We consider three possibilities: star formation, winds, and jets. Two of these possibilities attribute the radio emission to AGNs, and we include star formation as a comparison. The caveats in the previous section mean that our separation of radio emission from SF and AGN activity is not yet robust, and it is also possible that there is enhanced, compact star formation due to the AGNs. For each of these three possibilities, we construct a distribution of radio luminosities and compare this with the distribution of radio AGN luminosities in the  $T_b$ -identified samples of unclassified, SFG, and RQAGN sources.

For the star-formation distribution, we select all sources in the SFG class from Best et al. (in preparation) and remove the  $T_b$ -identified



**Figure 13.** Comparison of the measured  $L_{\text{AGN}}/L_{\text{total}}$  with the SED-estimated  $f_{\text{AGN}}$  values. The dashed line in each plot shows where  $L_{\text{AGN}}/L_{\text{total}} = f_{\text{AGN}}$ . Note that only sources which have a value of  $f_{\text{AGN}}$  reported are shown. Symbols are the same as in Fig. 12.

AGNs. The radio luminosity in these sources is therefore dominated by star formation, and we simply use the integrated radio luminosity calculated from the standard-resolution image. Similarly for jets, we select all HERGs and LERGs and use their integrated radio luminosity. For winds, we rely on the models in Nims, Quataert & Faucher-Giguère (2015), which provide a prescription for estimating the synchrotron spectrum from the shocked ambient medium of an AGN wind. We use their Eqn. 32 to estimate the radio luminosity at 144 MHz from this shocked ambient medium. We assume the same fiducial model as Nims et al. (2015), but we require a realistic distribution of bolometric luminosities. These we take from Woo & Urry (2002), who compiled values from the literature and estimated values for another  $\sim 200$  galaxies. From this sample we select only the lower luminosity Seyfert galaxies and radio-quiet AGNs and use this distribution of bolometric luminosities to estimate the wind contribution to the radio luminosity.



**Figure 14.** The top panel shows distributions for three possible contributions to AGN radio luminosity in radio-quiet AGNs: star formation, winds, and jets. The following three panels show these distributions overlaid on the distributions of AGN luminosity for the  $T_b$ -identified AGNs from the unclassified, SFG, and RQAGN samples.

Fig. 14 shows these three representative distributions in the top panel. The distributions of winds and star formation cover similar ranges, with only slight differences in the peaks of their distribution. The jet distribution is shifted towards higher radio luminosities, as expected. All three distributions are then overlaid on the distributions

of AGN luminosity for the unclassified, SFG, and RQAGN samples, for a qualitative comparison.

Both the SFG and RQAGN samples have distributions of  $L_{\text{AGN}}$  (calculated from the peak intensity in the high-resolution image) which fall within the distributions of all three potential sources of radio emission, while the unclassified sample has a slight overabundance at the very low  $L_{\text{AGN}}$  end, and also has peaks which seem to be consistent with the distributions of  $L_{\text{AGN}}$  from star-formation and jets, confirming their composite nature. It is possible that the unclassified sources have photometric redshifts very different from the actual redshifts, resulting in the inability of the SED fitting to properly classify them and yielding incorrect estimates of the AGN luminosity. Follow-up spectroscopic observations to secure redshifts for these sources will be critical for helping understand their nature.

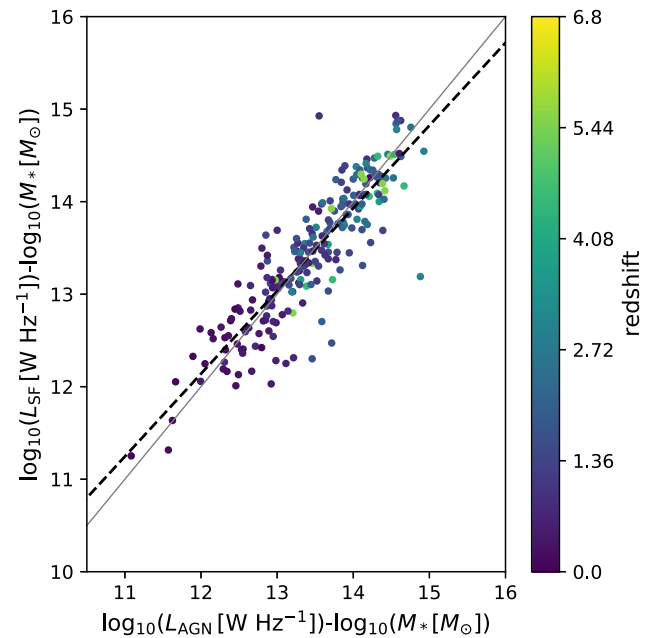
The SFG sample appears to have a single peak in its distribution at  $L_{\text{AGN}} \sim 10^{25} \text{ W Hz}^{-1}$ . We have seen that the SFG population has infrared colours on the borderline between AGNs and SFGs, and the population has median  $L_{\text{AGN}}/L_{\text{total}} = 0.44 \pm 0.14$ . The high- $T_b$  cores could be due truly to AGN activity, or intense starbursts in the nuclear region. However, Sweijen et al. (in preparation) examines a sample of hyper-luminous infrared galaxies (HyLIRGs) and finds that their brightness temperatures imply that 98 percent of the sample detected in the high-resolution image are likely to host radio AGNs. The fact that the SFG population appears to have a single Gaussian distribution of  $L_{\text{AGN}}$  implies that the radio emission mechanism is similar across the population, and the fact that the distribution is shifted slightly more towards the star formation and winds distributions could mean that it is less likely these galaxies host jets. Spatially resolved IFU data may be needed to confirm the presence of AGNs.

The RQAGN sample looks like it could be two overlapping distributions, one peaking around  $L_{\text{AGN}} \sim 10^{24.5} \text{ W Hz}^{-1}$ , and one around  $\sim 10^{25.5} \text{ W Hz}^{-1}$ . It is likely the population hosts a mix of different radio emission mechanisms, which is also supported by the fact that only 32 percent of the overall RQAGN population are identified as AGNs via their  $T_b$  values. There seems to be a significant fraction of the sample which is shifted towards higher  $L_{\text{AGN}}$ , so it is likely at least some of these host radio jets. However, the lower peak could be due to either winds or extremely compact starbursts. Nims et al. (2015) make predictions for the spectral index expected for radio emission from the shocked ambient medium due to winds, and future observations providing spectral information on the same spatial scales, e.g. with e-MERLIN, will be interesting to test this.

### 6.3 Co-evolution

Understanding the co-evolution of supermassive black holes and their host galaxies is linked to understanding the relative contributions of AGN activity and star formation to a galaxy's overall energy budget. Separating the radio emission from AGNs and star formation as we did in Section 5 affords us the opportunity to see how they are linked, although correlation does not mean causation. Fig. 15 shows the luminosity from star formation versus the luminosity from AGNs for unresolved, non-radio excess,  $T_b$ -identified AGNs. We normalize by the stellar mass to remove any redshift bias. A relation is clear, with a linear fit yielding:  $\log_{10}(L_{\text{SFR}}/M_*) = (0.89 \pm 0.03) \times \log_{10}(L_{\text{AGN}}/M_*) + (1.41 \pm 0.43)$ . The slope of the linear fit is slightly offset from unity.

Colouring the points by their redshift reveals that this relationship may still have some dependence on redshift. When restricting the linear fit to include only the 29 sources in the sample which have spectroscopic redshifts, the fit changes to:  $\log_{10}(L_{\text{SFR}}/M_*) =$



**Figure 15.** The stellar-mass normalized  $L_{\text{SF}}$  versus the stellar-mass normalized  $L_{\text{AGN}}$  for unresolved, non-radio excess,  $T_b$ -identified AGNs. The points are coloured by their redshift. The grey dashed line is a fit to the data, while the thin grey line shows the one-to-one line.

$(0.92 \pm 0.06) \times \log_{10}(L_{\text{AGN}}/M_*) + (1.19 \pm 0.83)$ . While the slope becomes slightly steeper, it is still consistent with the overall fit within the uncertainty on the parameter.

Overall, the tight correlation between the stellar-mass normalized radio luminosities due to star formation and AGNs supports the idea that supermassive black holes co-evolve with their galaxies. The caveats outlined previously in this section mean the exact form of this relationship should be approached with caution, and this issue will be explored further in future studies.

## 7 CONCLUSIONS AND FUTURE WORK

In this paper, we have demonstrated, for the first time, the use of brightness temperature measurements at low (144 MHz) radio frequencies to identify 940 AGNs. This is only possible due to the sub-arcsecond resolution of the ILT, which is critical for reaching the limiting values of brightness temperature generated by an SFG. We find that 89 percent of  $T_{b,\text{peak}}$  and 83 percent of  $T_{b,\text{total}}$  selected sources are also classified as AGNs via photometric or SED classification methods. This demonstrates that in the absence of detailed SED fitting (which provides the bulk of the AGN classifications, compared to photometric methods), we can reliably select AGNs using  $T_b$  at low frequencies. This will be crucial as we post-process the LOFAR Two-metre Sky Survey (LoTSS; Shimwell et al. 2022, 2019) at high resolution, as we will not have the exquisite ancillary data available in, e.g. the Lockman Hole region, for the entire Northern Sky.

Investigating the ancillary data, we find that infrared colour-colour plots show SFGs with a  $T_b$  AGN identification are either in or around the edges of the selection wedges, consistent with them being composite AGN/SFG sources. The majority of RQAGN lie within the selection wedges, while the unclassified sources split into two groups: one firmly within the AGN selection criteria, and one outside of it.

Using both the rms map from the high-resolution image and the peak brightness from the standard-resolution image, there are 13 439 potentially detectable sources of which only 2214 sources are detected.<sup>3</sup> The fraction of detected sources increases up to  $\sim 5$  mJy, then is fairly constant at higher flux densities, while below this the high- $T_b$  population splits from the overall population, indicating that we are detecting a mixture of SFGs and AGNs in the high-resolution image.

Dividing the  $T_b$ -identified AGN sample into subpopulations, we find qualitatively similar behaviour as in Best et al. (in preparation) for the LERG and HERG populations. The high- $T_b$  sources comprise 57 and 68 percent of these populations, respectively. Over half (61 percent) of the unclassified sources are high- $T_b$  sources. The high- $T_b$  RQAGN only make up 32 percent of the overall population, indicating that there is likely a mixture of radio emission mechanisms in the overall RQAGN population. Finally, only 20 percent of SFGs are identified as AGNs via their high- $T_b$ ; when considered in context of their placement on the infrared colour–colour plots these are highly likely to be composite AGN/SFG sources, consistent with the results of Strazzullo et al. (2010) for the Deep Swire Field, or LERGs which fall short of the radio excess cut-off. This is in contrast to a sample of HyLIRGs, in which Sweijen et al. (in preparation) find 98 percent of the sample to be  $T_b$ -identified AGNs.

With the simple assumption that the peak brightness of unresolved sources in the high-resolution map represents the AGN luminosity, we place the unclassified sources on the  $L_R$ –SFR relation and find that none have a radio excess. The ratio of AGNs to total radio luminosity in all non-radio excess, unresolved sources ranges from less than 0.1 to almost unity, with a median value of  $0.49 \pm 0.16$ , which again suggests composite systems. For RQAGN, this fraction is the closest to being correlated with  $f_{\text{AGN}}$  from the SED fitting, but the other populations show weaker correlations. We find only a very weak dependence on  $L_R$ , with large uncertainties.

This first step in using the unique observations from Sweijen et al. (2022) shows the potential for using the ILT as a tool for identifying AGNs based on their radio data alone, although there is more work to be done to understand these samples of AGNs. The new WEAVE multi-object survey spectrograph on the William Herschel Telescope will collect spectroscopic redshifts for the sample outlined here as part of the WEAVE-LOFAR survey (Smith et al. 2016), allowing us to refine the brightness temperature measurements. Complementary radio data on the same spatial scales but at GHz frequencies from an EVN/e-MERLIN project will help us understand their spectral properties (PI: McKean). In a future paper, we will use a forward-modelling approach to better constrain the separation of star-formation and AGN activity. This will allow us to fully exploit these new sub-arcsecond, low-frequency data sets to understand how AGNs co-evolve with their galaxies.

## ACKNOWLEDGEMENTS

The authors acknowledge the contributions of S.R. Ward, who worked on a related project that helped kickstart this one, and J.C.S. Pierce, who helped with early visual identifications of sources. This work was supported by the Medical Research Council [MR/T042842/1]. PNB is grateful for support from the UK STFC via grant ST/V000594/1. KJD acknowledges funding from the European Union’s Horizon 2020 research and innovation programme under the

Marie Skłodowska-Curie grant agreement No. 892117 (HIZRAD). IP acknowledges support from INAF under the SKA/CTA PRIN ‘FORECaST’ and the PRIN MAIN STREAM ‘SAuROS’ projects. WLW acknowledges support from the CAS-NWO programme for radio astronomy with project number 629.001.024, which is financed by the Netherlands Organisation for Scientific Research (NWO). RJvW acknowledges support from the ERC Starting Grant ClusterWeb 804208. This paper is based (in part) on data obtained with the International LOFAR Telescope (ILT) under project code LT10\_012. 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 that are collectively operated by the ILT foundation under a joint scientific policy. The ILT resources have benefited 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 work has used the Dutch national e-infrastructure with the support of SURF Cooperative through grant e-infra 180169.

## DATA AVAILABILITY

The data used in this paper come from Sweijen et al. (2022) and the LoTSS Deep Fields Data Release 1, for which the radio data are presented in Tasse et al. (2021) and made publicly available through both the Centre de Données astronomiques de Strasbourg (CDS) and through the LOFAR Surveys website (<https://lofar-surveys.org/deepfields.html>). Multiwavelength photometric catalogues and photometric redshifts come from Kondapally et al. (2021) and Duncan et al. (2021), respectively, both of which are also available through CDS and the LOFAR Surveys website. The radio data from Sweijen et al. (2022) are also available on the LOFAR Surveys website (<https://lofar-surveys.org/hdfields.html>) and we make additional use of the deconvolved sizes which are available upon request. These data were supplemented with SED fitting results presented in Best et al. (in preparation), which are available on request and will be published with the Best et al. paper, through <https://lofar-surveys.org/deepfields.html>. Supporting information was taken from tables published in Middelberg et al. (2013) and Woo & Urry (2002), and are available in csv format at <https://github.com/lmorabit/AGNdetect>, where the publicly available code for the analysis presented here can also be found.

## REFERENCES

- Best P. N., Longair M. S., Rottgering H. J. A., 1996, *MNRAS*, 280, L9  
 Bîrzan L., Rafferty D. A., McNamara B. R., Wise M. W., Nulsen P. E. J., 2004, *ApJ*, 607, 800  
 Bonzini M. et al., 2015, *MNRAS*, 453, 1079  
 Boquien M., Burgarella D., Roehly Y., Buat V., Ciesla L., Corre D., Inoue A. K., Salas H., 2019, *A&A*, 622, A103  
 Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, *MNRAS*, 370, 645  
 Burgarella D., Buat V., Iglesias-Páramo J., 2005, *MNRAS*, 360, 1413  
 Calistro Rivera G. et al., 2017, *MNRAS*, 469, 3468  
 Calistro Rivera G., Lusso E., Hennawi J. F., Hogg D. W., 2016, *ApJ*, 833, 98  
 Carnall A. C., Leja J., Johnson B. D., McLure R. J., Dunlop J. S., Conroy C., 2019, *ApJ*, 873, 44

<sup>3</sup>Specifically, there are 2214  $5\sigma$  sources after removing duplicates and those with no redshift information.

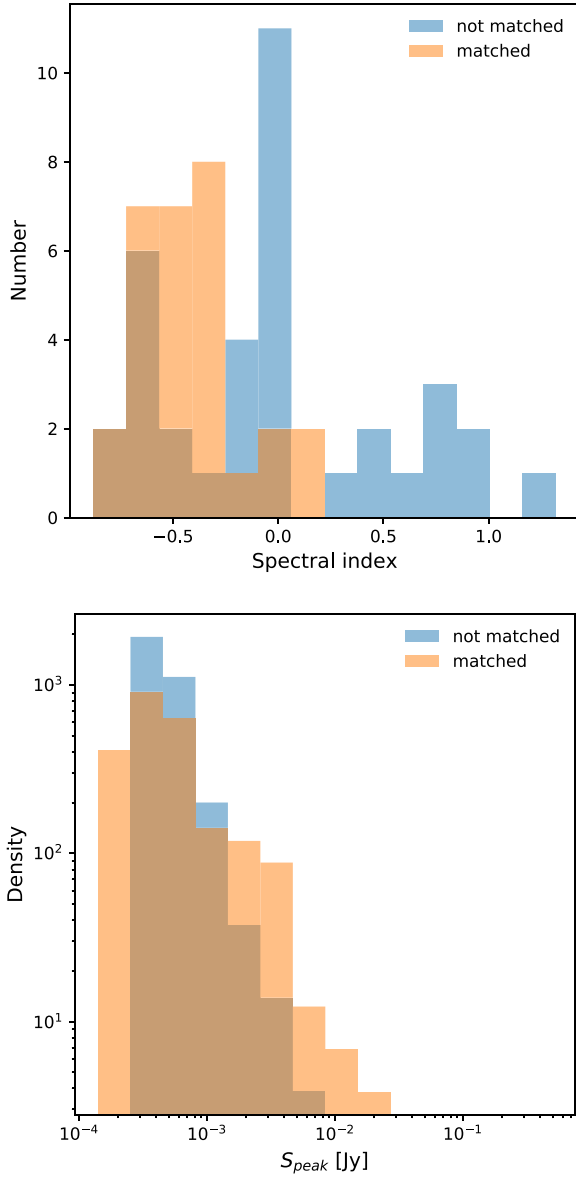
- Carnall A. C., McLure R. J., Dunlop J. S., Davé R., 2018, *MNRAS*, 480, 4379
- Chabrier G., 2003, *PASP*, 115, 763
- Chambers K. C., Miley G. K., van Breugel W., 1987, *Nature*, 329, 604
- Chi S., Barthel P. D., Garrett M. A., 2013, *A&A*, 550, A68
- Condon J. J., 1992, *ARA&A*, 30, 575
- Condon J. J., Kellermann K. I., Kimball A. E., Perley R. A., Ivezić Ž., 2013, *ApJ*, 768, 37
- Conway J. E., Elitzur M., Parra R., 2018, *ApJ*, 865, 70
- Croton D. J. et al., 2006, *MNRAS*, 365, 11
- da Cunha E., Charlot S., Elbaz D., 2008, *MNRAS*, 388, 1595
- De Young D. S., 1981, *Nature*, 293, 43
- Del Moro A. et al., 2013, *A&A*, 549, A59
- Deller A. T. et al., 2011, *PASP*, 123, 275
- Dey A., van Breugel W., Vacca W. D., Antonucci R., 1997, *ApJ*, 490, 698
- Donley J. L. et al., 2012, *ApJ*, 748, 142
- Drake C. L., McGregor P. J., Dopita M. A., van Breugel W. J. M., 2003, *AJ*, 126, 2237
- Duncan K. J. et al., 2021, *A&A*, 648, A4
- Ferrarese L., Merritt D., 2000, *ApJ*, 539, L9
- Flesch E. W., 2019, preprint ([arXiv:1912.05614](https://arxiv.org/abs/1912.05614))
- Fritz J., Franceschini A., Hatziminaoglou E., 2006, *MNRAS*, 366, 767
- Garrett M. A. et al., 2001, *A&A*, 366, L5
- Gebhardt K. et al., 2000, *ApJ*, 539, L13
- Greene J. E., Zakamska N. L., Ho L. C., Barth A. J., 2011, *ApJ*, 732, 9
- Groeneveld C. et al., 2022, *A&A*, 658, A9
- Gürkan G. et al., 2018, *MNRAS*, 475, 3010
- Herrera Ruiz N. et al., 2017, *A&A*, 607, A132
- Hinshaw G. et al., 2013, *ApJS*, 208, 19
- Ibar E., Ivison R. J., Biggs A. D., Lal D. V., Best P. N., Green D. A., 2009, *MNRAS*, 397, 281
- Jackson N. et al., 2022, *A&A*, 658, A2
- Kewley L. J., Heisler C. A., Dopita M. A., Sutherland R., Norris R. P., Reynolds J., Lumsden S., 2000, *ApJ*, 530, 704
- King A., Pounds K., 2015, *ARA&A*, 53, 115
- Kondapally R. et al., 2021, *A&A*, 648, A3
- Kukreti P. et al., 2022, *A&A*, 658, A6
- Lacki B. C., 2013, *MNRAS*, 431, 3003
- Lacy M. et al., 2004, *ApJS*, 154, 166
- Lobanov A. P., 2005, preprint ([arXiv:0503225](https://arxiv.org/abs/0503225))
- Macfarlane C. et al., 2021, *MNRAS*, 506, 5888
- McNamara B. R. et al., 2000, *ApJ*, 534, L135
- Middelberg E. et al., 2013, *A&A*, 551, A97
- Mohan N., Rafferty D., 2015, Astrophysics Source Code Library, record ascl:1502.007
- Morabito L. K. et al., 2016, *MNRAS*, 461, 2676
- Morabito L. K. et al., 2022, *A&A*, 658, A1
- Morgan J. S., Mantovani F., Deller A. T., Brisken W., Alef W., Middelberg E., Nanni M., Tingay S. J., 2011, *A&A*, 526, A140
- Muxlow T. W. B. et al., 2005, *MNRAS*, 358, 1159
- Nesvadba N. P. H., Bicknell G. V., Mukherjee D., Wagner A. Y., 2020, *A&A*, 639, L13
- Nims J., Quataert E., Faucher-Giguère C.-A., 2015, *MNRAS*, 447, 3612
- Noll S., Burgarella D., Giovannoli E., Buat V., Marcellac D., Muñoz-Mateos J. C., 2009, *A&A*, 507, 1793
- Ocran E. F., Taylor A. R., Vaccari M., Green D. A., 2017, *MNRAS*, 468, 1156
- Padovani P., Bonzini M., Kellermann K. I., Miller N., Mainieri V., Tozzi P., 2015, *MNRAS*, 452, 1263
- Panessa F., Baldi R. D., Laor A., Padovani P., Behar E., McHardy I., 2019, *Nat. Astron.*, 3, 387
- Prandoni I., Guglielmino G., Morganti R., Vaccari M., Maini A., R<sup>o</sup>ttinger H. J. A., Jarvis M. J., Garrett M. A., 2018, *MNRAS*, 481, 4548
- Radcliffe J. F. et al., 2018, *A&A*, 619, A48
- Radcliffe J. F., Barthel P. D., Garrett M. A., Beswick R. J., Thomson A. P., Muxlow T. W. B., 2021a, *A&A*, 649, L9
- Radcliffe J. F., Barthel P. D., Thomson A. P., Garrett M. A., Beswick R. J., Muxlow T. W. B., 2021b, *A&A*, 649, A27
- Ramírez-Olivencia N. et al., 2018, *A&A*, 610, L18
- Ramírez-Olivencia N., Varenus E., Pérez-Torres M., Alberdi A., Conway J. E., Alonso-Herrero A., Pereira-Santaella M., Herrero-Illana R., 2022, *A&A*, 658, A4
- Schreiber C. et al., 2021, *A&A*, 646, A68
- Shimwell T. W. et al., 2019, *A&A*, 622, A1
- Shimwell T. W. et al., 2022, *A&A*, 659, A1
- Silk J., 2013, *ApJ*, 772, 112
- Smith D. J. B. et al., 2016, in Reylé c., Richard J., Cambrésy L., Deleuil M., Pécontal E., Tresse L., Vauglin I., eds, Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics. p. 271
- Smith D. J. B. et al., 2021, *A&A*, 648, A6
- Sopp H. M., Alexander P., 1991, *MNRAS*, 251, 14P
- Stalevski M., Fritz J., Baes M., Nakos T., Popović L. C., 2012, *MNRAS*, 420, 2756
- Stalevski M., Ricci C., Ueda Y., Lira P., Fritz J., Baes M., 2016, *MNRAS*, 458, 2288
- Strazzullo V., Pannella M., Owen F. N., Bender R., Morrison G. E., Wang W.-H., Shupe D. L., 2010, *ApJ*, 714, 1305
- Sweijen F. et al., 2022, *Nat. Astron.*, 6, 350
- Tasse C. et al., 2021, *A&A*, 648, A1
- Timmerman R. et al., 2022, *A&A*, 658, A5
- van Haarlem M. P. et al., 2013, *A&A*, 556, A2
- van Weeren R. J. et al., 2016, *ApJS*, 223, 2
- van Weeren R. J. et al., 2021, *A&A*, 651, A115
- Varenus E. et al., 2015, *A&A*, 574, A114
- Ward S. R., 2022, *MNRAS*, 514, 2936
- Webster B. et al., 2021, *MNRAS*, 500, 4921
- White S. V., Jarvis M. J., Häußler B., Maddox N., 2015, *MNRAS*, 448, 2665
- Wilson A. S., 1988, *A&A*, 206, 41
- Wise M. W., McNamara B. R., Nulsen P. E. J., Houck J. C., David L. P., 2007, *ApJ*, 659, 1153
- Woo J.-H., Urry C. M., 2002, *ApJ*, 579, 530
- Zakamska N. L., Greene J. E., 2014, *MNRAS*, 442, 784
- Zhuang M.-Y., Ho L. C., Shangjun J., 2021, *ApJ*, 906, 38

## APPENDIX A: COMPARISON WITH VLBA OBSERVATIONS

We can compare our results with a mosaicked wide-field VLBI study of a portion of the Lockman Hole field (Middelberg et al. 2013, project code BM332). Although this area only covers 0.59 deg<sup>2</sup> with three overlapping pointing centres<sup>4</sup> rather than the 6.6 deg<sup>2</sup> from Sweijen et al. (2022), they do overlap which allows for a direct comparison. The Middelberg et al. (2013) study used the Very Long Baseline Array (VLBA) at 1.3 GHz, targeting 496 sources with integrated flux densities > 100 μJy, as measured with the Very Large Array (VLA). After calibration, the VLA peak brightness was compared to the local rms in the VLBI maps, and only 217 sources were deemed to be potentially detectable. Of these, there were 65 detections with 6σ significance. Note that at this frequency, any detection with the VLBA is considered to satisfy the  $T_b$  criteria for AGN identification. The  $T_b$ -identified AGN sample therefore implies a sky density of 109.96 sources per deg<sup>2</sup>. Within the same region of sky, there are 92  $T_b$ -identified AGNs from the high-resolution image, yielding a sky density of 155.64 sources per deg<sup>2</sup>. These numbers are approximately similar, although a direct comparison is difficult because the two studies are conducted at different frequencies. The spectral indices of the VLBA sample, calculated using VLA and

<sup>4</sup>from A. Deller, private communication:  
 Epochs A/D: 10<sup>h</sup>52<sup>m</sup>56<sup>s</sup>.0 + 57d29<sup>h</sup>06.0<sup>m</sup>  
 Epoch B: 10<sup>h</sup>52<sup>m</sup>08<sup>s</sup>.8 57d21<sup>h</sup>33.8<sup>m</sup>  
 Epoch C: 10<sup>h</sup>51<sup>m</sup>02<sup>s</sup>.0 + 57d13<sup>h</sup>50.4<sup>m</sup>





**Figure A1.** *Top:* Spectral index distributions of the VLBA sources which are matched (orange) and not matched (blue) to the  $T_b$ -identified sources in this study. *Bottom:* Peak brightness from the high-resolution image for sources which are matched (orange) and not matched (blue) to the VLBA sources.

GMRT observations, are diverse, which means scaling the sensitivity from 1.3 GHz to 144 MHz using a single spectral index is not reliable for predicting the outcome.

Cross-matching the  $T_b$ -identified samples at MHz and GHz frequencies, we find only 22 matches (15  $T_{b, \text{peak}}$  sources, and 7  $T_{b, \text{total}}$  sources). That means there are 43 unique VLBA detections, and 70 unique  $T_b$ -identified AGNs at 144 MHz. Using the VLA to GMRT spectral indices for the VLBA sources, the distribution of which is shown in Fig. A1, we can see that those which are not matched are predominantly flat-spectrum sources. However, it is worth keeping in mind that the spectral indices were calculated from lower resolution data, and may not reflect the spectral index of the compact AGN component. Looking at the peak brightness at 144 MHz we can see (Fig. A1) that the VLBA-matched sources are, on average, brighter.

One must consider that although both studies use  $T_b$  as an AGN identification tool, the beam solid angles at these different frequencies are an order of magnitude different, and therefore will be probing different physical scales. Future work exploiting the EVN and e-MERLIN will provide matched resolution to the high-resolution image studied here, and will be more appropriate for drawing meaningful conclusions.

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.